

USING SIMULATION TO EVALUATE CALL FORECASTING ALGORITHMS FOR INBOUND CALL CENTER

Guilherme Steinmann
Paulo José de Freitas Filho

Postgraduate Program in Computer Science
Dept. of Informatics and Statistics
Federal University of Santa Catarina
Caixa Postal 476 - 88.040-900
Florianópolis/SC, BRAZIL

ABSTRACT

The call center industry has expanded greatly over recent years and it is constantly striving to increase business efficiency and customer service effectiveness. Incoming call volume forecasting algorithms are used in inbound call centers to predict the demand for services and, as a result, to plan resource allocation. However, a number of phenomena can have an impact on incoming call volumes, meaning that classical forecasting algorithms will produce less than satisfactory results. When evaluating the performance of a forecasting algorithm, acquiring the data needed for research is not always straightforward. This article shows how simulation can be of use to generate data that can be used to evaluate incoming call forecasting algorithms.

1 BACKGROUND

In today's global economy, call centers are important points of contact between businesses and their customers, both existing and potential ones. Managing call centers involves administrating customer contacts (inbound call centers) and/or administrating the business's contacts with its customers (outbound call centers) (Mehrotra and Fama 2003).

The challenge is to become a world class call center, that is, to increase business efficiency and customer service effectiveness (Tanir and Booth 1999; Hall and Anton 1998; Borst et al. 2004; Brown et al. 2005). In inbound call centers, a key element for success is correct allocation of staff to meet time-varying demand for service (Aksin et al. 2007; Ibrahim 2012). A well-sized call center reduces operating and start-up costs and controls waiting times, agent idle times, busy tones and abandonment (Whitt 1999; Freitas Filho et al. 2007; Franzese et al. 2009; Avramidis et al. 2010).

The authors of this paper are conducting a research project to develop tools for predicting call volumes for inbound call centers. The project will work with "classical" forecast models, such as exponential smoothing, and enhance them with certain computational intelligence methods and techniques for decision-making under uncertainty. The idea is to allow an analyst to incorporate their tacit knowledge about a call center, in order to reduce the degree of prediction error in the output of prediction models. More accurate forecasts of the number of incoming calls can facilitate planning of department budgets and staffing requirements, such as how many workstations are required and how many lines are necessary (Cleveland and Mayben 1999; Steckley et al. 2005). Bayesian networks have proved to be the best alternative for achieving this. These networks allow analysts to adapt forecast models to incorporate the empirical knowledge they have acquired observing the behavior of demand for call center services by creating

countless events that can affect the forecast results. This makes it possible to refine the results of the classical models using Bayesian networks to incorporate tacit knowledge.

Classical forecast models tend to produce acceptable results when faced with relatively long and well behaved series, and this is particularly true of models based on analyses of time series data. However, the majority of time series data on demand for call-center services do not meet these criteria. The dynamics of these environments means that we are rarely in possession of consistent, homogenous data covering long periods (three years or more). On the contrary, the great majority of time series cover less than three years and lack uniformity. Short time series restrict the applicability of the most powerful statistical prediction tools, such as ARIMA for example. In turn, the lack of unity in a series indicates the presence of phenomena that very often disrupt the data considerably, leading to results that are less satisfactory than expected if classical forecast processes are employed.

This situation forces members of call-center planning teams to develop a multiplicity of empirical methods to supplement the classical forecasting models. This type of “knowledge” cannot be ignored. It is very often an important legacy that the prediction and planning teams need to rely on. According to Mehrotra and Fama (2003), call center forecasting is based on a combination of historical data, time series models and expert judgment.

Over the last 20 years it has become possible to employ computational tools to capture specialists’ knowledge in order to disseminate it and take advantage of it in multiple decision-making processes. Notwithstanding, in the great majority of corporate environments providing customer services, these types of resources have not yet been implemented.

To assess the impact of different phenomena in the forecasting algorithms, there is a need to simulate a variety of different scenarios, at varying degrees of intensity, on the behavior of the load on the systems under investigation. Thus, it is necessary to have an incoming calls dataset.

2 STATEMENT OF THE PROBLEM

Human resources account for 60-70% of call center operating costs (Gans et al. 2003; Borst et al. 2004). Managing an accurate incoming calls forecast involves reducing the error between forecast data and real data and correct resource planning. Call center managers aim for a monthly forecast error rate below 5%. In general, classical data forecasting methods are unable to deal with phenomena (random events) which disturb the expected values, leading to increased forecasting errors, simply because they do not take into account of the probability of such events occurring.

The first barrier to achieving the objectives of this research project was the difficulty involved in obtaining real incoming call center data that would make it possible to study the effects of random events on call volume forecasts. Such data are rarely made available by call center firms, even for the purposes of research. One solution to this problem is to use synthetic data.

Artificially generated data demand the use of stochastic models because the behavior of incoming calls volume varies randomly over time (Ibrahim 2012). Once the data have been generated, it is possible to evaluate forecasting models in a variety of simulated scenarios. The challenge is to use a simulator to construct a call generator capable of producing data that exhibit behavior as close as possible to the true variable. It is expected that the synthetic sample data replicates behavior associated with trend, seasonality and, most importantly, random fluctuations in call volumes.

3 FORECAST PROCESS

In order to understand the call forecasting process, it is first necessary to understand how incoming calls are dealt with in a call center. A model for an incoming call generator can then be constructed on the basis of this knowledge. Below is a summary of the steps involved in dealing with an incoming call to a call center (Tanir and Booth 1999; Gans et al. 2003; Avramidis and L’Ecuyer 2005):

- i) A customer dials the call center using a public service telephone network (PSTN).

- ii) If there is a trunk line free, the call is connected to a private automatic branch exchange (PABX) at the call center.
- iii) When the call is connected, an interactive voice response (IVR) unit may handle the call and send it to an automatic call distributor (ACD).
- iv) The ACD routes the call to an idle agent. If there is no idle agent available, the ACD may keep the call on hold in a queue until an agent is available. The customer may abandon the wait queue before being served if they become impatient.
- v) Once the customer is connected to an agent, the agent may use a computer or terminal to handle their requests.
- vi) At the end of the call, the agent may execute wrap-up tasks and then becomes available to receive another call.

During call center operations all of these steps are recorded in a log, from receipt of the incoming call right up to wrap-up, including data on waiting times, which agents took which calls and talk times. Together, all these detailed records (traces) contain a huge amount of information and illustrate a call center's performance. The ACD's operations are generally summarized and means are calculated for a set interval of time, normally 15 or 30 minutes (Brown et al. 2005; Shen and Huang 2008).

Reports can be generated showing the rate of incoming calls and their behavior as the call center operates, covering a variety of time periods: hourly, daily, weekly, monthly or even yearly. The incoming call rates may exhibit a degree of predictable variability or some type of seasonal pattern.

In the majority of cases, incoming call rates are forecast using a top-down approach (Gans et al. 2003), i.e. a forecast is made for a long period of time and then the forecast volume is divided across shorter periods. In general, the first step is to compile incoming call records in months and then these data are used for time series forecast methods such as exponential smoothing in order to produce month-by-month incoming call volumes.

However, the final forecast volume is very rarely the volume predicted by the mathematical model. It is more normal for the forecast volume to be corrected by one or more analysts. These analysts' experience and tacit knowledge are thus applied to the volumes predicted by the models, adjusting them upwards or downwards, as directed by these decision-makers' intuition, which in turn is based primarily on their expectations of behavior and events that are not captured, or only partially captured, by the mathematical forecast models.

With the forecasted incoming calls volume, the call center managers will plan the number of resources needed and allocate them affording the right demand for a certain period of time.

4 SIMULATION MODEL

Given the obstacle to analysis of historical call center data series created by the lack of available records, it was necessary to construct a synthetic data generator in order to set up a variety of different operating scenarios and analyze the performance of data forecasting methods. In order to achieve a consistent analysis, at least three years' historical data on the operations of a call center would be needed, which is very difficult to achieve from a technical perspective, but which can be easily generated by an incoming calls simulator.

A historical data archive that describes the activities of a customer services call center for credit card holders was used to test the synthetic call volume data simulation model.

These data are from a call center contract that lasted 44 months. The influence of factors such as trend and seasonality was clearly evident in this historical series, but it also exhibits certain anomalies, the causes of which were not immediately obvious, but which undeniably increased volumes beyond what would be expected otherwise. After more detailed study, it was found that there were two distinct factors that had strongly influenced the behavior of the variable call volume. The first of these was a 12-month marketing campaign encouraging customers to apply for credit cards. The second factor detected was that

the postal workers went on strike twice during the call center contract, for around 30 days both times. The postal strike forced customers to contact the bank in order to arrange an alternative method of receiving their credit card bills. By analyzing the logs from this credit card call center it proved possible to understand the historic behavior of its incoming call volumes and also the effect of these factors.

Based on the available information and on the results of the analysis of the credit card call center, a Monte Carlo simulator was developed as the first step towards evaluating forecast models and analyzing their sensitivity. The simulator is made up of three separate elements: a call generator which simulates incoming calls received by the call center; a random event generator which is responsible for changing the behavior of the variable number of calls generated depending on the type of event generated and its intensity; and a call logger, which generates incoming call statistics and lists factors and their impacts. Figure 1 is a simplified illustration of the components of the incoming call simulator.

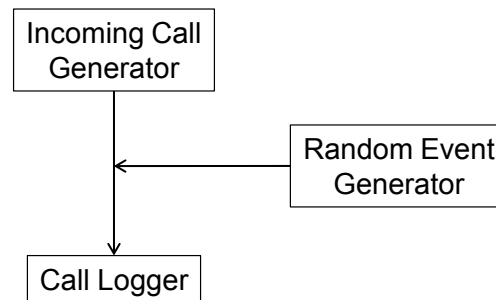


Figure 1: Diagram illustrating the elements of the incoming call simulator.

The simulator can be used to generate different types of incoming call and it is also possible to generate a variety of different random events with different types of impact on the number of calls received. A value can be attributed to the impact of each event, increasing or reducing the incoming call volume generated to a greater or lesser degree. The incoming calls generator has a module that modifies the number of incoming calls generated according to the day of the week and the occurrence of holidays and after holidays.

The call logger is responsible for generating statistics and provides an analyst with an overview of the call volume scenario and the respective influence of each event, enabling the analyst to understand the true impact on the volume predicted. Having understood the relationship between the events and incoming call volumes, it is then possible to calibrate the classical forecasting models.

The primary metric used to evaluate the forecast model employed was the mean absolute percentage error (MAPE) when the forecast volume is compared with the true volume. Jung and Min (2010) evaluated the call center service grade through a prediction model trained by a training dataset and assessed the forecast data, generated by the model, measuring the mean square error (MSE) with a test dataset. In order to conduct this comparison, the following sequence of steps was taken:

- i) Data for the historical series are generated for a period starting at time t_0 and lasting up to t_n . These data represent past calls and are used to train the forecast model.
- ii) Future data also have to be generated for a period t_{n+f} and will be used to test the forecast model.
- iii) Once the forecast model has been trained with historical data generated by the simulator, data for the future period t_{n+f} are generated using the forecast model algorithm.
- iv) The model is then verified by calculating the MAPE for the future volumes generated by the simulator against the data predicted by the forecast algorithm.

The periods covered by the training and test datasets are illustrated in Figure 2.

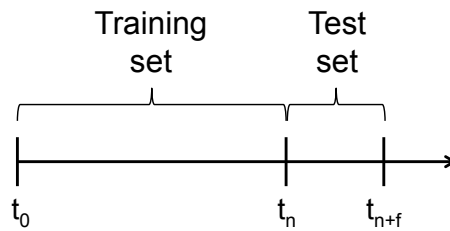


Figure 2: Diagram of the call data periods generated by the simulator.

5 EXPERIMENTATION

The data series from the credit card call center was used as a case study. In addition to analyzing the behavior of the variable itself, information on the occurrence of unknown factors was also investigated in order to assess the impact of these phenomena on the volume of incoming calls. In the case study, two factors (events) had a significant impact on call volumes during the observation period, increasing the predicted incoming call volume: (1) a marketing campaign and (2) two postal strikes. Figure 3 illustrates this data series, highlighting the peaks in demand caused by these two events.

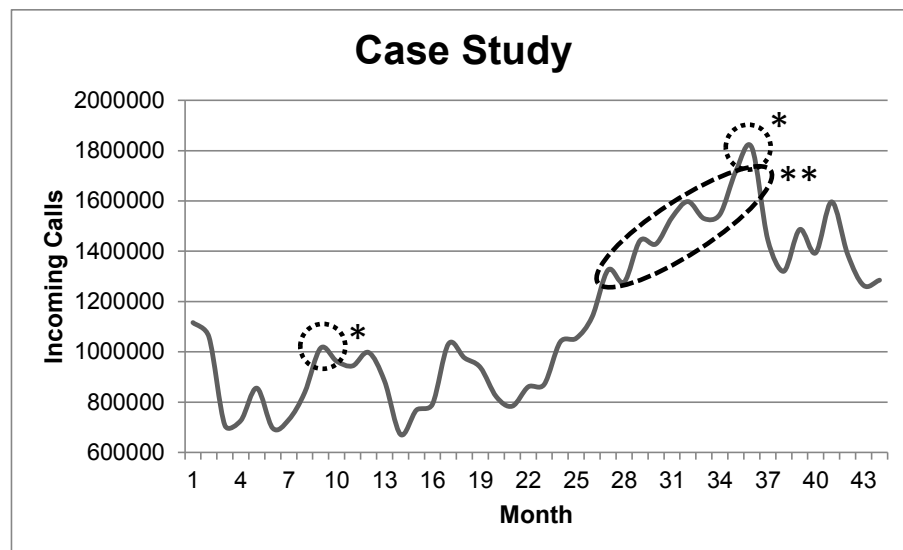


Figure 3: Historical monthly series of incoming call data from the case study. *period when postal workers were on strike. **period when the marketing campaign was running.

The objective of this analysis was to identify which components would need to be calibrated in the call simulator. After this study, less important parameters were discarded and the components trend, seasonality and random events were analyzed.

An incoming call simulator was then constructed to deal with the components chosen. The parameters for the call generator are functions for the random variables trend and seasonality of incoming calls. The parameters for the random event generator are functions for the random variables frequency and impact of factors.

During a first simulation, call volume data was generated taking account of the two events that impacted the call center. The simulation was run with the following parameters:

- a) Factor A occurs in months 09, 10, 34 and 35, and its random function of influence is: UNIF(1.15;1.19);

- b) Factor B occurs from month 27 to month 36, and its random function of influence is: UNIF(1.22;1.33).

These simulation parameters indicate that factor A will cause an increase of 15 to 19% in incoming call volumes in 4 of the months simulated. Factor B will cause an increase varying from 22 to 33% in 10 of the months simulated. There is a higher increase of the incoming calls in the months 34 and 35 due to a combination of the factors A and B.

Samples were generated with 5000 replications in order to validate the case. Figure 4 illustrates the means for these replications after the first simulation, indicating the influence of the two events. For ease of viewing, only the monthly means of the replications have been plotted. The mean confidence interval is not shown in Figure 4, because it was an average of 0.1% of the mean, which cannot be shown at this scale.

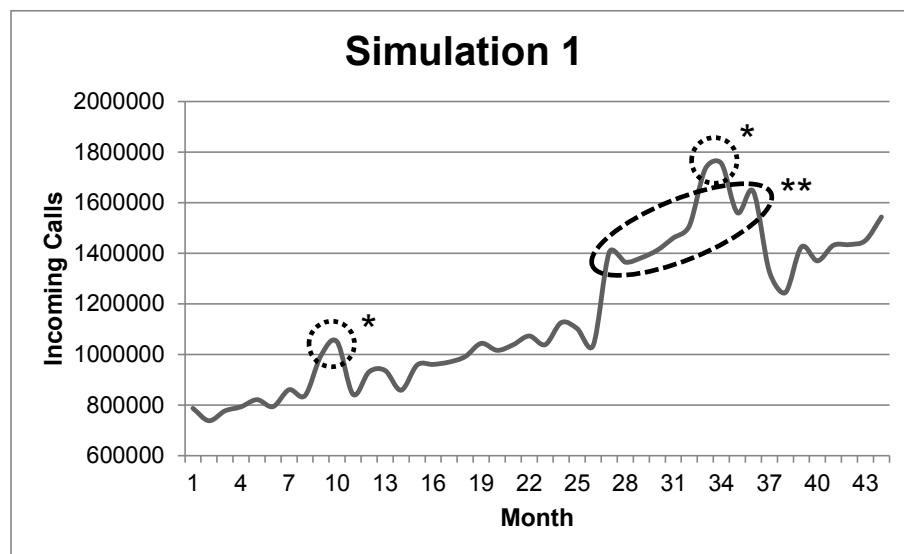


Figure 4: Monthly historical series of simulated incoming call volumes. *factor A. **factor B.

In addition to including random events that change the predicted call volumes, it is also possible to amplify these factors. The increase is applied to allow analysis of their impact on incoming call volumes and therefore on the forecast algorithm itself.

The next test was therefore to compare the effect on call volumes of changing the factors' impact levels. This test is based on the first simulation, with the following random function parameters for the two factors. The uniform distribution was used in both cases:

- a) Factor A: UNIF(1.33;1.41);
- b) Factor B: UNIF(1.48;1.73).

The impact parameters for the random factors were increased by 220% in order to highlight their effect on predicted call volumes. Figure 5 illustrates the mean monthly results for 5000 replications.

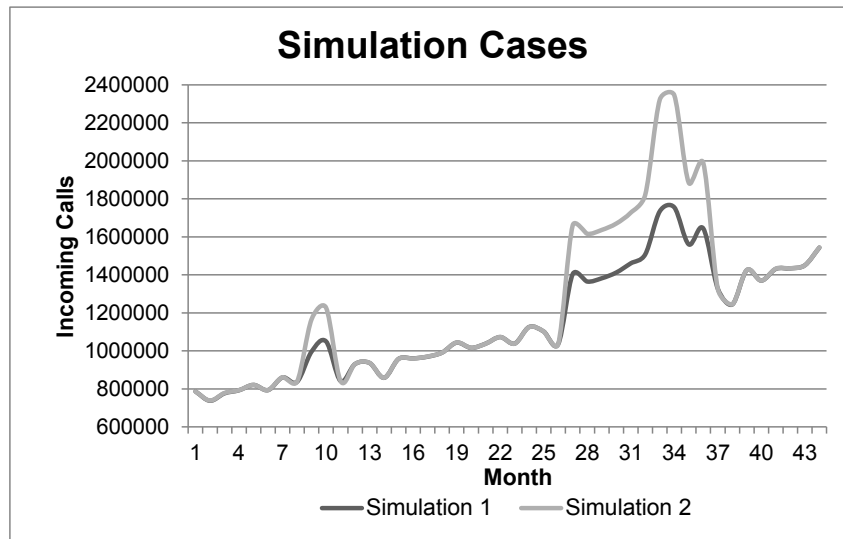


Figure 5: Comparison of two monthly historical series of simulated incoming call volumes.

6 CONCLUDING REMARKS

The major challenge for an inbound call center is to achieve more accurate incoming call volume forecasts and thereby allocate the correct amount of resources to a given operation. Call center managers require incoming call volume forecasting algorithms that afford minimum forecasting error. To achieve this objective, the algorithms must be capable of dealing with the variables involved in the dynamics of these environments. A lack of uniformity in a historical data series of incoming calls implies that phenomena are present which are capable of substantially affecting volumes, leading to forecast volumes that are above or below the true volumes if classical forecasting methods are used in isolation. The basic problem in assessing forecasting algorithms lies in the availability and quality of data.

Generating synthetic data could be an interesting alternative possibility for constructing and testing simulation models of call center operations. Countless different scenarios can be tested and random factors can be included with varying impacts. It is also possible to generate historical data for training the forecasting algorithm and future data for testing the forecast data, in order to assess the performance of different forecasting algorithms.

REFERENCES

- Aksin, Z., M. Armony, and V. Mehrotra. 2007. "The Modern Call Center: A Multi-disciplinary Perspective on Operations Management Research." *Production and Operations Management* 16:665-688.
- Avramidis, A.N., and P. L'Ecuyer. 2005. "Modeling and Simulation of Call Centers." In *Proceedings of the 2005 Winter Simulation Conference*, Edited by M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 144-152. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Avramidis, A.N., W. Chan, M. Gendreau, P. L'Ecuyer, and O. Pisacane. 2010. "Optimizing Daily Agent Scheduling in a Multiskill Call Center." *European Journal of Operational Research* 200(3):822-832.
- Borst, S., Mandelbaum, A., and Reiman, M. 2004. "Dimensioning Large Call Centers." *Operations Research* 52(1):17-34.
- Brown, L., N. Gans, A. Mandelbaum, A. Sakov, H. Shen, S. Zeltyn, and L. Zhao. 2005. "Statistical Analysis of a Telephone Call Center: A Queueing-science Perspective." *Journal of the American Statistical Association* 100(469):36-50.
- Cleveland, B., and J. Mayben. 1999. *Call Center Management on Fast Forward: Succeeding in Today's Dynamic Inbound Environment*. Annapolis, MD: Call Center Press.

- Franzese, L., M. M. Fioroni, R. C. Botter, and P. J. Freitas Filho. 2009. "Comparison of Call Center Models." In *Proceedings of the 2009 Winter Simulation Conference*, Edited by M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin and R. G. Ingalls, 2963-2970. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Freitas Filho, P. J., G. F. Cruz, R. Seara, and G. Steinmann. 2007. "Using Simulation to Predict Market Behavior for Outbound Call Centers." In *Proceedings of the 2007 Winter Simulation Conference*, Edited by S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 2247-2251. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Gans, N., G. Koole, and A. Mandelbaum. 2003. "Telephone Call Centers: Tutorial, Review, and Research Prospects." *Manufacturing and Service Operations Management (M&SOM)* 5(2):79-141.
- Hall, B., and J. Anton. 1998. "Optimizing Your Call Center Through Simulation." *Call Center Solutions Magazine*.
- Ibrahim, R., P. L'Ecuyer, N. Regnard, and H. Shen. 2012. "On the Modeling and Forecasting of Call Center Arrivals." In *Proceedings of the 2012 Winter Simulation Conference*, Edited by C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher, 256-267. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Jing, L., and G. Min. 2010. "Predicting Call Center Service Grade with Improved Neural Network Algorithm." *Intelligent Systems and Applications* 1-4.
- Mehrotra, V., and J. Fama. 2003. "Call Center Simulation Modeling: Methods, Challenges, and Opportunities." In *Proceedings of the 2003 Winter Simulation Conference*, Edited by S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, 135-143. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Shen, H., J. Z. Huang. 2008. "Interday Forecasting and Intraday Updating of Call Center Arrivals." In *Manufacturing and Service Operations Management (M&SOM)* 10(3):391-410.
- Steckley, S. G., S. G. Henderson, and V. Mehrotra. 2005. "Performance Measures for Service Systems With a Random Arrival Rate." In *Proceedings of the 2005 Winter Simulation Conference*, Edited by M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 566-575. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Tanir, O., and R. J. Booth. 1999. "Call Center Simulation in Bell Canada." In *Proceedings of the 1999 Winter Simulation Conference*, Edited by P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans, 1640-1647. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Whitt, W. 1999. "Dynamic Staffing in a Telephone Call Center Aiming to Immediately Answer All Calls." *Operations Research Letters* 24: 205-212.

AUTHOR BIOGRAPHIES

GUILHERME STEINMANN graduated from Federal University of Santa Catarina (UFSC) in 2007 with a degree in Computer Science and completed a Master in Computer Science at the UFSC in 2013. His research interests are call center analysis, modeling and simulation and artificial intelligence. His email address is guist@inf.ufsc.br.

PAULO JOSÉ DE FREITAS FILHO graduated from the Federal University of Rio Grande do Sul in 1978 with a degree in Metallurgical Engineering, completed a Masters in Production Engineering at the UFSC in 1985 and his doctorate in Systems Engineering was granted by UFSC and the University of South Florida in 1994. He is currently a Professor at UFSC. His experience is centered on Computer Science with an emphasis on Systems Performance and his teaching and research is primarily in the following areas: modeling and simulation, performance assessment and systems capacity planning. Over the past five years he devoted much of his time to research on modeling, simulation and data analysis of service engineering mainly in the call center area. His email address is freitas@inf.ufsc.br.