Anomaly Detection Using Digital Signature of Network Segment Aiming to Help Network Management

Mario Lemes Proença Jr., Bruno Bogaz Zarpelão and Leonardo de Souza Mendes

***Abstract*— The rapid and accurate identification and diagnosis of anomalies are a fundamental step for management of today's high speed and multiservice networks. This paper presents a model for anomaly detection based on the application of BLGBA model to characterize the traffic, on three levels of sensibility alarms and on the correlation of multiples SNMP objects. The obtained results validate the experiment and show significant improvement in networks management. The main contributions of this work are: (i) case studies for traffic characterization of network servers using BLGBA model and DSNS; (ii) a model for anomaly detection; (iii) several tests of the model using real data in four network servers.**

***Index Terms*—Alarm systems, Anomaly detection, Computer network management, Traffic characterization.**

I. Introduction

Computer networks are of vital importance nowadays for modern society, comparable to essential services like piped water, electricity and telephone. Its functionality cannot be interrupted due to its importance for the people that use its services. In this context, the automation of the network management becomes fundamental for reducing costs, avoiding performance bottlenecks and early detection of network failures [6][15].

The determination of network normal behavior is an important step to detect traffic anomalies [12]. In this work, we use the DSNS (Digital Signature of Network Segment) generated by BLGBA (Baseline for Automatic Backbone Management) model for traffic characterization. The DSNS can be defined as the set of basic information that shows the traffic profile in a server or segment of the network. This profile is raised through minimum and maximum thresholds of volume of traffic, quantity of errors, and types of protocols and services that flow through this server or segment along the day. This profile can also be defined as the baseline of the network server or segment.

In order to make the management decisions on problems that might be happening in the network more reliable and safer, it is necessary to obtain close to real forecast of traffic characteristics of the segments that make up the network backbone at a given instant.

Also, the traffic characterization is important for the security management of the network. The use of the DSNS can offer more precise information related to traffic behavior. The anticipated knowledge of traffic characteristics of a given segment or server is directly related to the profile of its use and this information can be used to identify anomalies, thus reducing network downtime and increasing network reliability [1][10][13][15].

The anomaly detection techniques known as profile-based or statistical-based do not require any previous knowledge about the nature and properties of the anomalies to be detected. Their main advantages are the effectiveness in detecting unknown anomalies and the easiness to adapt to new environments. The detection is accomplished by searching for significant behavior changes in the real traffic that are not coherent with the previously established profile for the network normal behavior [6][16][18].

The definition of which events represent an anomaly and therefore must be reported to the network administrators is still an open question [1][7][14]. These events can be characterized as a physical or a logical failure that can lead to the interruption or degradation of the service offered to the end user [16].

Roughan *et al.* [13] observed that some detected anomalies, despite not being informed in the systems' logs, needed to be reported to the network administrators. These events show a large variation in the behavior of the monitored data. Lakhina *et al.* [7] proposed that an event doesn't need to cause large disturbance in the network to be considered an anomaly. By causing even a small degradation in the service offered to the end user is enough to mark the event as an anomaly and justify its notification to the system administrator.

An important resource to be used in anomalies detection is the monitoring of different SNMP objects, trying to correlate the results obtained from the analysis performed for each one

Manuscript received September 19, 2006; revised June 8, 2007. This work is supported by The State of São Paulo Research Foundation (FAPESP).

M. L. Proença Jr is with the Computer Science Department, State University of Londrina (UEL), Londrina, PR, Brazil (e-mail: proenca@uel.br)

B. B. Zarpelão (corresponding author) and L. S. Mendes are with the School of Electrical and Computer Engineering (FEEC), State University of Campinas (UNICAMP), Av. Albert Einstein, 400, Cidade Universitária, CEP: 13083-852, Campinas, SP, Brazil (e-mails: {bzarpe, lmendes}@decom.fee.unicamp.br).

ISSN 1980-6604/$20.00 © 2008 SBrT/IEEE

of these objects. Each one of them offers a particular perspective of the problem. After the correlation these perspectives converge for a single notification containing the additional information useful to the network administrator. Besides, the correlation causes a reduction of the amount of notifications generated [3][17].

The anomaly detection model proposed in this work uses only data collected from MIB-II (Management Information Base) through SNMP (Simple Network Management Protocol), both standards for IP network management, in order to perform measurements of network traffic levels. The system doesn't require packet or flow instrumentation, thus. Our contribution is a lightweight and suitable approach to detect volume anomalies, based on management standards like SNMP and MIB-II, on the comparison of the real traffic to the DSNS performed through heuristics in a hysteresis interval and on the later alarm correlation performed according a unique rule. The idea is to notify the network administrator only about the events that really present some risk to the services reliability.

This paper is organized as follows. In section 2 we discuss related work. In section 3 we describe case studies for traffic characterization of network servers using DSNS and BLGBA model. In section 4 we describe our model for anomaly detection and present some initial results. Finally, in section 5, we present conclusions and suggestions for future works.

II. Related Works

Barford *et al.* [1] propose the application of signal analysis for the detection of several kinds of anomalies. The use of wavelet filters allows the separation of the signal into groups with different characteristics. From this approach the author verified, as we have done, that the network traffic presents very clear cycles with periods that vary on a daily or weekly basis.

Lakhina *et al.* [7] propose a technique to diagnose anomalies. This diagnostic includes, besides anomaly detection which is defined as a simple indication that the network is experiencing problems, the identification of the anomalies, which means the determination of the real location of the source of the events pointed as anomalous.

Wu *et al.* [18] characterize the normal traffic operation pattern through the factorial analysis. This statistical procedure allows the translation of large data sets into a smaller set of common factors. The anomalies are detected through the calculation of the Mahalanobis distance, which is applied to compare the factors that translate the normal traffic behavior with the real traffic found in the network. The results presented by the author are related to the detection of attacks and security issues of the information that travels on the monitored network.

Roughan *et al.* [13] assumed a simple approach to decrease the number of false alarms with the use of two data sources: the SNMP management protocol and the BGP routing protocol. Based on the premise that the false alarms found in the two data sources are not related, i.e., are not simultaneous, the system generates alarms only when it finds behavior deviations in the two data sources for the same situation. In our work, an anomaly is detected only when simultaneous alarms happen for more than one SNMP object in the same period of time.

Thottan *et al.* [16] point to the importance of monitoring and successful characterization of several SNMP objects, considering that each of them can respond to several kinds of anomalies. The paper suggests that an important condition for the anomaly detection is that an SNMP object can respond to more than one anomaly and several objects can be sensible to a single anomaly. The use of appropriate sets of SNMP objects and the relationship among these objects were pointed as success factors in anomaly detection. Wu *et al.* [19] also worked on anomaly detection using SNMP objects and correlation among them, but they focused on security issues. In this work, we search for making the correlation of several SNMP objects analysis aiming to improve anomaly detection.

Hajji [4] presents a proposal of a baseline for automatic detection of network anomalies that uses asymptotic distribution of the difference between successive estimates of a network traffic model. However, Hajji's model assumes that the training data is pure, i.e., with no anomalies. In the case of the model presented in this paper, the baseline is based on real data gathered from the network segment or server.

III. Traffic Characterization: BLGBA Model and DSNS

Our model for anomaly detection is focused on DSNS (Digital Signature of Network Segment) generated by BLGBA (Baseline for Automatic Backbone Management) model. The BLGBA model and the DSNS it generates were both proposed in [11]. The main purpose to be achieved with the construction of the DSNS is the traffic characterization of the segment or server that it refers to. This characterization should reflect normal behavior expected for the traffic along the day.

For the generation of the DSNS was used the BLGBA model, which was developed based on statistical analyses. The BLGBA model is used to perform analyses for each second of the day, each day of the week, respecting the exact moment of the collection, second by second for twenty-four hours, preserving the characteristics of the traffic based on the time variations along the day.

The BLGBA algorithm is based on a variation in the calculation of *mode*, which takes the frequencies of the underlying classes as well as the frequency of the modal class into consideration. The calculation takes the distribution of the elements in frequencies, based on the difference between the greatest G_{aj} and the smallest S_{aj} element of the sample, using only 5 classes. This difference divided by five forms the amplitude h between the classes, $h = (G_{aj} – S_{aj})/5$. Then, the limits of each L_{Ck} class are obtained. They are calculated by $L_{Ck} = S_{aj} + h*k$, where Ck represents the k class (k = 1...5).

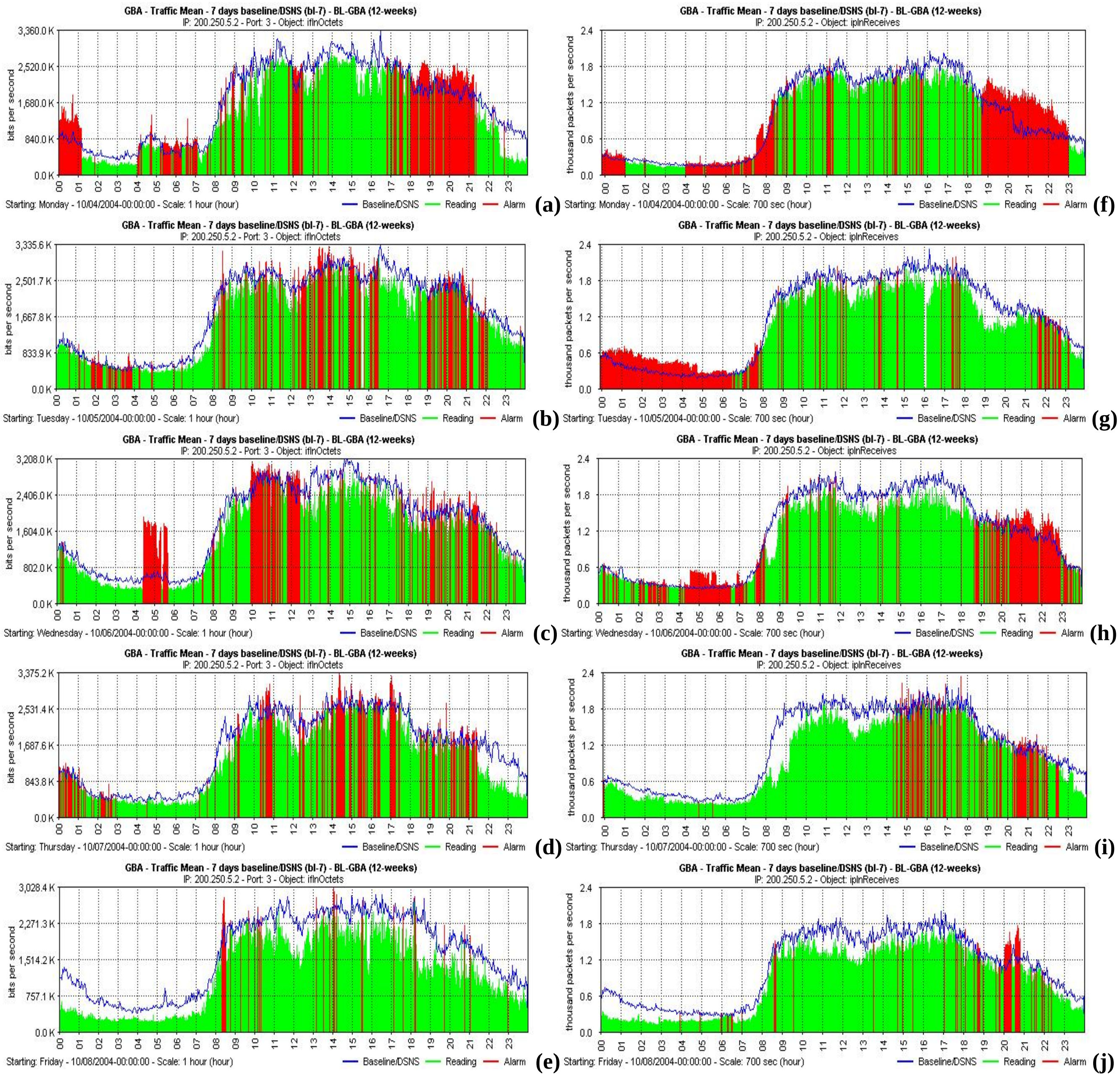

Figure 1 – DSNS and the daily movement for firewall server S_1 from 10/04/2004 to 10/08/2004.

The proposal for the calculation of the DSNS of each ***Bl_i*** second has the purpose of obtaining the element that represents 80% of the analyzed samples. The ***Bl_i*** will be defined as the greatest element inserted in class with accumulated frequency equal or greater than 80%. The purpose is to obtain the element that would be above most samples, respecting the limit of 80%.

Two types of DSNS were created, one called ***bl-7*** which consists of seven DSNS files, one for each day of the week, and the other one called ***bl-3*** which consists of three DSNS files, one for the workdays from Monday to Friday, one for Saturday and another one for Sunday. The choice for generating the DSNS separating the workdays of the week from Saturday and Sunday, was in order to minimize the margin of error in the final result, concerning the alterations in the volume of traffic that occur between the workdays and the other days.

A. *Case Studies for BLGBA Model and DSNS*

The data gathered by the GBA (Automatic Backbone Management) tool since 2003 up to the present were used for the tests and validation of our studies. This data concerning to the last three years was considered an important sample, characterized by periods of winter and summer vacation as well as holidays which contributed to the tests and validations of the ideas presented in this work. The analyzed data are related to the network segments and servers with traffic

TCP/IP based on Ethernet and ATM with LAN Emulation of State University of Londrina (UEL). We analyzed the following objects that belong to the MIB-II [9]: quantity of input and output of octets (*ifInOctets*, *ifOutOctets*), the number of IP datagrams received (*ipInReceives*) and the total number of TCP segments received from networks interfaces (*tcpInSegs*). The tests were carried out in the following segments and servers:

1. The first one which is called $\boldsymbol{S_1}$. It is the firewall of State University of Londrina (UEL) networks and gathers traffic of approximately 3000 computers to Internet;
2. The second one which is called $\boldsymbol{S_2}$. It is the main Web server of UEL;
3. The third one which is called $\boldsymbol{S_3}$. It is the Proxy server of UEL and interconnects all of 3000 computers of UEL network to Internet;
4. The fourth one which is called segment $\boldsymbol{S_4}$. It is responsible for interconnecting the ATM router to the others backbone segments of UEL networks and gathers traffic of approximately 3000 computers.

Several analytical tests had been carried out aiming to evaluate the reliability of the DSNS in relation to the real movement. We carried out tests from January 2003 to December 2005 using:

- Visual analysis of graphics containing the DSNS and its respective daily movement. As can be seen in figure 1, we found a good adjustment between the predicted (baseline/DSNS) and the real movement (reading);
- Linear Regression [5]: The results demonstrate a high correlation and adjustment between the movement that occurred in the days in relation to its DSNS;
- Test purposed by Bland & Altman [2]: Referred to the analysis of deviations occurred between the DSNS and the real movement. 95% of the deviations/errors observed during all days from January 2003 to December 2005, in servers $\boldsymbol{S_1}$, $\boldsymbol{S_2}$, $\boldsymbol{S_3}$, and $\boldsymbol{S_4}$ are between the required limits of $\overline{d} \pm 2 * s$, where $\overline{d}$ is the mean and s is the standard deviations of the differences between the DSNS and real movement. The reliability of the BLGBA model was confirmed;
- Hurst parameter (H): Carried out for the real movement and the DSNS generated by the BLGBA, using the statistical methods Variance-time, Local Whittle and Periodogram [8]. The analysis confirmed that the traffic is self-similar and the DSNS is also self-similar, however presenting a lower Hurst parameter. In most of the cases, these tests also allow us to notice that in segments with lower number of computers, the Hurst parameter presents a lower rate, between 0.6 and 0.7 and in segments with great aggregated traffic like the $\boldsymbol{S_1}$ and $\boldsymbol{S_4}$ it presents a rate between 0.8 and 1.0. Its utilization makes possible the evaluation of the DSNS quality in segments of different burstiness, indicating that the greater the burstiness of the segment, the bigger the Hurst parameter and the better the characterization shown by the DSNS, and the lower the burstiness of the segment, the smaller the Hurst parameter and worse the results shown by the DSNS. These results are corroborated by the other tests utilized to validate the DSNS that also indicate an increase of the DSNS quality in segments with a higher burstiness;
- Residual analysis [5]: The results showed that the BLGBA is a good model for predicting the traffic for analyzed segments and servers;

Figure 1 illustrates in the form of a histogram the daily movement of the server $\boldsymbol{S_1}$, and their respective DSNS. Figure 1 (a), (b), (c), (d) and (e) shows the workdays using SNMP object *ifInOctets* and figure 1 (f), (g), (h), (i) and (j) for SNMP object *ipInReceives* concerning to a week of October 2004.

We came to the following conclusions with the results of our experiences with traffic characterization of network servers:

1. In [11] the DSNS generated by BLGBA model was used for monitoring network switches and one router. In this work we utilize the DSNS to analyze network servers and one router. The SNMP objects analyzed in [11] were *ifInOctets* and *ifOutOctets*. In this work we analyze *ifInOctets, ifOutOctets, ipInReceives* and *tcpInSegs*. Here, we can conclude that the DSNS generated by BLGBA model is also applicable to these SNMP objects and for network servers. The obtained results show the validity of the BLGBA model for the generation of the DSNS, bearing in mind the performed analyses and the comparison with the real movement that occurred;
2. The DSNS is influenced by time factors which, in this case, are related to the working day that starts at 8:00 a.m. and finishes at 10:00 p.m.;
3. For the Web server $\boldsymbol{S_2}$ and Proxy server $\boldsymbol{S_3}$ it was performed the traffic characterization for the objects *ifInOctets*, *ifOutOctets*, *ipInReceives* and *tcpInSegs*. It was observed, in each server, that the DSNS and the real movement for the objects being monitored present results that are numerically different, considering the nature of each object, but that, otherwise, show a visually similar movement;
4. For $\boldsymbol{S_4}$, UEL’s router and $\boldsymbol{S_1}$, UEL’s firewall, the characterization was performed with the objects *ifInOctets*, *ifOutOctets* and *ipInReceives*. These network servers do not establish network

connection with their users, and this is why the object *tcpInSegs* was not used. The DSNS behavior with respect to the real movement for the analyzed objects presented good results, like those shown in Figure 1 for the firewall $\boldsymbol{S_1}$. However, it must be observed that the traffic variations, in relation to the DSNS, happen with different intensities for *ifInOctets* and *ipInReceives*, as can be seen, for example, from Figures 1 (c) and (h), at 05.00 am;

5. The generated DSNS fulfill their main objective which is the characterization of the traffic in the analyzed servers and segments;

IV. Anomaly Detection

In this paper, our goal is to inform network administrators about the exact moment of the occurrence of an anomalous event. A system for anomaly detection, which we call ADGBA (Anomaly Detection for Automatic Backbone Management), was developed to act in real time with the monitoring performed on SNMP objects.

Figure 2 presents the reference model, discussed in this work, of the anomaly detection system ADGBA. The model is composed of two modules. The first one, named Multilevel Alarms Module, is responsible for comparing the real data collected from SNMP objects with the DSNS and generating alarms with three levels of sensibility, *yellow*, *red* and *black*. The second one, named Correlation Module, is responsible for correlating the alarms generated on the multiple SNMP objects, in order to detect the occurrence of anomalies.

The anomaly detection system ADGBA will analyze the samples collected second by second from SNMP objects and inform the network administrator in case of an anomaly is detected.

The construction of ADGBA system was carried out based on the thresholds established by the DSNS and a hysteresis mechanism. The key idea is that the network administrator will only receive a notification if a significant deviation from network normal behavior occurs that justifies his/her attention.

For more accuracy in anomaly detection and to reduce false positive alarms, we create three levels of alarm sensibility. The first one called *yellow* is a level more sensible to behavior deviations. The second one called *red* is an intermediate level that can be used in situations that not require higher level of sensibility. The last one called *black* is harder monitoring that effectively can be used in situation where the network administrator will be informed only if significant changes in relation to the normal behavior of the traffic occur.

The anomaly detection mechanism establishes a window of time ***t*** for anomaly detection, that we call hysteresis window. In this window, deviations from the DSNS will be analyzed. The intent of hysteresis window is to reduce the probability of false alarms, generated from transient behavior of burst traffic. For *yellow* alarms ***t*** = 300 seconds, *red* alarms ***t*** = 600 seconds and *black* alarms ***t*** = 900 seconds.

The alarms are generated by the Multilevel Alarms Module only if the three following rules are broken simultaneously:

- **Rule 1:** the analyzed sample is higher or lower than the superior or inferior thresholds established in the DSNS. In (1), sample read is represented by ***x*** and DSNS threshold is represented by ***y***.
- **Rule 2:** the analyzed sample is higher or lower than the previous sample that broke rule 1 within the interval ***t***. In (1), previous sample is represented by ***v***.
- **Rule 3:** the quantity of samples that broke rules 1 and 2 is higher than δ. In (1), this quantity of samples is represented by ***z***.

$$(\forall x)(P(x,y)\wedge Q(x,t)\wedge P(x,v)\wedge P(z,\delta))\rightarrow x\in A \qquad (1)$$

In (1), the alarms' generation algorithm is showed as a predicate logic formula. The predicate ***P(a,b)*** is true if ***a>b*** and the predicate ***Q(a,t)*** is true if the timestamp of sample ***a*** is contained within the time interval ***t***. Finally, ***A*** is the set of real samples for which alarms were generated.

Rule number 3 was included in order to prevent excessive number of alarms occasioned by momentary bursts. The tests showed a relation which is inversely proportional to δ with respect to the quantity of alarms generated. In other words, the bigger the δ, the smaller will be the quantity of generated alarms. After several analyses, it was noticed that an acceptable value for the relation of alarms and problems that occurred implied in a δ equal to 130, 260 or 390, considering an interval of hysteresis of 5, 10 or 15 minutes. Figure 3 presents the automaton of the algorithm implemented for the three rules described above that compose the alarm system used for anomaly detection.

After implementation of the mechanism of hysteresis that forecasts the use of the δ accumulator for the generation of alarms, it was possible to notice the occurrence of a very small number of alarms per day, signaling effectively that something different was happening in the analyzed segment or server at that moment.

The Correlation Module receives information from the Multilevel Alarm Module. Its goal is to promote the correlation among alarms that occur within the same hysteresis interval and in different SNMP objects. Our approach for correlation is based on the similarity presented by the behavior of the monitored SNMP objects, that belong to three different MIB-II groups: *interface*, *ip* and *tcp*. An anomaly will be detected and notified if the total number of alarms generated for different objects in the same time interval be equal or higher than two.

Figure 4 presents the automaton for correlation system used by ADGBA. It is important to observe that in figure 4 the transition from state 8 to state 9 shows the application of spatial correlation and the transition from state 9 to 11 shows the application of temporal correlation.

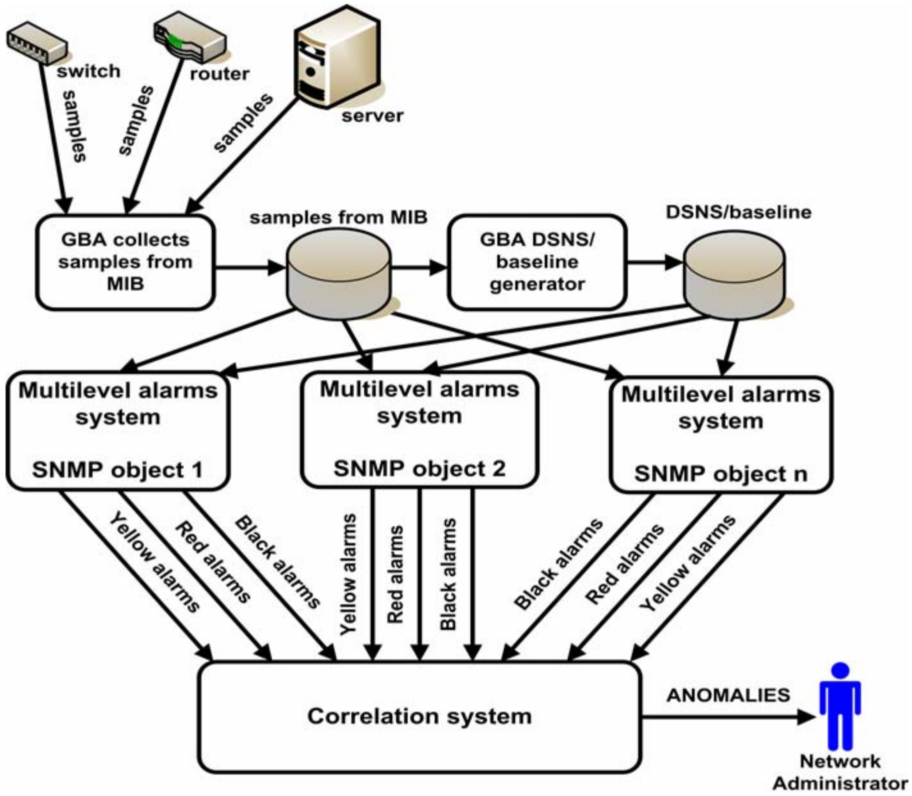

Figure 2 – Reference model for ADGBA

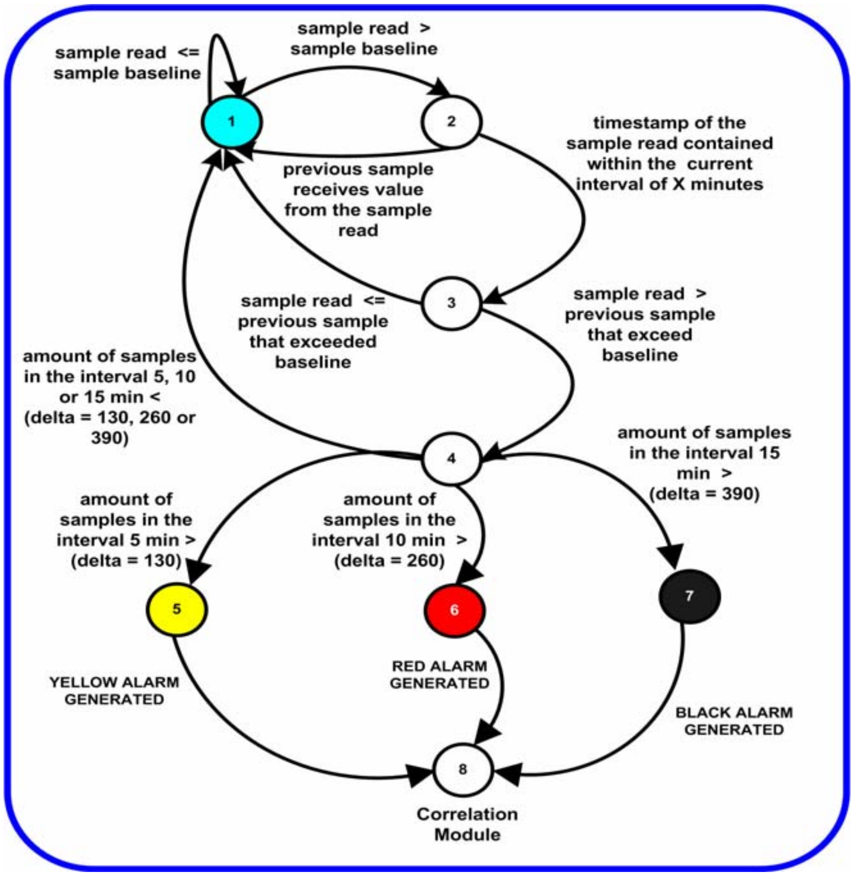
sample read <= sample baseline
sample read > sample baseline
1
2
previous sample receives value from the sample read
timestamp of the sample read contained within the current interval of X minutes
3
sample read <= previous sample that exceeded baseline
sample read > previous sample that exceed baseline
amount of samples in the interval 5, 10 or 15 min < (delta = 130, 260 or 390)
4
amount of samples in the interval 15 min > (delta = 390)
amount of samples in the interval 5 min > (delta = 130)
amount of samples in the interval 10 min > (delta = 260)
5
6
7
YELLOW ALARM GENERATED
RED ALARM GENERATED
BLACK ALARM GENERATED
8
Correlation Module

Figure 3 - Automaton for Multilevel Alarms Module using DSNS and the real movement.

A. Tests and Results for ADGBA System

Aiming to evaluate the effectiveness of ADGBA system, real data gathered in the State University of Londrina network environment since 2003 up to the present were used. The results obtained from these experiments are presented in this section.

In figures 5 and 6 are shown histograms of the daily mean of alarms that were generated by Multilevel Alarms Module from July to December 2004 in Web server S_2 and Proxy server S_3 for *ifInOctets*, *ipInReceives* and *tcpInSegs* objects. These figures show the daily mean of alarms *yellow*, *red* and *black* for many objects in the same server. It can be observed that the averages of alarms generated for different objects in the same server are quite similar, what suggests that they were generated for the same events. Thus, there is a correlation between the behavior of SNMP objects in the servers when they face anomalies.

In figure 7 is showed an example of situation at Proxy server S_3 in 11/5/2004, where we can see 23 *yellow* alarms, 9 *red* and 6 *black* for *ipInReceives* object and 16 *yellow* alarms, 9 *red* and 3 *black* for *tcpInSegs* object. In this example we can observe that a great difference between the real movement and DSNS occurred in all analyzed SNMP objects, which was informed by ADGBA as an anomaly. The differences are greater than 100% and occur simultaneously in different objects, indicating the occurrence of an anomalous state that had to be noticed by network administrator.

Figure 8 presents an example of a Web server S_2 state of operation, which happened in 12/1/2004, where it is possible to observe the occurrence of *yellow*, *red* and *black* alarms for the objects *ifInOctets*, *ipInReceives* and *tcpInSegs*, showing an operation in disagreement with the normal expected state. The number of alarms is not equal for all the objects and sensitivity levels, however, they do occur in the same hysteresis interval and suggest the simultaneous occurrence of an abnormality in the objects being analyzed. All the alarms, which were generated by Multilevel Alarms Module, converged to a unique and more complete anomaly notification, which were generated by the Correlation Module.

Tables I, II, III and IV present a summary of the daily mean of *yellow*, *red* and *black* alarms and anomaly notifications generated in the analyzed servers during six months. The anomalies column is only accounted when it was observed alarms in more than one object being analyzed. It’s important to notice that only a small quantity of anomaly notifications with respect to the daily mean number of alarms was observed. Thus, the correlation of alarms allows that a great number of alarms converge to a small number of anomaly notifications, decreasing the volume of notifications sent to the network administrator without losing accuracy on the detection.

Figure 9 presents a histogram of all occurrences of anomalies and of *yellow*, *red* and *black* alarms for the objects *ipInReceives* and *tcpInSegs* during September, 2004 in server S_3. It is possible to verify the occurrence of alarms for different objects and see that only 7 anomalies occurred during days 9, 12, 16, 21, 23, 26 and 30.

It was observed that the same objects in different servers presented diverse behaviors with respect to the anomalous events. For instance, the object *ifInOctets* presented a large sensitivity to variations in the traffic with respect to the DSNS in the Firewall server S_1, while in router S_4 a smaller sensitivity was observed. Now considering object *ipInReceives*, the opposite result was observed with the same servers. With respect to the servers that establish TCP connections with network clients, like servers S_2 and S_3, the behaviors of these SNMP objects were similar.

The results obtained for the objects *ipInReceives* and *tcpInSegs* were very similar for the servers S_2 and S_3. We believe that these results are due to their need for TCP connections.

V. Conclusions

This work presented a contribution related to the automatic generation of Digital Signature of Network Segment (DSNS) for network servers, which constitutes itself into an important mechanism for the characterization of the traffic of the analyzed servers, through thresholds that reflect the real expectation about the volume of traffic respecting the time characteristics along the day and the week.

With our experiences realized we can conclude that, apart from the behavior of the traffic of the Ethernet networks being random, self-similar and extremely influenced by the quantity of bursts, which intensify as the number of hosts connected to the segment increase, as shown in [8], the BLGBA model chosen for the characterization of the DSNS showed to be viable for the characterization of the traffic in backbone segments and servers that concentrate the traffic of a great number of hosts.

With the use of graphs such as the ones shown in figure 1 including information of the Digital Signature of Network Segment (DSNS), generated by the BLGBA model, and the daily movement, a better control in the segments follow-up was obtained.

Another contribution is the anomaly detection system ADGBA, which uses the DSNS and a Multilevel Alarms Module integrated with a Correlation Module to detect anomalies in a rapid and accurate way. The ADGBA makes it possible for the network administrator to be informed through notifications, at the exact moment a significant difference related to the expected traffic and the DSNS was found out.

Future works include the creation of a multiparametric model for anomaly detection aiming to find and locate the origin of anomaly in the network backbone. The key idea is to detect and localize the problem quickly in order to avoid network disrupt and reducing downtime.

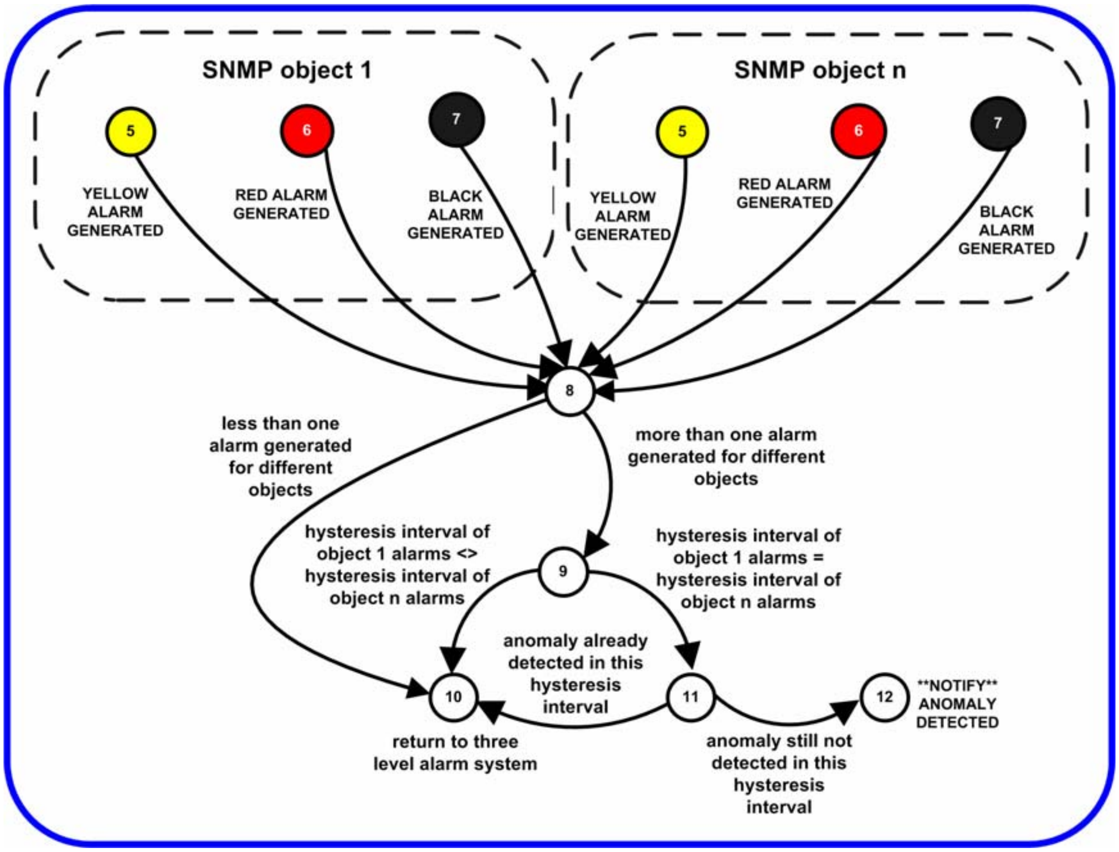
SNMP object 1
5
6
7
YELLOW ALARM GENERATED
RED ALARM GENERATED
BLACK ALARM GENERATED
SNMP object n
5
6
7
YELLOW ALARM GENERATED
RED ALARM GENERATED
BLACK ALARM GENERATED
8
less than one alarm generated for different objects
more than one alarm generated for different objects
9
hysteresis interval of object 1 alarms <> hysteresis interval of object n alarms
hysteresis interval of object 1 alarms = hysteresis interval of object n alarms
anomaly already detected in this hysteresis interval
10
11
12
NOTIFY ANOMALY DETECTED
return to three level alarm system
anomaly still not detected in this hysteresis interval

Figure 4 – Automaton for Correlation Module.

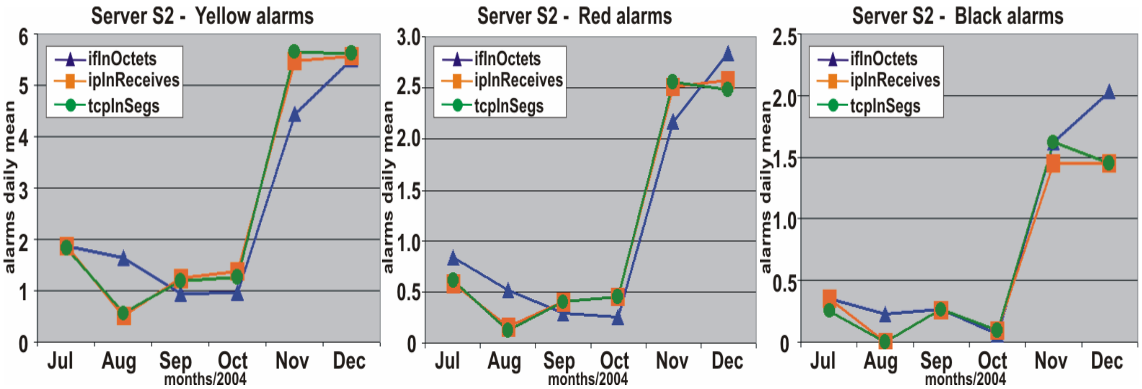

Figure 5 – Six month alarms mean for Web Server S_2.

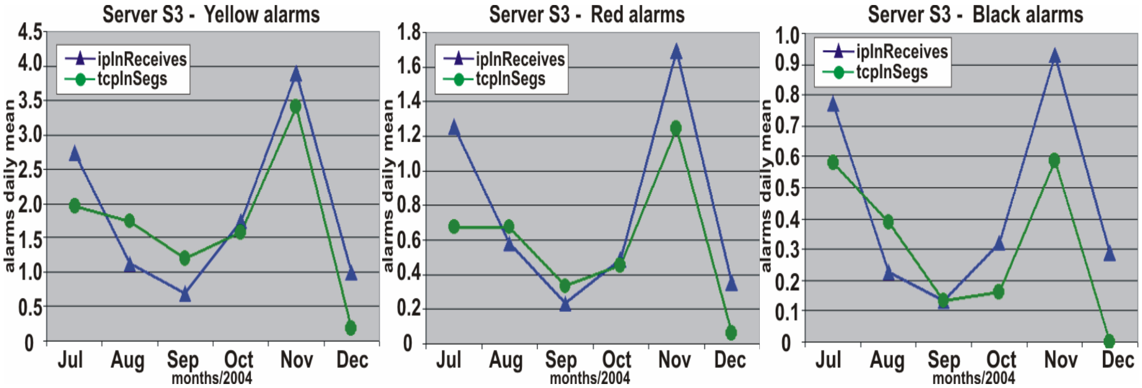

Figure 6 – Six month alarms mean for Proxy Server S_3.

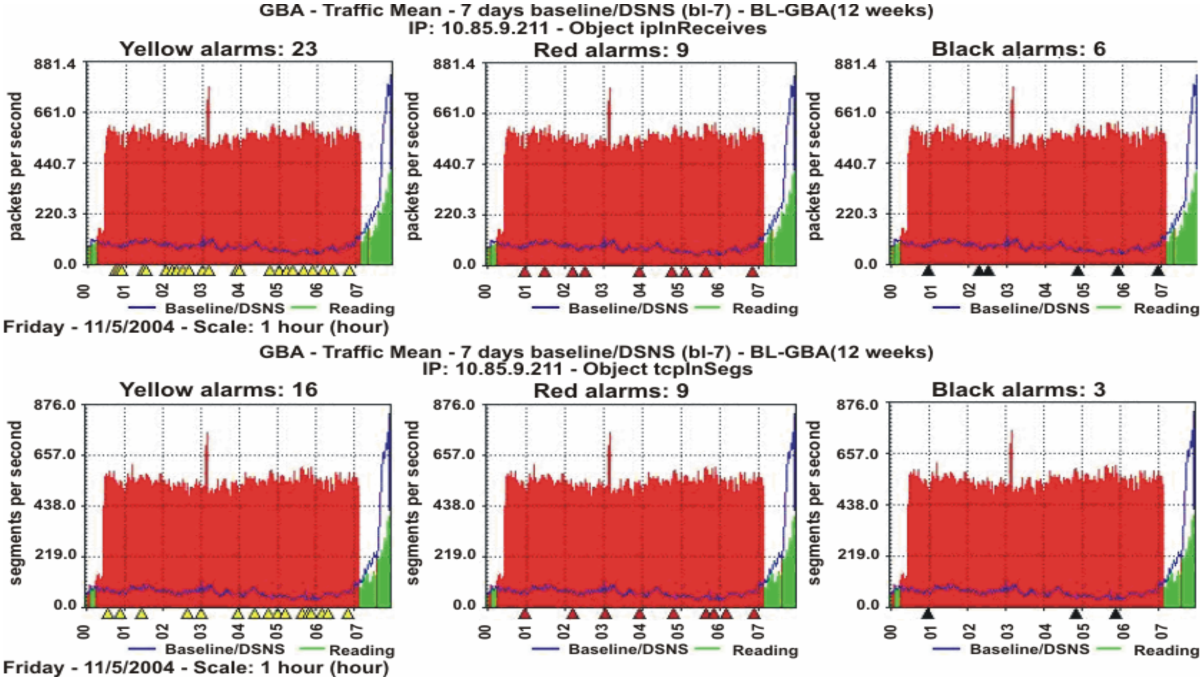

Figure 7 – Anomaly situation in Proxy server $\boldsymbol{S_3}$ and alarms generated

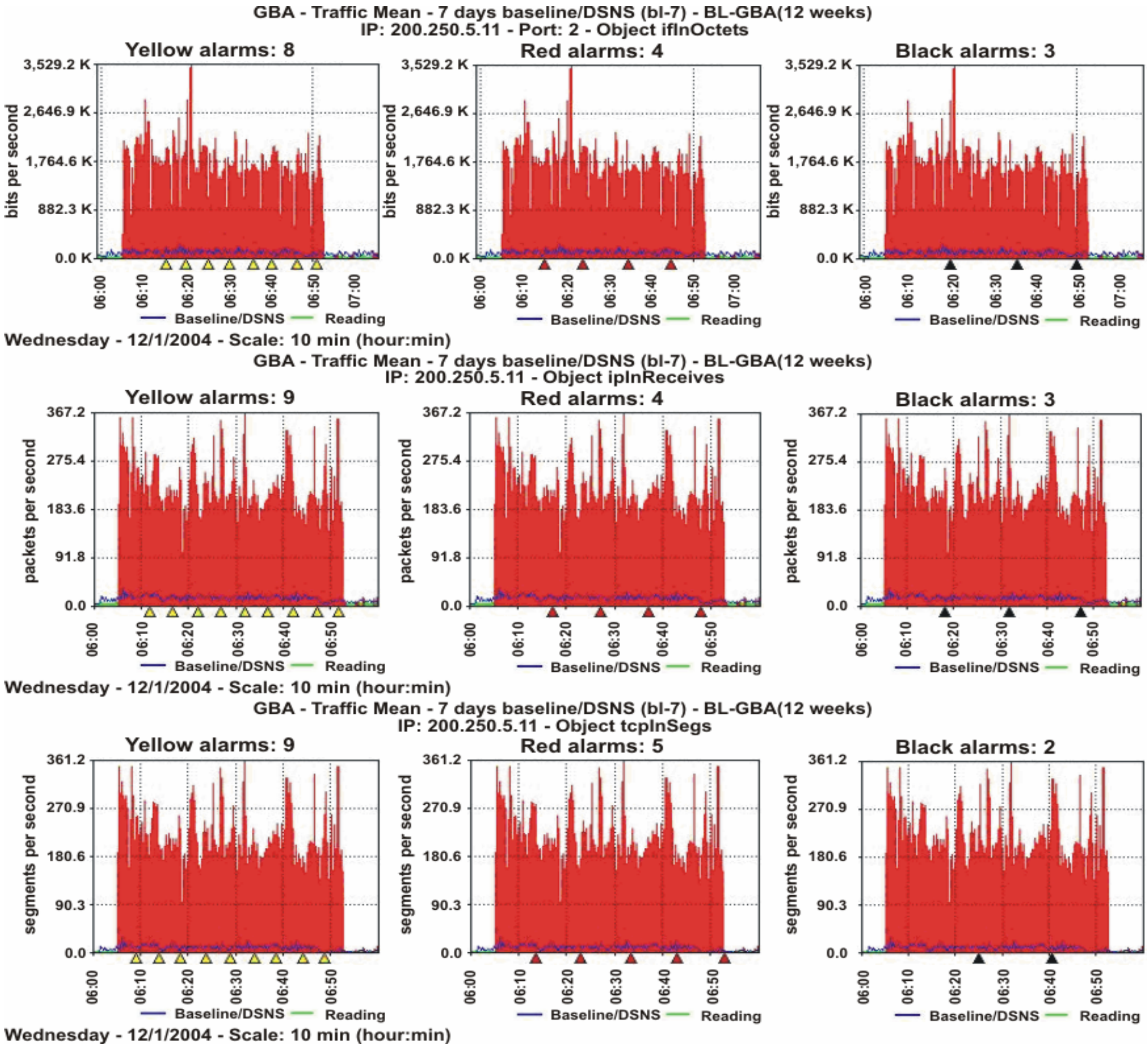

Figure 8 – Anomaly situation in Web server $\boldsymbol{S_2}$ and alarms generated.

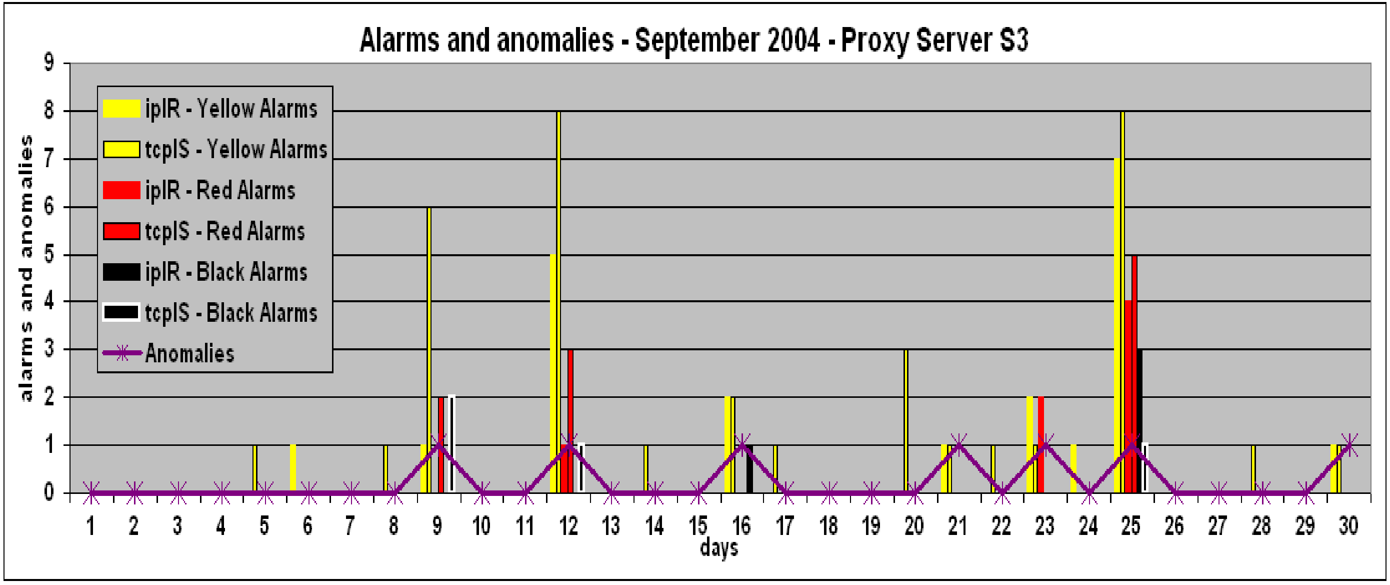

Figure 9 – Alarms and anomalies occurred in September 2004 for server S_3.

TABLE I
DAILY MEAN OF ALARMS AND ANOMALY NOTIFICATIONS FOR SERVER S_1

	Yellow ifIO	Yellow ipIR	Red ifIO	Red ipIR	Black ifIO	Black ipIR	Anomalies
Jul	2.84	1.45	1.39	0.81	0.77	0.42	0.06
Aug	1.35	1.23	0.42	0.45	0.13	0.39	0.03
Sep	8.07	1.40	3.50	0.67	2.33	0.47	0.07
Oct	7.48	0.71	3.26	0.26	2.00	0.13	0.03
Nov	3.90	0.31	1.72	0.10	1.07	0.00	0.10
Dec	2.61	1.32	1.23	0.52	0.87	0.29	0.06

TABLE II
DAILY MEAN OF ALARMS AND ANOMALY NOTIFICATIONS FOR SERVER S_2

	Yellow ifIO	Yellow ipIR	Yellow tcpIS	Red ifIO	Red ipIR	Red tcpIS	Black ifIO	Black ipIR	Black tcpIS	Anomalies
Jul	1.84	1.87	1.84	0.87	0.58	0.61	0.35	0.35	0.26	0.35
Aug	1.65	0.52	0.55	0.52	0.16	0.13	0.23	0.00	0.00	0.26
Sep	0.93	1.23	1.20	0.30	0.40	0.40	0.27	0.27	0.27	0.27
Oct	0.97	1.39	1.26	0.26	0.45	0.45	0.06	0.10	0.10	0.52
Nov	4.45	5.48	5.66	2.17	2.52	2.55	1.62	1.45	1.62	0.63
Dec	5.52	5.58	5.65	2.84	2.58	2.48	2.03	1.45	1.45	0.77

TABLE III
DAILY MEAN OF ALARMS AND ANOMALY NOTIFICATIONS FOR SERVER S_3

	Yellow ipIR	Yellow tcpIS	Red ipIR	Red tcpIS	Black ipIR	Black tcpIS	Anomalies
Jul	2.74	1.97	1.26	0.68	0.77	0.58	0.19
Aug	1.13	1.74	0.58	0.68	0.23	0.39	0.35
Sep	0.70	1.20	0.23	0.33	0.13	0.13	0.23
Oct	1.74	1.58	0.48	0.45	0.32	0.16	0.35
Nov	3.90	3.41	1.69	1.24	0.93	0.59	0.23
Dec	1.00	0.19	0.35	0.06	0.29	0.00	0.06

TABLE IV
DAILY MEAN OF ALARMS AND ANOMALY NOTIFICATIONS FOR SERVER S_4

	Yellow ifIO	Yellow ipIR	Red ifIO	Red ipIR	Black ifIO	Black ipIR	Anomalies
Jul	7.45	14.61	3.45	6.77	2.16	4.35	0.39
Aug	1.55	17.23	0.61	8.29	0.42	5.68	0.16
Sep	3.83	7.73	1.67	3.70	1.00	2.47	0.10
Oct	6.71	8.42	2.68	4.16	1.48	2.58	0.19
Nov	13.14	11.03	6.00	5.45	3.48	3.52	0.27
Dec	7.23	16.81	3.06	7.87	1.87	5.26	0.32

REFERENCES

[1] P. Barford, J. Kline, D. Plonka and A. Ron, "A signal analysis of network traffic anomalies", *Internet Measurement Workshop; Proceedings of the second ACM SIGCOMM Workshop on Internet measurement*, Marseille, 2002, pp. 71-82.

[2] J. M. Bland and D. G. Altman, "Statistical Methods For Assessing Agreement Between Two Methods of Clinical Measurement", *The LANCET,* pp. 307-310, February, 1986.

[3] R. D. Gardner and D. A. Harle, "Methods and Systems for Alarm Correlation", *Proceedings of Global Telecommunications Conference, 1996, GLOBECOM'1996*, London, 1996, p. 136-140.

[4] H. Hajji, "Statistical Analysis of Network Traffic for Adaptive Faults Detection", *IEEE Transaction on Neural Networks*, v. 16, n. 5, pp. 1503-1063, 2005.

[5] R. Jain, *"The Art of Computer Systems Performance Analysis, Techniques for experimental design, measurement, simulation and modeling"*, Willey Computing, 1991.

[6] B. Krishnamurthy, S. Sen, Y. Zhang and Y. Chen, "Sketch-based change detection: methods, evaluation, and applications", *Internet Measurement Workshop Proceedings of the 2003, ACM SIGCOMM conference on Internet measurement*, Miami Beach, 2003, p. 234 – 247.

[7] A. Lakhina, M. Crovella and C. Diot, "Diagnosing network-wide traffic anomalies", *Proceedings of ACM SIGCOMM'04*, Portland, Oregon, USA, 2004, pp. 219-230.

[8] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the Self-Similar Nature of Ethernet Traffic (Extended Version)", *IEEE/ACM Transactions on Networking,* volume 2, No 1, February 1994.

[9] K. McCloghrie and M. Rose, "Management Information Base for Network Management of TCP/IP-based internet: MIB-II", *RFC 1213*, mar 1991.

[10] S. Papavassiliou, M. Pace, A. Zawadzki, and L. Ho, "Implementing enhanced network maintenance for transaction access services: tools and applications", *IEEE International Conference on Communications ICC 2000*, New Orleans, 2000, pp. 211-215 .

[11] Mario Lemes Proença Jr., C. Coppelmans, M. Bottoli, A. Alberti and Leonardo de Souza Mendes "The Hurst Parameter for Digital Signature of Network Segment", *11th INTERNATIONAL CONFERENCE ON TELECOMMUNICATIONS - ICT 2004*, Fortaleza, 2004, pp. 772-781.

[12] Mario Lemes Proença Jr., Bruno Bogaz Zarpelão and Leonardo de Souza Mendes, "Anomaly Detection for Network Servers using Digital Signature of Network Segment", *Proceedings of IEEE Advanced Industrial Conference on Telecommunications 2005*, Lisbon, 2005, pp. 290-295.

[13] M. Roughan,T. Griffin, Z. M. Mao, A. Greenberg and B. Freeman, "IP forwarding anomalies and improving their detection using multiple data sources". *Proceedings of the ACM SIGCOMM workshop on Network troubleshooting: research, theory and operations practice meet malfunctioning reality*, Portland, Oregon, USA, 2004, pp. 307-312.

[14] R. Sekar, A. Gupta, T. Shanbhag, A. Tiwari, H. Yang, and S. Zhou, "Specification-based Anomaly Detection: A New Approach for Detecting Network Intrusions", *Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS'02),* Washington, 2002, p. 265-274.

[15] A. Soule, K. Salamatian, and N. Taft, "Combining Filtering and Statistical Methods for Anomaly Detection", *Proceedings of ACM SIGCOMM Internet Measurement Conference 2005 (IMC'05)*, Berkeley, 2005, pp. 317-330.

[16] M. Thottan and C. Ji, "Anomaly detection in IP networks", *IEEE Transactions on Signal Processing,* Volume: 51, Issue: 8, Aug. 2003.

[17] A. Valdes and K. Skinner, "Probabilistic Alert Correlation" *Proceedings of Recent Advances in Intrusion Detection : 4th International Symposium*, Davis, 2001, p. 54-68.

[18] N. Wu and J. Zhang, "Factor analysis based anomaly detection", *Information Assurance Workshop, 2003. IEEE Systems, Man and Cybernetics Society*, 2003, pp. 108-115.

[19] Q. Wu and Z. Shao "Network Anomaly Detection Using Time Series Analysis" *Proceedings of the Joint International Conference on Autonomic and Autonomous Systems and International Conference on Networking and Services (ICAS/ICNS 2005)*, 2005, pp. 42-47.

Mario Lemes Proença Jr. received his M.Sc degree in Computer Science from the Computer Science Institute of Federal University of Rio Grande do Sul, Porto Alegre, Brazil, in 1998 and his Ph.D. degree in Electrical Engineering from School of Electrical and Computer Engineering of State University of Campinas, Brazil in 2005. Also, he is a computer science professor since 1991 in State University of Londrina, Brazil. His research interests include Computer Network, Network Operations and Management and Security. He currently is leader of the group of research in computer networks of Computer Science Department of State University of Londrina.

Bruno Bogaz Zarpelão received his B.S. degree in Computer Science from State University of Londrina, Brazil. He is currently pursuing his Ph.D. in Electrical Engineering at School of Electrical and Computer Engineering from State University of Campinas, Brazil. His research interests include Computer Network Management and Operations and Anomaly Detection using SNMP and MIB-II.

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.