International Journal of Computational Intelligence Systems, Vol. 5, No. 5 (September, 2012), 868-877

Locator/Identifier Separation: Comparison and Analysis on the Mitigation of Worm Propagation

Ming Wan[*], Ying Liu, Jian-qiang Tang, Hong-ke Zhang
National Engineering Laboratory for Next Generation Internet Interconnection Devices, Beijing Jiaotong University
Beijing, 100044, China

Si-yu Lin
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University
Beijing, 100044, China
Department of Electrical and Computer Engineering, University of Victoria
Victoria, V8W 3P6, Canada

Received 30 November 2011

Accepted 15 June 2012

Abstract

As a basic prerequisite for worm detection based on computational intelligence in networks with locator/identifier separation, it is well worth considering the influence on worm propagation due to the incoming locator/identifier separation. In this paper, according to the characteristics of locator/identifier separation, we systematically analyze the mitigation of worm propagation in three aspects: address semantics, address space and mapping delay. By applying the classical AAWP and SIR worm propagation models, we give a quantitative comparison between today's Internet and networks with locator/identifier separation. In particular, our research results show that, the characteristics of locator/identifier separation can help to markedly mitigate worm propagation, and networks with locator/identifier separation are more resistant to worm propagation than today's Internet.

Keywords: locator/identifier separation; worm propagation; AAWP; SIR; computational intelligence

1. Introduction

In today's Internet, computational intelligence seems promising to detect a wide variety of threats and attacks, and has been increasingly applied in the area of network and information security, such as Bot-nets detection using network forensics and artificial intelligence techniques[1] and multi-agent intrusion detection system[2]. However, in networks with locator/identifier separation, since locator/identifier separation has changed the basic network architecture and communication mechanism, it will definitely exert some influences on network security, especially worm propagation. From this point, to develop worm detection based on computational intelligence more efficiently in networks with locator/identifier separation, the study on the mitigation of worm propagation due to the incoming architectural change is absolutely necessary and beneficial.

It has been pointed out in many recent researches that, the current Internet routing architecture is facing serious scalability issues[3,4], which are basically caused by the overloading of IP address semantics. That is, an IP address represents not only the identity but also the location of an end host. Therefore, a promising solution,

[*] Corresponding author: ming305.bjtu@gmail.com

Published by Atlantis Press

locator/identifier separation[5-8], has been proposed by many research communities to resolve the dual semantics of IP address. That is to say, the existing IP addresses are separated into two independent namespaces: a locator namespace and an identifier namespace. Literally speaking, the locator namespace, used in the network layer to forward packets, represents the location of an end hosts. On the other hand, the identifier namespace, used in the transport and upper layers to identify nodes, represents the identity of an end host. However, in networks with locator/identifier separation, a critical challenge is how to decouple locators from identifiers. Based on dissimilar hypotheses, a great many identifier-to-locator mapping services[9-14] have been put forward by both academia and industry to map locators onto identifiers. Also, all these mapping services have one common basic characteristic: the identifier-to-locator mappings of all end hosts are stored and maintained in one or some organized nodes which are referred to as the resolvers in the rest of this paper for ease of expression. When a tunnel router (TR) tries to resolve an identifier-to-locator mapping for some identifier, it first needs to send a mapping request to the corresponding resolver. Once the TR receives this identifier-to-locator mapping from the resolver, it can encapsulate and forward the packet by using the received locator. In this paper, we focus on worm spread in networks with locator/identifier separation, and systematically analyze the mitigation of worm propagation according to the ordinary characteristics of locator/identifier separation. In particular, we just utilize the commonness of networks with locator/identifier separation, and how to accomplish the separation or how to design a scalable mapping service is outside the scope of this paper.

When the paradigm of locator/identifier separation comes to being, it is worth arguing whether this paradigm can provide better security capability for the Internet. From this point, we pay more attention to the damage of worm spread, which cannot be evaded in either today’s Internet or networks with locator/identifier separation. Our main purpose is to demonstrate obvious mitigation of worm propagation by virtue of the incoming locator/identifier separation, and to provide a basic prerequisite for worm detection based on computational intelligence in networks with locator/identifier separation. In today’s Internet, worms exist as a great threat to its dependability due to their ability to infect millions of end hosts in a very short period of time[15], and the infected end hosts can be recruited as the bots or zombies who can be manipulated to cause unprecedented damage, such as DDoS attacks[16]. At the same time, the easy access and widespread usage of the Internet make it a powerful means for propagating the malicious worms. In general, worms are self-propagated programs throughout the Internet by exploiting the vulnerabilities or policy flaws of computers. When a worm aims at infecting other end hosts, it can find the vulnerabilities of these end hosts by mean of scanning self-generated IP address pools, and then directly comprises these end hosts by implanting the copies of worm into the vulnerable target computers. In addition, the newly infected end hosts will automatically and continuously attempt to infect other end hosts in the same way until the worm spreads to the entire IP address space. Given the fast spread and the substantial damage caused by worms, it is significant to develop monitoring and detection mechanisms against worms. In today’s Internet, computational intelligence has figured prominently in many solutions to the worm detection problem, and this prominence and popularity will continue to shine in networks with locator/identifier separation. Therefore, the study on the mitigation of worm propagation can help to develop worm detection based on computational intelligence.

In this paper，we concentrate our study on the scan-based worm propagation in today’s Internet and networks with locator/identifier separation. Indeed, the pronounced change of worm propagation caused by locator/identifier separation should deserve wide attention and discussion. In order to address this issue, this paper summarizes the characteristics of locator/identifier separation, and appropriately selects the classical Analytical Active Worm Propagation (AAWP) and Susceptible-Infected-Removed (SIR) models to illustrate the mitigation of worm propagation. By means of the quantitative comparison between today’s Internet and networks with locator/identifier separation, we can draw the conclusion: locator/identifier separation can be markedly conducive to alleviating worm propagation. To be precise, we systematically analyze the mitigation of worm propagation in the following three aspects:

Firstly, the change of address semantics can impact on worm propagation. In today's Internet, some types of worms can purposefully spread according to BGP routing information or Class A address space. However, locator/identifier separation resolves the dual semantics of IP address, and conceals the location information of end hosts. Besides, the feasible flat identifiers also make worms more difficult to spread.

Secondly, since IP address space has been divided into the identifier space and the locator space, worms are quite possible to scan the locator space which has no significance for worm propagation.

Thirdly, in networks with locator/identifier separation, when worms attempt to scan some identifier, the TR must send a mapping request to resolve the corresponding locator for the identifier. Therefore, the mapping delay may play an important role to mitigate worm propagation.

Not surprisingly, we also give some discussions on worm detection based on computational intelligence in networks with locator/identifier separation, and analyze some benefits provided by locator/identifier separation.

2. Related Work

In this section, we present some related work from both academia and industry. First of all, we briefly introduce various locator/identifier separation approaches and some corresponding mapping services. And then, we sum up the related researches on worm propagation, including the existing worms and worm propagation models.

2.1. *Locator/identifier separation*

Generally speaking, the existing locator/identifier separation approaches can be divided into two classes. One is host-based solution, for example I3[5] and HIP[6]. In order to accomplish locator/identifier separation, this solution must upgrade or modify the host's protocol stacks, and add an additional shim layer to identify end hosts. Although this class of approach is beneficial to host mobility and multi-homing, it is very difficult to deploy due to the requirement of host changes. The other is network-based solution which is also named as core-edge separation, such as LISP[7] and Ivip[8]. In this solution, IP addresses are divided into two parts: identifiers and locators. To be more exact, routing objects in edge networks are identifiers, but those in transit core are locators. Compared with host-based solution, this class of approach requires no change to the host and is quite convenient to deploy and maintain. However, network-based solution may need some other schemes to deal with host mobility or multi-homing. In a word, both these classes attempt to resolve the dual semantics of IP address, and they have their own strength and weakness.

A critical challenge for networks with locator/identifier separation is how to design an excellent mapping service. At present, a great many mapping services have been proposed by the sagacious pursuits. In order to achieve very low mapping delay, some approaches[8,9] advise to store and update the complete identifier-to-locator mappings for all end hosts in some routers or nodes. However, they are not scalable due to huge storage and maintenance effort. In addition, various overlay topologies have been applied to the design of the mapping service, for instance, LISP-DHT[10] and LISP+ALT[11]. Although these approaches can distribute the identifier-to-locator mapping information into different nodes, the side effect of it is that they need longer mapping delay. Given that locator/identifier separation can raise the opportunity for the flat identifiers, Ref. 12 accomplishes the mapping service by utilizing the domain name system (DNS), but it may increase the overhead on the DNS. Also, Ref. 13 uses a content-addressable network (CAN) approach to map the flat identifiers onto locators. In addition, Ref. 14 presents a new idea which allows the users of identifiers to choose their preferred/trusted mapping service providers.

2.2. *Worm propagation*

With complex Internet applications on the rise, worms have become one of the major threats to the Internet security. In 2001, by exploiting the buffer-overflow vulnerability in Microsoft's IIS web server, Code-Red worm[17] infected 359,000 end hosts in less than 14 hours, and the cost is estimated to be in excess of $2.6 billion. On January 25, 2003, Slammer[18] quickly spread throughout the Internet due to its superfast scan rate: 90% of vulnerable end hosts were infected within just 10 minute. At the same time, the enormous scan packets of Slammer caused a global-scale DoS attack to the

Internet. After only six months, Blaster[19] appeared and spread among more than 188,000 end hosts within two hours. In addition, Witty[20], detected in 2004, infected 100 and 160 end hosts in the first 10 seconds and the last 30 seconds, respectively. Ref. 21 studies worm spread over the Peer-to-Peer (P2P) networks, which make worm propagation more effective. In order to propagate faster, routing worm[22] can use BGP routing information to decrease the scanning space without ignoring any potential end hosts. In this way, the spread rate of routing worm increased by two or three times. In recent years, many other worms such as Conficker[23] and C-Worm[24] are more intelligent and have caused more substantial damage on the Internet.

In order to defend against worm propagation, worm propagation models which aim at characterizing and simulating the spread of worms have become an active research area. Currently, there are two main aspects in the study of basic worm propagation models. One is mainly based on the epidemiology model, including Susceptible-Infectious-Susceptible (SIS) model[25], Kermack-Mckendrick (KM) model[26], Two-Factor model[27], et al. This model provides a qualitative understanding of worm spread by using nonlinear different equations. In particular, KM model is also named Susceptible-Infectious-Removed (SIR) model, and is widely used as background research of other worm propagation models, SIRS model[28] for example. The other is mainly founded on the discrete time model, and a typical instance is the Analytical Active Worm Propagation (AAWP) model[29]. Properly speaking, this model can take advantage of deterministic approximation to describe the spread of active worms, and explain why virtually most worms will be slow in global prevalence to some extent.

3. Network Model

In this paper, we do not discuss how to accomplish locator/identifier separation or design a scalable mapping service in detail. By contrast, we only utilize the commonness of networks with locator/identifier separation, and give a general network model which is derived from the network-based solution. In particular, our study on the mitigation of worm propagation is based on this general network model. For ease of presentation, we use our network model and networks with locator/identifier separation interchangeably in the rest of this paper.

As illustrated in Fig. 1, our network model is founded on the core-edge separation[7,8]. That is to say, edge network which is composed of a great many customer networks (CNs) uses endpoint identifiers (EIDs) to forward packets, and core network which is composed of many provider networks (PNs) uses globally routing locators (RLOCs) to route packets. In addition, each tunnel router (TR) located at edge of either core network or edge network systematically administrates a portion of endpoint identifiers. Notice that identifiers in our network model may be either hierarchically structured[11] or flat[12,13], and we assume all resolvers belonging to all provider networks form an integrated mapping system.

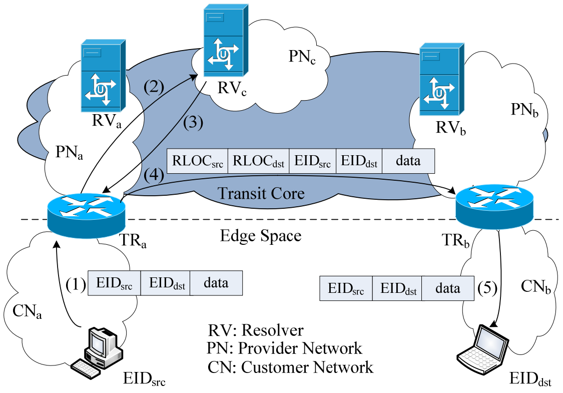

Fig. 1. Illustration for network model with locator/identifier separation.

As described by (1)~(5) in Fig. 1, the simple communication process is as follow: assuming that the end host with EID$_{src}$ in CN$_a$ wants to open a connection to the end host with EID$_{dst}$, the source host sends its first packet to its TR$_a$, using EID$_{src}$ and EID$_{dst}$ as the source and destination address of the packet respectively. When TR$_a$ with RLOC$_a$ receives this packet, if the mapping information for EID$_{dst}$ does not exist in its mapping cache, TR$_a$ sends a mapping request to the corresponding resolver RV$_c$ for the locator of EID$_{dst}$. After the mapping lookup of RV$_c$, TRa receives the locator RLOC$_b$ of EID$_{dst}$ from RV$_c$, and stores this mapping information into its mapping cache. Then, TR$_a$ encapsulates the received packet with a new routing header whose source address and destination address is RLOC$_a$ and RLOC$_b$ respectively. After that, TR$_a$ sends the newly-encapsulated packet into the transit core, and the packet will be routed and forwarded to TR_b. When TR_b receives the encapsulated packet, it simply strips the outer header and forwards the original packet to its destination host with EID_{dst}.

4. Modeling Mitigation of Worm Propagation

In this section, we mathematically model the mitigation of worm propagation between today's Internet and networks with locator/identifier separation. As stated previously, our study is mainly embodied with three aspects: address semantics, address space and mapping delay. We systematically analyze the mitigation of worm propagation by using AAWP and SIR models. Notice that, we use identifier and EID interchangeably in the rest of this paper, and the same holds for locator and RLOC. Besides, we assume the address space size of networks with locator/identifier separation is also 2^{32}, like that of today's Internet.

4.1. *Address semantics*

The dual semantics of IP address make worms faster to spread, because some worms can utilize the location information to reduce the scanning space, such as routing worm[22]. In this subsection, with the aid of the AAWP model, we illustrate the influence of address semantics on worm propagation by calculating the number of infected end hosts between today's Internet and networks with locator/identifier separation.

In today's Internet, according to the location information of IP address, we assume worms can define the size of scanning space as Ω. Therefore, the probability that any end host is hit by one scan is $1/\Omega$. Let m_i and n_i denote the total number of vulnerable end hosts (including the infected ones) and the number of infected end hosts at time tick i ($i \geq 0$) respectively. From Ref. 29, the number of the newly infected end hosts is $(m_i - n_i)[1-(1-1/\Omega)^{sn_i}]$, where s is the scanning rate. In addition, we assume d and p denote the death rate and the patching rate, respectively. Therefore, there will be $(d+p)n_i$ infected end hosts which will change to either vulnerable end hosts without being infected or invulnerable end hosts on the next time tick, and the total number of vulnerable end hosts (including the infected ones) will be reduced to $(1-p)m_i$. Therefore, on the next time tick the total number of infected end hosts will be:

$$n_{i+1} = (1-d-p)n_i + [(1-p)^i N - n_i][1-(1-1/\Omega)^{sn_i}]. \quad (1)$$

where $i \geq 0$, $n_0 = h$ represents the initial number of infected end hosts before worms spread, and $m_0 = N$ represents the number of vulnerable end hosts.

In networks with locator/identifier separation, since the dual semantics of IP address have been resolved by the paradigm of locator/identifier separation, the location information of end hosts can be sedulously concealed. In addition, the feasible flat identifiers[12,13] also make worms more difficult to surmise the identifier space, because they cannot find the corresponding addressing law. Although the change of address space described in the next subsection may decrease the size of identifier space, we also suppose the size of identifier space is 2^{32} in order to achieve the consistency. Therefore, on the next time tick the number of total infected end hosts will change to:

$$n'_{i+1} = (1-d-p)n'_i + [(1-p)^i N - n'_i][1-(1-1/2^{32})^{sn'_i}]. \quad (2)$$

where n'_i is the total number of infected end hosts at time tick i ($i \geq 0$) in networks with locator/identifier separation, and the assumption of other parameters is the same with that in the today's Internet.

4.2. *Address space*

Since IP address space has been divided into the identifier space and the locator space, worms are quite possible to scan the locator space which has no significance for worm propagation. In this subsection, we discuss the mitigation of worm propagation due to the change of address space. In particular, we assume worms do not know the location information of IP address, and they only use the random scanning to infect end hosts.

In today's Internet, in order to spread effectively, worms need to scan the entire IPv4 space. Therefore, Ω in Eq. (1) will be 2^{32}, and on the next time tick the total number of infected end hosts will be:

$$n_{i+1} = (1-d-p)n_i + [(1-p)^i N - n_i][1-(1-1/2^{32})^{sn_i}]. \quad (3)$$

where the assumption of the parameters is the same with that in Eq. (1).

In networks with locator/identifier separation, since a part of address space is used as the locators which do not represent end hosts, it is meaningless for worms to scan this address space. Therefore, for random scanning, the probability that any address is hit by one scan is $1/2^{32}$. Let q denote the identifier probability that the scanned address is an identifier. As a result, on the next time tick the total number of infected end hosts will be:

$$n'_{i+1} = (1-d-p)n'_i + [(1-p)^i N - n'_i][1-(1-q/2^{32})^{sn'_i}]. \tag{4}$$

where the assumption of the parameters is the same with that in Eq. (2). In reality, q can be computed by

$$q = \frac{\text{size of identifier space}}{\text{size of identifier space + size of locator space}}.$$

4.3. *Mapping delay*

In networks with locator/identifier separation, when a worm tries to scan some identifier, the TR must first achieve the mapping information for this identifier from the corresponding resolver. Therefore, the mapping delay may mitigate worm propagation. However, AAWP which is a discrete time model cannot felicitously embody the influence on worm propagation. In this subsection, we appropriately select the classical SIR model, which assumes that during the worm spread some infectious end hosts can either recover or die by patching or closing, and these end hosts are immune to the worm forever. Thus each end host stays in one of three states at any time: susceptible, infectious, or removed. The SIR model can be defined as

$$\begin{cases} \dfrac{dJ(t)}{dt} = \beta(\tau)J(t)[N - J(t)] \\ \dfrac{dI(t)}{dt} = \beta(\tau)S(t)I(t) - \gamma I(t) \\ \dfrac{dR(t)}{dt} = \gamma I(t) \\ J(t) = I(t) + R(t) = N - S(t). \end{cases} \tag{5}$$

in Eq. (5), $J(t)$ denotes the number of infected end hosts at time t, including the infectious and removed end hosts; $R(t)$ denotes the number of removed end hosts from previously infectious end hosts at time t; $I(t)$ denotes the number of infectious end hosts at time t; $S(t)$ denotes the number of susceptible end hosts at time t; N is the total number of vulnerable end hosts; γ is the removal rate of the infectious end hosts. In particular, $\beta(\tau)$ is the infection rate, which is dynamic and determined by the impact of infection delay τ. The infection delay represents the time required by a worm to infect another susceptible end host from some infectious end hosts.

Although $\beta(\tau)$ may be affected by other factors, such as $I(t)$, in order to strongly indicate the effectiveness of the mapping delay, we have no regard of other factors and only define $\beta(\tau)$ as

$$\beta(\tau) = \beta_0 f(\tau). \tag{6}$$

where β_0 is the initial infection rate, which is universal and constant; $f(\tau)$ is the function of the infection delay τ. In general conditions, when the infection delay is longer, the infection rate is lower. Therefore, according to Ref. 25, we define $f(\tau)$ as

$$f(\tau) = e^{-\eta\tau}. \tag{7}$$

where η is used to adjust the infection rate sensitivity. $\eta = 0$ means constant infection rate.

In today's Internet, in order to facilitate the analysis, we assume the infection delay is the constant u. Therefore, the infection rate will be:

$$\beta_I = \beta_0 f(u) = \beta_0 e^{-\eta u}. \tag{8}$$

In networks with locator/identifier separation, we also assume the mapping delay is the constant v. Since the mapping delay is the additional time for worm propagation, the corresponding infection delay can be calculated as $u+v$. Therefore, the infection rate will be:

$$\beta_S = \beta_0 f(u+v) = \beta_0 e^{-\eta(u+v)}. \tag{9}$$

5. Numerical Analysis and Discussion

In this section, by using some factual data, we give a quantitative comparison between today's Internet and networks with locator/identifier separation. We mainly analyze the numerical solutions of the above-mentioned equations and discuss the mitigation of worm propagation in networks with locator/identifier separation. Besides, we also give some discussions on worm detection based on computational intelligence in networks with locator/identifier separation.

5.1. *Address semantics*

In this simulation, we use the real parameters which simulate the Code Red v2 worm from Ref. 29. We assume there are 500,000 vulnerable end hosts in the Internet, and the worm starts on a single end host, namely $n_0 = 1$. In addition, the worm performs 2 scans per second and takes one second to infect an end host. We also set the death rate $d = 0.00002$ /second and the patching rate $p = 0.000002$ /second. In today's Internet, we assume the worm can use the information provided by Border Gateway Protocol (BGP) routing tables and Class A address allocations. Therefore, in accordance with the advice in Ref. 22, we can draw a conclusion that the size of scanning space Ω is $2^{32}/3.5$ and $2^{32}/2.21$, respectively. In networks with locator/identifier separation, since locator/identifier separation has resolve the dual semantics of IP address, the worm only use the random scan and the size of scanning space is 2^{32}. By Eq. (1) and Eq. (2), we compare the different numbers of infected end hosts in Fig. 2.

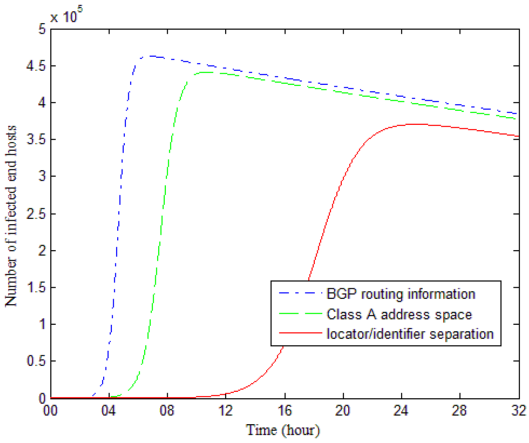

Fig. 2. Number of infected end hosts with different scanning spaces.

From Fig. 2, we can see that the total number of infected end hosts in networks with locator/identifier separation is smaller than that in today's Internet. When the scanning space is larger, the number of infected end hosts is smaller, and the worm propagation rate is slower. For example, it takes 20 hours for the worm in networks with locator/identifier separation to infect 300,000 end hosts, while it only takes 4 hours and 8 hours for the worms in the cases of BGP routing information and Class A address space to infect the same number of end hosts, respectively. In particular, if the worm also uses the unwise random scan, the curve in today's Internet will be the same with that in networks with locator/identifier separation. However, this is a particular case, and we believe that the sensible worms may try to avoid this situation in order to spread faster. Therefore, we can come to the conclusion that the change of address semantics can help to mitigate worm propagation.

5.2. *Address space*

In this simulation, we compare the number change of infected end hosts with different identifier probabilities. We assume the identifier probability q is 0.7, 0.8 and 0.9, respectively, and the assumption of other parameters is the same with that in the above simulation. By Eq. (3) and Eq. (4), we give the curves with different identifier probabilities in Fig. 3. Especially, in today's Internet we can consider the identifier probability as $q = 1$. From this figure, we can observe that in today's Internet, namely when the identifier probability is 1, the number of infected end hosts is the largest. However, when the identifier probability decreases from 0.9 to 0.7, the number of infected end hosts is also significantly reduced. That is to say, in order to rein in worm spread, we may attempt to reduce the identifier space. However, it may be inconsistent with the scarce address space. Therefore, we should find a trade-off between them in practice.

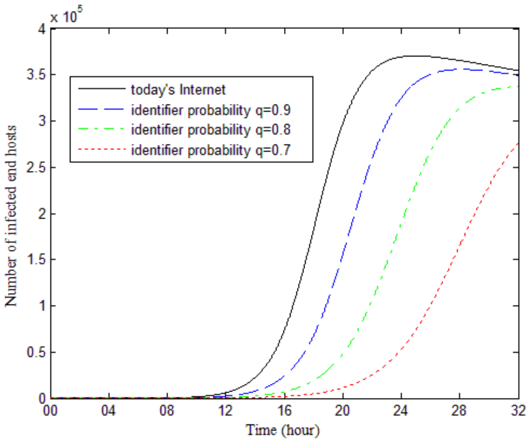

Fig. 3. Number of infected end hosts with different identifier probabilities.

5.3. *Mapping delay*

As mentioned before, the mapping delay can reduce the infection rate, and exert an influence on worm propagation. Like Ref. 27, we set $N = 1,000,000$, $\eta = 0.1$, $\gamma = 0.05$ and $\beta_0 = 0.8/N$. From Ref. 30 we can find that Slammer infected approximately 75,000 Microsoft SQL Servers, and the number of infected end hosts doubled every 8.5 seconds. Thus we assume the infection delay u is 8.5 seconds in today's Internet. In addition, the average mapping delay in each mapping service is significantly different. For example, it is reported in Ref. 31 that the median mapping delay in LISP+ALT mapping service, in LISP-DHT (recursive mode) mapping service and in LISP-DHT (iterative mode) mapping service is about 0.5 second, 1 second and 2 seconds, respectively. Therefore, in this simulation, we use the above median mapping delay as the mapping delay v in Eq. (9), and the corresponding infection delay in these three mapping services can be calculated as 9 seconds, 9.5 seconds and 10.5 seconds, respectively. By Eq. (5), Eq. (8) and Eq. (9), we show the number of infected end hosts in Fig. 4 and the number of infectious end hosts in Fig. 5.

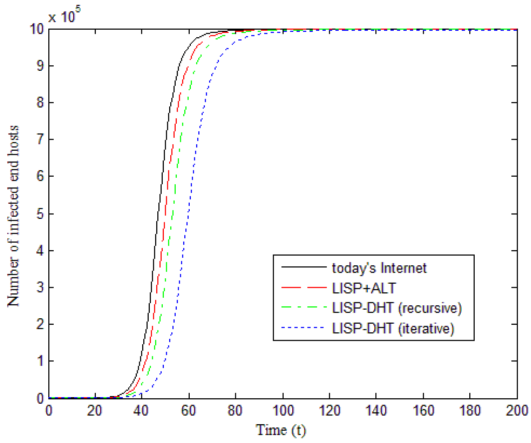

Fig. 4. Number of infected end hosts with different mapping delays.

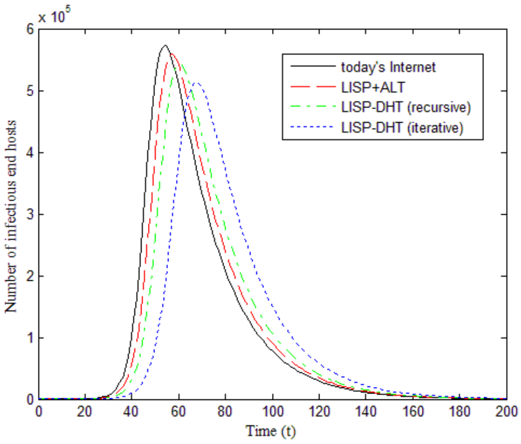

Fig. 5. Number of infectious end hosts with different mapping delays.

In Fig. 4, when the mapping delay is longer, the number of infected end hosts is smaller. For example, the number of infected end hosts at time tick 60 in today's Internet without any mapping service is about 950,000, while the number of infected end hosts with LISP-DHT (iterative mode) mapping service is reduced to approximately 500,000. Namely, the longer the mapping delay is, the lower the worm propagation rate becomes. At the same time, the number of infectious end hosts in Fig. 5 is also affected by the mapping delay: the longer mapping delay makes the maximum number of infectious end host smaller. That is because in Eq. (5) $\beta(\tau)$ is a main parameter which changes the number of infectious end host $I(t)$. Although the mapping delay can mitigate worm propagation, we do not advocate controlling worm spread by increasing the mapping delay, because it also has a significant influence on the normal and usual communications.

To sum up the above arguments, the characteristics of locator/identifier separation, including the changes of address semantics, address space and mapping delay, play an important role to mitigate worm propagation. Meanwhile, networks with locator/identifier separation are more resistive to worm propagation than today's Internet.

5.4. *Discussion on worm detection based on computational intelligence*

Although worm propagation can be obviously mitigated in networks with locator/identifier separation, worms will still exist and the mitigation of worm propagation may bring about some difficulties to detect worms. However, worm detection based on computational intelligence can also benefit from the paradigm of locator/identifier separation.

First of all, the tunnel router (TR) is the most crucial infrastructure in networks with locator/identifier separation, and almost all the communications among end hosts are accomplished by them. Therefore, multi-agent technology can be applied in the tunnel routers to detect worms. In order to make them more intelligent, these agents can be combined with different fields like artificial intelligence, neural networks, fuzzy logic, etc. By using co-operative intelligent agents distributed in the tunnel routers, worm detection can be more feasible and impactful.

Secondly, when worms or other threats want to launch an attack, the tunnel routers must first send the mapping requests to resolve the corresponding locators. Therefore, worms or some other anomalous behaviors can be identified and diagnosed by the mapping request traffic. Since the mapping request traffic, unlike the complicated and high-dimensional network traffic, is simple and single dimensional, the anomaly detection based on neural networks, machine learning or data mining could significantly reduce the false alarm rate.

Finally, the change of address semantics and address space in networks with locator/identifier separation can lower the detection limit of worm detection based on computational intelligence, and improve the detection efficiency.

6. Conclusion

This paper aims to argue the influence on worm propagation due to the incoming locator/identifier separation. The most significant benefit of our study is that it systematically analyzes the mitigation of worm propagation in networks with locator/identifier separation, and provides a basic prerequisite for worm detection based on computational intelligence in networks with locator/identifier separation. In this paper, we first introduce networks with locator/identifier separation by a general network model. Then, compared with today's Internet, we mathematically model the mitigation of worm propagation in networks with locator/identifier separation, with focus on the following three aspects: address semantics, address space and mapping delay. Last but not least, by the numerical analysis, we give a quantitative comparison between today's Internet and networks with locator/identifier separation. We find that, the characteristics of locator/identifier separation could contribute a lot to mitigating worm propagation in networks with locator/identifier separation.

Acknowledgements

This work is supported in part by National Nature Science Foundation of China (Grant No. 60833002, No. 60903150 and No. 61100219), in part by project from MIIT of China (Grant No. 2011ZX03002-005-03), and in part by the Fundamental Research Funds for the Central University (Grant No. 2011JBM016 and No. 2011JBM012). The authors are grateful to the anonymous referees for their insightful comments and suggestions.

References

1. I. Vural and H. S. Venter, Using network forensics and artificial intelligence techniques to detect Bot-nets on an organizational network, in *Proc. 7th int. Conf. Inf. Technol.: New Generations* (USA, Las Vegas, 2010), pp. 725-731.
2. G. Kolaczek and K. Juszczyszyn, Attack pattern analysis framework for multiagent intrusion detection system, *Int. J. Comput. Intelligence Syst.* **1**(3) (2008) 215-224.
3. D. Meyer, L. Zhang and K. Fall, Report from the IAB workshop on routing and addressing, *IETF Internet Standard, RFC 4984*, 2007.
4. T. Li, Recommendation for routing architecture, *IETF Internet Standard, RFC 6115*, 2011.
5. I. Stoica, D. Adkins, S. Zhuang, S. Shenker and S. Surana, Internet indirection infrastructure, *IEEE Trans. Networking* **12**(2) (2004) 205-218.
6. R. Moskowitz, P. Nikander, P. Jokela and T. Henderson, Host identity protocol (HIP), *IETF Internet Standard, RFC 5201*, 2008.
7. D. Farinacci, V. Fuller, D. Meyer and D. Lewis, Locator/ID separation protocol (LISP), *IETF Internet Draft, draft-ietf-lisp-15.txt*, 2011.
8. R. Whittle, Ivip (Internet vastly improved plumbing) architecture, *IETF Internet Draft, draft-whittle-ivip-arch-04.txt*, 2010.
9. E. Lear, NERD: A not-so-novel EID to RLOC database, *IETF Internet Draft, draft-lear-lisp-nerd-08.txt*, 2010.
10. L. Mathy and L. Iannone, LISP-DHT: towards a DHT to map identifiers onto locators, in *Proc. 2008 ACM CoNEXT Conf.* (Spain, Madrid, 2008), pp. 1-6.

11. V. Fuller, D. Farinacci, D. Meyer and D.Lewis, LISP alternative topology (LISP+ALT), *IETF Internet Draft, draft-ietf-lisp-alt-08.txt*, 2011.
12. O. Ponomarev and A. Gurtov, Embedding host identity tags data in DNS, *IETF Internet Draft, draft-ponomarev-hip-hit2ip-04.txt*, 2009.
13. H. B. Luo, Y. J. Qin and H. K. Zhang, A DHT-based identifier-to-locator mapping approach for a scalable internet, *IEEE Trans. Parallel Distrib. Syst.* **20**(2) (2009) 1790-1802.
14. H. B. Luo, H. K. Zhang and M. Zukerman, Decoupling the design of identifier-to-locator mapping services from identifiers, *Comput. Networks* **55**(4) (2011) 959-974.
15. P. K. Manna, S. Chen and S. Ranka, Exact modeling of propagation for permutation-scanning worms, in *Proc. 27th IEEE Int. Conf. Comput. Commun.* (USA, Phoenix, 2008), pp. 1696-1704.
16. C. C. Zou, N. Duffield, D. Towsley and W. Gong, Adaptive defense against various network attacks, *IEEE J. Sel. Areas Commun.* **24**(10) (2006) 1877-1888.
17. D. Moore, C. Shannon and J. Brown, Code-Red: a case study on the spread and victims of an Internet worm, in *Proc. 2nd Int. Meas. Workshop* (France, Marseille, 2002), pp. 273-284.
18. D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford and N. Weaver, Inside the Slammer worm, *IEEE Secur. Priv.* **1**(4) (2003) 33-39.
19. eEye Digital Security, Blaster worm analysis, *http://www.eeye.com/html/Research/Advisories/AL20030811.html*, 2003.
20. C. Shannon and D. Moore, The spread of the Witty worm, *IEEE Secur. Priv.* **2**(4) (2004) 46-50.
21. W. Yu, C. Boyer, S. Chellappan and D. Xuan, Peer-to-peer system-based active worm attacks: modeling and analysis, in *Proc. IEEE Int. Conf. Commun.* (Korea, Seoul, 2005), pp. 295-300.
22. C. C. Zou, D. Towsley, W. Gong and S. Cai, Routing worm: a fast, selective attack worm based on IP address information, in *Proc. Workshop Prin. Adv. Distrib. Simul.* (USA, CA., 2005), pp. 199-206.
23. R. McMillan, Conficker worm sinks French navy network, *http://www.pcworld.com/article/159224/conficker_worm_sinks_french_navy_network.html*, 2009.
24. W. Yu, X. Wang, P. Calyam, D. Xuan and W. Zhao, Modeling and detection of Camouflaging worm, *IEEE Trans. Depend. Secure Comput.* **8**(3) (2011) 377-390.
25. Y. Wang and C. Wang, Modeling the effects of timing parameters on virus propagation, in *Proc. WORM'03* (USA, Washington, 2003), pp. 61-66.
26. J. C. Frauenthal, *Mathematical Modeling in Epidemiology*, (Springer-Verlag, New York, 1980).
27. C. C. Zou, W. Gong and D. Towsley, Code Red worm propagation modeling and analysis, in *Proc. 9th ACM Symp. Comput. Commun. Secur.* (USA, Washington, 2002), pp. 138-147.
28. J. Kim, S. Radhakrishnan and S. K. Dhall, Measurement and analysis of worm propagation on Internet network topology, in *Proc. 13th Int. Conf. Comput. Commun. Networks* (USA, Chicago, 2004), pp. 495-500.
29. Z. Chen, L. Gao and K. Kwiat, Modeling the spread of active worms, in *Proc. 22nd Annu. Joint Conf. IEEE Comput. Commun. Societies* (USA, San Francisco, 2003), pp. 1890-1900.
30. M. Costa, J. Crowcroft, M. Castro, A. Rowstron, L. Zhou, L. Zhang and P. Barham, Vigilante: end-to-end containment of Internet worms, in *Proc. 20th ACM Symp. Oper. Syst. Prin.* (United Kingdom, Brighton, 2005).
31. F. Coras, CoreSim: a simulator for evaluating LISP mapping systems, *Technical Report, Technical University of Cluj-Napoca, http://lisp-2.cba.upc.edu/2009-fcoras-thesis.pdf*, 2009.