Failure Detection in Communication Systems

Gean D.Breda, Leonardo de S. Mendes and Maurício L.Bottoli

Campinas State University, Telecommunications Department, Faculty of Electrical Engineering, Brazil

Summary

The purpose of this paper is to demonstrate the Sample Space Algorithm that can be used to monitor the QoS and, while monitoring, detect the occurrence of failures in wireless and wireline communication systems. The algorithm is based in the analyzes of data stored in Call Detail Records. Each time a call is made in a voice system, VoIP or PSTN, a detailed record is generated. Detail Records are tickets whose data provide information related to the system elements involved, such as time and duration of the call, phone types and numbers, SS7 signaling, etc. The tickets are generated and stored either in PSTN switches or in VoIP gateways. For VoIP systems the detail records are usually called IPDR (Internet Protocol Detail Record). The algorithm was tested using data from real voice communication systems, in this case, a Brazilian cellular communication company. We Applied the algorithm to detect failures in four Base Transceiver Stations that uses the CDMA technology. So, our main goal here is to show, analyze and classify this algorithm according to its performance and use.

Key words:

QoS Monitoring, Performance Evaluation, Call Detail Records, Management Systems, Failure Detection.

Introduction

In the analysis and production of information performed by Telecom companies we often see a rather technical and immediate approach, frequently disregarding important information collected and stored in the Telecommunications Management Databases. An important example of such occurrence can be found when we analyze the use of Call Detail Record. Currently, their only function in Telecom Companies is dispose information to the billing system. CDRs are tickets whose data provide many information related to the call, such as time and duration, phone types and numbers, SS7 signaling, etc. The tickets are generated either in the PSTN switches or in VoIP gateways, in the case of Internet Protocol Detail Record, IPDRs.

The objective of this paper is to analyze an algorithm that can be used for monitoring the QoS and, in this process, detect failures in wireless systems (voice communication systems). It is based on a new approach to where the information contained in CDR is subjected to several treatments and analysis. For CDR we mean the Call Detail Records [1], for conventional networks, or the Internet Protocol Detail Record [2][3], for VoIP networks.

There are basically no conceptual differences between CDR and IPDR, therefore, the algorithm can be equally applied to both cases. Detail records have a complete range of information that contains the entire history of a call. It is unlikely that the information contained in the detail records can be found anywhere else on the telephone network. Some examples of information that a detail record contains are: switch's name and point code, in/out voice trunks, in/out voice time slots, origin and terminal BTS (base transceiver station) number, origin and terminal RF channels (Radio Frequency), switch peripheral components (through where the call passes inside the switch), calling and called phone numbers, serial phone number, dialed number, transferred number, phone features, starting and ending conversation time, call duration, signaling duration time, SS7 signaling information, internal call transit, type of response for the call, what happened to the call, etc. The majority of the elements contained in detail records can be monitored in order to detect failures.

Another characteristic of the detail records is reliability. This allows us to work with the detailed information contained in the CDR to perform critical tasks with large confidence in the results. In a broader view, we can consider the possibility of using CDR to perform from simple tasks like traffic monitoring [4][5] to complex ones, like the analysis of social and economic aspects of the system [6]. Such analysis can be performed once each call received or originated from the system has a correspondent detail record, making it possible to analyze the behavior of each user/element in the network. Therefore, the use of detail records, along with the algorithms presented here, can help decrease economic losses as well as lower complaints associated to a deficient Quality of Service [7].

There is only a handful of publications available about CDRs and IPDRs. Since the CDR and IPDR carry very strategic information for the operators and suppliers, it is understandable the reason why Telecom companies choose to restrict the information associated to it. There are some works developed for the use of CDR in Fraud Detection [8][9]. In these works, information is extracted from CDR and used to build up customer profiles. Other works that use CDR are related to data mining [10][11]. As far as we know, there are no publications using detail records to monitor the QoS and, consequently, no ways to detect failures in communications systems.

Manuscript received October 5, 2007
Manuscript revised October 20, 2007

The remainder of this paper is organized as follows: in section 2, we describe detail records classification; in section 3, the Sample Space Algorithm is introduced and its performance is analyzed; finally, in section 4, we present the conclusions of this work.

2. Records Classification

The classification of a detail record, which we call event, is a representation of what happened in a specific telephone call. It's much like attributing a badge or label to each possible detail record. For instance, if a call were successfully concluded, in which user "A" spoke to user "B" and the call was finalized by any of the users, we would have an OK call. It means that all detail records that represent an OK call will receive a label "OK". This classification is necessary in order to identify the system behavior in all of its range and paths where the call has been through.

Some examples of classification:

- CDMA Carrier Loss (CL);
- CDMA RF channel dropped (RFD);
- CDMA RF Channel Congestion (RFCC);
- Bad Peripheral Component (BPC);
- Bad Trunk (BT);
- Trunk Congestion (TC);
- User B does not answer (UA);
- User B busy (UB);
- Technical failure (TF);
- Incorrect Dialing (ID);
- Call OK (OK);
- Etc.

In some switches it is possible to classify a detail record in approximately 300 different ways according the call termination, this classification can be considered a highly detailed. As the classification increases more precise will be the detection and the diagnosis of the failure. On the other hand, a highly detailed classification will generate an additional work to create the table of events and their related data.
Figure 2.1 describes the hole process to obtain the results. In our case the results are the alarms represented by the failure detection.

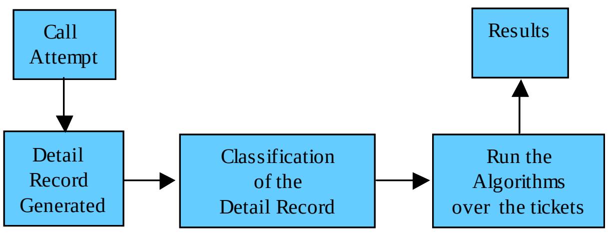

Figure 2.1 The hole process.

3. Algorithm

We are going to use the algorithm for the monitoring of the different resources in a wireless communication system. By resource, we refer to all the elements in the system, both logical and physical. A physical resource, as the name says, is related to a physical component of the system, such as:

- Switch name;
- Base Transceiver Station number;
- RF channel number;
- Base Station Controller number;
- Home Location Register id;
- Visitor Location Register id;
- Calling phone number;
- Called phone Number;
- Internal peripheral number;
- Internal sub-peripheral number;
- Incoming trunk;
- Incoming time slot;
- Outgoing trunk;
- Outgoing time slot;
- Etc.

A logical resource is a definition, such as:

- Country and area codes in the call direction monitoring;
- Switch software components;
- Translation tables;
- Etc.

The information about the resources are contained in the detail records. By monitoring these resources, we aim at following the behavior of all the events associated to that specific resource. A resource fails when one or more events associated to this resource fails. It means that when we are monitoring a resource in fact we are monitoring the QoS of each event related to that resource. In a general way, the QoS term is related with the reliability of the resources, but it can have a different meaning depending the resource that is being monitored.
Following, we present the algorithm called Sample Space Algorithm which can be used to detect failures using CDRs contained in database of telecommunication management systems of Telecom companies.

3.1 Sample Space Algorithm

The Sample Space Algorithm was developed based on two probability distributions: Binomial and Normal [12][13]. To explain this, we will start to model the events of the system as random variables. For example, let X be a random variable that represents an specific event in a Bernoulli experiment. The Sample Space of X can take two values

$$X = \begin{Bmatrix} 1 \\ 0 \end{Bmatrix}, \quad (1)$$

where the value X = 1 represents the occurrence of a specific event and 0 the occurrence of any other event. Let's also assume that the probability of X = 1 is equal to p.

The equation that express the Binomial distribution is

$$f_X(x; p, n) = \frac{(n)!}{(n-p)!(p)!} p^x (1-p)^{n-x}, \quad (2)$$

where n is the quantity of elements that will be analyzed or the size of the sample. This algorithm is not real time because it is necessary to wait until the sample is filled out to start the analysis.

We should also apply the equations above to assure that false positive alarms are restricted inside a margin of one in one million alarms. A false positive alarm that one that in the truth does not exist, there are no real problems.

Table 3.1 and Table 3.2 can assist us in understanding how to use this margin. To create Table 3.1 we use (2), n = 50, p = 0.017 or quality level of 1.7%. To create Table 3.2 we use (2), n = 400, p = 0.017 or quality level of 1.7%. On the first column we have the possible elements (β) of the sample. On the second column we have the probability of Binomial Distribution, $P(X = \beta)$, which represents the probability that in a group with n elements there's a quantity β of a specific event. The column $P(X \le \beta)$ is the Cumulative Distribution Function, which represents the probability that in a group of n elements there are 1 or 2 or 3 or, …, β of a specific event. Therefore, its values can be obtained from

$$\sum_{\beta=1}^{n} P(X = \beta), \text{ or } P(X \le \beta). \quad (3)$$

On the fourth column we have the values for the following equation

$$1 - \sum_{\beta=1}^{n} P(X = \beta), \text{ or } P(X > \beta). \quad (4)$$

This function represents the probability that in a group of n elements there are $\beta + 1$ or $\beta + 2$ or ,..., ∞ of a specific event.

The procedure to assure a limit to false positive alarms can be implemented as follows. Once we have a Binomial Distribution with n elements and the probability of an event occurrence is p, we look for the satisfactory element on Table I which satisfy the following equation

$$1 - \sum_{\beta=1}^{n} P(X = \beta) < 10^{-6}, \text{ or } P(X > \beta) < 10^{-6}. \quad (5)$$

This condition assures that the false positive alarms will be maintained within restricted limits, i.e., one in a million. In Table 3.1 and Table 3.2 the β values that satisfies the condition (5) are 8 and 22 elements, respectively.

3.2 Experiment

The algorithm's performance was tested using data from real voice communication systems, in this case, a Brazilian cellular communication company. This company has 5 million customers approximately and uses CDMA technology.

We applied the algorithm to analyze failures of various resources of the system, such as, Base Transceiver Station, RF channels, time slots, specific peripheral controllers, etc. The results shown in Figure 3.1 and Figure 3.2 synthesize the behavior of a Base Transceiver Station (BTS) of a cellular system with high traffic density.

Table 3.1: Sample Curve (Window' size of 50 elements)

Beta	P(X=k)	P(X<=k)	P(X>k)
0	4,24303E-01	4,24303E-01	5,75697E-01
1	3,66895E-01	7,91197E-01	2,08803E-01
2	1,55454E-01	9,46652E-01	5,33483E-02
3	4,30148E-02	9,89667E-01	1,03335E-02
4	8,74081E-03	9,98407E-01	1,59267E-03
5	1,39070E-03	9,99798E-01	2,01971E-04
6	1,80381E-04	9,99978E-01	2,15893E-05
7	1,96084E-05	9,99998E-01	1,98093E-06
8	1,82270E-06	1,00000E+00	1,58227E-07
9	1,47102E-07	1,00000E+00	1,11251E-08
10	1,04303E-08	1,00000E+00	6,94754E-10

Table 3.2: Sample Curve (Window' size of 400 elements)

Beta	P(X=k)	P(X<=k)	P(X>k)
0	1,05053E-03	1,05053E-03	9,98949E-01
1	7,26712E-03	8,31765E-03	9,91682E-01
2	2,50727E-02	3,33903E-02	9,66610E-01
3	5,75252E-02	9,09155E-02	9,09084E-01
4	9,87379E-02	1,89653E-01	8,10347E-01
5	1,35240E-01	3,24893E-01	6,75107E-01
6	1,53973E-01	4,78867E-01	5,21133E-01
7	1,49879E-01	6,28745E-01	3,71255E-01
8	1,27332E-01	7,56077E-01	2,43923E-01
9	9,59127E-02	8,51990E-01	1,48010E-01
10	6,48557E-02	9,16846E-01	8,31544E-02
11	3,97664E-02	9,56612E-01	4,33880E-02
12	2,22936E-02	9,78906E-01	2,10944E-02
13	1,15070E-02	9,90413E-01	9,58740E-03
14	5,50100E-03	9,95914E-01	4,08640E-03
15	2,44812E-03	9,98362E-01	1,63828E-03
16	1,01875E-03	9,99380E-01	6,19524E-04
17	3,97967E-04	9,99778E-01	2,21557E-04
18	1,46443E-04	9,99925E-01	7,51139E-05
19	5,09183E-05	9,99976E-01	2,41956E-05
20	1,67751E-05	9,99993E-01	7,42053E-06
21	5,24957E-06	9,99998E-01	2,17096E-06
22	1,56400E-06	9,99999E-01	6,06959E-07
23	4,44524E-07	1,00000E+00	1,62435E-07
24	1,20759E-07	1,00000E+00	4,16752E-08
25	3,14097E-08	1,00000E+00	1,02655E-08
26	7,83461E-09	1,00000E+00	2,43092E-09

The algorithm behavior was tested over different quality levels or probability p, which assumed the values 1%, 2%, 5%, 12%, 22%, 32%, 42%, 52%, 62% and 72% . For each level we find, by using formula (5), a number for detection of failure for that specific event. The quality level or probability p is used here as the Acceptance Quality Level (AQL).

The method adopted in the fault detection was to degrade the QoS of the BTS through random generation of problems in the RF channels. The troubleshooting was generated in a cumulative form, which means, a RF channel with a normal behavior starts to behave irregularly, presenting problems. In a second instant another channel starts to present the same failure and so on successively. As more channels present problems the QoS degrades. Each time the QoS degrades the algorithm is applied in order to detect any anomalies, failures.

The detection of this type of failure is complex, considering that the generation of these problems is purely random. It will be easier and faster to detect it if there is some order in the degradation of resources. An order presumes smaller entropy or a greater amount of information than just purely random occurrences.

3.3 Results

In Figure 3.1 we have the algorithm's performance on a sample of 50 elements. Figure 3.2 illustrates the performance on a sample of 400 elements. Comparing both pictures it is evident that increasing the sample's size decreases the level of degradation of the QoS to detect the failure, as well as the lower pattern deviation, meaning that the algorithm becomes more precise.

An important conclusion from these results is that the Sample Space Algorithm can be used to detect a great variety of problems, since problems of low impact to problems of high impact. This can be done using a variety of samples sizes. Once using different sizes a complementary detection is done, embracing all sorts of different behaviors.

Another important variable that should be measured is the amount of time needed to detect the fault. Figures 3.3 and Figure 3.4 shows that the algorithm's behavior related to time detection varies with the quality level as well as the QoS degradation level. Figure 3.3 represents a BTS behavior with quality level or *p* equals to 4% and with a degradation of 5.9%, 11.9%, 17.8%, 23.8%, 29.8%, 35.7% respectively, and the sample size equals to 50 elements. Figure 3.4 represents a BTS behavior with quality level or *p* equals to 4% and with a degradation of 5.9%, 8.9% respectively, and the sample size equals to 400 elements. The horizontal axes represents the time in which the failure was detected and the vertical axes the amount of times in which the failure was detected in the total of 4,000 experiments.

Inspecting the four graphs in Figure 3.3 and the two graphs in Figure 3.4 we can conclude that larger sizes of the sample means larger sensitivity and smaller behavior dispersion, i.e., the results are more precise. However, the larger the sample the longer the detection time will be. On the other hand, smaller samples mean smaller sensitivity and larger behavior dispersion, i.e., less precision. The advantage of a smaller window is that the detection time also decreases.

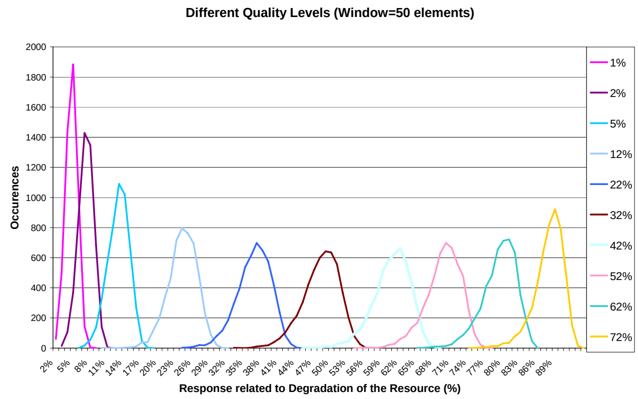

Fig. 3.1 Response related to resource degradation (50 elements).

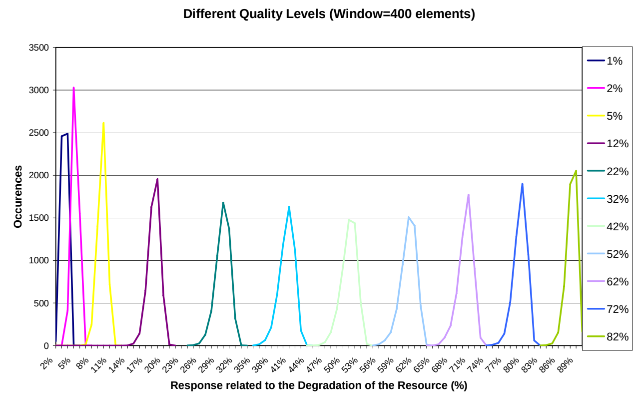

Fig. 3.2 Response related to resource degradation (400 elements).

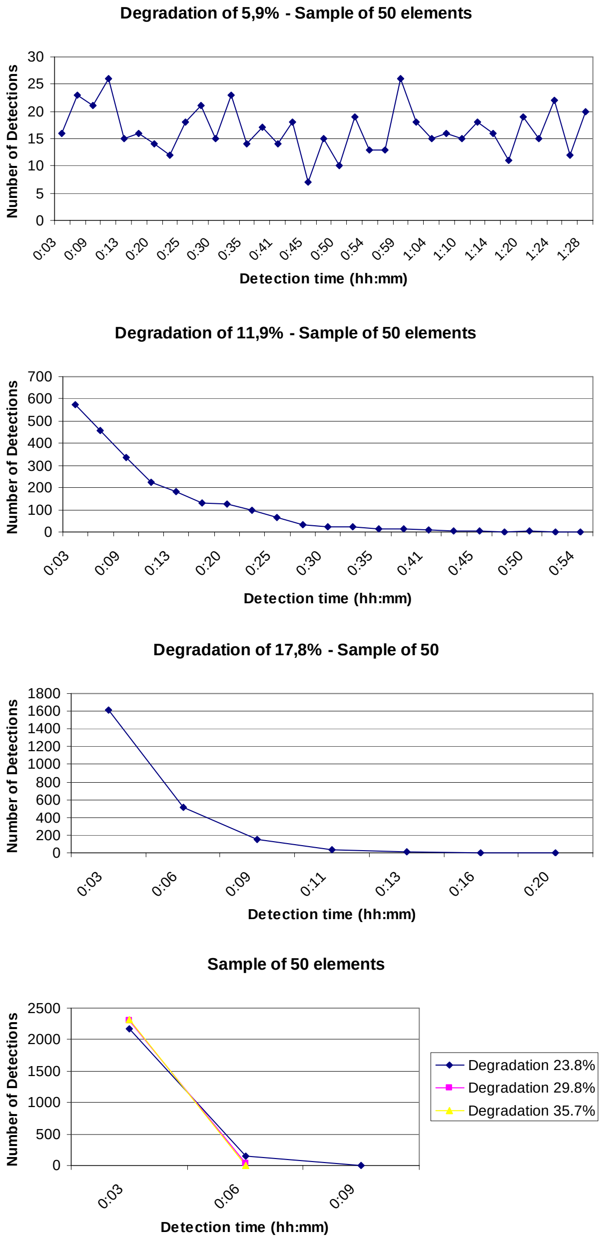

Fig. 3.3 Detection time related to resource degradation (50 elements).

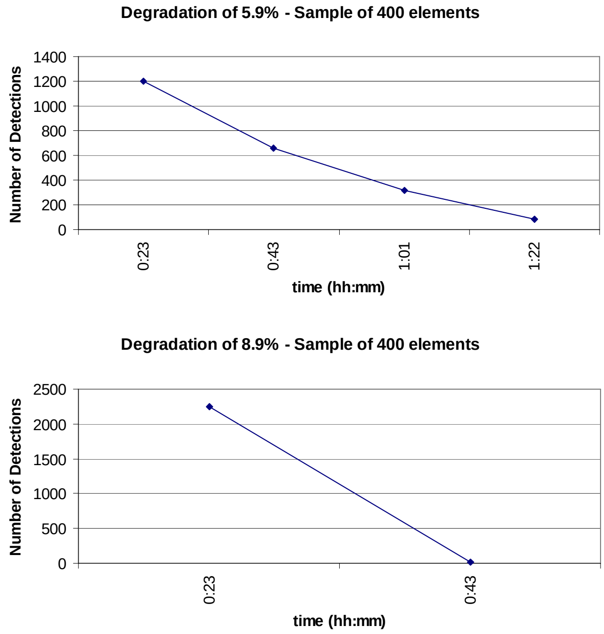

Fig. 3.4 Detection time related to resource degradation (400 elements).

4. Conclusions

Throughout this work we observe the development and behavior of Sample Space Algorithm. As we mention before the Sample Space Algorithm can be used to detect a great variety of problems, since problems of low impact to problems of high impact. This can be done using a variety of samples sizes. Larger sizes of the sample means larger sensitivity and smaller behavior dispersion, i.e., the results are more precise. However, a larger sample's size also means longer waiting time for filling it out and only then analyzing the data to detect the failures. On the other hand, smaller samples mean smaller sensitivity and larger behavior dispersion, i.e., less precision, but smaller samples also means decrease the detection time.

As future work we intend to construct others algorithms. Currently we are developing algorithms based on Neural Networks. This new algorithm will work along with Sample Space Algorithm.

Acknowledgments

Our thanks to The State of São Paulo Research Foundation (FAPESP) that supports this work.

References

[1] P. Ginzboorg, "Seven Comments on charging and billing", in Communication of the ACM, Vol.43, No II, pp: 89-92, November 2000.

[2] IPDR Organization, “Business Solution Requirements -Network Data Management-Usage (NDM-U)”, Version 3.5.0.1, November 2004.
[3] Sandra Borthick, “Call accounting and billing for IP services”, Business Communications Review, pp: 28-33, August 2001.
[4] Q. Chen, M. Hsu, U. Dayal, “A data-warehouse/OLAP framework for scalable telecommunication tandem traffic analysis”, in Proc. International Conference in Data Engineering, pp: 201-210, March 2000.
[5] M. Sestak, A.R. Kaye, “Traffic monitoring in a network of PBXs”, in Global Telecommunications Conference, 1990, vol. 1, pp: 402-407, December 1990.
[6] L. Yan, M. Fassino, P. Baldasare, “Predicting customer behavior via calling links”, in Proc. IEEE International Joint Conference, vol. 4, pp: 2555-2560, August 2005.
[7] M. C. Mozer, R. Wolniewicz, D. B. Grimes, E. Johnson, and H. Kaushansky, “Predicting subscriber dissatisfaction and improving retention in the wireless telecommunications industry”, in IEEE Transactions on Neural Networks, vol.1, pp: 690-696, 2000.
[8] W. Dong; W. Quan-yu; Z. Shou-yi; L. Feng-xia; W. Da-zhen, “A feature extraction method for fraud detection in mobile communication networks”, in Intelligent Control and Automation, vol.1, pp: 1853-1856, June 2004.
[9] S. Rosset, U. Murad, E. Neumann, Y. Idan, G. Pinkas, “Discovery of fraud rules for telecommunications-challenges and solutions”, in Proc. ACM SIGKDD International Conference on Knowledge discovery and data mining, pp: 409-413, August 1999.
[10] C. Cortes, K. Fisher, D. Pregibon, A. Rogers, F. Smith, “Hancock: A language for analyzing transactional data streams”, in ACM Transactions on Programming Languages and Systems (TOPLAS), Vol. 26, Issue 2, pp: 301-338, March 2004.
[11] S. Trisolini, M. Lenzerini, D. Nardi, “Data integration and warehousing in Telecom Italia” in proc. ACM SIGMOD international conference on Management of data SIGMOD, Vol. 28, Issue 2, pp: 538-539, June 1999.
[12] G. Nunes, “Probability Theory and Renewal Theory”, Curse of Probability theory applied to Engineering, 2001, unpublished.
[13] Montgomery, D.C., “Applied Statistics and Probability for Engineers”, Chapter III and IV, John Wiley & Sons, Inc, 1994.

Gean Davis Breda received the B.S. degrees in Electrical Engineering from Federal University of Santa Maria, Brazil, in 1993. Received his Master degree in Electrical Engineering from School of Electrical and Computer Engineering of State University of Campinas, Brazil, in 1996. During 1997-2002, he worked as a system engineering in CDMA cellular systems where he was responsible for the system support and operation. Also, since 2002 he is pursuing his Ph.D. in Electrical Engineering at School of Electrical and Computer Engineering from State University of Campinas, Brazil. His research interests include QoS in Telecommunication Systems, Billing.

Leonardo de Souza Mendes received his B.S. degree in 1985 from the Gama Filho University, Rio de Janeiro, his M.S. degree in 1987 from the Catholic University of Rio de Janeiro, and his Ph.D. degree in 1991 from Syracuse University, all in Electrical Engineering. In 1992 he joined the School of Electrical Engineering of the State University of Campinas, Brazil. Prof. Mendes’s recent R&D focus is in the studies and development of Communications Engineering applications for metropolitan IP networks. Prof. Mendes created, at UNICAMP, the Laboratory of Communications Network (LaRCom), from which he is now the Director and also the main coordinator. At LaRCom , Prof. Mendes and his group have developed or are developing the following projects: 1) an optical system simulator to help in the analysis of optical networks; 2) an environment for the simulation of systems using event driven technique which allows the development of ATM, IP and CDMA simulators; 3) development of Internet set top boxes using J2ME for small devices; 4) communications description of Internet devices using CORBA component modules for Telecommunications; 5) development of e-Learning objects for the PGL project.

Mauricio Luiz Bottoli received the B.S. degrees in Electrical Engineering from Federal University of Santa Maria, Brazil, in 1996. Received his Master degree in Electrical Engineering from School of Electrical and Computer Engineering of State University of Campinas, Brazil, in 1999. Since 2001 he is working as a telecommunication system engineering. Also, since 1999 he is pursuing his Ph.D. in Electrical Engineering at School of Electrical and Computer Engineering from State University of Campinas, Brazil. His research interests include Optical Networks, Metropolitan networks, VoIP.