Greynets: A Definition and Evaluation ofSparsely Populated Darknets

Warren Harrop, Grenville Armitage
Centre for Advanced Internet Architectures, Swinburne University of Technology
Melbourne, Australia
{wazz,garmitage}@swin.edu.au

ABSTRACT
Darknets are often proposed to monitor for anomalous, externally sourced traffic, and require large, contiguous blocks of unused IP addresses - not always feasible for enterprise network operators. We introduce and evaluate the Greynet - a region of IP address space that is sparsely populated with 'darknet' addresses interspersed with active (or 'lit') IP addresses. Based on a small sample of traffic collected within a university campus network we saw that relatively sparse greynets can achieve useful levels of network scan detection.

Categories and Subject Descriptors
C.2.3 [**COMPUTER-COMMUNICATION NETWORKS**]: Network Operations - Network monitoring

General Terms
Management, Measurement, Performance, Security, Verification.

Keywords: Greynet, Network Security, Sparse Darknets, Darknet, Network Telescope, Intrusion Detection Systems

1. INTRODUCTION
A darknet is a large block of unused-yet-valid IP addresses monitored for inbound IP packets that have no business arriving at such addresses. Enterprise networks usually have unevenly distributed clusters of free IP addresses rather than large blocks, so we believe that enterprise-level darknets need a modified set of definitions. We thus introduce the concept of a *greynet* - that is, a region of network address space that is sparsely populated with 'darknet' addresses. This paper has two primary goals - to introduce a set of terminology for classifying and describing greynets, and to summarise our analysis of real-world data to show the efficacy of greynets of varying sparseness.

2. BACKGROUND
Intrusion detection systems do not require darknets to be effective - examples such as Bro [1][2] monitor traffic at central points in a network and use signature analysis of observed traffic flows between active IP addresses to identify likely intrusion candidates. By contrast, darknets detect speculative network scanning by making a key assumption – any and all packets heading towards a darknet IP address should not be doing so.

SIGCOMM'05 Workshops, August 22–26, 2005, Philadelphia, PA, USA.
ACM 1-59593-026-4/05/0008.

2.1 Previous Work
Much previous darknet research has focused on accurately inferring wider Internet activity [3] [4] [5] [7] [8]. The darknet is typically made as large as possible. Infection vectors and possible attacks are inferred from scan patterns. Denials of service attacks are inferred from back-scatter [9]. The Internet Motion Sensor (IMS) [5][6] distributes variably sized darknets (down to /24 nets) around the Internet. The 'spinning cube of potential doom' [10] [11] visualises darknet data by representing darknet space as a three dimensional cube. Scans reveal themselves by the patterns they form, easily recognisable by the human observer. There is currently work in the area of small darknet deployments and their effectiveness in a LAN context [12], focused on creating effective methods to report zero day worms as quickly as possible based on scanning patterns.

2.2 Enterprise and Campus Darknets
Internal darknets can provide an early warning of hosts launching hostile scans against other parts of an enterprise network. Scans and probes originating internally are of far greater concern to enterprise networks than externally originated probes, as the internal source is already inside the network's outer defenses.

3. Defining and Characterising a Greynet
Greynets are collections of non-contiguous blocks of IP addresses that are 'dark' in the classical darknet sense, but interspersed between groups of 'lit' IP addresses – active addresses belonging to real hosts on the enterprise network. Interspersing 'darknet' addresses among valid hosts makes it harder for malware to avoid hitting a greynet address while it searches for infection targets.

3.1 Terms and Definitions
Greynet: A mix of 'lit' (used) and 'dark' (unused) IP addresses.

Potentials: 'P', the set of 'dark' IP addresses that may potentially be monitored and are otherwise unused. The symbol 'P_m' represents a set of Potentials covering m actual (but not necessarily contiguous) IP addresses. P_m is a subset of the entire greynet, visualized as a set of IP addresses around the perimeter of a circle. (e.g. If P was 192.168.10/24 then 192.168.10.1 would be 'next to' both 192.168.10.2 and 192.168.10.255).

Listeners: 'L', the set of 'dark' IP addresses being monitored. 'L_n' represents n listeners, where $L_n <= P_m$ ($n <= m$, L_n is a subset of P_m fully contained within P_m). A greynet will have $L_n << P_m$. The greynet 'sees' those packets that head towards members of L_n.

Distribution of listeners: Members of L_n may be variously distributed throughout P. For example, a block of n contiguous addresses (type A) or n addresses spaced uniformly around the circumference of P (type B). We introduce the use of 'L_nX' to represent n listeners in distribution style X across the space P.

Orientation of listeners: A particular L_nX has a rotational orientation relative to P called θ. A set of listeners within P is fully described by coordinate (L_nX, θ).

Figure 1 illustrates this on a simple greynet where there are two listeners uniformly distributed (a type B) around the set P (of all potential listeners) and offset at an angle θ. We recognise that some combinations of Type B distributions, 'n' and 'θ' are redundant. For example, in Figure 1 θ=0 and θ=180 degrees create identically distributed sets of listeners.

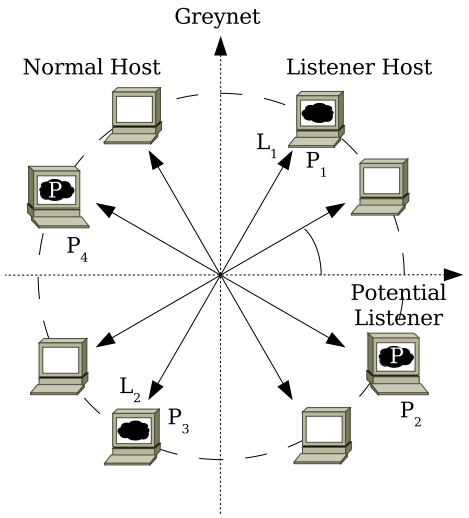

Figure 1. A greynet with an even distribution of potential and listener hosts within it, with offset θ

3.2 Establishing the efficacy of a Greynet

For each type of greynet (A and B) we wish to evaluate two things: how small we can make L_n while still seeing intrusion traffic 'quickly enough', and how our intrusion detection depends on the orientation of the listener set L_n. We define two metrics – median inter-event interval, and 'time to detect' (TTD). The former reflects the median time between seeing an inbound event for a particular (L_nX, θ). The latter (TTD) reflects the time difference between when a full darknet (L_n == P_m) would see a particular inbound event, and when a sparser greynet of listeners (L_nX, θ) would see an equivalent event.

The efficacy of a particular greynet structure can be found by evaluating median inter-event intervals and TTD as a function of greynet sparseness and orientation. We chose to analyse data gathered experimentally from a live darknet, selectively filtering the dataset to simulate various greynet configurations.

We used 238 contiguous 'dark' IP addresses open to the Internet and campus network, gathering traffic from June and September 2004. Two examples stand out. To detect sasser-infected hosts in less than 200 seconds we needed only 30 listeners, a decent interval for identifying (and optionally isolating) infected hosts. TCP scans that move linearly across the greynet space are detected very quickly (low TTD) with only small numbers of listeners. Even with only one listener the median TTD for externally sourced TCP scans was less than 2.5 seconds.

4. CONCLUSION

Greynets are 'dark' IP addresses sparsely distributed among 'lit' (active) IP addresses. We have introduced terminology to classify and define various types of greynet structures – Potentials, Listeners, distributions of Listeners and 'angle' of the Listener distribution. We fully describe a particular greynet with (L_nX, θ) notation.

A darknet with 238 addresses was operated for 3 months in 2004. With this real-world data we simulated the operation of greynets with less than 238 addresses and different address configurations. We demonstrated that two metrics, median inter-event interval, and 'time to detect', provide a good tool for evaluating the efficacy of particular greynet configurations relative to different types of network scanning patterns.

We finish by noting that greynet address assignment rules should be integrated into DHCP servers where applicable, and that VLAN-based enterprise networks could use a single host (running e.g. FreeBSD) on a VLAN trunk switch port to instantiate greynets simultaneously spanning multiple subnets.

5. REFERENCES

[1] "Bro: A System for Detecting Network Intruders in Real-Time", V. Paxson, Proceedings of the 7th USENIX Security Symposium, January 26-29, 1998

[2] "Bro", http://www.icir.org/vern/bro-info.html, August 2004

[3] D. Moore, C. Shannon, G. M. Voelkery, S. Savagey, "Network Telescopes: Technical Report", CAIDA, April 2004

[4] Telescope Analysis, http://www.caida.org/analysis/security/telescope/, April 2005

[5] M. Bailey, E. Cooke, "Tracking Global Threats with the Internet Motion Sensor", Nanog 32, September 7th, 2004

[6] University of Michigan Internet Motion Sensor, "http://ims.eecs.umich.edu/", April 2005

[7] E. Cooke, M. Bailey, Z. M. Mao, D. Watson, F. Jahanian, D. McPherson, "Toward Understanding Distributed Blackhole Placement", Conference on Computer and Communications Security, Proceedings of the 2004 ACM workshop on Rapid malcode, 2004

[8] The Team Cymru Darknet Project, "http://www.cymru.com/Darknet/", April 2005

[9] D. Moore, G. Voelker, S. Savage, "Inferring Internet Denial-of-Service Activity," 2001USENIX Security Symposium August 2001

[10] S. Lau, "The Spinning Cube of Potential Doom", LBNL Computer Protection Brown Bag seminar, Jan 2004

[11] S. Lau, "http://www.nersc.gov/nusers/security/TheSpinningCube.php", April 2005

[12] G. Gu et al, "Worm Detection, Early Warning and Response Based on Local Victim Information", ACSAC, December 2004