Overall-transparent Dynamic Identifier-mapping Mechanism against Network Scanning in SINET

Linyuan Yao[1, a], *Ping Dong[1, b], and Hongke Zhang[1, c]

[1] School of Electronic Information and Engineering, Beijing Jiao Tong University, P. R. China

{[a]11111020, [b]pdong, [c]hkzhang}@bjtu.edu.cn

Keywords: Dynamic identifier-mapping, Network scanning, Smart identifier network

Abstract. Static assignment of *ip* addresses or identifiers can be exploited by an adversary to attack a network. However, existing dynamic *ip* address assignment approaches suffer from two limitations, namely: participation of terminals in the assignment and inadequate network server management. Thus, in this paper, we propose an overall-transparent dynamic identifier-mapping mechanism to manage the identifier of network nodes to defend against scanning in the smart identifier network. We establish the selection and allocation constraints. The non-repetition probability and cover cycle allow us to evaluate the defense efficiency against scanning. Simulation results and theoretical analysis show that the proposed method effectively reduces the detection probability of routing identifiers.

1. Introduction

The static assignment of *ip* addresses or identifiers (*ip*s/*id*s) for a network device can be exploited by attackers to scan a network to inform his/her attack strategies. For example, when a device responds to the probe sent by the attacker, the adversary can easily obtain information such as *ip*s/*id*s of the target and conduct an attack, using, say distributed denial of service (*ddos*) [1].

Existing approaches to dynamic *ip* address assignments can be broadly categorized into the following two categories.

In the first approach, terminals help to complete the mutation process. Applications that participate in their own defense (*apod*) [2] were developed to defend against network-based attacks. The tunnel mode based on random mutation of the address and the port is used in *apod*. Network address space randomization (*nasr*) [3] operates by specifically hardening networks against hitlist worms. In other words, *nasr* is an address mutation method that updates the dynamic host configuration protocol. Both *apod* and *nasr* resist sniffing or worms only when the network server works collaboratively with the terminals. It is, thus, necessary for the terminals to participate in the whole process.

A second approach is to deal with the mutation process using the network without the terminals. For example, dynamically network address translation (*dnat*) [4] was designed to prevent man-in-the-middle attacks. In *dnat*, an *ip* address is translated before it is forwarded to a public network. The moving target defense technique in [5] enables host-to-*ip* binding for each destination host to vary randomly across the network based on spatial randomization and temporal randomization.

Random host mutation (*rhm*) and openflow *rhm* [6] are two mutation methods to host *ip* address in order to thwart scanning by the attacker. The mutation process is transparent to the terminal. However, the address of the terminal is changed randomly during the process.

In addition to designing solutions that are resilience to active attacks, it is important that the solutions allow us to effectively manage and secure the network server. In other words, there is a need for an efficient dynamic *ips*/*ids* management network-based mechanism for both terminals and network server in order to effectively respond to cyber-attacks.

Therefore, in this paper, we propose an overall-transparent dynamic identifier-mapping mechanism (*odim*) to manage the identifier of the network nodes. *Odim* is designed to defend against scanning in smart identifier network (*sinet*) architecture [7].The *odim* process is transparent to the terminals. During the mutation process, terminals do not take part in the communication, and no additional software or protocol is required to be installed on the terminals. Also, in our approach, the terminals and the network forwarding nodes, such as access router node (*arn*), and the mapping service node (*msn*), are managed. We then present a selection algorithm and an allocation algorithm to solve the constraints in *odim*. The non-repetition probability and cover cycle are proposed to evaluate the defense efficiency of *odim* against scanning.

2. Smart Identifier Network

Sinet comprises three vertical layers and two horizontal domains, as reported in [7]. The vertical layers are the smart pervasive service layer (*l-sps*), the dynamic resource adaption layer (*l-dra*), and the collaborative network component layer (*l-cnc*). The two domains are the entity domain (*d-en*) and the behavior domain (*d-be*). This paper only discusses one of the communication modes of *l-cnc*. We refer interested reader to [7,8] for other details about *sinet*.

As shown in Fig.1, the network topology can be divided into the access network and the core network in the *sinet*. Each node in the access networks connects to the core network via *arn* with a unique node identifier (*nid*). The *arn* is responsible for mapping the *nid* to the node-behavior description (*nbd*). One of the many attributes of the *nbd* is *rid* or the location of the node. In the core network, the *msn* is responsible for managing the mapping relationships. The *rid* and *nid* are used in the core network and the access network, respectively.

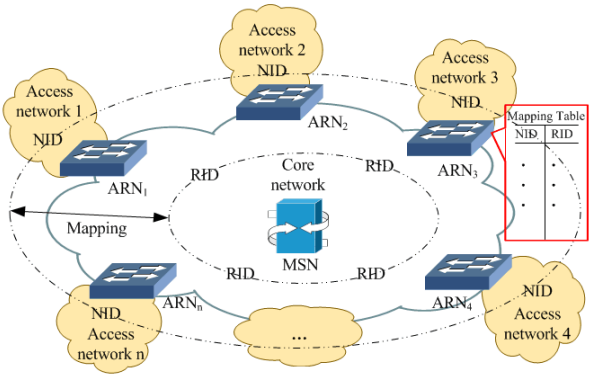


Fig.1 Diagram of data transmission

3. Overall-Transparent Dynamic Identifier-Mapping Mechanism

Problem Definition

The management of identifier-mapping is based on selection constraint and allocation constraint. The selection constraint is used to filter the *rid* mapping pool. The *rid* in the filtered mapping pool can

be used to map the *nid*s of the terminals and network nodes. The allocation constraint is responsible for matching the *nid*s of the terminals and network nodes with *rid*s in the filtered mapping pool.

1) Selection Constraints

Priority constraint: Prioritize *rid*s as in Table 1 and arrange *rid*s in order from high to low. If an *rid* has low priority, then it will be banned in the next mapping period. If the number of *rid*s with high and middle priority is less than the number of *nid*s that need to be mapped, then *rid*s with low priority will be activated and changed to middle-priority according to the order of being attacked.

Table 1 Priority list of rid

Priority	Monitored	Attacked
High	×	×
Middle	√	×
Low	√	√

Number constraint: The *rid*s are mapped to *nid*s for the terminals and the network nodes. The mapped *rid*s belong to different network segments. The number of *rid*s must be larger than the sum of the access segments and the core segment. An *rid* mapped to a terminal in an access segment must be in the same network segment. Meanwhile, the number of *rid*s in a network segment is no less than that the number of *nid*s. At least two *rid*s in every network segment are allocated to the network node.

2) Allocation Constraints

Uniqueness constraint: Each *rid* can be mapped to only one port in a terminal or network node in each mapping period.

Space consistency constraint: To reduce the size of route entry in the core network, an *rid* mapped to an access segment should belong to the same network segment. To separate the access and core networks, *rid*s mapped to terminals and network nodes must be in different network segments.

Difference constraint: Each new *rid* should be completely different from the old *rid*. For example, if the old *rid* belongs to network segment 2, then the new *rid* can belong to network segment 1.

Number optimization constraint: The chosen deviation between the number of *rid*s and required number of *rid*s should be the smallest over optional *rid* network assignments.

Basic Assumptions

The basic data transmission process is shown in Fig.1. There are n access networks, and each access network is connected to the core network by an *arn*. Access network 1 corresponds to arn_1. The terminal in an access network uses its *nid* to communicate. If a user in access network 1 sends packets to another user in access network 2, then the mapping process will occur in arn_1 and arn_2. When the data is forwarded from the access network to the core network, the *nid* is mapped to an *rid*. In contrast, the *rid* is mapped to an *nid*.

The main goal is to find the best *rid* for each *nid* to reduce the scanned probability. For ease of analysis, the assumptions are as follows.

1) Each *arn* is responsible for connecting one access network to the core network. For example, access network i corresponds to arn_i.

2) If the numbers of *arn*s and *msn*s are n and q respective, then $q \leq n$. When the size of the network is small, q is 1.

3) Each port of network nodes is allocated one *nid*. Each *nid* is mapped to one *rid* in each cycle.

4) The label of an access network is determined according to the number of *nid*s in the access network. The access network with the fewest *nid*s is access network 1.

5) The label of an *rid* network segment is determined by the number of *rid*s in the segment. The *rid* network segment with the fewest *rid*s is *rid* network segment 1.

6) The number of *rid*s in each *rid* network segment allocated to a network node is no less than 2.

4. Evaluation and Simulation

In this section, we present two evaluation metrics, non-repetition probability P and cover cycle T, to evaluate the defense efficiency of *odim* against network scanning.

Evaluation Metrics

The number of ports sniffed in the network node can be one or more. We first consider the situation where an adversary can sniff only one port in the network.

We set τ as the mapping cycle of *odim* and ϑ as the number of *rid*s sniffed in one cycle.

The non-repetition probability P is the probability that the *rid*s monitored during two adjacent mapping cycles are entirely different. The value of P is limited by the difference constraint as in Equation 1. Where $f(\vartheta)$, defined in Equation 2, denotes the times when all *rid*s sniffed in the old mapping will be mapped to new *nid*s in the new mapping. These are totally different from the *nid*s in the old mapping. N_{RID} is the total number of the available *rid*.

$$P=\frac{C_{N_{RID}-\vartheta}^{\vartheta}}{C_{N_{RID}-\vartheta}^{\vartheta}-f(\vartheta)}, \tag{1}$$

$$f(\vartheta)=\vartheta!\left(\frac{1}{2!}-\frac{1}{3!}+\frac{1}{4!}+\cdots+\frac{(-1)^{\vartheta}}{\vartheta!}\right). \tag{2}$$

The function $f(\vartheta)$ is derived from the sieve formula. To simplify calculation, $f(\vartheta)$ can be simplified to $f(\vartheta)=\{\vartheta!/e\}$. $\vartheta!/e$ denotes the integer closest to $\vartheta!/e$.

The cover cycle T is the time required for the adversary to sniff all available *rid*s.

We first calculate the mean mapping cycle time for a chosen *rid*. This is used to determine the time required to sniff all *rid*s. Using mathematical induction, we can calculate the probability of any *rid* sniffed for the first time, after the kth mapping. The probability is given in Equation 3.

$$P_k=\frac{A_{N_{RID}-1}^{\vartheta}}{A_{\vartheta}^{\vartheta}}\cdot\left(\frac{A_{N_{RID}-1}^{\vartheta}}{A_{N_{RID}-1}^{\vartheta}}\right)^{k-2}\cdot\frac{C_{N_{RID}-1}^{\vartheta-1}A_{\vartheta}^{\vartheta}}{A_{N_{RID}-1}^{\vartheta}}. \tag{3}$$

The mathematical expectation for P_k is Equation 4. The cover cycle T is Equation 5.

$$E=\sum_{k=1}^{\infty}kP_k=\frac{N_{RID}}{\vartheta}-\frac{(N_{RID}-\vartheta)(N_{RID}-\vartheta)!}{(N_{RID}!)\vartheta}. \tag{4}$$

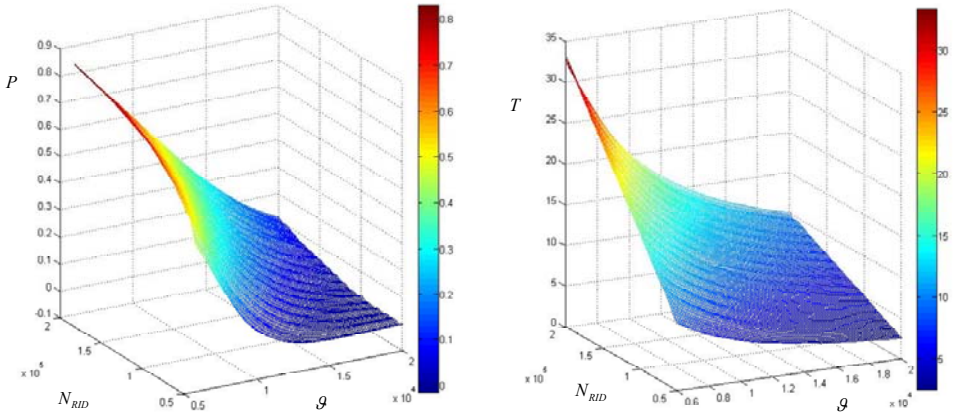
$$T=\tau\times\left(\frac{N_{RID}}{\vartheta}-\frac{(N_{RID}-\vartheta)(N_{RID}-\vartheta)!}{(N_{RID}!)\vartheta}\right). \tag{5}$$

The formulas above are for situations where the adversary can sniff only one port in the network. When the adversary is able to sniff more than one port, the number σ of *rid*s sniffed will be greater than the number ϑ of *rid*s in one sniffed port. Because the network topology and the *rid* segment pool are unchanged, we can use σ to calculate P and T rather than ϑ. Therefore, the formulas above are also useful for analyzing a more complex monitoring environment.

Numerical and Simulation Evaluation

Fig.2(a) shows the theoretical value of P with N_{RID} and ϑ. The value of N_{RID} is the total number of *rid*s available, while ϑ is the number of *rid*s sniffed in one cycle. Fig.2(b) shows the theoretical value of T with N_{RID} and ϑ. We set the time of mapping as the unit for T. In Fig.2(a), we observe that P decreases as ϑ increases when N_{RID} is constant, while P decreases as N_{RID} increases when ϑ is

constant. Note that T shows the same trend as P, which peaks when the difference between N_{RID} and ϑ is as large as possible.



(a) Probability distribution (b) Cycle

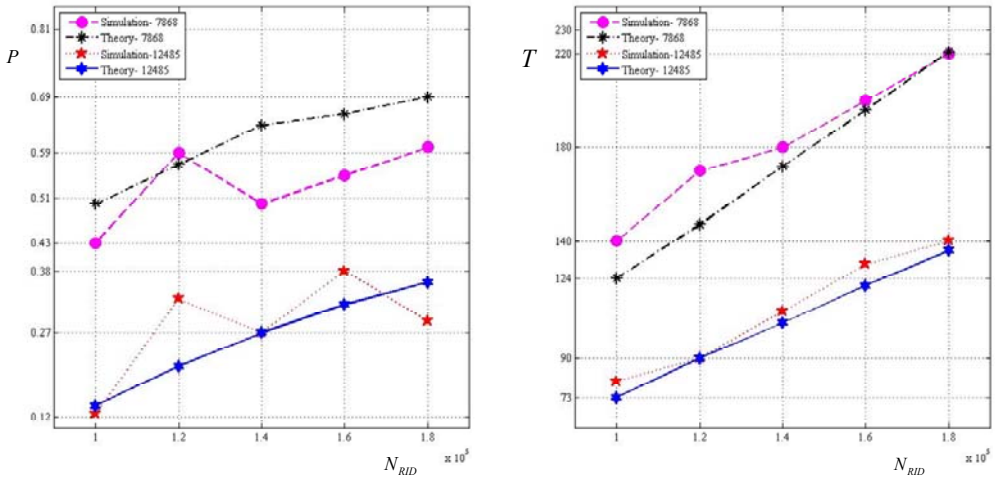
Fig.2 Theoretical values of P and T

From the theoretical results in Fig.2, we see that increasing N_{RID}will increase P and T as well as reducing the efficiency of collecting network information to improve the network security.

Table 2 Parameters for the Topologies

Parameter	Topology One	Topology Two
The number of *nid* being used, N_{NID}	61440	89112
The number of link in the core network, $\|E\|$	42	65
The number of the available *rid* , N_{RID}	No less than 62000	No less than 90000
The number of the sniffed *rid* in one cycle, ϑ	7868	12485

In our simulation, we build two topologies to observe and examine the changes of P and T. The parameters for the topologies are shown in Table 2. We again set D to be equal to N_{NID} and N to be infinite. Fig.3(a) and (b) show the simulation results for P and T respectively as N_{RID} ranges from 10^5 to $1.8*10^5$. From figures, we can see that the simulation results echo the theoretical analysis.



(a) Probability distribution (b) Cycle

Fig. 3 Comparison of simulated and theoretical data

5. Summary

In this paper, we presented *odim* to manage the identifiers of network nodes for defending against scanning and worm propagation in the *sinet*. To the best of our knowledge, *odim* is the first mechanism that manages *nid*s for the network nodes in the *sinet*. We established the selection and allocation constraints, and proposed the selection and allocation algorithms to solve the constraints. To evaluate the defense efficiency against scanning, we defined non-repetition probability and cover cycle. Simulation results and theoretical analysis demonstrated that *odim* can effectively reduce the detection probability of *rid*s.

Acknowledgements

This work was supported in part by the National High Technology Research and Development Program 863 (2015AA016103), in part by Chinese National Natural Science Foundation (61301081), in part by National 973 Program of China (2013CB329100).

References

[1] Yan Q, Yu F R, Gong Q, et al. Software-defined networking (SDN) and distributed denial of service (DDoS) attacks in cloud computing environments: A survey, some research issues, and challenges[J]. IEEE Communications Surveys & Tutorials, 2016, 18(1): 602-622.
[2] Jones R A, Horowitz B. A System-Aware Cyber Security architecture[J]. Systems Engineering, 2012, 15(2): 225-240.
[3] Antonatos S, Akritidis P, Markatos E P, et al. Defending against hitlist worms using network address space randomization[J]. Computer Networks, 2007, 51(12): 3471-3490.
[4] Francis P. Network Address Translation (NAT)[J]. ACM SIGCOMM Computer Communication Review, 2015, 45(2): 50-50.
[5] Jafarian J H H, Al-Shaer E, Duan Q. Spatio-temporal address mutation for proactive cyber agility against sophisticated attackers[C]. Proceedings of the First ACM Workshop on Moving Target Defense, Scottsdale, AZ, USA, 2014: 69-78.
[6] Jafarian J H, Al-Shaer E, Duan Q. An effective address mutation approach for disrupting reconnaissance attacks[J]. IEEE Transactions on Information Forensics and Security, 2015, 10(12): 2562-2577.
[7] Zhang H, Quan W, Chao H, et al. Smart identifier network: A collaborative architecture for the future internet[J]. IEEE Network, 2016, 30(3): 46-51.
[8] Zhang H, Su W, Quan W. Security Technologies of SINET[M]. Smart Collaborative Identifier Network. Springer Berlin Heidelberg, 2016: 223-249.