

# LOAD FORECASTING IN WIFI ACCESS POINTS OVER THE LTE NETWORK

Hugo Marques<sup>1</sup>, Pedro M. B. Torres<sup>2</sup>, Paulo Marques<sup>3</sup>, Rogério Dionísio<sup>4</sup>, Jonathan Rodriguez<sup>5</sup>

<sup>1,2,3,4</sup> Instituto Politécnico de Castelo Branco, Castelo Branco, Portugal

<sup>3,4</sup> ALLBESMART, Avenida do Empresário, Centro de Empresas Inovadoras 1, Castelo Branco, Portugal

<sup>5</sup> Instituto de Telecomunicações, Campus de Santiago, Aveiro, Portugal

[hugo@ipcb.pt](mailto:hugo@ipcb.pt); [pedrotorres@ipcb.pt](mailto:pedrotorres@ipcb.pt); [paulomarques@ipcb.pt](mailto:paulomarques@ipcb.pt); [rdionisio@ipcb.pt](mailto:rdionisio@ipcb.pt)

---

**Abstract.** The concept of smart cities grew with the need to rethink the use of urban spaces based on the constant technological advances and respecting sustainability. Today the urbanism and the methodologies to think about the city are changing, as citizens want more access to digital information on almost everything. Therefore, cities need to be planned and equipped with infrastructures that enable connectivity between the citizens' devices and the digital information. This challenge raises technological problems, such as traffic management, in an attempt to guarantee fair network access to all users. Solutions based on wireless resource management and self-organizing networks are key when design the connectivity for these smart cities. This paper presents a study on forecasting the daily load of Wi-Fi city hotspots, taking also in consideration the weather conditions. This is particularly interesting to predict the network load and resource requirements needed to ensure proper quality of service is provided to the hotspot users. The study was performed in a Wi-Fi hotspot located in the city of Castelo Branco, Portugal. The results show the ARIMA model is capable of identifying and forecasting seasonality events for one week in advance including its capability to correlate the number of hotspot users with weather conditions.

**Keywords:** Smart Cities, Iot, Wi-Fi, Forecasting, First Section.

---

## 1. Introduction

Citizens want more access to digital information on almost everything, global mobile data traffic grew 63 percent in 2016, reaching an astounding 7.2 exabytes per month at the end of 2016[1]. Mobile data traffic has grown 18-fold over the last 5 years and the fourth-generation (4G), also known as Long Term Evolution (LTE), traffic accounted for 69% of mobile traffic in 2016. Although 4G connections represented only 26 percent of mobile connections in 2016, they accounted for 69 percent of mobile data traffic, while 3G connections represented 33 percent of mobile connections and 24 percent of the traffic. Furthermore, 60 percent of the total mobile data traffic was offloaded onto the fixed network through Wi-Fi or femtocells in 2016. In total, 10.7 exabytes of mobile data traffic were offloaded onto the fixed network each month in 2016. Almost half a billion (429 million) mobile devices and connections were added in 2016.

Smartphones accounted for most of that growth, followed by M2M (machine to machine) communications. By looking at 2016, on an average, a smart device generated 13 times more traffic than a non-smart device.

Mobile network (cellular) connection speeds grew more than 3-fold in 2016. Globally, the average mobile network downstream speed in 2016 was 6.8Mbps (up from 2Mbps in 2015). Also, according to [1], monthly global mobile data traffic is expected to increase up to 49 exabytes by 2021, and annual traffic is expected to exceed half a zettabyte. Mobile will represent 20 percent of total IP traffic by 2021. The number of mobile-connected devices per capita is expected to reach 1.5 by 2021. The average global mobile connection speed will surpass 20Mbps by 2021. The total number of smartphones will be over 50 percent of global devices and connections by 2021.

Given these numbers, solutions based on big data analytics and machine learning are expected to play a major role when planning communication networks. Forecasting network behaviour and load are, therefore, imperative for the short, middle and long term. Network load forecasting is already a technique used by mobile operators, where accuracy is of great importance for the self-organized networks (SON) functions achieve their goal.

Most forecasting methods use numerical techniques or Artificial Intelligence (AI) algorithms such as regression, neural networks, fuzzy logic, etc.

Depending on the time-series forecasting, the results can be categorized as seasonal, where data increases or decreases with season, trend on certain events and random for no specific cause. There are some studies in the literature that address the problem of forecasting and analysing time series to analyse and predict network load. In [2] the authors use Support Vector Machines (SVM) as a forecasting technique to address the problem of forecasting a function of the link load, such as the peak load or percentiles of its distribution, during an arbitrary time interval.

The Support Vector Regression are used with the purpose of link load prediction, investigating the impact of several parameters on the forecast accuracy and considering several real-work traces. In [3] the authors use the ARIMA model for forecasting Wi-Fi data network traffic values.

A six order Auto Regressive Integrated Moving Average (ARIMA) traffic model was obtained, which predicted traffic with relatively small mean square error values for an 18-day term. In [4] a novel stochastic model for the user throughput prediction in mobile networks is proposed. The model takes into consideration the most relevant sources of prediction inexactness, such as random phenomena (e.g., fast fading) or imprecise information (e.g., user location).

In this paper, the ARIMA forecasting method is used to evaluate its applicability in forecasting the average number of users per hour in a Wi-Fi city hotspot according to weather conditions. Figure 1 provides an aerial photo of the location (a city park) and identifies the location of the two access points installed.



Fig. 1. - a) "Violetas" park with two Wi-Fi hotspots; b) Wi-Fi hotspot

The remaining of the paper is organized as follows: Section II describes the methods and procedures; Section III presents the experimental results, and, finally; Section IV presents conclusions and future work.

## 2. Methods and Procedures

The goal is to forecast the average hourly number of Wi-Fi hotspot users for one entire week, using one month of collected historic data (including number of users and weather conditions).

The approach is to evaluate the ARIMA statistical method [6] for time series forecasting, to obtain a forecast model based on the collected dataset. The model is Autoregressive because it uses a dependent relationship between an observation and some number of lagged observations. Integrated because it differentiates raw observations (e.g. subtracting an observation from an observation at

the previous time step) to make the time series stationary. Moving Average because it uses the dependency between an observation and a residual error from a moving average model applied to lagged observations.

Typically, the ARIMA model can be specified as  $arima(p,D,q)$  function, where  $p$  indicates the non-negative integer indicating the degree of the non-seasonal auto-regressive polynomial.  $D$  is a non-negative integer indicating the degree of the non-seasonal differencing polynomial (the degree of non-seasonal integration) and  $q$  is a non-negative integer indicating the degree of the non-seasonal moving average polynomial.

The general forecasting equation can be expressed according equation 1, where  $e(t)$  are the innovation terms with  $P$  auto-regressive terms and  $Q$  moving average terms.

$$y(t) = c + a_1 \cdot y(t-1) + \dots + a_p \cdot y(t-P) + e(t) + b_1 \cdot e(t-1) + \dots + b_q \cdot e(t-Q) \quad (1)$$

Next section presents the results obtained.

### 3. Experimental Results

The dataset used was obtained by collecting user access and climate data during one month. The location is the one identified in Section 1 and depicted in Figure 1. The referring Wi-Fi hotspot uses LTE as backbone connection to reach the Internet and works with a battery set fed by a solar energy panel, the solution was provided by [5] under a pilot municipality project, “Castelo Branco Smart City”. It was therefore crucial to understand traffic and user patterns and their relationship with temperature to better design future deployments in other city locations. The trial was also important to identify and understand other seasonal events that occur in the city.



Figure 1 – Number of network accesses per hour

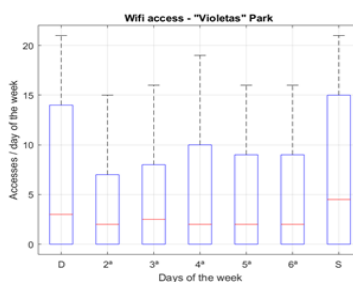


Figure 2 – Number of network accesses per day of the week

Another important aspect is the fact that on Wednesdays a higher number of users was observed than on other working days. This fact was credited to afternoon school breaks which always occur on this day of the week. It was therefore possible to preliminarily identify seasonality events and hotspot usage for this location. Using this dataset, the number of accesses in the first week of February was predicted based on the ARIMA forecast method.

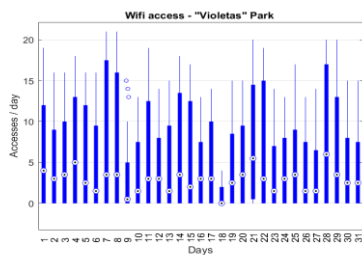


Figure 1 – Number of network accesses per day during the month of January 2017

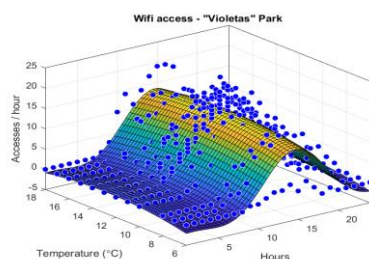


Figure 3 – Correlation between the number of accesses per hour and the temperature

The forecasting results, provided in Figure 7, show similar behaviour based on the past results, also predicting high affluence to the park during the peak hours, weekends and on the Wednesdays afternoons. These results provide a good indication of the ARIMA model suitability for predicting network usage.

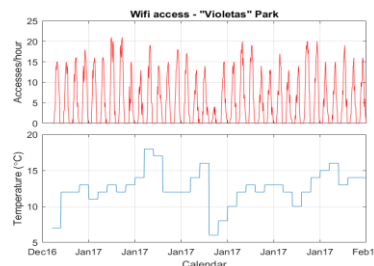


Figure 2 – Relationship between the number of accesses per hour and the average daily temperature in January 2017

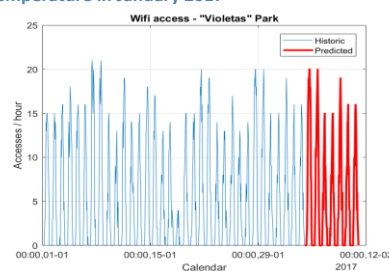


Figure 4 – Results of one week forecasting

## 4. Conclusions

This paper presents a case study where the time series forecasting method ARIMA is evaluated to predict the number of Wi-Fi hotspot users in a city.

The used dataset was based on one month of real collected measurements in the city of Castelo Branco, Portugal. The ARIMA model has proved to be a feasible approach capable of identifying and forecasting seasonality events for one week in advance. Another important aspect is its capability to correlate the number of hotspot users with weather conditions, confirming that in colder days the number of accesses is smaller. The dataset used refers to a small park in a small city in a winter month, as such its validity is limited. However, the methodology has been proved to be valid and potentiates expanding the study to other areas of the city.

## Acknowledgments

This work is funded by the European Structural Investment Funds (ESIF), through the Operational Competitiveness and Internationalization Programme (COMPETE 2020) [Project Nr. 17787 (POCI-01-0247-FEDER-MUSCLES)].

## References

- [1] Cisco Mobile VNI, 2017, Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021, March 28 (2017).
- [2] Bermolen, P. and Rossi, D.: Network forecast with support vector machines, Proceedings of the International Workshop on Traffic Management and Traffic Engineering for the Future Internet, FITraMEn 08, Porto, Portugal (2008).
- [3] Suarez, C., Parra, O. and Díaz, A.: An ARIMA model for forecasting Wi-Fi data network traffic values”, *Revista Ingeniería e Investigación*, Vol. 29, Issue 2, p. 65-69 (2009).
- [4] Bui, N., Michelinakis, F. and Widmer, J.: “A Model for Throughput Prediction for Mobile Users”, 20th European Wireless Conference, Barcelona, Spain, pp. 1-6 (2014).
- [5] ALLBESMART, <http://allbesmart.pt/?lang=en>, accessed in May (2017).
- [6] Hamilton, J.: *Time series analysis*, Vol. 2. Princeton: Princeton university press, (1994).