Network Telescope Metrics

Barry Irwin
Security and Networks Research Group, Department of Computer Science
Rhodes University
Grahamstown, South Africa
b.irwin@ru.ac.za

***Abstract*—Network telescopes are a means of passive network monitoring, increasingly being used as part of a holistic network security program. One problem encountered by researchers in the sharing of the collected data form these systems. This is either due to the size of the data, or possibly a need to maintain the privacy of the Network address space being used for monitoring. This paper proposes a selection of metrics which can be used to communicate the most salient information contained in the data-set with other researchers, without the need to exchange or disclose the data-sets. Descriptive metrics for the sensor system are discussed along with numerical analysis data. The case for the use of graphical summary data is also presented.**

***Index Terms*—network telescope, network management, network, security, metrics**

I. Introduction

ONE of the problems faced when dealing with large numbers of data points is deciding which means of data summarisation best conveys the overall picture of the data being analysed in such a way that meaningful conclusions can be drawn from it. The approach taken in this chapter is to present a number of metrics that can be used to provide aggregate and summarised data for a network telescope data-set. A series of common metrics specific to network telescopes are proposed which can be used to quantitatively compare data-sets from differing sources/organisations without the need to transfer increasingly large raw capture sets, which in themselves have issues relating to privacy of the payloads, and possibly the anonymity of the actual sensor.

The remainder of this paper consists of a brief introduction to Network Telescopes in Section II. The discussion of the proposed metrics is broken down into two categories — those describing the telescope/sensor itself (Sensor Metrics) in Section III, and those describing the characteristics of a particularly sampled data-set (Dataset Metrics) in Section IV. A single network telescope may produce several types of datasets, for example one may be processed at monthly intervals while other systems may produce weekly or even annual bundling of data. Provided the operating conditions remain the same, the sensor metrics should remain consistent and only the data-set metrics change.

The graphical representation of these metrics is presented in Section V, which provides useful trending information without too much detail. The paper concludes with a discussion relating to proposed common metrics that could be used for comparative analysis of network telescope data sets, both for intra-institutional and inter sensor use.

This paper is not intended to provide a detailed overview of the field of metrics and information security related metrics in particular. Andrew Jaquith's *Security Metrics* book [1] provides an excellent overview of the topic. More detail on the visualisation of security metrics can be found in *Security Data Visualization* [2] and *Applied Security Visualization* [3]. Further resources can also be found on the http://securitymetrics.org website and accompanying mailing list.

II. Network Telescopes

The concept of Network Telescopes as a means of monitoring activity on the Internet at large, has developed progressively since early 2000. In essence a network telescope is a passive sensor system that collects incoming traffic or 'radiation' from the Internet. This radiation is constituted from multiple source systems and traffic types. The analysis of this collected data can provide useful insight into the operation of the Internet, or even particular events such as worms or distributed denial of service (DDoS) attacks. Over the last few years researchers have focused on using telescopes for DDoS analysis as discussed in [4]. The RUSCOPE1 data-set referred to in this paper was collected by the author from August 2005 to September 2009, using a /24 netblock allocation.

III. Sensor Metrics

Sensor metrics are those metrics that describe the configuration of a particular network telescope (sensor). These are applicable to a system for a given period of time and may change as the operation and configuration of the sensor system evolves. They should, however, remain relatively constant across multiple datasets at smaller time scales, such as months or quarters, but may be stable over much longer periods. The purpose of reporting these values is to describe the operation and configuration of the network sensor in such a manner that easily allows comparison of results with data from other organisations or even with other sensors in the same organisation.

A. Lens Size and Shape

The size and type of the IP address space being used for monitoring can have a large bearing on the volume and quality of data collected. Knowledge of this is important in order to be able to perform comparative analysis among different datasets from differing collection sources. It is proposed that the size of the address space (as an aggregate total if smaller components

are used) — being analogous to the lens size in a traditional astronomical telescope — be expressed in CIDR (also referred to as 'slash') notation [5]. What is important is that the makeup of the address block be made clear – in essence the shape or form of the collecting mechanism used. Details as to whether the address space is a single contiguous block or an aggregate block comprised of smaller sized portions of address space should be disclosed.

Information that is useful, but not critical, would be how the sensor space was dispersed in the event of it being comprised of smaller constituent parts, and if the monitored block is part of a larger allocation used by an organisation. For example a block of size /29 (eight addresses) in each subnet in an organisation with a /16 allocation may be monitored. While there may be a relatively wide coverage (dependent on overall subnet size), there is still little address space in use, ranging from the equivalent of a /21 if subnets of size /24 are to used to the equivalent of a single contiguous /23 block if subnet assignments of size /22 are utilised. The fact that the blocks are not contiguous can have an impact on certain kinds of analyses such as that investigating the scanning algorithms behind observed scan traffic as described in [6], [7], [8].

Should the sensor be a portion of a larger block this could be expressed in a suitable format, preferably as a fraction rather than a percentage. As an example, two /29 networks within a larger /23 netblock may be used for network telescope monitoring. This amounts to a /28 in total, which in effect represents $\frac{16}{512}$ addresses or $\frac{1}{32}$ of the address space, which is a more understandable measure than when expressed as a percentage: 3.125%.

B. Mode of operation

The mode of operation of the network sensor(s) is essential when processing either historical data or data obtained from another organisation. Four common classes of operations for network telescopes are described below, and it should be stated as to which of these best fits the mode of operation of the sensor being described.

- **Passive:** The sensor operates as a traditional network telescope, being completely passive, only logging traffic and affording no response on the monitored address space.
- **Low Interaction:** Some form of low interaction system is used in order to enable the observation of data payloads in the subsequent first data packet. This is particularly important if TCP payloads are of interest to the researcher as these will only be present for active traffic once the initial TCP handshake is completed. This is usually performed by the software system responding to incoming TCP requests, and completing the 3-way handshake. The connection is then dropped after the first data packet has been received. Payload capture can be of value in determining the nature of the inbound traffic and may be of use when combined with some form of NIDS.
- **Full – Medium Interaction:** A general honeypot system such as honeyd or dionaea, or a protocol specific tool such as kippo (a SSH honeypot) is used to emulate live systems on given addresses/ports or a combination of the aforementioned. The advantage of this is that TCP payloads can be obtained beyond the first packet, as in the low interaction systems.
- **Live system**: Live physical or virtualised systems are used as endpoints for the monitored address space.

The mode of operation has a major bearing on the type of traffic and consequently data that is likely to be observed. Thus it is important to disclose to other researchers the mode of operation used by the sensor in order to capture traffic (and to record for historical use).

C. Sampling

Details of the means and frequency of sampling for the data collection process should be documented. Disclosing this type of information may allow for more accurate selection of datasets for comparison, or even possible extrapolation of missing data. This would include the following aspects:

- **Snap length:** The value of how many bytes are recorded for each packet should be reported. Larger telescopes, or those recording large volumes of traffic, may elect to only capture the first *n* bytes of a datagram, rather than all bytes as seen 'on the wire'. It is worth noting that in its default mode of operation, tcpdump uses a *snaplength* (capture size per packet) of 64 bytes which is sufficient to capture the Ethernet header, IP header and the appropriate transport layer headers for TCP/UDP/ICMP, but not any actual payload, although portions of ICMP and UDP data may be present. The volume of data captured has direct bearing on the storage requirements and the potential flexibility of the data in future research. An alternate approach is that used by the University of Michigan, which calculates the MD5 checksum of payloads, and only stores the payload if it is not already present in the data store as described in [9].
- **Capture Interval:** This metric is a statement of how frequently data is sampled and stored from the sensor system. The CAIDA backscatter project [10], [11], for example, produces a set of captures for one week of every month. Although the interval may be driven largely by the storage requirements, interval sampling is different from a network telescope in its simplest form, which captures on a continuous basis. Another important item of information is detail of how the actual packets are sampled. In the case of data collected by the University of Wisconsin, only every 10^{th} packet was recorded for later analysis [12].

D. Noise Suppression and Filtering

It may be to the benefit of other researchers for publishers of network telescope datasets to disclose if there is any known noise suppression or other filtering that is performed by the sensor. Traffic originating due to service discovery, infrastructure management or monitoring can potentially generate huge volumes of traffic which may well skew results obtained in later analysis. In addition there may well be a number

of known misconfigured devices, or even active poisoning of the data-set traffic which can be regarded as noise and thus removed. This could be performed either at capture time through the use of suitable filters or as part of post-processing prior to archiving or further analytical processing. There may also be the case where certain ports, or even protocols, are missing from captures due to filtering by upstream providers or through choice. An example of this could be a sensor system dedicated to investigating and monitoring SSH scanning activity. In this case all that may be relevant is traffic destined for 22/tcp and ICMP Type 8 (ping) packets. The suggested preferred means of expressing filters used is in the commonly used BPF syntax used by libpcap and many other common tools for working with packet traces and captures. This should be augmented by suitable notes where clarification may be needed. When publishing filters, care should be taken not to disclose potentially sensitive information relating to address space, which could result in future pollution of the sensor.

E. Meta-data

There are various other pieces of relevant meta-data relating to the network telescope which are worth recording and including in published descriptions of the network telescope. The first two of these deal with the location of the network from a physical (geographic) and topological perspective. While it is recognised that there is some need for keeping the exact specifics of the network location hidden, it is useful for researchers publishing datasets to be able to provide these details below.

- **Geographic Region:** The geographic region that a network resides in can have a bearing on a number of factors, such as the available bandwidth and the amount of network address space available. In the case of the latter, North American and European organisations traditionally have relatively large IP address allocations while those in developing countries have been much more restricted. The regions could be described in general, or preferably at the level of the country or possibly even countries that a sensor may operate in. This can also be of use when measuring local geographic bias of traffic that may occur.
- **Topological Location:** Similarly, it may be of use to other researchers to disclose which major networks provide upstream Internet connectivity for the network sensor. Sites with better peering to Tier-1 ISPs may receive better coverage of traffic. Networks could be described through the use of the name of the upstream organisation (e.g. Sprint, AT&T, and JANET) or via the registered Autonomous System (AS) Number(s) allocated to the provider by IANA through the Regional Internet Registries (RIRs).

Contact information for the group or person(s) operating the network telescope and/or publishing the data is important to enable queries regarding the datasets produced by the sensor in question. Sensible information would also include the organisation (and group if appropriate). The final two items of meta-data would be if the telescope/sensor is still operational or not, and the date that the information was last updated. In the case of a sensor that is no longer in operation, the dates of operation could be stated.

IV. Dataset metrics

Dataset metrics would be used to describe a particular data-set as produced through the logging of packets on a network telescope or sensor network. A single data-set could be regarded as an aggregated collection of captures, or a single contiguous series of captures over a defined temporal period. The purpose of these proposed metrics is to be able to communicate the more useful and salient features of the data-set in a format that allows for easy comparison with other similar datasets, both from the same sensor, and ideally from other sources.

A. Top Items

The purpose of presenting these metrics is to provide an overview of the data-set,and highlight the significant components. In terms of ease of comparison, it is suggested that a relative scoring scheme be used in preference to raw numeric data (although both can be provided). Two possibilities are available, the first being a simple percentage. This would indicate, for example, that traffic destined to 445/tcp amounted to 41.4% of the observed packets in the period and 50.75% of TCP traffic. The advantage of this approach is it provides a basis for relative comparison without having to worry about scaling issues when dealing with sensors of differing lens sizes and configurations. A second method is to use an index based scoring scheme with some value set as a starting value, and all others normalised against this. Common choices for such values could include the values when monitoring was first started (which then provides some idea of the growth in observed traffic volumes over a period of time), or probably less useful, as it is problematic for inter-dataset comparisons, is an average score of sort. The use of such index based metrics is discussed in the following Section V where they are used for producing plots. Whichever of the above means of displaying the quantitative information is chosen, the final choice is what number of values should be presented — the value of N. While a full detailed analysis is useful, in practice the majority of useful information is contained in the top 10-20 items for port based traffic and that analysed at a network block level. When presenting lists of 'top N' items, the total percentage of traffic represented by the values should ideally be disclosed. Items that are likely to be useful as summaries and to other researchers are:

- **Top Hosts/Networks:** The top network sources observed as aggregated at differing levels of granularity (/8, /16, /24 and /32) can provide some indication as to the distribution of traffic.
- **Top Geopolitical:** The origins of the observed traffic. Bearing in mind that traffic can be trivially spoofed, these can provide some indication of hot-spots of malicious activity, or prevalence of malware. For example common wisdom in the network security community is that generally there is very little legitimate traffic originating out of countries such as South Korea (KR) for organisations not

dealing directly with clients there. There are published block lists[1,2,3] to allow for easy blocking of traffic from South Korea and of other countries. Most importantly in this section it should be disclosed how the geolocation was performed, generally describing the method and libraries/sources used to perform the correlation.

- **Top Topological**: While network level aggregation may have some use, when looking at the traffic from a topological perspective, it is possible to further aggregate networks into the groupings by Autonomous System (AS) number, whereby they are routed on the Internet by the BGP protocol. One may find, for example, that while there may be a small number of hosts coming from individual netblocks, when these are aggregated by AS number, certain organisations or network providers may show up as hot-spots.
- **Destination Ports:** These represent the targeted destinations of traffic and allow for the monitoring of emerging threats and trends in scanning and other malicious activity. Summary data should be produced for both TCP and UDP as relevant for the datasets. Any specific filtering (such as backscatter or active only) should be noted.
- **Source Ports:** While destination port traffic generally receives much attention, an analysis of the source ports can be interesting. For while the spread of these is much wider than with destination ports, certain ports do stand out. Source ports can generally be trivially controlled on the sending side and may be used for evading IDS systems or attempting to bypass firewall rules. An example of this is the very common use of 80/tcp as a source port. Source ports also reveal other interesting information when viewed from the perspective of backscatter, as they allow a sensor operator to ascertain to some extent possible DDoS activity which commonly involves spoofed addresses, which in turn are reflected back to the sensor. As with Destination ports, processing and reporting should be done for both TCP and UDP.
- **Protocols:** A breakdown of the composition of traffic by protocol is useful to other researchers when evaluating the data-set fitness for use in their own research. For example, the CAIDA backscatter datasets contain no UDP traffic, as it has been filtered out as part of the backscatter isolation process. While one would expect the common three protocols of ICMP, TCP and UDP to dominate, there may well be anomalies worth highlighting. It is suggested that traffic composition is reported as primarily percentages.

For all of these basic, statistical information such as minimum, maximum, mean, median and mode can be reported. Due care should be taken when working with these values as many of the distributions of the traffic are not normal, often tending towards Poisson or multi-modal distribution. More detailed statistical metrics than these are in most cases better included in separate reports rather than in a metric summary report. A partial summary of the metrics described above is shown in Table I which is a sample taken from the Rhodes University RUSCOPE1 data-set for traffic captured during July 2009.

[1]http://www.countryipblocks.net/
[2]http://www.blockacountry.com/
[3]http://www.ipdeny.com/ipblocks/

B. Temporal Aspects

The disclosure of period of time which the data-set covers is critical when making data available for comparative analysis. Date and timestamps should either be adjusted to UTC/GMT or if local time is used, suitable offsets from indicated (taking into account any local daylight saving). Along with the period captured, the duration should be noted along with the percentage of this time that the address space was actually monitored. Outages from both network and equipment perspectives do occur and it can be useful if these are recorded. It is suggested that a coverage score be disclosed as a percentage of when the sensor was actually active during the overall capture period. The granularity at which this was calculated should also be disclosed (days, hours, minutes, seconds). Of these units, days and hours are probably the more useful units of measure on which to base the calculation.

C. Active/Backscatter Ratio

Although methods to discriminate between active and passive or so called backscatter traffic [4], [13] are imperfect, particularly for UDP traffic (which is why collections such as CAIDA backscatter remove it completely), it can provide an interesting analysis; in particular what the ratio of active to backscatter traffic was. Ratios for TCP are relatively easy to calculate, with ICMP a clear cut issue. Traffic composition should be expressed as a percentage-wise ratio of the total traffic for the protocol under consideration.

V. Graphical Metrics and Temporal Sequences

This section builds on the metrics previously described in Section IV. In many cases being able to represent data in a graphical format provides a far more succinct means of conveying a large volume of information. The most common format for doing this is as a line plot with time being represented on the *x*-axis and volume or count on the *y*-axis. While maintaining the temporal progression on the *x*-axis, the granularity is markedly decreased, as the purpose is to convey trends rather than detailed information. To this effect raw numeric data is converted into index based representations, which allow for the plotting of comparative data on the same set of axes. The value of including the types of non-textual information described in this section is that they provide context in which the data can be interpreted. In the case of large bundled datasets, they can show the inherent trends within the data, which in many cases may be the information of greatest value to other researchers.

The issue of the granularity at which data is plotted needs to be considered. In most network datasets, the level of granularity goes down to sub-second accuracy. For most practical evaluation, particularly over longer time spans, a temporal bin size of days or weeks may be preferable, particularly if trends are being looked at rather than the minutiae of the data. As

Table I
EXAMPLE DATASET METRICS FOR RHODES UNIVERSITY TELESCOPE: JULY 2009

Protocols	Ports	N	Top 10	Source Netblocks	
TCP	Source	56 406	6000 80 25511 5641 6667 3306 25521 6005 10000 4496	/32	703 855
	Destination	9 844	445 135 22 1433 2967 5900 139 25 80 4899	/24	250 937
UDP	Source	13 808	1859 1231 3106 1029 3302 2373 1102 53 4659 4150	/16	1 4135
	Destination	2 683	1434 137 38293 5060 33435 33436 33437 33438 22105 7548	/8	181
ICMP	Type:Code		8:0 11:0 3:3 0:0 8:204 8:74 3:10 8:196 8:7 8:225	Total Packets	2 426 940
Traffic Composition			TCP 95.453% UDP 3.985% ICMP 0.561%		
Geopolitical	Regions	191	"CN" "RU" "BR" "US" "TW" "IT" "DE" "RO" "KR" "IN"		

the granularity decreases, the jitter factor when plotting data becomes less, however, the overall shape of the graph however remains the same.Appropriate levels of granularity should be chosen to convey data that is deemed to be important or significant, particularly when a reader may not have access to the raw data-set for further exploration. Summary numeric data should ideally be provided along with such plots, to allow for more detailed analysis and comparison.

A. Index plots

An index plot provides a useful means of plotting multiple data series onto the same set of axes, particularly where the raw volumes represented may differ by orders of magnitude. While not intended to provide detailed information, it serves to convey trends and facilitates comparison of these trends of different data series relative to each other. An example is shown in Figure1. In this plot, the raw values for each series are mapped to be relative to a starting index of 100, although depending on the level of variance, other values such as 1 000 or more could be used. This is the same as the system commonly used to denote stock market performance. In this case the base values used in the calculation were those for 2005 Q3 were 114, 8 626 and 106 012 for networks aggregated respectively by /8, /16 and /24 masks. As can be seen these values differ quite substantially, and would be difficult to plot with any type of discernible meaning on a standard line graph even when using logarithmic scaling. The remainder of the data is calculated as a score relative to the starting index as in:

$$Index_{T=N} = (Count_{QuarterN}/Count_{Q1}) \times 100.$$

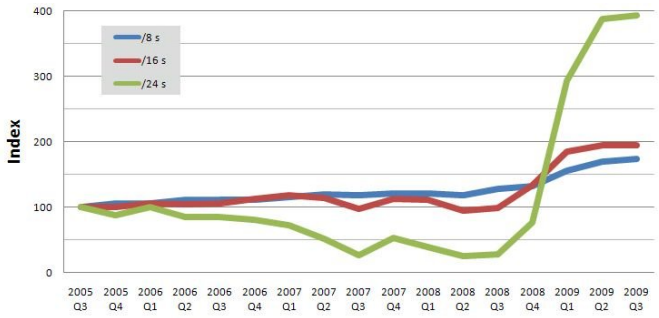


Figure 1. Traffic by Network Size

What is apparent from Figure 1 is that by the end of 2009 Q3 more than four times as many distinct /24 networks were being observed quarterly than four years previously, while twice as many /16 blocks were being observed. The effects of the Conficker Worm in late 2008 [14], [15], [16] can be clearly seen in both this Figure, and in Figure 2, which provides a similar index based plot broken down by protocol. Figure 2 also shows a clear drop in the real volume of ICMP traffic, ending on an index of 24.92. Viewing these two Figures in conjunction, one can observe a widespread increase in TCP activity. This activity is most likely attributable to the Conficker Worm.

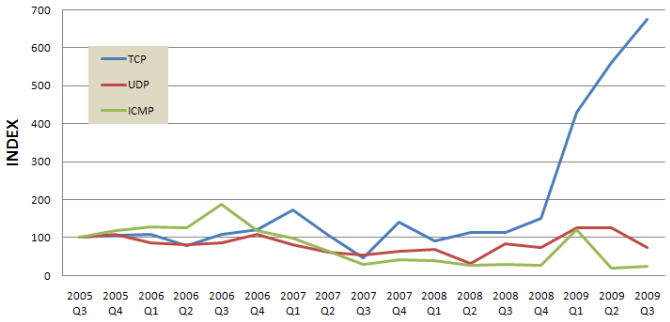
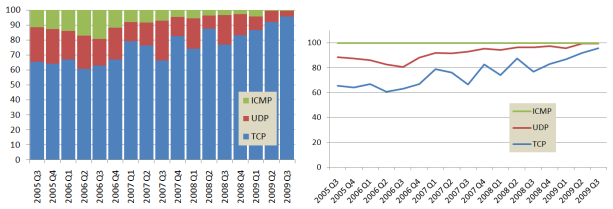


Figure 2. Traffic by Protocol

B. Proportional plots

As discussed in Section IV, there is merit in sometimes representing the data in terms of percentage contribution of the parts to the whole. The use of percentage representation is similar to the index plots above but is a far coarser, yet often serves as a better understood means of communication. The intention of the proportional plot format discussed in this Section is to provide an overview of relative constitution of a data-set sample. Examples of this are shown in Figure 3 which illustrates two common methods of representing the percentage composition of the observed traffic from the perspective of the three primary constituent protocols of ICMP, TCP and UDP. In both cases, the sub figures illustrate the contribution of the various series to the whole. Depending on the data being displayed, either the bar or line formats may be more appropriate. These should be seen to be complementary to the index based plots. Although they also allow for a comparison of the relative contribution of each protocol to the traffic composition, it is at a fairly low level of granularity. Comparing Figure 3 with Figure 2, one can observe that the proportion of ICMP traffic observed is decreasing, most likely due to the large increase in TCP traffic in the last two plotted quarters. In real terms the volume of ICMP traffic is dropping too as shown by its lower index value in Figure 2.



(a) Bar Plot (b) Line plot

Figure 3. Percentage composition of traffic

C. Sparkline plots

The final form of plot to be considered is that of the Sparkline, popularised by Edward Tufte [17]. This is a a simple 'word sized' line graph. Sparklines are most often used to show trends over time and are generally dimensionless in that there are no values attributed to the axes. One variation to this plotting scheme is that the current or most recent point may have a value attached along with high and low markers. The intention of these plots is to provide succinct trend information in a dense format. As a general rule this plotting technique is used only to show a single series per plot. Common uses of the Sparkline plot are in-line in text and in dashboard type information displays and reports, where trend information is of more value than specific point data. Examples of Sparkline plots produced from the RUSCOPE1 data-set are shown in Figure 4. Considering the similar shapes of Figures 4a and 4b, the significance of the contribution to traffic destined to 445/tcp to the whole can be seen.

(a) Total traffic (b) 445/tcp (c) 1434/udp

Figure 4. Sparkline plots: 3^{rd} August 2005 to 30^{th} September 2009

VI. Summary

This paper presents a set of guidelines and motivation for the use of a number of different metrics which can be used to suitably describe both the network telescope and data collected. The use of graphical summaries allows for trend information to be conveyed in a very succinct format. It is hoped that the adoption of metrics such as those described, will allow for easier comparison among multiple datasets. The publication of metrics may in some cases also obviate the need to have access to full packet captures — which have other issues such as privacy and anonymity to consider, quite aside from the potential problems of moving multiple gigabytes of data.

Acknowledgment

The author wishes to acknowledge the funding received from the Rhodes University JRC, NRF Thutuka Program Grant number 69018, and the Telkom Centre of Excellence in the Department of Computer Science, which was instrumental in carrying out this research.

References

[1] A. Jaquith, *Security Metrics: Replacing Fear, Uncertainty, and Doubt*. Addison-Wesley Professional, April 2007. ISBN 978-0321349989.

[2] G. Conti, *Security Data Visualization: Graphical Techniques for Network Analysis*. No Starch Press, 1st ed., October 2007. ISBN 978-1593271435.

[3] R. Marty, *Applied Security Visualization*. Addison-Wesley Professional, August 2008. ISBN 978-0321510105.

[4] D. Moore, G. Voelker, and S. Savage, "Inferring internet denial-of-service activity," in *In Proceedings of the 10th Usenix Security Symposium*, pp. 9–22, 2001.

[5] V. Fuller and T. Li, "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan." RFC 4632 (Best Current Practice), Aug. 2006.

[6] J.-P. van Riel and B. Irwin, "Identifying and investigating intrusive scanning patterns by visualizing network telescope traffic in a 3-d scatter-plot," in *Proceedings of 6th Annual Information Security South Africa (ISSA)*, (Balalaika Hotel, Sandton, South Africa), 5–7 July 2006.

[7] R. J. Barnett and B. Irwin, "Towards a taxonomy of network scanning techniques," in *Proceedings of the 2008 annual research conference of the South African Institute of Computer Scientists and Information Technologists on IT research in developing countries: riding the wave of technology*, SAICSIT '08, (New York, NY, USA), pp. 1–7, ACM, 2008.

[8] R. Barnett and B. Irwin, "A framework for the rapid development of anomaly detection algorithms in network intrusion detection systems," in *8th Annual Information Security South Africa (ISSA) Conference*, 6-9 July 2009. School of Tourism & Hospitality, University of Johannesburg, Auckland Park, Johannesburg, South Africa.

[9] M. Bailey, E. Cooke, F. Jahanian, J. Nazario, and D. Watson, "The Internet Motion Sensor: A distributed blackhole monitoring system," in *Proceedings of Network and Distributed System Security Symposium (NDSS '05)*, (San Diego, CA), Feb. 2005.

[10] D. Moore, C. Shannon, G. M. Voelker, and S. Savage, "Network telescopes," tech. rep., CAIDA, 2004.

[11] C. Shannon, D. Moore, and E. Aben, "The CAIDA Backscatter-2004-2005 Dataset - May 2004 - November 2005, (collection)." Online, , CAIDA Network Telescope Project - Backscatter, 2005.

[12] A. Kumar, V. Paxson, and N. Weaver, "Exploiting underlying structure for detailed reconstruction of an internet-scale event," in *Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement*, IMC '05, (Berkeley, CA, USA), pp. 33–33, USENIX Association, 2005.

[13] E. Cooke, M. Bailey, Z. Mao, D. Watson, F. Jahanian, and D. McPherson, "Toward understanding distributed blackhole placement," in *WORM'04 - Proceedings of the 2004 ACM Workshop on Rapid Malcode*, (Arbor Networks), pp. 54–64, 2004.

[14] E. Aben, "Conficker/Conflicker/Downadup as seen from the UCSD Network Telescope." Online, CAIDA Network Telescope Project - Backscatter, February 2009.

[15] Microsoft, "Virus alert about the Win32/Conficker worm (KB962007)." Online, August 18 2008. Last Review: December 1, 2010 - Revision: 10.0.

[16] P. Porras, H. Saidi, and V. Yegneswaran, "An analysis of conficker's logic and rendezvous points," tech. rep., SRI International, 4 February 2009. Last Update 19 March 2009.

[17] E. Tufte, *Beautiful Evidence*. Graphics Press, 2004.

Barry Irwin Has an interest in passive network monitoring with a particular focus on Network Telescopes and Honeypot systems. He heads the Security and Networks Research group at Rhodes University, and is co-chair of the South African Chapter of the Honeynet Project..