IDENTIFYING AND INVESTIGATING INTRUSIVE SCANNING PATTERNS BY VISUALIZING NETWORK TELESCOPE TRAFFIC IN A 3-D SCATTER-PLOT

Jean-Pierre van Riel and Barry Irwin

Security and Networks Research Group
Department of Computer Science, Rhodes University

g02v2468@campus.ru.ac.za
B.Irwin@ru.ac.za
Department of Computer Science
Rhodes University
P.O. Box 94
Grahamstown
6140
South Africa

ABSTRACT

Detecting and investigating intrusive Internet activity is an ever-present challenge for network administrators and security researchers. Network monitoring can generate large, unmanageable amounts of log data, which further complicates distinguishing between illegitimate and legitimate traffic. Considering the above issue, this article has two aims. First, it describes an investigative methodology for network monitoring and traffic review; and second, it discusses results from applying this method. The method entails a combination of network telescope traffic capture and visualisation. Observing traffic from the perspective of a dedicated sensor network reduces the volume of data and alleviates the concern of confusing malicious traffic with legitimate traffic. Complimenting this, visual analysis facilitates the rapid review and correlation of events, thereby utilizing human intelligence in the identification of scanning patterns. To demonstrate the proposed method, several months of network telescope traffic is captured and analysed with a tailor made 3D scatter-plot visualisation. As the results show, the visualisation saliently conveys anomalous patterns, and further analysis reveals that these patterns are indicative of covert network probing activity. By incorporating visual analysis with traditional approaches, such as textual log review and the use of an intrusion detection system, this research contributes improved insight into network scanning incidents.

KEY WORDS

Computer networks; network monitoring; network security; intrusion detection; network telescope; network visualisation; 3D scatter-plot.

IDENTIFYING AND INVESTIGATING INTRUSIVE SCANNING PATTERNS BY VISUALIZING NETWORK TELESCOPE TRAFFIC IN A 3-D SCATTER-PLOT

1 INTRODUCTION

There has been a substantial worldwide increase in the accessibility, application, and reliance on information communication technology (ICT) [Gray 2006][1]. Consequently, the burden of monitoring and defending networks and their systems is greater. The challenge is further compounded by an escalating number of security threats. Over the past eleven years (1995-2005), vulnerability reports have shown a rising trend [CERT]. In addition to this, new Win32 virus and worm variants have arisen at an increasing rate [Symantec 2003, 2006]. Consequently, there is pervasive amount of malicious Internet activity. A study conducted by Yegneswaran *et al*. infers that, in 2002, intrusion attempts reached the order of 25 billion incidents on some days, and projections for the average daily number of non-worm scans increased from 6.5 billion to 8.2 billion over a three month observation [Yegneswaran 2003].

The above factors pose serious challenges to monitoring and auditing network security. While increased levels of network traffic generate larger amounts of data, the number of vulnerabilities and exploits accumulate. Hence, the complexity of protecting and monitoring systems is expanding. Moreover, it is questionable whether conventional approaches such as firewalls and intrusion detection systems will improve at a rate sufficient to scale with these issues.

With specific focus on intrusion detection and analysis, Section 2 overviews techniques and issues specific to network monitoring. Section 3 outlines an investigative strategy to deal with some of these difficulties. The strategy entails monitoring Internet activity from the perspective of a dedicated sensor network, and advocates the use of network visualisation (in combination with conventional methods). Throughout the above sections, related work is discussed in context. Section 4 then offers an account of applying the suggested investigative techniques and documents results of particular interest. The conclusion summarises the key arguments and contributions made by this research, and outlines further applications for the investigative approach.

2 MONITORING NETWORKS AND IDENTIFYING INTRUSIVE ACTIVITY

There are a number of motivations for monitoring networks; one is assessing the value of security measures, a second is identifying security breaches, and a third is characterising threats. This section is mostly concerned with the third motivation, and outlines a number of aspects involved with gaining insight as to how intrusive activity probes through networks.

2.1 Network Data Volumes

Monitoring sizeable networks can generate unmanageable amounts of network log data in the order of gigabytes per day. Supposing one were to perform complete traffic capture, even a modest Internet link with a 512 Kbit/s connection operating at a 50% utilization average would transfer 2.7GB in a day. Network security research may warrant full packet traces, but for typical production networks, complete traffic capture is impractical, if not excessive. Often monitoring systems simply log the history of successful and unsuccessful connection attempts, discarding other

---

[1] Additional figures supporting this claim can be found at “Internet World Stats”, <http://www.internetworldstats.com/stats.htm> (20/04/2006).

traffic data such as the packet payloads. Despite reducing the amount of data recorded, connection information for sizeable production networks can result in thousands of log entries per day. This is an intractable amount for a network administrator or security officer to read line by line.

2.2 Intrusion Detection Systems

Network intrusion detection systems (NIDS) offer a collection of algorithms suited to real-time traffic inspection and filter out innocuous traffic by logging network alerts. Typically, these algorithms attempt to match traffic with intrusion signatures or detect anomalies. The signatures describe a series of bytes or transmission sequences known to be indicative of malicious activity. Anomaly detection relies on a characterisation of normal traffic whereby abnormal activity can be identified by monitoring thresholds, detecting protocol violations, or employing statistical analysis.

One criticism of signature based detection methods is that they are only adept at uncovering known attacks. For anomaly based detection methods, the criticism applies to a lesser degree, but anomaly detection will fail to detect both known and unknown intrusion attempts that fall within the characterisation of normal traffic, and tends to have higher a false positive rate. As a second criticism cited against NIDS, the ratio of legitimate traffic versus illegitimate traffic is a dominating probability factor that dictates the likelihood of false positives. In the case of production networks, most traffic is innocuous, thus requiring intrusion detection systems to be extremely accurate [Axelsson 2000, 2004]. In addition to this, attackers can intentionally agitate signature based intrusion detection by crafting packets to trigger an overwhelming number of alerts [Yurick 2002]. A similar affect can be achieved against anomaly detection measures by disrupting traffic; for example, initiating denial of service attacks (DoS). Thresholds and alert suppression are designed to counteract this, but diversion tactics will still succeed by either obfuscating the actual attack, or effectively disabling a detection rule. To counterbalance the weaknesses of each approach, NIDS are becoming hybridised. Snort NIDS is an example of what was formally considered a signature based IDS that has evolved to incorporate some anomaly detection [Beale 2004, Snort].

2.3 Textually Based Network Traffic Review

Textual log inspection and packet capture review tools still serve their purpose when detailed examination is desired. Many of these tools provide filtering mechanisms. The capability to filter allows events of interest to be isolated, but requires prior knowledge of what to look for. An example of a basic command line tool is tcpdump, which can record, filter, and output packet information [TCPDump]. Ethereal offers a graphical interface and can perform some basic statistical analysis [Ethereal].

2.4 Visualisation and Graphical Representation of Network Traffic

Humans read and understand text in a sequential manner though the auditory cognitive modality. For this reason, it is difficult to correlate numerous attributes and facets when data is presented in a textual format. Contrary to this, the visual/spatial cognitive modality is highly parallel and preattentive [Wickens 1983]. Therefore, human vision is well suited to pattern recognition, and unlike signature based intrusion detection, facilitates the observation of unexpected patterns [Ball 2004]. “By incorporating human perception into the data-mining process, researchers can detect patterns in data missed by traditional automatic data mining methods” [Yin 2004].

Visualisation can excel at providing the viewer with a rapid overview of network traffic, but does so at the cost of diminishing detail. For this reason, several visualisations provide various methods to focus more closely on facets of interest. Navigation and zooming can make 'details on demand' accessible [Scanmap3D, Fisk 2003]. The ability to 'visually drill down' into separate subcomponent views is another option allowed by visualisation [Yin 2004]. A further mechanism of accessing detail is the ability to select graphical objects and raise textual information on demand [Scanmap3D, Fisk 2003, Etherape].

One serious challenge facing the practical application of network visualisation is the issue of limited scalability. Large amounts of data can overwhelm a visualisation, as too many graphical objects clutter the display and render the image unintelligible. In many visualisations, lines are an intuitive representation of connection [Ball 2004, Scanmap3D, Fisk 2003, Teoh 2004, Etherape, Yin 2004]. However, as argued in a previous article, line-based representations suffer from issues such as crossover and costly display-space utilisation per event [van Riel 2006]. Point based representations offer better scalability. For this reason, Stephen Lau's 3D “Spinning Cube of Potential Doom” visualisation is a primary design reference for this work (Section 3.2 follows with more details) [Lau 2004].

Many visualisations plot connections in a manner that visually distinguishes the internal home network domain from external Internet domains [Ball 2004, Scanmap3D, Fisk 2003, Lau 2004, Yin 2004]. Some visualisations further distinguish the direction of traffic as inbound or outbound [Yin 2004].

2.5 Network Telescopes and Honeynets

Production networks pose two problems for network monitoring, namely large amounts of legitimate traffic and comparatively low volume of illegitimate traffic. Consequently, the illegitimate traffic is obfuscated by the legitimate traffic - a classic 'needle in the haystack' problem. As discussed in Section 2.2, this can result in intrusion detection systems producing an overwhelming number of false alarms. One solution is to remove the haystack from the needles. As dedicated sensor networks, network telescopes and honeynets offer a clearer perspective of intrusive network activity. These networks are designated regions of IP address space where no legitimate production services or client hosts reside, and therefore, all traffic targeting the address range is unsanctioned. This confines the possibility of false positives to mistaking unintentional traffic as actual intrusion attempts, where unintentional traffic is typically the result of network miss-configuration and errors.

Network telescopes passively monitor incoming traffic without offering any response. The main disadvantage is that this will only observe initial probing packets or single packet exploits, which substantially limits the information available for analysing intrusion attempts. Alternatively, honeynets actively monitor the network range by containing one or more hosts that respond to incoming traffic. By responding and interacting with incoming traffic, the general intent is to gain more information about intrusion attempts through observing further stages in communication. Actively responding does come with associated risks such as amplifying malicious activity. Due to the necessary exposure, a honey pot system also faces more risk of actually being compromised (as opposed to merely posing as vulnerable). Another concern is that attackers may be able to infer the presence of a particular honeynet from the characteristics of its responses. Therefore, the measurements made by a honeynet can be misleading. Contrary to this, it is impossible to differentiate an unresponsive empty address space from a network telescope.

A few other caveats warrant mentioning. While specialist sensor networks are well suited to identifying reconnaissance activity, they do not provide information about direct attacks focused on a particular production host. Secondly, their effectiveness and range of observation can depend on the size and locality of the sensor network [Moore 2004]. Lastly, considering that IPv4 address space is a costly and limited resource for most organisations, the required empty network segments will conflict with the policy of maximal address utilization.

3 INVESTIGATIVE METHODOLOGY

This section offers an account of the investigative methodology used to establish the findings presented in Section 4. Firstly, the network monitoring scenario and data collection process is described. This is followed by a brief description of the original concepts behind InetVis - an academic visualisation project developed for Internet traffic visualisation – as well as an overview

of key features found to be valuable in the investigation processes adopted. The section then closes with a brief account of ad hoc tools used to perform further analysis.

3.1 Data Collection

A designated class C address space has been allocated for collecting raw unfiltered Internet Traffic samples. Due to passive monitoring restrictions on the address range, the sensor is a network telescope. Since the beginning of August 2005, packet traces of all IP traffic has been captured and recorded in the libpcap binary format. Network tools such as Ethereal, Snort and Etherape [Etherape] can read this format. Capture files were separated by month, as large files tend to cause difficulties when opened with Ethereal. The capture dataset used in this investigation is the 8-month period between August 2005 and the end of March 2006. For some months, capture files were as large as 117MB, and contained in excess of 800,000 packets. In total, 4.7 million packets were captured over eight months and amounted to 666MB of data.

3.2 Visual Traffic Review with InetVis

Given the substantial quantity of data to be analysed, the promising scalability of Stephen Lau's "Spinning Cube of Potential Doom Visualisation" made it a suitable visualisation concept to build upon. At the time of commencing research, no such visualisation was publicly available and necessitated a completely independent implementation. (Recently [Doomcube] has been released as a basic clone of Lau's visualisation, but currently lacks several of the features described in Section 3.2.2). This project's custom implementation is named InetVis (*Inter*net *Visualisation*), and includes a significant number of important extensions to enhance the original concept. The aim of the project is to produce a viable visual analysis tool. A brief description of its concept, design and key features will follow. (For a more detailed account of the added contributions to Lau's original concept, their motivating considerations and benefits, refer to [van Riel 2006].)

3.2.1 Concept and Design

Computer graphics offers several possibilities beyond traditional 2-D static graphs. From the onset, the design objective is to take full advantage of graphical computer capabilities and maximise the amount of information that can be visually conveyed. Various techniques such as animation, colour, transparency, and varied size can provide mechanisms to extend the number of data attributes represented. InetVis makes use of these methods and is a fully dynamic time animated visualisation that plots network events as points in a three-dimensional scatter plot. The plotting scheme extends Stephen Lau's work with the addition of the ICMP plane as illustrated in Figure 1. This scheme is suited to convey address scanning across the network range as well as target host port scans (and a treatment of scanning classification follows in Section 4.1).

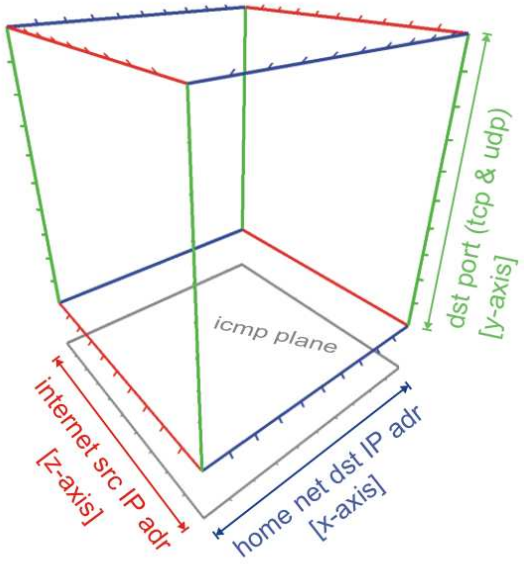


Figure 1: InetVis plotting scheme

3.2.2 Key Features

The majority of network traffic tends to occur in the well used lower service port regions. With a linear plot along the y-axis, this can cluster and obscure traffic concentrated in the lower port regions. An important extension to the plotting scheme is a log plot function. It has a single parameter to adjust the level of spatial expansion at the lower ranges (with a proportionate level of contraction at the higher ranges).

The temporal order in which events occur is paramount to identifying and understanding scanning patterns. InetVis is intuitively time-animated and able to replay events in the time order they originally occurred. Providing the viewer with the ability to adjust the replay position, playback rate, and time window enables highly refined temporal control for traffic review. An increased time scale (playback rate) enables rapid review for finding events of interest while slowing playback speed down is useful for carrying out meticulous inspections of specific events [Fisk 2003]. The size of the 'time window' (or time frame) is linked to the replay position and implicitly performs dynamic time filtering [Yin 2004]. During playback, the continuous introduction and removal of points according to the moving time window conveys the temporal order that network packets occurred. Linked to the time window size, the transparent fading of points makes new events fully opaque in contrast to older events that are gradually faded out [Ball 2004]. Added to this, momentarily bulging new points provides a pulse effect for the viewer to distinguish recent events.

The ability to focus on events of interest is provided in several ways. Immersive navigation is made possible by translation (moving), rotation, and scaling (zooming) [Scanmap3D, Fisk 2003, Lau 2004]. Added to this, the capability of setting a source address range, destination address range, and destination port range offers a mechanism to visually ‘drill down’ into a sub-set of the data and see regions of interest in isolation [Yin 2004]. Furthermore, traffic can be filtered with BPF expressions, which offer a flexible control to remove uninteresting traffic [McCanne 1993]. Various colour schemes can be chosen to investigate other attributes in the data, such as colouring points according to destination port, source port, source address, protocol, or packet size.

3.2.3 Implementation and Performance

Interactive 3-D graphics can place strenuous performance demands on systems, and poor performance is another factor that can limit the scalability of a visualisation. For this reason, InetVis was implemented in C++ with OpenGL, and directly interfaces with libpcap to read data from the capture files. With a fair number of performance optimizations in place, InetVis renders at a stable rate of 25 frames per second for up to 500,000 events (points), and has been tested with 4.7 million events where an acceptable level of interaction was maintained[2].

3.2.4 Identifying and Investigating Events

The procedure outlined here is a process whereby the strengths of visual cognition are used to perform pattern recognition and identify events, thereby incorporating human intelligence into the detection process. During the development of InetVis, tests were conducted with Nmap [Nmap] to produce visual signatures as a proof of concept [van Riel 2006]. These signatures also serve as references for identifying common scanning techniques found in the network telescope traffic.

The initial step toward reviewing a full month’s capture is to form an overview of all the events. Typically, a very high replay rate of 86400*x* speedup (one day per second) can be combined with a time window of seven days. This allows a month’s traffic to be skimmed over in roughly 30 seconds where each event would then be represented for seven seconds. Although this tends to be too fast to identify specific events of interest, it provides a quick and broad chronological impression of events. The static view of an entire month’s traffic (with a 31 day time window) is useful for observing patterns that are randomly formed over longer periods of time, and experimenting with various colour schemes may reveal subtle correlations.

The strategy for identifying and isolating events of interest follows an iterative approach. Begin with a fast replay rate and large time window to identify events that progress slowly. Then gradually reduce the replay rate and time window to allow the details of faster events to become more evident. From experiences with the tool, a rate of 3600*x* (one hour per second) and a time

---

[2] Tests performed on an Intel 3.0GHz Pentium 4, 1GB RAM, NVidia GeForce 6600GT (running Ubuntu 5.10)

window of 24 hours is suitable for the discovery of related scanning events (provided the events do not progress excessively slowly). Once an event of interest is identified, the source address and destination port ranges can be reduced to drill down into the visualisation (place a subset of the data into full view) and obtain a clearer perspective of the event. For rapid events, further reduction of the time window and replay speed will improve the viewer's sense of timing between the occurrences of packets. Once the resolution (in terms of port numbers and addresses) is sufficiently refined, filters can be applied to isolate and refine a clear view of the event. At any stage, colour schemes can be experimented with to reveal links between attributes of different events. The isolated event can then be recorded to a capture file for further analysis with tools such as Ethereal and Snort.

3.3 Detailed Investigation of Specific Events

Once an event has been isolated into its own capture file, analysis with tools such as Ethereal and Snort can provide more detail about the event. Ethereal allows the reviewer to perform low-level packet analysis, and can be used to report some simple statistics. Snort can be used for automated analysis that either provides a description and classification of an intrusive event, or fails to identify the event. In the case of failure, either the human reviewer has falsely identified the event as intrusive, or it is an instance of a false negative for the IDS. One thing to note is that only the network scanning detection module of Snort is likely to be of use, due to the network telescope only capturing initial probe packets. A final step in the research is then to correlate the event with vulnerability and exploit advisories.

4 ANALYSIS AND FINDINGS

The investigative tools and techniques outlined in Section 3 are used to establish the findings and analysis presented in this section. The discussion commences with a categorisation of the common types of scanning and follows with a more detailed account of select events. The chosen events of interest are considered out of the ordinary, anomalous, or good examples that emphasise the advantages of visual investigation. The section then concludes with a broad report of figures summarising monthly network telescope activity.

4.1 Characteristic Types of Network Scanning

Within the literature, various synonymous terms are employed to describe two common types of network scanning, namely horizontal 'port-sweeps' and vertical 'port-scans'. In the network telescope traffic reviewed, horizontal lines across the network range are numerous and prominent (Figure 3 provides examples in due course). The lines occur either in the ICMP plane, or within the cube at height corresponding to the TCP/UDP destination port. Aptly relating to this visual metaphor, Yegneswaran *et al*. use the term 'horizontal-scanning' to describe the probes that emerge from these horizontal lines [Yegneswaran 2003]. The Snort documentation classifies the scan as a 'port-sweep', defining it as a single source host targeting multiple destination network addresses using a particular destination port [Snort]. Analogous to a horizontal-scan/port-sweep, another simple probing technique is called a 'vertical-scan', or 'port-scan', whereby a single source address targets a single destination address on a multiple number of destination ports.

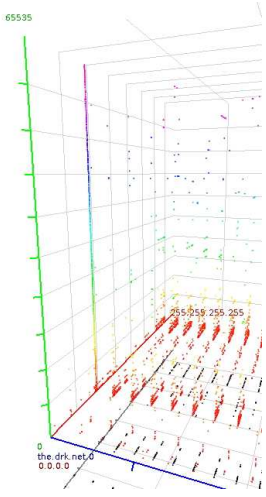
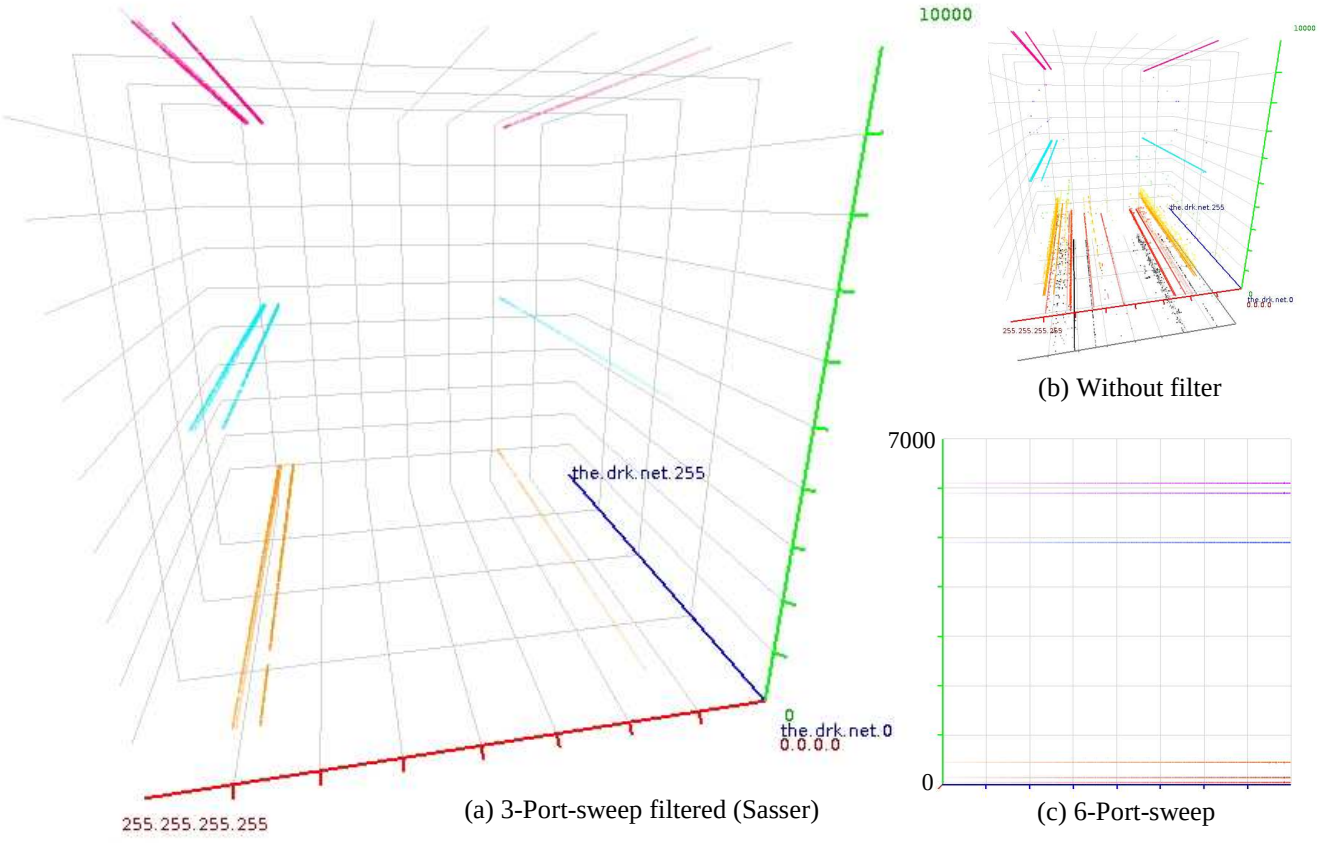


Figure 2: Vertical port-scan

4.2 Scanning Incidents and Anomalies of Interest

Throughout the eight-month capture, only one extensive vertical-scan is noted. This is shown in Figure 2 where each point is coloured according to the destination port and results in the rainbow coloured line[3] as each

[3] Note that in Fig. 2, destination address axis (blue) is scaled to 64 addresses, whereas the default is 256.



(a) 3-Port-sweep filtered (Sasser)

(b) Without filter

(c) 6-Port-sweep

Figure 3: Coordinated horizontal scans across multiple ports

consecutive port is probed. The scan randomly progresses over the full port range taking 23 days to complete during September and October 2005, and proves slow enough to evade detection with Snort 2.2.4. Snort testing was conducted with the 'sfportscan' pre-processor module. The default sensitivity was changed from low to high and detection of all scanning types was explicitly enabled – all other defaults were preserved [Snort].

Generally, an attacker would first establish if the target were present by sending a single probe (such as an ICMP ping) before expending time to scan multiple ports on a host that may be offline or not exist. Since network telescopes do not respond to any incoming traffic, these reconnaissance probes always fail, and the rarity of vertical port scans is expected.

4.2.1 Related Horizontal-Scans

Several instances of related port-sweeps are readily evident in the network telescope traffic. Figure 3(a) and 3(b) show sets of corresponding lines that originate from four distinct sources, and sweep across three specific ports; in the figures, two of the sources are close together producing the thicker line on the left. The characteristic visual scanning impressions left by this particular 3-port-sweep-scan is prevalent throughout the 8 months of traffic, indicating that it is replicated and may be viral activity. The scans from all four hosts complete within one minute, but do not occur simultaneously. One minute is orders of magnitude faster than their regular observation frequency, and this suggests that there may be some relation between the incidents (or an unlikely time coincidence). Each source completes scanning ports 1023, 5554, and 9898 in just over 5 seconds.

The three ports are associated with Sasser and Dabber viral activity [SANS]. Each scan begins with port 5554 and port 1023 scan follows a very slight step behind. By comparison, the port 9898 probe is noticeably delayed. The apparent explanation is that this is Dabber virus probing activity. Sasser uses port 5554 to open an ftp service for downloading the worm binary, and similarly port 1023 is used by the Sasser.E variant. Dabber sequentially scans IP address space on

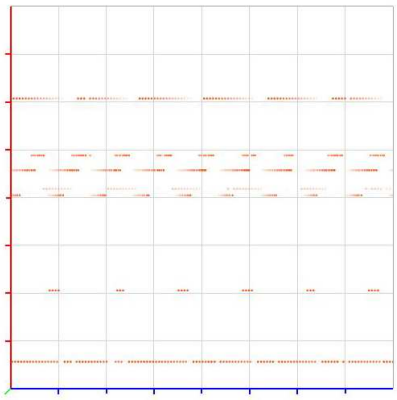


Figure 4: 'Creepy crawly' scans – Sapphire/Slammer

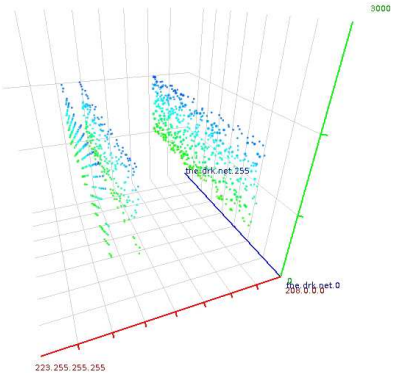


Figure 5: Psudo-random traffic

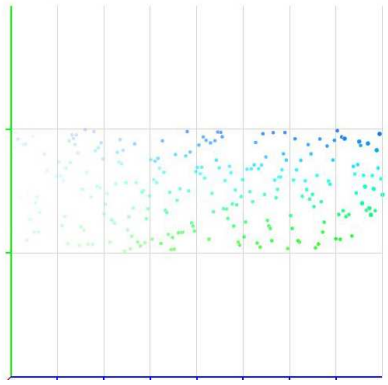


Figure 6: Fast random scan

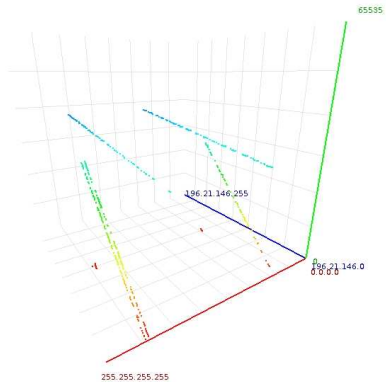


Figure 7: Anomalous diagonals

port 5554 and 1023 with the objective of further exploiting systems previously compromised by Sasser. The attack waits for newly inserted exploit code to be injected into buffer overflow vulnerability and execute, installing a backdoor on port 9898.

Figure 3(c) is an orthographic front view of the destination address and port range. The image illustrates a different type of multiple port-sweep scan, and unlike the 3-port-sweep, it simultaneously probes across six ports (as is shown by the effect of transparent decay). The scan's progression time is also notably slower and takes 149 seconds to complete (as apposed to 5 seconds).

4.2.2 'Creepy Crawly' Horizontal Scans

In the eight months of traffic capture, the timing and manner in which port-sweeps are conducted is diverse. Some scans are random, eventually filling out the address range, whereas other scans progress in a sequential manner. In Figure 4 an unconventional, yet prevalent port-sweep scans the address range in small line segments and is called a 'creepy crawly' due to the characteristic crawling motion it produces when time animated. Figure 4 presents an orthographic top view image of traffic from September 20th 2005, and is taken with the time window set to two days (recall that the red axis is the source IP, and the blue axis the destination IP). As can be seen, the timing and spacing for different instances of the 'creepy crawly' vary considerably, and the very top specimen in Figure 4 progresses in the opposite direction to the rest.

The 'creepy crawly' scan is conducted on UDP port 1434 which is associated with several critical MS-SQL vulnerabilities and the infamous Slammer/Sapphire worm of 2003. All the packets concerned had a characteristic IP size of 404 bytes confirming that these scans are Slammer worm activity. The resultant scanning patterns seen in the image are presumably the effect of this worm's poor pseudo random number generator implementation. Slammer selects addresses in a manner that keeps the 25th and 26th bits of an IP address constant for a given execution of the worm [Moore 2003]. This and various other issues with its random number generation explain why the pattern produced fails to appear random at all. Once more, Snort 2.2.4 failed to detect this as scanning activity, despite being set to high sensitivity.

4.2.3 Random Distributed Scans and Anomalous Diagonals

Figure 5 and Figure 6 exhibit examples of pseudo-random activity occurring within the bounds of destination port 1000 to port 2000. All packets received are from TCP source port 80, the standard HTTP port, and most packets have SYN/ACK flags set (except for one scan instance that has RST/ACK flags set). The SYN/ACK packets are usually sent as the second step in a TCP handshake response to a SYN packet that attempts to initiate the TCP connection. A SYN/ACK indicates that the target port is

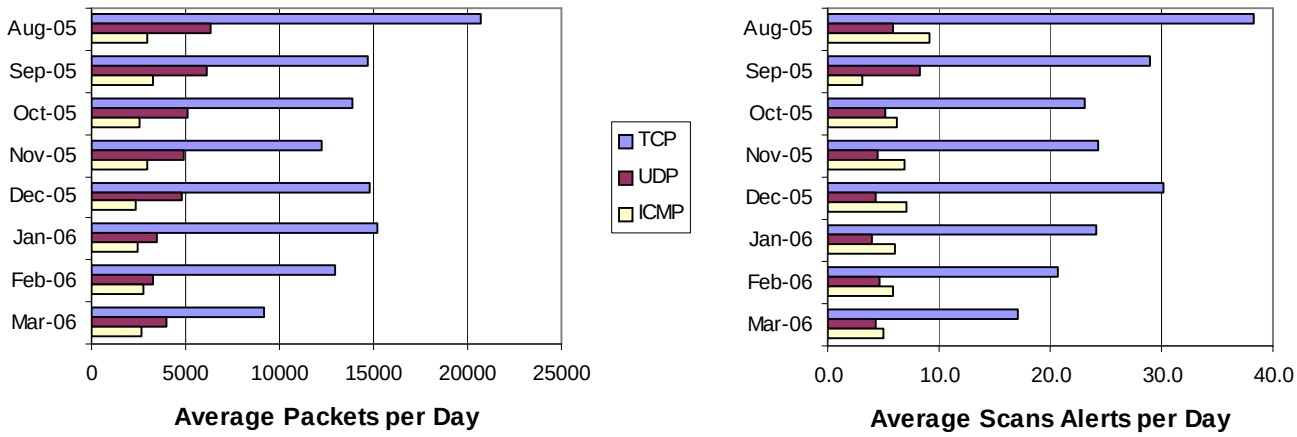


Figure 8: Packet and Snort alert counts by month

open while RST/ACK packets are the standard response when the port is closed. One interpretation of the images is that they could be the result of a denial of service (DoS) attempt, because many of the random patterns occurred in a very short time window – in one particular case, 233 SYN/ACK arrived from 61.145.127.92 within in 45 milliseconds (Figure 6). However, this IP address and the other addresses in question (Figure 5) are not registered web servers making the DoS explanation less tenable.

With a fifty-millisecond time window and the transparent decay, Figure 6 displays increased levels of opacity toward the right. This suggests that the occurrence of events is not entirely random due to the time progression of points from left to right. As an alternative explanation, the activity can be considered a deceptive network sweep masquerading as an HTTP connection (which could appear normal on a network populated with numerous clients making use of HTTP web services). In support of this hypothesis, the 233 packet count is nearly enough to cover the class C address space with one packet per address. Upon closer inspection, it becomes evident that this is indeed the case - each packet targets a distinct address. As can be expected, Snort 2.2.4 did not detect this novel stealthy scan. Arguably, a signature general enough to detect this pseudo-random activity would generate too many false positives.

Figure 7 shows a filtered view of anomalous diagonals with packets that are also allegedly sourced from port 80. Unlike the fast but stealthy random scan in Figure 7, these diagonals take more than a month to form. In supposing that the diagonals are not the result of some obscure network error, the foreseeable purpose of diagonally scanning is to discover hosts while attempting to evade intrusion detection. In the case of Snort 2.2.4, this traffic did not trigger any alerts.

4.3 Summary of Alert and Packet Counts

The two graphs in Figure 8 report the average of daily counts for packets and alerts for each month. TCP is evidently the favoured protocol for performing scanning, and correspondingly, TCP packets form the greater share of the traffic. For all months UDP packets are more numerous than ICMP packets, yet the number of ICMP scans detected with Snort are greater than the number of UDP scans detected (with the exception of August). This implies at least two possible explanations. Either a greater proportion UDP traffic is benign, or the detection rate for ICMP scans is higher than that of UDP scans. In support of the second explanation, ICMP does not have any port information, and therefore, scans cannot be hidden by randomly diffusing the probe packets between port numbers (see Figure 6).

5 CONCLUSION

The increase in network use obviously results in larger volumes of network data. Coupled with this, the accumulating number of security threats further complicates the task of tracking vulnerabilities and detecting exploits. Drawing from Section 2's outline of strengths and weakness for various

network monitoring approaches, Section 3 puts forward a description of the methodology used to conduct forensic network traffic analyses, and in doing so suggests some answers to the aforementioned challenges. The key strategic part is removing legitimate traffic by employing dedicated sensor networks, which greatly reduces the volume of data and significantly reduces the probability of false positives. The second key part is incorporating human intelligence in the detection process. The merits of visual cognition provide the reviewer with far more insight than an obscure 'black box' NIDS producing hundreds, if not thousands of textual alerts. As seen in Section 4, a number of interesting and covert incidents were able to evade the NIDS, but could be visually discovered and analysed. Although visualisation was used as the primary detection method, and provisioned many cues and insights, other tools were needed to conduct further detailed analysis (namely Ethereal and Snort). Hence, this paper does not suggest visualisation should replace such tools, but rather, that it should function as a supplemental analysis tool.

The research presented here reaches two general conclusions. Firstly, the use of dedicated sensor networks is strongly advocated. Such a network can be assigned alongside a production network and acts as a clear indicator of intrusive activity. This in turn provides a valuable reference for exposing false positives in the production network and can also indicate the occurrence of false negatives. Secondly, the use of visualisation is recommended for conducting traffic review. Whilst it may not be a practical way to perform full time monitoring, its suited application is forensic auditing of network scanning activity, and may also prove useful for evaluating other security measures – for example, revealing what the NIDS fails to uncover.

6 ACKNOWLEDGEMENT

This work was undertaken in the Distributed Multimedia Centre of Excellence at Rhodes University, with financial support from Telkom SA, Business Connexion, Comverse, Verso Technologies, Tellabs and StorTech THRIP, and the National Research Foundation.

7 REFERENCES

[Axelsson 2000] Axelsson, S. The Base-Rate Fallacy and the Difficulty of Intrusion Detection, ACM Transactions on Information and System Security, Vol. 3 No. 3, (August 2000). ACM Press, New York. 2000. p.186-205.

[Axelsson 2004] "Combining a Bayesian Classifier with Visualisation: Understanding the IDS". In Proceedings of the 2004 ACM workshop on Visualisation and data mining for computer security. ACM Press, New York. 2004. p. 99-108.

[Ball 2004] Ball, R., Fink, G.A., North, C. "Home-centric visualisation of network traffic for security administration". In Proceedings of the 2004 ACM workshop on Visualisation and data mining for computer security, ACM Press, New York. 2004. p. 55-64.

[Beale 2004] Beale J., Baker, A.R., Caswell, B., Poor, M., Alder, R., Babbin, B., Doxtater, A., Foster, J.C., Kohlenberg, T., Rash, M. Snort 2.1 Intrusion Detection, Second Edition. Syngress Publishing, Inc. Rockland, 2004. Chapter 6, p. 196-265.

[CERT] CERT/CC – CERT® Coordination Center. Carnegie Mellon Software Engineering Institute. Carnegie Mellon University. 2006 <http://www.cert.org/stats/cert_stats.html> (20/04/2006).

[Scanmap3D] Clark, D. "Scanmap3D" open source visualisation for Snort. Scanmap3D-2.1b and Scanmap3D-3.0. <http://scanmap3d.sourceforge.net/> (29/05/2005).

[Doomcube] Kershaw, M. The GPL Cube of Potential Doom. <http://www.kismetwireless.net/doomcube/> (05/04/2006).

[Fisk 2003] Fisk, M., Smith, S.A., Weber, P.M., Kothapally, S., Caudell, T.P. "Immersive Network Monitoring." In proceedings of PAM2003 – Passive and Active Measurement 2003, NLANR/MNA (National Laboratory for Applied Network Research / Measurement and Network Analysis Group). 2003. <http://public.lanl.gov/mfisk/papers/pam03.pdf> (25/05/2005).

[Ethereal] Ethereal Protocol analysis tool <http://www.ethereal.com> (07/11/2005).

[Gray 2006] Gray, V., Magpantay, E., Thompson, H., de Ridder, J., Southwood, R. World Telecommunication/ICT Development Report 2006 on Measuring ICT for Social and Economic Development. International Telecommunication Union, 2006. <http://www.itu.int/ITU-D/ict/publications/wtdr_06/material/WTDR2006_Sum_e.pdf> (20/04/2006).

[Lau 2004] Lau, S. "The Spinning Cube of Potential Doom". In Communications of the ACM archive, Volume 47, Issue 6, ACM Press, New York. 2004. p. 25-26.

[McCanne 1993] McCanne, S., Jacobson, V. "The BSD packet filter: A new architecture for user-level packet capture". Proceedings of the Winter 1993 USENIX Conference. USENIX Association. January 1993. p. 259-269.

[Moore 2003] Moore, D., Paxson, V., Savage, S., Shannon, C., Staniford, C. Weaver, N. "The Spread of the Sapphire/Slammer Worm" Cooperative Association for Internet Data Analysis (CAIDA). 2003. <http://www.caida.org/publications/papers/2003/sapphire/sapphire.html> (20/04/2006).

[Moore 2004] Moore, D., Shannon, S., Voelker, G.M, Savage, S. "Network Telescopes: Technical Report" Cooperative Association for Internet Data Analysis (CAIDA). 2004. <www.caida.org/publications/papers/2004/tr-2004-04/tr-2004-04.pdf> (2006/04/20).

[Nmap] Nmap "Network Mapper" security scanner. <http://www.insecure.org/nmap/> (07/11/2005).

[SANS] SANS Internet Storm Centre <http://isc.sans.org/port_details.php> (24/06/2006).

[Snort] Roesch, M. Snort network intrusion prevention and detection system. "Snort Users Manual 2.4.0RC1". Sourcefire, Inc. <http://www.snort.org/> (17/04/2006).

[Symantec 2006] Turner, D. (exec. Ed.), Entwise, S. (Ed.), *et al*. Symantec Internet Security Threat Report, Volume IX, March 2006. Symantec Corporation. 2006. <http://www.symantec.com/enterprise/threatreport/index.jsp> (20/04/2006).

[Symantec 2003] Higgins, M. (Ed.), *et al*. Symantec Internet Security Threat Report, Volume III, February 2003. Symantec Corporation. 2003. <http://www.symantec.com/enterprise/threatreport/index.jsp> (20/04/2006).

[van Riel 2006] van Riel, J-P., Irwin, B. "InetVis, a visual tool for network telescope traffic analysis". Afrigaph '06: Proceedings of the 4th international conference on Computer graphics, virtual reality, visualisation and interaction in Africa. ACM Press, New York. 2006. p. 85-89.

[TCPDump] tcpdump packet capture utility and libpcap packet capture library. <http://www.tcpdump.org> (06/11/2005).

[Teoh 2004] Teoh, S.T., Ma, K-L., Wu, S.F., Jankun-Kelly, D.T.J. "Detecting Flaws and Intruders with Visual Data Analysis". In IEEE Computer Graphics and Applications, Volume 24, Issue 5. IEEE Computer Society Press, Los Alamitos, CA. 2004. p. 27-35.

[Etherape] Toledo, J. *et al*. EtherApe: a graphical network monitor. <http://etherape.sourceforge.net/> (2006/04/23).

[Wickens 1983] Wickens, C., Sandry, D., Vidulich, M. "Compatibility and resource competition between modalities of input, central processing, and output". Human Factors, Vol. 25, No. 2. 1983. p. 227-247.

[Yegneswaran 2003] Yegneswaran, V., Barford, P. Ullrich, Y. "Internet Intrusions: Global Characteristics and Prevalence". Proceedings of the 2003 ACM SIGMETRICS international conference on Measurement and Modelling of Computer Systems, June 2003. ACM Press, New York, 2003. p. 138-147.

[Yin 2004] Yin, X., Yurcik, W., Treaster, M., Li, Y., Lakkaraju, K. VisFlowConnect: netflow visualisations of link relationships for security situational awareness. In Proceedings of the 2004 ACM workshop on Visualisation and data mining for computer security, ACM Press, New York. 2004. p. 35-44.

[Yurick 2002] Yurick, W. "Controlling Intrusion Detection Systems by Generating False Positives: Squealing Snort Proof-of-Concept". Proceedings of the 27th Annual IEEE Conference on Local Computer Networks (LCN'02). IEEE Computer Society, Washington, DC, USA. 2002. p. 134-135.