On the Impact of Dynamic Addressing on Malware Propagation

Moheeb Abu Rajab Fabian Monrose Andreas Terzis
Computer Science Department
Johns Hopkins University
{moheeb,fabian,terzis}@cs.jhu.edu

ABSTRACT

While malware models have become increasingly accurate over the past few years, none of the existing proposals accounts for the use of Network Address Translation (NAT). This oversight is problematic since many network customers use NAT in their local networks. In fact, measurements we collected from a distributed honeynet show that approximately 19% of the infected hosts reside in NATted domains. To account for this fact, we present a model that can be used to understand the impact of varying levels of NAT deployment on malware that spread by preferentially scanning the IP space. Using this model, we show that NATting impedes malware propagation in several ways and can have a significant impact on non-uniform scanning worms as it invalidates the implicit assumption that vulnerable hosts reside in densely populated subnets.

Categories and Subject Descriptors

D.4.6 [**Operating Systems**]: Security and Protection—*Invasive Software*

General Terms

Security, Measurement

Keywords

Network Security, Internet worms, Network Address Translation, Private Address Space

1. INTRODUCTION

The research community has been on a quest over the past several years to discover ways to accurately capture the spreading behavior of malware on the Internet. Understanding the intricacies of such behavior continues to be an important problem because the resulting insights are invaluable when designing and evaluating malware countermeasures. Indeed, analysis of past outbreaks has lead to a deeper understanding of malware dynamics and the findings have already been incorporated in a number of analytical models (*e.g.*, [8, 15, 20]).

WORM'06, November 3, 2006, Alexandria, Virginia, USA.

However, all of the models that have been presented thus far assume that the infection views on both sides of a network boundary are identical. Unfortunately, the widespread deployment of firewalls coupled with the use of Network Address Translation (NAT) severely distort these two views, and can lead to inaccurate model predictions. In this paper, we explore the influence of NAT on the spreading of malware that use non-uniform and localized scanning to spread. Our exposition is based on a refined model that incorporates the fact that many vulnerable hosts are deployed in private address spaces.

To gauge the impact of address translation, we first estimate the number of infected sources located in private address spaces by analyzing traces collected from a conglomeration of network telescopes. As we show, dynamic addressing is a fairly common practice — approximately 19% of the sources in our trace reside in NATted domains. The model we develop shows that, at this level of usage, address translation techniques introduce significant skew in the prediction capabilities of existing malware spreading models. These predictions will increasingly depart from reality as NAT usage grows.

The rest of the paper is organized as follows: In Section 2 we elaborate on the impact of NAT on malware infections and the challenges it creates for accurate forensic analysis. Section 3 presents our data collection efforts and the methodology we use to infer the prevalence of NATted sources. In Section 4 we provide the analytical model and use it to examine the impact of varying levels of NAT deployment on malware spreading in Section 5. We present related work in Section 6 and conclude in Section 7.

2. OVERVIEW

It should come as no surprise that the use of private address space and network address translation techniques influences how malware spreads. First, NAT devices reduce the percentage of vulnerable hosts that are globally reachable. The reason is that these devices block connection attempts that originate from the outside by default, thus protecting internal vulnerable hosts from external infections. Even when port forwarding is enabled — usually to allow specific services to be accessible from the global Internet — only a subset of the potentially vulnerable hosts is visible to external malware scans. Second, when a new host inside a private address space is compromised, NATting affects how efficiently this host can find other vulnerable hosts. This is especially true for malware that spread through preferential scanning, including *non-uniform scanning* (*e.g.*, CodeRed-II [6], Nimda [7], and MSBlaster [11]) and *localized scanning*, in which infected hosts (predominantly) scan their local address prefix. Recently, Rajab *et al.* [17] showed that localized scanning is widely used by botnets, and hence models that capture localized behavior may become increasingly important in

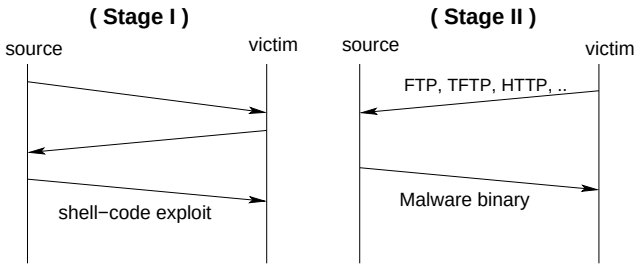

Figure 1: A Multi-stage Malware infection.

the near term.

The mere fact that NATted hosts are usually located in large address spaces (*e.g.*,10/8, 192.168/16) causes preferential scanning malware to divert the majority of its scans towards the NATted space rather than the globally routable IP space. While the infected machine can still contaminate other vulnerable hosts within the private address space, locating these hosts can take a prohibitively long time. This slowdown in infection speed arises because the density of active hosts within private address spaces is orders of magnitude lower than the host density in the global address space. This is certainly the case when a private /8 address prefix (*e.g.*, 10/8) is used. Networks that use /16 private address spaces induce another interesting behavior; preferential scans from infected hosts in those networks will not only target the NATted space, but also contact the encompassing routable /8 prefix. The net effect is that these parts of the IP space will receive a disproportionate percentage of scans by several kinds of malware [1]. While this creates an attractive measurement hot-spot as reported in [4, 9], the increased traffic is an annoyance to the networks operating in that prefix.

The use of NAT poses another obstacle to malware that employ a multi-stage infection process. This multi-stage infection process, shown in Figure 1, is a common occurrence in botnets [17]. In the first stage, a vulnerability that is remotely exploitable is used to transfer a shellcode that instructs the victim to initiate a connection back to the infector's IP address to download the actual malware binary. The download constitutes the second stage of the exploit and usually occurs through a file transfer protocol such as TFTP. If however the infector is located behind a NAT device then the provided address points to a globally unreachable IP address, thereby causing the second-stage transfer to fail.

Aside from slowing the spread of malware, NATting poses several challenges to forensic analysis of malware [13, 16]. These challenges are related to the difficulty of uniquely identifying NATted hosts in the absence of explicit information (*e.g.*, [3, 4]). On one hand, a group of infected hosts behind a NAT device with a single public address will appear at a network monitor as a single infected host thereby under-estimating the number of infected hosts. Conversely, few hosts behind a NAT device with a large number of external addresses can inflate the estimation because subsequent scans from the same infected host will most likely be mapped to several source addresses as they are re-written by the NAT device. Shannon *et al.* conjectured that this was indeed the case for a set of addresses observed in the Witty worm outbreak [18].

In the next section we derive an initial estimate of the prevalence of NAT in malware traces. In Section 4 we quantitatively analyze the impact of NAT on the spreading of different classes of malware.

[1]For example, *all* locally addressed Code Red II victims that use 192.168/16 addresses send half of their scans to the 192/8 prefix.

3. ON THE EXTENT OF NAT USAGE

Estimating the prevalence of malicious (or infected) sources that use NAT is a challenging task in its own right. Fortunately, inferring whether or not an infected source uses address translation can be relatively easy for certain types of malware. As a case in point, Casado *et al.* [4] showed that one can detect NAT usage by leveraging the information from malicious traffic traces captured at carefully located distributed monitors. The authors specifically exploited the scanning behavior of CodeRed II to detect infected hosts residing behind NAT devices that use the 192.168/16 prefix. Based on the observation that CodeRed II sources send 50% of their scans to the encompassing 192/8 prefix. Casado *et al.* inferred that more than 60% of the sources, in their 48-hour darknet traces, were NATted.

Similarly, by exploiting the fact that the Witty worm [10] used a fixed source port (4000) to send its packets, it is fairly straightforward to detect Witty victims residing behind NAT devices as these devices rewrite the packets' source port [18]. Using the Witty worm packet traces obtained from CAIDA [5], we extracted all sources that sent Witty packets with source ports other than (4000). We observed that from roughly 60,000 unique Witty source IP addresses, 4,643 ($\approx 7\%$) had their source ports re-written.

While these results indicate that the use of NAT is fairly common, neither the result of Casado *et al.* [4] nor that derived from the Witty dataset can be used to reliably infer a global estimate of NAT usage[2]. Doing so would require both longer monitoring period and more diverse vantage points (especially as Code Red II used a non-uniform scanning strategy). More importantly, the technique used to gauge the prevalence of NAT usage should not be tied to any specific vulnerability.

In what follows, we provide another estimate of NAT usage by examining malware collection logs captured at a distributed honeynet platform. While we make no claim that our approach leads to a closer approximation of the global ratio of NAT usage, we believe our preliminary results are more general than those presented elsewhere as they do not suffer from the previously outlined shortcomings.

3.1 Inferring NAT usage

Our approach is based on inferring the presence of NATted hosts from malware traces captured at a number of distributed active responders. In particular, we deployed a modified version of the Nepenthes malware collection tool [2] to a /24 prefix in our institution and 14 smaller monitors running on PlanetLab nodes [14] with access to darknet space (covering from 4 to 12 IP addresses). Nepenthes emulates a number of vulnerable services and collects exploits sent via these services.

As noted earlier, the initial exploit in multi-stage infections is likely some form of MS-Windows shellcode containing a URL that hosts the malware binary. In most cases, this URL points back to the source that sent the exploit in the first stage. We can therefore determine which sources are located behind NAT devices by parsing the log of collected URLs and extracting those sources that use local IP addresses in the URL sent to the victim. Over a period of one month, we observed a total of 14,651 unique sources that initiated first stage exploits in the local /24 network. Among them, 2,782 ($\approx 19\%$) were NATted. Furthermore, the distributed nodes observed 3,850 malicious sources during a monitoring period of one week, 710 ($\approx 18.5\%$) of which used local addressing. These results provide further evidence that NAT usage is fairly common.

[2]Casado *et al.* themselves acknowledged that the inferred NAT ratio did not include hosts that use 10 /8 and 172.16 /12 addresses and is not generalizable to the overall NAT usage on the Internet.

4. ANALYSIS OF THE IMPACT OF 'NAT' ON MALWARE PROPAGATION

We present a model that predicts the evolution of malware infections, accounting for the effect of NAT deployment in the Internet. The proposal is an extension to a model we previously developed to study the impact of vulnerable population distributions on Internet infections [15]. We consider the general case in which malware instances apply preferential scanning, using different probabilities to locate and exploit victims in their surrounding /16 and /8 prefixes as well as random scanning to find victims in the global Internet.

We account for the effect of NATting by dividing the vulnerable population into two categories: *(i)* the publicly reachable vulnerable population including vulnerable hosts with public IP addresses in addition to NATted vulnerable hosts which are however publicly reachable (*e.g.*, due to port forwarding), and *(ii)* the vulnerable population that resides behind NAT devices and is inaccessible from the public Internet.

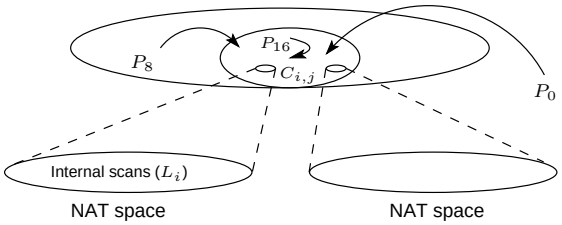

Figure 2: The incoming scanning activity to a single /16 prefix with NATted domains.

Figure 2 illustrates the malware preferential scanning activity that reaches a routable /16 prefix containing a number of NATted domains. The number of incoming scans in this case is simply the sum of the scanning components from infected hosts within that prefix (indicated as P_{16} in Figure 2), from infected hosts within the encompassing /8 prefix (indicated as P_8), and from the entire infected population (the P_0 component). Observe that in the case of the NATted infectees, the encompassing prefixes will be those of the private address rather than their external routable space. As a result, preferential scans from these hosts will be diverted towards private (un-routable) space. For this reason, the number of incoming scans into each routable /16 prefix excludes any preferential scanning activity originating from NATted hosts. Using the notation from Table 1, the sum of the three scanning components above can be written as:

$$C_{i,j} = P_{16}s(I_{i,j} - \mathcal{N}_{i,j}) + \frac{P_8 s(I_{i,j}^{(/8)} - \mathcal{N}_{i,j}^{(/8)})}{2^8} + \frac{P_0 s I_i}{2^{16}} \quad (1)$$

in this case, P_{16}, P_8, P_0 are the probabilities that an infected host will send a scan to the encompassing /16, /8 prefix, and the entire Internet, respectively [3]. $I_{i,j}$ is the number of infected hosts within the j^{th} /16 prefix at time i; $I_{i,j}^{(/8)}$ is defined similarly for the surrounding /8 prefix. $\mathcal{N}_{i,j}$ is the total number of infected NATted hosts that are publicly reachable within the j^{th} /16 prefix, and $\mathcal{N}_{i,j}^{(/8)}$ is the total number of infected NATted and reachable hosts in the /8 prefix surrounding the j^{th} /16 prefix.

[3]For example, in the case of a host infected with Code Red II, $P_{16} = 0.375$, $P_8 = 0.5$ and $P_0 = 0.125$.

I_i	Total number of infected hosts at time i
s	Average scan rate per infected host
P_0	Probability of scanning a random address
P_8	Probability of scanning an address within the same /8 prefix as the infectee
P_{16}	Probability of scanning an address within the same /16 prefix as the infectee
V_j	Initial number of vulnerable and reachable hosts in the j^{th} /16 prefix
$I_{i,j}$	No. of infected hosts in the j^{th} /16 prefix at time i
$C_{i,j}$	Total number of incoming scans into the j^{th} /16 prefix at time i
T_j	Total number of NATted networks within the j^{th} routable /16 prefix
L_i	Total number of scans within a particular NATted network at time i.
f	Initial number of vulnerable hosts in a particular NATted network
d_i	Number of infected hosts in a particular NATted network at time i

Table 1: Infection Model Notation.

The $C_{i,j}$ scans will infect members of the first population category. The expected number of infected hosts in the j^{th} /16 prefix at time $i+1$ is then equal to the number of infected hosts in the previous interval plus the new infections due to scans that reached vulnerable hosts[4]. This is expressed as:

$$I_{i+1,j} = I_{i,j} + (V_j - I_{i,j})\left[1 - \left(1 - \frac{1}{2^{16}}\right)^{C_{i,j}}\right] \quad (2)$$

in which, V_j is the initial number of vulnerable hosts in the j^{th} /16 prefix.

In addition to the infections due to the scanning activity in the public IP space, infected hosts within NATted domains will infect other vulnerable hosts within the same private space, including vulnerable hosts from the second population category (*i.e.,* publicly inaccessible vulnerable hosts)— assuming, of course, that no internal countermeasures, such as "hard-LANs" [19], are locally deployed.

If we consider NATted domains that use /8 private addresses[5] and assume, for simplicity, that hosts in these private spaces are co-located in the same /16 private address prefix, then the number of scans within that network domain can be written as:

$$L_i = sd_i\left(P_{16} + \frac{P_8}{2^8} + \frac{P_0}{2^{16}}\right)$$

in which, d_i is the number of infected hosts within a given NATted network. Therefore, the number of additional infections within a single private address space can be expressed as:

$$d_{i+1} = d_i + (f - d_i)\left[1 - \left(1 - \frac{1}{2^{16}}\right)^{L_i}\right] \quad (3)$$

in which, f is the initial number of vulnerable hosts in a given NATted network.

Then, the total expected number of infected hosts at time step $i+1$ is simply the sum of the infected hosts in all 2^{16} /16 prefixes,

[4]To isolate the impact of NAT we do not consider the node removal rate due to patching or failure.

[5]As we show in Section 5, using NAT with /16 address space (*e.g.* 192.168 /16) causes the P_8 scan component from *all* NATted infected hosts to target a single prefix (*e.g.* 192 /8).

Number of Vulnerable hosts	632,472
Average scanning rate (s) per infected host	10 scans/sec
Size of initial Hit List	100
Local domain size	/24
Number of publicly accessible hosts per NATted domain	1
Number of runs	10 per experiment

Table 2: Analysis parameters.

including the infections that occur within all NATted networks that are part of each /16 prefix:

$$I_{i+1} = \sum_{j=1}^{2^{16}} \left(I_{i+1,j} + \sum_{l=1}^{T_j} d_{i+1,l} \right) \quad (4)$$

in which T_j is the number of NAT domains in the j^{th} /16 prefix. Notice that the private space is usually sparsely populated so the infection rate within that space is substantially slower than the global infection rate. In the next section, we show that this effect reduces the overall propagation speed of malware and could have a considerable impact as NAT deployment increases.

Finally, the observant reader will note that for malware that uniformly scans the entire IP space, the address of any particular infected machine does not impact its scanning behavior. The only impact of NATting in this case is that it decreases the reachable vulnerable population. Therefore, for the remainder of the paper we only evaluate the impact of address translation on the evolution of malware that use preferential scanning.

5. EVALUATION

We now make use of the model presented in Section 4 to evaluate the impact that NAT has on malware spreading. As we have previously shown in [15], analytical models must use realistic vulnerable population distributions if they are to accurately model the behavior of worm outbreaks. For this reason, we drive our evaluation with a vulnerable population distribution extracted from a real dataset. In particular, the dataset is provided by DShield [12] and contains intrusion traces collected over a period of three months from over 1,600 intrusion detection systems distributed around the globe. Given that the logs were obtained from IDS reports, it is safe to assume that they represent unwanted traffic originating either from compromised hosts or active scanners. We construct a vulnerable host set by extracting the sources that attempt connections to port 80[6]. Overall, the data contains 632,472 such sources.

We emulate the impact of NAT by segmenting the set of vulnerable hosts into different network domains. For simplicity, we assume that all domains reside in equal sized /24 public address prefixes. We acknowledge that this is not necessarily the case in the Internet today and different domain sizes can alter the rate of malware evolution[7]. Incorporating more realistic domain size distributions is part of our ongoing work.

We assume that each NATted network has one vulnerable host that is publicly accessible. This reflects common network administration practices in which a number of hosts behind a NAT device are made accessible so that certain services (*e.g.,* web servers) are publicly available. The remaining vulnerable hosts are unreachable by external scans and can therefore only be contaminated by scans from infected hosts within the private network. We consider varying degrees of NAT deployment by assuming different percentages of domains that use private NAT spaces.

[6]We assume in this case that these sources are infected and trying to spread the infection using the same vulnerability.

[7]For example, malware local spreading will be faster within larger and more densely populated NATted networks.

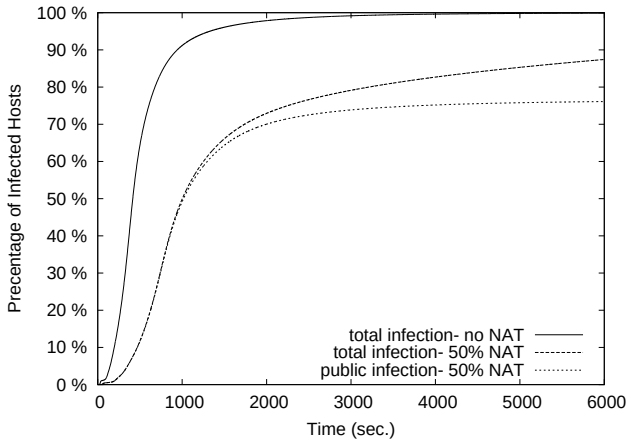

Figure 3: Non-uniform scanning worm spreading when no NATting is used and when 50% of vulnerable hosts are located in private address spaces.

5.1 Non-uniform Malware Spreading

We first present results for non-uniform scanning worms by simulating a Code Red II worm using the parameters shown in Table 2. Figure 3 illustrates the general impact of NAT on the spreading of the worm. The graph compares the worm propagation when there is no NATting versus when 50% of the domains are NATted. As the graph indicates, NAT generally slows down worm propagation. This slow down is primarily caused by the fact that *(1)* NAT shrinks the overall vulnerable population exposed to global worm scans. This effect is evident in Figure 3, in which infection from public scans reaches saturation before the worm is able to infect the entire vulnerable population, and *(2)* NAT expands the effective search space of the worm within the perimeter of the NATted network domain. Networks that employ private address spaces are more sparsely populated compared to the routable public space and so it takes longer to find and infect the next vulnerable host within a private network. As shown in Figure 3, the overall added infection resulting from the infected NATted hosts grows at a much slower rate compared to the global infection rate.

Next, we gauge the impact of different levels of NAT deployments on the spreading of the worm. We consider several NAT deployments including the actual percentage observed in our traces ($\approx 20\%$). For this spectrum of deployments, we examine two cases: first, we assume that NATted domains use 10/8 private addresses and second, we consider NATted domains using the 192.168 /16 private address space. Figure 4 graphs the evolution of the worm for the first case under varying degrees of NAT deployment. Taken together, these results show that NAT significantly affects the model's predictions, and must be taken into account to reflect the behavior of the worm accurately. Moreover, these results confirm our intuition that the deployment of NAT acts as an impediment to malware spreading.

Our results for NATted domains that use /16 private address spaces (omitted due to space constraints) show very similar trends to those in Figure 4. Intuitively, one would expect that malware propaga-

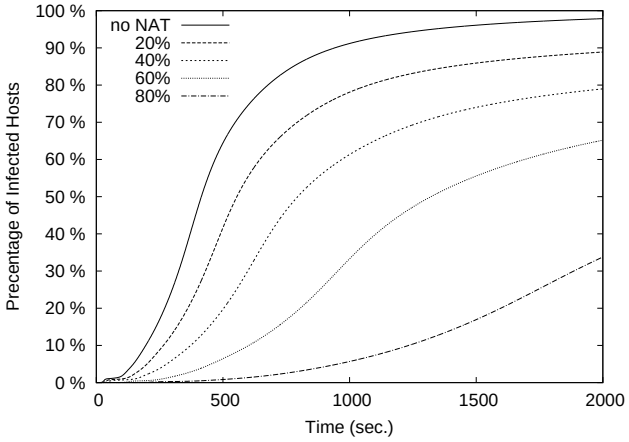

Figure 4: Non-uniform scanning worm spreading for different levels of NAT deployment.

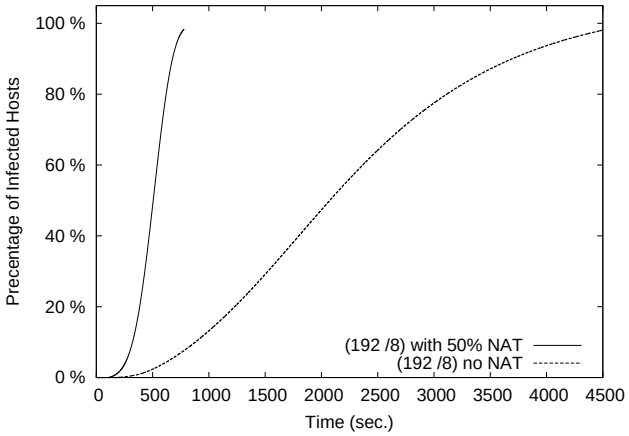

Figure 5: Non-uniform scanning worm spreading in the 192/8 prefix in the case of zero and 50% NATted domains.

tion in this case would be somewhat faster, because the /8 random scanning component will be released to the public Internet. However, because all these scans will be directed at a single /8 prefix (namely, 192/8), the overall increase in the speed of the worm is minimal. That said, a disproportionate percentage of scans originating from all NATted infected hosts will consequently target the 192/8 prefix. The resulting outcome is that the worm propagates much faster in that prefix compared to the rate of spread observed in other parts of the IP address space (see Figure 5).

5.2 Localized Malware Spreading

As mentioned earlier, localized scanning, in which each infected host scans its local address prefix, represents an important infection vector in botnets [17]. Therefore, it is also important to understand the impact of network address translation on the spread of these malware strains.

Figure 6 represents the spreading behavior of botnets that scan the encompassing /8 prefix of each infected host using the parameters listed in Table 2. It is evident from the graph that the infection spread is slower than the non-uniform case. This can be explained

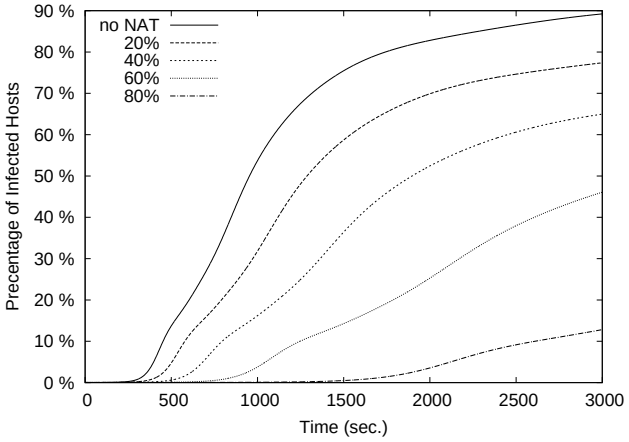

Figure 6: Spreading behavior of malware using /8 prefix localized-scanning under varying degrees of infected NATted hosts.

by the fact that unlike non-uniform scanning worms, the localized scanning malware has no "island hopping" component that allows the infection to move across different prefixes. As a result, malware instances uselessly scan the same prefix after all its vulnerable hosts have been infected. More importantly, the impact of address translation is amplified in this case since NAT devices completely contain the scanning activity within the perimeter of the sparsely populated private networks.

Finally, for malware classes that spread via the multi-stage infection process illustrated in Figure 1 (*e.g.,* botnets), NAT poses another obstacle; regardless of the scanning technique used by the malware, an infected host behind a NAT device will not succeed in transferring the malware binary to a new infectee outside the network perimeter. Therefore, we conjecture that increasing NAT deployment will impede botnets that spread by active scanning.

6. RELATED WORK

Worm models have undergone a series of refinements over the past few years, leading to increasingly accurate representations of worm behavior in the wild. For example, Zou *et al.* presented a "two-factor" worm model that extended the classical epidemic model to account for the removal of infected hosts (due to patching or failure) and demonstrated how accounting for that factor more accurately reflects the infection dynamics of Code Red I [20]. Chen *et al.* subsequently presented the "AAWP" model which was the first attempt to model non-uniform scanning worms [8]. More recently, Rajab *et al.* demonstrated the significance that the distribution of vulnerable hosts has on the spreading of non-uniform scanning worms and presented an extended model that accounts for this factor [15]. However, none of these models account for the skew introduced by NAT and evaluate its impact on malware spreading.

The development of techniques for reliably detecting hosts behind a NAT device remains an open problem. Bellovin [3] presented a technique to count the number of hosts behind a NAT device by exploiting the evolution sequence of the `IP_ID` field in the outbound packets. Shannon *et al.* pointed to the potential skew introduced by NAT and DHCP and its subtle implications in the analysis of the spreading behavior of the Witty worm [18]. More recently, Casado *et al.* suggested the possibility of using unwanted

traffic to measure important Internet-wide characteristics [4]. The authors showed that one can infer NAT usage by studying the scanning behavior of Code Red II sources captured by carefully located distributed network monitors. Similar insights were also noted by Cooke *et al.* who showed that the non-uniformity in the scanning behavior of infected hosts, due in this case to flaws in the worm's random number generator and side-effects of NATting, can create worm "hot-spots" [9]. Our work complements these efforts by exploring another avenue for estimating NAT usage by examining malicious traces and studying the failed-connection rate of multi-stage infections.

Finally, some distantly related work is that of Antonatos *et al.* that illustrated the potential of address space randomization to protect against hit-list worms by continuously changing the IP addresses of active hosts [1]. Our work, on the other hand, is focused on illustrating the overall impact of NATting as an impediment to malware spreading, and we argue that it is an important factor that must be considered in modeling non-uniform malware spreading.

7. SUMMARY

In this paper, we show that the widespread use of network address translation has significant implications on how different families of malware spread on the Internet. Using analytical modeling, we quantitatively show that NATting acts as an impediment to the propagation of malware that spread by preferentially scanning the Internet. This effect is due to the fact that NAT effectively increases the address space that active scanners must explore. Moreover, NATting decreases the density of the vulnerable host population residing in network domains that use private address space and in doing so, negates the advantage that non-uniform scanning provides. Finally, we note that the use of NAT causes multi-stage infections to fail at a high rate since the URLs transmitted in these infections hold private network addresses that are unreachable from the public Internet.

Acknowledgments

This work is supported in part by National Science Foundation grant CNS-0627611. We thank DShield and CAIDA for graciously providing access to their IDS logs and Witty Worm dataset, respectively. We also extend our gratitude to the anonymous reviewers for their insightful comments.

8. REFERENCES

[1] S. Antonatos, P. Akritidis, E. P. Markatos, and K. G. Anagnostakis. Defending Against Hitlist Worms Using Network Address Space Randomization. In *WORM '05: Proceedings of the 2005 ACM workshop on Rapid malcode*, pages 30–40, 2005.

[2] Paul Baecher, Markus Koetter, Thorsten Holz, Maximillian Dornseif, and Felix Freiling. The Nepenthes Platform: An Efficient Approach to Collect Malware. In *Proceedings of the 9^{th} International Symposium on Recent Advances in Intrusion Detection (RAID)*, to appear Sept. 2006.

[3] Steven M. Bellovin. A Technique for Counting NATted Hosts. In *Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurment (IMW)*, pages 267–272, 2002.

[4] Martin Casado, Tal Gankel, Weidong Cui, Vern Paxson, and Stefan Savage. Opportunistic Measurement: Extracting Insight from Spurious Traffic. In *Proceedings of the 4^{th} ACM Workshop on Hot Topics in Networks (HotNets-IV), College Park, MD*, November 2005.

[5] Colleen Shannon and David Moore, The CAIDA Dataset on the Witty Worm - March 19-24, 2004, See `http://www.caida.org/passive/witty/`.

[6] CERT. Code Red II: Another Worm Exploiting Buffer Overflow in IIS Indexing Service DLL. See `http://www.cert.org/incident\_notes/in-2001-09.html`.

[7] CERT. Nimda Worm Advisory CA-2001-26. See `http://www.cert.org/advisories/ca-2001-26.html`.

[8] Zesheng Chen, Lixin Gao, and Kevin Kwiat. Modeling the Spread of Active Worms. In *Proceedings of IEEE INFOCOMM*, volume 3, pages 1890 – 1900, 2003.

[9] Evan Cooke, Z. Morley Mao, and Farnam Jahanian. Hotspots: The Root Causes of Non-Uniformity in Self-Propagating Malware. In *Proceedings of the International Conference on Dependable Systems and Networks DSN*, pages 179–188, November 2004.

[10] Symantec Corporation. Witty worm advisory. See `http://securityresponse.symantec.com/avcenter/venc/data/w32.witty.worm.html`.

[11] Symantec Corporation. MS Blaster Worm Advisory, W32.Blaster.Worm. See `http://securityresponse.symantec.com/avcenter/venc/data/w32.blaster.worm.html`, August 2003.

[12] The Distributed Intrusion Detection System (DShield). See `http://www.dshield.org/`.

[13] Abhishek Kumar, Vern Paxson, and Nicholas Weaver. Exploiting Underlying Structure for Detailed Reconstruction of an Internet-scale Event. In *Proceedings of ACM SIGCOMM/USENIX Internet Measurement Conference (IMC)*, pages 351–364, October 2005.

[14] Larry Peterson, Tom Anderson, David Culler, and Timothy Roscoe. A Blueprint for Introducing Disruptive Technology into the Internet. *SIGCOMM Computer Communication Review*, 33(1):59–64, 2003.

[15] Moheeb Abu Rajab, Fabian Monrose, and Andreas Terzis. On the Effectiveness of Distributed Worm Monitoring. In *Proceedings of the 14^{th} USENIX Security Symposium*, pages 225–237, August 2005.

[16] Moheeb Abu Rajab, Fabian Monrose, and Andreas Terzis. Worm Evolution Tracking via Timing Analysis. In *Proceedings of ACM Workshop on Rapid Malware (WORM)*, pages 52–59, November 2005.

[17] Moheeb Abu Rajab, Jay Zarfoss, Fabian Monrose, and Andreas Terzis. A Multifaceted Approach to Understanding the Botnet Phenomenon. In *Proceedings of ACM SIGCOMM/USENIX Internet Measurement Conference (IMC)*, to appear Oct., 2006.

[18] Colleen Shannon and David Moore. The Spread of the Witty Worm. *IEEE Security and Privacy Magazine*, 2(4):46–50, July 2004.

[19] Nicholas Weaver, Dan Ellis, Stuart Staniford, and Vern Paxson. Worms vs. Perimeters: The Case for Hard-LANs. In *Proceedings of the 12^{th} Annual IEEE Symposium on High Performance Interconnects*, pages 70–76, August 2004.

[20] Cliff Zou, Weibo Gong, and Don Towsley. Code Red Worm Propagation Modeling and Analysis. In *Proceedings of ACM Conference on Computer and Communication Security (CCS)*, pages 138–147, 2002.