Anomaly Detection using Spatio-Temporal Measures

Syed Azahad R. Lakshmi Tulasi

Abstract:
With the development of network technology and growing enlargement of network size, the network structure is becoming more and more complicated. Mutual interactions of different network equipment, topology configurations, transmission protocols and cooperation and competition among the network users inevitably cause the network traffic flow which is controlled by several driving factors to appear non-stationary and complicated behavior. Because of its non-stationary property it cannot easily use traditional way to analyze the complicated network traffic. We present different approaches to characterize traffic: (i) a model-free approach based on the method of types and Sanov's theorem, (ii) a model-based approach modeling traffic using a super statistics theory (iii) another model –based approach using Markov modulated process. Using these characterizations as a reference we continuously monitor traffic and employ large deviations and decision theory results to "compare" the empirical measure of the monitored traffic with the corresponding reference characterization, thus, identifying traffic anomalies in real-time. According to the super statistics theory, the complex dynamic system may have a large fluctuation of intensive quantities on large time scales which cause the system to behave as non-stationary which is also the characteristic of network traffic. Partitioning the non-stationary traffic time series into small stationary segments which can be modeled by discrete Generalized Pareto (GP) distribution. Different segments follow GP distribution with different distribution parameters which are named slow parameters. Throughout, we compare these two approaches presenting their advantages and disadvantages to identify and classify temporal network anomalies. We also demonstrate how our framework can be used to monitor traffic from multiple network elements in order to identify both spatial and temporal anomalies. We validate our techniques by analyzing real traffic traces with time-stamped anomalies.

Index terms:
Large deviations, Markov processes, method of types, Super statistics, Pareto distribution, network traffic.

1. Introduction

With the fast increase of network connections, the problem of intrusion detection becomes more and more important [1]. Although internet service can provide useful information due to its open property, it should also be noticed that the number of network intrusions increases faster than before, which introduces a lot of inconvenience to the users [2]. Network anomaly detection approaches can be broadly grouped into two classes: signature-based anomaly detection where known patterns of past anomalies are used to identify ongoing anomalies (e.g., see [1], [2] for intrusion detection), and anomaly detection which identifies patterns that substantially deviate from normal patterns of operation [3]. Earlier work has showed that systems based on pattern matching had detection rates below 70% [4],

In this work we focus on anomaly detection and in particular on statistical anomaly detection, where statistical methods are used to assess deviations from normal operation. Our main contribution is the introduction of a new statistical traffic anomaly detection framework that relies on identifying deviations of the empirical measure of some underlying stochastic process characterizing system behavior. In contrast with other approaches [1], [2], [6], we are not trying to characterize the abnormal operation, mainly because it is too complex to identify all the possible anomalous instances (especially those that have never been observed).

The main advantages of network traffic anomaly detection based on the characteristic quantity are as follows: the number of the characteristic quantity is far lower than the original network flow, so it only spends less time to complete the detection.

Network traffic anomaly detection based on statistical model establishes the statistical model first with comprehensive consideration of all of properties of network traffic, and then predicts network flow according to the model, finally detects on the basis of the difference between the prediction results and the actual results.

More specifically, we propose methods to characterize normal behavior: (a) a model-free approach employing the method of types [7] to characterize the type (i.e., empirical measure) of an independent and identically-distributed (i.i.d.) sequence of appropriately averaged system activity, and (b) a model-based approach where system activity is modeled using a Markov Modulated Process (MMP). Given these characterizations, we employ the theory of Large Deviations (LD) [7] and decision theory results to assess whether current system behavior deviates from normal. LD theory provides a powerful way of handling rare events and their associated probabilities with an asymptotically exact exponential approximation.

Non-stationarity in system activity can also cause problems to our approach as it may be responsible for legitimate distributional differences between past and current activity. However, as long as stationarity holds over relatively short periods of time one could often update the reference trace with more recent and relevant activity, thus, reducing the possibility of misdetections and false alarms.

Manuscript received July 5, 2013
Manuscript revised July 20, 2013

The abnormal traffic flow caused by attack, the network flow exhibits some basic characters, such as non-stationary, heavy-tailed property, Long-Range Dependence (LRD), abrupt[5,6,7,8], etc. In the posterior part of the experiment, the non-stationary is proved according to the autocorrelation function, which shows that the parameters change in the process.

The model-free approach aggregates traffic over short time intervals to which we will refer to as time buckets. Although the correlation between samples in short time scales is significant, it reduces rapidly between aggregates over a time bucket. Hence, we consider the sequence of traffic aggregates over a time bucket as an i.i.d. sequence and employ the method of types to characterize its distribution. Our model-based approach uses an MMP process to model legitimate traffic during some time-of-day interval. Earlier work has shown that MMP models can accurately characterize network traffic [9], [10], at least for the purposes of estimating important quality-of-service metrics.

Therefore, aiming at these kinds of complicated problems, the super statistics theory [11, 12, and 13] has been put forward to relate with the network flow, which is suitable for the change of the statistical parameters. We propose to use a more complex method which comprise the conception of 'statistics of statistics' (that is super statistics', SS) to model the network traffic.

Our experimental results demonstrate that our characterizations, based on low-dimensional models, do a respectful job in identifying anomalies and also show that the abnormal traffic flow is a kind of complicated changing process, which is non-stationary, random and abrupt. Network traffic abnormity detection can be completed through the research on the decisive distribution parameters which are named slow parameters and the adaptive detection method.

2. Datasets for desired analysis

We validate our methodology against real traffic traces from a backbone network. Our source of data is the IP-level traffic flow measurements collected from every point of presence (PoP) in the Abilene Internet2 backbone network. Abilene is the major academic network, connecting over 200 universities in the US, and peering with other research networks in Europe and Asia. Abilene has 11 PoPs resulting in 121 origin–destination flows. The data we are using is sampled flow data from every router of Abilene for a period of one week (April 7 to 13, 2003). Sampling is random capturing 1% of all packets entering every router. Three different representations (features) of sampled flow data are used, a time-series of the number of bytes (B), of packets (P) and of flows (F). In order to avoid synchronization issues, the measurements are aggregated into 5-minute bins. The issue of how packets are sampled is an important one but we do not consider it here because, in most cases, packet sampling is predetermined by the monitoring instrumentation.

Our experiments used actual network traffic taken from the MIT Lincoln Laboratory [14, 15, and 16]. This is second (1998, 1999, 2000) in a series of data sets created at MIT, under a DARPA sponsored project to evaluate intrusion detection systems, and to guide research directions. The DARPA1999 dataset was created by group of the MIT Lincoln Laboratory to conduct a DARPA-sponsored comparative evaluation of different IDS. It is the reference dataset in the evaluation of IDS performance. The dataset is made of five weeks of network traffic traces extracted from a simulated military department network. It consists of two components, seven weeks of training data with labeled attack and two weeks of unlabeled test data. Each data set includes tcpdump file, tcpdump list file which is labeled with attack information, Solaris BSM audit data, and ps monitoring data. Only tcpdump files are used that record the network traffic information to analyze anomalies. We use the DARPA1999 datasets in weeks 1 and 2 as the normal traffic and detect the abrupt change of the traffic data in weeks 3 and 4 by taking the package counts per second as the observation.

In Fig.1, we show the original data on week 4, day 3. It is overlapped data containing normal and anomaly data. The network is shown in Fig. 1 and its Autocorrelation Function (AF) is increases with the time delay, the relevance of the series is still significant, and there is not convergence to 0, the data series has obvious LRD property.

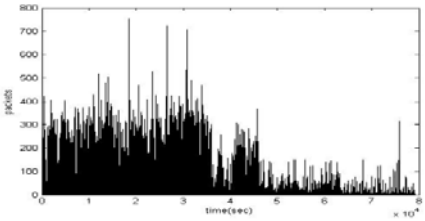

Fig.1 The original data of DARPA on week 4, day 3

Fig.1 The original data of DARPA on week 4, day 3

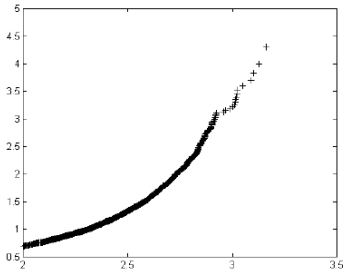

Fig 2 inverse cumulative distribution function log-log plot of darpa on week 4, day 3

3. A MODEL-FREE APPROACH: Sanov's Theorm

In this section we discuss our model-free approach and provide the structure of an algorithm to detect temporal network anomalies. We assume that the *traffic trace* we monitor (in bits/ bytes/ packets/ flows per time unit), corresponding to a specific time-of-day interval, can be characterized by a stationary model over a certain period.

Consider a time series $X_1,\ldots,X_n$ of traffic activity (say, in bits/bytes/packets/flows per sample). Let Y_t^b the *partial sum* (or aggregate traffic) over the time bucket starting at (t – 1)b and containing b samples, namely, Y_t^b t= $\sum_{i=1}^{b} X_{(t-1)b+i}$. The crucial assumption we make is that is an i.i.d. sequence for some appropriate bucket size. This is a reasonable assumption in many settings as temporal correlations tend to become weaker over longer time intervals. We quantize the values of the partial sums Y_t^b mapping them to the finite set $\sum = \{ \alpha_1, \ldots, \alpha_N \}$ of cardinality N . For the rest of the paper, we will be referring to $\sum$ as the *underlying alphabet*. The quantization is done as follows: we let$[r_0, r_N]$ be the range of values Y_t^b =takes, divide it into N subintervals $[r_0, r_1], \ldots, [r_{N-1}, r_N]$of equal length, and map$[r_{i-1}, r_i]$ to α_i for i= 1,…, N. . To select the appropriate size of the alphabet we follow the approach of [10] and use the so called Akaike's Information Criterion (AIC) [14].

A. Measuring Large Deviations of the Empirical Measure

Combinatorial methods can be applied for the empirical measures Of $\sum$ -valued process. Let $Y_t^{b} = (Y_t^{b*}{}_{-w+1}, \ldots, Y_t^{b*})$ be the trace of the w most recent partial sums using a bucket size b. We assume that the elements of Y_t^b are i.i.d., following a Law $\mu \in M_1(\sum)$, where $M_1(\sum)$ denotes the space of all probability measures on the alphabet $\sum$. Let also, $\sum\mu$ denote the support of the law μ , i.e., $\sum\mu = \{ \alpha i : \mu(\alpha i) > 0 \}$.

Define the *type* (empirical measure) of Y_t^{b*} as

$$\varepsilon_{\omega}^{y_t^{b^*}}(\alpha i) = \sum_{j=1}^{\omega} 1\alpha_i \ (Y_t^{b*}{}_{-w+j)} \ , \ i = 1, \ldots. N,$$

Where is 1 αi the indicator function $Ytb^*{}_{-w+j}$ of being of type αi. Namely, $\varepsilon_{\omega}^{y_t^{b^*}}$ αi is the fraction of occurrences of αi in the Sequence Y_t^{b*} . Let $\varepsilon_{\omega}^{y_t^{b^*}} = (\varepsilon_{\omega}^{y_t^{b^*}}(\alpha 1), \ldots, \varepsilon_{\omega}^{y_t^{b^*}}(\alpha N))$. The next theorem, which is due to Sanov, establishes a large deviations result for $\varepsilon_{\omega}^{y_t^{b^*}}$.

Theorem II.1: For every let $v \in M1(\sum)$ let

$I_1(v) = H(v|\mu)$ Where $H(v|\mu)$ is the relative entropy of the probability vector v with respect to μ.

$$H(v|\mu) \sum_{i=1}^{-N} v(\alpha i) \log \frac{v(\alpha i)}{\mu(\alpha i)}$$

More intuitively, Theorem II.1 states that for a long trace (i.e., large) its empirical measure is "close to" with probability. We will be referring to exponents such as the *exponential decay rate* of the corresponding probability.

B. Anomaly Detection

Theorem II.1 can be used to identify anomalies. Specifically:

1) From an anomaly-free trace construct the alphabet $\sum = \{ \alpha 1, \ldots, \alpha N\}$ and the empirical measure (law) μ induced by this sequence.

2) For each time t let $Ytb^{=}(Y_t^{b*}{}_{-w+1}, \ldots, Y_t^{b*})$ be the trace ω of the most recent partial sums using a bucket size b .Compute its empirical measure and let be the result. Based on Theom. II.1, $\rho\ t,w\ e^{-wI1(vt,w)}$ approximates the probability that the trace Y_t^b is drawn from the probability law μ. Thus, if is consistently low over some observed time interval, we can conclude that the observed trace deviates from the anomaly-free trace, which indicates an anomaly.

C. A Formal Anomaly Detection Test

Theorem II.1 rigorously identifies a distance metric—the exponent $I_1(v)$—between the two measures μ and *v* , constructed
as specified in Steps (1) and (2) of Section II-B. The key question we wish to answer is whether Y_t^b is generated from μ or
from some other law. This is known as a *composite hypothesis*
problem as one hypothesis (no anomaly) has a known law μ
while the alternative hypothesis is characterized by a family of
laws (all laws other than μ). Hoeffding ([8]; see also [7]) has
suggested an optimality criterion for these problems and a rule
that is optimal. It is also well known, that the empirical measure
$V_{t,w}$ of Y_t^{b*} of is a sufficient statistic.

4. MODEL-BASED APPROACH: Markov Modulated Process

One potential disadvantage of this aggregation is that it increases the response time to an anomaly since data is being processed on the slower time-scale of time buckets. In this section, the question we are seeking to answer is whether it is possible to process data on the timescale we collect them. To that end, and because the i.i.d. assumption will no longer hold, we will impose some more structure on

the stochastic nature of the traffic time-series. In particular, we will assume a Markovian structure as it is tractable and has been shown to represent traffic well [9], [10], at least for the purpose of estimating distribution-dependent metrics like loss probabilities.

A. An MMP Model

We start again with a time series X1,…., Xn of traffic activity

during a small time interval (several hours) which we will model as an MMP process. Such a process is characterized by an underlying Markov chain with transition probability matrix {P (i,j)} Mi,j =1 . To each state i, i=1,…, M. we associate an interval [ri-1, ri] of real numbers from which traffic activity observations are drawn. That is, when the MMP is in state at time t then Xt takes values in [ri-1, ri] . (For the application we are considering we do not need to specify how observations are drawn from; in general they can follow some probability distribution.) MMPs, when the state is "hidden", are also known in the literature as hidden Markov models (HMMs) [10].

We restrict ourselves to models in which the ranges of possible observations corresponding to different states are disjoint. Thus, an observation can be uniquely associated to an MMP state and the state is no longer hidden. To model the traffic trace as a MMP [r0,rM] we let be the range of all observations we make, split[r0,rM] into M subintervals of equal length, and assign state i ,i= 1,… M,, to interval [ri-1, ri] . To select the appropriate number of states M we use the AIC as in Section II. Given, the transition probabilities are obtained via maximum likelihood estimation. Specifically, let Y denote a sequence of Y1, Y2,..,Y_n of states that the Markov chain visits. A maximum likelihood estimator of the transition probabilities is given by

$$Pn(I,j) = qnf(j|i)^{\frac{qn(i,j)}{qn1(i)}} \quad I,j=1\ 1,\ldots., M.$$

where qn (I,j) denotes the fraction of transitions from i to j in the sequence Y and qn_1 (i) the fraction of transitions out of i. We assume that is large enough to have for all. As, with probability one (w.p.1).

5. Model Based Method: GP Distribution Superstatistics

a. Statistical Model of Network traffic

As is known, normal network traffic data possess stable statistical properties such as stationary mean and variance over a period of time. However, when an attack occurs, these statistics will change, and hence, they can be used to detect network abnormal. A good traffic model should be accurate enough to capture the statistical characteristics of actual traffic, and at the same time should be computationally efficient.

The GP distribution introduced by Pickands (1975) [8], is widely used for modeling extreme values in hydrology [9]. Let X be a GP distribution random variable; then the cumulative distribution function (CDF) of X is

$$F(x;b,k) = 1- (1+k \frac{x}{b})^{-1/k}$$

Where b is a positive scale parameter and k is a shape parameter. The range of X is $0 \leq x < \infty$ for $k < 0$ and $0 \leq x \leq b/k$ for $k > 0$. It is readily seen that when $k > 0$, the sample space of X is a finite interval with b/k as its upper bound. It is readily seen that when $k > 0$, the sample space of X is a finite interval with b/k as its upper bound. In this case, the GP distribution is short-tailed. On the other hand, when $k < 0$, the GP distribution is sometimes simply called Pareto.

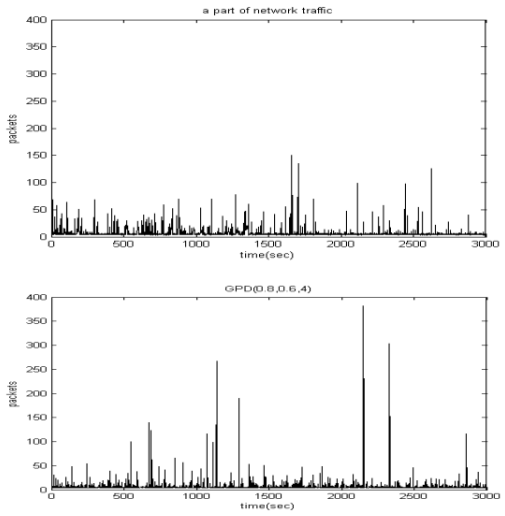

Fig 3. Original traffic and GP distribution (0.8, 0.6,4)

In Fig.1 the upper traffic is original traffic of DARPA 1999 and the lower traffic is the sample series generated by GP distribution (0.8, 0.6, 4). We find that the original traffic is similar to the GP distribution series with almost the same mean, variance and other statistic property.

B. Parameter Estimation of Distribution

Once a distribution function is assumed or selected for study at hand, it remains to estimate its parameters. The methods of maximum likelihood (ML)[9], of moments (MM) and of probability weighted moments (PWM)[10], are some of the main methods used to fit the GP distribution model. A newer method proposed and analyzed for the GP distribution model by Rasmussen (2001) is the method of generalized probability weighted moments (GPWM) [8].

The maximum likelihood estimators (MLE) may be numerically intractable. Algorithms for computing the MLE are given by Davison and Smith (1990) and Grimshaw (1993). Smith (1987) has shown that estimating GP distribution parameters with MLE is a non-regular problem for

k ≥0.5. The probability weighted moments (PWM) are easily computed and more efficient in general compared to MLE. Kaplan-Meier estimate of the cumulative distribution function (CDF), also known as the empirical CDF. Then the scale parameter b and the shape parameter k can be calculated.

c. Partition Algorithm

In this paper we treat the network traffic time series as a superposition of different segments which can be modeled by discrete GP distribution. To analyze the slow parameter we propose an algorithm to partition the traffic series into small segments. The sliding window has an initial size of 50 according to the requirement of the GPWM method and the size of window is increased by 1 after estimating the parameter without moving the start Position.

The BG method begins with the partition in following steps. We move a sliding pointer from the left to the right along the signal. At each position of the pointer, we compute the mean of the subset of the signal to the left of the pointer m1 and to the right m2. To measure the significance of the difference between m1 and m2, the statistic t=|(m1-m2)/sd is computed, s1 and s2 are the standard deviations of the data to the left and to the right of the pointer, respectively, and N1 and N2 are the number of points to the left and to the right of the pointer.

d. Testing of Generalized Pareto Distribution

When the parameters of the model are estimated, it is then desirable to access how well the distribution fits the observed data. Goodness of fit test is often essential to reveal departures from the assumed model. In part 2.3 parameters are estimated by modified GPWM method and in the part the Kolmogorov-Smirnov (K-S) test for GP distribution is used. And we also have adopted the probability chart testing method of Pareto distribution plans. Fig. 4 is a chart of test results. We found that all points basically in a straight line fitting, in line with that of the GP distribution distribution test.

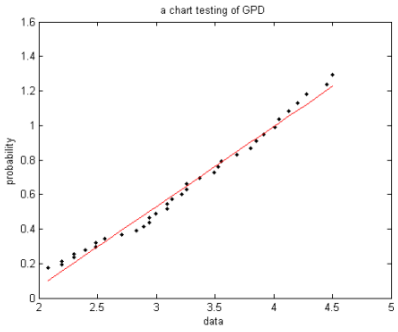

Fig 4. Probability chart testing

e. Modeling superstistics theory

According to the abnormal network traffic, particularly because of attacks caused by the abnormal flow of non-stationary and the basic characteristics of the sudden - abnormal flow is a complex non-linear or random change in the process of application of superstatistics, traffic monitoring statistical parameters of the statistical series features real-time network traffic anomaly detection. The GP distribution is parameterized with a scale parameter sigma, and a shape parameter k. k is also known as the "tail index" parameter, and determines the rate at which the distribution falls off. So shape parameter is the slow parameter compare with the fast change parameter according to the superstatistics theory.

Several basic theoretical algorithm used in our method to model network traffic flow have been discussed in previous sections. The idea that views the time series of traffic flows as a non-stationary superposition of segments obeying discrete GP distribution associated with superstatistics theory provides us a novel method to partition the non-stationary time series into stationary segments which can be modeled by discrete GP

Distribution in certain time scales. By implementing the partition and parameter estimate algorithm mentioned.

The AR model has already been widely applied to analysis and forecast of the time series. For model selection, order selection and parameter fitting, there is already a set of complete method. The most commonly used time series model is the Auto Regression model (AR), the Moving Average model (MA) and the Auto Regression Integrated with Moving Averages (ARMA). The autocorrelation function of Time Series (ACF) and partial autocorrelation function (PACF) are usually used to determine the type of models. The ACF and PACF charts of the slow variable series show that the two functions have a character of obvious tails, but neither is truncated. Therefore, the slow variable series should choose the ARMA model.

The concrete practice of the Generalized Maximum Likelihood Ratio (GMLR) is such that the two contiguous time windows of R(t) and S(t) in the test sequence are considered first. During the Real-Time Detection Process, both of them move ahead step by step, so they are called Sliding Windows. Using the Likelihood Ratio Test method, abnormal changes between R (t) and S (t) can be tested.

6. Conclusion

Many studies show the network flow presents a different character in a different time scale. The network traffic flow in second scale is studied in this paper and the network flow shows abruptness in the local segment. But with the increase of the time scale, abruptness will be decreased. Therefore, a further study should be continued using other distribution models. The parameter series is chosen to research on the change of the network flow in the

method and a good effect has achieved. Obviously the abnormal changes of the network flow can be found by the parameter series.

We introduced a general distributional fault detection scheme able to identify a large spectrum of temporal anomalies from attacks and intrusions to various volume anomalies and problems in network resource availability. Our proposed anomaly detection frameworks are able to identify temporal or spatial anomalies [12], we are able to identify both as we preserve both the temporal and spatial correlation of network feature samples.

We provided different approaches, a model-free and a model-based one. The model-free method works on a longer time-scale processing traces of traffic aggregates over a small time interval. Using an anomaly-free trace it derives an associated probability law. Then it processes current traffic and quantifies whether it conforms to this probability law. The model-based method constructs a Markov modulated model of anomaly-free traffic measurements and relies on large deviations asymptotics and decision theory results to compare this model to ongoing traffic activity.

According to the characters of the network flow, a network traffic model based on superstatistics and Markov modulated process is developed, which describes the characters of the actual network flow quite well such as: non-stationary, heavy-tailed property, LRD and abruptness. Using the superstatistics theory to analyze the model, the parameter series which reflect the abnormal changes of the network flow are studied in this method, and the method achieves the goal of analyzing the whole model system at last. In addition, the number of the parameter series is far lower than the original network flow, so it can increase the calculation speed, and reduce the computational complexity to a certain extent. Obviously the abnormal changes of the network flow can be found by the parameter series. As a whole, it is more visual than ever before. This method has obtained a very good effect through a lot of experiments. Our method is of low implementation complexity (only an additional counter is required), and is based on first principles, so it would be interesting to investigate how it can be embedded on routers or other network devices.

References

[1] R. A. Kemmerer and G. Vigna, Intrusion detection: A brief History and review, Computer, vol. 35, no. 4, pp. 27–30, Apr. 2002.

[2] B. Mukherjee, L. T. Heberlein, and K. N. Levitt, Network Intrusion detection, IEEE Netw., vol. 8, no. 3, pp. 26–41, May/Jun. 1994.

[3] Di He Leung, H. Network Intrusion Detection Using CFAR Abrupt-Change Detectors, Instrumentation and Measurement, IEEE Transactions on, Volume: 57, pp: 490-497, Mar 2008

[4] R. Lippmann, D. Fried, I. Graf, J. Haines, K. Kendall, D. McClung, D.Weber, S. Webster, D. Wyschogrod, R. Cunningham, and M. Zissman,"Evaluating intrusion detection systems: The 1998 DARPA off-line intrusion detection evaluation," in Proc. DARPA Information Survivability Conf. and Expo., Los Alamitos, CA, Jan. 2000, pp. 12–26.

[5] R. Lippmann, J. W. Haines, D. J. Fried, J. Korba, and K. Das, "The1999 DARPA off-line intrusion detection evaluation," Computer Net-works, vol. 34, no. 4, pp. 579–595, 2000.

[6] V. Yegneswaran, J. T. Giffin, P. Barford, and S. Jha, "An architecture for generating semantics-aware signatures," in USENIX Security Symp., Baltimore, MD, Jul. 2005, pp. 97–112.

[7] A. Dembo and O. Zeitouni, Large Deviations Techniques and Applications, 2nd ed. New York: Springer-Verlag, 1998.

[8] W. Hoeffding, "Asymptotically optimal tests for multinomial distributions," Ann. Math. Statist., vol. 36, pp. 369–401, 1965.

[9] I. Paschalidis and S. Vassilaras, "On the estimation of buffer overflow probabilities from measurements," IEEE Trans. Inf. Theory, vol. 47, no. 1, pp. 178–191, 2001.

[10] Rasmussen, P., Ashkar, F., Rosbjerg, D., Bobe´e, B., 1994. The POT method for flood estimation: A review. In: Hipel, K.W. (Ed.), Stochastic and Statistical Methods in Hydrology and Environmental Engineering. Kluwer Academic Publishers, pp. 15–26.. pp. 71–86, Feb. 1997.

[11] Spatio-Temporal Network Anomaly Detection by Assessing Deviations of Empirical Measures Ioannis Ch. Paschalidis, and Georgios Smaragdakis IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 17, NO. 3, JUNE 2009.

Syed. Azahad currently pursuing M.tech in computer science at QIS College of Engineering and Technology which is affiliated under JNTU, kakinda. He received the M.Sc in (IT) from Vinayaka Missions university. He received his O, A, B levels from DOEACC.

R. Lakshmi Tulasi currently working as HOD (Head of the Dept.) in QIS College of Engineering and Technology which is affiliated under JNTU, Kakinada which is permanent NBA accredited Institute. She has 12 years of experience. She has published 10 papers in national and international journals has published 8 papers in the conferences.