RELEVANCE OF TIME SERIES MODELS IN NETWORK TRAFFIC MODELING: IN THE CONTEXT OF TCP SYN DOS ATTACK

A THESIS

submitted by

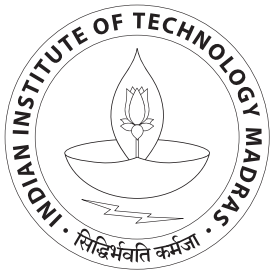
CYRIAC JAMES

for the award of the degree

of

MASTER OF SCIENCE

(by Research)



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY MADRAS.

November 2012

THESIS CERTIFICATE

This is to certify that the thesis titled **Relevance of Time Series Models in Network Traffic Modeling: In the Context of TCP SYN DoS Attack**, submitted by **Cyriac James** to the Indian Institute of Technology Madras, for the award of the degree of Master of Science (by research) is a bonafide record of research work carried out by him under my supervision and guidance. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Dr. Hema A. Murthy

Chennai - 600 036

Date:

ACKNOWLEDGEMENTS

I am deeply indebted to my advisor Dr. Hema for her guidance and support at all stages of my research work. Her energy and enthusiasm has been a constant source of inspiration to me. But for the confidence she has shown in me and her unlimited patience, this work would never have completed.

I thank the members of my Graduate Technical Committee, Dr. Krishna Sivalingam, Dr. Nitin Chandrachoodan and Dr. Kamala Krithivasan for their constructive feedback during my interactions with them.

My thanks to all members of the Computer Science department for providing a stimulating atmosphere for conducting research. I also thank the entire IIT Madras system for the rewarding influence it had on me.

I am thankful to all members of Donlab for all the great times we had, inside and outside the lab. The network and speech groups have provided a good environment for informal discussions. I thank Dr. C. Chandra Sekhar and other members of these groups for their patient listening and feedback on my work.

There are many colleagues and friends who have helped me in so many ways during my life as a graduate student. I refrain from naming each one, lest I omit somebody.

Last, but not the least, none of these would have happened without the continuous support from my family. I am grateful to the understanding shown by my parents and the constant motivation provided by my wife Anila.

ABSTRACT

Keywords: *Network traffic characterisation; Stochastic modeling; Time series models; Monitoring and forecasting; Anomaly detection*

Network traffic characterisation and modeling using time series models is an area which has been extensively studied in the past. *Coarse-grained* (aggregated traffic) time series analysis using parametric approach, primarily carried out at the backbone network over a long time period (of the order of days to months), show strong deterministic cyclic trends, while the *fine-grained* (at the packet or flow level) counterpart, done mostly at edge network over small time period (of the order of few minutes), exhibit *self-similar* behaviour. This thesis is an attempt to study the *fine-grained* time series characteristics of network traffic at an edge network, observed over a long period (of the order of days and weeks), using parametric approach. The analysis is carried out in the context of anomaly detection. Most of the earlier attempts in this direction followed a non-parametric approach, by either using adaptive mechanism (model adapted to evolving time series trend) or non-adaptive mechanism (assumes stationarity, usually after normalising the series), whose performance is found to be extremely *sensitive* towards empirically determined parameters of the model and hence difficult to determine. Also, the model parameters need to be recomputed at regular intervals of time (of the order of few seconds to minutes). To some extent, this make such algorithms less attractive in terms of generality and practical implementation.

The thesis initially discusses the statistical characteristics of such *long range* net-

work time series. These are found to exhibit *structural breaks* apart from *transient shocks*. It is observed that the stationary component in the series can be decoupled from the non-stationary one by an *absolute first difference* transformation and the former can be approximated by an AR model. The stationarity and stability are understood by studying the higher order statistics. In the later part, the analysis is further motivated by applying the model for detecting TCP SYN flood Denial of Service (DoS) attacks. The efficacy of the model proposed is evaluated, by conducting extensive trace driven simulations for the detection of *low intensity* TCP SYN attacks. The feature to be modeled for detection of SYN attacks is chosen based on two criterion; Firstly, it has to exhibit different properties during normal and attack period. Secondly, it should be as independent as possible with respect to the access pattern. The difference between incoming SYN packets and outgoing SYN/ACK packets (called *Half-Open Count*) is found to be a good candidate that satisfy the above two conditions. The solution being proposed for the edge router at the victim side is scalable to Distributed DoS (DDoS) attacks as well.

Performance is measured in terms of false positives, false alarm time, detection rate and detection delay. Experiments are performed on actual traffic traces collected from one of the edge networks of IIT Madras over a period of three months and for various sampling intervals (10s, 60s, 120s). Comparative studies with adaptive and non-adaptive methods are carried out to demonstrate the relevance of the proposed model. It is observed that the proposed method gives better performance with 100% detection accuracy for false positive as low as 0.9%. Further, it is proposed that the stability of the system during normal and attack period can be used as a measure for detection of anomalies in the network traffic. This is apart from the traditional method (also used in this work) of applying threshold on the forecasted error.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

HOC	Half-Open Count
RTT	Round Trip Time
CUSUM	Cumulative Sum
AR	Auto-Regressive
MA	Moving Average
ARMA	Auto-Regressive Moving Average
ARCH	Auto-Regressive Conditional Heteroskedastic
FARIMA	Fractional ARIMA
FBM	Fractional Brownian Motion
ACF	Auto-Correlation Function
PACF	Partial ACF
SRD	Short Range Dependent
LRD	Long Range Dependent
FP	False Positive
FAT	False Alarm Time
DR	Detection Rate
DD	Detection Delay

CHAPTER 1

Introduction

Understanding network traffic behaviour for different network sizes and at different scales is an area which is actively being researched. In this context, modern time series analysis found applications in network traffic modeling and characterisation. Broadly, there are two types of approaches followed in time series modeling; parametric and non-parametric. Parametric approach is used when a model is valid over the entire input space. In other words, the model coefficients are said to parametrise the distribution from which the input time series is drawn. This is possible when the series is either deterministic or stationary[1]. The model coefficients are estimated from the training data (a subset of input space) to build a single global model. If the above conditions are not satisfied, then local models are computed using training samples which are close by in the series. The model coefficients are empirically determined at regular intervals of the order of few seconds to minutes to counter the *non-deterministic and non-stationary* trend in the traffic or made constant by assuming *stationary* traffic behaviour. They are referred to as adaptive (also called quasi-stationary) and non-adaptive mechanisms respectively later in the thesis. In either case, model coefficients are not estimated from the entire training space, unlike in the case of parametric modeling.

[1]The first and second order moments are finite and time invariant, also called weak stationarity

1.1 Time Series Modeling of Network Traffic

In backbone networks, parametric time series models are used at a *coarse-grained* level (aggregated traffic data, for eg. link utilisation), to understand the long term trend in the traffic, which report strong deterministic daily and weekly cyclic trends [1, 2]. Also, it is not feasible to collect the Internet traffic at a finer level (packet or flow) due to link speed (OC-48 and higher) and huge amount of data to be analysed in real-time. Hence, *fine-grained* analysis at the backbone network is done at a limited scale (in terms of time and the number of links observed) from which generalisations are made [3–5]. Unlike in backbone network, *fine-grained* time series analysis (at the level of packet or flow) are extensively carried out at the edge network [4–13], which is found suitable for application like anomaly detection. Some of these studies proposed parametric models for traffic traces collected over small duration (of the order of few minutes to hours). The work in [6] modeled 1 hour traffic trace collected at different times of the day, month and year. Work proposed in [10], carried out study on simulated traffic and hasn't investigated the stationarity of the traffic. Hence, practicality of this approach cannot be established. Most of the studies reported in literature use non-parametric time series methods [5, 7–9, 11–13] to model *fine-grained* network time series extending to days and weeks. This approach is generally suitable when the time series does not show any structure, in which case parametric methods are likely to fail. More details are discussed in chapter 2. To the best of our knowledge, there are very few studies on the long term characteristics of network time series collected at the *fine-grained* level, using parametric method. Some of the advantages of parametric approach over quasi-stationary and non-parametric ones, provided one can find structure in such *long range* network time series, are; Firstly, modeling is simple if the series follows a known distribution or model. Secondly, there will be only limited number of

model parameters, which can be estimated straightforward using standard parameter estimation techniques, rather than empirically determining (performance *sensitive* to these values) it as in the case of non-parametric ones. Thirdly, the re-computation of model parameters at regular intervals of time can be avoided, making it attractive for online Intrusion Detection System (IDS) with respect to lesser computational complexity.

Hence, an attempt is made in this direction, by systematically studying the traffic characteristics at an edge router using network traffic collected continuously over 3 months. The analysis is done in the context of detecting *low intensity*[2] TCP SYN DoS attacks [3, 11, 12, 14, 15]. Studying such *fine-grained* time series becomes all the more important in the context of such attacks, that are shown to affect the dynamics of the system at fine scales [16]. These dynamics get averaged out during aggregation in the case of *coarse-grained* feature(s) and can be studied only by a *fine-grained* analysis. Such finer level dynamics, which are generally difficult to model, is studied extensively in electrical systems for understanding the oscillatory and stability nature of the system [17]. The importance of modeling short-term or transient behaviour in the context of *low intensity* DoS attacks have been discussed in [16] and [18] from a control theoretic approach. Hence, seasonality (or diurnal pattern) if any in the series is not studied in this work. The general usefulness of the modeling technique and the detection mechanism proposed, are also discussed. The thesis also briefly discusses the two important phenomena in network traffic, namely, *Self-Similarity* and *Long Range Dependence (LRD)* [6, 19], though a detailed discussion is beyond the scope of this thesis and is left as future work.

Despite numerous studies over the past two decades, network anomaly detection is

[2]Attacks with rate comparable to normal traffic

still an immature area and continues to draw attention of researchers across the world. The existing IDS techniques can be broadly classified into two [20, 21]: (a) Signature based and (b) Anomaly based mechanism. In signature based detection mechanisms [22–28], various events in the network are kept track of and matched with the attack signatures. The main advantage of this approach is that it is straightforward to implement. But, the approach suffers from following disadvantages. Firstly, exhaustive list of all possible normal sequence of events have to be defined while building the state machines. Secondly, the method is stateful, which means keeping track of past, present and future states of the system, hence not the best approach in terms of space utilisation. Thirdly, the approach is not scalable with increasing network/traffic size. Fourthly, it can detect only known attacks. Anomaly based detection techniques, which is also the focus of this thesis, build models for the normal traffic, and if the incoming traffic does not conform to the model, it is marked as anomalous. Though this approach can detect future attacks, major challenges are to reduce false positives (particularly for *low intensity* attacks) and obtain good detection rate.

1.2 TCP SYN Attack

TCP SYN attack is a protocol based attack. It exploits the vulnerabilities in the TCP protocol. TCP connection establishment consist of a three way handshake. The client which is in need of a service, actively initiates a connection by sending a SYN packet to the server (entity which provides the service). Server which is in the LISTEN state initially, on receiving the SYN packet, updates its state to SYN RECEIVED. After storing information about the connection request in a backlog queue (of limited size), server replies with an SYN-ACK packet. Client on receiving this, acknowledges with

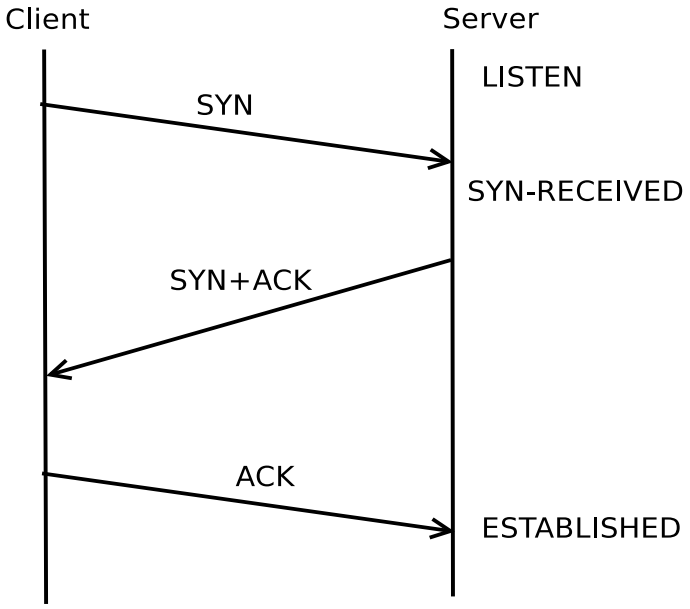


Fig. 1.1: Three way Handshake

an ACK packet. Once the ACK packet is received at the server, TCP connection state becomes ESTABLISHED and the corresponding entry for the connection is removed from the backlog queue. So, the backlog queue essentially consist of connection requests which the server has accepted but not established. Such connections are also called *Half-Open Connections* and the number of such outstanding connection requests at any point of time is called *Half-Open Count*. If the queue is full, subsequent incoming requests are rejected by the server. Figure 1.1 shows this three way hand-shake and the sever state transitions.

During an attack, the attacker or a group of attackers, generate connection requests to the server with *spoofed* IP addresses (non-existent). Server acknowledges these requests by sending SYN-ACK packets to destinations which are not reachable. TCP protocol defines a retransmission timer which is computed based on Round Trip Time (RTT). RTT for a TCP connection is the time taken by a TCP segment to reach the receiver and for the acknowledgement from the receiver to reach the sender. As

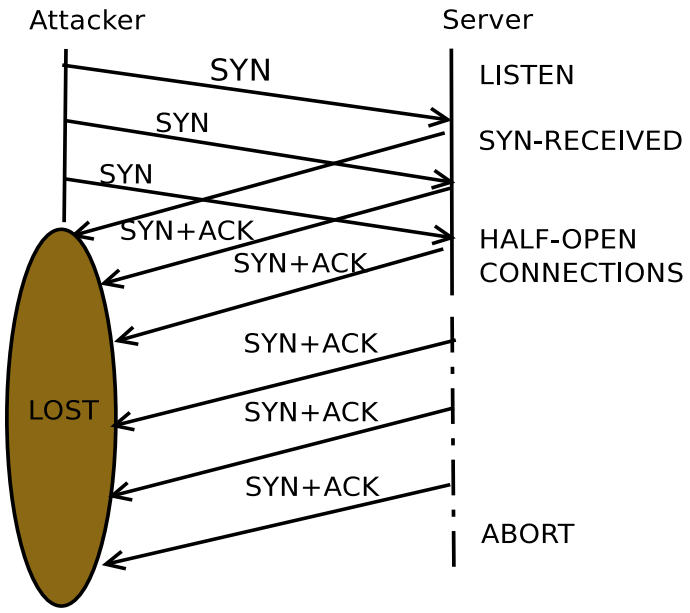


Fig. 1.2: SYN Attack

per the RFC [29], the initial value of the Retransmission Time Out (RTO) is set to $3s$ at both the ends until RTT is measured. Hence, the server retransmits the SYN-ACK packet after RTO of $3s$. Due to the exponential back-off property of TCP, after each time-out, timer is doubled. Also, the minimum number of re-transmissions by default is 5 in numbers. Hence, the value of the timer varies as 3s, 6s, 12s, 24s and 48s, after which connection attempt is actively terminated from the server side. So, the information about any connection request is kept in the backlog queue for about 93s. When the queue is filled with attack requests, it denies service to legitimate requests. An attacker can design attacks which are *low intensity* in nature from different geographical locations, to partially or completely fill the backlog queue, thereby adversely influencing the service. A typical attack scenario is shown in Figure 1.2.

The earliest solutions to detection of TCP SYN DoS attacks are *signature* based approaches. This is possible since TCP is a stateful protocol. Syncookies [26], Syn-Defender [27], and Synkill [28] are to mention a few. For reasons discussed in Section

1.1, we will study *anomaly based* detection techniques in this work.

1.3 Objective

The objective of this work is twofold: (a) Understand and model the long-term *fine-grained* traffic characteristics of network time series at the edge network for different time scales and (b) Application of the model developed in *low intensity* DoS attack detection (TCP SYN attack in specific), by studying the stability and stationarity of the underlying process.

1.4 Major Contributions of the Thesis

The following are the major contributions of this research thesis:

- It is observed that network time series over a long period at the edge network exhibit structural breaks apart from transient ones. These structural breaks causes non-stationarity (non-deterministic as well) in the data, which is generally difficult to model.
- It is shown that the stationary component can be decoupled from the non-stationary one by an *absolute difference* operation and the former can be approximated by an Auto-Regressive (AR) Model.
- The model proposed is tested for its adequacy and its efficacy in detecting *low intensity* TCP SYN DoS attack is extensively studied with trace driven simulation. Comparative analysis with an adaptive [5] and non-adaptive [7] method is conducted to demonstrate the relevance of the proposed model.
- Stability of the system during normal and abnormal period is studied by es-

timating the poles of the transfer function in the Z-plane. A new measure is proposed for detecting anomalous behaviour in Internet traffic like TCP SYN attacks, by tracking the stability of the system at regular intervals of time. This is apart from the traditional method (also used in this work) of applying threshold on the forecasted error.

1.5 Organization of the Thesis

The rest of the thesis is organized as follows. Chapter 2 gives a background on the related literature and time series techniques applied in this research work. *Long range* time series characteristics of *HOC* is discussed in Chapter 3. Stability and stationarity of the underlying data generating process, model selection, prediction, validation and detection of *low intensity* TCP SYN attack are studied in Chapter 4. Finally the thesis is concluded in Chapter 5, with discussion on the drawbacks and future directions of this work.

CHAPTER 2

Background and Related Work

In the last chapter, we have discussed the topic that is the focus of this thesis and have given a broad outline on the background literature. A more detailed discussion on the relevant related work and time series models used in our study are explained in this chapter.

2.1 Related Work

Long term time series characteristics and forecasting of network traffic have been studied earlier with application on capacity planning and traffic engineering. Hence, these analyses are done at larger time scales and on *coarse-grained* features. In the work reported in [2] and [1], authors have studied the long term time series trend in a backbone network at a Tier-I ISP (Internet Service Provider), by analysing the aggregate *traffic rate* sampled at every 5 and 90 minutes, respectively. These studies demonstrated the strong daily and weekly deterministic cyclic trends, which can be modeled using non-stationary models. Short term events like flooding attacks (eg: TCP SYN attack) and re-routing of traffic due to link failures, can cause outliers in the aggregate traffic. Since it is not feasible to do a *fine-grained* analysis because of reasons explained in Section 1.1, exact cause of these sudden peaks in traffic rate cannot be identified. Hence, such points in the time series are discarded as outliers without much reasoning [1, 2]. Moreover, flooding attacks may get buried in the huge background traffic at a back-

bone network [1,3]. This resulted in developing computationally intensive algorithms (in terms of space and time) to detect such attacks, by investigating *fine-grained* and multivariate time series, in the backbone network [3, 30, 31], employing dimensionality reduction techniques like Principal Component Analysis (PCA) and Sketch data structure. It is observed in [32] that PCA techniques are hard to implement and is near to impractical in real network scenarios. The algorithm in [3] though reduces the space complexity to some extent by randomly aggregating the flows, its detection accuracy is heavily dependent on the size of the data structure. Also, re-computation of Least Mean Square (LMS) filter coefficients and chi-square divergence tests are done at every 1 minute, resulting in greater computational cost. Even from an economic point of view, proposals for such large investments in the network architecture have always faced stiff resistance from ISPs [33]. We believe that, even though changes in Internet are inevitable, rather than adopting such 'clean-slate' approach (as known in literature), it is better to follow a middle path between traditional design (complex end points and simple backbone network) [34] and 'clean-slate' design. For example, the *low intensity* attack scenario discussed in [3], can be detected with much more simpler methods at the edge network, where *fine-grained* analysis is more practicable.

The non-parametric methods (commonly followed for *fine-grained* analysis at the edge network) either used adaptive or non-adaptive (i.e assuming stationarity) mechanisms. The approach discussed in [7,11] identify the change points where the time series become statistically non-homogeneous, by conducting a sequential test on it. This is achieved by applying non-parametric Cumulative Sum (CUSUM) algorithm after normalising the time series. The predefined threshold used in these algorithms make it vulnerable to higher false positives due to inherent non-stationarity in the traffic. Adaptive methods proposed in [5,8,9,11–13], employ variants of moving average tech-

nique to compute the local model. The major disadvantage of these non-parametric methods are that the performance of these methods are extremely *sensitive* towards empirically determined model parameters and hence difficult to determine. This fact has been acknowledged in all these work with detailed discussion in [11, 13]. To some extent, this makes such algorithms less attractive in terms of generality (towards different types of attack and network scenarios) and practical implementation. A parametric time series model for prediction on *fine-grained* series over long time period is reported in [4]. But, the work discussed here goes way beyond it in proposing a mathematical model for such network time series, investigating the stationarity of the series, applying the technique in the context of anomaly detection and by doing a comparative analysis to understand the efficiency of the proposed method.

A method to decouple the stationary and non-stationary (non-deterministic as well) components in the *long range* network time series is proposed in this work. The stationary component is then approximated by an AR model, which is shown to be efficient (low false positives and high detection rate) in detecting *low intensity* TCP SYN flood DoS attacks. The results are compared with two non-parametric approaches [5, 7].

2.2 Time Series Models

Time series techniques discussed in this thesis are based on the studies reported by G.U.Yule in 1927 [35] and later by H.Wold [36] in 1938, where they proposed *linear regression* models for stationary processes which has time series with correlation between adjacent values. These time series values are obtained by sampling or polling the feature under study at regular intervals and hence can also be called *discrete* time

series. By analysing the observation values, underlying data generating process can be studied and modeled.

Stationary time series models are developed under the assumption that the series of random variables are identically distributed (also called *strong stationarity*). This being impractical,only the first and second order moments (i.e expected mean and covariance) are considered constant (called *weak stationarity*). Since we have only one realisation of the underlying stochastic process, in order to empirically prove that the random process is stationary, it is sufficient to show (by the assumption of ergodicity) that sample mean and covariance over sample values at arbitrary equidistant set of time points $T = (t_1, t_2, ..., t_n)$ is same as at time points $T' = (t_1 + t, t_2 + t, ..., t_n + t)$, where T' is some arbitrary translation in time.

According to the linear filter model proposed in [35, 36], a weakly stationary process with successive values highly dependent can be represented as a combination of independent and identically distributed random variables (pure random process),

$$x_t = \sum_{i=0}^{\infty} \psi_i a_{t-i} \tag{2.1}$$

where x_t is a mean differenced series, ψ_i being model coefficients with $\psi_0 = 1$ and a_t is a pure random process with,

$$E[a_t] = 0 \tag{2.2}$$

$$E[a_t a_{t-k}] = \begin{cases} \sigma_a^2 & \text{if } k \text{ is } 0 \\ 0 & \text{Otherwise} \end{cases} \tag{2.3}$$

For all practical purposes, the distribution of a_t is assumed to be normal. The mean and variance of the above process can be estimated as follows:

Mean:

$$E[x_t] = E[\sum_{j=0}^{\infty} \psi_j a_{t-j}] = \sum_{j=0}^{\infty} \psi_j E[a_{t-j}] = 0 \tag{2.4}$$

Variance:

$$Var(x_t) = E[x_t x_t] = E[(\sum_{j=0}^{\infty} \psi_j a_{t-j})(\sum_{j=0}^{\infty} \psi_j a_{t-j})] \tag{2.5}$$

$$= \sigma^2 \sum_{j=0}^{\infty} \psi_j^2 = \gamma_0 \tag{2.6}$$

Hence, for the variance to be finite,

$$\sum_{i=0}^{\infty} \psi_i^2 < \infty \tag{2.7}$$

This condition is also called the stability of the process. Stability of a system can be visualised as an inverted cone balancing with its tip on the surface and is said to be in an equilibrium position. Due to unknown random disturbances acting on the system, it moves from its equilibrium position. But, as long as the variance of the response of the system (which are the observed values) is converging or finite, the system comes back to its original equilibrium position and hence said to be in stable state. Else, the cone will swing widely and topples i.e system becomes unstable and finally fails if no corrective mechanisms are taken.

Similarly, auto-covariance (covariance between random variables of the same stochastic process) for lag k can be represented as:

$$\gamma_k = E[x_t x_{t+k}] = E[(\sum_{j=0}^{\infty} \psi_j a_{t-j})(\sum_{j=0}^{\infty} \psi_j a_{t+k-j})] \tag{2.8}$$

$$= E((\psi_0 a_t + \psi_1 a_{t-1} + ...)((\psi_0 a_{t+k} + \psi_1 a_{t+k-1} + ...)] \tag{2.9}$$

$$= \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k} < \infty \tag{2.10}$$

Above equation can be written in terms of auto-correlation function as,

$$\rho_k = \frac{\gamma_k}{\gamma_0} \tag{2.11}$$

where $\rho_0 = 1$. For a stationary process, first and second order moments must be time invariant and finite (converging). Above model is useful in practice only when it is parsimonious (i.e number of model coefficients are limited). Models derived from the linear filter model, which found wide application in forecasting of economic time series are discussed in the subsequent sections.

2.3 Definitions

2.3.1 Auto-Regressive (AR) Model

In AR model [37–40], value of the series at time t is written as a linear combination of past p values and a random shock at time t, where p is the order of the model.

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + ... + \alpha_p x_{t-p} + a_t \tag{2.12}$$

where x_t is the mean subtracted time series, α_i(i = 1,2,....p) are the model coefficients and a_t is the random shock or residual error. Residual values are pure random process (white noise process). For simplicity, in most cases, it is assumed to be normally distributed with mean zero and variance σ_a^2. A stable AR process can be written as an infinite but converging series in present and past random shocks (for parsimonious representation, it is written otherwise),

$$x_t = \sum_{j=0}^{\infty} \psi_j a_{t-j} \tag{2.13}$$

where

$$\sum_{i=0}^{\infty} \psi_i^2 < \infty \tag{2.14}$$

From the above equation, auto-covariance at lag k can be derived as,

$$\gamma_k = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k} < \infty \tag{2.15}$$

where σ_a^2 is the variance of a_t. Then, auto-correlation at lag k can be written as,

$$\rho_k = \frac{\sigma_a^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k}}{\gamma_0} < \infty \tag{2.16}$$

So, the auto-correlation values decays fast, but extent *infinitely* (theoretically). Hence, an AR system is also called an Infinite Impulse Response (IIR) system. Stability of the AR process is discussed in detail in Section 2.5. The coefficients of an AR model can be estimated using *Yule-Walker* method [37,41]. For demonstration purpose, consider a second order AR model,

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + a_t \tag{2.17}$$

On multiplying with x_{t-k} and taking expectation on both sides, it becomes,

$$E(x_t x_{t-k}) = E(\alpha_1 x_{t-1} x_{t-k} + \alpha_2 x_{t-2} x_{t-k} + a_t x_{t-k}) \tag{2.18}$$

For values of $k > 0$, x_{t-k} and a_t are uncorrelated. Hence, above equation can be written as,

$$\gamma_k = \alpha_1 \gamma_{k-1} + \alpha_2 \gamma_{k-2} \tag{2.19}$$

On dividing by γ_0,

$$\rho_k = \alpha_1 \rho_{k-1} + \alpha_2 \rho_{k-2} \tag{2.20}$$

Above homogeneous equation is called *Yule-Walker* equation, which can be used to estimate the model coefficients by replacing ρ_k with estimated autocorrelation values.

2.3.2 Moving Average (MA) Model

MA models [37–40] are time series models in which value at time t of the series is represented as the weighted average of present and past q random shocks, where q is the order of the model.

$$x_t = a_t - \psi_1 a_{t-1} - ..... - \psi_q a_{t-q} \tag{2.21}$$

where x_t is the mean subtracted time series, ψ_i (i = 1, 2, ..., q) are the model coefficients and a_t is the pure random process as in the case of AR model. MA model in Equation 2.21 is nothing but *finite series* linear filter model discussed in Section 2.2 and hence are inherently stable. Such systems are also called finite impulse response systems since the auto-correlation values become theoretically zero after lag q. To demonstrate this, consider a MA(1) process,

$$x_t = a_t - \psi_1 a_{t-1} \tag{2.22}$$

On multiplying with x_{t-k} and taking expectation on both sides,

$$\gamma_k = E(a_t x_{t-k} - \psi_1 a_{t-1} x_{t-k}) \tag{2.23}$$

Since auto-covariance is symmetric, consider values of $k \geq 0$,

$$\gamma_0 = E(a_t x_t - \psi_1 a_{t-1} x_t) \tag{2.24}$$

$$\gamma_0 = E(a_t(a_t - \psi_1 a_{t-1}) - \psi_1 a_{t-1}(a_t - \psi_1 a_{t-1}) \tag{2.25}$$

$$\gamma_0 = \sigma_a^2 + (\psi_1)^2 \sigma_a^2 \tag{2.26}$$

where σ_a^2 is the variance of the random shock a_t with $E[a_t] = 0$. Auto-covariance for larger lags can be written as follows,

$$\gamma_1 = E(a_t x_{t-1} - \psi_1 a_{t-1} x_{t-1}) \tag{2.27}$$

$$\gamma_1 = E(a_t(a_{t-1} - \psi_1 a_{t-2}) - \psi_1 a_{t-1}(a_{t-1} - \psi_1 a_{t-2}) \tag{2.28}$$

$$\gamma_1 = -\psi_1 \sigma_a^2 \tag{2.29}$$

$$\gamma_2 = E(a_t x_{t-2} - \psi_1 a_{t-1} x_{t-2}) \tag{2.30}$$

$$\gamma_2 = E(a_t(a_{t-2} - \psi_1 a_{t-3}) - \psi_1 a_{t-1}(a_{t-2} - \psi_1 a_{t-3}) \tag{2.31}$$

$$\gamma_2 = 0 \tag{2.32}$$

$$\tag{2.33}$$

Similarly, $\gamma_3, \gamma_4, ...$ are all zero. Hence, the auto-correlation function becomes,

$$\rho_k = \begin{cases} \frac{-\psi_1}{1+\psi_1^2} & \text{for } k = 1 \\ 0 & \text{for } k > 1 \end{cases}$$

Similar but more complex equations can be derived for higher order MA model. The simple *Yule-Walker* approach used in the case of AR model is not available in the case of MA model for estimating the model coefficients. By substituting estimated values of autocorrelation in the above non-linear equation, model coefficients can be estimated. Unlike in the case of AR model, coefficients have multiple values. To choose unique values for the model parameters, *invertibility* condition is defined for the MA model. If the MA model is invertible, it can be represented as an infinite but converging series of past values of the time series and the present value of random shock.

$$x_t = \sum_{j=1}^{\infty} \beta_j x_{t-j} + a_t \tag{2.34}$$

Detailed explanation on invertibility is set aside for later discussion in Section 2.5.

2.3.3 Auto-Regressive Moving Average (ARMA) Model

ARMA model can be represented as a combination of an AR component, which is a linear representation of past values and an MA component, which is a linear representation of past and present random shocks. An ARMA(p,q) model can be represented as,

$$x_t = \phi_1 \, x_{t-1} + \, .. \, + \phi_p \, x_{t-p} - \psi_1 \, a_{t-1} - \, .. \, - \psi_q \, a_{t-q} + a_t \tag{2.35}$$

where p and q are coefficients of AR and MA components respectively. Let L be a lag operator which when operated on any variable at time t gives the value of the variable at time $t-1$. Applying this on Equation 2.35, we get,

$$(1 - \phi_1 L - \, .. \, - \phi_p L^p) x_t = (1 - \psi_1 L - \, .. \, - \psi_q L^q) a_t \tag{2.36}$$

For a process to be represented by an ARMA(p,q) model, its AR and MA components must be stable and invertible respectively. The model coefficients of an ARMA model can be estimated in two stages; AR stage and then MA or vice-versa. The time series to be modeled is fed as input for the first stage and the residual error from it is fed as the input for the second stage. The residual error of the second stage should ideally be a pure random process. This condition is used to check the model adequacy, later in the thesis.

2.4 Extensions of ARMA Model

2.4.1 Auto-Regressive Integrated Moving Average (ARIMA) Model

If the process is mean varying (i.e non-stationary), a difference operator ∇ is applied [37], which takes the difference of adjacent values to form a new time series. The differencing operation is repeated till the resultant series is stationary and the number of times it is done is called the degree of the differencing d ($d = 0,\ 1$ or 2 for most of the cases). For $d = 1$, the difference series y_t can be written as,

$$y_t = \nabla x_t = x_t - x_{t-1} \tag{2.37}$$

Applying the lag operator,

$$y_t = \nabla x_t = (1 - L)x_t \tag{2.38}$$

From Equations 2.36 and 2.38, a general form of such non-stationary model can be represented as,

$$(1 - \phi_1 L - \ .. \ - \phi_p L^p)\nabla^d x_t = (1 - \psi_1 L - \ .. \ - \psi_q L^q)a_t \tag{2.39}$$

Model in Equation 2.39 is called ARIMA(p,d,q) model, which is an extension of ARMA(p,q) model in the non-stationary space. It has to be noted that the difference operation works only when the process is mean varying, but variance and covariance stationary.

2.4.2 Auto-Regressive Conditional Heteroskedastic (ARCH) Model

ARMA model or ARIMA model in general, apart from assuming stationarity, also assumes conditional variance to be constant i.e residuals or random shocks of the process are homoskedastic. For a better understanding on this, consider an ARMA(1,1) model

$$x_t = \phi_1 x_{t-1} - \psi_1 a_{t-1} + a_t \tag{2.40}$$

The expected value of x_t, given past values (conditional mean) is

$$E(x_t|x_{t-1}, x_{t-2}...) = E_{t-1}(x_t) = \phi_1 x_{t-1} - \psi_1 a_{t-1} \tag{2.41}$$

This value is predicted by ARMA model after estimating the AR and MA coefficients. The expected variance is then given by,

$$V(x_t|x_{t-1}, x_{t-2}...) = E[(x_t - E_{t-1}(x_t))^2|x_{t-1}...] \tag{2.42}$$

$$= E[(a_t)^2] = \sigma_a^2 \tag{2.43}$$

which is a constant. Hence an ARMA model is incapable of representing the changing conditional variance. Above result can be easily extended to any ARMA(p,q) model. Such variations can be modeled by applying Auto-Regression on the square of the residuals, which is called as Auto-Regressive Conditional Heteroskedastic (ARCH)

model [42]. An ARCH (q) model can be represented as

$$v_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \ .. \ + \alpha_q a_{t-q}^2 \tag{2.44}$$

where v_t^2 is the conditional variance. $\alpha_0 > 0$, α_1, ... , $\alpha_{q-1} \geq 0$ and $\alpha_q > 0$, so that conditional variance is strictly positive. While ARMA models are linear regression on first order moments, ARCH model can be considered as an extension of this in the second order moments.

2.4.3 Fractional ARIMA (FARIMA) Model

ARIMA model cannot appropriately represent stationary processes with persistent and anti-persistent behaviour. Both these processes are characterised by slowly decaying auto-correlation function and will be classified as unstable processes, if one follows ARIMA model. It has been shown that such behaviour can be effectively modeled by *self-similar* models like fractionally differenced white noise process proposed in [43] (or Fraction Gaussian Noise as proposed in [44]), by extending the difference operator to take non-integer values, where $-\frac{1}{2} < d < \frac{1}{2}$ (later extended by Odaki to $d > -1$ in [45]). Similar process in the continuous domain is called Fractional Brownian Motion (FBM) [46]. *Self-similar* behaviour in time series acronym can be defined as follows; a process is *exact* second order *self-similar* if the Auto-Correlation Function (ACF) structure is preserved across different sampling intervals. The difference operation can then be expanded as a binomial series,

$$\nabla^d = (1-L)^d = \sum_{i=0}^{\infty} \binom{d}{i} (-L)^i = \tag{2.45}$$

$$1 - dL - \frac{1}{2}d(1-d)L^2 - \frac{1}{6}d(1-d)(2-d)L^3 - ..... \tag{2.46}$$

The autocorrelation of such processes for lag $k \to \infty$ is

$$\rho_k \sim k^{2d-1} \tag{2.47}$$

For values of $0 < d < \frac{1}{2}$ and $-\frac{1}{2} < d < 0$, the autocorrelation values decays at a rate lesser than exponential and is said to be *asymptotically self-similar*, which can also be called long range persistence and anti-persistence behaviour respectively. In such cases, combining fractional differencing with Short Range Dependent (SRD) models (like ARMA,ARCH), give a separate class of models, which can model long range persistence/anti-persistence and short range dependencies respectively, called *Fractional* ARIMA (p,d,q) model [43, 47–49]. When $d = 0$, it becomes a pure white noise process, which are completely uncorrelated.

2.5 Stability and Invertibility of an ARMA process

For a process to be represented by an ARMA(p,q) model, its AR and MA operators must be stable and invertible respectively. To explain the stability of the ARMA process, consider an AR(p) model which can be written as:

$$x_t - \phi_1 x_{t-1} - \ .... \ - \phi_p x_{t-p} = a_t \tag{2.48}$$

Applying lag operator, above equation can be written as:

$$x_t(1 - \phi_1 L - \ ... \ - \phi_p L^p) = a_t \tag{2.49}$$

$$x_t = \frac{a_t}{1 - \phi_1 L - \ ... \ - \phi_p L^p} \tag{2.50}$$

Assuming distinct roots for the polynomial in L,

$$1 - \phi_1 L - \ ... \ - \phi_p L^p = 0 \tag{2.51}$$

Equation 2.50 can be written as:

$$x_t = \frac{a_t}{(1 - \alpha_1 L)(1 - \alpha_2 L)...(1 - \alpha_p L)} \tag{2.52}$$

where $\alpha_1, \alpha_2, ..., \alpha_p$ are the inverse of the roots of the polynomial in L. If the absolute value of the root(s) are outside the unit circle (in the complex plane), then Equation 2.52 can be written as a converging series in random shocks, which is the condition for stability of the process [37]. It also means that an AR component of any ARMA model can be represented as an infinite but converging series in MA. But for parsimonious representation, it is written otherwise. ACF of such series decay exponentially with increasing lag. Since, estimations are always done on a finite number of samples, absolute value of the roots of the AR polynomial will always be greater than unity. Hence, if the absolute value of the roots are very near to unit circle, process should be classified as unstable. Auto-correlation function of such processes decay at a rate less than exponential and shows significant correlation for larger lags. Yet, there is a class of stationary processes whose autocorrelation decay very slowly such that,

$$\sum_{j=0}^{\infty} |\rho_k| = \infty \tag{2.53}$$

Such processes are called Long Range Dependent (LRD) or persistent processes (earlier mentioned in Section 2.4.3). Since there are infinite number of MA coefficients, it cannot be represented by a finite number of ARMA(p,q) coefficients. Fractional ARIMA(p,d,q) models with $0 < d < \frac{1}{2}$ are better in modeling such long range stationary trend, compared to its Integer ARIMA(p,d,q) counterpart with $d = 0, 1, 2..$.

To understand the invertibility condition for an ARMA process, consider an MA(q) model:

$$x_t = a_t - \psi_1 a_{t-1} - \ .... \ - \psi_q a_{t-q} \tag{2.54}$$

Since, it is represented in terms of finite number of past random shocks, MA(q) process is always stable. By using the lag operator explained earlier,

$$x_t = a_t(1 - \psi_1 L - \ ... \ - \psi_q L^q) \tag{2.55}$$

$$a_t = \frac{x_t}{1 - \psi_1 L - \ ... \ - \psi_q L^q} \tag{2.56}$$

Assuming distinct roots for the polynomial in L

$$1 - \psi_1 L - \ ... \ - \psi_q L^q = 0 \tag{2.57}$$

Equation 2.56 can be written as

$$a_t = \frac{x_t}{(1 - \beta_1 L)(1 - \beta_2 L) \ ... \ (1 - \beta_q L)} \tag{2.58}$$

where $\beta_1, \beta 2, ..., \beta_p$ are the inverse of the roots of the polynomial in L. If the absolute value of the root(s) are outside the unit circle (in complex plane), then Equation 2.58 can be expanded as a converging series in the past values of x_t. This condition is called the invertibility of the MA model [37]. Otherwise, it becomes a non-converging series in past x_t and hence cannot be meaningfully represented by finite number of ARMA(p,q) coefficients. But, there is a class of processes which are stationary and invertible with infinite number of AR coefficients. These are called anti-persistent or mean reverting processes, which can be modeled better using FARIMA, where $-\frac{1}{2} < d < 0$. Partial ACF (PACF) of these processes decay at a rate lesser than exponential, but the process does not possess any *LRD* property.

Stability and invertibility can also be explained in terms of the transfer function of the process. For this, consider an ARMA(p,q) model represented in Equation 2.35,

$$x_t - \phi_1 \ x_{t-1} - \ .. \ - \phi_p \ x_{t-p} = a_t - \psi_1 \ a_{t-1} - \ .. \ - \psi_q \ a_{t-q} \tag{2.59}$$

Taking *Z-transform* on both sides,

$$X(z)(1 - \phi_1 z^{-1} - \ .. \ - \phi_p \ z^{-p}) = A(z)(1 - \psi_1 z^{-1} - \ .. \ - z^{-q}) \tag{2.60}$$

The transfer function can then be written as

$$\frac{X(z)}{A(z)} = \frac{(1 - \psi_1 z^{-1} - \ .. \ - z^{-q})}{(1 - \phi_1 z^{-1} - \ .. \ - \phi_p \ z^{-p})} \tag{2.61}$$

which can be reduced to

$$\frac{X(z)}{A(z)} = \frac{(1-\beta_1 z^{-1})(1-\beta_2 z^{-2}) \ ... \ (1-\beta_q z^{-q})}{(1-\alpha_1 z^{-1})(1-\alpha_2 z^{-2})...(1-\alpha_p z^{-p})} \tag{2.62}$$

where $\alpha_1, \alpha_2, ..., \alpha_p$ and $\beta_1, \beta 2, ..., \beta_p$ are the poles and zeros of the transfer function respectively. For the process to be stable and invertible, absolute value of the poles and zeros must lie inside the unit circle (in the complex plane) [17]. Above condition can be obtained by replacing L with z^{-1} in Equations 2.51 and 2.57, which gives two new polynomials. Hence, the condition for stability discussed in the subsequent chapters will be, *absolute values of the roots of the AR polynomial or poles of the transfer function, must lie within the unit circle.* Similarly, condition for invertibility will be, *absolute values of the roots of the MA polynomial or zeros of the transfer function, must lie within the unit circle.* Estimation of the poles/zeros of the transfer function, ACF and PACF of a stationary process, provide sufficient *evidence* regarding the presence of persistent and anti-persistent behaviour of the process.

2.5.1 Duality Property

From the above discussions on the properties of AR, MA and ARMA models, aspects of duality between AR and MA process can be drawn as follows:

- In a stationary AR model of order p, the x_t can be written as infinite but converging weighted sum of past $a's$. Similarly, in an invertible MA model of order q, a_t can be written as infinite but converging weighted sum of past $x's$.
- Hence, in a stationary and invertible ARMA model, AR component can be written as an infinite but converging series in MA and vice-versa. But for parsimonious representation, it is written otherwise.
- Theoretically, the autocorrelation function of a stationary AR model of order p

decays at an exponential rate, but extent *infinitely*, while partial autocorrelation function become zero after lag p. Similarly, partial autocorrelation function of an MA model of order q decays exponentially, but extent *infinitely*, while autocorrelation function become zero after lag q.

Above properties are useful in identifying the presence and order of the AR/MA components, while building models for a given time series. More details of these models can be found in [37–39].

2.6 Summary

In this chapter, we have discussed the literature related to our work and have given a background on the time series models explored in this work. Following chapters addresses the parametric modeling of *fine-grained* characteristics of network time series over a long period of time, in the context of detecting *low intensity* TCP SYN DoS attacks.

CHAPTER 3

Characteristics of Long Range Network Time Series

The *fine-grained* time series analysis of network traffic collected over short period (few minutes to an hour) at different times of day, month and year, reported "bursty" and "self-similar" behaviour [6, 19]. Often the term *self-similarity* is used along (or rather used interchangeably) with *LRD* since *self-similar* process like FBM (continuous domain) or FARIMA (discrete domain) are widely seen to exhibit long memory. A stationary process is said to be LRD, if the absolute auto-correlation values as lag $k \to \infty$, are non-summable. The *Hurst Exponent* value (also called as H value or parameter) [50] lies within the range, $\frac{1}{2} < H < 1$ for the process to exhibit LRD behaviour. Nevertheless, *self-similarity* does not always mean LRD, as in the case of Brownian motion (i.e Weiner process) with $H = \frac{1}{2}$ and for anti-persistence process with $0 < H < \frac{1}{2}$. Further, the term *stationarity* needs to be emphasised since it is difficult to identify the presence of LRD if the process is non-stationary, as both cases show similar *evidences* like slowly decaying autocorrelation values and spectral density having a peak at zero frequency [51,52]. In the remaining chapter, an attempt is made to characterise the network time series at the finer level, collected over a long time range (of the order of days and months) from an edge router. We have also investigated briefly the *self-similarity* and *LRD* of such *long range* series, though a detailed analysis is beyond the scope of this study.

3.1 Model Parameter

The feature to be modeled for the detection of SYN attacks is chosen based on two criterion; Firstly, it has to exhibit different properties during a normal and attack period. Secondly, it should be as independent as possible with respect to the access pattern. Since the onus is on the service provider to detect and take corrective measures against these kind of attacks, the solution is proposed for the edge router at the victim side and hence is scalable to *Distributed* DoS (DDoS) attacks as well. Edge router connects the Local Area Network (LAN) within an organisation to the Internet. The number of such edge router(s) vary depending on the size of the organisation, type and demand of service it is providing to the Internet users. The *HOC* is found to be a good candidate satisfying the above two conditions. It can be defined as the difference in the number of incoming[1] TCP SYN packets and the number of outgoing[2] TCP SYN/ACK packets. During an attack, for each attack SYN request, six SYN/ACK response packets are generated by the server (for the default configuration discussed in chapter 1), which can be captured by the HOC. As per the TCP protocol, client receives the SYN/ACK response within one RTT, which is normally in the order of milliseconds. Hence, there is a strong correlation between the incoming SYN and outgoing SYN/ACK. Due to these reasons, HOC is studied in earlier literature as well [7, 9, 53].

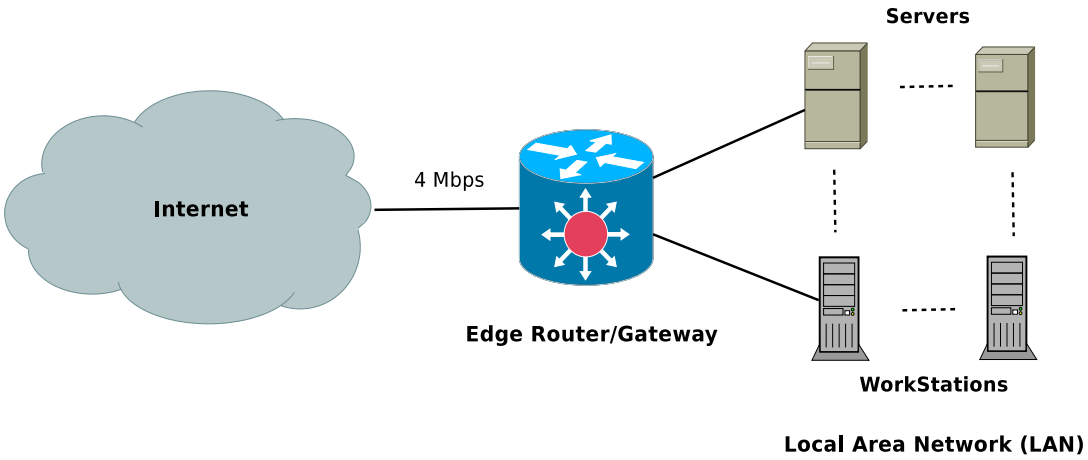


Fig. 3.1: Edge Network Scenario

3.2 Data set used in the Study

The network scenario studied in this work is shown in Figure 3.1. The edge router connects the Tenet network of IIT Madras [54] to the Internet. The bandwidth of the link connecting the gateway to the service provider is 4Mbps. The network trace is collected over a period of three months from July 15^{th} 2010 to September 30^{th} 2010, from the edge router using TCPDUMP utility [55]. From this, traffic to/from four servers which provide HTTP and SMTP services are identified and HOC feature is extracted at equally spaced time intervals of the order of seconds. Sampling interval in the order of seconds are chosen for two reasons. Firstly, the sampled *HOC* values were insignificant for smaller time intervals. Secondly, the exponential back-off and retransmission of SYN/ACK packets during a SYN attack happens in the order of seconds. It is this property of TCP which is being exploited for modeling the mechanism of a SYN attack [7, 9, 53]. For simplicity and to provide a representation of the traffic over 3 months, three data sets are created, each consisting of 5 days (i.e 24 hrs x 5

[1]Traffic entering from outside network

[2]Traffic leaving to outside network

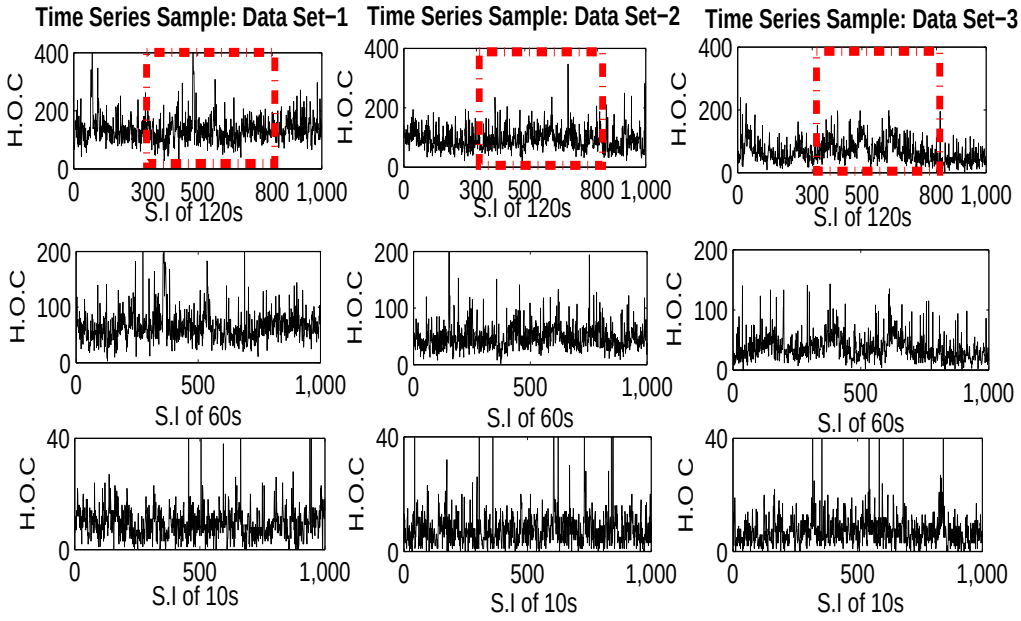


Fig. 3.2: Half-Open Count Characteristics

days) of continuous traffic, one from each month.

- Data Set-1: 26^{th} July 2010 to 30^{th} July 2010
- Data Set-2: 23^{rd} August 2010 to 27^{th} August 2010
- Data Set-3: 20^{th} September 2010 to 24^{th} September 2010

These data sets are checked for any existing TCP SYN DoS attack by verifying whether all the connection requests are valid. This is done to demonstrate the effectiveness of our approach in detecting SYN DoS attack, discussed in Section 4.3.

3.3 Time Series Analysis of HOC

For any time series analysis, visual inspection of the data can be very useful. The traffic behaviour with respect to the *HOC* for different sampling intervals are shown in Figure

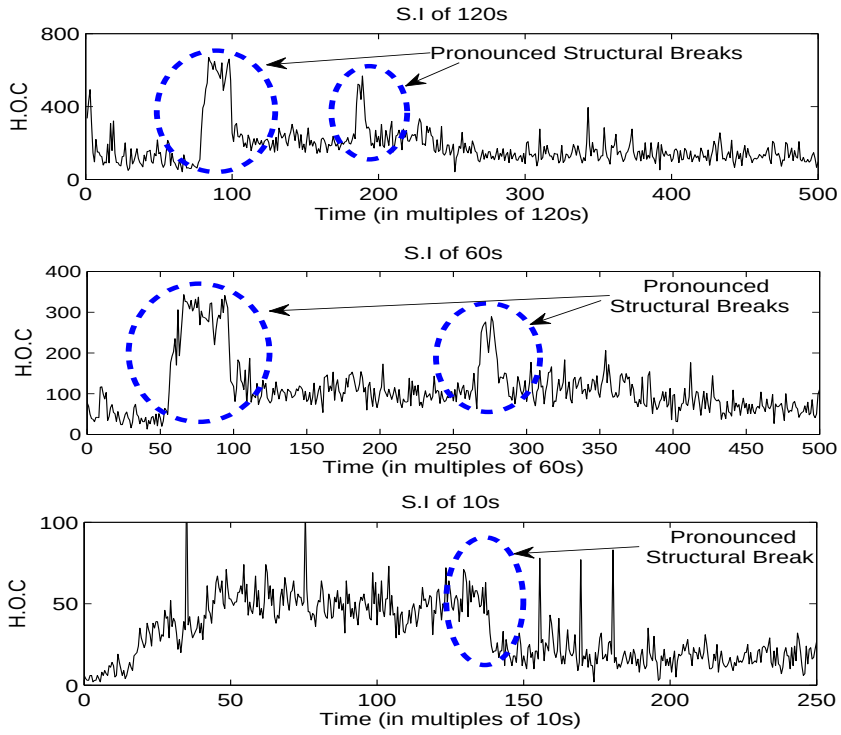


Fig. 3.3: Visible Structural Breaks in the Series

3.2[3]. An interesting observation is the similarity in the plots across different sampling intervals, which indicates the possibility of *self-similar or scale invariant* behaviour [6, 19]. The red dotted line in the case of sampling interval of 120s is expanded wholly for 60s and partially for 10s. More intriguing evidences on scale invariant behaviour can be seen in the statistical analyses discussed later.

A closer look shows *transient shocks*[4] as well as *structural breaks*[5] in the time series, irrespective of the sampling interval. On a formal note, the ACF of the series with structural breaks decays slowly while that containing only transient shocks decays at an exponential rate. Figure 3.3 shows some of the pronounced structural breaks across various sampling intervals. This trend is seen across all three data sets. Such structural

[3]S.I means Sampling Interval. Same abbreviation used in later figures as well

[4]Shocks whose effect last in the process for a small time duration are called Transient shocks

[5]Shocks whose effect last in the process permanently are called Structural breaks

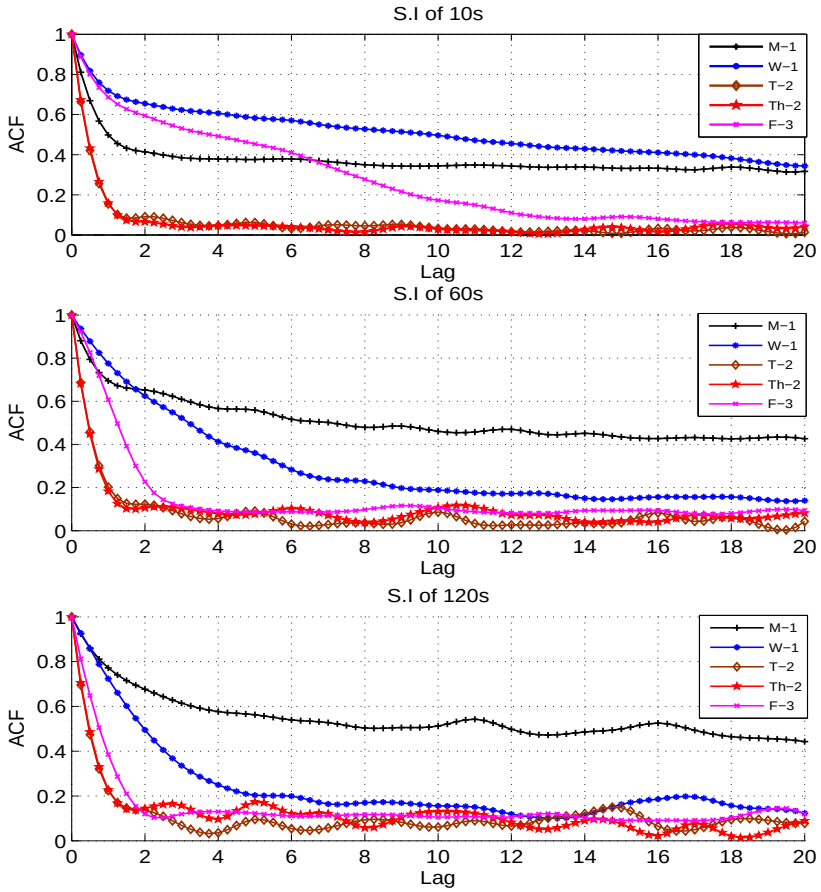


Fig. 3.4: Sample ACF Showing Non-Stationarity

breaks cause permanent *mean shifts* in the series and introduces non-stationarity in the data, as can be observed from the variations in the ACF plots across different data sets in Figure 3.4[6]. A more closer observation clearly shows (refer Figure 3.5) slowly decaying estimated auto-correlation values, which are artifacts of the *structural breaks* in the data and can sometimes be mis-interpreted as *LRD* behaviour. Blue lines shows the 95% confidence interval which is $\pm\frac{2}{\sqrt{N}}$, where N is the number of data points in the series. Coefficients within this window are considered insignificant or zero. This is also called Bartlett's approximation [37–39]. Non-stationarity can also be observed in

[6]M-1, T-1, W-1, Th-1 and F-1 correspond to Monday, Tuesday, Wednesday, Thursday and Friday data of data set-1. Similar naming convention applies to other data sets as well

the estimation of the 'H' parameter, which irrespective of the sampling interval, shows large deviations when estimated over different set of equal sized samples, as shown in Figure 3.6 (results from Periodogram estimator). Since different Hurst estimators may give conflicting results and no estimator is perfect in all conditions [50], we have used three estimators, namely R/S, Periodogram and Absolute value methods for estimating the Hurst exponent. Even though all estimators gave slightly different 'H' values in each case, overall trend remained as shown in Figure 3.6.

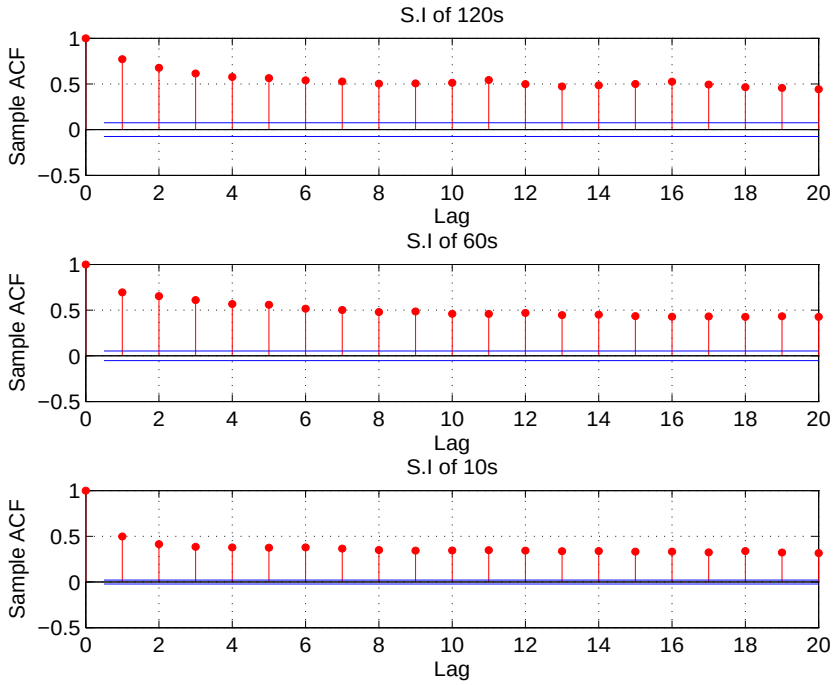


Fig. 3.5: Slowly Decaying ACF

Identifying an appropriate operation which can decouple the stationary component from the non-stationary component of the process helps in modeling the stationary component efficiently.

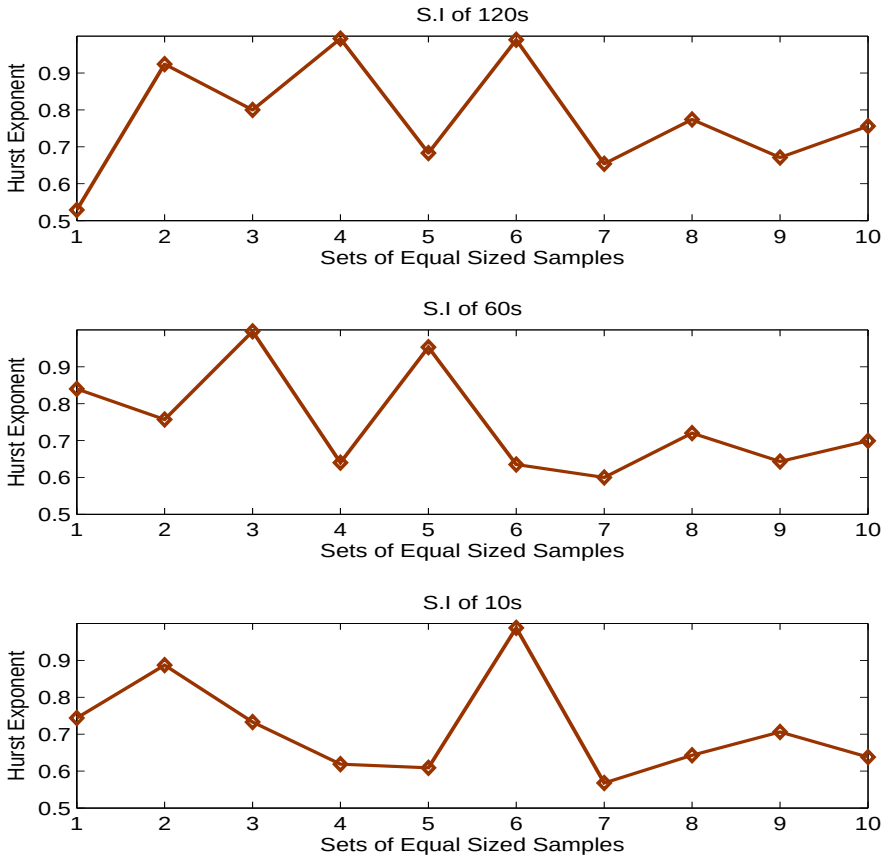


Fig. 3.6: Hurst Estimate of Half-Open Count

3.3.1 Modeling Structural and Transient Breaks

Let us consider the time series under study, which is a sample realisation of the process we need to model, as a combination of *structural breaks* and *transient shocks*[7] (originally proposed by Chung Chen and George C. Tia in [56]):

$$y_t = s_t + z_t \tag{3.1}$$

$$s_t = s_{t-1} + \rho_t \, \delta_t \tag{3.2}$$

where y_t is the stochastic process, which is a combination of non-stationary structural break component (s_t) and stationary transient break component (z_t). The transient

[7]Terms *breaks* and *shocks* are used interchangeably

shocks can either be a pure random process (white noise process, a_t) with no serial correlation or may have serial dependencies which decay exponentially and can be modeled using SRD models like ARMA, ARCH (or a combination of more than one model), etc. ρ_t follows an *I.I.D* (Independent and Identically Distributed) binomial distribution with,

$$\rho_t = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases}$$

δ_t is assumed to follow an *i.i.d N(0,*σ_δ^2*)* distribution and is responsible for *structural breaks* in the series. When the probability p becomes one, the model in Equation 3.1 represents a combination of random walk process (without drift) [40], which has a break at every sampling period and a stationary process. If zero, it reduces to a stationary process. But, what happens if $0 < p < 1$? In other words, how to model the process when the *structural breaks* are significant, but not occurring at every time interval?

To better understand how the first and second order moments of the random level shift process discussed earlier varies with probability of structural breaks, consider Equations 3.1 and 3.2, which can be re-written as,

$$y_t = s_0 + \sum_{j=1}^{t} \rho_j \ \delta_j + z_t \tag{3.3}$$

where s_0 is the initial value of s_t. For ease of explanation, consider z_t as a pure random process a_t with no serial dependencies (*N(0,*σ_a^2*)*) and initial values $s_0 = 0$, $\rho_t = 0$, $\delta_t = 0$ and $a_t = 0$ for $t \leq 0$. Further, a_t, ρ_t and δ_t are considered mutually independent. Then, the mean and variance of the process can be derived as,

$$E(y_t) = 0 \tag{3.4}$$

$$Var(y_t) = t \ p \ \sigma_\delta^2 + \sigma_a^2 \tag{3.5}$$

Variance of the process is time variant and hence is not stationary. The covariance for lag k, where $k > 0$ is

$$\gamma_k = E(y_t \; y_{t+k}) = E((\sum_{j=1}^{t} \rho_j \; \delta_j + a_t)(\sum_{j=1}^{t+k} \rho_j \; \delta_j + a_t)) \tag{3.6}$$

Since, a_t, ρ_t and δ_t are mutually independent and identically distributed, above equation becomes

$$\gamma_k = t \; p \; \sigma_\delta^2 \tag{3.7}$$

Then, auto-correlation value for any lag k, can be estimated as

$$\lambda_k = \frac{t \; p \; \sigma_\delta^2}{\sqrt{t \; p \; \sigma_\delta^2 + \sigma_a^2} \; \sqrt{(t+k) \; p \; \sigma_\delta^2 + \sigma_a^2}} \tag{3.8}$$

$$\lambda_k \propto tp \tag{3.9}$$

where tp is the expected number of structural breaks during time duration t. Hence, with increase in number of structural breaks, auto-correlation values also increases. As $t \to \infty$, if $p > 0$, $\lambda_k \to 1$, for any lag value k. Since, for all practical purposes, the number of data samples being finite, one have to look for *evidences* from the time series data, to understand whether the *structural breaks* are significant. Sample ACF across different data sets in Figure 3.4 shows non-deterministic and slowly decaying trend with increasing lag values. This shows existence of significant *structural breaks* in the data. In order to verify and quantify this further, *Canny Edge Detector Algorithm* [57] is used to find the probability of structural breaks with increasing time duration. Algorithm first removes the high frequency component (which contributes to the frequent spikes as can be observed in Figure 3.2) from the time series using Gaussian filtering before finding the point(s) (local maxima or minima) where the gradient reaches a peak value, as shown in Figure 3.7. The red dotted lines show the points at which the algorithm has detected structural breaks. Figure 3.8 shows that across all sampling intervals,

the probability of structural break is consistently greater than zero and approaches a constant value with increasing number of samples. This confirms the significant presence of *structural breaks* in the series.

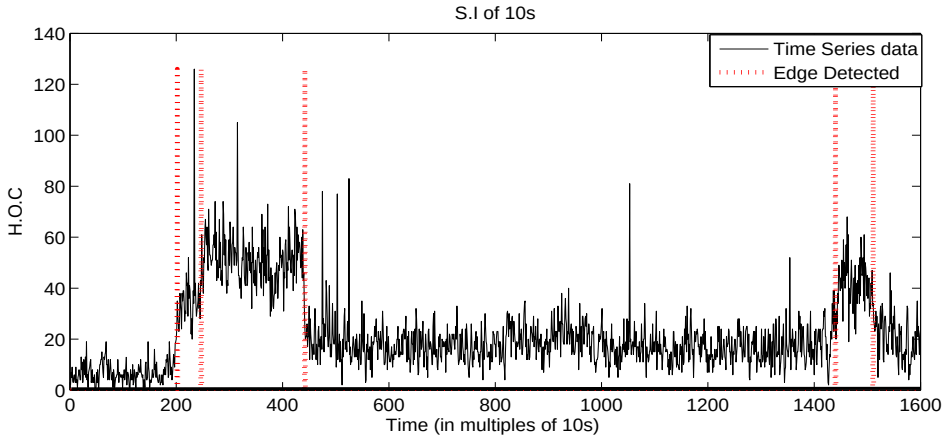


Fig. 3.7: Edge Detection in Time Series: S.I of 10s

3.3.2 Discussion

Identifying an appropriate operation which can decouple the stationary component from the non-stationary component of the process helps in modeling the stationary component efficiently. By taking absolute difference of the adjacent values, the non-stationarity induced by the *structural breaks* can be removed and only the effects of *transient shocks* are retained. The idea is to model the dependencies in the *transient shocks* by considering *structural breaks* to be *i.i.d.* Since, we are interested only in the absolute value of the shocks and not in the sign, an absolute difference operation is performed rather than traditional differencing operation discussed earlier in Section 2.4. Certainly, there will be a penalty in doing an *absolute difference* transformation, which will be justified during model diagnosis and TCP SYN detection stages, explained in the next chapter. Also, it has been observed from the studies done in core

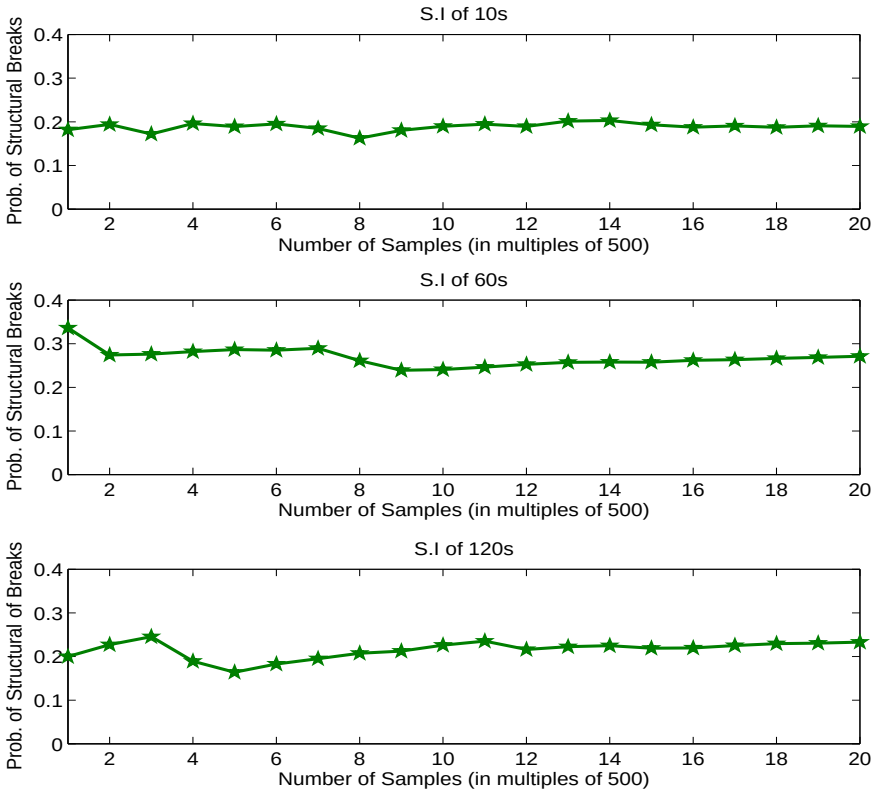


Fig. 3.8: Probability of Structural Breaks

network that the effect of *low rate* attacks is larger on the short term dynamics than on the trend component. This is another reason for modeling the short term dynamics (transient shocks) by removing the trend induced by the structural breaks.

Further, network traffic is known to exhibit volatility clustering [4, 53, 58], which means that large variations are followed by large variations and small variations are followed by similar small ones, regardless of the sign of the variation. The characteristics of traditional differencing and absolute differencing on *long* network time series are shown in Figure 3.9. Clusters of larger and smaller variations can be seen in the case of latter, while it is not so evident in the case of the former. Similar trend is observed for other sampling interval as well. This also motivated for taking absolute value of the difference operation rather than traditional differencing operation. The presence

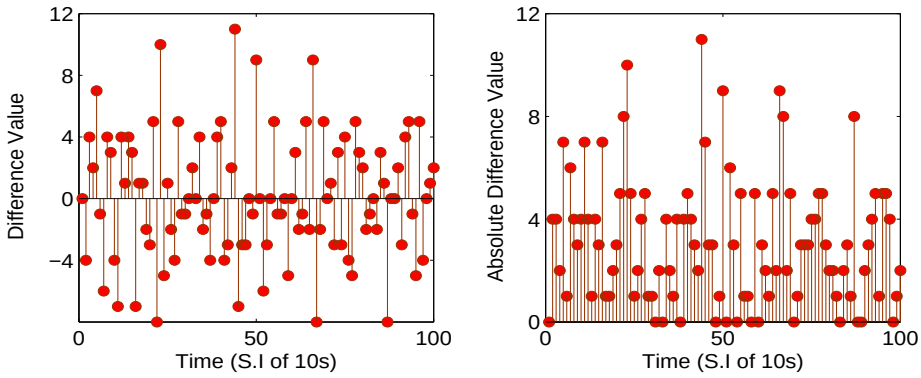


Fig. 3.9: Difference Vs Absolute Difference

of clustering behaviour has to be formally identified during model identification and diagnosis, which will be done in the subsequent chapter. It is interesting to note that because of similar volatility behaviour, absolute value of the *returns* are extensively studied in econometrics [59–61].

3.4 Summary

It is observed that the *HOC* time series is a combination of *structural breaks* and *transient shocks*. Further, it is also shown that such *structural breaks* are significant in the time series data and need to be taken into account while modeling the underlying data generating process. An *absolute differencing* operation is done to remove the non-stationary effects caused by the *structural breaks* and model the *transient* shocks. The derived series can be modeled using regression based time series models provided the adjacent values in the series have significant correlations and the process is stable as well as stationary. Modeling the derived series by studying the stationarity and stability is discussed in the subsequent chapter, with discussion on the *evidences* of *self-similarity* and *LRD*. Efficacy of the model proposed for detecting TCP SYN DoS

attacks are also studied.

CHAPTER 4

Modeling and Performance Analysis

Earlier work using ARMA models and its extensions [2, 4, 8–10, 53] haven't studied in detail the stationarity and stability of network time series. This chapter discusses these properties with respect to the *HOC* and *Difference*[1] series, by studying the first and second order moments of the same. Later, based on this analysis, appropriate model is proposed, which is then used for detection of *low intensity* TCP SYN attacks.

4.1 Stationarity and Stability

4.1.1 *HOC Series*

Table 4.1 lists the sample mean calculated for different days across all three data sets for all sampling intervals. It shows variations as large as 50% within as well as across data sets. Similarly, ACF plotted for various days across all three sampling intervals do show any consistency within and across data sets (refer Figure 4.1). This indicates that the *HOC* series is non-stationary. Also, it is worth mentioning that the relatively significant correlation for larger lags in this case should not be misunderstood as *LRD*, since *LRD* can be defined *only* when the process is stationary [50, 62]. Studies in [10], [12] and [63] define stationarity of the time series in terms of *one* sample ACF. It is stated in these studies that if a process exhibits a slowly decaying sample ACF, it's non-stationary, else stationary. According to this definition, from the sample ACF of

[1]The series derived from absolute differencing operation will henceforth be referenced as *difference* series

Day	Data Set-1	Data Set-2	Data Set-3
Monday	13.41	7.90	8.16
Tuesday	11.43	8.16	6.70
Wednesday	14.09	8.41	4.99
Thursday	14.37	8.44	4.71
Friday	13.40	8.27	6.00

(a) Sampling Interval: 10s

Day	Data Set-1	Data Set-2	Data Set-3
Monday	78.93	45.37	45.80
Tuesday	66.93	47.02	37.27
Wednesday	82.44	48.25.9	27.24
Thursday	84.32	47.97	25.52
Friday	79.24	47.51	32.94

(b) Sampling Interval: 60s

Day	Data Set-1	Data Set-2	Data Set-3
Monday	157.74	90.65	91.41
Tuesday	133.57	93.91	74.08
Wednesday	164.66	96.36	53.84
Thursday	168.52	95.60	50.60
Friday	158.37	94.97	65.35

(c) Sampling Interval: 120s

Table 4.1: Mean: *HOC* Series

HOC series for Tuesday and Thursday of data set-2, the series is stationary. But, similar consistent behaviour is not seen for sample ACF of other days. Hence, it's difficult to conclude about stationarity of the process from just *one* sample ACF. On the other hand, it is already discussed in Section 3.3.1 that such trends are induced by the *structural breaks* in the series.

Stability of a process has often been interpreted as stationarity (under the assumption of a time invariant system) in many time series studies [37–39]. According to this assumption, a stable process is always stationary and vice-versa. Hence, stability of the *HOC* series is also studied to understand its stationarity property. In order to do this, *HOC* series for each day are modeled as an AR process. It is observed experimentally that a model order of three and one day's data is adequate for modeling. The model coefficients can be computed using Yule-Walker estimation as discussed in Section 2.3.1. The roots of the characteristic equation discussed in Section 2.5 for each

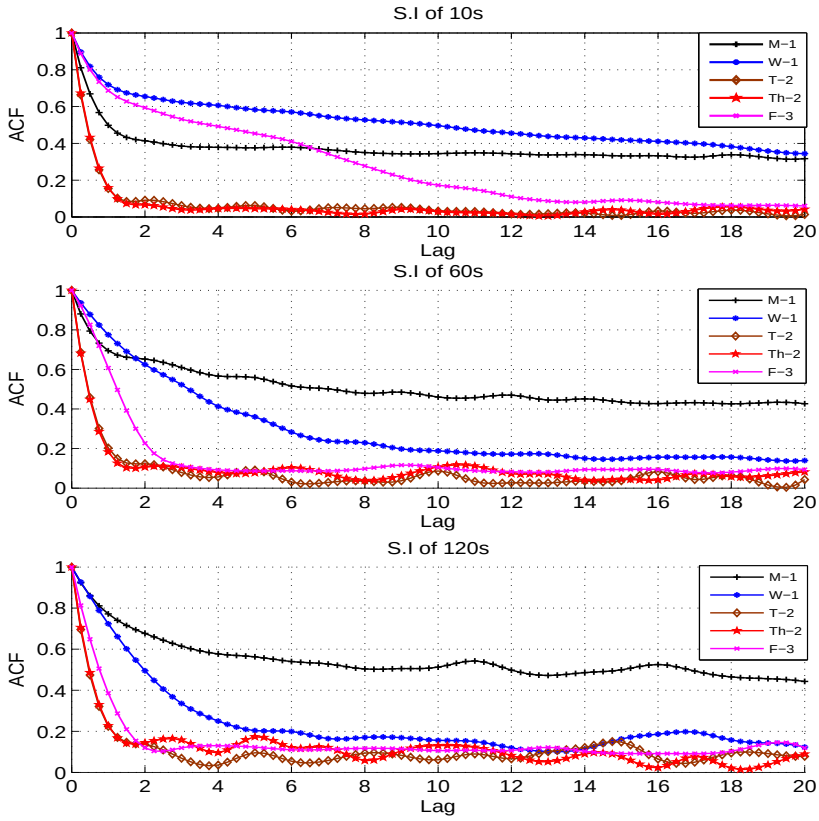


Fig. 4.1: Sample ACF of *HOC* Series

day are computed. Roots of the *HOC* series form two clusters as shown in Figure 4.2, with one cluster well within the unit circle bound, but the other cluster near to unity. Similar trend is repeated for all three sampling intervals. Hence, it can be concluded that the *HOC* series is unstable and non-stationary.

4.1.2 Difference Series

Table 4.2 lists the mean for the difference series for various days across the data sets and sampling intervals. It can be observed that the means show consistent behaviour within and across three data sets. Similar consistent behaviour is seen in the case of sample ACF plotted for different sampling intervals in Figure 4.3, implying stationarity in the data. In order to cross examine this claim, stability of the difference series is

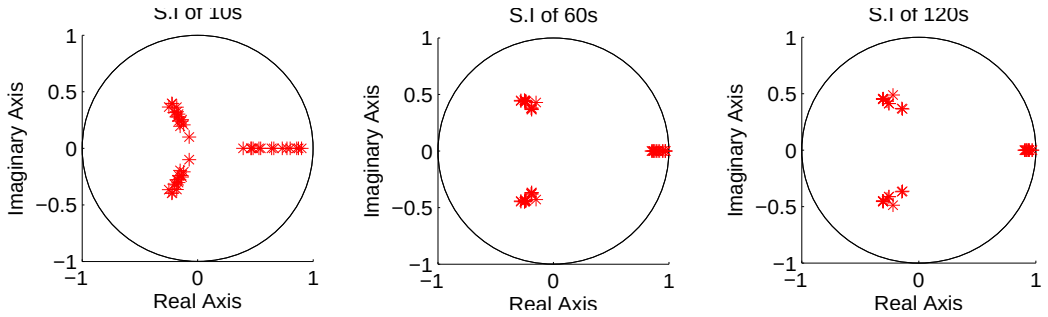


Fig. 4.2: Roots of HOC Series

studied. The roots of the characteristic equation for all days are estimated as discussed in Section 2.5. Figure 4.4 shows the roots of the difference series and it can be observed that the roots are well within the unit circle compared to that of the *HOC* series, for all three sampling intervals. Hence, the absolute differencing transformation has made the series stable and time invariant (stationary).

4.2 Modeling and Prediction

4.2.1 Model Selection

The initial approximation on the model orders for the AR and MA part is obtained from the estimates of ACF and PACF. Absolute values of the roots of the AR polynomial being well within the unit root circle (refer Figure 4.4),fastly decaying sample ACF (refer Figure 4.5) and PACF (refer Figure 4.6), rules out the possibility of long range persistence or anti-persistence. Further, MA component is found to be absent in the series. Hence, FARIMA model is not considered in this work.

Day	Data Set-1	Data Set-2	Data Set-3
Monday	5.5403	5.0499	4.8475
Tuesday	5.0008	5.0958	4.2121
Wednesday	5.6499	5.1832	3.7834
Thursday	5.7435	5.0452	3.6704
Friday	5.3730	5.2722	3.9951

(a) Sampling Interval: 10s

Day	Data Set-1	Data Set-2	Data Set-3
Monday	23.77	20.41	20.22
Tuesday	21.76	20.87	17.00
Wednesday	25.34	25.39	17.34
Thursday	23.04	20.07	16.52
Friday	21.09	21.02	17.55

(b) Sampling Interval: 60s

Day	Data Set-1	Data Set-2	Data Set-3
Monday	42.18	30.13	33.65
Tuesday	37.36	33.88	27.31
Wednesday	49.54	45.75	28.85
Thursday	43.38	31.75	27.72
Friday	37.48	34.53	32.99

(c) Sampling Interval: 120s

Table 4.2: Mean: Difference Series

4.2.2 Modeling, Prediction and Validation

It is observed empirically that an ARMA(3,0) or simply AR(3) model can efficiently represent the serial correlation in the conditional means. The model coefficients are estimated using *Yule-Walker* method as discussed in Section 2.3.1. Also, it is observed that one day data is sufficient for training the model (i.e estimating the model coefficients), since there was no significant improvement in the prediction error for larger training data. To justify this, using each day's data for training, 15 models are built (there are 15 days across all three data sets) and tested by using it for prediction on the respective training data itself. If the model is adequate, autocorrelation of the residual errors during prediction (refer Equation 2.12 of Section 2.3) will behave like that of a pure white noise process with 95% of values within $\pm\frac{2}{\sqrt{N}}$ (also called Bartlett's approximation), where N is the number of predicted samples [37,38]. This bound will

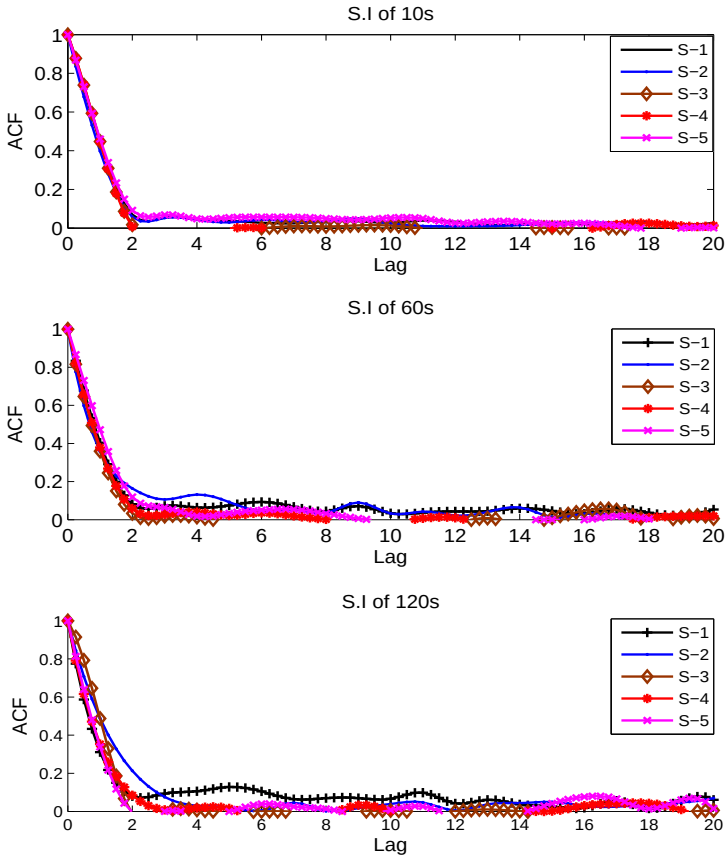


Fig. 4.3: Sample ACF of Difference Series

henceforth be called the *95% confidence interval*. In order to study the spread of the autocorrelation values of the error obtained during testing, boxplot of the same for all days and across all sampling intervals are plotted in Figures 4.7, 4.8 and 4.9. The blue box represents the inter-quartile range with red line within it representing the median of the correlation values. It is found that for all days, 99% of the correlation values are spread within the *95% confidence interval*. This means that the residual error obtained during prediction is totally uncorrelated in the first order moments and hence satisfy the equation for the AR model discussed in Section 2.3. *T-test* and *Chi-Square* tests conducted on the autocorrelation values of the residual, accepted the null hypothesis that 95% autocorrelation values are not significantly different from zero. Results of

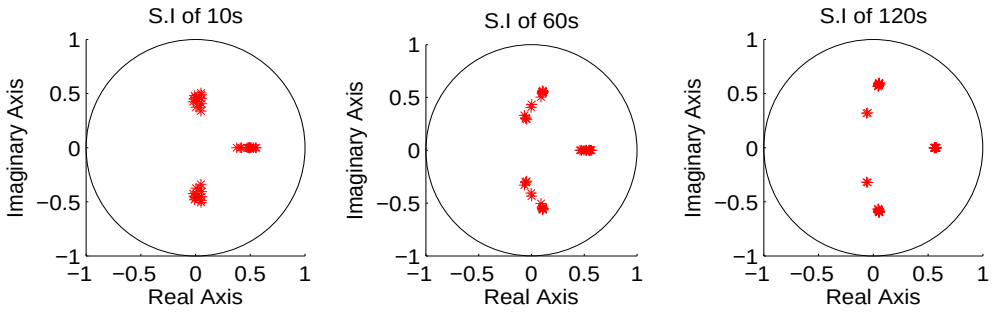


Fig. 4.4: Roots of Difference Series

Table 4.3: RMSE (Sampling Interval of 10s)

Day	Model using F-1	Model using F-2	Model using F-3
M-1	8.78	8.76	8.76
Th-1	7.45	7.46	7.47
M-2	7.56	7.52	7.52
Th-2	7.59	7.56	7.56
M-3	7.75	7.75	7.74
Th-3	6.87	6.83	6.82

N-fold cross validation, by testing model built on each of the day with other days, also shows consistent performance in terms of prediction error. A snapshot of this validation, by estimating the Root Mean Square Error (RMSE) of forecast error, using models built on all three Friday data, to test on other days across three data sets are shown in Table 4.3. It can be observed that, on any test day, RMSE is consistent across all three models. It is worthwhile to note that the square of the residuals show some correlation for smaller lags as observed from Figure 4.10. This signifies the volatility clustering discussed in the previous chapter. It is experimentally observed that ARCH model [42] of order 1 can efficiently model this correlation in the second order. Since, there is no significant improvement in the detection efficiency by including the ARCH stage, it is not included in the proposed model.

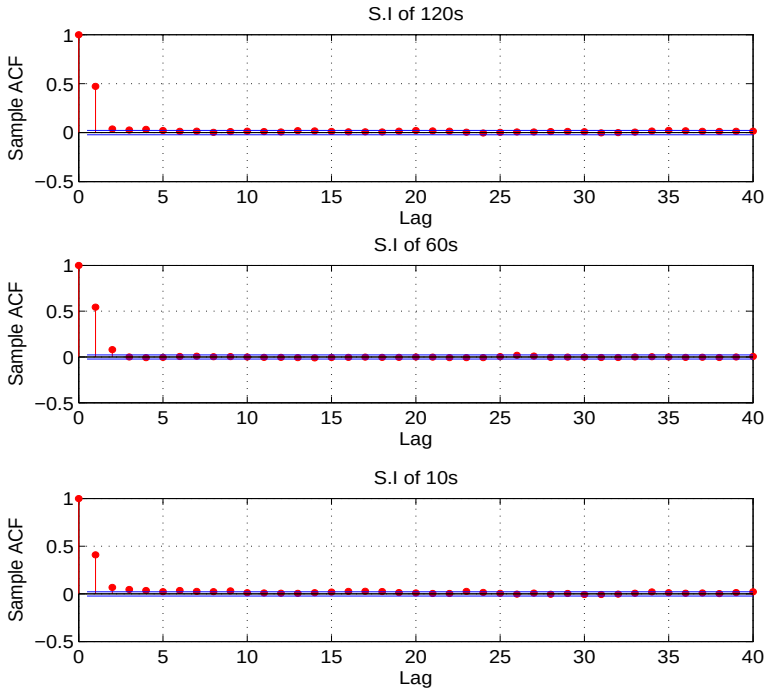


Fig. 4.5: Sample ACF of Difference Series

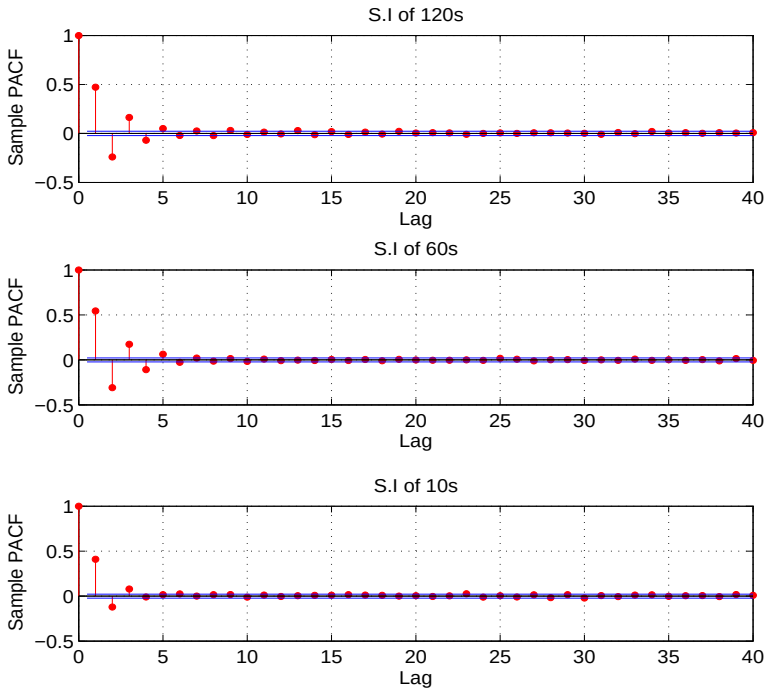


Fig. 4.6: Sample PACF of Difference Series

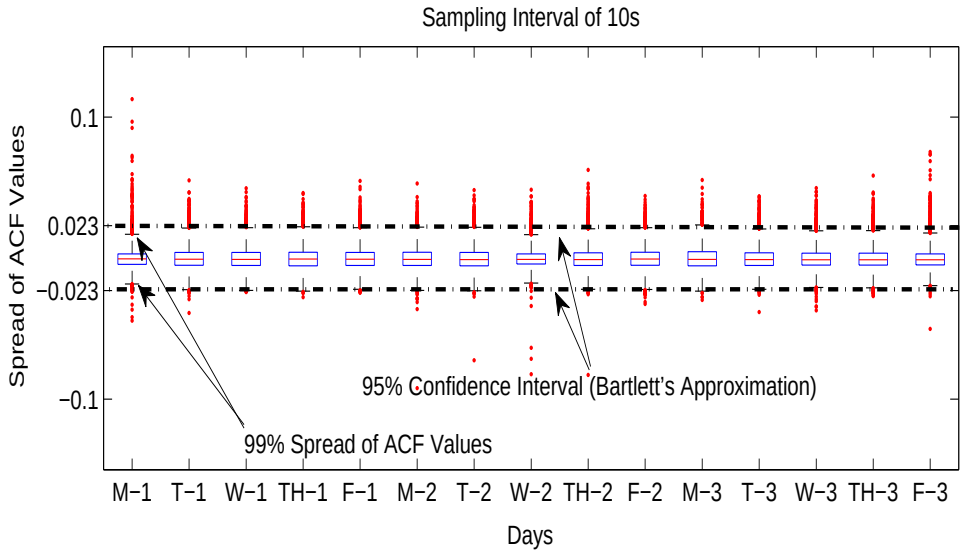


Fig. 4.7: ACF Spread of AR Res.: Sampling Interval of 10s

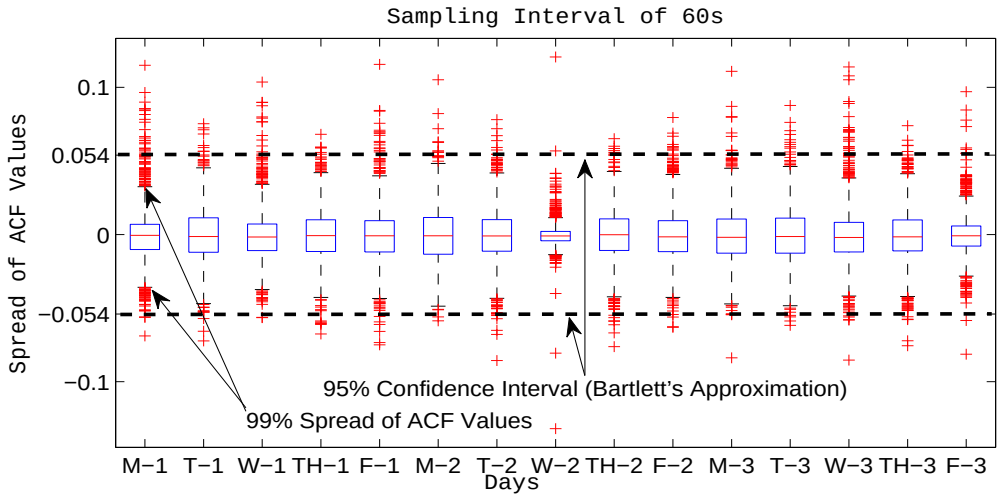


Fig. 4.8: ACF Spread of AR Res.: Sampling Interval of 60s

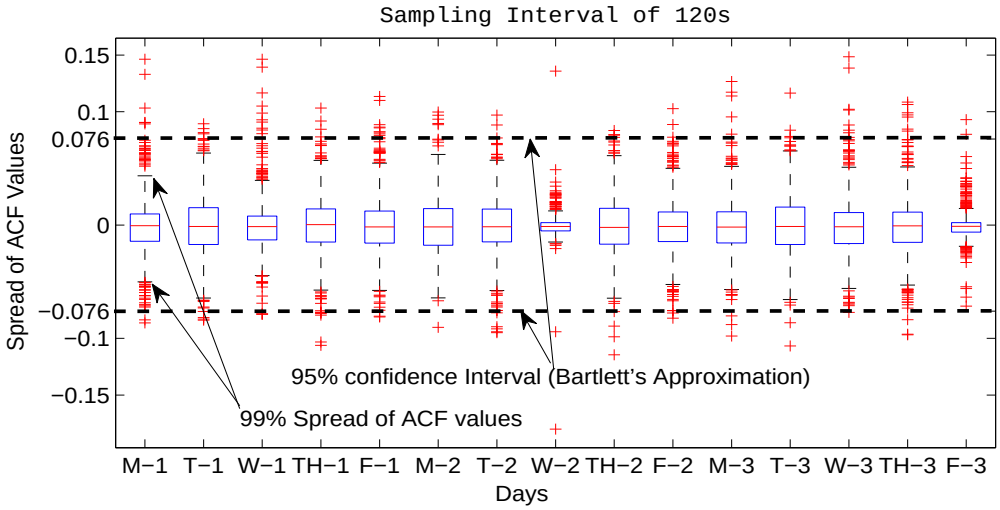


Fig. 4.9: ACF Spread of AR Res.: Sampling Interval of 120s

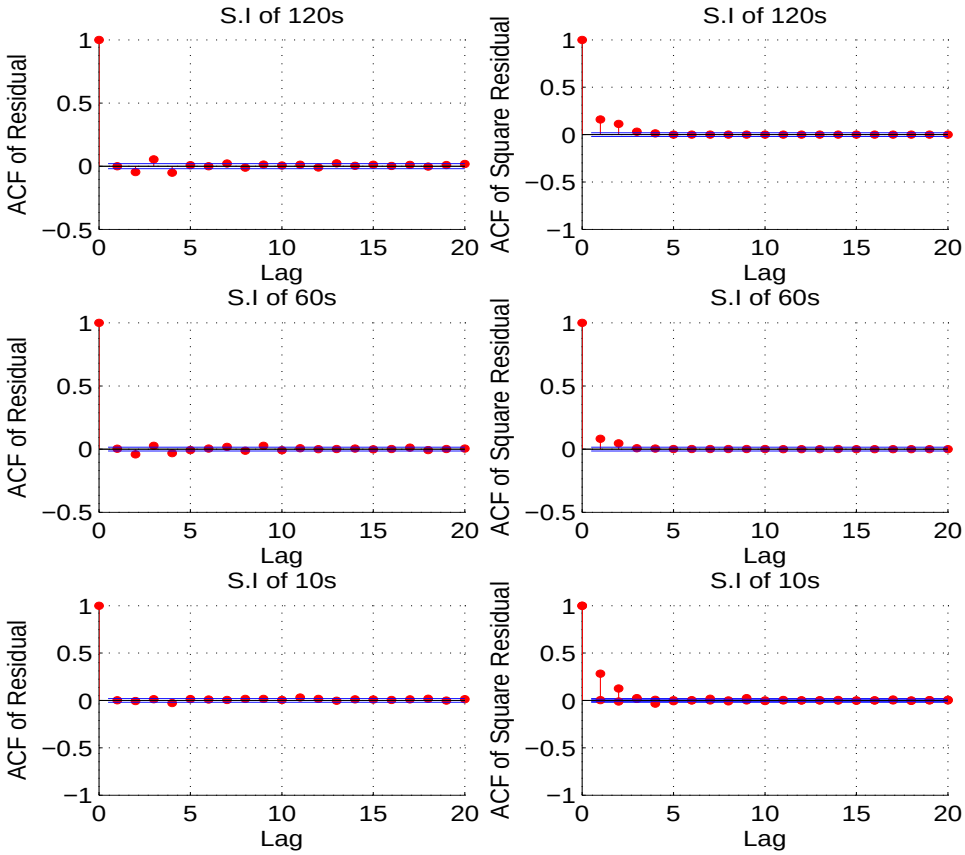


Fig. 4.10: ARCH Effect

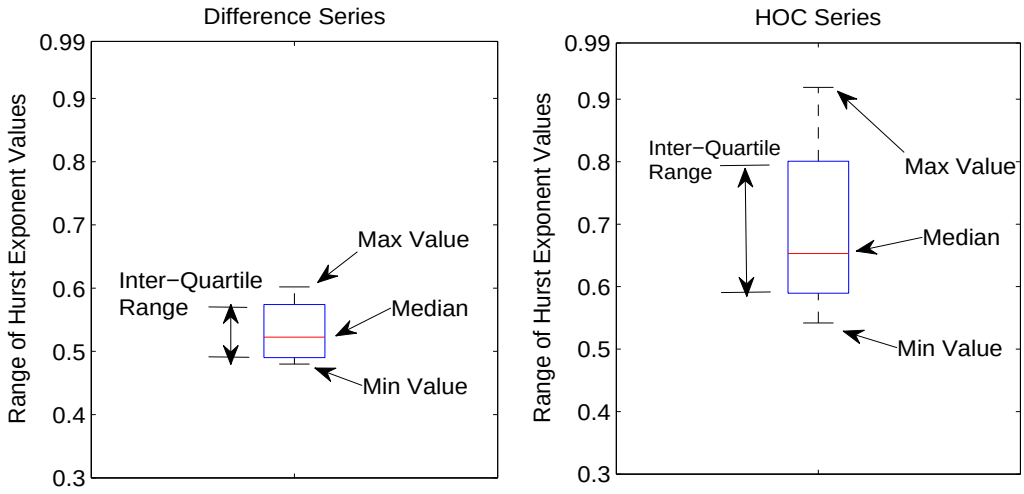


Fig. 4.11: Hurst Estimate: Sampling Interval of 10s

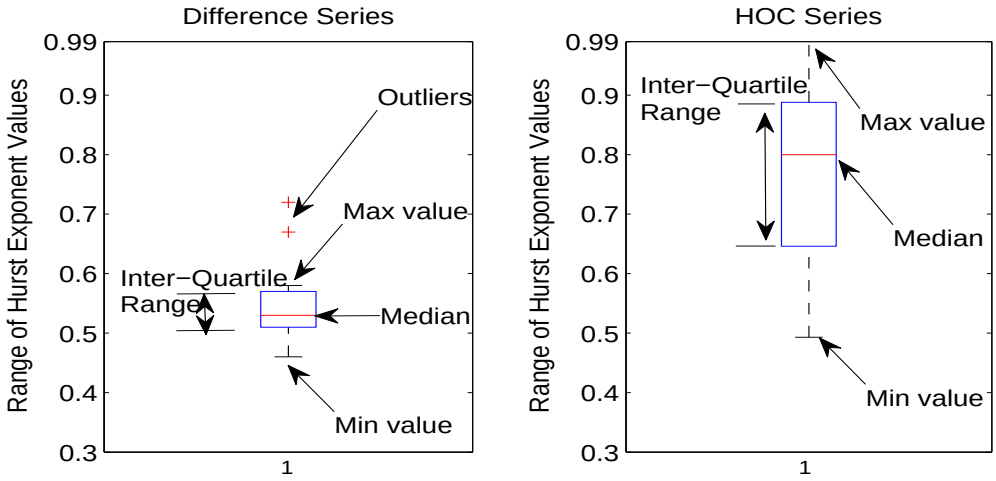


Fig. 4.12: Hurst Estimate: Sampling Interval of 60s

4.2.3 Discussion

The sharp changes in the time series values (called as *structural breaks* in this thesis) is reported as a passing reference in the context of RTT series [64]. Though structural breaks have been extensively studied in economic time series [56,61], it has not received much attention in the analysis of network time series. Further, similar models being

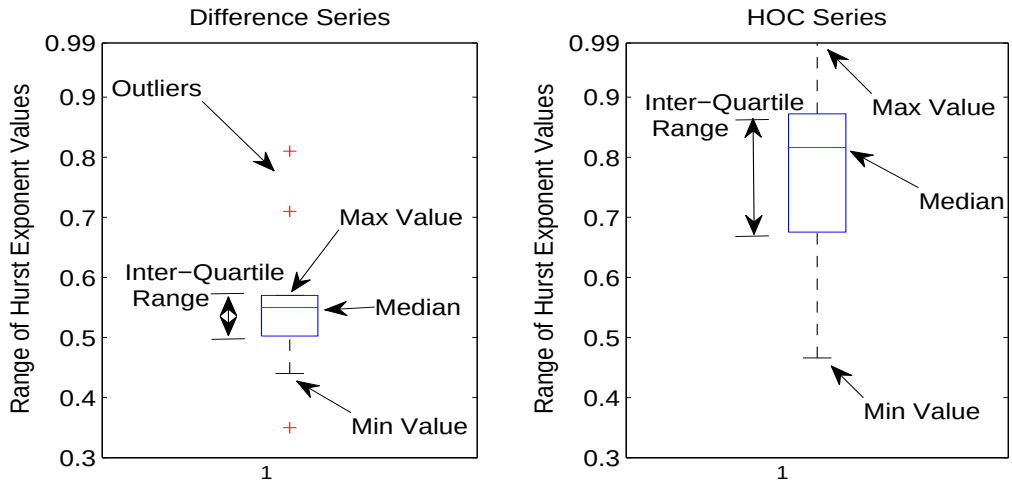


Fig. 4.13: Hurst Estimate: Sampling Interval of 120s

found useful in the context of forecasting 'packet size' time series in [4], demonstrate the generality of the proposed approach. The *evidences* of *self-similar* behaviour can be seen throughout various analyses explained in the previous Sections. For sampling intervals of 10s, 60s and 120s: Similar visual characteristics can be seen as shown in Figures 3.2 and 3.3, similar behaviour with respect to probability of structural breaks as seen in Figure 3.8, similar behaviour of the roots of the AR polynomial as shown in Figure 4.4, similar modeling behaviour and finally *exact self-similarity* [6] in the structure of ACF as shown in Figure 4.5. The Hurst Exponent estimation of the absolute differenced series (over different samples) from various estimators shows values *near* to 0.5 as shown in Figures 4.11, 4.12 and 4.13 (results from Periodogram estimator), which means that the realisation is very close to a Fractional Gaussian Noise (FGN) with $H \approx 0.5$ [43, 50] (The large variations in Hurst value for *HOC* series shows its non-stationary behaviour). This probably explains the self-similarity property seen across sampling intervals. At the same time, it is important to note that no evidence of *LRD* is observed. This is evident from the roots of the AR

polynomial shown in Figure 4.4, exponentially decaying ACF (refer Figure 4.5) and PACF (refer Figure 4.5) structure. These findings are similar to the discussions in [4, 65, 66]. Absence of "strong" *LRD* is reported in [4], while [65] and [66] discusses that *long tailed* distribution (for eg. Pareto distribution), which is commonly used to represent *LRD*, is not necessary to model network traffic. On the other hand, *log-normal* distribution is found to be appropriate. This means that the argument for *LRD* in network traffic is weak. A possible reasoning can be as follows; given a stationary time series, SRD models like ARMA or ARCH, has relevance in modeling the data generating process, as long as the value of H parameter is close to 0.5. We believe that network traffic belongs to this category. As the value of H moves towards 1 (but strictly less than 1), the stationary process shows long term persistent behaviour and if it moves towards 0 (but strictly greater than 0, later extended to $d > -1$ in [45]), the stationary process shows long term anti-persistent behaviour. Either of these cases can be better modeled using *fractional difference* models like FARIMA(p,d,q), where the fractional difference value $d = H - 0.5$, as explained in Figure 4.14. But, still there are several unanswered questions like "how far the value of H can vary on either side of 0.5 to be considered for modeling using SRD models?". Perhaps the answer lies in the ACF and PACF trend. A deeper analysis on this topic is beyond the scope of this report and is left as future work. A snapshot of the different stages of modeling is shown in Figure 4.15.

4.3 TCP SYN Attack Detection

The rationale behind our approach is that the prediction error during a normal period will be different from that of the abnormal period. The prediction errors obtained from

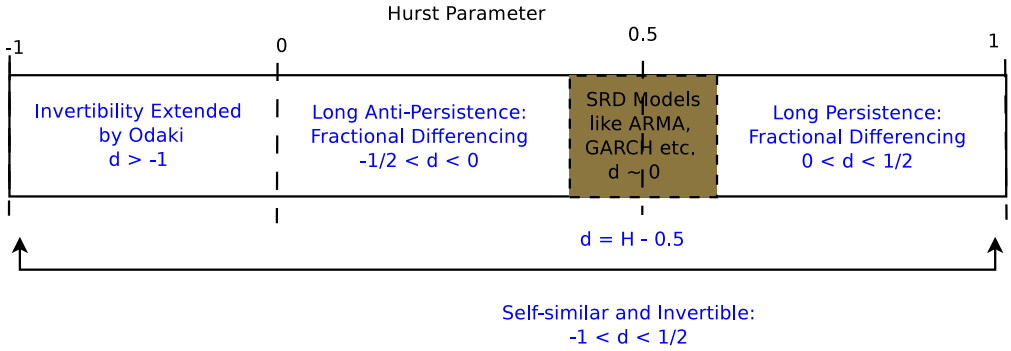


Fig. 4.14: Effect of 'H' on a Stationary Process

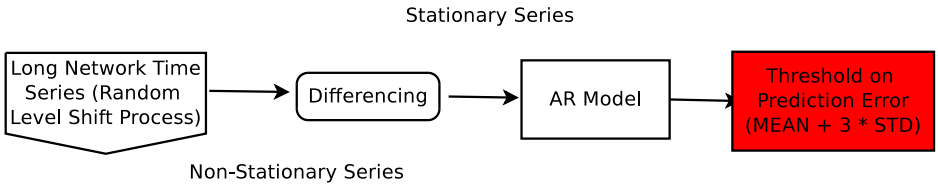


Fig. 4.15: Snapshot of Different Stages in Modeling

the model is found to follow a normal distribution with mean 0. A threshold is fixed on the absolute value of the prediction error to detect the attacks. This is equivalent to folding the distribution of error at the mean value, which is also called *half-normal* distribution. The mean and variance of the *half-normal* distribution, given that the original normal distribution has mean μ and variance σ^2 is:

$$\mu_{half-normal} = \sigma\sqrt{2/\pi} \tag{4.1}$$

$$\sigma^2_{half-normal} = \sigma^2(1 - 2/\pi) \tag{4.2}$$

The value of the threshold is estimated as sum of the mean and three times the standard deviation.

4.3.1 Attack Simulation

Low intensity TCP SYN DoS attacks, similar to the ones discussed in [11] and [12] are studied in this work. In such attacks, attacker aims to degrade the efficiency of the service rather than denying the service completely, by flooding the network with large number of spoofed SYN packets and remains undetected in the background traffic. It is observed that SYN arrival during normal period follows a *log-normal* distribution (refer Figure 4.16) similar to what has been observed for the distribution of 'file sizes' in [65] and 'inter-arrival packet time' in [66]. The mean and standard deviation of the number of SYN arrivals per second is observed to be 4.45 and 4.32 respectively.

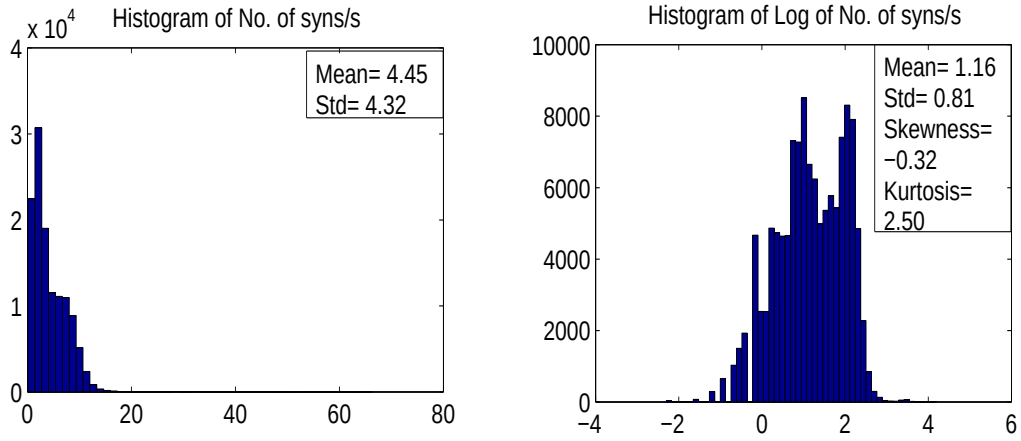


Fig. 4.16: Histogram Plot

Attacks of different rates and nature are simulated: (1) Attacks with rate uniformly varying between 10 to 20 syns/s (mean rate is 13.26 syns/s) (2) Attacks with rate uniformly varying between 2 to 8 syns/s (mean rate is 3.19 syns/s) and (3) Slowly increasing/decreasing attacks with rate varying between 0.5 to 5 syn/s. The rate is incremented/decremented in a step fashion by 0.005 between 0.5 and 5, after sending each *SYN* packet (mean rate is 0.91 syns/s). This is a modification on the *stealth* attack discussed in [12], where the attack rate always increases. Each type of attack is

Method	FP(%)	DD(sampling units)	DR(%)
AR	2.79	1	100
CUSUM	6.39	1	100
Entropy	0.5	37.9	96.67

(a) Sampling Interval: 10s

Method	FP(%)	DD(sampling units)	DR(%)
AR	1.5	1	100
CUSUM	6.84	1	100
Entropy	0.5	1	95

(b) Sampling Interval: 60s

Method	FP(%)	DD(sampling units)	DR(%)
AR	0.9	1	100
CUSUM	6	1	100
Entropy	0.5	1	80

(c) Sampling Interval: 120s

Table 4.4: SYN Attack of 10 to 20 syns/second

simulated for over 50 times with duration of nearly 3 hours. Traffic trace at the victim server is collected and is mixed with normal traffic for sampling intervals of 10s, 60s and 120s across all three data sets to create the attack traffic.

4.3.2 Comparative Study

For a meaningful analysis, the proposed model is compared with two other algorithms, (1) Non-Adaptive mechanism [7] and (2) Adaptive mechanism [5]. These two algorithms represent the two main approaches followed in non-parametric modeling of network traffic at the edge network as discussed in Section 2.1. The non-parametric cumulative sum y in [7] (earlier discussed in Section 2.1) is estimated at every time t by sequentially observing the normalised time series of HOC $(X(n))$ as follows,

$$G(t) = X(n) - a \tag{4.3}$$

$$y_t = \begin{cases} y_{t-1} + G(t) & \text{if } G(t) > 0 \\ 0 & \text{Otherwise} \end{cases} \tag{4.4}$$

Method	FP(%)	DD(sampling units)	DR(%)
AR	2.79	1	100
CUSUM	6.39	1	100
Entropy	0.5	-	0

(a) Sampling Interval: 10s

Method	FP(%)	DD(sampling units)	DR(%)
AR	1.5	1	100
CUSUM	6.84	1	100
Entropy	0.5	1	85

(b) Sampling Interval: 60s

Method	FP(%)	DD(sampling units)	DR(%)
AR	0.9	1	100
CUSUM	6	1	100
Entropy	0.5	1	80

(c) Sampling Interval: 120s

Table 4.5: SYN Attack of 2 to 8 syns/second

If $y_t > N$, then an alarm is raised to notify a possible attack. The values of a and N are fixed as 0.35 and 1.05 respectively, since it is claimed to be independent of the access pattern and time of the day [7]. HOC value is normalised with SYN packet count in each sampling period. The cumulative sum technique is quite popular and is employed in other work reported in literature [12, 67]

Based on the adaptive algorithm proposed in [5], randomness of the HOC series is estimated by computing the entropy over a sliding window, which is shifted 1 sample at a time (overlapping window). A major challenge in terms of practical implementation of this algorithm is to find an appropriate window size for efficient attack detection and at the same time to keep the amount of re-computation to the minimum. A window of size 100 samples is used in this case. The HOC is divided into n bins and the entropy is computed using the formula:

$$\chi(i) = \sum_{j} p(x_j) log(p(x_j)) \tag{4.5}$$

where $p(x_j)$ is the probability of occurrence of an event belonging to the j^{th} bin. The

Method	FP(%)	DD(sampling units)	DR(%)
AR	2.79	24.1	100
CUSUM	6.39	4.7	100
Entropy	0.5	-	0

(a) Sampling Interval: 10s

Method	FP(%)	DD(sampling units)	DR(%)
AR	1.5	35.19	100
CUSUM	6.84	1.6	100
Entropy	0.5	1.2	50

(b) Sampling Interval: 60s

Method	FP(%)	DD(sampling units)	DR(%)
AR	0.9	53.8	100
CUSUM	6	1.6	100
Entropy	0.5	1	50

(c) Sampling Interval: 120s

Table 4.6: SYN Attack of 0.5 to 5 syns/second

Method	10s	60s	120s
AR	44.39 (7.39 min)	64.85 (64.85 min)	146.31 (292.62 min)
CUSUM	17.17 (2.86 min)	15.59 (15.59 min)	16.40 (32.8 min)
Entropy	250.25 (41.70 min)	177.54 (177.54 min)	93.65 (187 min)

Table 4.7: False Alarm time for different Sampling Intervals

intuition is that during an attack, randomness of the series will be less compared to a non-attack period. As discussed in [5], the mean (μ) and standard deviation (σ) of the entropy values are re-computed in each window, in order to adapt to the inherent non-stationarity in the traffic. If the entropy value falls below ($\mu - 3\sigma$), an alarm is raised indicating a possible attack.

Performance of all three models are compared based on four measures: False Positives (FP), False Alarm Time (FAT), which is the mean distance between two consecutive false alarms, Detection Rate (DR) and Detection Delay (DD). Tables 4.4, 4.5 and 4.6 list these measures (except FAT which is listed in Table 4.7 expressed in sampling units as well as in minutes within brackets) for the three techniques across different at-

tack types and sampling intervals. Some of the important observations and inferences are as follows: Firstly, in terms of FP, the AR method performs better than the other two with 100% DR in all the cases. FP is as low as 0.9% for sampling interval of 120s. It may be noted that DD is relatively higher for AR. This need not be an issue given the fact that the attacks are low rate in nature and will not bring down the service immediately. Secondly, CUSUM method also gives 100% DR (with smaller detection delay compared to AR), but with larger FP of around 6%. Higher false alarms can be attributed to the non-deterministic trend of the *HOC* series. Thirdly, the false alarms are wide apart in the case of AR (larger FAT values), which is a desirable feature for any sequential decision technique. Fourthly, entropy method is found to be generally inefficient in detecting *low intensity* attacks with maximum DR of 50% in the case of 0.5 to 5 syns/s. DR is better in other two cases. Fifthly, there is a varying improvement in performance across all three models with increase in sampling interval. This can be attributed to the *smoothing* that happens on increasing the sampling interval. It is shown in [68] that a process can have *self-similar* behaviour (often discussed in the context of *burstiness* nature) and still be *smooth*. The *smoothness* is measured in terms Relative Variance (RV) i.e *variance normalised by the mean* and Mean Squared Difference (MSD) i.e *mean of the square of adjacent values*. The estimated values for RV and MSD, after the time series of various sampling intervals are uniformly scaled (one second in this case), are listed in Table 4.8. Smaller the values, greater is the *smoothness*. It can be seen that *smoothness* increases with increase in sampling interval. This shows that, given discrete data, sampling interval is crucial for the determination of structure in the data that can be modeled. Also, the relatively poor performance of entropy method (which models the trend) compared to AR and CUSUM (which model the stochastic variations) implies that the effect of *low intensity*

attacks are more on the short term dynamics rather than on the long term trend. Further, it is observed that the performance of CUSUM and Entropy method is sensitive to the model coefficients a, N and α. Since the proposed technique models the short term dynamics of the system, it is a strong contender for detecting *low rate* attacks discussed in [14, 15]. By choosing an appropriate feature, which can be a 'fairness measure' in the case of [14] and 'packet drops' for [15], detection of such attacks using the proposed technique seems plausible.

Method	**Relative Variance**	**Mean Squared Difference**
10s	1.6151	0.9073
60s	1.3543	0.4449
120s	1.2412	0.3817

Table 4.8: Measure of Smoothness

4.4 Stability as a Measure for Anomaly Detection

In addition to using forecasting error, which is one of the traditional estimate for anomaly detection in time series analysis, we propose a new estimate called the poles of the transfer function, which measures the stability of the discrete linear system under study, for detecting anomalies in the network traffic. By sequentially tracking one or more roots of the AR polynomial at regular intervals of time, one can identify the normal period and the anomalous period. During an attack, there will be change in the expected mean and variance, which can be captured by estimating the roots over each window of samples.

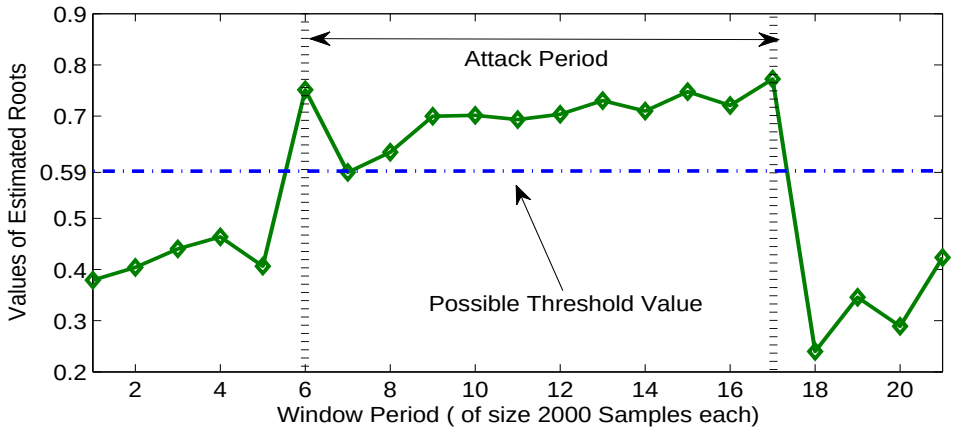


Fig. 4.17: Sequential Behaviour of Roots

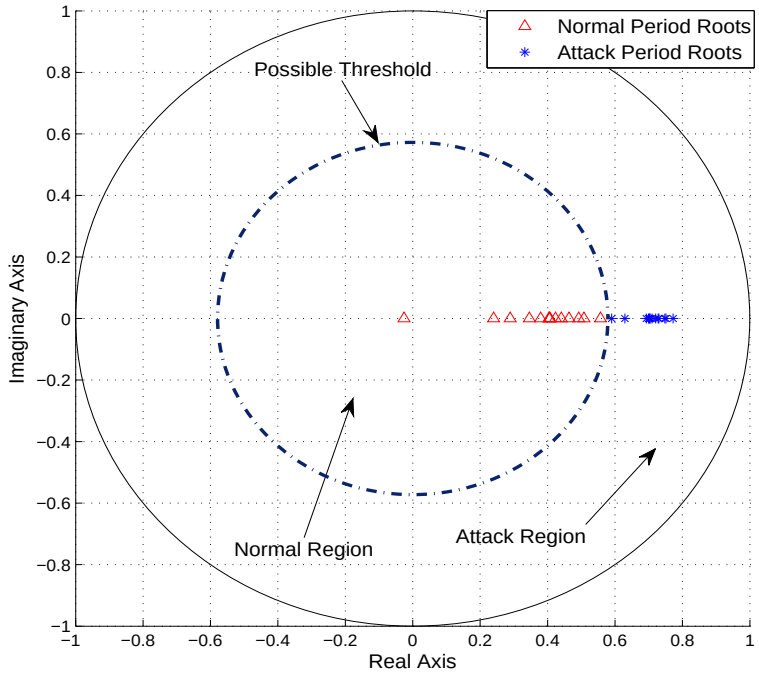


Fig. 4.18: Region of Normal and Attack Period

4.4.1 Proof of Concept

To demonstrate this concept, analyses are conducted on sampling interval of 10s and for attacks with rate varying between 10 to 20 syns/s. The roots of the AR polynomial are computed over a window size of 2000 samples at a time, which is shifted by 1000 samples during the next estimation. The sequential change in the estimated values of the root, during an attack and non-attack period is shown in Figure 4.17. The absolute value of the roots during an attack is relatively higher compared to that of a non-attack period. The spread of the roots during attack and non-attack period in a complex plane is shown in Figure 4.18. Since the attack simulated are not high rate in nature, system does not become unstable (i.e service getting completely disrupted), but shows tendency to move towards the unstable region (i.e efficiency of the service is affected). This may cause some misclassification due to overlap of attack and non-attack region, which can be observed in Figure 4.18, where a number of points belonging to both the regions lie near to the boundary of separation. Further analyses on an efficient classification technique for the roots is left as future work.

4.5 Summary

Analyses discussed in this chapter have shown that it is possible to derive a stationary series from a network time series observed over a long period by an *absolute difference* operation, which can be efficiently modeled using linear time series models. The efficiency of the model in detecting *low intensity* TCP SYN attacks are also studied by comparing its performance with an adaptive and non-adaptive mechanism. Further, a proof of concept is shown to demonstrate the feasibility of using stability as a feature for detecting anomalies in the network traffic.

CHAPTER 5

Conclusion

This chapter summarises the work carried-out in the thesis. Criticism of the work and scope for future work are also discussed.

5.1 Summary

A new insight into the characteristics and parametric modeling of *fine-grained* network time series observed over long period at an edge network, using previously developed methods, is attempted in this work. Similar to the studies reported in econometric time series, network time series is shown to exhibit *structural* breaks apart from the *transient* breaks and can be approximated by a stationary AR model after an appropriate transformation (*absolute differencing*). The analysis is conducted in the context of detecting TCP SYN DoS attack by systematically studying the stationarity/stability of the underlying process. The model proposed is validated using diagnostic tests and by comparing its efficiency with similar algorithms in detecting *low intensity* TCP SYN DoS attacks. Further, it is shown that the stability of the system during normal and attack period can be used as a measure for detection of anomalies in the network traffic. Since the experiments and analyses are carried out on three different data sets, taken across three months, these results scales with time. This means that the model coefficients need not be re-estimated in real time, as is the case with quasi-stationary and non-parametric approaches, and hence more practically suitable.

5.2 Criticism of the work

The biggest draw back of this work is that the detection of TCP SYN attack is studied on simulated attack traffic. Even though the efficiency of the proposed model is validated using comparative studies, detection of real attacks would have added value to the proposed technique. Since, extracting the *half-open count* feature from the traffic trace needs information on the direction of traffic flow (inbound or outbound) and TCP flags of each packet, the present work couldn't make use of the publicly available traffic trace repositories like [69] and [70].

5.3 Scope for Future Work

Following topics are worth exploring as part of future work:

- For a better understanding on the scalability of this study, the work needs to be extended to other traffic traces, not necessarily in the context of TCP SYN attack.
- A detailed study on the behaviour of the *poles* of the transfer function can give more insight in proposing a better classification mechanism for detecting anomalous behaviour using stability measure.

REFERENCES

[1] M. Roughan, A. Greenberg, C. R. Kalmanek, M. P. Rumsewicz, J. Yates, and Y. Zhang, "Experience in measuring Internet backbone traffic variability: Models, metrics, measurements and meaning," in *Proceedings of the International Teletraffic Congress (ITC)*, 2003, pp. 379–388.

[2] K. Papagiannaki, N. Taft, Z. Zhang, and C. Diot, "Long-term forecasting of Internet backbone traffic," *IEEE Transactions on Neural Networks*, vol. 16, no. 5, pp. 1110–1124, Sep 2005.

[3] O. Salem, A. Makke, J. Tajer, and A. Mehaoua, "Flooding Attacks Detection in Traffic of Backbone Networks," in *Local Computer Networks (LCN)*, 2011, pp. 441–449.

[4] B. Zhou, D. He, and Z. Sun, "Traffic predictability based on ARIMA/GARCH model," in *Next Generation Internet (NGI)*, 2006, pp. 200–207.

[5] L. Arshadi and A. H. Jahangir, "Entropy based SYN flooding detection," in *Local Computer Networks (LCN)*, 2011, pp. 139–142.

[6] 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*, vol. 2, no. 1, pp. 1–15, Feb 1994.

[7] H. Wang, D. Zhang, and K. Shin, "SYN-dog: Sniffing SYN flooding sources," in *Proceedings of the 22nd International Conference on Distributed Computing Systems (ICDCS)*, 2002, pp. 421–428.

[8] W. U. Qing-tao and S. Zhi-qing, "Detecting DDoS attacks against web server using time series analysis," *Wuhan Univesity Journal of Natural Sciences*, vol. 11, no. 1, pp. 175–180, 2006.

[9] D. M. Divakaran, H. A. Murthy, and T. A. Gonsalves, "Detection of SYN flooding attacks using linear prediction analysis," in *International Conference on Networks (ICON)*, 2006, pp. 1–6.

[10] G. Zhang, S. Jiang, G. Wei, and Q. Guan, "A prediction-based detection algorithm against distributed denial-of-service attacks," in *Proceedings of the 2009 International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly (IWCMC)*, June 2009, pp. 106–110.

[11] V. A. Siris and F. Papagalou, "Application of anomaly detection algorithms for detecting SYN flooding attacks," in *Global Telecommunications Conference (GLOBECOM)*, 2004, pp. 2050–2054.

[12] H. Liu and M. S. Kim, "Real-time detection of stealthy DDoS attacks using time-series decomposition," in *International Conference on Communications (ICC)*, 2010.

[13] N. Ye, S. Vilbert, and Q. Chen, "Computer intrusion detection through EWMA for autocorrelated and uncorrelated data," *IEEE Transactions on Reliability*, vol. 52, no. 1, pp. 75–82, March 2003.

[14] K. Papagiannaki, N. Taft, Z. Zhang, and C. Diot, "Evaluation of a low-rate DoS attack against iterative servers," *Journal of Computer and Telecommunications Networking*, vol. 51, no. 4, pp. 1013–1030, March 2007.

[15] A. Kuzmanovic and E. W. Knightly, "Low-rate TCP-Targeted denial of service attacks and counter strategies," *IEEE/ACM Transactions on Networking*, vol. 14,

no. 4, pp. 683–696, Aug 2006.

[16] M. Guirguis, A. Bestavros, I. Matta, and Y. Zhang, "Reduction of Quality (RoQ) Attacks on Internet End-Systems," in *International Conference on Computer Communication (INFOCOM)*, vol. 2, 2005, pp. 1362–1372.

[17] B. P. Lathi, *Principles of Linear Systems and Signals*. Oxford University Press, 2009.

[18] M. Guirguis, A. Bestavros, and I. Matta, "Exploiting the transients of adaptation for RoQ attacks on Internet resources," in *International Conference on Network Protocols (ICNP)*, 2004, pp. 184–195.

[19] V. Paxson and S. Floyd, "Wide- Area traffic: The failure of Poisson modeling," *IEEE/ACM Transactions on Networking*, vol. 3, no. 3, pp. 226–244, June 1995.

[20] W. Zhang, Q. Yang, and Y. Geng, "A Survey of Anomaly Detection Methods in Networks," in *International Symposium on Computer Network and Multimedia Technology (CNMT)*, 2009, pp. 1–3.

[21] E. Alpaydin, *Introduction to Machine Learning*. MIT Press, Cambridge, 2004.

[22] V. Paxson, "Bro: A system for detecting network intruders in real-time," *Computer Networks*, vol. 31, pp. 2435–2463, 1999.

[23] M. Roesch, "Snort - Lightweight Intrusion Detection for Networks," in *Proceedings of the 13th USENIX Conference on System Administration (LISA)*, 1999, pp. 229–238.

[24] M. Bellaiche and J. Gregoire, "Avoiding DDoS with active management of backlog queues," in *International Conference on Network and System Security (NSS)*, 2011, pp. 310–315.

[25] P. Yi, Y. Jiang, Y. Zhong, and S. Zhang, "Distributed intrusion detection for mobile ad hoc networks," in *Proceedings of the 2005 Symposium on Applications and the Internet Workshops (SAINTW)*, 2005, pp. 94–97.

[26] D. J. Bernstein and E. Schenk, "Linux kernel syn cookies firewall project," http://cr.yp.to/syncookies.html, last accessed on November 2012.

[27] Check Point Software Technologies Ltd, "Syndefender: http://www.checkpoint.com/press/1996/synattack.html," last accessed on November 2012.

[28] C. L. Schuba, I. V. Krsul, M. G. Kuhn, E. H. Spafford, A. Sundaram, and D. Zamboni, "Analysis of a denial of service attack on TCP," in *Proceedings of IEEE Symposium on Security and Privacy*, May 1997, pp. 208–223.

[29] V. Paxson and M. Allman, "RFC 2988 - Computing TCPs Retransmission Timer," Nov 2000.

[30] A. Lakhina, M. Crovella, and C. Diot, "Diagnosing Network-Wide Traffic Anomalies," in *Proceedings of the 2004 conference on applications, technologies, architectures, and protocols for computer communications (SIGCOMM)*, 2004, pp. 219–230.

[31] L. Huang, X. L. Nguyen, M. Garofalakis, J. M. Hellerstein, M. I. Jordan, A. D. Joseph, and N. Taft, "Communication-Efficient Online Detection of Network-Wide Anomalies," in *International Conference on Computer Communication (INFOCOM)*, 2007, pp. 134–142.

[32] H. Ringberg, A. Soule, J. Rexford, and C. Diot, "Sensitivity of PCA for traffic anomaly detection," in *International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS)*, 2007, pp. 109–120.

[33] J. Rexford and C. Dovrolis, "Future Internet Architecture: clean-slate versus evolutionary research," *Communications of the ACM*, vol. 53, no. 9, pp. 36–40, Sept 2010.

[34] D. Clark, "The design philosophy of the DARPA Internet protocols," in *Symposium proceedings on communications, architectures and protocols (SIGCOMM)*, vol. 18, no. 4, 1988, pp. 106–114.

[35] G. U. Yule, "On a method of investigating periodicities in distributed deries, with special reference to wolfer's sunspot numbers," *Philosophical Transactions of the Royal Society of London*, vol. 226, 1927.

[36] H. Wold, *A Study in the Analysis of Stationary Time Series*. Almqvist & Wiksell, 1938.

[37] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*. Pearson Education, 1994.

[38] A. Pankratz, *Forecasting With Univariate Box-Jenkins Models: Concepts and Cases*. John Wiley & Sons, 1983.

[39] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*. Springer, 2008.

[40] G. Kirchgassner and J. Wolters, *Introduction to Modern Time Series Analysis*. Springer, 2007.

[41] J. Makhoul, "Linear prediction: A tutorial review," in *Proceedings of the IEEE*, vol. 63, no. 4, April 1975, pp. 561–580.

[42] R. F. Engle, "Autoregressive Conditional Heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, vol. 50, no. 4, pp. 987–1007, July 1982.

[43] J. R. M. Hosking, "Fractional differencing," *Biometrika*, vol. 68, pp. 165–176, 1981.

[44] R. F. Engle, "Computer experiments with fractional gaussian noises," *Water Resources Research*, vol. 5, pp. 228–267, 1969.

[45] M. Odaki, "On the invertibility of fractionally differenced ARIMA processes," *Biometrika*, vol. 80, no. 3, pp. 703–709, 1993.

[46] B. B. Mandelbrot and J. W. Van Ness, "Fractional brownian motions, fractional noises and applications," *Society for Industrial and Applied Mathematics (SIAM)*, vol. 10, no. 4, pp. 422–437, Oct 1968.

[47] F. Sowell, "Modeling long-run behaviour with the Fractional ARIMA model," *Journal of Monetary Economics*, vol. 29, pp. 277–302, 1992.

[48] C. Diebolt and V. Guiraud, "A note on long memory time series," *Quality and Quantity*, vol. 39, no. 6, pp. 827–836, 2005.

[49] Y. Shu, Z. Jin, L. Zhang, L. Wang, and O. Yang, "Traffic prediction using FARIMA models," in *International Conference on Communications (ICC)*, June 1999, pp. 891–895.

[50] T. Karagiannis, M. Molle, M. Faloutsos, and A. Broido, "A nonstationary Poisson view of Internet traffic," in *International Conference on Computer Communication (INFOCOM)*, 2004.

[51] C. Diebolt and V. Guiraud, "Long memory time series and fractional integration. a cliometric contribution to French and German economic and social history," *Historical Social Research*, vol. 25, no. 314, pp. 4–22, 2000.

[52] T. Karagiannis, M. Molle, M. Faloutsos, , and A. Broido, "A Nonstationary Poisson View of Internet Traffic," in *International Conference on Computer Commu-*

nication (INFOCOM), 2004.

[53] N. Ranjan, H. A. Murthy, and T. A. Gonsalves, "Detection of SYN flooding attacks using Generalized Autoregressive Conditional Heteroskedasticity (GARCH) modeling technique," in *National Conference on Communications (NCC)*, Jan 2010, pp. 1–5.

[54] The Telecommunications and Computer Networking Group, "http://www.tenet.res.in/," last accessed on November 2012.

[55] TCPDUMP, "http://www.tcpdump.org/," last accessed on November 2012.

[56] C. Chen and G. Tiao, "Random level-shift time series models, ARIMA approximations, and level-shift detection," *Journal of Business and Economic Statistics*, vol. 8, no. 1, pp. 83–97, Jan 1990.

[57] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 8, no. 6, pp. 679–698, 1986.

[58] Z. Sun, D. He, L. Liang, and H. Cruickshank, "Internet QoS and traffic modelling," *IEEE Proceedings on Software*, vol. 151, no. 5, pp. 248–255, Oct 2004.

[59] R. Cont, "Long Range Dependence in financial markets," *Fractals in Engineering*, vol. 4, pp. 159–179, 2005.

[60] C. Granger and Z. Ding, "Some properties of absolute return: An alternative measure of risk," *Annals of Economics and Statistics*, no. 40, pp. 67–91, 1995.

[61] C. Granger and N. Hyung, "Occasional structural breaks and long memory," *Discussion Paper 99-14, University of California, San Diego*, 1999.

[62] T. Karagiannis, M. Molle, and M. Faloutsos, "Long-range dependence ten years of Internet traffic modeling," in *IEEE Internet Computing*, vol. 8, Sept 2004, pp.

57–64.

[63] J. Cheng, J. Yin, C. Wu, B. Zhang, and Y. Liu, "DDoS attack detection method based on linear prediction model," in *International Conference on Emerging Intelligent Computing Technology and Applications (ICIC)*, 2009, pp. 1004–1013.

[64] T. Elteto and S. Molnar, "On the distribution of round-trip delays in TCP/IP networks," in *Local Computer Networks (LCN)*, 1999, pp. 172–181.

[65] A. Downey, "The structural cause of file size distributions," in *International Symposium in Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS)*, 2001, pp. 361–370.

[66] A. Bhattacharjee and S. Nandi, "Statistical analysis of network traffic inter-arrival," in *International Conference Advanced Communication Technology (ICACT)*, 2010, pp. 1052–1057.

[67] S. Wang, Q. Sun, H. Zou, and F. Yang, "Detecting SYN flooding attacks based on traffic prediction," *Security and Communication Networks*, vol. 5, no. 10, pp. 1131–1140, 2012.

[68] M. Roughan and J. Gottlieb, "Large-scale measurement and modeling of backbone internet traffic," in *ITCOM*, 2002.

[69] MAWI working group traffic archive, "http://mawi.wide.ad.jp/mawi/," last accessed on November 2012.

[70] The Internet Traffic Archive, "http://ita.ee.lbl.gov/," last accessed on November 2012.

PUBLICATIONS BASED ON THE THESIS

1 Cyriac James and Hema A. Murthy, "Time Series Models and its Relevance to Modeling TCP SYN Based DoS Attacks," in *7th Euro-NF Conference on Next Generation Internet (NGI)*,Germany, June 2011, pp. 1-8.

2 Angelos Marnerides, Cyriac James, Alberto Schaeffer-Filho, Saad Yunus Sait, Andreas Mauthe, Hema Murthy, "Multi-Level Network Resilience: Traffic Analysis, Anomaly Detection and Simulation," *ICTACT Journal on Communication Technology: Special Issue On Next Generation Wireless Networks and Applications*, vol. 2, no. 2, pp. 345-356, June 2011.

3 Cyriac James and Hema A. Murthy, "Decoupling Non-Stationary and Stationary Components in Long Range Network Time Series in the Context of Anomaly Detection," in *37th Annual IEEE Conference on Local Computer Networks(LCN)*, U.S.A, October 2012, pp. 76-84

GTC MEMBERS

Chair:

Dr. Kamala Krithivasan (HoD's Nominee),

Professor,

Department of Computer Science and Engineering,

IIT Madras, Chennai

Thesis Advisor:

Dr. Hema A. Murthy,

Professor,

Department of Computer Science and Engineering,

IIT Madras, Chennai

Other Members:

Dr. Krishna Moorthy Sivalingam,

Professor,

Department of Computer Science and Engineering,

IIT Madras, Chennai

Dr. Nitin Chandrachoodan,

Associate Professor,

Department of Electrical Engineering,

IIT Madras, Chennai

CURRICULUM VITAE

ACADEMIC BACKGROUND:

- B.Tech in Computer Science and Engineering from Cochin University of Science and Technology, Kerala, 2000-2004
- M.S in Computer Science and Engineering from Indian Institute of Technology Madras, Chennai, 2009-2012 (thesis submitted)

PROFESSIONAL BACKGROUND:

- Scientist/Engineer, ISRO-ISTRAC, Bangalore, from 29-09-2007 to 17-10-2008
- Software Engineer, NeST, Technopark, Trivandrum, from 21-01-2006 to 07-09-2006
- Junior Research Fellow, C-DAC, Trivandrum, from 26-07-2004 to 29-12-2005

CONTACT INFORMATION:

GN-43, Palapparampil,

Gandhiji Nagar,

Nalanchira P.O,

Trivandrum, Kerala

email: cyriac83@yahoo.com