Statistical analysis of Snort alarms for a medium-sized network

Kitti Chantawut [chantawut@gmail.com], Bogdan Ghita [bghita@plymouth.ac.uk]
Centre for Security, Communications, and Network Research
University of Plymouth, UK

Abstract

Statistical analysis of network intrusions has been an active topic for researches for many years. However, due to the complexity and security concerns associated with the Internet, this area of research remains challenging, from the monitored networks and methodology used to the focus of the analysis and presentation of the results. This paper aims to provide additional insight into this area by analysing a set of IDS alarms collected over a period of three months from the external interface of the edge router at the University of Plymouth. The motivation of this study is to quantitatively classify and understand the nature of current Internet threats, as observed at a medium stub network, leading to long-term analysis of trends and recurring patterns of attacks. In the study, fundamental features of intrusions activities are investigated through a number of characteristics, from the daily volume of intrusion attempts to the source/destination of the intrusion attempts as well as the specific attack type. The results of the study show high levels and wide variety of intrusion attempts. It also shows that the attacks reflect daily timescales and the on/off patterns exhibit recurrence of correlated behaviours. Furthermore, the Slammer worm appears to feature on the Internet long after its original release. Deeper investigation reveals that the sources of attacks spread uniformly, apart from a large proportion of intrusions generated by a small number of IP addresses located in China.

Keywords

Trend analysis, Intrusion detection system, Snort, Slammer

1. Introduction

Prevention is universally recognized as the best strategy to protect critical information infrastructure from malicious attacks. As part of their duties, security staff and network administrators must understand the trend of the threats in order to be able to prevent any significant damage to their networks. Understanding the trend of the attacks would help organisations to determine the current fitness of their security systems and the budget for improving the system to defense against cyber attacks due to the uncertainty of the occurrence of the attacks.

Currently there are a number of tools that can be use to analyse intrusion data, such as BASE (Basic Analysis of Security Engine) or SnortSnarf. These tools provide querying and presenting the intrusion analysis in easy to use graphical mode, but they offer only a basic set of analytic options to users. For instance, they do not provide the support to perform in depth statistical analysis from geolocation mapping the source of attacks to forecasting the attack trends in the future. The aim of this study is specifically to conduct such analysis on the set of collected alarms.

The methodology used included an online alarm collection stage, where snort was run on the monitoring machine, followed by offline analysis of the alarms to gain an understanding of the behaviour of the threats. The analysis included statistical profiling of the alarms, such as the source countries, number of attacks per unique IP, the distribution of attack on the targets, and profiling the nature of some of the threats using normal statistic and time series analysis theory.

2. IDS and alarm profiling

2.1. Network based Intrusion Detection System (NIDS)

NIDS monitors the traffic in specific network segment or subnet. NIDS looks for anomalies in the traffic, such as port scanning or denial of service attacks. In order for NIDS to be effective, it has to be located where it can monitor the most traffic that an organization deems critical. Therefore, placement is critical to the success of uncovering of anomalous traffic or behaviour in the monitored area.

There are, typically, two types of detection mechanisms using by NIDS which are “signatures (or rules) based detection” and “anomaly based detection”. In signature based detection, the NIDS look in bytes codes and expressions to match any known attacks expressions (signatures or rules). When it matches any intrusion, flags an alarm. Signature based detection is useful for detect known threats but it cannot detect new unknown threats or even the variants of already defined attacks.

In anomaly based detection the NIDS first establishes a normal activity model (a baseline) after training the system for specific period of time. Then the NIDS will use the baseline model to detect suspicious events that deviate from normal activity model. If an anomaly is detected, the will flag alerts for an intrusion. The benefit of anomaly detection is that it

can detect unknown threats without having to understand the cause of the threats. The main problem for this approach is that it is prone to false alarms.

2.2. Time Series Analysis

In general, the analysis of a time series will be based on the fact that observations close together in time will be more closely related than observations further apart and values in a series for a given time will be expressed as deriving in some way from past values, rather than from future values.

2.2.1. Stationary Time Series

In stationary time series, the random variables fluctuate about a fixed mean level, with constant variance over the observational period. This property is a requirement for time series analysis because there must be a sort of regularity exists in the time series so that the behaviour of the time series can be predicted. Various levels of stationarity exist; however, in a context of univariate time series, the time series must satisfy the assumption of **"**weakly stationary"**,** that the mean is a constant, independent of any time shift.

2.2.2. Autocorrelation (ACF)

ACF is a statistical measure that captures the correlation between different time shift samples (or lag) of the process. (NIST/SEMATECH, 2006) has summarised the main purposes of ACF into two points. The first purpose is to detect the non-randomness in time series and the other is to identify an appropriate time series model if the data are not random.

2.2.3. Long Range Dependency (LRD) and Self Similarity (SS)

A stationary process is said to exhibit Long Range Dependency if it has a high degree of correlation between distantly separated data points, even across large time shifts. Whereas in short range dependence processes, the correlation between values at different times decreases rapidly as the time difference (lag) increases. Self Similarity is a property of an object whose appearance remains unchanged regardless of scale of which it is viewed. Self similarity detected in the intrusion data could explain certain characteristics and behaviours of a specific intrusion attempt. It is also useful to note that some self-similar processes may exhibit LRD, but not all processes have LRD are self-similar. The degree of SS and LRD can be estimated by the calculation of the Hurst parameter H. For a self-similar process with LRD, the value of H will be in the range of $0.5 < H < 1$. As $H \rightarrow 1$, the degree of both self-similar and LRD increases. Any pure random processes would have $H = 0.5$. The rescale adjusted range statistics (R/S) method can be used to evaluate H. R/S states that, for a self similar data set, the rescaled range or *R/S* statistic grows in proportion to a power law with exponent *H* as a function of and the time-aggregation block size (*m*). As a result, a log-log plot of *R/S* against *m* has a gradient equal to the estimate value of *H*.

2.3. Related Studies

The trend analysis of network attacks is an important subject in IDS. A number of studies focused on the field of statistical analysis of Internet intrusion trends. Many of the studies are based on packet level investigation of intrusion logs generated by either firewalls or IDS.

On worldwide scale, a number of well known projects were set up to collect large scale attack data from firewalls and intrusion detection logs. Two of the most recognised such initiatives are the SANS Internet Strom Center and Symantec's Deep Sight system. On the research side, (Moore, 2004) provides the analysis that gives a better understanding of the nature and the long term trend and recurring patterns of the denial-of-service (DoS) attacks on the Internet. The study concludes that DoS attacks are a common threat for those depending on the Internet. Similarly, (Yegneswaran, 2003) investigates fundamental features of intrusions activities by evaluating the log data along a number of dimensions such as the daily volume of intrusion attempts, the source and destination of the intrusion attempts, and specific types of intrusion attempts.

On the trend identification area, the studies of (Jouni, 2009; Kim, 2007; Wu, 2005) focus on finding best-fit forecasting model for anomaly detection analysis, more specifically representing the dynamics of intrusion and attack activities including the comparison of the model accuracy.

3. Examining Snort Alarms

3.1. Data Collection

The data for this study was collected from the Internet gateway of the University of Plymouth network. The collection used a SPAN port of a core switch; the mirrored port was located in front of the university firewall, allowing traffic capture prior to filtering. This ensured that the overall behaviour of Internet attacks on internal network can be studied; the purpose of the study was not to test the efficiency of the university firewall, but to observe the amounts, structure, and evolution in time of the alarms.

Capturing network traffic on high speed network requires very large storage. Therefore, in order to overcome this problem, Snort, running in packet logger mode, was used to capture the network traffic. The filter was set to capture only the traffic destined to the internal subnet of the University of Plymouth (UoP) network. This limited somehow the scope of the analysis, not allowing observation of packet exchanges following the attack, nor any possible attacks mounted by hosts within the university network. The traffic traces were collected in 2009, between 24th of April and 28th of July. Due to user privacy and network security concerns, the destination IP addresses of traces (the IP addresses within the university) were anonymised using a prefix-preserving technique prior to the analysis, so that the traces would still contain the structure and distribution of local IP addresses. The anonymisation tool used (AAPI) and the associated procedure can be found in (Koukis, 2006; Foukarakis, 2007).
Then the next step was the analysis the anonymised traces. For this study, Snort (version 2.8.4.1 and VRT Certified Rules released on 21/07/09) was run in off-line mode to extract the detailed alerts in the captured traffic traces using signature based detection method. As the large amount of traffic traces were analysed, the unified alert record format was set as the output in the snort.conf because of the small size and the completeness of details contained in the output. Subsequent to successfully creating the complete unified alerts, Barnyard was used to convert all the unified alert files generated by snort into a single output file in *csv* format.

3.2. Distribution of Alerts

A total of 71 types of alerts were detected, distinguished by unique Snort's signature ID (SID), detected by Snort which produced a total of 157747024 alerts. Only 3 types of alerts, which are triggered by Snort Signature IDs (SID) 2003, 2004, 2050 represent almost 95% of all the attacks. All the three alerts are triggered by the MS-SQL (Slammer) worm attacks; the 2003 and 2004 SIDs have the same trigger (only differ by the review code), while 2050 differs slightly but appears to register the same packets.

Name/SID	Number of events
Slammer (2003)	50057346
ICMP NMAP(469)	4904803
ICMP L3(466)	1830953
ICMP CyberKit(483)	322414
MS-SQL probe (2329)	116738
Double decoding (2)	63623
Experimental TCP (58)	50601
IIS Unicode/FTP encrypt (7)	43793
SNMP trap (1419)	37321
ICPMP fragment (255)	30176
ICMP redirect (472)	24685
TCP offset<5 (46)	22827
Truncate TCP options (55)	20396
TCP port 0(524)	20005
Other* (55 alarms)	86642
Total* (filtered)	57632323

Table 1 - Distribution of alerts in the dataset

Table 1 lists the main contributors to the dataset. The Total and Other entries in the table do not include the 2004 and 2050 SIDs which were matching very closely the figure for the 2003 SID. After removing the duplication, by removing the 2004 and 2050 SIDs, the total number of alarms went down to 57632323, of which 50057346 (86.9% of the alarms) were due to Slammer. The next type of alarm to dominate the dataset is the class of ICMP scanning attacks: ICMP NMAP PING (SID469), ICMP L3retriver (SID466) and ICMP PING CyberKit (SID463). Based on these preliminary statistics, the trend analysis focused on Slammer (SID2003) and ICMP scanning (SID469) scanning, given their contribution to the dataset.

3.3. Analysis of Slammer

The analysis of Slammer attacks showed that the attacks came from 125 countries. While the geographical distribution appeared to be uniform, almost 95% of the alarms were from five countries: China (88.6%), Korea (1.8%), United States (1.6%), Romania (1.5%) and Mexico (0.8%). Interestingly, there was no significant correlation between the ranking of the total number of alarms and the number of source hosts, as it is shown in Table 2, but China remains the highest contributor, with 25181 hosts out of a total of 56019 attackers.

Source Country	Slammer alerts (%)	Number of Hosts
China	88.6	25181
Korea	1.8	43
United States	1.6	3133
Romania	1.5	339
Mexico	0.8	127

Table 2 - Distribution of Slammer alerts and number of sources

3.3.1. Alerts per source IP address

The 50 millions Slammer attacks were created by a total of 56019 unique IP addresses. The cumulative distribution function (CDF) plot of the number of attacks during the observation period per unique IP is shown in Figure 1

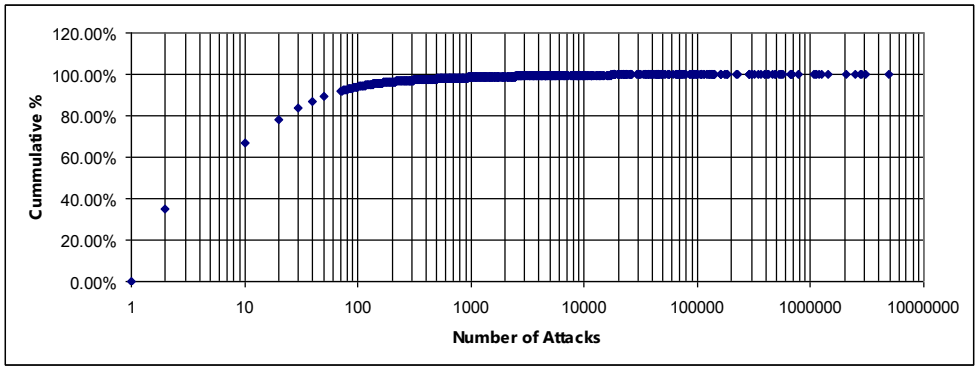

Figure 1 - CDF plot of the Number of Attacks per IP of Slammer

As shown by the CDF, most of the hosts (90%, equivalent to approximately 50000 hosts) sent out attacks less than 50 times, 232 hosts sent out the attacks more than 10000 times, 66 hosts sent out attacks more than 100000 times and a number of 12 hosts (all of them from China) created each more than 1000000 alerts.

3.3.2. The Targets

In order to determine whether any specific hosts were targeted by the Slammer attacks, statistical analysis was applied to the number of alarms per IP address. The target systems consisted of 65536 IP addresses which covered the whole class B of the University of Plymouth's allocated space. One of the typical characteristics of the Slammer worms is that, once the host is infected, the worm will generate random IP addresses and send itself to the random targets. This means that the Slammer worms had been broadcasted all over the IP range with no specific target. Hence, the expected amount of number of attacks on each target IP would be random.

To confirm the above statement, the histogram and the probability density function plot (PDF) of the number of attacks per target IP address will be examined to verify the randomness in the numbers. From the dataset, the mean is equal to 763.81 and the standard deviation is 58.74. The following figures show the histogram of the number of attacks per IP and PDF plots of the number of attacks compare with the normal distribution plot using the mean and standard deviation of the samples.

As it can be seen, most of the targets experienced approximately 760 attacks, with a small number of targets experiencing smaller or higher number of attacks. The PDF plot of the number of attacks is bell-shaped with a peak at the centre of the curve. Comparing this PDF plot against the normal distribution curve with the same mean and standard deviation, it can be seen that the PDF of the dataset is almost perfectly fitted to the normal curve. Therefore, it can be concluded that the Slammer attacks on targets were truly random and automated processes.

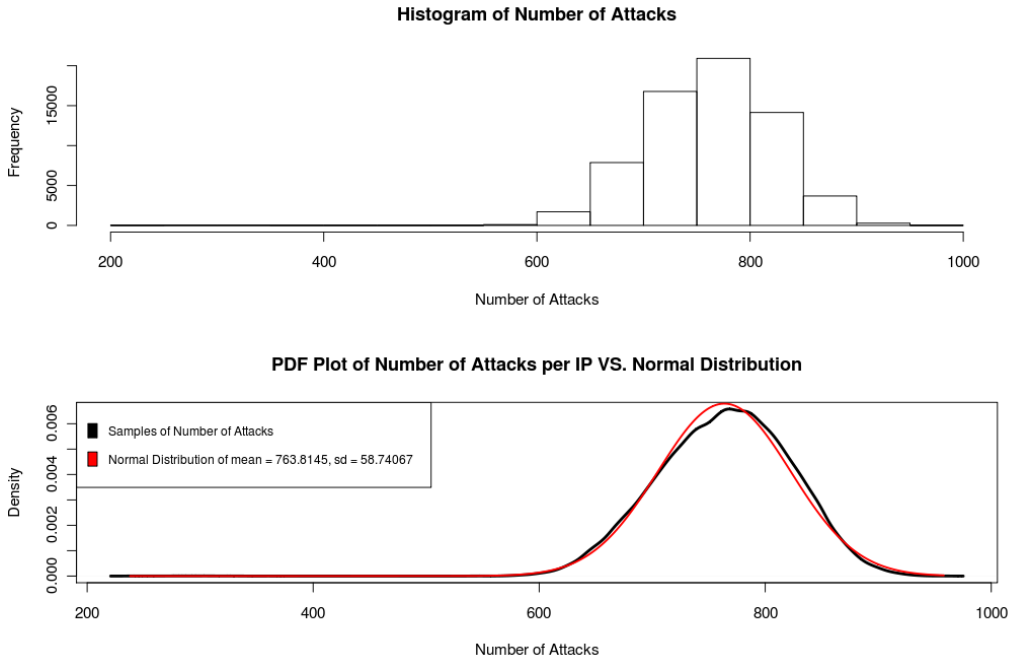

Figure 2 - Histogram and PDF Plots of Number of Attacks

3.3.3. Pattern of attacks

Aside from the total figures, the analysis also investigated the pattern of attack from high contributing countries belonging to different time zones on the university's network. The aim of this analysis was is to determine whether attacks from the same geographical area exhibit different periodicity patterns, determined by the local time of the source hosts.

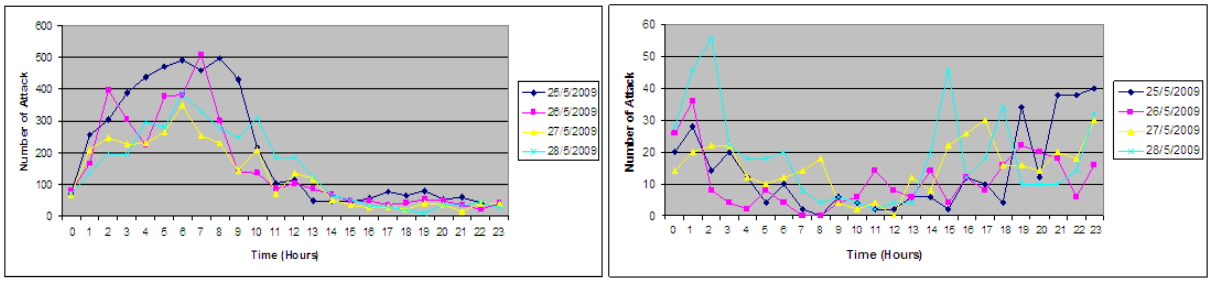

Figure 3 - Numbers of attacks against Hours of a day from China and the USA

Data of the hosts from two different countries (US and China)/continents were analysed. The number of attacks from China and the number of attacks from the USA on two consecutive days (25/05/09 and 26/05/09) were extracted from the log file and plotted on hourly basis as shown in Figure 3.The hourly plot exhibits pattern in every 24 hours. The plots of the USA data show less clear trend of periodic pattern due to the small number of attacks per hour. However, they both still show a periodic pattern with peaks at about 9 am to 10 am and 4 pm until midnight and a dip between 10 am to 4 pm, whereas, the plots from Chinese IP addresses data clearly show peaks during midnight to 10 am and low from 10 am until 4:30 pm.

As mentioned earlier, Slammer infected hosts would broadcasting the code automatically as fast as they can at all time using single UDP packets as the mean of transport. Therefore, the worm is able to broadcast scans without requiring responses from the potential targets. From the aggressive nature of the UDP transport mechanism, the propagation of the worm would interfere with its own growth because there is no congestion control built in the protocol. As a result, each of the infected hosts must compete with each other for limited Internet access bandwidth. Hence, these competitions for available bandwidth may contribute to the diminishing of the entire growth rate of Slammer, issue also identified in (Moore, 2003).The apparent reduction of the number of attacks detected during office hours might also have been caused by the local Internet traffic congestion as well as packet sniffer (the sensor) missing dropping some of the traffic. Unfortunately, no statistics were collected regarding the performance of snort or the number of dropped packets (anywhere between the mirrored port on the core switch and the TCP/IP stack on the monitoring host) throughout the collection period. It is acknowledged that, in order to obtain more precise view of the attack pattern, a network tap could be implemented to connect the sensor at the network gateway instead of connecting the sensor to the mirrored port.

3.3.4. Time Series Analysis

Time series is a sequence of data where the same measurement is taken at regular time intervals. The series can be analysed using various methods, such as autocorrelation function (ACF) analysis, time series modelling and forecasting. There are two main goals of time series analysis. Firstly, it used is to identify relationship of values between different points in time and formalise (model) the dynamic behaviour observed in time series data. Secondly, once the time series model has been derived, the model is used to predict the future events of the time series based on the idea that the past behaviour has some correlation (relationship) with the future behaviour and can be used to estimate (or predict) the future values of the time series. This is known as time series forecasting. The time series used for this analysis is the measurement of the aggregate number of Slammer attacks (SID2003) on the university's network sampling at regular interval of 1 hour.

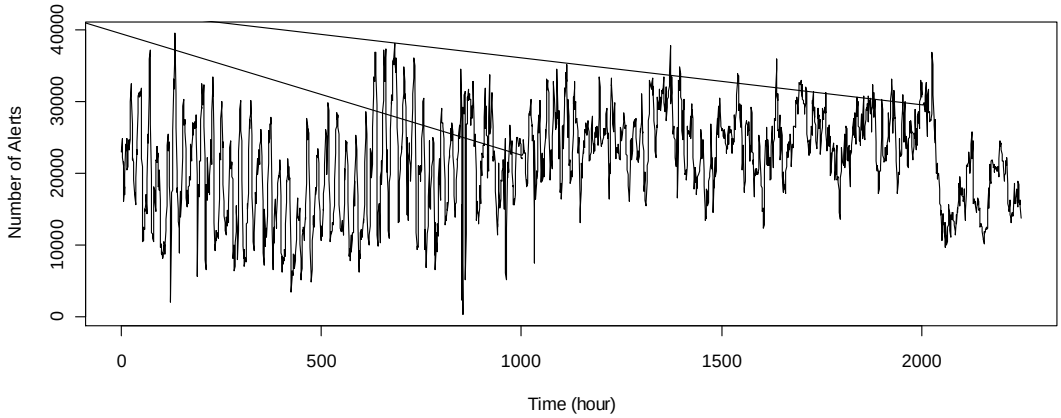

Figure 4 -Time Series of Slammer attack

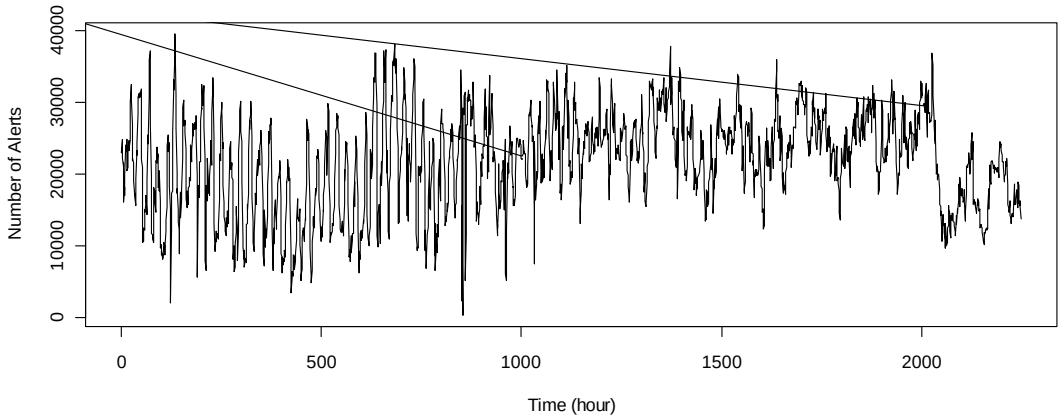

Figure 4 shows a time series of number of Slammer that occurred per hour during observation period. With the first glance on the time series plot, it could be noticed that the time series is not stationary as its means and variance are not constant over time. The sample's Auto-Correlation Function (ACF) and Partial Auto-Correlation (PACF) plots of the time series are investigated to confirm the non-stationary property of the time series as shown in Figure 5.

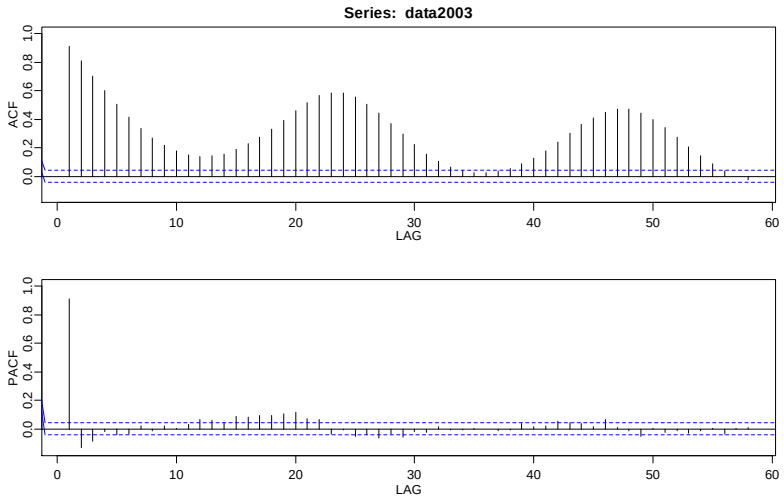

Figure 5 - ACF and PACF of Slammer

It can be seen that the ACF plot shows repetitive patterns at lag 24, 48, etc. with relatively slow decay of its peaks. This confirms the findings above, showing strong correlation and nonstationary character of the time series (seen in the slow decaying of the ACF and the pattern repeating itself every 24 lags/hours).

3.3.5. Self Similarity (SS) and Long Range Dependency (LRD)

Slow decaying of variance (hence, autocorrelation) is a property of a self similarity process. From Figure 5, it can be seen that the ACF plot of Slammer time series shows very slow decay of autocorrelation. This is why this time series is needed be tested to see if there is any sign of any self similarity (SS) with long range dependence (LRD). The R/S method was applied to the dataset to estimate the value of H. Figure 6 shows the plot generated by the R/S method. The R/S method gives the estimation of $H = 0.8011294$, which indicates strong degree of SS and LRD in the time series.

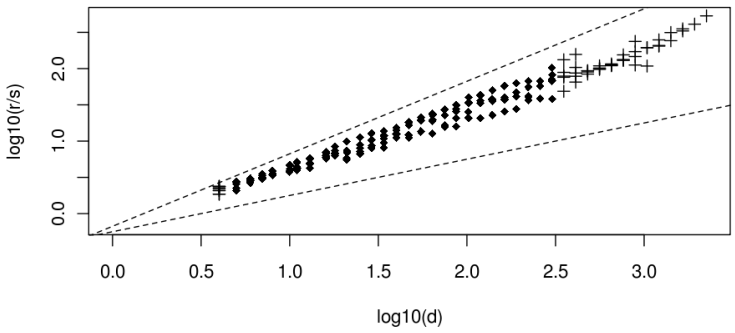

Figure 6 -R/S plot of Slammer

3.4. Analysis of ICMP NMAP PING

The list of target systems included a total of 64310 IP addresses; this is very close to the whole class B (65535 IP addresses) allocated to the University of Plymouth. This is a good indication that the targeting was indiscriminative, given that the network includes a significantly lower number of hosts/allocated IP addresses. And, indeed, as it is revealed by CDF in Figure 7, the hosts that were attacked 20 times or less during the observation period represent more than 99% of the total number of targeted hosts (approximately 64000 hosts).

Further investigation shows that there were only 12 hosts heavily targeted by ICMP NMAP PING almost every day, accounting for 95% of all attacks. Each of these 12 targets was attacked 4000 times per day or more. Given that the IP addresses from the trace were anonymised, the 12 hosts cannot be identified, but they are likely to be servers rather than workstations. The rest of the targets (up to 64298 hosts) received just overt 250k alarms during the observation period.

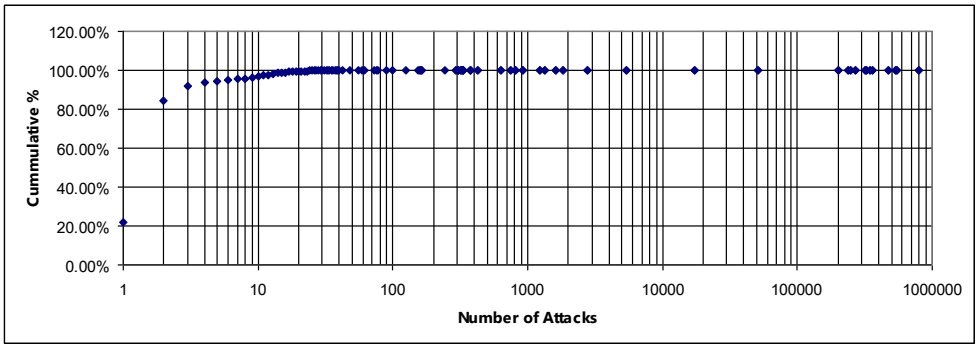

Figure 7 - CDF plot of the Number of Attacks per Target IP

Unlike Slammer, the ICMP attacks came mostly from the US. Although there were a number of contributors with higher numbers of attacks, the distribution was not dominated by any subset of hosts. As for Slammer, the R/S method was used to determine the self similarity and long range dependency within the ICMP attacks. The method produced an estimation of $H = 0.7696193$, indicating a strong degree of self similarity.

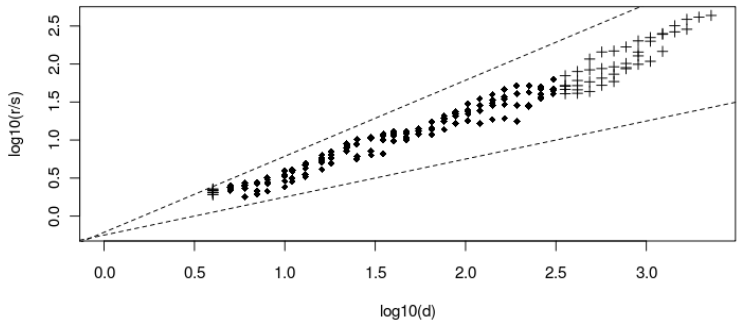

Figure 8 - R/S plot of ICMP NMAP PING

4. Comparison of the results to previous studies

The analysis shows the similar finding as the two prior studies (Moore, 2004; Yegneswaran, 2003) in the way that DoS activities are numerous, a very small collection of sources are responsible for a significant fraction of intrusion attempts and there is a little sign of reduction of such intrusion attempts. This could mean that the behaviour of intrusion attempts on the Internet, especially DoS attacks, has not been changed very much since the earlier studies of the attacks and the situation tends to be going on as common threats on the Internet for very long period of time.

5. Conclusion

The analysis showed that the majority of the alarms in the dataset were due to Slammer. The Slammer attacks were governed by small clusters of hosts each automatically and constantly broadcasting a substantial amount of traffic, therefore the number of attacks did not correlate with the number of the attacking hosts. Additionally, it was found that the periodic pattern of attacks in every 24 hour might cause by the competition for Internet bandwidth among the Slammer sources and the overloading of the mirrored port during peak hours may have also contributed to the results. There was no specific target of the attacks as the number of attacks on each unique IP was normally distributed all over class B of the University of Plymouth's allocated space. The analysis of the time series of number of the attacks per hour showed that the time series was nonstationary, however, strong degree of correlation and periodicity could be spotted from the ACF plot which was supported by further analysis on Self Similarity (SS) and Long Range Dependence (LRD) through the high estimated value of the associated Hurst parameter.

6. References

HIDESHIMA, Y. and KOIKE, H. (2006). STARMINE : A Visualization System for Cyber Attacks. In Proc. *Asia Pacific Symposium on Information Visualisation (APVIS2006)*, Tokyo, Japan. CRPIT, 60. MISUE, K., SUGIYAMA, K. and TANAKA, J., Eds. ACS. 131-138.

JOUNI, V., HERV, D., LUDOVIC, M., ANSSI, L. & MIKA, T. (2009) Processing intrusion detection alert aggregates with time series modeling. *Information Fusion,* 10**,** 312-324.

KIM, D., LEE, T., JUNG, D., IN, P. H., LEE, H. J. (2007) Cyber Threat Trend Analysis Model Using HMM. *Information Assurance and Security, International Symposium on,* The Third International Symposium on Information Assurance and Security.

KOUKIS, D., ET AL., (2006) A Generic Anonymyzation Framework for Network Traffic. *Communications, 2006. ICC '06. ,* 5**,** 2302-2309.

NIST/SEMATECH (2006) e-Handbook of Statistical Methods. The National Institute of Standards and Technology (NIST), http://www.itl.nist.gov/div898/handbook/eda/section3/eda35c.htm (Accessed on 30/06/09).

WU, Q., SHAO, Z (2005) Network Anomaly Detection Using Time Series Analysis. *Proceedings of the Joint International Conference on Autonomic and Autonomous Systems and International Conference on Networking and Services.* IEEE Computer Society.

YEGNESWARAN, V., BARFORD, P., ULLRICH, J. (2003) Internet intrusions: global characteristics and prevalence. *Proceedings of the 2003 ACM SIGMETRICS international Conference on Measurement and Modeling of Computer Systems.* San Diego, CA, USA, ACM.