MODELING A DATA NETWORK CONSIDERING PRIORITIES AND TRANSMISSION PROBABILITIES USING MATLAB AND PYTHON SOFTWARE

Gabriel Cerqueira Pires[1][0009-0008-3031-4349], Leonardo Grando[2][0000-0002-0448-209X],

José Roberto Emiliano Leite[3][0000-0001-8368-7386], Edson Luiz Ursini[4][0000-0002-1597-4057]

[1,2,3,4] FT UNICAMP, Limeira/SP, 13484-332, Brazil
g256728@dac.unicamp.br
lgrando@gmail.com
joserel@unicamp.br
ursini2@unicamp.br

Abstract: This paper presents a general proposal for planning and dimensioning a multi-service IP network with voice, video, and data. It presents the methodology for dimensioning link capacity and packet delay, with QoS requirements regarding general distributions, using discrete event simulation. Furthermore, our paper includes a procedure for optimizing transmission probabilities between different nodes, aiming to increase throughput and reduce delay, considering the traffic loads offered to the network. This procedure applies different transmission probabilities to the listed services. Admission control is also implicit to regulate the number of active stream services. A Jackson network is used to validate the simulation model. Once validated using the Jackson network, other probability distributions are used to explore the available delay and throughput values, and the result of the suboptimal optimization is obtained.

Keywords: Modeling, MATLAB, Python, Networks, Probability, Discrete Event Simulation.

1. Introduction

In modern telecommunications networks there is a significant increase in data from sensors or different types of machines (IoT) that travel through data networks and may present characteristics such as self-similarity (bursts), due to the characteristics of machine-to-machine traffic, IoT sensors, Digital Twins, and other typical characteristics of Industry 4.0. In addition to these traffic, such as file transfers or database queries are also present in current networks. This first type of traffic is what we will call elastic traffic and will be modeled by a Pareto distribution with infinite variance but based on data obtained by measurements in a real network (in this network, it is the traffic of Node 1). On the other hand, other stream services with severe loss requirements also share these kind of networks. To consider these data, characteristics close to real ones were considered, such as the types of CODECs and their durations (in this network, they will be Nodes 3 and 5). These networks may present bottlenecks that can compromise the sending of information, especially in critical cases, and, mainly, delays that must be avoided. This project proposes a typical network, which cannot be dimensioned by analytical models alone (e.g., Jackson Networks) because they require simplifications that can mask the intended results. Therefore, a simulation approach is essential. In general, one wants to evaluate the delay or throughput of the network. To demonstrate the methodology, this simulation project addresses a case study consisting of a network with six nodes, in which the data flowing through the network must be sent to a known destination. In this project, the Jackson network will validate the simulation model since its results are well known. To “stress” the network, three services are placed in Nodes 1 (elastic), 3, and 5, respectively, and the remaining nodes are only for traffic passage, one elastic and two streams. Then, to evaluate other situations such as different service probability distributions, whether of lower variance (e.g., deterministic) or long tailed (e.g., Lognormal), the Jackson network does not produce adequate results, and simulation becomes essential. We also want to evaluate the effect that different priorities given to services have on both delays (seconds) and throughput (bits/s, bps) of traffic to the destination. In addition, this work also aims to evaluate the best (sub-optimal) probabilities that can be allocated to improve delays and throughput of data delivery to the mediator. As much as possible, the project worked with measured data from the existing network of the University of Campinas, School of Technology, containing elastic traffic (sensitive to losses) and made assumptions very close to the real values known in the literature for stream traffic (sensitive to delays). Results show that both throughput and delay can be significantly improved by adopting this type of approach. Section 2 shows some related works, Section 3 addresses the services to be considered, in an application-level approach, and their corresponding throughputs, Section 4 shows the proposed model, at the physical and link level, with its initial probabilities for data transfer between nodes and its solution by the Jackson Network, Section 5 evaluates the performance using simulation and comparison with the Jackson model, and Section 6 presents the conclusions and future work.

2. RELATED WORK

Since their inception, IP networks have been limited in terms of QoS requirements, which can be overcome with appropriate planning and dimensioning. The work by Tavares, 2010 presents Analysis and simulation of packet-switched communication networks in MATLAB. The article by Almeida, 2014 presents the Probabilistic Queue Models M/M/1, M/G/1. The article Trindade et al, 2003, presents a proposal for planning a multi-service IP network with QoS, identifying the main components of a service model with network mechanisms and protocols, calculation methods used in dimensioning, and the description of the main traffic parameters. planning of a multi-service IP network with QoS and a case study. This paper presents a proposal for planning and dimensioning a multi-service IP network with voice, video, and data. The methodology for dimensioning link capacity and packet delay, with QoS requirements, using discrete event simulation, is presented. This model can be used for the reliability of real networks because it is initially validated analytically. The applicability of this method for typical multi-service IP network scenarios is demonstrated. The work of Leite, J. R. E. (a)., 2019, present an AdHoc network simulation model contemplating 7 clusters, 14 nodes and 3 gateways, forwarding packets to an IoT Mediator (the end node); one of the Gateways is an emergency one; it was created for situations of overload in the main gateways, diverting priority traffic to a decongested route. Traffic studies were carried out without considering traffic types. Scenarios and use cases were created with treatment of real situations of Electricity Consumption on University Campuses and Natural Disasters, for example, using discrete event simulation as in Leite, J. R. E. et al.(b). Some of these nodes (six) were used in the proposed article. Our work has different characteristics from those previously mentioned, as it designs the services that will share the network (applications) and then designs the network to accommodate them to the end node.

3. NETWORK MODEL AND SERVICES

A six-node network was proposed using a key system to transfer files between nodes. This keying was done based on a predefined probability distribution, as shown in Figure 2 and the R matrix (Jackson Network).

- The keys that exit Nodes 1, 2, and 5 have their initial probabilities shown (they are the only nodes that route traffic).
- The simulation ends at Node 6 (in the context of an IoT network, it is the mediator), where it ensures that all files generated by nodes one, two, and five reach their final destination (node six).

The assumption is that Node 1 generates only elastic, bursty, self-similar (Pareto) traffic. Node 2 generates VoIP stream traffic, and Node 5 generates Video on Demand stream traffic. Nodes 2 and 5 have Admission Control to establish the maximum number of active calls, both for VoIP and Video on Demand. Node 1 is only a call generator, but Nodes 2 and 5 are call generators and also outflow traffic from other nodes. Nodes 3 and 4 only outflow traffic, and Node 6 receives all traffic from the network.

3.1. Elastic traffic generated on node 1 (HTTP)

To evaluate a real network, a School of Technology State University of Campinas HTTP service real measurement, which the timely downlink and uplink data received. Figure 1 shows the day with the highest use and the time with the highest peak.

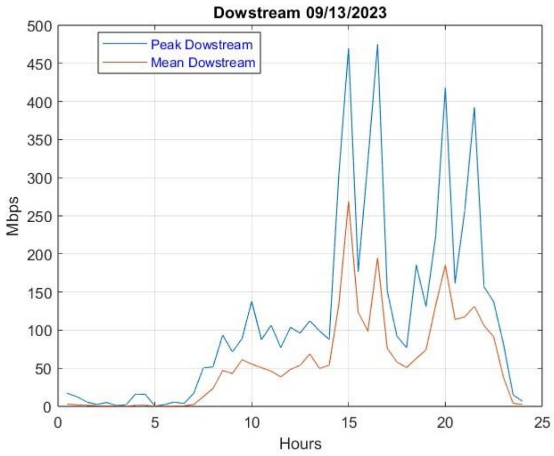


Fig. 1. Graph showing peak downstream data on 9/13/2023 - Shows that there was a data peak of around 440 Mbps at 3:00 p.m. (choose the second-highest value).

Next, some calculations are necessary to obtain the traffic volume generated at node 1 (bps). Considering $E[T(x)] = 15s$ as the maximum time the file can be sent in up to 95% of cases and initially considering a rate for elastic traffic (Ce) of 300 kbps.

From a poll of 45 websites that students access, a $x_m = 88\ kb$ is the size of the second smallest site, where the vast majority of sites are larger than a x_m . The Pareto distribution is used to identify the minimum file size that will be sent with certainty as in Trindade et al., 2003. It is worth mentioning that, in practice, it is necessary to investigate the data flows to evaluate the best distribution to be adjusted. The use of Pareto is almost a worst-case assumption because it assumes a large variability in the packets (bursts). For this value α=1.5, the distribution has infinite variance.

$$X_m = x_m \,.\, 0.05^{-\frac{1}{\alpha}} \quad (1)$$

Pareto distribution (95% of sites are smaller than X_m)

$$\alpha = 1.5 \text{ (2)}$$

$$X_m \ = \ 648.390kb \ \text{ (3)}$$

(Maximum file size in 95% of cases)

$$h \ = \ 64 \ kbps \text{ (4)}$$

(Default transmission rate for elastic services, or the minimum available channel size)

$$E[x] = \ \alpha \,. xm \ - 1 = \ 1{,}5 \,.88 \ k0{,}5 = 3xm \ = \ 264 \ kb \text{ (5)}$$

(mean of files and mean of Pareto distribution)

$$\lambda \ = 300k \,. 264k \ \ = \ 1.13 \, files \ / \ s \ \text{ (6)}$$

$$m \ = 1{,}13.264 \ k \ = (files/s).(bits/file) = \ \ 300 \ kbps = 37{,}5kB/s \text{ (7)}$$

(traffic offered or average bandwidth to carry out the service).

It is defined:

$$Fr \ = \ E[T(x)] \ .hx \ \ = \ 1.48 \ = \ 1 + \frac{E_2(R,R_p)}{R.(1-\rho)} \text{ (8)}$$

(Service delay factor).

Therefore,

$$0{,}48 = \frac{E_2(R,R_p)}{R.(1-\rho)} \text{ (9)}$$

$E_2 \ (N, A)$ is the Erlang expectation distribution (Erlang C or Erlang second formula).

$$R\rho \ = \frac{m}{h} = \ 4.68 = \ A \qquad \text{(10)}$$

ρ is the use of a server and A is the traffic offered.

$$R = \frac{C_e}{h} \text{ (11)}$$

Note that

$$\rho = \frac{m}{C_e} = 0.83 \text{ (12)}$$

that is, the occupancy of each server. The equations $(10), (11)$ and (12) are used to find C_e, which is the required bandwidth and the number of servers, each with 64kbps, Roberts et al., 1996.

To find the value of C_e the Newton-Raphson method and the Erlang wait formula were used.

$$C_e \ = \ C_{elastic} \ = 361.17 \ kbps \text{ (13)}$$

The Pareto distribution was not initially used in node 1 of the network, but only to evaluate the amount of bits/s that would be generated by this distribution, whose Pareto approximation takes into account the characteristic of elastic traffic (bursts), Trindade et al., 2003.

3.2. Stream traffic generated on Node 3 (Video on Demand)

For Stream Video on Demand traffic, AV.1 was chosen (with a required bandwidth of 10 Mbps), (Y. Chen *et al.*, 2018). N = 15 calls each one with holding time ½ hour time and T = 24h all day.

$$\Sigma \mathrm{t}_{calls \ duration} = 15 \,. ½ \ hour \ = \ 27000 \ s \text{ (16)}$$

$$\lambda = NT = 1524.3600 \ = \ 0.00017361 \ call/s \text{ (17)}$$

$$N° \ services \ = \ 10A \ = \ \lambda \,. tr. N° \ services \ = \ 10 \ Erlangs \text{ (18)}$$

Applying Erlang B formula results in a 3.65% blocking rate that satisfies this requirement.

$$Bandwidth\ =\ C_{video} = services\ .10\ M = 10\ .\ 10\ M\ =\ 100\ Mbps\ (19)$$

3.3. VoIP Traffic stream generated on node 5

The codec G.711 has been assigned (with the necessary bandwidth of 64 kbps) for VoIP. The data N, T, t and services number, were assigned through assumptions that must/can be improved for each application and each new situation must be properly adapted,.N = 500 calls and T = 24h, period in which the calls were generated.

During this period, measurements indicated that there were 50,000 s of occupancy.

$$\frac{\Sigma_{calls\ duration}}{N} = \frac{50000}{500} = tr = 100\ s/call\ (20)$$

$$\lambda\ = NT =\ 50024.3600\ = 0.0058\ call/s\ (21)$$

$$Services\ number\ =\ 20\ (22)$$

$$A\ =\ \lambda.tr.20 =\ 11.6\ Erlangs\ (23)$$

Since the number of services is 20, the Erlang formula B, for blocking= 0.739%<1%, satisfies the blocking requirement.

$$Bandwidth\ = C_{VoIP}\ = Channel.64\ k = 20\ .\ 64\ k = 1.280\ Mbps\ (each\ service\ uses\ 5.64\ kbps).\quad (24)$$

The modeling defined the virtual servers for each of the nodes: on node 1 with 5.6 virtual servers of 64 kbps. Which is related to equation 11. On node 3 with 225 virtual servers of 10 Mbps (225 CODECs), and on node 5 with 20 virtual servers with 64 kbps (20 CODECs). Thus, these are the models that defined the bandwidths for nodes 1, 3, and 5. The real network will have to forward this volume of traffic. The total bandwidth of the traffic stream (C_s) é $C_s\ = C_{Video}\ + C_{VoIP} =\ 100\ M\ +\ 1,28\ M = 101.28\ Mbps.$ Therefore, the total bandwidth (C_t) is: $C_t = C_s + C_{elastic=}\ 101.28\text{M}\ + 361.17\text{k} = 101.64\ \text{Mbps}.$ This amount of traffic gained at the application level is drained at the physical level.

4. SIMULATION MODEL

The simulation model follows the topology presented in Figure 2, with the elastic and stream services (Nodes 1, 3, and 5) being generated as described in Section 3. Each node in the real network has only one server (processor).

4.1. Jackson Network

The next step is to distribute the traffic obtained in Section 3. According to the values of the rates generated in the three types of services (Nodes 1, 3, and 5), it is possible to identify the transmission rate of each node. The traffic is distributed throughout the network according to the routing probabilities of each link. Figure 2 shows the rates generated in each node and the routing probabilities that were stipulated for the network.

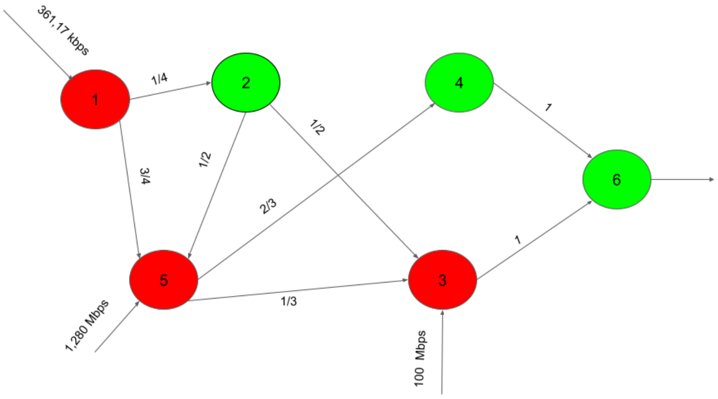


Fig. 2. Simulation model with the respective routing probabilities

To calculate the effective transmission rate of each node (the sum of the generated rate adding to the data passing rate), to determine the appropriate service rate (μ_i) for statistical equilibrium of the system, the service rate must be greater than or equal to the arrival rate at each node. According to Jackson's model, a matrix R was created with these probabilities and a matrix $\Lambda = [1\ 2\ \ldots n]$ with the values generated at each node. From the values of the matrix R with the routing probabilities and the values generated at each node, the effective rates at each node are obtained ($\Gamma = [1\ 2 \ldots n]$).

Using Matlab to calculate the matrix which shows the rate of each of the nodes according to the following procedure:

$$R = [0\ 0.25\ 0\ 0\ 0.75;\ 0\ 0\ 0.5\ 0\ 0.5;\ 0\ 0\ 0\ 0\ 0; 0\ 0\ 0\ 0\ 0; 0\ 0\ 0.33\ 0.67\ 0]\ (26)$$

$$\Lambda\ =\ [361170\ 0\ 100000000\ 0\ 1280000]\ (27)$$

$$\Gamma = \Lambda.[I_{5X5} - R]^{-1}\ (28)$$

$$\Gamma = [360000\ 90000\ 100580000\ 1060000\ \ 1600000]\ (29)$$

which shows the rate of each of the nodes according to the following procedure: with effective λ_i, it is possible to determine the values of ρ_i (the processing rate of each node), whose only limitation is that they must be greater than the $\rho_i = \frac{\lambda_i}{\mu_i} < 1$ at each node, so that the queue is stable (however, the higher the value of ρ_i, the lower the node delay). From a practical point of view, the value of ρ_i as being 70 percent of the value of $\frac{\lambda_i}{\mu_i}$ (whatever value you make $\frac{\lambda_i}{\mu_i} < 1$ could be used), except Node 1 with the elastic service and $\rho = 0.83$. Although the assumption is that each node has exponential arrivals, the characteristic of Nodes 1 (Pareto), 3, and 5 (Poisson arrival with blocking), makes the network performance different from the Jackson network. Therefore, the network behavior is then analyzed for the Lognormal (long tail) and Constant (zero variance) service distributions.

4.2. Lognormal G in General Services

In the case of the Lognormal distribution, the assumption is that the meaning of the values of the associated normal distribution is equal to the meaning of the exponential distribution, and its standard deviation is four times the standard deviation of the exponential distribution.

The Jackson Network exists only for M/M/1 networks; for the lognormal distribution, two parameters are required: μ_i and σ_i. Therefore, these values were assumed and calculated from the following equation:

$$\mu_{imean} = \frac{1}{\lambda_i} \quad (30)$$

$$\sigma_i = \frac{4}{\mu_{i(service\ rate)}} \quad (31)$$

This can be visualized with the software shown in Appendix B. Being essential, the multiplication of the standard deviation by four so that the distribution would become truly long-tailed. To arrive at the values that would be used in MATLAB - Python (Tavares 2010, Almeida 2014).

4.3. Constant G in General Services

Thus, as in the previous section, a new simulation model was proposed in which the service was made from constant distribution. Therefore, to use it, an adaptation was made, since Matlab (Tavares 2010, Almeida 2014) has only uniform distribution. However, the uniform distribution becomes a constant distribution when its initial and final parameters are the same. Therefore, the initial and final values of each node were assumed to be equal to:

$$G = \frac{1}{\mu_i} \quad (32)$$

5. Simulation performance evaluation

To evaluate the network performance, as previously described, it is necessary to use a simulation program. A reference for the network nodes is that they should have a utilization of 70% for the M/M/1 model as a reference. For this, the Python programming language (version 3.12.8) was used, combined with the Simpy library with a simulation time of 1000 s. This tool was used due to its ease of use and the fact that it is a free software. The exponential services are calculated by the Jackson Network. Keeping the same average rates, Figure 3 shows the situation in which the services are exponential with priority in Nodes 3 and 5 in relation to Node 1 which has less priority.

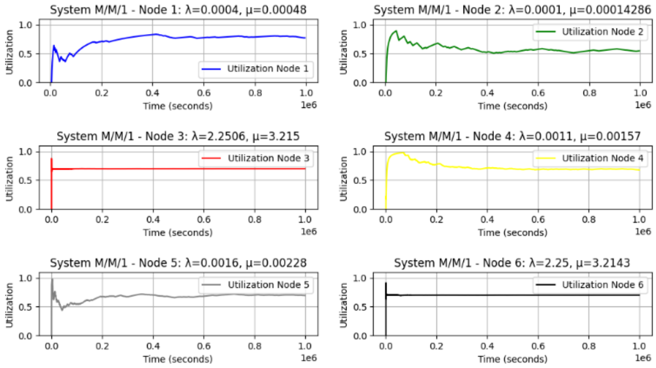
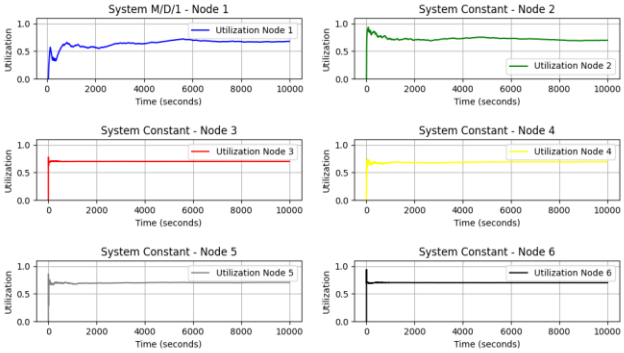


Fig. 3. Utilization of server from M/M/1 with priority

Figure 4 shows the situation in which the services are deterministic with priority in Nodes 3 and 5 in relation to Node 1 which has less priority.


System M/D/1 - Node 1

System Constant - Node 2

System Constant - Node 3

System Constant - Node 4

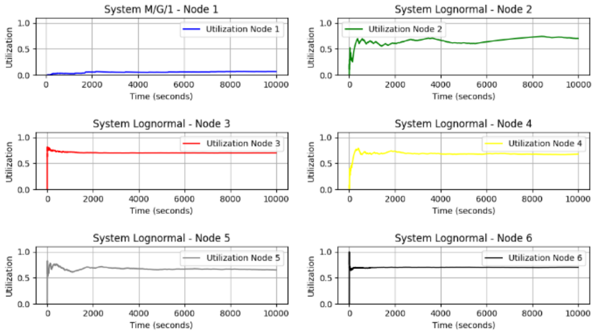
System Constant - Node 5

System Constant - Node 6

Fig. 4. Utilization of server from deterministic queue with priority

Figure 5 shows the situation in which the services are lognormal with priority in Nodes 3 and 5 in relation to Node 1 which has less priority.

.


System M/G/1 - Node 1

System Lognormal - Node 2

System Lognormal - Node 3

System Lognormal - Node 4

System Lognormal - Node 5

System Lognormal - Node 6

Fig. 5. Utilization of server from lognormal queue with priority

After checking how the network behaves to nodes with different service distributions, the next step was to analyze the time it takes for each file to reach its destination, with the same rates with and without priority. This time, however, MATLAB and Python were used. The values obtained concern the network, however, it would still be possible to improve these data using the method of "optimizing" the rates by varying the probabilities of the three nodes that inject rates into the network. Several simulations were performed with five probability values (1/6, 1/5, 1/4, 1/2, and 1/3) in which each switch received one of these probabilities and all the values were tested for the others, making a total of 3^5 =125 simulations to find the one that would make the delay time "optimized". The probability that they did this, and their packet delivery time are seen in Table 1. The probability for output S_1, p_{s_1}, is given since the probability of output $S_2$is

$$p_{s_2} = 1 - p_{s_1}. \quad (33)$$

This is a process of work, and Table 1 shows the simulation results when the initial relay odds of Nodes 1, 2, and 5.

Table 1: Results with initial relay probabilities $\frac{1}{6}, \frac{1}{6}$ and $\frac{1}{6}$.

Service	**Exponential**	**Constant**	**Lognormal**
Delay (s)	0.0188	0.0069	0.0631
Throughput (Mbps)	101.6(*)	103.96	101.78 (*)

(*) The differences found for delays are as expected for the three types of service. However, the lognormal service should have a lower throughput than the exponential service, but this value is very close and it is due to the very high load value at Node 3.

6. CONCLUSIONS

To determine a communication network based on the OSI model (application and physical-link levels), two essential factors must be analyzed: the delay and throughput caused by the network. This could be seen from the probabilistic characteristics of each node. The use of the server is important to ensure the stability of the nodes with $\rho < 1$. If the delay is a determining factor, the best option will be the deterministic service, but the best distribution should be the one obtained through measurements. If the distribution is long-tailed, it may be necessary to increase the network server rates to comply with. In future work, we intend to execute these methods also in software R. The fact that there is an imbalance in the network, as Node 3's load is much greater than that of Nodes 1 and 5, does not provide a better understanding of the algorithm employed. The idea is that future work will lead to a more balanced network load. New tests will be carried out with more appropriate values to better illustrate the method. Another important aspect may be assigning different probabilities to services and observing the impact on delay and throughput.

ACKNOWLEDGMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior — Brasil (CAPES) — Finance Code 001.

References

1. Almeida, J. M. 2014. *Probabilistic Models: M/M/1, M/G/1 Queues*. Belo Horizonte: Federal University of Minas Gerais, Department of Computer Science.
2. Chen, Y., J. Han, J. Lin, S. Wang, and C. Chiang. 2018. "An Overview of Core Coding Tools in the AV1 Video Codec." In *Proceedings of the 2018 Picture Coding Symposium (PCS)*, San Francisco, CA, 41–45.
3. Leite, José Roberto Emiliano. 2019 a. *Modeling and Simulation of IoT, AdHoc, and RFID Networks*. PhD diss., University of Campinas (UNICAMP), School of Technology.
4. Leite, J. R. E; Massaro, F. R., Martins, P. S.; Ursini, E. L..2019. Reducing Power Consumption in Smart Campus Network Applications Through Simulation of High-Priority Service, Traffic Balancing, Prediction and Fuzzy Logic, Winter Simulation Conference, 2019 b.
5. Riedl, A., Bauschert, T., and J. Frings. 2002. "Framework for Multi-Service IP Network Planning." In *Proceedings of NETWORKS 2002*, Munich, 183–190.
6. Roberts, J., U. Mocci, and J. Virtamo. 1996. *Broadband Network Teletraffic – Performance Evaluation and Design of Broadband Multiservice Networks* (Final Report of Action 242). Lecture Notes in Computer Science 1155. Berlin: Springer-Verlag.
7. Tavares, J. M. 2010. *Analysis and Simulation of Packet-Switched Communication Networks Using Matlab*. Master's thesis, University of Aveiro, Department of Mathematics.
8. Trindade, M. B., M. S. Medrano, and A. C. Lavelha. 2003. "Planning of Multi-Service IP Links Considering QoS Requirements." Paper presented at the National Congress on Information and Communication Technology – SUCESU 2003, Salvador, Brazil.
9. Ursini, E. L., P. L. Martins, V. S. Timóteo, and F. R. Massaro Jr. 2015. *Modeling and Simulation Applied to Link Dimensioning of Stream IP Traffic with Incremental Validation*. Campinas: University of Campinas (UNICAMP).

APPENDIX A