Profiling of a large-scale municipal wireless network

Carlos Kamienski1 · Juliano Ratusznei1 · Andrea Trindade1 · Dave Cavalcanti2

Published online: 16 June 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract
Pervasive connectivity is an essential underlying substrate for smart cities, leading municipalities to start programs where Wi-Fi is the fundamental building block to develop public Municipal Wireless Networks. Even though hundreds of cities around the world offer some form of Wi-Fi access, there are no widely available results regarding the network Quality of Service (QoS), user Quality of Experience (QoE), and overall utilization profile. The Municipality of São Paulo operates a free Wi-Fi Internet program in 120 public spaces called digital squares. We collected user connection data, network performance, and service availability for more than 2 years from the 120 squares and undertook experiments with video streaming in five squares. We used this unique large dataset to evaluate the impact of current admission control practices in public Wi-Fi networks on the network QoS, user QoE and service availability, also providing insights into the most common QoS/QoE issues and their causes. We also leveraged the data set to establish and verify a correlation between the number of users access in the network and specific events occurring in the area.

Keywords Municipal Wireless Networks · Performance of public Wi-Fi · Video streaming over Wi-Fi · Event identification

1 Introduction
In the last 15 years, several cities in all continents started programs of different magnitudes and coverage for providing Internet access as a public utility like water and power. Most, if not all, make use of Wi-Fi as the fundamental building block to develop a public Municipal Wireless Network. In the early 2000s, it was widely believed that cities and communities would provide Internet access to their citizens via Municipal Wireless Networks, in order to promote digital inclusion and create new services towards smarter cities [6, 33]. Particularly in the United States, where 3G was deployed late compared to other countries, hundreds of cities started projects to provide free Wi-Fi access, also inside private and public buildings. By the end of the decade, most of them had been discontinued for different reasons, such as regulatory or technical issues, high cost, and low demand [12]. Nowadays, there is an ongoing discussion about the advantages and disadvantages of Municipal Wireless Networks [35], focusing on the most appropriate business models for the deployment and operation of those networks [21].

Nevertheless, Public Wi-Fi coverage is even becoming part of the urban planning of a smart city. In China, about 300 municipalities have projects to become smart cities counting on free Wi-Fi network coverage integrated into the overall urban development plan [42]. Hundreds of cities around the world offer some form of Wi-Fi access. When it comes to large networks with a large number of hotspots,
we can include as examples Barcelona,\textsuperscript{1} Buenos Aires,\textsuperscript{2} Adelaide,\textsuperscript{3} Paris,\textsuperscript{4} Seoul,\textsuperscript{5} New York\textsuperscript{6} and Curitiba (Brazil).\textsuperscript{7} However, most cities limit the connection time or simply do not provide free access. Regardless of the type and scope of service, there are no widely available results providing insights on the network Quality of Service (QoS), user Quality of Experience (QoE), and overall utilization profile of the service, as we provide in this paper.

The Municipality of São Paulo City in Brazil started the WiFiLivreSP\textsuperscript{8} program in 2014, aimed at making the Internet more accessible to the citizen by providing free of charge access to a municipal Wi-Fi network covering 120 public spaces called digital squares. This service is provided by Internet Service Providers via a public bidding process and covers the five regions of the city: Center, North, South, East, and West. The Wi-Fi modes supported are IEEE 802.11a, 802.11b, 802.11g, and 802.11n. By contract, ISPs had to cover 70\% of the area of digital squares, except for large parks.

We established a partnership with the Municipality of São Paulo to study the performance of this large-scale public Wi-Fi network. The main research goals are twofold. First, we leverage the access to data from all 120 digital squares to evaluate network QoS and user QoE. As part of the evaluation, we performed controlled experiments in some of the locations to evaluate QoS and QoE data for video streaming use cases. Secondly, we evaluate the feasibility of using Wi-Fi user access data to identify specific events happening in the area. This paper provides our key findings related to this municipal Wi-Fi network at São Paulo, which is not available in the literature and could apply to many large Wi-Fi networks.

We collected data of user connection, network performance, and service availability for more than 2 years from the 120 squares and undertook experiments with video streaming in five squares. Our goal was to validate three hypotheses: Hypothesis (#1) static admission control based on a priori estimates may unnecessarily deny service to users thus diminishing the benefits of this public policy; Hypothesis (#2) traffic shaping based on a priori estimates of average per-user throughput may unnecessarily prevent users from accessing their preferred applications; Hypothesis (#3) the number of Wi-Fi users can be employed to identify people crowding in public places. Hypothesis number #1 and #2 were confirmed, while hypothesis #3 was partially confirmed. Our key conclusions, also the most important contributions of this paper, can be summarized as:

(a) Admission control is used to block a new connection that may cause general performance degradation. Even though there has been significanct amount of work on advanced admission control for wireless networks, we found out that in practice, static admission control is still used by service providers. The data shows that static admission control would not be useful, but it also shows that in some cases, a high number of users negatively impacted the QoS/QoE, suggesting that dynamic measurement-based admission control should be implemented for avoiding poor user experience.

(b) Traffic shaping is a packet queuing technique used to enforce a maximum bandwidth for a network connection to avoid general performance degradation. In a large-scale public Wi-Fi network, no traffic shaping allows users to use high bandwidth-demanding services, such as video streaming, whenever the network is underutilized. A static and strict control over system resource usage may block users from accessing services that, from their point of view, make the public policy to be a success. We found out that there is a considerable variation in the perception of the quality of experience (QoE) for single users trying to watch videos over Wi-Fi. In some squares, users obtain much more than the 512 Kbps estimate, whereas, in others, they obtain much less than this. In other words, whenever the network is underutilized, users could watch videos with a coding rate of up to 1200 Kbps. The QoE was measured based on video interruptions that negatively affect the user experience. This behavior makes a case for dynamic traffic shaping where a maximum per-user bandwidth is only enforced when the network resources start to become scarce.

(c) The number of users connected to a public Wi-Fi network in a digital square can be used to identify people crowding in public places for events that occur in specific squares, but not generic events that occur in multiple squares at the same time. We found out that for events that occur in multiple places at the same time, there is no discernible correlation between the instant number of Wi-Fi users and the number of users accounted in the previous month. In this case, people tend to spread over multiple locations and dilute the increase in the number of Wi-Fi users in digital squares. On the other hand,
events that occur in specific squares show a clear increase in the number of users, because they cause people to stay in the same place for a long time. This finding opens up new opportunities for authorities to use Wi-Fi hotspots to perform early identification of people crowding in unpredictable events, such as unexpected traffic jams or demonstrations.

In the remainder of the paper, Sect. 2 presents background and discusses related work. Section 3 presents research methods and design, and results are shown in Sect. 4. Section 5 discusses the main lessons learned in our long term study and opportunities for future work, and finally, Sect. 6 draws some conclusions.

2 Background and related work

The emergence of smart cities requires enhanced, widespread, and 24 × 7 connectivity that is pushing the development of new technologies for mobile and wireless communications. Examples are the ever-evolving cellular standards that are now going towards 5G [41] and high-speed Wi-Fi with the newest IEEE 802.11ay standard to be released in 2020 [13], but also Low Power Wide Area Network (LPWAN) IoT wireless solutions [27], or a combination of them, such as 5G-powered IoT [26]. For now, Wi-Fi is the leading technology in this field and thus the focus of this paper. Also, nowadays, most client applications are deployed as smartphone apps that interact with a server located in a cloud datacenter. Soon, they are also likely to communicate with a server infrastructure located in the network edge, thus closer to the user, via fog computing, in order to provide lower latency for real-time applications [28].

2.1 Municipal Wireless Networks

Municipal Wireless Networks, or Municipal Wi-Fi, are built by or for governments to provide free or low-cost Internet access based on the IEEE 802.11 standards to citizens, public agencies, and businesses [18]. It encompasses different models, varying from infrastructures built and run by the local government, partnerships between the local government and private companies, or even only regulated by governments and explored by entrepreneurs. There is an ongoing discussion about the advantages and disadvantages of such networks focusing on the most appropriate business models to deploy them [21, 35]. Regardless, it is nowadays widespread in the world, and commercial businesses have been setting up Wi-Fi infrastructures [33]. Community Wireless Networks are alternative models to provide Wi-Fi Internet access for free by encouraging members to share their networks with higher quality and lower cost for the public authority [39]. Some initiatives combine public and private Wi-Fi networks to create a vast collaborative network [16, 50].

For a subject that received so much attention in the last 15 years, it might be curious that most information available in the literature focuses on the benefits and drawbacks, public policies, and funding models. Currently, most literature reports alternative uses of Municipal Wireless Networks data, such as understanding urban vitality [23], classifying users [32, 33], and planning future smart cities [43]. Long term and large-scale quality and performance data of Municipal Wi-Fi Networks, as we show in this paper, are still challenging to find, maybe because it does not interest municipal authorities. Biswas et al. [8] conducted a large-scale study involving millions of Wi-Fi hotspots in two separate weeks in January 2014 and January 2015. Their main conclusion is the continuous and fast-growing demand for Wi-Fi, the increase in the number of connected devices, and the use of video streaming applications such as YouTube and Netflix. Although they studied networks in cafeterias and bars that differ from our focus here, their findings contribute to highlight the importance of providing public services that fulfill the needs of citizens.

Oliveira et al. [31] analyzed and compared the behavior of users of a university campus and an urban Wi-Fi network, mostly focusing on user connection behavior. Compared to previous studies, we did not have access to individual user connections and accessed applications due to privacy concerns of the municipal authorities. Thus, our dataset contains aggregate metrics for digital squares, such as the number of connected users, incoming throughput, RTT, and packet loss. Thus, our results focus on aggregate user behavior, not covered by these previous studies. Ghosh et al. [14] conducted a large-scale study of public Wi-Fi in bars, companies, and cafeterias in two cities in the United States and derived a data mining based predictive model of user utilization. Our study is complementary to this one because it reports results of publicly available Wi-Fi networks inside private organizations, and ours focuses on a real public Wi-Fi network. Ishikawa et al. (2019) developed a QoS Estimation method for Wi-Fi networks based on null function data frames, which require a specific firmware to be installed in access points and tested it in a large conference hall [17]. Unlike this study, here, we report the results of a more detailed performance analysis of a large-scale municipal Wi-Fi program and some insights about using the same Wi-Fi connection data for identifying the occurrence of events.

While Wi-Fi is becoming increasingly pervasive, resource control aimed at providing advanced network QoS and user QoE is still a challenge. The variable and
sometimes unpredictable user base makes it difficult to deliver performance and fairness, mostly in large-scale public wireless networks. Previous studies mostly focused on providing QoS guarantees based on the IEEE 802.11e standard [29, 51]. Admission control, traffic shaping, and load balancing are mechanisms used to provide enhanced performance in Wi-Fi networks. Most Wi-Fi access points allow static admission control where user access is blocked whenever a maximum predetermined number of users are connected.

Dynamic measurement-based admission control mechanisms rely on the instant availability of resources, i.e., wireless and wired bandwidth [51], although not yet fully supported by commercial products. Also, traffic shaping has been proposed for providing QoS in both wired and wireless networks, which limits the throughput that users can use in an attempt to provide consistent performance and fairness for a higher number of users [3]. Currently, Wi-Fi performance in top-notch commercial products focuses on features such as wireless channel management and load balancing among access points and frequencies [2, 10]. The current state of the art and the technique mostly refer to dynamic admission control for IEEE 802.11 networks, i.e., focusing on estimates of resources available in the wireless segment of the network [7, 45]. However, for a public Wi-Fi network user, the wired and wireless segments must be analyzed together in order to provide adequate performance levels. In other words, the requirements of dynamic resource control for large-scale Municipal Wi-Fi networks, as we show in this paper, are not fully met by the existing literature.

2.2 WiFiLivreSP: a municipal wireless network in São Paulo

The Municipality of São Paulo started the WiFiLivreSP program in 2014, aimed at making the Internet more accessible to the citizens by providing a free of charge access to a municipal Wi-Fi network covering 120 public spaces called digital squares. The Wi-Fi modes supported are IEEE 802.11a, 802.11b, 802.11g, and 802.11n. IEEE 802.11ac was not required since the standard was not ready along the year 2013 while the public bidding was open. Figure 1 depicts some images of the WiFiLivreSP program. Figure 1(a) shows a map with the locations of the digital squares that can also be found online\(^9\) and Fig. 1(b, c) show a sign that tells the citizen the availability of free Wi-Fi and an Access Point, respectively.

\(^9\) prefeitura.sp.gov.br/cidade/secretarias/inovacao/inclusao_digital/index.php?p=216921.

2.3 Video quality

The widespread prevalence of mobile devices and wireless communications (LTE and Wi-Fi) makes them the preferred user technological combo for video streaming, which encourages researchers to push the current state of the art. For example, Burger et al. [9] conducted a simulation-based investigation to assess the effects of Wi-Fi offloading on the Quality of Experience (QoE) of video streaming and concluded that 3G/4G provides a higher user experience compared to Wi-Fi. However, no performance analysis results that evaluate the video streaming QoE in public Municipal Wi-Fi networks have been found.

The direct assessment of the user perception of QoE has been preferred in the last years compared to the older approach of mapping network Quality of Service (QoS) metrics into QoE ones [4]. Usually, the measurement of video quality uses metrics such as Peak Signal to Noise Ratio (PSNR) [40], Structural Similarity (SSIM) [46] and, Video Quality Metric (VQM) [34]. While these metrics help the process of understanding the user experience, they do not directly capture the user sentiment towards the quality of the watched video, and consequently, new metrics that directly measure the QoE have been proposed [5]. Video quality tends to have a high correlation with the behavior of the buffer, such as the playback start time, the number and duration of interruptions (video freezing) and, the place in the video where the interruptions occur [4, 19].

Here we opt for evaluating the video quality based on the user perception with the assumption that video interruptions caused by buffering mechanisms are the critical factor that negatively affects user experience. Besides, adaptive streaming schemes that decrease the coding rate during low network performance periods also affect user perception, as shown by Kim and Kim [22]. However, we used fixed coding videos in order to be fair in our analysis when comparing different digital squares, days, and times.

2.4 Event identification

The very essence of any monitoring activity is the perception that an event of interest is occurring so that the anticipated reactions can be triggered, such as starting corrective procedures upon the confirmation of an unwanted event. The primary goals of event detection are [30]: (a) identify the occurrence of an event of interest; (b) characterize the event; (c) identify the affected data subgroup by the features that better describe the event (e.g., spatial region and time duration); (d) assess the severity of the event; (e) detect the event accurately; (f) detect the event in an agile and timely manner. The last two goals are of particular importance to act proactively to minimize the
impact of events, i.e., to minimize the error and delay in event detection. Yan et al. [49] published a comprehensive survey on analytical techniques for detecting complex events generated by applications for Smart Cities. They suggest the combined use of models, approaches, and automated orchestration of techniques, with the semantic perception of the application context, which is also supported by Agrawal et al. [1].

Since technology has become pervasive in urban spaces, plenty of methods for analyzing and estimating crowd behavior based on event detection have been developed, many of them relying on users carrying Wi-Fi-enabled devices [20]. Schauer et al. [38] showed the feasibility of estimating crowd density by capturing Wi-Fi management frames for 2 days in an airport and comparing the data with boarding pass scans. Xi et al. [48] estimated the number of people moving in indoor and outdoor places by monitoring the variation of the wireless channel state of a Wi-Fi access point [48], as the Channel State Information (CSI) is highly sensitive to the movement of people. Crowd monitoring in indoor places also has been shown possible by tracking Wi-Fi devices using dozens of Wi-Fi scanners mounted in the ceiling of a large event center [47]. During almost 2 weeks, they counted more than 300,000 individual devices and observed an error of less than 20% compared to the ground truth coming from video images. Toch et al. [44] recently published a survey on machine-learning methods for mining and understanding large-scale human mobility data, which may be used as a basis for event detection. They categorize methods into three classes, namely, user modeling, place modeling, and trajectory modeling. Demographic and census inferences can also be performed using Wi-Fi data. Kontokosta and Johnson [24] showed that by comparing data collected from 53 APs in New York to official census information, they were able to create an hourly population estimate of residents, workers, and visitors. Li et al. [25] showed that demographic information, such as gender and education level, can be inferred from Wi-Fi data using machine learning. They analyzed a dataset with almost 30,000 users covering 5 months.

In this paper, we consider event identification in public spaces by evaluating Wi-Fi user connection information and comparing it with crowd formation due to events in public squares. There are some differences between our approach in this paper and the abovementioned literature. Firstly, we did not need to be physically present in the squares to monitor user connection patterns, since we used management data provided by the ISPs in charge of the Wi-Fi service. Secondly, we compared the observed people crowds measured by the number of connected users with events of interest that happened in the squares. Thirdly, we evaluated events of interest for almost 2 years in 120 digital squares of a Municipal Wireless Network, while existing literature usually reports small scale crafted experiments or using public datasets.
3 Research methods and design

Our endeavor of profiling user behavior and performance of the digital squares is divided into three parts, which are network usage, video streaming, and event identification. Table 1 shows the 29 squares selected to be deeply analyzed, namely 15 for usage, 4 for video, and 11 for event. Squares are referenced by acronyms to facilitate the understanding for the general reader, and the name is provided to readers familiar to São Paulo. For network usage, three squares from each region were selected, representing different levels of the expected maximum number of users. For video streaming, which required on-site measurements, one square per region was selected based on the criteria of usage, safety, and logistics. For event identification, the 11 selected squares hosted significant events. Two squares were used for two different analyses, Square C4 for video and event and Square S3 for usage and event. The number of squares is different in each region because of the event identification analysis that required us to select squares where events occurred, which were mostly in the Center, South, and East regions.

Also, the predetermined maximum number of users is shown for each square in Table 1. This maximum number of users was calculated for each square based on estimates of the number of people who visit the squares daily, which has been done by technicians of São Paulo municipality. This number plays a vital role in the wireless and wired network resource provisioning of the squares, even though it was not explicitly enforced via admission control.
policies. In other words, this is a theoretical and static number that was guessed by some offline process, used only as a reference for the ISPs to provision their networks, as users are not denied access when this number is achieved.

Figure 2 depicts the selected squares, as a subset of Fig. 1, in Google My Maps, so that readers can have a better understanding of square location and density.

### 3.1 Network usage

Network usage is defined by access, availability, and performance metrics using both passive and active measurement techniques. A critical factor that affects performance and quality of experience is the number of users allowed to connect to any Wi-Fi access point in a digital square. We collected data from June 2014 to August 2016 from 119 squares (one was partially closed during this period), and 15 were analyzed individually for network usage, as presented in Table 1. Three squares from each region were selected, representing different levels of the expected (predetermined) maximum number of users: low, medium, and high. The choice for these particular 15 squares was made together with the municipal authorities based on the subjective criteria of importance and representativeness. In other words, these 15 squares properly represent the

---

10 Google My Maps allows users to create custom maps that can overlay Google Maps (mymaps.google.com).
behavior of 119 squares in the five regions of São Paulo. The collected data comes from two different sources:

- **SIMETBox** The performance of all digital squares in São Paulo is monitored by SIMETBox, an open-source middleware that can be installed in different models of Wi-Fi Access Points. SIMETBox is developed by the Brazilian Network Information Center (nic.br) and performs tests in Wi-Fi networks connected to the Internet. For the WiFiLivreSP program, every square was equipped with one or more APs for providing Wi-Fi service and also one exclusive AP for running SIMETBox. Once installed, SIMETBox performs measurements every hour by sending a sequence of packets and calculating throughput, latency, and packet loss.

- **ISPs** ISPs are requested to provide management data in XML from the squares they operate, such as the number of connected users and the incoming bitrate.

Network usage metrics collected from SIMETBox and ISPs shed light into three aspects that directly affect the experience of the citizens when using the Wi-Fi service:

- **Access** In an attempt to be inclusive and to understand the user behavior, the maximum predetermined number was not enforced via admission control mechanisms, i.e., user access was not blocked when this number was achieved, thus making it a theoretical maximum. As metrics, we collected the number of users and calculated service utilization:
  - **Number of users** the instantaneous number of users connected at each square. In larger squares with more than one Wi-Fi Access Point (AP), the number of connected users is the sum of all APs.
  - **Service utilization** calculated as the percentage of the number of users in relation to the predetermined maximum. Since no admission control is enforced, this metric can exceed 100%.

- **Availability** Squares were designed to provide Internet service at a 24 x 7 basis, and ISPs have to keep availability above 96%. We used two metrics, based on ISP and SIMETBox data.
  - **Availability ISP** for all data collection events, the service is considered active if values reported by SIMETBox are not null. The percentage of the active data collection events divided by the total is the availability reported by SIMETBox.
  - **Performance** Squares must provide service quality levels compatible with QoS metrics such as throughput, delay, and packet loss. The capacity of Wi-Fi access points and links connecting to the Internet must be configured in such a way to offer at least 512 Kbps per user, an average latency of 5 ms, and “low” packet loss. Among the many performance metrics available, we used the following:
    - **Throughput** aggregate incoming bitrate reported by ISPs.
    - **Number of users** provided by ISPs, it is used for computing other performance metrics, such as the per-user average throughput.
    - **Latency** the round trip time (RTT) measured between a SIMETBox located in every square and the main Internet Exchange Point in São Paulo operated by Equinix under the PTTMetro program of the Brazilian Internet Steering Committee (CGI.br). SIMETBox sends every hour a group of packets from the squares to the Internet Exchange, and the computed metric represents the combined delay of both wireless and wired networks.
    - **Packet loss** the two-way packet loss between the SIMETBox and the Internet Exchange Point, presented as a percentage. The same packets used to compute the delay yield the packet loss metric.

Data from SIMETBox and ISPs were regularly collected from October 2014 to August 2016. Data is missing for some periods due to different reasons, such as lack of connectivity, server crash, data unavailability, or unexpected changes in the data formats. We did not have access to the user or device-level data (MAC and IP layers) due to user privacy concerns, but only to aggregate data. The data processing required the collection of data of ISPs and SIMETBoxes and pre-processing for eliminating missing, duplicate, or outlier data, as well as equalization of metrics.

### 3.2 Video streaming

Video usage has been consistently growing and will account for 82% of Internet traffic by 2021, according to Cisco [11]. Therefore, it is expected that users try to watch...
videos whenever they have connectivity provided by the WiFiLivreSP program. For the analysis of video streaming, we hypothesized that traffic shaping based on a priori estimates of average per-user throughput might unnecessarily prevent users from accessing their preferred applications.

Video analysis requires on the spot active measurements, and therefore we narrowed down our universe of digital squares to five, one in each region. The choice of the five squares was based on three criteria:

- **Usage** appropriate squares were considered those with a significant number of connected users over a certain period compared to their predetermined maximum number of users;
- **Safety** since the video measurements required a physical presence on the spot, the criminality rate of the area was considered an important factor limiting the number of candidate squares;
- **Logistics** the convenience of access via public transport and an efficient route between the squares to visit all of them on the same day.

After an iterative process where the three criteria were applied and screened most squares, the five squares shown in Table 1 were selected, namely, squares C4 (Center), N4 (North), S4 (South), E4 (East), and W4 (West). Metrics used to analyze video streaming performance belong to two groups: network QoS and user QoE.

- **Network QoS Metrics** traditional metrics used by the networking community, coming from both SIMETBox and ISPs at the exact time of the video measurements: per-user throughput, latency, packet loss, already defined in Sect. 3.1.
- **User Video Streaming QoE Metrics** we adopted objective QoE metrics [19] that influence the user experience:
  - **Playback start time** the time duration before a video starts, also called initial delay. Longer initial delays represent low QoE because users get bored and tend to leave the video before it starts.
  - **Number of interruptions** interruptions, also called freezing, occur when the video is stalled due to an empty buffer. Users tend to leave the video prematurely when the number of interruptions is high;
  - **Duration of interruptions** the longer the duration of each interruption, the lower the QoE.
  - **Received bitrate** the actual video incoming bitrate as received by the application, which must match the video encoding rate for positive user experience.

The correlation and causality between network QoS metrics and user QoE metrics are straightforward. When throughput is low, and latency and packet loss are high, video streaming experience higher initial delays and higher number and duration of interruptions. In other words, low network quality of service causes low user quality of experience. On the other hand, high quality of service allows a high quality of experience, given that no other factors affect quality, such as poor video quality and problems in video servers or client devices.

The measurement of video streaming QoE according to user perception based on video interruptions required the development of a tool. We developed a video client [16] using VLC [17] (v2.2.1), an open-source multiplatform media player that allows different video coding formats and provides libraries for customizing client applications. When launched, this video client automatically connects to the VLC server hosted in our university lab and starts the experiment, where it also measures the instantaneous end-to-end RTT and packet loss. The VLC library allows the client to identify whenever a video streaming is started, interrupted, and resumed, which makes it possible to count the number and duration of interruptions. Also, it allows the client to obtain the received bitrate (throughput).

Unlike network usage data measurements, for the video streaming, we performed on the spot measurements, which required a more elaborated work plan with well-designed experiments. We performed video streaming experiments twice to understand the variation of the time of the year in the user QoE, in November 2015 and January 2016. Three video coding bitrates (420 Kbps, 840 Kbps, and 1200 Kbps) were tested for the five selected squares from Sunday to Saturday. In November 2015, we used the 840 Kbps and 1200 Kbps video bitrates, but after low-performance levels observed in some squares, in January 2016, we switched from 1200 to 420 Kbps. Each experiment was replicated 30 times, and 99% asymptotic confidence intervals were computed. The same 1-min long video was played 30 times with two video bitrates. Also, considering the time spent with changing videos, intervals between experiment replications, and occasional connection problems, experiments lasted approximately 1 h and 40 min. The video excerpt has 1 min [18] and was edited from a VLC promotional video at YouTube with a high number of views that mixes different scenes, light and dark and with different speeds. [19] Our video streaming experiments are based on HTTP and do not provide any adaptation mechanism so that we could effectively understand the influence of the square in the user QoE with a specific

---

[16] gist.github.com/ratusznei.
[17] videolan.org/vlc.
[18] youtu.be/05i6WZg8v8c.
[19] youtu.be/nKLu9yen5nc.
Table 2 Parameters

| Parameter       | Value   |
|-----------------|---------|
| Distance from the AP | $\geq 10$ m |
| Target server   | University lab |
| Video duration  | 60 s   |
| Metric capture  | 1 s    |

Table 3 Factors and levels

| Factor      | Level                  |
|-------------|------------------------|
| Square      | Center4, North4, South4, East4, West4 |
| Bitrate     | 1200 Kbps, 840 Kbps, 420 Kbps |
| Hour        | 7 am to 6 pm           |
| Weekdays    | Sun, Mon, Tue, Wed, Thu, Fri, Sat |

encoded quality if the same bitrate is maintained during the whole video session.

Table 2 lists the fixed parameters used in the experiments. In order to avoid bias, the researcher was located at 10 meters of the Access Point in all experiments. This value has been chosen after some preliminary experiments and after empirically observing the user behavior in the squares trying to obtain a better connection. We did not vary the distance due to physical, time, and weather limitations. Table 3 shows the factors (parameters varied) and their levels (values), as explained previously.

3.3 Event identification

During our analysis, it became clear that after some adaptation process, the user behavior in terms of Wi-Fi connection habits enters a predictable and steady state. Moreover, understanding the typical user behavior is the first step for being able to detect abnormalities. Here we hypothesize that the number of Wi-Fi users can be used to predict people crowds in public places. In this paper, due to scope limitations, we deal only with planned and known events—such as celebrations, concerts, and religious festivals—as the first step in automatic event identification using Wi-Fi access.

The idea is to understand whether Wi-Fi user access in digital squares can be effectively used to identify user-crowding patterns in a large city. We analyzed two types of events: general events that potentially occur in all squares, such as New Year and Carnival, and those that happen in specific squares, such as music concerts and religious festivals. In a large metropolis like São Paulo, with more than 12 million inhabitants, thousands of significant events occur every year. When selecting the squares that received events, we prioritized the older ones according to the square installation schedule. Our analysis covered generic events that presumably have a higher chance to occur in multiple squares:

- **Launch of Pokémon GO in Brazil** Pokémon GO was a fever worldwide and also in Brazil, where it was officially launched on August 3rd, 2016. It was the trigger that called our attention to understand the relationship of Wi-Fi usage with the occurrence of significant events.
- **Carnival Week in 2015 (13–18/Feb/2015) and 2016 (05-10/Feb/2016)** by Carnival Week we understand the period between the Friday before and the Wednesday after Carnival (the latter is also called Ash Wednesday) when most Brazilians enjoy holidays, and a part of them (usually younger ones) go to the streets to dance to the music of bands and Trios Elétricos (electronic music cars). There is a non-negligible overlap between places where people crowd for Carnival celebrations and digital squares.
- **New Year 2015 and 2016** This is a traditional celebration and might lead users to access the Internet via the free Wi-Fi connection in digital squares.
- **Cultural Overnight 2015 and 2016 (21–22/May/2016)** Cultural Overnight (Virada Cultural) is an annual 24-h cultural festival that occurs in São Paulo, including live music concerts, plays, and a variety of other cultural activities and performances.

Also, we identified squares where significant specific events happened, shown in Table 4, which are:

- **Anniversary of São Paulo 2016** a variety of celebrations occur in São Paulo on its anniversary every year on February 25th.
- **Independence of Brazil 2015** on September 7th, the Independence of Brazil is celebrated with parades.
- **Cultural and thematic events** we identified three additional events that generated some increase in the use of Wi-Fi in the digital squares: Chinese New Year (13–14/Feb/2016), Jim Beam History Fest (15/Jul/2016) and Street Samba (Pagode da 27) (28/Aug/2016), a samba singing and dancing event.

The comparative analysis needed for event identification was helped by the development of software tools using the R software environment for statistical computing. Data was loaded from the SQL database created for Sect. 3.1 into the RStudio tool, and the required scripts were developed and executed to generate the charts presented in Sect. 4.3.
4 Data analysis

The results from the data analysis are divided into three parts, as presented in Sect. 3.

4.1 Network usage

The contract signed between the municipality of São Paulo and ISPs does not oblige the latter to disclose the network infrastructure used to provide the Internet service. Therefore, regular measurements are required to evaluate service compliance with the contracted SLA in terms of QoS and QoE metrics. Here we present results for 2015 that are representative of the whole period.20

4.1.1 User access

All squares were planned to provide Internet access service to a given number of users according to an initial estimate of the potential number of citizens that would make use of the digital squares. Table 5 shows statistics for the whole year of 2015 for the average and the maximum number of connected users, the theoretical estimated maximum number of users, and average and peak service utilization, for the 15 squares listed in Table 1. It is possible to observe that the maximum number of accesses is not enforced in practice. Therefore, it is called a theoretical maximum number of users.

A characteristic example shown in Table 5 is Square C3, where up to 501 users were simultaneously connected, but the initial estimate was 250. The last two columns show the average and peak service utilization, i.e., the ratio of the number of users by predetermined maximum. It reveals that for most squares, the estimates were too low since no admission control was used, and the utilization reached more than 100% or even 200% of the planned accesses. On average, the number of users in the squares is low, except for some specific cases. The same behavior can be observed in most squares shown in Table 5 and in most of the 119 squares, which are not shown for the sake of conciseness because the entire table is lengthy. Generally speaking, we observed that the squares located in the Center Region received a higher number of accesses because it is a transit and work area, where many people cross squares going/coming from work or school. A higher number of connected users was also observed in the East Region, a very densely populated region and a dormitory suburb. Squares located near residential buildings consistently kept a high number of users throughout the day. For those cases, an investigation showed that many households canceled their contract with Internet Service Providers after the WiFiLivreSP program started.

Figure 3 provides a visual representation of the geographical location of the 15 selected squares from Table 5, which provides a way of understanding the density of users per region of the city. The metric in the circles is the maximum number of users observed for each square.

Figure 4 expands the data provided in Table 5 by showing the average hourly behavior of the number of users (global and for each region—Fig. 4a) for five selected squares, one in each region (Fig. 4b). The squares were selected based on receiving the highest number of users in each region, namely C3 (Center), N1 (North), S1 (South), E3 (East), and W3 (West). We classified all squares into

| Square | Event                                      |
|--------|--------------------------------------------|
| C4     | Cultural overnight 2015 ✓                  |
| C5     | Cultural overnight 2016 ✓                  |
| C6     | Anniversary São Paulo 2016 ✓                |
| C7     | Independence Brazil 2015 ✓                 |
| C8     | Cultural and thematic ✓                    |
| C9     |                                            |
| S3     |                                            |
| S5     |                                            |
| S6     |                                            |
| E5     |                                            |
| E6     |                                            |

20 The full technical report is available (in Portuguese) at prefeitura.sp.gov.br/cidade/secretarias/upload/servicos/inclusao_digital/WiFiLivreSP/relatoriosufabc/Wi-Fi-Livre_R2-9.pdf.
### Table 5 User access to the digital squares in 2015

| Square | Predetermined max | Observed Average | Max | Service utilization (%) Average | Peak |
|--------|-------------------|------------------|-----|--------------------------------|------|
| C1     | 75                | 28               | 93  | 37                             | 124  |
| C2     | 150               | 45               | 270 | 30                             | 180  |
| C3     | 250               | 223              | 501 | 89                             | 200  |
| N1     | 50                | 14               | 72  | 28                             | 144  |
| N2     | 75                | 15               | 80  | 20                             | 107  |
| N3     | 100               | 14               | 76  | 14                             | 76   |
| S1     | 50                | 8                | 36  | 16                             | 72   |
| S2     | 100               | 24               | 179 | 24                             | 179  |
| S3     | 250               | 40               | 254 | 16                             | 102  |
| E1     | 50                | 26               | 102 | 52                             | 204  |
| E2     | 100               | 25               | 134 | 25                             | 134  |
| E3     | 150               | 134              | 303 | 89                             | 202  |
| W1     | 50                | 6                | 31  | 12                             | 62   |
| W2     | 75                | 13               | 73  | 17                             | 97   |
| W3     | 125               | 25               | 123 | 20                             | 98   |

![Geographical visualization of user access density (2015)](image)

**Fig. 3** Geographical visualization of user access density (2015)
four categories of function according to usage pattern, region, and seasonality of usage days and hours: sport and leisure, passage, stopover, and religious. Most squares present seasonality during weekdays and hours of the day, which is influenced by factors such as region and usage category. For example, passage squares tend to have more access during weekdays in business hours, whereas sport and leisure squares tend to receive more users in the evening and weekends. The former is typical for the Center, where the latter for the other regions, except for parks, whose behavior is similar to Center. This typical behavior that tells apart squares in the Center and the other regions can be well observed in Fig. 4b, where the square in the Center has this typical bell-shaped behavior with the peak hours between 10 am and 3 pm. The square in the East shows an increase during the day and a decrease in the evening. The squares in regions North, South, and West also have the same behavior, though not so noticeable due to the scale of the chart. These findings may be used to improve network management capabilities, as configuring different admission control and traffic shaping policies based on the square category according to its function.

Figure 5 provides the same information from Fig. 4b for 15 squares in a base-2 logarithmic scale, which shows the fact that squares present a typical behavior during the day for each region.

Throughout this section, in order to show specific behaviors that can be generalized for most squares in the city, we will use two squares as examples, namely Square C3 in the Center Region that experienced a high number of accesses, and Square N1 in the North Region where a low number of accesses was observed. These two squares were selected because they faithfully represent different user access behaviors concerning the functions that public places play in the dynamics of a large city. Square C3 is located in the Center and has movement mainly during the day so that the user count follows a bell-typical shaped curve. On the other hand, Square N1 in the North Region is in the middle of a dormitory zone, where movement in the square grows towards the evening.

Figure 6 depicts the number of connected users in 2015 averaged by week for Square C3 and Square N1. Figure 6a shows that Square C3 kept a high number of connected users for the whole year, and for 17 weeks in a row, the average was even higher than the predetermined maximum. Even when the average was below the theoretical maximum, during the day, there are peaks where the actual access demand exceeded the maximum. Figure 6c zooms in 3 weeks of June 2015 for Square C3 and highlights this behavior. It can also be observed that we had no data for week 31 caused by a change made in the configuration of the XML files that took us some time to realize and fix the problem. For Square N1, the number of accesses was considerably lower than expected, showed in Fig. 6b, d.

The policy of not enforcing admission control wherein squares allow users to connect when the predetermined maximum number of users is exceeded can be considered a fortunate fact or not, depending on the network capacity provisioning. Intuitively, admission control must be enforced in order to prevent service degradation. On the other hand, service degradation will only occur when system capacity is effectively exceeded as measured by a real-time monitoring system and not when estimated static system capacity is exceeded. If admission control prevents users from connecting before the instant system capacity is exceeded, this is an unfortunate event, mainly for public policy. On the other hand, a successful action is when the system blocks users just before their connection would cause performance degradation. For example, Square C3 has a predefined maximum of 250 users and should be provisioned with at a link of at least 128 Mbps to fulfill this requirement. Table 6 in Sect. 4.1.3 reveals that the average downlink throughput in 2015 was 305.62 Mbps, more than twice the configuration required by contract.
4.1.2 Service availability

Availability is critical for the success of any public policy for providing Internet access via Wi-Fi wireless networks since it defines the first impression of the users regarding the service quality. However, service availability was the leading cause of poor user quality of experience, even though ISPs were expected to keep the availability above 96%, as mentioned in Sect. 3.1. Figure 7 depicts service availability in the 15 selected squares, according to ISPs and SIMETBox measurements.

A first observation is that no square was available 100% of the time. The ISP serving the Center and East regions tends to report more variability among squares, whereas the ISP serving the North, South, and West regions reports the same availability for all squares. Also, there is a considerable variation of the availability reported by the ISPs and SIMETBox. On average, availability reported by ISPs was 95%, while from SIMETBox was only 66%. We cannot
Table 6 Performance of the Wi-Fi service (average values for 2015)

| Square | Throughput (Mbps) | Users | Throughput per user (Mbps) | RTT (ms) | Packet loss (%) |
|--------|-------------------|-------|---------------------------|----------|-----------------|
| C1     | 4.78              | 28.13 | 0.17                      | 48.37    | 0.96            |
| C2     | 305.62            | 223.08| 1.37                      | 67.82    | 1.13            |
| C3     | 17.38             | 44.56 | 0.39                      | 60.92    | 1.41            |
| N1     | –                 | 101.87| –                         | 60.17    | 0.99            |
| N2     | 2.45              | 15.15 | 0.16                      | 81.31    | 1.89            |
| N3     | 4.61              | 13.93 | 0.33                      | 117.03   | 2.68            |
| S1     | 2.95              | 14.21 | 0.21                      | 72.06    | 1.65            |
| S2     | 2.68              | 10.29 | 0.26                      | 36.35    | 1.93            |
| S3     | 11.69             | 40.30 | 0.29                      | 42.20    | 0.21            |
| E1     | 8.13              | 24.01 | 0.34                      | 61.57    | 1.78            |
| E2     | 1.51              | 7.71  | 0.20                      | 47.16    | 0.54            |
| E3     | 0.76              | 10.10 | 0.08                      | 89.56    | 1.19            |
| W1     | 8.97              | 24.93 | 0.36                      | 79.43    | 0.93            |
| W2     | 5.20              | 26.00 | 0.20                      | 64.22    | 0.56            |
| W3     | 125.24            | 134.67| 0.93                      | 80.02    | 1.60            |

Fig. 7 Availability of the Internet Service (ISPs and SIMEBOX)
conclude that ISPs inflate the availability levels because SIMETBox occasionally went offline in many squares, which may be the reason they report low availability levels. In some cases, the availability showed by ISPs and SIMETBox is the opposite, as for Square N1, where it is higher for SIMETBox. We consider that SIMETBox gives the right figure for this square since it is an independent measurement infrastructure. When a square is out of service, in no way packets sent by SIMETBox would receive a response. Therefore, after analyzing availability metrics for more than 2 years and talking to maintenance and onsite measurement personal, we conclude that ISPs and SIMETBox complement each other.

4.1.3 Performance

Table 6 shows results for performance metrics for the 15 selected squares during 2015. It raises compelling observations, among which some deserve to be highlighted. Usage varies considerably among squares both in terms of the number of users and aggregate throughput. On the other hand, the average throughput per user is kept under 400 Kbps for most squares, with one exception (Square C3). This level of usage cannot be simply associated with low performance caused by an underprovisioned network because this behavior also holds for underutilized squares. The latter means that some squares have an average low
service utilization that can be observed by the number of connected users and the low average throughput per user.

Figure 8 shows the evolution of the number of user accesses and QoS metrics in digital squares, averaged by region and the global average, for five periods of 3 months covering the whole year of 2015 (starting in October 2014 and ending in January 2016). We can observe in Fig. 8a that after an initial period when accesses increased in all regions, the average user count started to decrease. The number of active squares reached the final number (119, with one blocked), and service availability stabilized at a higher level only in the second semester of 2015, as shown by Fig. 8e, f, respectively. In other words, after a period of excitement when users were exposed to the Wi-Fi service and faced unstable network conditions, the user base entered in a steady state.

Figure 8b shows the aggregate throughput entering the digital squares, dominated by the traffic in the Center that suffered a decrease in the period from June to August and resumed the increase in the following periods. The decrease in throughput follows the same behavior of the user count shown in Fig. 8a. The East region had a steady increase during the whole period, and the North, South, and West regions increased initially but showed a trend to achieve a steady state in the end. The RTT (Fig. 8c) and packet loss (Fig. 8d) showed a general trend to decrease in all regions, except for some spikes caused mostly by service unavailability (Fig. 8e) and ended with much lower values. The packet loss is of particular relevance because it stabilized at a considerably high level of 0.94%. We identified that the critical factor affecting this performance level is service unavailability because when the square is not connected to the Internet, SIMETBox frequently reported losses. In other words, the measurement system is negatively affected by the measured system. Figure 8f shows that the leading cause of reduced performance levels in the second period (April 2015) is the increase in the number of squares from about 80 to 120. In the following periods, the performance showed a general trend to improve considering throughput, RTT, packet loss, and availability.

The average RTT varies considerably among squares, but most of it lies in the tens of milliseconds range. The maximum and minimum average RTT were 356.63 ms and 3.09 ms, respectively. Average RTT is within a reasonable range for many applications, even though it could be lower because the delay budget for end-to-end applications is higher than what has been measured (until an Internet Exchange). The average packet loss for most squares is in the range of 0.5% and 2%, with 4% being the maximum. The effect of RTT (two-way delay) and packet loss is further analyzed in more detail. Figure 9 depicts results for the delay in Square C3 and Square N1, in four different ways: week by week comparing delay with the number of connected users (a and b), hourly daily view (c and d), histogram and ECDF of delay (e and f) and the delay variation according to the level of service usage (g and h). The critical point highlighted by Fig. 9 is that while the number of connected users negatively influences the delay, the precise way it occurs is not always obvious, or it may not be understood only with local wireless square data disregarding the wired network. ISPs do not provide information about their backbone network and links connecting it to the squares, which would be useful to have a precise view of the delay budget, i.e., whether it is caused by the wireless network, by the wired network, or by both. Crossing this data with our on the spot video streaming measurements helped us to build the understanding that the primary source of poor performance is the wired network part since it happens even with low service utilization (Sect. 4.2). Fundamentally, it shows that while access control must be enforced, dynamic thresholds should be used according to instant QoS metrics, rather than the static, predetermined maximum number of users. Also, if dynamic access control cannot be enforced, static one should not be used in order to unnecessarily deny user service while the network still can accept connections with reasonable QoS.

As reported in Table 6, the delay is considerably higher than it would be expected, which is some squares such as C6 may cause annoyances for its users. However, Fig. 9a, b show that the number of connected users does not necessarily influence the weekly average since the correlation is low (0.13 and 0.24, respectively). In other words, delay variations do not necessarily follow the same pattern as the number of users increases and decreases. When daily averages for the whole year are compared, results differ significantly, as shown by Fig. 9c, d, where delay increases together with the number of users during business hours and decreases at night and early hours of the day. The same happens with the behavior of the aggregate incoming bitrate curve (throughput, not shown in charts), which follows both the number of users and delay. Figure 9e, f show the RTT histogram and ECDF for the two selected squares. For Square C3, 60% is less or equal to 4 ms, 80% is less or equal to 16 ms, and 90% is less or equal to 30 ms. On the other hand, for Square N1, 60% is less or equal to 8 ms, 80% is less or equal to 50 ms, and 90% is less or equal to 100 ms. A comparative analysis between Fig. 9a–f reveals that Square C3 has a higher weekly and hourly
mean delay than Square N1, but a lower median delay, affecting different users in different ways.

Figure 9g, h show how delay varies according to service utilization, which allows us to observe the effect of the policy of not blocking users after the predefined maximum.
is reached. For Square C3, it is clear that the delay becomes much worse after 100% of the predetermined users are connected (250). For Square N1, there is a peak of delay at about 110% (55), from where it abruptly decreases and increases again. That behavior does not mean that the service quality improves as the number of connected users increases.

Fig. 10 Packet loss for two selected squares (Center and North) in 2015
increase. Instead, it means that beyond 100%, the data is not statistically reliable because there were only three sharp peaks during the period where the number of connected users surpassed the maximum allowed (50).

Packet loss must also be kept low in order to provide acceptable user QoE. Again, we are interested in understanding the effect of the increased number of connected users in the packet loss rate. Unlike delay, packet loss does not have a linear relationship with the number of users. Figure 10 depicts the weekly average packet loss for 2015, the hourly average, the histogram, and ECDF and the behavior of packet loss according to service utilization. Although Fig. 10a reveals peaks up to 12% for Square C3, it is more an exception than a rule. An interesting observation is that for weeks 7 to 12, 17 and 37, where packet loss was high, in the subsequent week, the number of users decreased significantly (not shown in charts). In other words, for Square C3, whenever users face poor network conditions for some time, they give up using the Wi-Fi connection for a while and resume using it afterward. For Square N1, Fig. 10b shows that the packet loss rate kept oscillating between 1% and 4% up to week 27 and stabilized close to zero after week 31, after the period with missing data. Looking closer into the data of our dataset and correlating loss and availability, it became clear that most of the loss is due to service unavailability. Up to week 27 (end of June 2015), there were many days where loss amounted to 60%, 80%, or even 100%. On average, the packet loss lay between 1% and 4% because on many days had 0% loss.

Figure 10c, d also reveal a low correlation between packet loss and the hours of the day, as shown for delay. Even though there is a trend of increasing packet loss during business hours, it also oscillates during the day. For Square C3, the packet loss rate curve forms a finger shape, and for Square N1, it sometimes follows the increase of connected users, and sometimes it does the opposite. The latter can be observed between 8 am, and 10 am. Figure 10e, f show that most of the time, packet loss is zero (83% for Square C3 and 72% for Square N1). As for the behavior of packet loss according to service utilization, Fig. 10g reveals a slight trend for a gradual increase, which is expected since the higher the number of users in a square, the higher the probability of congestion and packet loss to occur. On the other hand, for Fig. 10h, the behavior is less conclusive, since it shows a sharp fall after 80% (around 40 users). This behavior shows that packet loss for this square is not related to the number of connected users and corroborates the conclusion in the previous paragraph that the main offender is service unavailability.

The combined analysis of Figs. 9 and 10 reveals that when the network reaches its maximum capacity, even though enforcing admission control may deny access, the accepted users obtain higher quality service levels. However, the critical challenge is that in order to be precise about when starting to deny access to users, one needs to know the instant dynamic maximum number of users and not a predetermined theoretical one.

4.2 Video streaming

After observing the annual macro view of network-related performance metrics in Sect. 4.1, this session goes into a detailed view of user-related video streaming experience metrics. First, we analyze QoS metrics collected together with video measurements in order to broaden our understanding of the video streaming results. Please recall that these results come from our measurements. Due to space limitations, here we present results for only two (out of five) squares, namely Square E4 in East, and Square W4 in West, which reveal the worst and the best results, respectively.

Figure 11a shows that Square E4 had only 10% of its maximum predefined users on Sunday, due to the measurement time (7 am), unlike Tuesday, when it reached up to 40% because the measurements took place at 1 pm. For Friday, data is missing for 1200 Kbps and 840 Kbps (in November 2015) due to a failure in data capture. Generally speaking, Square E4 had a far lower utilization ratio compared to Square W4 (Fig. 11b), which should provide higher QoS and QoE. However, we found exactly the opposite. These inconsistent results corroborate with the conclusions for Figs. 9 and 10, showing that dynamic, rather static, access control should be enforced. Alternatively, in other words, there is a mismatch between the maximum predetermined number of users and the network resource provisioning of the digital squares, which is only revealed by our analysis. This conclusion is also corroborated by the results of Figs. 12, 13 and 14.

Figure 11(c, d) reveal a significant difference between the two squares when it comes to the per-user throughput, higher for Square W4 compared to Square E4. A joint analysis of service utilization and throughput tells us that Square W4 achieved higher throughputs even with higher service utilization, which might seem counterintuitive but can be explained by the network infrastructure installed by the two ISPs. Also, for different video bitrates on the same day of the week, there is a variation because they happened on different days of the month and year. The same behavior was observed for the other squares in the Center, North, and South.

Figure 11(e–h) shows a considerable difference between Square E4 and Square W4 in terms of delay (RTT) and packet loss, which reveals a noticeable case of wired network underprovisioning in Square E4 and the East Region in general that we found in our measurements. For service
utilization, around 30%, RTT goes up to 500 ms and packet loss to 10%. Average RTT reached a peak of almost 1 s, while packet loss was up to almost 30% for Square E4, which is unacceptable. It is notorious that even with low service utilization (Fig. 11a) and low throughput (Fig. 11c), the values of RTT (Fig. 11e) and packet loss (Fig. 11g) were high. Also, considering that the RTT and packet loss were lower on Sundays, and other squares in the East Region also face poor performance with the same characteristics, make us conclude that the wireless network is not to be blamed, but a sever underprovisioning in the wired network (backbone and link to the square). The low RTT and packet loss on Sundays rule out the possibility of a simple Wi-Fi misconfiguration since the same behavior

Fig. 11 Network performance metrics during video experiments (East and West)
was observed in November 2015 and January 2016. On the other hand, for Square W4, the RTT was kept under 50 ms (Fig. 11f) and packet loss under 2% (Fig. 11h) in most cases.

Next, we explore video start time, interruptions, and throughput, supported by the network performance results of Fig. 11. Figure 12 shows the video playback time as a percentage of the total video duration. For Square E4, one can observe that for most days of the week, except Sunday, there is a significant delay before the video starts. The same behavior, but somewhat softer, was found in the North, South, and Center regions. The opposite happened in Square W4, where the playback start time was about 2% (or 60 s). The network QoS results corroborate the video QoE results, i.e., higher throughput, lower delay, and packet loss lead to improved video quality.

Figure 13 shows the number and duration of interruptions during playback, also known as freezing, that indeed harms video quality as perceived by the user. Figure 13a shows the worst case for the number of interruptions, where the average value varied between two and four, which was slightly lower on Sundays. These results hold for the three different bitrates (1200 Kbps, 840 Kbps, and 420 Kbps) and the two periods (November 2015 and January 2016). It can be noticed in Fig. 13b that there is not always a correlation between the average number of interruptions and the video bitrate. In other words, the intuition says that the higher the bitrate, the more susceptible to interruptions will be the video because it occupies a larger share of the bandwidth. In practice, it can be observed that, for example, the 840 Kbps video experienced more interruptions on average than the 1200 Kbps.
for Sunday, Tuesday, Thursday, and Friday. However, the explanation is straightforward. Firstly, the two experiments were executed in sequence, and the network conditions vary significantly in small timescales sometimes. Secondly, the TCP protocol tries to treat all flows with fairness, but as the IP protocol provides only the best effort service model, frequently simultaneous flows obtain different levels of QoS. Thirdly and more importantly, we repeated each experiment 30 times and computed confidence intervals. As the confidence intervals overlap in all cases (observed in the error bars), statistically, the results do not differ from each other.

The duration of interruptions is also significantly high, with a 30 s peak (on average) for the 1200 Kbps video on Tuesday, shown in Fig. 13c. The QoS results in Fig. 11(a, c, e, g) explain the reason for the poor performance on Tuesday. On the other hand, among the five squares, Square W4 achieved the best results, where the average number of interruptions was below two (Fig. 13b), of which most lasted less than 1 s (Fig. 13d). The explanation is also a straightforward effect of network QoS shown by Fig. 11(d, f, h).

The received bitrate (goodput) complements and corroborates the previous results. Figure 14a shows that whenever the average bitrate reaching the receiver is consistently lower than the transmitted bitrate, the video quality is severely affected. For example, on Tuesday, the average goodput was about 200 kbps for a video coding bitrate of 1200 Kbps. On the other hand, the best result was observed in Square W4 (Fig. 14b), where the bitrate sent and received was equivalent for most cases.

### 4.3 Event identification

After analyzing network QoS and video streaming performance in the previous sections, we now turn to observe the behavior of users when significant events occur. The analysis of general events that potentially may happen in multiple squares shows that there was some increment in the number of users in a certain number of squares compared to the previous month, with an emphasis on the launch of Pokémon GO that caused an increase in 77 squares. However, contrary to our expectations, the increase was not significant since it was not higher than the standard deviation. We chose the previous month as the baseline after analyzing different periods. For example, comparing with the same period of the previous year yields completely uncorrelated results, and the next month was just similar to the previous month. Also, when forecasting future behavior, next month’s data will not be available.

This behavior holds for the five general events analyzed, which are the launch of Pokémon GO in Brazil in August 2016, the New Year celebrations of 2015 and 2016, and Carnival 2015 and 2016 (both in February). The New Year
2016 registered only a small increase in 17 squares. This phenomenon is illustrated in Fig. 15 that presents the average and standard deviation of connected users in all South Region squares, where an increase was observed during the launch period of Pokémon GO and the previous month. In other regions, the behavior was similar. It can be observed that these events did not cause a significant increase in the number of connected users, maybe by their decentralized nature that diluted the potential impact. Also, we compare the 2015 and 2016 editions of the Cultural Overnight, a yearly event in São Paulo with multiple attractions in public venues that spans 24 h in a selected weekend. Figure 16 depicts the number of connected users for five squares that hosted this event in both years, all of them in the Center Region: C4, C5, C6, C7, and C8.

Figure 16 reveals a lack of pattern for people crowding. In some cases, it presents normal behavior (Fig. 16b, g, h), while in other cases the event caused an increase in the number of users during the night (Fig. 16d–f, i, j), but in two cases it decreased (Fig. 16a, c). Secondly, the user behavior followed a similar pattern for 2015 and 2016 for all squares, with small variations. For example, in Square C7, the average usage pattern was similar to the previous month. On the other hand, in 2016, it raised to the upper bound of the standard deviation showing a noticeable increase. An unexpected decrease in the number of connected users was observed for Square C4 and Square C6. In some squares, the Cultural Overnight literally engaged the users “overnight” and changed the typical bell-like usage pattern, observed in Square C6 and Square C8. Finally, a significant increase of almost five times occurred during the night and early hours of the day for Square C8 in 2016.

This behavior leads us to conclude that similarly to generic events (e.g., New Year, Carnival), significant events that happen in multiple squares, such as the Cultural Overnight in São Paulo as measured for five squares in 2 years, do not necessarily show correlation with the number of Wi-Fi users.

When it comes to events that happened in specific squares, substantial differences arise. Figure 17 depicts the Wi-Fi access behavior of users during events that happened in specific squares, which cause an increase in the number of users while the event is taking place and sometimes a decrease in different hours. The celebration of the Anniversary of São Paulo in 2016 caused an increase in the number of users in Square E6 (Fig. 17a) and Square C7 (Fig. 17b). Even though the attractions (concerts) were at night, users started to crowd many hours before. Square C7 had an increase in the number of users from noon up to 10 pm.

The celebration of the Chinese New Year in 2016 brought to Square C9 (a well-known Japanese and Chinese stronghold in Brazil) a high number of users during the whole day, up to 2.5 times the regular number (Fig. 17c). The same happened during Street Samba in Square S6 on a Sunday (Fig. 17d). The celebration of the Independence of Brazil in 2016 brought to Square S3 (Independence Park), a significant number of additional users during the morning, which is when the parade and other attractions take place (Fig. 17e). After the celebrations were over, users left the place at a higher rate than the normal behavior. The Jim Beam History Fest 2016 was held in Square C4 from 3 pm on (Fig. 17f), and at this time, the user crowding at the place started to increase and reached a high number at a time when it usually would start to decrease. Remarkable is that the same square received a lower number of users in the hours before the event started, showing that this event caused a very atypical behavior during the whole day. Should an unexpected event occur in such a square, authorities could promptly detect it by the atypical behavior of Wi-Fi users.

5 Discussion

We performed a comprehensive study of the Municipal Wi-Fi Network in São Paulo, collecting data for more than 2 years. Our analyses taught us valuable lessons that might improve the public policy of providing free Internet connection to the citizens in São Paulo and other cities around the world.

The joint analysis of data from network usage and video streaming measurements makes us conclude that some squares suffer from severe wired network underprovisioning that causes general poor performance showed by low throughput, high delay, and high packet loss rate. The analysis of Fig. 9 revealed that while the number of connected users negatively influences the delay, the precise way it occurs may not be understood by the configuration of the Wi-Fi network installed in the squares. On the other hand, Fig. 11 reveals that poor performance also occurs in squares with a small number of users (service utilization of 10% to 20%), and in some days (Sundays, notably), the performance increases. This fact rules out the possibility of a simple Wi-Fi misconfiguration since the same behavior was observed in different periods.

This finding provides strong evidence that the implementation of Municipal Wireless Networks requires in-depth management of the network infrastructure, both wired and wireless. Moreover, that is valid even when the operation is outsourced to private ISPs, as in São Paulo.
The lack of topology, technology, configuration, and management data from ISP backbone network and links connecting it to the squares, leaves public authorities with no mechanism to verify the performance in the Wi-Fi access network. However straightforward this recommendation might be, authorities must exert higher control levels over the wired and wireless networks. The monitoring of the end-to-end QoS and QoE as apparent to the users is not enough to enforce higher quality. Also, better measurement tools are needed to understand the behavior of the wired network. We used Pathload and Pathrate, but they did not provide reliable results and therefore were not useful to bring new evidence on the quality of provisioning and operations of the ISPs’ wired networks.

In the following sub-sections, we discuss the most important findings of the three categories of analyses of the Municipal Wireless Network in São Paulo we show in this paper.

5.1 Network usage

The results presented in Sect. 4.1 confirm Hypothesis #1, leading us to conclude that a large-scale public Wi-Fi network should adopt measurement-based dynamic admission control for limiting the number of users.
connected to Access Points when, and only when, accepting them would cause poor performance for all users. If dynamic mechanisms are not possible, then no admission control is preferred than static admission control based on a priori estimates that incorrectly guess the number of users who connect to the Wi-Fi AP, causing either underprovisioning or overprovisioning. If the static maximum number of users is too low, the system will mistakenly block users when there are still plenty of network resources, just contradicting the very nature of the public policy and public investment to the benefit of the people. On the other hand, if the static maximum number of users is too high, additional users will be accepted when system resources are already depleted, thus causing poor quality of experience for all users. The WiFiLivreSP program chose not to enforce admission control, and our performance results show that for most squares, the static maximum number of users was too low, indicating that a bad public policy was avoided. On the other hand, for some squares, admission control should be enforced as users faced poor performance due to the lack of it.

In practice, network administrators usually maintain the factory setting, which may cause performance problems. Also, as in the WiFiLivreSP program, since no admission control has a similar effect of setting up an excessively high number of users, network administrators typically set up a very low one to be on the safe side, thus unnecessarily blocking user access. If either case, the current state of practice is to use static admission control. However, as it becomes clear from our data if admission control is to be used, dynamic measurement-based mechanisms should be considered, rather than static ones. In this case, the critical challenge is to set up and manage this measurement infrastructure and to choose QoS metrics to enforce the admission control that both can be measured and provide performance levels that, on average, outperform the static option. Also, literature frequently focuses on admission control for IEEE 802.11 wireless networks, using small-scale testbeds and simulations. We showed performance metrics that combine both wireless and wired network segments, which may be used by future work to understand and provide solutions to real-world problems involved in setting up and managing dynamic techniques. Besides, the classification of squares into different categories according to their function (sport & leisure, passage, permanence, and religious) can be effectively used to improve network management capabilities, such as configuring different admission control and traffic shaping policies.

Other conclusions that remained relatively invariant for the whole period analyzed can be summarized as:

- Service availability is the fundamental cause of poor user experience since the absence of service causes many inconveniences and lack of trustworthiness. We confirmed this assumption based on two distinct measurements (ISP and SIMETBox) crossing different performance metrics. Also, the lack of reliability in data collection is yet an open challenge.

- The observed per-user throughput is usually not as high as the 512 Kbps target, which is mostly due to the applications that do not demand this bitrate over a long period, except for video streaming. Results from Sect. 3.1 reveal a low correlation between throughput per user, RTT, and packet loss, which represents the findings for all squares (not shown in this paper due to space constraints). In other words, an increase in the network bandwidth used by one user, or a couple of users, does not necessarily harm the other users. Again, this provides evidence that a well-managed Municipal Wireless Network can provide high quality levels to its users.

- The delay (RTT) does not necessarily follow the increase in the number of connected users, where the correlation between these two variables is close to zero. While there is a trend for the RTT to increase as the service utilization increases (Fig. 9), this is more likely to be caused by underprovisioning issues because it starts before 100% is achieved. Up to 60% of measurements reported RTT values below 15 ms.

- The packet loss is highly influenced by service unavailability due to the characteristics of the SIMETBox measurement system. This fact corroborates our previous comment that improved measurement tools are needed to provide higher QoS and QoE levels in Municipal Wireless Networks.

### 5.2 Video streaming

The results presented in Sect. 4.2 confirm Hypothesis #2, leading us to conclude that in a large-scale public Wi-Fi network, no traffic shaping allows users to use high bandwidth-demanding services, such as video streaming, whenever the network is underutilized. Traffic shaping aims at guaranteeing the reasonable quality of experience, preventing high demanding users from consuming system resources beyond their fair share. However, static and strict control of system resources may block users from accessing services. We performed on the spot active measurements for video streaming in five selected squares and found out that there is a considerable variation in the perception of the quality of experience for single users trying to watch videos over Wi-Fi.

We chose to measure QoE metrics for video streaming based on buffer management that considers interruption as the principal offender for the user experience. As expected,
we observed a tight connection between network QoS metrics (throughput, delay, packet loss) and video QoE that confirms the intuition. For example, service utilization influences the throughput, delay, and packet loss, which influence video reception rates, and consequently, cause video interruptions. Since our performance analysis study used active measurements, it might have caused service degradation for other users connected to the Wi-Fi at the same time. Practitioners may use Software-Defined Networks (SDN) to enhance the performance of video streaming at times where the digital square is underprovisioned to provide the throughput for adequate QoE [15].

Even squares having typical network conditions that are not appropriate for video streaming can provide this service to their users on certain days of the week. For example, even in Square E4, the worst performance, proper results we observed on Sundays, as shown in Figs. 11a, 12a, 13a and 14a. In the best case for Square E4, the received video bitrate as close as 1200 Kbps for the higher quality video in November 2015 and matched 840 Kbps and 420 Kbps for the medium and lower quality video in January 2016. Based on our experience, we could derive an empirical RTT threshold of about 50 ms and less than 3% of packet loss for guaranteeing an acceptable QoE for video streaming. There is a high correlation between these values and playback start time below 2 s and only one interruption of at most 2 s.

These days the Internet traffic is dominated by video streaming [36], and even for the worthiest uses of a Municipal Wireless Network, such as learning, video is the preferred application [37]. However, even though our measurements showed evidence that the user QoE is reasonable in most cases (in four out of five squares), there are still issues to observe in the provisioning and configuration of large-scale public Wi-Fi networks. As we showed in this paper, high service availability and adequate provisioning of the wired network are keystones for the success of such networks.

5.3 Event identification

The results presented in Sect. 4.3 partially confirm Hypothesis #3, which states that the number of Wi-Fi users can predict people crowding in public places. Generic events that occur in multiple squares simultaneously (e.g., New Year, Carnival, Cultural Overnight) do not necessarily provide clues about any increase in the number of users compared to the previous month. On the other hand, when it comes to events that occur in specific squares (e.g., Anniversary of São Paulo, Independence of Brazil, Chinese New Year), the number of connected users increases compared to the past. This finding is very positive since unpredictable people gatherings tend to happen in specific places. This behavior is clear when comparing Figs. 16 and 17. So, this finding gives an insight into people’s behavior that might be very useful for public agencies in charge of law enforcement, traffic management, health care, and rescue, that might use it for dealing with unexpected people crowds, such as in atypical traffic jams and unplanned demonstrations.

We also analyzed the effectiveness of the Wi-Fi network in digital squares to hold events that cause people to gather while keeping adequate quality levels. The results of the network QoS metrics (not shown in this paper) for the time and duration of the events show that the network provisioning was adequate to support such activities since no discernible differences exist in per-user throughput, delay, and packet loss rate.

Future work may consider identifying and automatically classifying unexpected events based on patterns of abrupt increase or decrease in the number of users in a digital square. An important follow-up of this paper is the identification of general rules for people crowding considering a higher number of both planned and unplanned events. An extensive analysis of events can contribute to derive general models (e.g., machine learning) to identify and classify an event as soon as it starts. An important extension of this future work is the automatic identification of root causes of unplanned events, which requires the understanding of people crowding patterns and mobility patterns and associating them to information about what happened in the specific places in the city. Such models and root cause analysis may, in turn, be used to the development of a monitoring system for public authorities to be informed and able to act in case of emergencies detected by the behavior of users in digital squares.

Last by not least, given the serious moment we are living in 2020 with the coronavirus pandemic and the need for social isolation, counting the number of Wi-Fi users in digital squares may significantly help authorities to identify anomalous and uncalled for people aggregates.

6 Conclusions

The Municipality of São Paulo operates the WiFiLivreSP program since 2014, providing free Wi-Fi Internet access in 120 public spaces called digital squares. We performed a comprehensive study of this program collecting data for more than 2 years and analyzing: (a) the performance of the Wi-Fi network focusing on Quality of Service parameters for 119 digital squares in the city of São Paulo; (b) the Quality of Experience for video streaming over Wi-Fi measuring five squares of every region of the city, and;
(c) the suitability of the number of connected users to identify the occurrence of events in the squares.

We learned valuable lessons that might be useful to improve the public policy of providing free Internet connection in a large-scale public Wi-Fi network, such as: (a) no admission control is preferred to static admission control based on a priori estimates, which can unnecessarily deny service to many users; (b) no traffic shaping allows users to use high bandwidth-demanding services, such as video streaming, whenever the network is underutilized; (c) The number of users can identify people crowding in public places for events that occur in specific squares, rather than generic events that occur in multiple squares at the same time.

Such large-scale and long-term studies of Municipal Wireless Networks are difficult to find in the literature, if not impossible at all. In this paper, due to scope and size limitations, we present data analysis of user access, video quality, and event identification. However, a variety of different techniques and approaches exist for dynamic access control and automatic event identification, such as time series forecast and machine learning, which are left for future work. For example, a foreseeable next step encompasses the detection of unplanned events—such as traffic jams caused by a car crash or undetected demonstrations—that may require the attention of government agencies in charge of traffic control, law enforcement, healthcare, and rescue.

Acknowledgements Authors would like to acknowledge the institutions that contributed by providing support to the execution of this work, namely UFABC, CNPq, and the Municipality of São Paulo.

Authors’ Contributions JR collected datasets, preprocessed data, and generated results for network usage and video streaming. AT generated results for event identification. CK supervised JR and AT and worked on manuscript writing. DC worked on manuscript writing and revision.

Funding This research was funded by the Municipality of São Paulo, Brazil.

Availability of Data The data that support the findings of this study are available from the corresponding author, CK, upon request.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no competing interests.

References

1. Agrawal, S., & Agrawal, J. (2015). Survey on anomaly detection using data mining techniques. Procedia Computer Science, 60, 708–713.

2. Aruba. Aruba 510 series. Retrieved April 2020, from https://www.arubanetworks.com/products/networking/access-points/510-series.

3. Backholm, A., & Luna, M. (2015). Distributed implementation of dynamic wireless traffic policy. U.S. Patent 9,077,630.

4. Baik, E., Pande, A., Stover, C., & Mohapatra, P. (2015). Video acuity assessment in mobile devices. In IEEE INFOCOM 2015.

5. Balachandran, A., Sekar, V., Akella, A., Seshan, S., Stoica, I., & Zhang, H. (2013). Developing a predictive model of quality of experience for internet video categories and subject descriptors. In ACM SIGCOMM 2013.

6. Bar, F., & Park, N. (2005). Municipal Wi-Fi networks: The goals, practices, and policy implications of the US case. Communications and Strategies, 61, 107–126.

7. Bazzi, A., Diolaiti, M., & Pasolini, G. (2005). Measurement based call admission control strategies in infrastructured IEEE 802.11 WLANs. In IEEE 16th international symposium on personal, indoor and mobile radio communications.

8. Biswas, S., Backed, J., & Wong, E. (2015). Large-scale measurements of wireless network behavior. In ACM SIGCOMM 2015.

9. Burger, V., Seufert, M., Kaup, F., Wichtlhuber, M., Hausheer, D., & Tran-Gia, P. (2015). Impact of Wi-Fi offloading on video streaming QoE in urban environments. In IEEE international conference on communication.

10. Cisco. Cisco Aironet 4800 access points. Retrieved April 2020, from https://www.cisco.com/c/en/us/products/wireless/aironet-4800-access-points.

11. Cisco. (2017). Cisco visual networking index: Forecast and methodology, 2016–2021. White paper.

12. Fraser, E. (2010). A postmortem look at citywide Wi-Fi. Journal of Internet Law, 14(2), 1.

13. Ghasempour, Y., da Silva, C. R., Cordeiro, C., & Knightly, E. W. (2017). IEEE 802.11 ay: Next-generation 60 GHZ communication for 100 Gb/s Wi-Fi. IEEE Communications Magazine, 55(12), 186–192.

14. Ghosh, A., Jana, R., Ramaswami, V., Rowland, J., & Shankaranarayanan, N. K. (2011). Modeling and characterization of large-scale Wi-Fi traffic in public hot-spots. In IEEE INFOCOM 2011.

15. Go, S. J., Festin, C., & Tan, W. (2019). An SDN-based framework for improving the performance of underprovisioned IP Video Surveillance networks. Journal of Network and Computer Applications, 132, 49–74.

16. Heer, T., Hummen, R., Viol, N., Wirtz, H., Götz, S., & Wehrle, K. (2010). Collaborative municipal Wi-Fi networks—Challenges and opportunities. In 8th IEEE international conference on pervasive computing and communications workshops (PERCOM workshops).

17. Ishikawa, N., Ohishi, Y., & Maeda, K. (2019). Nulls in the air: Passive and low-complexity QoS estimation method for a large-scale Wi-Fi network based on null function data frames. IEEE Access, 7, 28581–28591.

18. Jassem, H. C. (2010). Municipal Wi-Fi: The coda. Journal of Urban Technology, 17(2), 3–20.

19. Juluri, P., Tamarapalli, V., & Medhi, D. (2016). Measurement of quality of experience of video-on-demand services: A survey. IEEE Communications Surveys & Tutorials, 18(1), 401–418.

20. Kaiser, M. S., Lwin, K. T., Mahmud, M., Hajializadeh, D., Chaipimonplin, T., Sarhan, A., et al. (2017). Advances in crowd analysis for urban applications through urban event detection. IEEE Transactions on Intelligent Transportation Systems, 19(10), 3092–3112.

21. Kaiser, T. W., Verbrugge, S., Van der Wee, M., & Colle, D. (2017). An overview of different business models for public Wi-Fi and their implications on indirect revenue. In 28th European regional ITS conference.
22. Kim, S., & Kim, C. (2018). XMAS: An efficient mobile adaptive streaming scheme based on traffic shaping. *IEEE Transactions on Multimedia*, 21, 442–456.

23. Kim, Y. L. (2018). Seoul’s Wi-Fi hotspots: Wi-Fi access points as an indicator of urban vitality. *Computers, Environment and Urban Systems*, 72, 13–24.

24. Kontokosta, C. E., & Johnson, N. (2017). Urban phenology: Toward a real-time census of the city using Wi-Fi data. *Computers, Environment and Urban Systems*, 64, 144–153.

25. Li, H., Zhu, H., & Ma, D. (2018). Demographic information inference through meta-data analysis of Wi-Fi traffic. *IEEE Transactions on Mobile Computing*, 5, 1033–1047.

26. Li, S., Da Xu, L., & Zhao, S. (2018). 5G internet of things: A survey. *Journal of Industrial Information Integration*, 10, 1–9.

27. Mekki, K., Bajic, E., Chaxel, F., & Meyer, F. (2018). A comparative study of LPWAN technologies for large-scale IoT deployment. *ICT Express*, 5, 1–7.

28. Mukherjee, M., Shu, L., & Wang, D. (2018). Survey of fog computing: Fundamental, network applications, and research challenges. *IEEE Communications Surveys & Tutorials*, 20(3), 1826–1857.

29. Nafaa, A. (2005). Provisioning of multimedia services in 802.11-based networks: Facts and challenges. *IEEE Wireless Communications*, 14(5), 106–112.

30. Neill, D. B., & Wong, W. K. (2009). Tutorial on event detection. In 15th ACM SIGKDD international conference on knowledge discovery and data mining (KDD 2009).

31. Oliveira, L., Obrazcka, K., & Rodriguez, A. (2016). Characterizing user activity in Wi-Fi networks: University campus and urban area case studies. In 39th ACM international conference on modeling, analysis and simulation of wireless and mobile systems.

32. Picco-Schwendener, A., Reinhold, H. J., & Cantoni, L. (2019). Being smart with data: Automatic classification of users of a free municipal Wi-Fi network (MWN) using a two-tiered classification tree. In *ACM 20th annual international conference on digital government research*.

33. Picco-Schwendener, A., Reinhold, H. J., & Cantoni, L. (2017). Usage practices and user types of a municipal Wi-Fi network: The case of Wi-Fi Lugano. In *10th international conference on theory and practice of electronic governance*.

34. Pinson, M., & Wolf, S. (2004). A new standardized method for objectively measuring video quality. *IEEE Transactions on Broadcasting*, 50(3), 312–322.

35. Potts, J. (2014). Economics of public Wi-Fi. *Australian Journal of Telecommunications and the Digital Economy*, 12(1), 1–9.

36. Sandvine. (2019). The global internet phenomena report.

37. Scagnoli, N., Choo, J., & Tian, J. (2019). Students’ insights on the use of video lectures in online classes. *British Journal of Educational Technology*, 50(1), 399–414.

38. Schauer, L., Werner, M., & Marcus, P. (2014). Estimating crowd densities and pedestrian flows using Wi-Fi and Bluetooth. In *11th international conference on mobile and ubiquitous systems: Computing, networking and services*.

39. Schmidt, T., & Townsend, A. (2003). Why Wi-Fi wants to be free. *Communications of the ACM*, 46(5), 47–52.

40. Serral-Gracia, R., Cerqueira, E., Curado, M., Yannuzzi, M., Monteiro, E., & Masip-Bruin, X. (2010). An overview of quality of experience measurement challenges for video applications in IP networks. In *International conference on wired/wireless internet communications (WWIC 2010)*.

41. Shafi, M., Molisch, A. F., Smith, P. J., Hausteim, T., Zhu, P., Silva, P. D., et al. (2017). 5G: A tutorial overview of standards, trials, challenges, deployment, and practice. *IEEE Journal on Selected Areas in Communications*, 35(6), 1201–1221.

42. Tang, X., Tan, Z., Hu, S., & Geng, H. (2018). Evaluating spatial service and layout efficiency of municipal Wi-Fi facilities for Smart City planning: A case study of Wuhan city. *China. Socio-Economic Planning Sciences*, 65, 101–110.

43. Tang, X., Tan, Z., Hu, S., & Geng, H. (2019). Evaluating spatial service and layout efficiency of municipal Wi-Fi facilities for Smart City planning: A case study of Wuhan city, China. *Socio-Economic Planning Sciences*, 65, 101–110.

44. Toch, E., Lerner, B., Ben-Zion, E., & Ben-Gal, I. (2019). Analyzing large-scale human mobility data: A survey of machine learning methods and applications. *Knowledge and Information Systems*, 58(3), 501–523.

45. Trudeau, P., & Hewlett Packard Development Co LP. (2013). Call admission control for Wi-Fi. U.S. Patent 8,374,622.

46. Wang, Z., Lu, L., & Bovik, A. (2002). Video quality assessment using structural distortion measurement. In *IEEE international conference on image processing (ICIP 2002)*.

47. Weppner, J., Bischke, B., & Łukowicz, P. (2016). Monitoring crowd condition in public spaces by tracking mobile consumer devices with Wi-Fi interface. In *ACM international joint conference on pervasive and ubiquitous computing (UbiComp 2016)*.

48. Xi, W., Zhao, J., Li, X.-Y., Zhao, K., Tang, S., Liu, X., & Jiang, Z. (2014). Electronic frog eye: Counting crowd using Wi-Fi. In *IEEE INFOCOM 2014*.

49. Yan, Y., Yang, Y., Meng, D., Liu, G., Tong, W., Hauptmann, A. G., et al. (2015). Event oriented dictionary learning for complex event detection. *IEEE Transactions on Image Processing*, 24(6), 1867–1878.

50. Yu, H., Cheung, M. H., & Huang, J. (2017). Cooperative Wi-Fi deployment: A one-to-many bargaining framework. *IEEE Transactions on Mobile Computing*, 16(6), 1559–1572.

51. Zhai, H., Chen, X., & Fang, Y. (2005). How well can the IEEE 802.11 wireless LAN support quality of service? *IEEE Transactions on Wireless Communications*, 4(6), 3084–3094.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Carlos Kamienksi received a B.S. degree in computer science from the Federal University of Santa Catarina (Florianópolis, Brazil) in 1989, a M.S. degree from the State University of Campinas (Campinas, Brazil) in 1994 and a Ph.D. in computer science from the Federal University of Pernambuco (Recife Brazil) in 2003. He is full professor of computer science at the Federal University of ABC (UFABC) in Santo André SP, Brazil, where currently he also holds the positions of head of the NUVEEM Strategic Research Group. His current research interests include Future Internet, Internet of Things, Smart Cities, Smart Agriculture, network softwarization and fog computing.
Juliano Ratusnei received a B.S. degree in computer science from the State University of Center-West (Guarapuava, Brazil) and a M.S. degree in information engineering from the Federal University of ABC (Santo André, Brazil) in 2016. He is currently a lecturer of programming languages and databases at the State Center for Technological Education Paula Souza. His main research interests include wireless networks and performance analysis.

Andrea Trindade received a B.S. degree in software engineering from São Paulo State Technological College (Santos, Brazil) and a M.S. degree in information engineering from the Federal University of ABC (Santo André, Brazil) in 2018. She is currently a lecturer at the State Center for Technological Education Paula Souza working with distance learning. Her current research interests are in event identification and time series modeling.

Dave Cavalcanti received a B.S. degree in electrical engineering from the Federal University of Pernambuco (Recife, Brazil) in 1998, a M.S. degree in computer science from the Federal University of Pernambuco (Recife, Brazil) in 2001 and a Ph.D. in computer science and engineering from the University of Cincinnati in 2006. He is currently a Principal Engineer at Intel Corporation. His key interests are next generation wireless connectivity and networking technologies and their applications in autonomous, time-sensitive systems. He develops Wireless Time-Sensitive Networking (TSN) technologies over next generation 802.11 and 5G networks. He currently serves as the chair of the Wireless TSN working group in the Avnu Alliance.