Statistical Analysis of Snort Alerts

O.B.Remi-Omosowon and B.V.Ghita

Centre for Security, Communications and Network Research,
Plymouth University, Plymouth, UK
e-mail: info@cscan.org

Abstract

Intrusion detection systems are used to monitor information systems, creating large number of alerts which are difficult to respond to. Many of these alerts do not present threats as they merely report the normal working condition of the system. These information systems are often used for specific tasks that are repetitive or consistent over the period of its use; hence a pattern is expected for these alerts. A single alert may have no significance by itself but can be part of a bigger threat and Analysing these alerts individually can be very tedious and time consuming. Such threats will alter the normal course of the system's statistics. This paper focuses on the processing of high volumes of alerts generated by snort analysing the trend of the hourly alert intensities triggered at the edge of the Plymouth University network. The analysis is conducted on real world data. The goal of this analysis is to identify the true positive alerts that signify actual intrusion attempts. This paper also presents a way to share the statistic property of alerts with researchers without sharing the actual traffic source or alerts that can be mined for information about an enterprise. The research also reveals that the top 6 alerts contribute over 99.6% of the entire alerts. Security administrators and other researchers will benefit from the findings in this research paper.

Keywords

Intrusion detection, alert analysis, snort alarms, trend analysis, network security, STL

1 Introduction

Intrusion detection is the process of monitoring events on a system and analysing these events for the occurrence of a security incident or malicious activities (Bace, 2000). Intrusion detection systems (IDS) do not simply detect intrusions, but instead detect events in traffic that may or may not be intrusions (Endorf et al. 2003). It creates excessive alerts that trigger on detection of malicious activities. The alerts tend to become cumbersome to analyse and respond to, as a number regular system activities also trigger the alerts (Dong et al. 2008). Consequently it is almost impossible for administrators to use intrusion detection systems appropriately.

The alerts triggered by normal system activities are known as false positives and tend to be repetitive in the alert files. IDS systems aim to filter out the anomalies or specific attacks that have been identified in the past by signatures. IDS research has been evolving to improve its performance by reducing the false acceptance rate and reduce its false rejection rate but these systems are still far from perfect at identifying only intrusions in a system. The alerts generated are, in most cases, elementary and difficult to interpret by security administrators; hence, the need for a specialist to

decipher the content. IDS systems generate large number of alerts which are consists of several false positives, false negatives, and true positive alerts (Cuppens, 2001; Li and Tian, 2010). This leaves the response to these alerts as a job for the busy security administrator who is unable to interpret the cumbersome alerts.

Individually, these alerts may be insignificant, and are sometimes discarded by the IDS operator. The significance of an alert cannot be measured in most cases, as its significance often depends on its past behaviour and presence. Collectively, these alerts can provide more information about the state of the protected system (Viinikka and Debar, 2004). Thus, alerts accumulated over a period of time are invaluable, and can provide information about the nature of intrusion for the purpose of fingerprinting. The alerts can be aggregated together as they occur sequentially, and examined as an hourly, hence reducing the time required to review the alerts. This creates a time series of the alert intensities, and emphasizes the need to store alerts over a long period of time.

The main objective of this research is to identify the false positives and actual alerts in using a variety of statistical analysis tools ranging from trend analysis to other tools such as the box and whisker plot. This relies on the assumption that the false positives will generate a certain amount of alerts always as the normal activities that trigger it are in most cases repetitive and thus, generates a pattern in the trend. This paper involves trend analysis of the alerts originating from different countries generated by snort at the edge of the Plymouth University network using Seasonal decomposition of Time series by Loess (STL).

Due to privacy concerns, it is also difficult to get real world data for researchers to analyse. The source IP addresses in traffic sources can be anonymised but it will still contain a reasonable amount of information that can be extracted to determine the sites frequently visited but users in the enterprise. Most research is aimed at exploiting the statistical properties contained in the alert sets. Thus, extracting these properties removes the need to share the alert files. This paper suggests a framework for making the statistics available to researchers.

2 Statistical Analysis of Alerts

The statistics of these alerts can provide a basis for identifying the characteristics of the normal system activity. Spathoulas and Katsikas(2010) shows that false alerts can be identified by the frequency with which the alert signatures trigger false positives. Thus, changes in the trend of the alert intensity will also reveal the actual intrusion attempts from the alert sets. There is no single solution to distinguishing the actual intrusion attempts from the normal system activities, but numerous statistical procedures have been used in the past.

Algorithms including REDUCE, which determines the periodicity of the normal system events using Fourier analysis of time series of the alerts, and CLUSTER have been used in some research to correlate the alerts correlation (Julisch, 2003; Viinikka et al. 2006; Dong et al. 2008). The CLUSTER algorithm groups together similar alerts based on the attack patterns. Clifton and Gengo (2000) characterises false

alarms using the frequent episodes algorithm by identifying the alarms sequential alarms that occur frequently over a period of time.

Ye et al. (2002) suggests procedures for obtaining efficient results using EWMA control charts. Ye et al. (2003) actualises two of these procedures and tests the performance using EWMA. Viinikka et al.(2004) shows that EWMA does not provide successful results at all times. Stationary autoregressive models (AR) is seen to generate a more detailed result in Viinikka et al. (2006) but with the necessity for removing the trend and periodic components which is tends to introduce artifacts in the series. For this reason, Viinikka et al. (2009) proposes the continuous use of EWMA, instead of the stationary AR model. Debar and Wespi, (2001) propose an algorithm that can be used in aggregation and correlation components of intrusion detection systems. Chantawut(2009) also uses time series approach using autoregressive integrated moving average (ARIMA) method.

The normal system alerts are known to have a periodicity, and hence the time series approach is widely used. Trend analysis of the alerts over a period of time can reveal the behaviour and pattern of false positive alerts over a period of time. The normal state of the system can be observed, and changes to this can be easily spotted. This paper involves the trend analysis using STL to decompose the alert. Unlike AR, STL does not introduce artifacts (Cleveland et al. 1990) and overcomes the shortfall of AR identified by Viinikka et al. (2009).

3 Trend Analysis

A time series is a sequence of successive observations which are ordered in time. It is simply a set of observations (in this case the alert intensity) each one being recorded at a specific time (hourly). Univariate time series analysis, as the name implies involves analysing the collection of observations for a single variable over a period of time, whereas a multivariate Time series involves doing the same for multiple variables. This paper will focus on univariate analysis of the alert intensities observed. A univariate time series, z_t, usually takes the form in equation 1.

$$z_t = f(z_t, z_{t-1}, \ldots, z_1) + a_t$$

Equation 1: Univariate Time series (src: Pena et al. 2001)

Where $f(z_t, z_{t-1}, \ldots, z_1)$ a function of is previous values of the series, and a_t is a sequence of identically distributed variables. It is sometimes written in the form as seen in equation 2.

$$z_t = m_t + s_1 + y_t$$

Equation 2: Univariate Time series (src: Brockwell and Davies, 2001)

where m_t is the trend of the series, s_t is the periodic seasonal component of the series, and y_t is the residual noise component. Time series is usually written as a mathematical model to make it easy to decompose it into its systematic component and noise component. This property of time series makes it suitable for the correlation of alerts.

This paper focuses on the use of STL for decomposing the time series; although STL does not generate a mathematical model. It uses localised models at each data point combining the simplicity of the linear square regression and the flexibility of the non-linear regression (NIST, 2010). This simplicity allows analysis of the properties of the procedure and allows fast computation, even for very long time series and large amounts of trend and seasonal smoothing. This removes the need for determining an appropriate model for each series that is plotted, thus no expertise required in understanding statistical procedures. This makes this a suitable model for implementation by a security administrator.

4 Program Framework

Numerous researches are based on the alert intensity per unit of time, or alert flows based on the IP addresses. Enterprises have to maintain the privacy of its users when sharing information; thus the real world data that is sometimes made available is anonymised to avoid reconstruction of the packets. The payload of the packets is in most cases stripped, and most of the header and trailer information as well. This still leaves the concern of enterprises with concerns of data leakage and data mining. It is often a question of how much information can be distributed that will not present information about the network users.

Most organisations archive the alerts after a long time for legal reasons, or intrusion detection. This practice can provide useful the research community, since the alerts are needed over a long duration. The alert intensities can be computed hourly or daily for individual alert types over different durations before the long-term organisational archiving and backup. IP-to-Country APIs can be used to determine the destination countries to allow for correlation of the alerts based on the source countries. An example is the free Geolocation API available from Maxmind that has an accuracy of 95% for most countries. The frame work will generate index files for the countries and alert types to help researchers identify the alerts and also facilitate a time of day correlation from the respective countries.

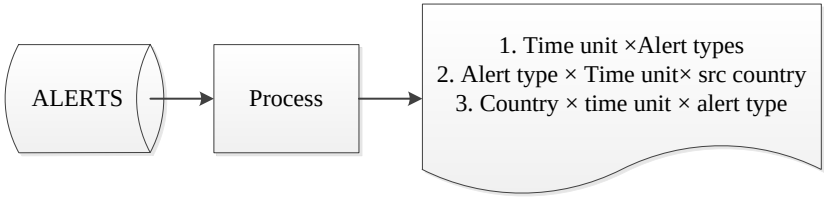

Figure 1: Program framework for extracting statistics of alerts

This frame work proposes that 3 types of files, excluding the program with the index of countries and alert types, are needed. The first stores the alert intensity within each time unit (hourly and daily) for different alert types. Next, different files for each alert type present containing the intensities of the alerts during each time unit from all source countries. The third type of file will include files for each source country of the alert containing the alert intensity of different alert types for at every time unit.

5 Experimental Results

Viinikka et al. (2006) concluded that few signatures account for most of the alerts generated by intrusion detection systems. Basic analysis of the alert set shows that 6 of the top alerts contribute up to 99.6% of the entire alerts in the set. Table 1.0 shows the alert distribution for 2 different datasets obtained from the Plymouth University's network.

Dataset 1			**Dataset 2**		
Signature ID	**Intensity**	**Cum %**	**Sig. ID**	**Intensity**	**Cum %**
Slammer (2050, 2003, 2004)	50057355,50057346 50057346	95.2	2	331422	60.3
ICMP (469,466, 483)	4904803, 1830953 322414	99.7	7	115166	81.3
MS-SQL overflow (2329)	116738	99.7	14	47948	90.0
2	63623	99.8	15	39338	97.2
58	50601	99.8	6	7862	98.6
7	43793	99.8	9	6718	99.8
1419	37321	99.9	4	558	99.9
255	30176	99.9	3	258	100.0
472	24685	99.9	2464	8	100.0
All alerts	157747024	100		549288	100.0

Table 1: Signature ID Present in Dataset in order of intensity

Analysis of HTTP Double decoding alerts (SID2) originating from United Kingdom

Snort alerts with ID value of 2 represents HTTP double decoding alerts, invalid FTP commands, Back Orifice client detected, Tear drop attacks, as well as encrypted Telnet sessions. This alert type is present in both datasets and originates from various countries. In 2009, we have United Kingdom, China, United States, Taiwan, and Netherland contributing the following percentages of the entire SID 2 alerts: 75.1, 9.5, 5.4, 0.7 and 0.6 respectively. In 2010, United Kingdom appears as the source of just 60.2% of the alerts while unspecified countries in Europe originate 26%. Italy, United States and Sweden have 11.9%, 1.3% and 0.4 % respectively. There are few alerts from China in 2010 contributing less than 0.01% as compared to the results from 2009. This shows that the trend of the alerts on the internet is changing.

It is also observed that the slammer and ICMP alerts which were massively present in 2009 are no longer present in the dataset captured in 2010. The signature rule-sets used for the experiment were the latest available at the time of the research procedure. It was noticed that just 1 of the 3 slammer alerts was enabled but previous study by Chantawut(2009) showed that the 3 slammer alerts all triggered on the same packets, hence the modifications to the rule set removes two-third of the slammer

occurrence if present. We can conclude that snort has been tuned appropriately to trigger the slammer alarms when it detects a more significant occurrence of the worm, or it could also be that the worm is being blocked at a higher layer in the network topology as a result of on-going research to block and remove traces the worm on the internet.

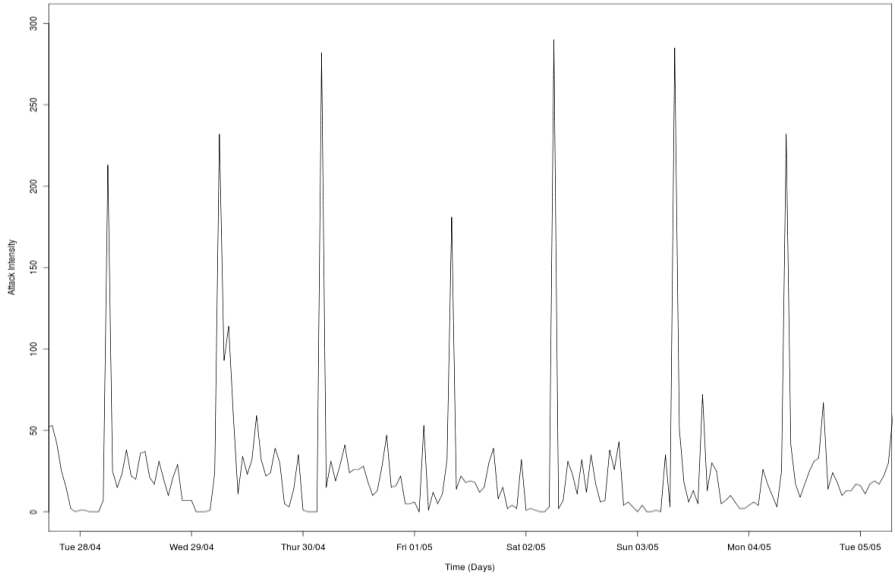

Figure 2: Time series plot of SID2 from United Kingdom from Fri Apr 24 to May 5

The time series plot of the hourly alert intensity reveals a characteristic behaviour for the normal system activities on the network. A pattern can be observed in figure 2 with varying peaks at different times of the day. This depicts the regular system activity for the first week in the dataset. The time of the level shift differs each day, but the pattern continues for the first week. The pattern alternates randomly on different days, with 6 different levels each day at different hours but the intensity remains on the same level. A large increase in the trend can be seen in figure 3.This suggests that the alerts observed on the consistent alert patterns observed are false positives as the alerts are repetitive. The obvious true positive alert here is the event signified by the sharp rise on 20 July 2009 between 3AM and 5AM. The trend of the series also reveals that the intensity varies at each hour of the day.

The box and whisker plot of the daily intensity of the alert in figure 3 also shows that the alerts on the 92 out of 96 days in the alert set have similar values of intensity. It can be observed that the data has a normal distribution as the median lies in the middle of the box

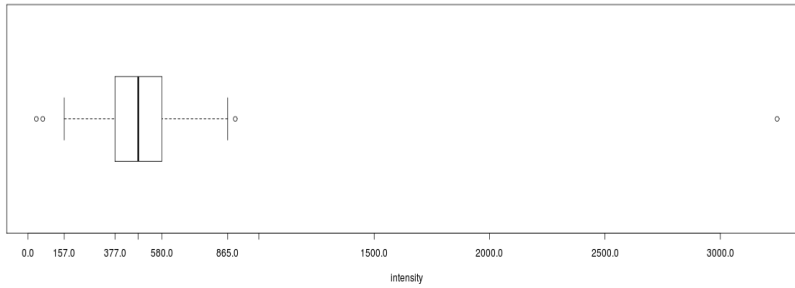

Figure 3: Boxplot of SID2 alerts from United Kingdom

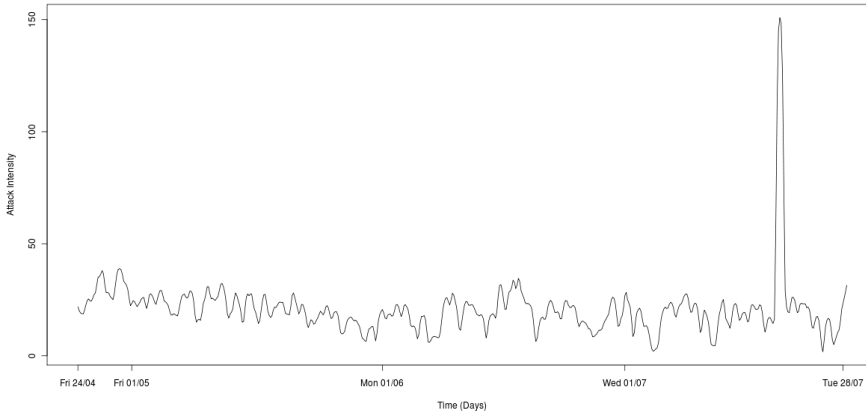

Figure 4: Trend analysis of SID2 from United Kingdom

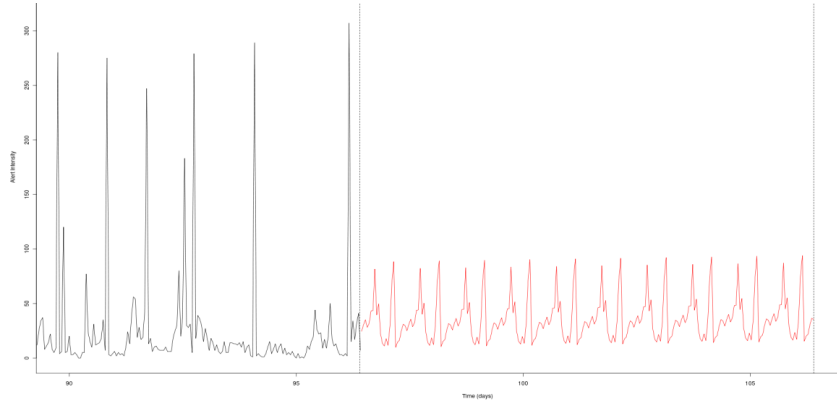

Figure 5: Forecasting 10 additional days of SID2 using the operator module in R

6 Conclusions and Future

This paper presents various statistical analyses of alerts triggered by snort. It shows how false positives can be identified from the trend of the series. This paper shows that the trend can be used to determine levels of false positives in the alert set, periodicity, as well as identify malicious activities. The trend of the series can provide information to support Analysis also showed the lack of ICMP and Slammer alerts in 2010, which may be due to improvements to snort in the last year.

STL is able to decompose the series into the respective components, but it is unable to generate a suitable forecast of the trend. The AR model or ARIMA models should be used for subsequent analysis to allow for significant prediction of the trend.

The program framework can be implemented and combined for use with the major IDS solutions in use in the Industry. This will provide a way for more research to be done and improve the quality of IDS systems faster than the current evolution trend.

7 References

Bace, R. G. (2000), Intrusion detection, Macmillan Publishing Co., Inc.

Chantawut, K., (2009), "Trend Analysis of Snort Alarms", MSc Thesis, Plymouth University.

Clifton C, Gengo G.(2000), Developing custom intrusion detection filters using data mining, 21st century Military Communications MILCOM 2000, vol. 1; 2000. p. 440–3.

Cleveland, R., Cleveland, W., Mcrae, J. & Terpenning, I. 1990. STL: A Seasonal-Trend Decomposition Procedure Based on Loess, *Journal of Official Statistics*, 6, 3-73,

Cuppens, F. (2001), Managing alerts in a multi-intrusion detection environment, In: Computer Security Applications Conference, 2001, Proceedings 17th Annual, 2001. 22-31

Debar, H. & Wespi, A. (2001), Aggregation and Correlation of Intrusion-Detection Alerts, In: Lee, W., Mé, L. and Wespi, A. (eds.) Recent Advances in Intrusion Detection, Springer Berlin / Heidelberg.

Dong, L., Zhitang, L. & Jie, M. (2008), "Processing Intrusion Detection Alerts in Large-scale Network", In: Electronic Commerce and Security, 2008 International Symposium on, 3-5 Aug. 2008 2008. 545-548

Endorf, C., Schultz, E., and Mellander, J., (2004), Intrusion Detection & Prevention, McGraw-Hill, USA

Julisch, K. (2003), Clustering intrusion detection alarms to support root cause analysis, ACM Transactions on Information and System Security, 6, 443-471.

Li, W. and Tian, S. (2010), An ontology-based intrusion alerts correlation system, Expert Systems with Applications, 37, 7138-7146.

Spathoulas, G. P. and Katsikas, S. K. (2010), Reducing false positives in intrusion detection systems, Computers and Security, 29, 35-44.

Viinikka, J. and Debar, H. (2004), Monitoring IDS Background Noise Using EWMA Control Charts and Alert Information, In: JONSSON, E., VALDES, A. & ALMGREN, M. (eds.), Recent Advances in Intrusion Detection, Springer Berlin / Heidelberg.

Viinikka, J., Debar, H., Mé, L. and SéGuier, R., (2006), Time series modeling for IDS alert management, Proceedings of the 2006 ACM Symposium on Information, computer and communications security. Taipei, Taiwan: ACM.

Viinikka, J., Debar, H., Mé, L., Lehikoinen, A. and Tarvainen, M, (2009), Processing intrusion detection alert aggregates with time series modeling, Inf. Fusion, 10, 312-324.

Ye, N., Borror, C., and Zhang, Y., (2002), EWMA techniques for computer intrusion detection through anomalous changes in event intensity, Quality and Reliability Engineering International, 18, 443-451.

Ye, N., Vilbert, S. and Qiang, C. (2003), Computer intrusion detection through EWMA for autocorrelated and uncorrelated data, Reliability, IEEE Transactions on, 52, 75-82.