MIB-ITrace-CP: An Improvement of ICMP-Based Traceback Efficiency in Network Forensic Analysis

Bo-Chao Cheng[1], Guo-Tan Liao[1], Ching-Kai Lin[1], Shih-Chun Hsu[1], Ping-Hai Hsu[2], and Jong Hyuk Park[3]

[1] Dept. of Communications Engineering, National Chung Cheng University, Taiwan
[2] Information and Communications Research, ITRI, Taiwan
[3] Dept. of Computer Science and Engineering, SeoulTech, Korea
bcheng@ccu.edu.tw, becker@itri.org.tw,
{loboyoh,hisa918203}@gmail.com,
{ganes0503,parkjonghyuk1}@hotmail.com

Abstract. A denial-of-service (DoS) / distributed-denial-of-service (DDoS) attack may result in rapid resource depletion along the attack path. For stepping-stone and masquerading techniques typically used in DoS/DDoS attacks such as internet protocol (IP) or Media Access Control (MAC) address spoofing, tracing the intrusion back to the true attacker becomes a challenging task for network security engineers. Although the Internet Engineer Task Force (IETF) has proposed an Internet Control Message Protocol (ICMP) based Traceback solution, it faces severe difficulties in practice in regard to justifying the interoperability of deployed routers as well as the correctness of Traceback with multiple attack paths. This research proposes a novel approach to embed the essence of a management information base (MIB) into iTrace messages, named MIB-ITrace-CP, in order to improve the accuracy and efficiency of the original ICMP-based Traceback. Through our implementations on a Testbed@TWISC platform, we validated our approach and demonstrated the feasibility of practical network forensics.

Keywords: DoS, Spoofing, Forensics, Traceback, ITrace-CP.

1 Introduction and Background

For Internet service providers (ISP), a denial-of-service (DoS) / distributed-denial-of-service (DDoS) attack is one of the more difficult problems faced. This is because attackers can attack either using a real or a fake internet protocol (IP) address, rendering it difficult to find the real attacker and prevent the attack. Identification of the attackers' IP address is even not possible for computer forensics [1] nowadays. In conventional IP networks, there are three famous Traceback models that provide fundamental foundations on a range of different researches: Logging [2], Probabilistic Packet Marking (PPM) [3] and ICMP-based Traceback [4] discussed by IETF [5].

With respect to various ICMP-based Traceback [6-8], we found that there are some weaknesses. For example, mobile attacker, multiple attack sources, multiple attack

J.J. Park et al. (Eds.): NPC 2012, LNCS 7513, pp. 101–109, 2012.

paths and spoofing attack make Traceback more difficult to correctly locate the attacker's address over multiple paths. Here, we show the progressive improvement of ICMP-based Traceback mechanisms as below. Initially, this ICMP Traceback (ITrace) idea was presented as an industry standard of IETF (Internet Engineering Task Force) on 2000, available at draft-ietf-itrace-00.txt. In 2001, S. Felix Wu et al. proposed intention-driven ITrace [9], whose concept is to use an extra intention bit in the routing (e.g., the community attribute in BGP routing table) for controlling the forward ITrace option to achieve a much better tracing performance about the statistic problem of ITrace. The latest version of ITrace draft "draft-ietf-itrace-04.txt" [4] was updated in 2003. Then in 2007 A. Izaddoost et al. [10] proposed an accurate ICMP Traceback model based on intention-driven ITrace to reconstruct attack paths accurately by generating more effective ICMP Traceback packets. The ITrace message is emitted randomly by routers along the path and sent randomly to the destination (to provide useful information to the attacked party) or to the origin (to provide information to decipher reflector attacks).

On the other hand, Henry C.J. Lee et al. in 2003 proposed a so-called ICMP Traceback with Cumulative Path (ITrace-CP) [6], which is an enhancement of the ITrace approach. ITrace-CP messages are made to carry a part of entire attack path information, the same as ITrace messages, so as to accelerate the attack path construction in the event of a DDoS attack. In ITrace-CP, it proposed three kinds of schemes, where the scheme 3 is "Hash-based Packet Identification with indicator bit" involving three mechanisms: 1) Basic Packet Identification (BPI), 2) Hash function, 3) Indicator bit. About scheme 3 of ITrace-CP [6], it solves the problem of how to identify corresponding IP packets and ITrace-CP messages, and reduces the storage requirement.

With reference to ITrace draft [4], the probability of Traceback generation should not be greater than 1/1000, adjustable by the operator of the router. In 2005, V. Thing [7] proposed the distribution of generation probability in an exponential manner. For example, the probability at each router is computed by: $p=d^x/c$ where d is the distance from current router to the victim, x is the exponent and c is a constant. The enhanced ITrace-CP [7] shows an idea that the furthest router from the victim has the highest probability, and has better Traceback performance than ITrace when the distance between an attacker and a victim is near the diameter of a network. Inspired from that Traceroute works by increasing the Time To Live (TTL) value of each successive set of packets sent until the destination host receives the packets and returns an ICMP Echo Reply message, H. Tsunoda et al. [8] propose a countermeasure against TTL spoofing by TTL-based calculation of generation probability of ITrace-CP messages.

As mentioned above, conventional ICMP-based Traceback approaches have weaknesses in the issues of multiple attack sources and multiple attack paths to correctly identify the attacker's location. We proposed MIB-ITrace-CP approach, which embeds Management Information Base (MIB) [11] information into ITrace-CP message combined with benefits of various ICMP-based Traceback schemes including the Hash-based Packet Identification scheme [6], intention-driven model [9], TTL-based calculation of probability [8]). The goal of MIB-ITrace-CP is to improve the ICMP-based Traceback's correctness and efficiency, and we have used Testbed@TWISC [12] as the platform for our experiments.

2 System Model and Assumption

In the original ITrace-CP method, due to a flaw in the algorithm, it does not work well in a diversified topology. In particular, serious mistakes can result with multiple paths. Thus, to improve the original solution, we propose our MIB-ITrace-CP that not only applies the Hash-bashed Packet Identification (HPI) with indicator bit, the same as intention-driven model, but also embeds extra MIB data into original ITrace-CP message. We consider a scheme for the generation probability of MIB-ITrace-CP message. Firstly, the main assumptions are as follows:

(**A1**) All MIB-ITrace-CP messages (m) should store the source (S), the destination (D), the initial originator ($\overline{S}$) of m and which interface (ifName), a path (from $\overline{S}$'s previous node (N_P) to the last expected node), and two hash values for a packet (P). In addition, if $\overline{S}$ is just the source S, the N_P field will be filled in with itself ($\overline{S}$). The intention-driven hash $H_R^1(P)$ is varied with the current generator of MIB-ITrace-CP message (N_G), but the flow-classification hash $H_R^2(P)$ is unique and determined by $\overline{S}$ during the transmission to D.

(**A2**) Furthermore, using the Management Information Base (MIB) [RFC 1156, 1213, 2863], it enables routers to handle more efficient information. Besides of the interface identifier mentioned in A1 (such as "ifName", N_P and N_N (the next node of the current node from the routing table)), we use another two external pieces of data which are "ifEntry::ifSpeed" and "ifEntry::ifOutUcastPkts" provided by MIB module. By providing more information for the victim to perform better judgments, it also increases the accuracy in determining the attacker's real address and would not cause a problem for computer forensics. The two data is detailed as follows:

- ifEntry::ifSpeed (interfaceSpeed): Provide every connected device's bandwidth and take "bits per second" as the unit of measurement. ifSpeed (S_i) will represent the current operational speed of the interface in bits per second. In other words, the ifSpeed object defined in MIB-II's interfaces table provides "an estimate of the interface's current bandwidth in bits per second".
- ifEntry::ifOutUcastPkts: Provide the number of packets in the high-level protocol. What we want is to get the packet number (P_i) sent between two subsequent MIB-ITrace-CP messages originated from the same interface that can be calculated by a register).

(**A3**) The additional packet marking information, embedded into a packet (P), is the indicator bit, which indicates that a MIB-ITrace-CP message has been generated for a specific IP packet P, as well as the intention-driven hash $H_R^1(P)$ when the indicator bit (*ITRACE_CP_DONE*) is set.

Typically, the so-called Basic Packet Identification (BPI), set by the source ($\overline{S}$), is a value ($UN_{TimeWindow}$) that must be unique for that source-destination pair and protocol for the time the packet will be active in the Internet. HPI is the hash of BPI, used to reduce the storage requirement instead of storing the BPI of a packet. In this paper, we use two HPI of different context information for intention-driven model and flow

classification. We define the 1st hash value, calculated in the router R and whose four inputs are N_G, D, *Protocol* from m and the new N_N from the routing table of R for P, is the so-called intention-driven hash $H_R^1(P)$. The 2nd hash value, calculated in the router R and whose four inputs are S, D, *Protocol* of the packet P and $UN_{TimeWindow}$ for the flow of P, is the so-called flow-classification hash $H_R^2(P)$. And, for intention-driven check in each router, there is a table T kept for a short time period and composed of two attributes, the intention-driven hash value ($H_R^1(P)$) and the next hop (N_N) corresponding to a packet (P).

We propose MIB-ITrace-CP to improve the ICMP-based Traceback. If the victim experiences DoS attacks in traffic, MIB-ITrace-CP method can still determine the attacker's real address with high accuracy and the path it passed through. The packets received by a victim can be separated into three kinds according to their types: 1) packets without the indicator bit set, 2) packets with the indicator bit set, 3) MIB-ITrace-CP messages with MIB information for attack graph reconstruction. And, the system architecture of MIB-ITrace-CP involves three parts, namely "Originating", "Forwarding" and "Path reconstruction", described as follows.

- Implementation of Generation Probability

Originally, all routers use the same probability for generating Traceback messages in ITrace. In order not to cause heavy overhead, it is suggested that the probability should not be greater than 1/1000. As the average maximum diameter ($H_{\max}$) of the Internet is 20 hops, a default value of probability about 1/20000 is suggested. Inherited from the improvement of previous generation probability [7, 8], we adopt the idea of TTL value to determine a probability p. And we propose a practical method by the assessment of $H_{\max}$ to calculate the probability for generating MIB-ITrace-CP messages. In theory, the equations are formulated as Eq. (1), and we specially focus on the exponent x=1 and modify as shown in Eq. (2).

$$p = d^x / c, \quad \sum_{d=1}^{H_{\max}} (d^x / c) = 1/1000 \tag{1}$$

$$p = d / c, \quad c = 500 \cdot H_{\max} \cdot (H_{\max} + 1), \quad d = H_{\max} - d_{src} \tag{2}$$

As for easy calculation for adjusting $H_{\max}$, using the case x=1 of Eq. (1) has its advantages. The near from the flow source, the higher probability for generating MIB-ITrace-CP messages. And, the source node also adopts the highest probability as ${H_{\max}}^x / c$. In overall, using the view of d_{src} to set p is better than that from d_{dst} (to destination) because it can promote the efficiency in terms of Traceback time. Furthermore, it can integrate with judgment of the regular TTL value (255, 128, 64 or 32) to defeat against TTL spoofing.

- Forwarding of Traceback Message

Moreover, a key point is to achieve the intention-driven model by simply comparing intention-driven hash values of the packet and the corresponding MIB-ITrace-CP message (m). After identification of the expected m, a router R decides to send a new

m' embedded *R*'s address to N_N if matched or send a new message *m*' to *D* without making any changes to the payload of *m* if mismatched. Especially when the new expected MIB-ITrace-CP message (*m'*) is generated, it should be transmitted to N_N of *T* as far as possible. In detail, the *m*' of transmitted to the expected next hop keeps the original information of *m*, including *S*, *D*, $\overline{S}$, *Protocol*, $H_{\overline{S}}^{2}(P)$, P_i, S_i and the route path from $\overline{S}$'s previous hop to *R*'s previous hop (N_P), and then also embeds *R*'s address and $H_R^1(P)$.

- Path Reconstruction

Firstly, speaking about two categories of MIB-ITrace-CP messages received at a victim, one category (C_1) with the corresponding intention-driven hash $H_R^1(P)$ value is sent triggered from the previous node (NP), and the other (C2) with mismatched $H_R^1(P)$ is forwarded by N_P. Fig. 1 shows how path reconstruction is done through the MIB-ITrace-CP system, how bandwidth is set to judge whether DoS attacks have occurred or not, and how collected information from the victim is used, threshold calculation, data comparison in MIB-ITrace-CP messages, and filtering of the MIB-ITrace-CP messages produced by DoS attacks. Using this pseudo code, we thus improved on the ITrace-CP, enabling it to work with multiple paths and increasing its efficiency. With the information of MIB-ITrace-CP messages (C_1), it can help reconstruct more exact attack graph. On the other hand, as a mobile attacker sends a huge number of packets to the victim, the victim can infer that attacker's movement by those MIB-ITrace-CP messages (C_2).

After the victim has finished collecting MIB-ITrace-CP messages, it performs path reconstruction and tries to find the attacker's real address using the information given by MIB-ITrace-CP messages. MIB-ITrace-CP is an algorithm based on ICMP-based Traceback's basic structure and addresses the limitations of ICMP-based Traceback. Inside this part, apart from path reconstruction, it also contains a filter. The filter's main function is to filter the MIB-ITrace-CP messages (*m*) produced by DoS attacks. It is useful, by filtering the message m_i with Ψ_i lower than a threshold value Ψ_θ, for finding the correct attack path from multiple paths and filters using Ψ_θ as shown in Eq. (3) where α represents the weighted factor that is defined and executed by the administrator or the intrusion detection system, P_{avg} represents the average of P_i and S_{avg} represents the average of S_i. For example, for SYN attacks, Teardrop attacks or DoS-like attacks, if SYN flood attacks are found, α=0.8. This is because SYN flood attacks will produce a large amount of packets. If the α value is higher, it means that P_{avg} weight will be higher and the selection will be more accurate. Conversely, if Teardrop attacks occur, then α=0.2. This is because Teardrop attacks may create a series of IP fragments with overlapping offset fields in a high network traffic flow. As the α value is lower, the S_{avg} weight ($1-\alpha$) is also higher and the threshold value will be more accurate. For MIB-based calculation, each MIB-ITrace-CP message (m_i) includes the two MIB information, P_i and S_i.

$$\begin{cases} \Psi_\theta &= \alpha \cdot P_{avg} + (1-\alpha) \cdot S_{avg} \\ \Psi_i &= \alpha \cdot P_i + (1-\alpha) \cdot S_i \end{cases} \tag{3}$$

Path reconstruction procedure at victim V:
Two categories of MIB-ITrace-CP messages: C_1 and C_2
C_1 is triggered from the previous node N_P;
C_2 is forwarded by N_P with failure of intention-driven check
List a table sorted according to the master column "S-D pair" (S, D), $H_R^2(P)$ and C_i
Get the suspect ITrace-CP messages with key $H_R^2(P)$ and (S, D)
Select the path of C_1 whose entry has $H_R^2(P)$ to form an attack graph G_1 (reliable)
Select the path of C_2 whose entry has (S, D) to form an attack graph G_2 (inferable)
Execute MIB-based calculation of Ψ_i for C_1 in according to the parameter α of Ψ_θ by Eq.(3)
Select the paths, filtered by comparing Ψ_i with Ψ_θ, to form an attack graph G_3 (reasonable)
Identify attack paths and attackers' location by G_1, G_2 and G_3

Fig. 1. Pseudocode of attack path reconstruction by MIB-based ITrace-CP at a victim

- Example

For example, as shown in Fig. 2, there are two source nodes that generate DoS attacks (Node A) and normal traffic (Node H) respectively. The victim collects all the packets and five MIB-ITrace-CP messages (m_i) with path information in the example, as shown in Table 1. The convention ICMP-based approaches are not able to identify the real attach path (A→B→C→D→E→F) based on these five ITrace-CP messages because the victim collects a normal traffic flow path (H→I→J→C→D→E→F) at the same time. Now, we show how MIB-ITrace-CP to solve this multiple-path problem via extra MIB information (P_i and S_i) embedded in MIB-ITrace-CP messages. First, we can obtain Ψ_θ and Ψ_i based on Eq.(3) as α=0.8. Since Ψ_3 and Ψ_5 are less than Ψ_θ, m_3 and m_5 would be skipped. As such, m_1, m_2 and m_4 would be noticed as the attack graph G_3 (reasonable) for path reconstruction.

In summary, it yields an attack graph (G_1) which can trace back to node A and node H. It is possible to find out the real attacker as well as finding a regular source which makes huge traffic. Moreover, MIB-ITrace-CP with the availability of more information would make the analysis more effective, as only the attaching path (A→B→C→D→E→F) in the attack graph G_3.

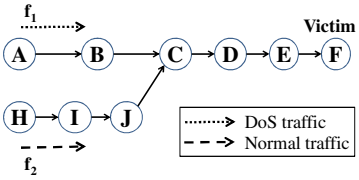

Fig. 2. Example of path reconstruction

Table 1. Example of MIB model calculation

m_i	S-D	$H_R^2(P)$	C_i	$\overline{S}$	*Interface*	*TTL*	P_i	S_i (bps)	Ψ_i	Cumulative Path
1	A-F	x	C_1	B	B_1	61	135000	20000	112000	A→**B**→C→D→E
2	A-F	x	C_1	C	C_1	62	130000	21000	108200	B→**C**→D→E
3	I-F	x	C_1	C	C_2	62	60000	5000	49000	J→**C**→D→E
4	I-F	x	C_1	E	E_1	64	170000	26000	141200	D→**E**
5	I-F	x	C_1	I	I_1	60	50000	6000	41200	H→**I**→J→C→D→E
							109000 (P_{Avg})	15600 (S_{Avg})	90320 (Ψ_θ)	← Ψ_θ (as α=0.8)

3 Experiment and Analysis

In this section, we first discuss the research application and the flow of the experiments. Our research used Testbed@TWISC [12] as the experimental environment. Testbed@TWISC (Taiwan Information Security Center) provides an integrated lab environment that fulfills the requirements of being quarantined, closed, recordable, controllable and storable for researchers. In accordance with the MIB-ITrace-CP system, the experiment was built to verify that path reconstruction under DoS attacks will be reconstructed effectively. The simulated environment uses a multiple path topology as shown in Fig. 3, similar with [10]. Next, we make experiments and show the analysis of effectiveness on multiple traffic flows. And, MIB-ITrace-CP is conducted under the parameters x=1, H_{max}=20 and c=210000 in Eqs. (1)–(2).

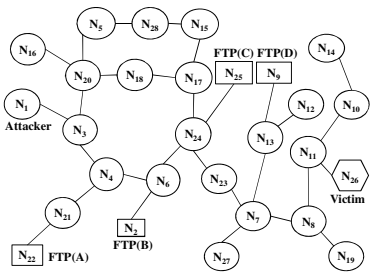

Fig. 3. Topology of the experimental network

In our experiments, we only discuss DoS attack by MIB-ITrace-CP with α=0.8. For the experiment of multiple paths, an attack path from the attacker (N_1) to the victim (N_{26}) is $N_1 \rightarrow N_3 \rightarrow N_4 \rightarrow N_6 \rightarrow N_{24} \rightarrow N_{23} \rightarrow N_7 \rightarrow N_8 \rightarrow N_{11} \rightarrow N_{26}$, and there are four FTP servers (N_2, N_9, N_{22} and N_{25}) in Fig. 3. The FTP traffic is simulated by the download of data with 10Mbps from each FTP server to the victim. And, the attack traffic is launched by SYN flood attack of 2500 SYNs/sec. Firstly, as shown in Fig. 4(a), the x-axis represents the number of FTP servers and the y-axis represents the number of packets received at the victim until it gets an MIB-ITrace-CP message with full path for tracing the DoS attacker. The performance metric is expressed as the convergence

time as the minimum threshold number of packets required. In this scenario, because the minimum hop from the source (N_1) to the victim (N_{26}) is only 9, so the results of ITrace-CP are better than that of the Enhanced ITrace-CP as the parameters x=1, H_{max}=20 and c=210000. On an average, the MIB-ITrace-CP method has the best performance and can get the key Traceback message of full path in the shortest time.

Secondly, as shown in Fig. 4(b), the x-axis represents the number of FTP servers (b/w=10Mbps) and the y-axis represents the percentage of effective Traceback messages, which are triggered from the attack flow, among total Traceback messages received at the victim as the attacker sends out 150,000 SYN packets. Here, please note the performance evaluation for MIB-ITrace-CP is based on the results which are filtered by MIB-ITrace-CP's path reconstruction algorithm presented in Fig. 1. In sum, the MIB-ITrace-CP method has the best performance due to efficient filter although it can quickly gather more Traceback messages for path reconstruction. The results show that MIB-ITrace-CP can perform good Traceback performance and support effective path identification.

The main advantage of MIB-ITrace-CP is that could trace back to multiple attackers quickly and effectively as well as under the condition that normal traffic occurred simultaneously by a weighted α value for different types of attack to facilitate Traceback identification.

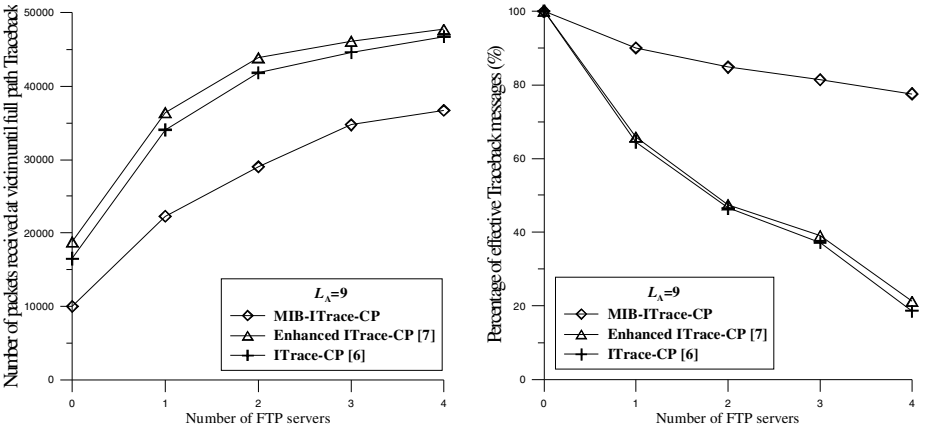

Fig. 4. Efficiency analysis of multiple flows among SYN attacks (a) Traceback convergence time (b) percentage of effective Traceback message

4 Conclusion

Upon being attacked on the Internet, all victims desire to locate the real attacker. For network forensics, in this paper, we use information provided by MIB embedded in ITrace-CP message. As DoS attacks usually generate a large amount of traffic flow or packets to debilitate a server's connection, the additional information enables filtering of Traceback packets and is useful for promoting the efficiency of ITrace-CP. The

addition of the MIB system is not only lightweight, but also helps the victim gather more information to forensic teams. Further, this paper proposes a flow-classification concept, working with intention-driven, to promote the accuracy of attack path reconstruction. And, the experiments on Testbed@TWISC platform ensures that the results obtained would be closer to real-world conditions.

Acknowledgement. This research was supported by the Industrial Technology Research Institute, Taiwan.

References

1. US-CERT, Computer Forensics (2008), http://www.us-cert.gov/reading_room/forensics.pdf
2. Snoeren, A.C., Partridge, C., Sanchez, L.A., Jones, C.E., Tchakountio, F., Kent, S.T., Strayer, W.T.: Hash-Based IP Traceback. In: SIGCOMM 2001 (August 2001)
3. Savage, S., Wetherall, D., Karlin, A., Anderson, T.: Network Support for IP Traceback. IEEE/ACM Transactions on Networking (TON) 9(3), 226–237 (2001)
4. Bellovin, S., Leech, M., Taylor, T.: ICMP Traceback Messages. Internet Draft (February 2003), http://www.ietf.org/proceedings/03mar/I-D/draft-ietf-itrace-04.txt
5. Internet Engineer Task Force (IETF), http://www.ietf.org/
6. Lee, H.C.J., Thing, V.L.L., Xu, Y., Ma, M.: ICMP Traceback with Cumulative Path, An Efficient Solution for IP Traceback. In: 5th International Conference on Information and Communications Security, pp. 124–135 (October 2003)
7. Thing, V.L.L., Lee, H.C.J., Sloman, M., Zhou, J.: Enhanced ICMP Traceback with Cumulative Path. In: IEEE 61st Vehicular Technology Conference (VTC 2005-Spring), vol. 4, pp. 2415–2419 (2005)
8. Tsunoda, H., Tochiori, T., Waizumi, Y., Kato, N., Nemoto, Y.: Improving the Efficiency of DoS Traceback Based on the Enhanced ITrace-CP Method for Mobile Environment (Invited Paper). In: Third International Conference on Communications and Networking in China (ChinaCom 2008), pp. 680–685 (2008)
9. Mankin, A., Massey, D., Wu, C.L., Wu, S.F., Zhang, L.: On Design and Evaluation of Intention-Driven ICMP Traceback. In: IEEE Int' 10th Conf. Computer Communications and Networks, pp. 159–165. IEEE CS Press (2001)
10. Izaddoost, A., Othman, M., Rasid, M.F.A.: Accurate ICMP Traceback Model under DoS/DDoS ATTACK. In: 15th International Conference on Advanced Computing and Communications (ADCOM 2007), pp. 441–446 (December 2007)
11. IEEE Draft Standard for Management Information Base (MIB) Definitions for Ethernet. P802.3.1/D3.0 (November 2010)
12. Testbed@TWISC, Network Emulation Testbed, http://testbed.ncku.edu.tw/