RC24433 (W0711-168) November 21, 2007
Computer Science

IBM Research Report

Identifying and Tracking Suspicious Activities through IP Gray Space Analysis

Yu Jin, Zhi-Li Zhang
University of Minnesota

Kuai Xu
Yahoo! Inc.

Feng Cao
Cisco Systems, Inc.

Sambit Sahu
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 704
Yorktown Heights, NY 10598

Research Division
Almaden - Austin - Beijing - Cambridge - Haifa - India - T. J. Watson - Tokyo - Zurich

Identifying and Tracking Suspicious Activities through IP Gray Space Analysis

Yu Jin
University of Minnesota
yjin@cs.umn.edu

Zhi-Li Zhang
University of Minnesota
zhzhang@cs.umn.edu

Kuai Xu
Yahoo! Inc.
kuai@yahoo-inc.com

Feng Cao
Cisco Systems, Inc.
fcao@cisco.com

Sambit Sahu
IBM Research
sambits@us.ibm.com

ABSTRACT

Campus or enterprise networks often have many unassigned IP addresses that collectively form IP gray space within the address blocks of such networks. Using one-month traffic data collected in a large campus network, we have monitored a significant amount of unwanted traffic towards IP gray space in various forms, such as worms, port scanning, and denial of service attacks. In this paper, we apply a heuristic algorithm to extract the IP gray space in our campus network. Subsequently, we analyze the behavioral patterns such as dominant activities and target randomness, of the gray space traffic for individual outside hosts. By correlating and contrasting the traffic towards IP gray addresses and live end hosts, we find the gray space traffic provides unique insight for uncovering the behavior, and intention, of anomalous traffic towards live end hosts. Finally, we demonstrate the applications of gray space traffic for identifying SPAM behavior, detecting malicious scanning and worm activities that successfully compromise end hosts.

Categories and Subject Descriptors

C.2.0 [**General**]: Security and protection (e.g., firewalls)

General Terms

Security

Keywords

IP Gray Space, profiling, network traffic analysis, entropy, anomaly detection

1. INTRODUCTION

In this paper we apply the novel notion of IP *gray space* analysis [1] to monitoring, identifying and tracking suspicious activities in a large campus network. IP gray space analysis is motivated by the observation that within a typical (large) campus/enterprise network which owns one or multiple address blocks (e.g., class B or /16 address blocks), not all IP addresses are likely to be assigned to "active" hosts (i.e., actual machines such as servers, desktops, laptops, etc.) at any give time period. We refer to these IP addresses within the campus network that are *not assigned* to any host throughout a given time period, say, an hour or a day, as "inactive" or *gray* IP addresses. In contrast, the IP addresses within the same address blocks that are assigned to hosts at any point within the time period are referred to as *active* IP addresses. The inactive IP addresses collectively form the so-called IP *gray space* within the address blocks, while active addresses the *active space*. By definition, IP *gray* and *active* space within a campus/enterprise network are time-dependent – in other words, they are not fixed and vary over time.

Unlike the well-studied IP "dark space" analysis techniques (see, e.g., [2, 3, 4, 5, 6, 7, 8, 9]) which are inherently *ex situ* and can potentially be evaded [10], IP gray space analysis is *in situ* and provides us with a direct means to monitor, identify and track suspicious and potentially harmful activities launched by outside hosts. In particular, we can observe the traffic generated by outside hosts towards both the IP *gray* space and *active* space of a network, and correlate them to infer the nature of activities engaged by the outside hosts and isolate potentially harmful ones. After all, it is live hosts (behind active IP addresses) that outside attackers are interested in! In [1] we developed a simple heuristic algorithm for extracting the IP gray space within a campus/enterprise network, and applied IP gray space analysis for dissecting and classifying various scanning activities.

Built on the study in [1], this paper focuses on the development of a novel three-step methodology for identifying and tracking potentially harmful hosts by correlating traffic towards both IP gray and active spaces of a campus network. Using the traffic towards the IP gray space, we first extract a candidate set of potentially suspicious outside hosts, and infer the dominant destination (or source) ports seen in the traffic towards the IP gray space that are used by an outside host for scanning, worm infection or other attack activities. Using these dominant ports, we then extract all the relevant scanning and other suspicious traffic towards both gray and active inside hosts. (This first step is described in more details in [1].) In the second step we zero in on and further extract *bad scanners* and *focused hitters* with likely harmful

activities (from the candidate set of potentially suspicious outside hosts) by analyzing the target "footprint" of dominant suspicious activities and correlating them with *other flows* generated by these hosts towards live hosts in active space. In most cases, these other flows cause security concerns, since they could reflect harmful activities from outside hosts, such as follow-up behaviors after successfully compromising an inside host. In the third step, we track all flows generated by these bad scanners and focused hitters before, during and after the (detected) dominant suspicious activities, and perform an in-depth analysis of these flows through a variety of means including inference using evidences obtained from other (possibly external) sources.

Using a month-long NetFlow data collected at our campus border router, our investigation reveals that i) many bad scanners successfully compromised inside end hosts, such as infecting them with new worms; ii) a number of bad scanners collects targets information for further exploit activities, e.g., sending spam messages after locating active SMTP servers or performing specific scanning after ICMP probing. For focused hitters, we rely on additional information and evidence, such as DNS lookup, behavior tracking, active probing to obtain a plausible explanation. An important finding is the prevalence of spam behaviors from outside hosts. The majority of focused hitters are outside hosts that persistently send spam messages to inactive mail servers as well as active mail servers. Through querying a widely-used spam database, we find nearly 80% of such hosts targeting SMTP ports are indeed spammers. We believe the remainder are also very likely spammers, since they share strong similarities with known spammers.

The remainder of this paper is organized as follows. Section 2 briefly introduces the concept of IP gray space and how we use IP gray space analysis technique to identify potential harmful hosts and further classify them into bad scanners and focused hitters. In Section 3, we examine the behavior, strategies and potential threats of bad scanners. Section 4 studies the behavior of focused hitters, in particular, the dominant spammer behaviors. Section 5 concludes this paper.

2. IP GRAY SPACE ANALYSIS

In this section we first briefly introduce the definition of *gray* IP address and our heuristic for obtaining the collection of gray IP addresses – collectively referred to as the (IP) *gray space* – for a given network. We then present our IP gray space analysis technique, which utilizes the characteristics of the IP gray space to identify potential harmful hosts.

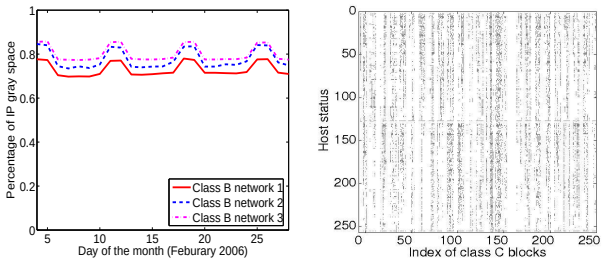
2.1 Identifying IP Gray Space

Let I denote the collection of all IP addresses of a network under consideration, and t_0 the starting time of a time period of interest, and T the length of the period. We say an (inside) IP address $g \in I$ is a *gray* (or inactive) address over the time period $[t_0, t_0+T]$ if and only if no traffic *originating* from g is observed during $[t_0-\tau, t_0+T+\tau]$ for some fixed τ [1]. We use G to denote the collection of all gray IP addresses within the time period, or *IP gray space*. The complementary set, $A = I - G$, is referred to the *active* space. In other words, for any $a \in A$, there is traffic originating from a at some time during $[t_0-\tau, t_0+T+\tau]$; thus a is likely assigned to an active host during the time period. In this study, we set T to be 24 hours, t_0 the zeroth hour of a day, and τ one hour.

[1]In this definition, to be conservative, we also require no traffic originating from g for a period of τ before and after the time period of interest to provide additional assurance that g is indeed unlikely to be assigned to any host over the said time period

2.2 Characteristics of IP Gray Space

We apply the above heuristic to the NetFlow data collected at the border router of the University of Minnesota campus network during February 2006. The data set includes all unsampled traffic flows between inside hosts and outside hosts during the entire month.



(a) IP gray space changes in one month

(b) Distribution of gray IP addresses

Figure 1: IP gray space properties

Our campus network owns three class B (1/16) IP address blocks, with a total $3 \times 2^{16} = 196608$ IP addresses. Among these many IP addresses, we found that in each day of Feb 2006 over 70% of the addresses are *gray* ("inactive") over the entire day (Fig. 1[a]). Fig. 1[b] illustrates the distribution of gray IP addresses in the 256 /24 address sub-blocks ("class C") within one of the class B (/16) address blocks. The x-axis represents each class C sub-block, while y-axis represents each host in a corresponding sub-block. A point on the graph stands for an active host on 02/06/2006. All the blank space belongs to IP gray space. We observe that the gray IP addresses are unevenly distributed among different /24 address sub-blocks.

Since no traffic is observed to *originate* from a gray IP address to any outside host (in the rest of the Internet) for an entire day, it is likely that the address is not assigned to any live host during that day. *Ideally* one would expect no traffic from any outside host either. This is in general not true at all. We observe that within four hours from the start (zeroth hour) of everyday, all gray IP addresses are "touched" (i.e., as the destination IP of an incoming flow) by at least one outside host! This is not surprising, because without *a priori* knowledge of which IP address is gray or active in a network, an outside attacker must perform some kind of reconnaissance activities such as scanning to identify vulnerable hosts or other targets of interest. Such activities, in particular, when done randomly, would result in touching the IP gray space with high probability given the large size of IP gray space in our campus network. Hence gray flows are a strong indication of likely suspicious activities, and the corresponding outside hosts that generate them warrant some scrutiny. To study the activities of the outside hosts generating gray flows, our methodology consists of three steps as explained below.

2.3 Extracting Outside Hosts Suspicious Activities

Our first step is to identify the ports appearing repeatedly in those gray flows from an outside host. Those ports, referred to as *dominant scanning ports* (DSP's), represent the likely services or exploits that the outside host is scanning for. With those DSP's, we then separate the scanning activities of the said outside hosts from other (if any) traffic from the same host: this is done by excluding any incoming flow from the outside host that does not use any of the DSP's as the corresponding source or destination ports (see [1]).

Let $\mathcal{O}_S$ be the set of outside hosts that we are interested in [2]. For any $h \in \mathcal{O}_S$, let $GF(h)$ denote the collection of gray flows generated by h. The destination ports (*dstPrt* in short) used by gray flows in $GF(h)$ induce an empirical distribution: for each *dstPrt* i, $p_i := m_i/m$ where m_i is the number of gray flows in $GF(h)$ with *dstPrt* i, and m is the total number of gray flows in $GF(h)$. We apply an information theoretical metric – *Relative Uncertainty* (RU) or *standardized entropy* [11] – defined below to the destination port distribution of h to identify dominant scanning (destination) ports (if they exist):

$$RU(dstPrt) := \frac{-\sum_{i \in dstPrt} p_i \log p_i}{\log m} \in [0,1], \qquad (1)$$

where $-\sum_{i \in dstPrt} p_i \log p_i$ is the entropy of the *dstPrt* distribution, and $\log m$ is its maximum entropy. We have $RU(dstPrt) \in [0,1]$. $RU(dstPrt)$ close to 0 suggests one or a few *dstPrt*'s dominate in the gray flows; while $RU(dstPrt)$ close to 1 signifies that there is no dominant *dstPrt*'s. Similarly, we can define $RU(srcPrt)$, for the source port (*srcPrt*) distribution of $GF(h)$. Hence $RU(srcPrt)$ and $RU(dstPrt)$ allow us to determine the existence of DSP's in the gray flows of an outside host, and if they exist, identify them using Algorithm 1 below.

Algorithm 1 Identifying Dominant Scanning Ports

```
Parameters: GF(h); β = β_0;
Initialization: DSP := ∅;
compute pro. dist. P_prt and θ := RU(prt) from GF(h);
while θ ≤ β and |GF(h)| >= 10 do
    find prt_i with highest P_prt_i;
    DSP := DSP ∪ prt_i;
    remove flows associate with prt_i from GF(h);
    remove P_prt_i from P_prt;
    compute θ := RU(prt) from GF(h);
end while
```

Algorithm 1 presents a heuristic procedure for extracting DSP's from either the destination or source port distribution P_{prt} of host $h \in \mathcal{O}_S$ (the same procedure applies to both *dstPrt* and *srcPrt*). The algorithm starts with an empty DSP set. It iteratively finds the port with the current highest probability, adds the port into DSP and removes all the flows associated with it from $GF(h)$. The algorithm terminates when there are not enough flows left ($|GF(h)| < 10$) or the ports in the rest of the flows are nearly uniformly distributed ($RU(prt) > \beta_0$, where we choose $\beta_0 = 0.7$).

Using the algorithm, we extract dominant destination and source ports for all outside hosts in $\mathcal{O}_S$. The DSP's (Fig. 2) include ICMP scanning (port 0), and well-known exploit (UDP/TCP) ports such as 137,139,445,1025 and 1434 as well as a few service ports such as 25,80,443.

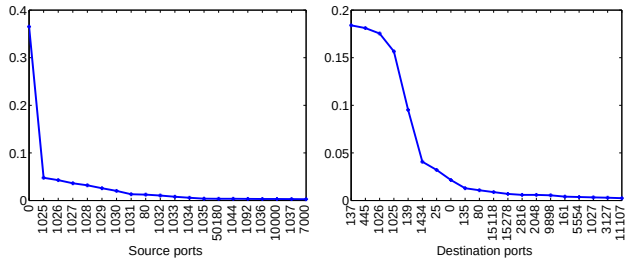


Figure 2: Top 20 DSP source ports and destination ports

Using those identified DSP's, we can separate incoming flows into two categories: *scanning flows* and *other flows*. Scanning flows are the flows associated with corresponding source or destination DSP's; while the remaining flows are considered as other flows.

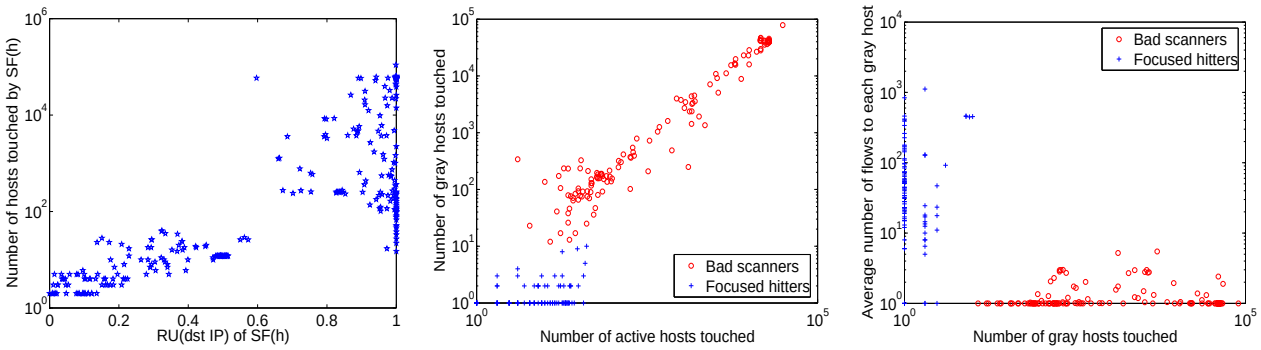
There are many reasons that a gray outside host produces other flows. In many cases, these other flows can be part of normal activities of the host, e.g. an outside host that interacts with some inside hosts normally could be infected by worms or other malware that generate the scanning flows. In other cases, these other flows may cause more concerns, as they could signify more harmful activities involving the outside host, such as follow-up activities after successfully compromising an inside host. It is these outside hosts with potentially harmful activities that we will focus our attention on in the remainder of this paper.

2.4 Zeroing In On (Potentially) Harmful Outside Hosts

In the third step we propose an effective method to pick out potential harmful outside hosts by correlating the scanning flows and other flows generated by these hosts. For $h \in \mathcal{O}_S$, let $A_s(h)$ denote the set of active IP addresses touched by the scanning flows of h (i.e., they appear as the destination IP (*dstIP*) addresses of the scanning flows in $SF(h)$. Let $A_o(h)$ denote the set of active IP addresses touched by other flows of h, namely, the set of *dstIP*'s in $OF(h)$. We include h in the candidate set of *potentially harmful* outside hosts, denoted by $\mathcal{O}_H$, using the following simple criterion: $h \in \mathcal{O}_H$ if and only if $A_s \cap A_o \neq \emptyset$, namely, there is an active IP address that is touched by both scanning flows and other flows. The intuition behind is that if activities embodied by the scanning flows and other flows are uncorrelated, they likely involve distinct inside hosts. When they touch the same inside hosts, it is highly probable that they are correlated, thus such hosts warrant additional special scrutiny.

Using the flow data of 02/06/2006, out of 7468 outside hosts in $\mathcal{O}_S$, there are only 284 outside hosts with scanning flows and other flows both touching the same active IP addresses inside our campus network, namely, $|\mathcal{O}_H| = 284$.

[2]We narrow our interest to those outside hosts with sustained suspicious activities, i.e. the outside hosts that generate at least 100 incoming flows over a day, and 10% of which are gray flows. Our analysis shows that a small portion of outside hosts generate a large portion of gray flows. For example, on 2/6/2006, although only 2% of the outside hosts generate more than 100 flows, of which 10% touching the IP gray space, they contribute to 98% of the total gray flows



(a) No. of unique `dstIP`'s hosts vs. $RU(dstIP)$

(b) No. of unique active hosts vs. gray hosts

(c) No. of unique gray hosts vs. avg. number of flows to each gray host

Figure 3: O_H classification and the properties of bad scanners and focused hitters

In the following, we examine the behavior characteristics of these hosts from several perspectives. First we investigate these hosts based on the suspicious flow activities, in particular, the following two questions: i) how many inside hosts (gray or active) are touched by the scanning flows of each host in $\mathcal{O}_H$? and ii) how varied are the inside hosts being touched, e.g., whether each host being touched once or repeatedly? The first question tells us the size of the "footprint" of the suspicious activities of the outside host, while the second provides some clue on the nature of the suspicious activities.

For the first question, we count the number of distinct inside IP addresses (i.e., *dstIP*'s) in $SF(h)$ for each $h \in \mathcal{O}_H$. For the second question, we measure the *relative uncertainty* (RU) of the *dstIP*'s of flows in $SF(h)$, $RU(dstIP)$, defined using eq.(1), where for each *dstIP* i, p_i equals the relative frequency (or probability) of i (the number of flows in $|SF(h)|$ with i as *dstIP* divided by $|SF(h)|$). Fig. 3[a] plots the number of distinct *dstIP*'s vs. $RU(dstIP)$ for each $h \in \mathcal{O}_H$. Except for one outlier, the points on the plot are grouped together. By applying a simple clustering algorithm, we obtain two separate clusters. One cluster consists of the points in the left lower corner which have low $RU(dstIP)$'s, and touch a small number of inside hosts. The points of the right (upper) corner form the second cluster, which have a relatively large number of distinct *dstIP*'s and high $RU(dstIP)$'s. The outlier is caused by a mixed behavior in which an end host randomly performs ICMP probing, while simultaneously sends legitimate web traffic. We refer to the first group as *focused hitters* and the second *bad scanners* (while the third *mixed intruders*). There are totally 101 focused hitters, and 150 bad scanners. We will focus mostly on these two groups.

To further understand the behaviors of focused hitters and bad scanners, we examine the number of gray and active hosts that they touched as well as the number of flows sent to gray hosts. Figs. 3[b] compares the number of unique active hosts vs. gray hosts touched by $SF(h)$. Clearly, the points of the bad scanners are around the line, $y = 0.7x$, where 0.7 is the average gray ratio of the 3 class B networks (the total number of gray addresses divided by the total number of active addresses in the campus network). It indicates that the bad scanners randomly choose the $SF(h)$ targets. On the other hand, the focused hitters have much low gray ratios as they only access a small number of gray hosts. In addition, Fig. 3[c] shows the difference between the number of unique gray hosts accessed by $SF(h)$ vs. the average number of $SF(h)$ towards each gray host. For bad scanners, they send one or a few flows to each gray hosts, which are the typical patterns of scanners. However, for focused hitters we find that they tend to send tens or hundreds of flows to each gray host, which indicate that these focused hitters are likely unaware of the destinations becoming gray hosts. In the next two sections, we will investigate these two groups in depth.

3. BAD SCANNER ANALYSIS

In this section, we investigate the activities of the bad scanner group identified earlier and try to assess the potential threats they pose to the campus network by correlating the scanning and other flow activities as well as resorting to evidences gleaned from additional sources (such as DNS records, active probing). We use the flow data of 02/06/2006 as an example to illustrate the results. Based on analysis of the scanning flows and other flows to the common (inside) IP's that a bad scanner touches, we can roughly divide the 150 bad scanners of 02/06/2006 into the following subgroups for each of which we perform a more in-depth analysis below:

i) The first sub-group includes bad scanners that employ ICMP probes in their dominant scanning activities, and upon receiving responses to the ICMP probes from live hosts (from the active space), they follow up with TCP/UDP scanning activities (seen in their other flows). On 02/06/2006, 48 (33%) out of 150 bad scanners belong to this sub-group, and nearly all of them are searching for well-known service ports such as ports 22, 25, 80. By examining these further scanning activities, we find that these bad scanners receive few successful responses from live inside hosts. This is not surprising given that most hosts inside the campus network are client machines. For a few inside hosts (servers) that respond to the scanning activities, no further interactions are observed. Although the bad scanners of this sub-group do not seem to pose any immediate threat, they still warrant

to be tracked for possible future activities.

ii) A bad scanner in this sub-group scans using TCP/UDP probes on a variety of ports, many of which are exploit or service ports; furthermore, after responding to the TCP/UDP probes, a few live inside hosts in return *initiate* an ICMP ping or a TCP connection request on port 113 (the IDENT protocol) to which the scanners respond back. Out of 88 bad scanners in this sub-group, 77 receive an ICMP ping in return upon being scanned from a total of 32 inside hosts. Comparing with the other active inside hosts, these 32 inside hosts are being scanned by the bad scanners more than 8 times on average, compared to the average of 1-2 times for the other active inside hosts. Furthermore, these active inside hosts are being scanned on a variety of ports including service, exploit, and high TCP/UDP ports, to which they all respond successfully. The corresponding names of these 32 inside active IP's (via reverse DNS lookup) reveal that they are names for DHCP assigned machines (in dormitory and other non-departmental subnets); none of them are servers (although some respond to port 25, 80 when scanned on these ports). Further inspection of activities from both these inside hosts and the bad scanners that touch them seems to suggest that the client machines behind these 32 inside IP addresses may be infected with malware, or even part of a botnet. We are still conducting on-going analysis and tracking of these hosts (both inside and outside).

The other 11 bad scanners in this sub-group trigger the inside hosts scanned to initiate a TCP connection request on port 113. All of them are scanning for service ports 22, 23 and 25. Analyzing the DNS names of the inside hosts (a total of 804 hosts) that initiate the TCP port 113 connection requests in turn, they are Unix or Linux machines in computer labs in Computer Science or other engineering departments, which are configured with the IDENT protocol to "identify" any user who accesses these ports. No other flows (except for the response to the IDENT request) are initiated from these bad scanners, suggesting that they are not able to pursue any further activities. Most interestingly, for the 3 outside hosts that scan for port 25, a query of their IP addresses in the spammer database [12] reveals that they are listed as known spammers.

iii) The third and last sub-group include 11 bad scanners that also scan using the TCP/UDP probes and receive responses from some live inside hosts; furthermore, they also have *other* TCP/UDP connections with these live inside hosts that are *initiated by them*. Correlating the scanning activities with other activities (as indicated by the flows in $OF(h)$), we find that most of the other connections initiated by these bad scanners occur *after* the scanning activities (SF flows) – suggesting possible follow-up activities; while a couple of them occur beforehand, as possible precursor activities. In one case the scanning activities occur during a series of connections between a bad scanner and a live inside host. Detailed investigation of this latter case suggests that a hacker from Romania has likely broken into this inside hosts, make a series of connections on TCP ports 4489, 17864, and 80, which lasts about 9 minutes. In the midst of these suspicious activities, he/she also launches a TCP port 80 scanning which also touches the inside host he/she has broken into. In the two bad scanners with the precursor activities, they are performing queries to an inside DNS server, and then launch scanning for TCP port 80. The remaining bad scanners in this sub-group are engaged in some kind of follow-up activities. Further investigation of the scanning and subsequent follow-up activities of these bad scanners strongly suggests that these bad scanners have successfully compromise or infiltrate some inside hosts. For example, one bad scanner from Japan performs sequential scanning on TCP port 445, UDP port 1023, TCP port 5554 and TCP port 9898; 96 inside hosts respond to the TCP port 9898 scanning. The bad scanner then proceeds with TCP connection attempts with these hosts on TCP port 8967, with successful connections with 6 of them. Such traffic patterns exactly match the signature of the recent Sasser worm [13]. In another case, a bad scanner scans on UDP port 38293, and upon receiving responses, follow up with connections on various UDP ports as a part of worm infection process.

4. FOCUSED HITTER ANALYSIS

We now turn our attention to the focused hitter group. We first examine the DSP's of the focused hitters and separate them into sub-groups of likely similar behaviors. We then perform an in-depth analysis of the focused hitters in each sub-group, often relying on additional information and evidences from other sources, and thereby attempt to infer the nature of activities these focused hitters are engaged in – in particular, obtaining a plausible explanation for why a focused hitter touches the gray IP addresses.

In general, we find that DSP's of focused hitters typically belong to a small number of applications, especially, SMTP, Web and peer-to-peer. For example, using the 02/06/2006 data, out of the 101 focused hitters, 69 target the destination port 25, namely, attempting to access email servers, while while 12 target the web service ports, 80 and 443, and 3 target the destination ports such as 6881 (BitTorrent) and 6364 (Gnutella) that are typically associated with peer-to-peer applications, 4 targets X windows service port 6000, while the rest of 13 focused hitters target various high ports.

We first perform an in-depth analysis of the biggest sub-group, the 69 focused hitters that attempt to access email servers. There are two likely explanations for why these outside hosts repeatedly access the IP gray space: the inside hosts (email servers) are temporarily down or the outside hosts are spammers. The 69 focused hitters access a total of 19 gray IP addresses, each touching multiple gray addresses, often repeatedly throughout the day. We perform a DNS look-up (with MX option) of these 19 inside gray IP addresses: all of them have a legitimate DNS record and the associated MX record indicates that they are email servers, although only a few have "mail" in their DNS names. Further investigation on flow data reveals that except for one, 18 stay *gray* throughout the entire February, indicating that they are likely old email servers that have been taken out of service, however their DNS records have not been updated. The persistence of the 69 focused hitters in accessing these likely out-of-service email servers strongly suggests that they are likely email spammers who have harvested DNS records for email servers. To confirm this, we query a widely-used online spammer database [12], and find that 48 out of these 69 outside hosts are listed as known spammers in the database!

For the remaining 21 hosts (referred to as *unknown hosts* hereafter), we conduct a detailed *comparative* analysis between these hosts and the known spammers such as DNS MX records, active port 25 probing, patterns of unique gray and active IP addresses touched, email traffic temporal fre-

Table 1: Feature similarity of known spammers and unknown hosts

Type	no.	MX records	open port 25	receive SMTP traffic	avg. active hosts accessed	avg. gray hosts accessed
Spammers	48	20	28	7	12.1	1.4
Unknown hosts	21	13	18	7	16	1.6

quency and reciprocity (i.e., whether they receive email traffic from inside hosts) analyses. Some of the results are summarized in table 1. The third column ("MX records") shows that 20 of 48 known spammers are listed as email servers based on DNS MX record look-up, while 13 of 21 unknown hosts are. The fourth column ("open port 25") shows that 28 of 48 known spammers accept `telnet` connection probing on port 25, while 18 of 21 unknown hosts also do. 7 of 48 known spammers receive SMTP traffic from inside hosts, while 7 of 20 unknown hosts also do (Column 5 "receive SMTP traffic"). The last two columns show the average active and gray IP's touched by known spammers and unknown hosts. These comparative analyses lead us to believe the 21 unknown hosts are very likely spammers that have not been included in the spammer database.

For the 12 hosts targeting web service ports (80, 443), we also perform a DNS look-up of the gray IP addresses touched. Surprisingly, all of the gray hosts correspond to dynamic hostnames containing "dial-up", "wireless" or "dhcp". A detailed analysis on the traffic of these gray hosts before or after 02/06/2006 reveals that unlike typical web servers these hosts provide temporally web service for a short time duration. There are two possible explanations for this behavior. First, these hosts are likely compromised and controlled by outside attackers who occasionally turn on web services for file sharing or communications. Second, these hosts might provide web services in small communities. For both cases, we believe the behaviors of focused hitters are likely suspicious and warrant further inspections. Analysis of the hosts targeting the peer-to-peer ports suggests that the outside hosts are peer-to-peer clients with stale peer cache, attempting to access gray IP addresses that were at one time dynamically assigned to inside users running peer-to-peer applications (incidentally the address blocks touched by these outside hosts belong to the student dormitory subnet.) Study of the remaining focused hitters targeting on various destination ports seems to suggest that those hitters have anomalous behaviors.

We have done similar analyses using flow data from other days. Again we find that except for a number of peer-to-peer clients that repeatedly touch the gray IP addresses (due to stale peer cache), the majority of focused hitters tend to be email spammers, active attackers. Thus using our approach, we are able to zero in on these "bad hosts" and thus subsequently monitor and track them for potentially harmful activities.

5. CONCLUSIONS

With the IP gray space information extracted by applying the heuristic algorithm in [1], we develop a noval three-step approach for identifying and tracking potentially harmful hosts by correlating their traffic towards both gray space and active space. Using one-month traffic data collected in a large campus network, we find i) many outside hosts successfully compromised inside end hosts, such as infecting them with new worms; ii) a number of scanners collects targets information for further exploit activities, e.g., sending spam messages after locating active SMTP servers or performing specific scanning after ICMP probing, and iii) hundreds of spammers persistently send email traffic towards inactive mail servers as well as active mail servers during every day.

Acknowledgement

The project was supported in part by NSF grants CNS-0435444 and CNS-0626812, a University of Minnesota Digital Technology Center DTI grant, a Cisco gift grant and an IBM Faculty Partnership Award.

6. REFERENCES

[1] Y. Jin, G. Simon, K.Xu, Z.-L. Zhang and V. Kumar. Gray's Anatomy: Dissecting Scanning Activities Using IP Gray Space Analysis. In *Proc. of the Second Workshop on Tackling Computer Systems Problems with Machine Learning Techniques (SysML'07)*, 2007.

[2] D. Moore, C. Shannon, G. M. Voelker, and S. Savage. Network Telescopes. Technical report, CAIDA, 2003.

[3] M. Bailey, E. Cooke, F. Jahanian, J. Nazario, and D. Watson. The Internet Motion Sensor: A distributed blackhole monitoring system. In *Proc. of Network and Distributed System Security Symposium*, 2005.

[4] E. Cooke, M. Bailey, F. Jahanian, and R. Mortier. The Dark Oracle: Perspective-Aware Unused and Unreachable Address Discovery. In *Proc. of the 3rd ACM/USENIX Symposium on Networked Systems Design and Implementation (NSDI'06)*, May 2006.

[5] Team CYMRU. The Darknet Project, June 2004. http://www.cymru.com/Darknet/.

[6] V. Yegneswaran, P. Barford, and D. Plonka. On the Design and Use of Internet Sinks for Network Abuse Monitoring. In *Proc. of Symposium on RAID*, 2004.

[7] V. Yegneswaran, P. Barford, and V. Paxson. Using Honeynets for Internet Situational Awareness. In *Proc. of the ACM/USENIX Hotnets IV*, November 2005.

[8] R. Pang, V. Yegneswaran, P. Barford, V. Paxson and L. Peterson. Characteristics of Internet Background Radiation. In *Proc. of ACM SIGCOMM IMC*, 2004.

[9] The Honeynet Project, 2005. http://www.honeynet.org.

[10] J. Bethencourt, J. Franklin and M. Vernon. Mapping Internet Sensors with Probe Response Attacks. In *Proc. of 14th USENIX Security Symposium*, 2005.

[11] K. Xu, Z.-L. Zhang and S. Bhattacharyya. Profiling Internet Backbone Traffic: Behavior Models and Applications. In *Proc. of ACM SIGCOMM*, August 2005.

[12] Real-time Spam Black Lists. http://openrbl.org/.

[13] ENSRT Incident Note ETS-i-2004-10860. http://www.enterasys.com/support/security/incidents/2004/05/10860.html.