# Self-adaptive End-point Mutation Technique Based on Adversary Strategy Awareness

Yong Zhao<sup>†</sup>, Cheng Lei<sup>\*‡</sup>, Yingjie Yang<sup>\*‡</sup>, Hongqi Zhang<sup>\*‡</sup>,

Tong Yang<sup>§</sup>, Zongyi Zhao<sup>†</sup>, Xiaomei Sun<sup>†</sup>

\*China National Digital Switching System Engineering and Technological Research Center, Zhengzhou 450001, China

<sup>†</sup> Department of Computer Science and Technology, Tsinghua University, Beijing, China

<sup>‡</sup>Henan Key Laboratory of Information Security, Zhengzhou, China

<sup>§</sup>Department of Computer and Science, Peking University, Beijing, China

 $zhaoyong 11 @mails.tsinghua.edu.cn\ leicheng 12150 @gmail.com\ zhq 37922 @126.com$ 

yyje@126.com yang.tong@pku.edu.cn {sunxm15, zhaozy13}@mails.tsinghua.edu.cn

Abstract-Moving target defense is a revolutionary technology to change the pattern of attack and defense, and end-point information mutation is one of the hotspots belonging to this field. In order to counterpoise the defense benefit of end-point information mutation and service quality of network system, the self-adaptive end-point mutation technique based on adversary strategy awareness is proposed. Directed at the blindness problem of mutation mechanism in the course of defense, adversary strategy awareness based on Sibson entropy algorithm is proposed for guiding the choice of mutation mode by discriminating the scanning attack strategy. Aimed at the low availability problem caused by limited network resource and high mutation overhead, satisfiability modulo theories are used to formally describe the constraints of mutation. Finally, theoretical and experimental analysis shows the ability to resist scanning attack and mutation overhead.

Index Terms— End-point Information Mutation; Self-adaptive Adjustment Mechanism; Adversary Strategy Awareness; Software Defined Network; Satisfiability Modulo Theories

# I. INTRODUCTION

With the development of means of attack, new types of attacks, such as zero-day exploit attack and Advanced Persistent Threats, are on the rise. On the one hand, network scanning attack uses different scanning strategies in different kinds of network information system architectures. On the other hand, the static and certainty feature of existing network information system decreases the difficulty of scanning attack. Traditional defense methods cannot defense proactively without ascertaining all kinds of network attacks[1]. Therefore, the contrast between the targeted of network scanning attack strategy and the certainty of information system makes the traditional security methods not adequate to new threats.[2].

In order to improve the advancement and effectiveness of defensive mode, network based moving target defense[3] comes into being. It tricks, evades and prevents scanning attack by changing network configuration and status dynamically. End-point mutation technique[4] is one of the key techniques. NASR[5] prevents connection requests not within service period by using address transition of packet header and update of net-flow table based on DHCP update. SDNA[6] confuses scanning attack by virtual hopping, which deploys a hypervisor node in each subnet to ensure mutation consistency. OF-RHM[7] proposed virtual end-point mapping mechanism based on Openflow<sup>1</sup>. It converts real IP addresses to virtual IP addresses so as to implementing end-point hopping. Jafarian *et. al*[8] proposed ST-RHM mutation mechanism. It can resist cooperative scanning attack effectively by using temporal-spatial mixed mutation method based on SDN.

The key contributions of self-adaptive End-point Mutation Technique based on Adversary Strategy Awareness (SEMT) is shown as follows:

1) Adversary strategy awareness based on Sibson entropy is proposed. It guides the choice of mutation mode by discriminating the scanning attack strategy, which enhances the defense capability.

2) End-point mutation based on satisfiability modulo theories is proposed. It uses satisfiability modulo theories[10] to formally describe the constraints of mutation in order to ensure the low-overhead of mutation implementation, which in turn increases the availability of mutation mechanism.

# II. ARCHITECTURE OF SEMT

The overall architecture of SEMT is shown in Fig. 1. It implements scanning attack strategy identification and endpoint mutation by using hopping switch (HS) and hopping controller (HC) collaboratively. In HC, the function of network scanning attack strategy analysis engine is to analyze scanning attack strategy. SMT calculator is to obtain the required end-point information set. Hopping mapping engine is used to convert Endpoint Information (EI) to hopping Endpoint Information (hEI). The function of HS is to update information of hopping pattern and net-flow tables. Besides, HS is used for detecting, filtering and collecting network topology changes and illegal connection requests.

# III. SELF-ADAPTIVE END-POINT MUTATION ALGORITHM

# A. Adversary Strategy Awareness Based on Sibson Entropy

Suppose the hEI space is divided into  $m_B m_L$  blocks, the total number of failed requests in the  $t^{th}$  hopping period is

<sup>1</sup>OpenFlow is the first standard communications interface defined between the control and forwarding layers of an SDN architecture

© 2016, Yong Zhao. Under license to IEEE. DOI 10.1109/LCN.2016.99





Fig. 1. Architecture of SEMT

 $N_{fail}$ . The number of failed request packets in the  $i^{th}$  divided hEI space is denoted as  $N_{fail}^i$ . Eq. 1 is used to calculate the probability distribution of the source and destination address of failed requests in one hopping period denoted. Eq. 2 indicates that the Sibson entropy of the source address probability distribution of the failed request packets in the two consecutive low-frequency hopping period. By comparing with the set threshold, we can determine whether the attacker uses follow-up scanning strategy.

$$P_i^j(\pi) = \pi_k \cdot (\sum_{k=1}^{N_{fail}} \pi_k)^{(-1)}$$
(1)

$$D_{S}(P_{t-1}^{Src}(\pi), P_{t}^{Src}(\pi)) = \frac{1}{2} \{ D_{i}[P_{t-1}^{Src}(\pi), \overline{P^{Src}}] + D_{i}[P_{t}^{Src}(\pi), \overline{P^{Src}}] \}$$
(2)

If the attacker uses the blind scanning strategy, the average number of scanned times of every end-point information is  $N_{fail}/m_Bm_L$  in the ideal condition. However, because the attacker cannot always complete a random scan of the whole network address space within one hopping period, the Sibson entropy directly calculated based on the distribution of failed request packets of destination address and  $N_{fail}/m_Bm_L$  in one hopping period will be larger. Therefore, we use Chauvenet criterion shown as Eq. 3 to eliminate the abnormal HTHR. On the basis of the Eq. 4, we can calculate the destination address probability distribution of the failed request packets. By comparing with the setting threshold, we can determine whether the attacker uses blind scanning strategy.

$$\frac{N_{fail}^{i} - N_{fail}/m_{B}m_{L}}{(m_{B}m_{L})^{2}/12} < \zeta$$
(3)

$$D_{S}(P_{t}^{Dst}(\pi), \frac{N_{fail}}{m'_{B}m'_{L}}) = \frac{1}{2} \{ D[P_{t}^{Dst}(\pi), \overline{P_{t}^{Dst}}] + D[\frac{N_{fail}}{m'_{B}m'_{L}}, \overline{P_{t}^{Dst}}] \}$$
(4)

Furthermore, if attackers use mixed scanning strategies, SEMT implements corresponding mutation strategy according to the discrimination of scanning attack strategy so as to proactive defense actively.

# B. End-point Mutation based on SMT

Firstly, SMT solver is used to obtain the required hEI set, which is to meet the constraints. Although SMT solving is still N-P problem, the existing SMT solver, such as Z3[10] can reach millions of orders of magnitude, thus can be used effectively to obtain required hEI set.

If the network nodes needed protected are represented of by  $\{h^1, h^2, ..., h^l\}$ , which are distributed in subnets  $\{S^1, S^2, ..., S^k\}, k \leq l$ . The availability hEI set is represented as  $hEI \land \neg (hEI_1 \lor ... \lor hEI_k \lor EI_1 \lor ... \lor EI_l)$ . The available hEI set in base hopping period (TBHR) is divided into  $m_B$  number of Base Hopping Range (BHR) according to the number of subnet and its scale. Then, each BHR is divided into  $m_L$  number of Low-frequency Temporal Hopping Range (LTHR) in each low-frequency mutation period ( $T_{LTHR}$ ) according to the number of nodes in subnet and its resource value. It has  $T_{BHR} = c \cdot T_{LTHR}, c \in Z^+$ . What's more, each LTHR is divided into  $m_H$  number of High-frequency Temporal Hopping Range (HTHR), which contains  $n_{HTHR}$ hEIs.

Since end-point information mutation implementation needs HS and HC in collaboration, the mutation constraints can be divided into routing constraints and end-point information constraints. The details are as follows:

1) Capacity constraint: It refers to the maximum net-flow volume of each hopping routes can carry. The details are shown in equation 5 and 6.

$$\forall hR_j, C(hR_j) - B_f - \sum b_{ij} \cdot B_d \ge B_{th}, \text{ so } b_i^k = 0$$
 (5)

$$\sum_{k} \sum_{j_1} \sum_{j_1 \neq j_2} D^k_{j_1, j_2} \ge \Phi$$
 (6)

2) Quality of service constraint: It means that the maximum length of routing path cannot exceed the threshold  $L_{max}$ , which is shown in Eq. 7.

$$\sum b_i^k \le L_{max}, i \in \{Src, hR_1, ..., Dst\}$$
(7)

3) Mutation rate constraint: This constraint sets  $T_{EHP}$  for every nodes according to its resource value firstly. Based on that, it assigns mutation space according to mutation rate and repetition rate, which is specified in Eq. 8 and 9.

$$T_{EHP}^{i} = \frac{T_{LTHR}}{V^{i}} \tag{8}$$

$$N_{LTHR}^{i} = \left\lceil \frac{T_{LTHR}}{T_{EHP}^{i}} \right\rceil n_{HTHR}, N_{BHR} = \sum_{i \in S_x} N_{LTHR}^{i} \qquad (9)$$

4) Mutation space selection constraint: This constraint ensures every node can be assigned required hEI, and improves the unpredictability of mutation, which is shown in Eq. 10.

$$\sum_{i < j \le M} b_i^j \ge 1 \quad \sum b_i^j \ge \frac{N_{LTHR}^i - 1}{2\delta_3 n_{HTHR}} \quad \forall h EI \in F, b_i^j = 0 \quad (10)$$

In order to improve the defense benefit of end-point mutation, SEMT selects different mutation strategy on the basis of scanning attack strategy analysis after meeting the SMT constraints.

1) If it is blind scanning attack, weighted random hopping strategy is used. Eq. 11 is used to calculate the weighted value  $w_i^{EI}$  of hEI in each  $T_{EHP}$ . The higher the  $w_i^{EI}$  is, the more possibility of hEI might be selected in the following hopping period.

$$\begin{cases} 0, & \text{if } l \leq \frac{l}{M} - \frac{M^2}{6} \\ 1 - \frac{1}{\delta^2} \min\left(\sqrt{D_S}, \delta_2\right), & \text{else} \end{cases}$$
(11)

$$hEI_{\alpha} \in hEI_{n_{HTHR}},$$
  
$$\alpha = H_f(SrcIP, SrvID, K_S) \mod n_{HTHR} + 1 \qquad (12)$$

2) If it is follow-up scanning attack, SEMT introduces hopping period stretching based on weighted random mutation. The so-called hopping period stretching means that the hopping period of hEI is stretched according to both network environment and scanning attack frequency change.

The magnitude of  $T_{EHP}$  to decrease is shown in Eq. 13. It means the  $t + 1^{th}$  hopping period  $T_{EHP}^{t+1}$  is determined by the  $t^{th}$  hopping period and scanning attack frequency, which must be larger than TTL.

$$T_{EHP}^{t+1} = \max[\alpha T_{EHP}^t d/n_{fail}^t + (1-\alpha) T_{EHP}^t, T_{EHP}^{lb}]$$
(13)

The magnitude of  $T_{EHP}$  to increase is shown in Eq. 14. It means the  $t + 1^{th}$  hopping period  $T_{EHP}^{t+1}$  is determined by network delay, which must be smaller than  $T_{HP}^{ub}$ .

$$\begin{cases} T_{EHP}^{t} + t'_{d} \text{ if } T_{EHP}^{t} + t'_{d} \leq T_{EHP}^{ub} \\ T_{EHP}^{ub}, \text{ else} \end{cases}$$
(14)

3) If it is half-blind scanning attack, SEMT uses reverse mutation strategy to defend. It constructs directed graph DG(V, E), network nodes are divided into different groups according to Weakly Connected Component (WCC). The reversed mutation strategy assigns hEIs without weakly connected relationship to end-point in the following hopping period, which can reduce the success rate of scanning attack. Transition parameter  $ts_t^i$ is used to obtain the possibility of  $h^j$  to be scanned in time, which is shown in Eq. 15.

$$ts_{t+1}^{i} = \alpha ts_{t}^{i} + (1 - \alpha)ts_{0}^{i} \tag{15}$$

## IV. THEORETICAL ANALYSIS

# A. Security Analysis

Suppose there are  $n_1$  targets node in the network, the endpoint information space is m, scanning width is w, and the scanning frequency is  $1/T_{SCN}$ , the number of the end-point information scanned by the attack is  $1/T_{SCN}$ ; the ratio of scanning attack scanning frequency and mutation frequency is  $r = T_{EHP}/T_{SCN}$ .

1) The capability of resisting blind scanning attack

Since the blind scanning attack are even scanning without repetition, in static network  $T_{EHP} = \infty$ , the probability that

an attacker has successfully scanned x end-point information obeying the hyper-geometric distribution can be expressed as  $P_b(x) = (C_{n_l}^x \cdot C_{m-n_l}^{n_s-x})/C_m^{n_s}$ . Therefore, in static network when an attacker performs successful blind scanning, the success rate is  $P_b^{static}(x > 0) = 1 - C_{m-n_l}^{n_s}/C_m^{n_s}$ . In ST-RHM[9] and SEMT network, the success rate that the attacker successfully scanned x end-point information within one  $T_{EHP}$  obeys Bernoulli distribution, and it can be expressed as  $P_b(x) = C_{n_l}^x [n_l/(n_l+m)]^x [1-n_l/(n_l+m)]^{n_s-x}$ . Therefore, the probability that the attacker successfully launches blind scanning is  $P_b(x > 0) = 1 - [1 - rwn_l/(mn_l + mrw)]^{n_s}$ . Particularly when r = 1, the scanning attack frequency is the same as the mutation frequency, the probability that an attacker successfully launching blind scanning is  $P_b(x > 0) = 1 - [1 - rwn_l/(mn_l + mrw)]^{n_s}$ .

2) The capability of resisting half-blind scanning attack

When the attacker launch follow-up scanning attack, there is  $r \geq 1$ . Suppose attackers can repeat scanning b times in one  $T_{EHP}$ . The success rate of attackers in ST-RHM is  $P_{fu}(x > 0) = 1 - [1 - (bn'_l - n_r)/(n'_l + \phi mb)]^{n_s}$ . Since SEMT use "fast decreasing, slow increasing" policy to increase mutation frequency, it makes  $r \leq 1$  eventually. Therefore, the attacker's success rate is  $P_{fu}(x > 0) = 1 - [1 - (rn'_l - n_r)/(n'_l + \phi m)]^{n_s}$ 

3) The capability of resisting follow-up scanning attack

Since the half-blind scanning attack will repeat scanning physically adjacent end-point information space, it may be assumed that an attacker can repeat scanning *a* times, and the end-point information space can be scanned is  $\phi m$ ,  $\phi \in (0, 1)$ . In static network, the success rate is  $P_{hb}^{static}(x > 0) = 1 - aC_{\phi m-n_l}^{n_s/a}$ . For ST-RHM and SEMT, due to the implementation of cheat mutation and reverse mutation, there are  $n_r = r|WCC(h^i)|$  invalid hEIs in the next hopping period. Therefore, the success rate to scan *x* hEI is  $P_{hb}(x) = C_{n_l}^{r_l}[(n_l' - n_{\gamma})/(n_l' + \phi m)]^x[1 - (n_l' - n_{\gamma})/(n_l' + \phi m)]^{n_s - x}$ . Therefore, after *a* times half-blind scanning, the success rate is  $P_{hb}(x > 0) = 1 - a[1 - (wrn_l' - \phi mn_{\gamma} + \phi mwr)]^{n_s}$ .

# B. Mutation Overhead

The overhead of static networks, ST-RHM and SEMT mutation is shown in Table I. Assuming the number of host nodes in a subnet is  $n_t$ , hEI space is  $n_m$ , and EI can be aggregated is  $n_a$ .

TABLE I MUTATION COST

Mutation mechanism	Computational complex	Average transmission delay	Net-flow table size
Static network	O(1)	$t \times L_S$	$n_l$
ST-RHM	$O((\gamma n_h)^2)$	$t \times L_S$	$1 + n_m m_H$
SEMT	$O((\gamma n_h)^2)$	$t \times L_S$	$1 + m_H n_m / n_a$

# V. EXPERIMENTS AND ANALYSIS

In order to verify the feasibility and effectiveness of SEMT, we use Mininet[11] to build simulation network topology and adopt Erdos-Renyi model for random network topology generation. We choose Openflow1.3 as HS and POX as HC. The configuration of source and destination nodes is shown in Table II. Besides, hEI is composed of Class B IP address pool and  $2^{16}$  size port pool,  $\sigma = 5$ ,  $\delta_1 = 0.05$ ,  $\delta_2 = 0.075$ ,  $\delta_3 = 0.05$ ,  $\gamma = 0.4$ ,  $\lambda = 0.02$ ,  $L_{max} = 32$ ,  $T_{LTHR} = 50s$ ,  $\xi = 2.0$ .

 TABLE II

 CONFIGURATION OF EXPERIMENTAL NETWORK

Nodes	OS	corresponding HS	V
Web Server	Linux	16	5
FTP Server	Linux	16	3
Clients	Windows XP	1	2

## A. Resisting of Scanning Attack

In practical environments[2][4], the attacker often filtered EI through blind scanning. The success rate of mixed scanning attack is shown in Fig. 2. SEMT can effectively reduce about 29% scanning attack compared with ST-RHM.



Fig. 2. Success rate of mixed scanning attack

# B. Overhead

Since the routing forwarding complexity in SEMT mutation is proportional to the size of net-flow table to be updated, the size of net-flow table to be updated is tested in our experiment to analyze routing load caused by SEMT mutation. Fig. 3 shows that SEMT routing capacity constraints can effectively reduce the size of net-flow table.



Fig. 3. Experiments of flow table size

## VI. CONCLUSION

End-point information mutation is a kind of proactive network defense technique. Aimed at the blindness problem and low availability of mutation mechanism in the course of defense, self-adaptive end-point mutation technique based on adversary strategy awareness is proposed. Adversary strategy awareness based on Sibson entropy is designed in order to guide the choice of mutation mode by discriminating the scanning attack strategy, which enhances the defense benefit. Based on the discrimination of scanning attack strategy, end-point mutation based on satisfiability modulo theories is proposed. We use satisfiability modulo theories to formally describe and solve the constraints of mutation in order to ensure the low-overhead of mutation, which decreases the defense cost. Theoretical analysis and simulation experiments show SEMT can disrupt almost 90% scanning attack even in mixed scanning strategy. Consequently, SEMT has good defensive performance, and ensures the low overhead of mutation at the same time.

### VII. ACKNOWLEDGEMENT

This work is supported by Foundation Items: The National High Technology Research and Development Program of China (863 Program)(2012AA012704, 2015AA016106) and Zhengzhou Science and Technology Talents (131PLKR-C644) and Primary research and development plant of China (2016YFB1000300).

#### REFERENCES

- Quigley K, Burns C, Stallard K. Cyber Gurus: A rhetorical analysis of the language of cybersecurity specialists and the implications for security policy and critical infrastructure protection[J]. Government Information Quarterly, 2015, 32(2): 108-117.
- [2] Bou-Harb E, Debbabi M, Assi C. Cyber scanning: a comprehensive survey[J]. Communications Surveys & Tutorials, IEEE, 2013, 16(3): 1496-1519.
- [3] Sun K, Jajodia S. Protecting enterprise networks through attack surface expansion[C]//Proceedings of the 2014 Workshop on Cyber Security Analytics, Intelligence and Automation. ACM, 2014: 29-32.
- [4] Jun Xu, Pinyao Guo, Mingyi Zhao, et al.. Comparing different moving target defense techniques[C]. Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, Scottsdale, Arizona, 2014: 97-107.
- [5] 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.
- [6] Yackoski J, Xie P, Bullen H, et al. A self-shielding dynamic network architecture[C]//MILITARY COMMUNICATIONS CONFERENCE, 2011-MILCOM 2011. IEEE, 2011: 1381-1386.
- [7] Jafarian J H, Al-Shaer E, Duan Q. Openflow random host mutation: transparent moving target defense using software defined networking[C]//Proceedings of the first workshop on Hot topics in software defined networks. ACM, 2012: 127-132.
- [8] 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. ACM, 2014: 69-78.
- [9] Sibson R. Information radius[J]. Zeitschrift fr Wahrscheinlichkeitstheorie und verwandte Gebiete, 1969, 14(2): 149-160.
- [10] Bjørner N, De Moura L. Z310: Applications, enablers, challenges and directions[C]//Sixth International Workshop on Constraints in Formal Verification. 2009.
- [11] Lantz, B., B. Heller, and N. McKeown. A network in a laptop: rapid prototyping for software-defined networks. in Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks. 2010