E-RIHT: Enhanced Hybrid IP Traceback Scheme with 16-bit marking field

Chaitanya Kumar Singh [#1], Srinivas Koppu [#1], V Madhu Viswanatham[#2]

[1#] School of Information Technology and Engineering, VIT University
[2#]School of Computing Science and Engineering, VIT University
Vellore, Tamilnadu, India
[1] srinukoppu@vit.ac.in
[1] chaitanyajadon@gmail.com
[2]vmadhuviswanatham@vit.ac.in

Abstract—**Internet is a worldwide network and used in almost every field of work such as industrial, educational, military etc. Based on the use, its security needs differ. Few applications may need less security and few may need high security. Today various internet attacks are being developed every day, such as viruses, DoS (Denial of Service), spoofing, etc. Spoofing is a kind of attack in which attacker masks itself under some other user's IP address. Thus, it is difficult to find the original attacker. IP traceback technique is used to detect the DoS attack. This paper is based on IP traceback, which will help to detect the spoofing attacker by using packet marking and packet logging technique. In packet marking technique router marks identification information of its own into the forwarded packets. In packet logging, routers keep the digest information regarding the forwarded packets. Proposed scheme is termed as E-RIHT (Enhanced Routers Interface Hybrid Traceback) in this, memory requirement will be less because we are using marking field of 16-bit, which will also solve the packet fragmentation problem.**

Keyword- DoS(Denial of Services), packet marking, packet logging, IP trace back.

I. INTRODUCTION

Internet is growing day by day and the users using it are increasing exponentially. Security becomes an important issue, as internet is being used for the exchange of data transactions and confidential information etc. Among various attacks on internet, this paper is focused on DoS attack. DoS attack is classified into flooding attacks and software exploit [1][5]. In flooding attack, large number of packets flooded to the victim machine. Due to highly distributed nature of DoS attacks, victim can be overwhelmed quite easily, even if individual attackers send low number of packets to the victim. Software exploit attack, attacks a host using host's vulnerabilities with few packets. Software exploit attacks include IP spoofing attack. Spoofed packets are traced via traceback scheme by augmenting the packets with partial information called as packet marking and also by storing the packet digest or signature at intermediate routers called as packet logging [2][8]. In such type of schemes, large number of packets required at victim to traceback the path. There are various problems associated with the IP traceback scheme such as the requirement of high storage on logged routers. Traceback scheme cannot avoid the false positive and false negative problem.

Scheme proposed in [10], real source of flooding based attack can be traced using link test in which UDP service is used to generate access load to the upstream links. The excess load, works against the attack packets and disturb the attack packets traffic. Through the excess load attack traffic can be traced, which passes via upstream router. Link test scheme required access packets, Bellovinetal [18] suggested iTrack scheme. iTrack scheme uses ICMP packet having forward and backward link of the router to control the triggering packet. ICMP messages are collected at victim host, using that attack path is constructed. iTrack scheme make use of packet marking which is divided into two categories, probabilistic packet marking (PPM) and deterministic packet marking (DPM). DPM traceback scheme proposed by Belenky and Ansari [15],[16] states marking the IP addresses of packets passing through the border routers. Full IP address cannot be stored using IP header's identification field alone because of this, IP address is divided into several segments by border router and its digest is computed. Segments and digests are chosen randomly to mark on packets passing. When enough packets are received at destination host, segments are assembled using the digest. Other traceback schemes are probabilistic pipelined packet marking scheme (PPPM) by Al-Duwari and Govindarasu [5], PPM scheme with edge sampling called as FMS by Savage et al. [17], practical packet marking scheme by Gong and Sarac [7]. Probability based schemes that are used for marking of partial path information on the packets passing through the router with a probability. If enough number of marked packets gets collected at victim side then full attack path can be reconstructed. Flooding based traceback schemes requires access amount of attack packets to trace the attack path. These schemes are not favored for tracing attack path in software exploit attacks.

In this paper, we are using IP traceback technique to trace the location of the DoS attacker. IP traceback techniques are classified under two heads, packet marking and packet logging. In packet marking scheme,

routers write their identification in the header field of forwarded IP packets. Packet marking includes little overhead at routers and it requires flow of marked packets to form the network path towards origin. In logging scheme, packets are logged into routers on the network path towards destination. Previously, packet logging was not in practice but now it is being implemented after Hash-Based IP traceback came into existence.

II. LITERATURE SURVEY

There are many techniques for Denial-of-services (DoS) detection. They are activity profiling, change point detection and wavelet-based signal analysis. Following techniques shows, how the DoS flooding attacks can be identified successfully. But all these detection techniques show the desirable result within a limited testing environment. None of the technique is able to solve the DoS attack problem completely [10]. Denial of Services attack can be described as attempt by an individual or a group to disable an online service. DoS attacks can cause major disaster for organizations which are based on online availability to do business like eBay, Amazon, etc. Various large scale attacks have been detected which were targeted on several high profile websites [11]. A lot of research and efforts were made to define various detection mechanisms to control the attacks.

First DoS attack took place in the year 1998. It has become a well known attack due to the losses occurred by affecting the online services. DoS attacks can be broadly classified into three forms, a) Attacks exploiting some vulnerability or implementation bug in the software, implementation of a service to bring that down. b) Attacks that use up all the available resources at the target machine. c) Attacks that consume all the bandwidth available to the victim machine [10]. The third type of attack is called bandwidth attack.

DoS attacks are much more suited to the distributed network because in distributed network, number of hosts can be used to generate huge amount of traffic for the targeted machine [11]. It becomes more difficult and one of the major problem in today's internet to provide protection for such large scale distributed network attack.

Various approaches are used for tracing back the path of DoS attack, so that the attacker can be identified. There is a hybrid approach, proposed by B. Al-Duwairi and M. Govindarasu [5] which is called as DLLT (Distributed Link List Technique). In DLLT track is kept for a subnet of routers involved in the forwarding packets via temporary link between routers using distributed link list. Packets are marked, stored and forwarded based on probability. Benefits of DLLT is the reduced overhead involved as it recovers the path with lesser packets than marking based approach and low storage overhead than logging based approach. DLLT has a demerit as it is based on probabilistic nature, it cannot trace single packet. So, DLLT is useful for DoS traceback only.

PPPM (Probabilistic Pipelined Packet Marking) is proposed by B. Al-Duwairi and M. Govindarasu [5]. PPPM is based on grouping the packets having similar destination addresses. Packet carries path fragments, which are left in the router by previous packets having similar destination. Path fragments are stored in the marking field of packets. There is no requirement for query processing as the path fragments are gathered at destination. It has lower storage overhead as the part of traffic is logged in routers. But PPPM has similar demerit to that of DLLT, it is based on probabilistic nature thus it cannot trace single packet.

C. Gong and K. Sarac [8] proposed another hybrid approach termed as HIT (Hybrid IP Trace back). HIT is based on recording the path fragment of each packet. Marking and logging is done at each router. While every packet passes via router, marking of packets is done deterministically where as logging is being done alternately. Encoding pattern of path fragments in HIT is also different from DLLT and PPPM. DLLT and PPPM make use of 32-bit IP address, which is added with other marking information and used for path recovery that makes it 34-bits for DLLT and 57-bits for PPPM.

Snoeren et al [3] proposed two schemes called packet integrity and SPIE (Source Path Isolation Engine). These schemes are based on hash function. Bloom Filter [6] is used to record hash value and packet time interval, also used to reduce logging storage. So that the logging storage can be reduced and the required storage is about 0.5% of the total flow.

Modified SPIE scheme was proposed by Gong and Sarac [9]. The concept of modified SPIE scheme is that, the transmitted packets will log data in router while passing through every two routers and Its required half of the storage space as compared to Snoeren et al [4] SPIE scheme.

Single packet IP traceback based on routing path proposed by Ning Lu, Yulong Wang, Fangchun Yang, Maotong Xu[14]. It shows the decrease in storage overhead on routers by making traceback path using the label switching method.

RIHT (Routers Interface Hybrid Traceback) was proposed by Ming-Hour Yang and Ming-Chien Yang [12]. In RIHT scheme interface of the routers are marked on the packets, so that path of the packet can be traced. Hash tables are used to store the marking field of the packets due to limitation of marking field on each packet. Packet marking/logging process is repeated till the packet reaches its destination.

Light weight traceback scheme was proposed by the Yan Fen,Zhu Hui , Chen Shuang-shuang , Yin Xin-chun[20]. First 8-bit of ID field and 64-bit optional field values are used as marking field. TTL value is copied

into the first 8-bit of ID field. Starting IP and End IP are copied into the optional field. Then TTL is update at every router and used for marking and traceback.

In PPM scheme [13], 32-bit AS (Autonomous System) number is used for IP traceback. 16-bit identification field is used for packet marking. AS number is converted to hash as probability and use to traceback the attacker.

III. PROPOSED SCHEME

In this paper, proposed IP traceback scheme is termed as E-RIHT which is similar to RIHT [12] with some modification. E-RIHT is based on both packet marking and packet logging technique. In E-RIHT, marking and logging of packets is done on the border router and core router. RIHT make use of 32-bit marking field while E-RIHT uses 16-bit marking and logging field in IP packet which solves the packet fragmentation problem. Also RIHT uses router degree for the calculation of marking value where as E-RIHT make use of router id for the calculation of marking value. If attack packet is received by the victim then it automatically discard the packet and send the path reconstruction request to upstream router. For path reconstruction, identification field is used and path towards the attacker is constructed. 16-bit hash table is used for logging the 16-bit marking field. Information from the hash table can be retrieved very easily by using the marking field because of the index in the marking field. It is easy to search the particular information and trace back the path.

TABLE I
Notation Table

Rid	Router ID
UIi	Upstream Interface
P	Received Packets
H()	Hash Function
M	Size of Hash Table
c1, c2	Constant
HT	m entries in hash table.
HT[index]	Entry of the hash table with the address index. HT[index].mark: mark field. HT[index].UI:UI field.
%	The modulo operation.

Algorithm 1:

Packet marking and logging

Input: P, UI_i

Begin

1. $mark_{new} = P.mark \times (R_{id} + 1) + UI_i + 1$
2. *if* $mark_{new}$ *is overflow then*
3. $index = h = H(P.mark)$
4. $probe = 0$
5. *while not (HT[index] is empty or HT[index] is equal to* $(P.mark, UI_i)$ *)*
6. $probe++$
7. $index = (h + C1 \times probe + C2 \times probe2\ ) \% m$
8. *endwhile*
9. *if HT[index] is empty then*
10. *HT[index]. Mark = P.mark*
11. $HT[index].UI = UI_i$
12. *endif*
13. $mark_{new} = index \times (R_{id} + 1)$
14. *endif*
15. $P.mark = mark_{new}$
16. *Forward the packet to the next router*

End

In the following algorithm P and UI_i are taken as input in the beginning $mark_{new}$ is calculated using the formula P.mark × (R_{id} + 1) + UI_i +1. Here R_{id} is the router id which is basically the IP address of the router that will be converted from 32-bit to 8-bit . Conversion is done by breaking an IP address into four 8-bit segments say S1, S2, S3 and S4. OR function is applied to the segments and 8-bit resultant is generated.

Now incase $mark_{new}$ is overflow then index = h = H(P.mark) and probe = 0. After that checking HT[index] is empty or HT[index] is equal to (P.mark, UI_i) then set probe++ and index = (h + C1 × probe + C2 × probe2) % m. Now incase HT[index] is empty then set HT[index]. Mark = P.mark and HT[index].UI = UI_i. Then set $mark_{new}$ = index × (R_{id} + 1). At last set P.mark = $mark_{new}$ and Forward the packet to the next router.

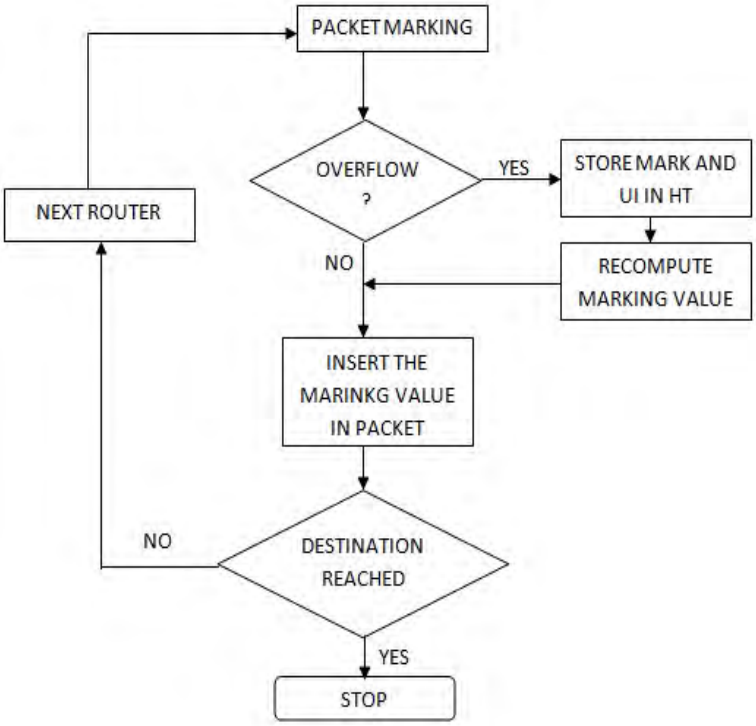

Fig. 1. Flowchart for packet marking and packet logging.

Algorithm 2:

Path Reconstruction

Begin

1. $UI_i = mark_{req} \%(R_{id} + 1) - 1$
2. *if* $UI_i = -1$ *then*
3. *index* = $mark_{req} / (R_{id} + 1)$
4. *If not index = 0 then*
5. $UI_i = HT[index].UI$
6. $mark_{old} = HT[index].mark$
7. *Send reconstruction request with* $mark_{old}$ *to upstream router by* UI_i
8. *else*
9. *This router is the nearest border router to the attacker*
10. *Endif*
11. *else*
12. $mark_{old} = mark_{req} / (R_{id} + 1)$
13. *Send reconstruction request with* $mark_{old}$ *to upstream router by* UI_i
14. *endif*

end

In the following algorithm first UI_i is calculated using the formula $mark_{req}$ %(R_{id} + 1) – 1. Now incase if UI_i = -1 then set index = $mark_{req}$ / (R_{id} + 1) and if not index = 0 then set UI_i = HT[index].UI and $mark_{old}$ = HT[index].mark. Send reconstruction request with $mark_{old}$ to upstream router by UI_i. This router is the nearest border router to the attacker. Otherwise set $mark_{old}$ = $mark_{req}$ / (R_{id} + 1) and send reconstruction request with $mark_{old}$ to upstream router by UI_i.

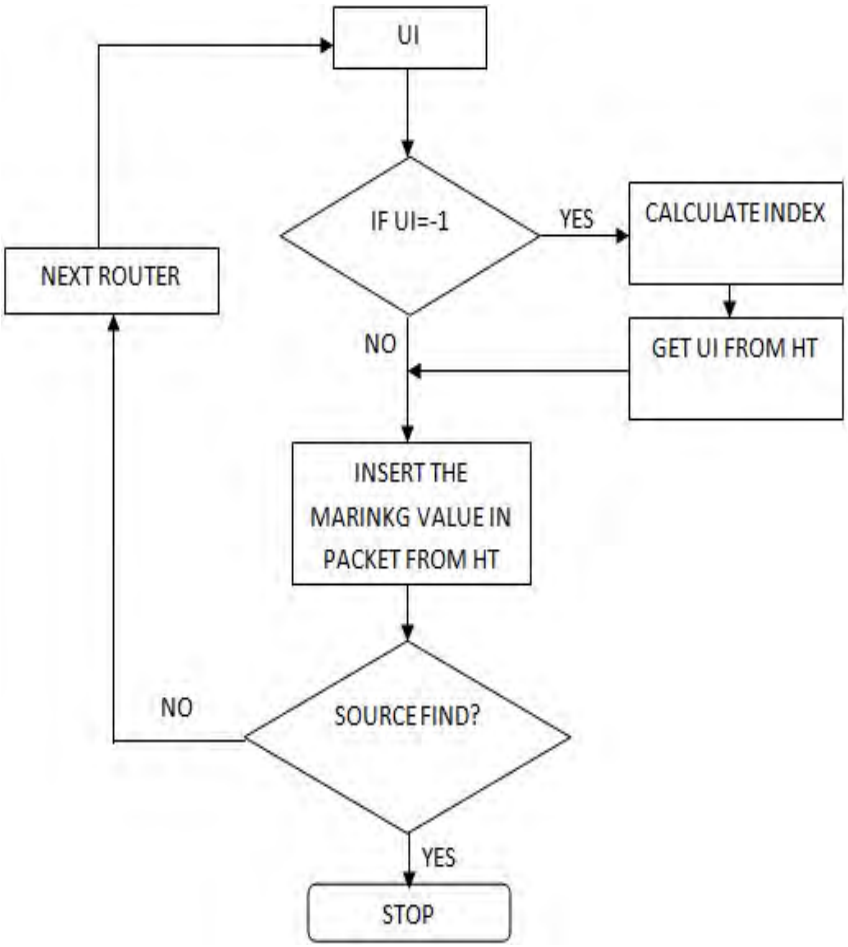

Fig 2. Flowchart for path reconstruction.

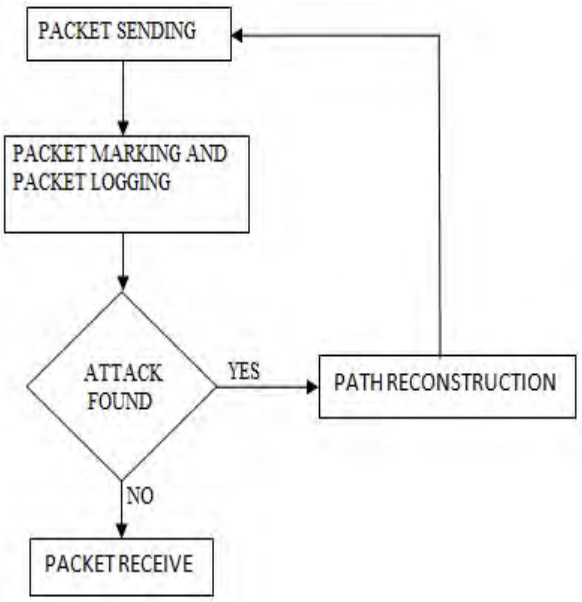

Fig. 3. Flowchart for packet marking, logging and path reconstruction.

E - RIHT can traceback single IP packet and multiple IP packets and reconstruct the path to attackers. It can efficiently traceback the DoS attacks. This technique has zero false positive and false negative results in path reconstruction.

VI. RESULT AND DISCUSSION

Table 2 shows the comparison between several IP Traceback schemes under several heads, they are packets required, marking overhead and storage overhead, traceback overhead, marking bits.

TABLE III
Comparision Between Several Hybird IP Traceback Schemes

IP Traceback schemes	Packets Required	Marking Overhead	Storage Overhead	Traceback Overhead	Marking Field
DLLT and PPPM	Multiple	High	High	Medium	DLLLT 34-bits, PPPM 57-bits
HIT	Single	Low	Low	Medium	16-bits
RIHT	Single	Low	Low	Low	32-bits
Enhanced RIHT	Single	Low	Low	Medium	16-bits

Packet Required:

Packet required means the number of packets needed for the IP traceback. DLLT and PPPM require more number of packets as compared to HIT. HIT require only single packet for the IP traceback.

Marking Overhead:

Marking overhead means number of bits required to mark the IP packets. DLLT requires 34-bits and PPPM requires 57-bits for packet marking. HIT needed only 16-bits each for packet marking.

Storage Overhead:

Mainly marking field value is to be stored at router that creates storage overhead. In DLLT, each router maintain individual MIT table using bloom filter where as in PPPM different routers share MIT table with the neighbor routers. HIT uses bloom filter. HIT make use of hast table to store packet digest and each router keeps digest table for all the neighboring routers.

Traceback Overhead:

Traceback overhead can be defined as the time consumed while accessing hash table for retrieving the marking value of various packets. HIT generates maximum traceback overhead while DLLT, PPPM and E - RIHT generates medium traceback overhead. RIHT gives least traceback overhead.

Marking Field:

Marking field comprises of number bits required to mark the packets passing through routers. Different schemes may use different number of marking bits such as DLLT uses 34-bits, PPPM uses 57-bits, RHIT uses 32-bits, HIT and E - RIHT uses 16-bits.

E-RIHT can trace single attack packet i.e. attack path can be traceback using only one attack packet. In this, excess of attack packets are not required. This scheme is quite similar to RIHT except that in this only 16-bit marking field is used. In RIHT, router degree is used for calculating marking value where as in E-RIHT router id is used for marking value calculation. The fault in RIHT is that, in case if router degree changes while traceback then attack path reconstruction gets failed and some false node is treated as attacker node. While E-RIHT is using router id which is the IP address of router that changes very rarely as compared to router degree.

V. CONCLUSION

In this paper, new scheme for IP traceback is proposed which is termed as E-RIHT. In E-RIHT, router id is used and 16-bit marking field is used to provide efficient IP traceback using single attack packet. E-RIHT make use of router id which is basically IP address of router, there are very less chances of having false traceback. 16-bit marking field is used, so packet fragmentation problem will be reduced. E-RIHT uses 16-bit marking field thus, marking and storage overhead is very less.

REFERENCES

[1] A. Hussain, J. Heidemann, and C. Papadopoulos, “A Frame-work for Classifying Denial of Service Attacks,” Proc. ACM SIGCOMM ’03,Aug. 2003
[2] A. C. Snoeren, C. Partridge, L. A. Sanchez, C. E. Jones, F. Tchak-ountio, B. Schwartz, S. T. Kent, and W. T. Strayer, “Single-packet IP traceback,” IEEE/ACM Trans. Networking, vol. 10, no. 6, pp. 721–734, Dec. 2002.
[3] A. C. Snoeren, C. Partridge, L. A. Sanchez, C. E. Jones, F. Tchakountio, S. T. Kent, and W. Strayer. “Single - Packet IP Traceback”. IEEE/ACMTransactions on Networking, Vol. 10, pp. 721–734, 2002.
[4] A. C. Snoeren, C. Partridge, L. A. Sanchez, C. E. Jones, F. Tchakountio, Stephen T. Kent, and W. Timothy Strayer. “Hash-based IP traceback,” Proceedings of ACM SIGCOMM 2001, pp.3–14, 2001.
[5] B.Al-Duwariand M.Govindarasu,“Novel hybrid schemes employing packet marking and logging for IP traceback,” IEEE Trans. Parallel Distributed Syst. , vol. 17, no. 5, pp. 403–418, May 2006.
[6] B. H. BLOOM, “Space/time trade-offs in hash coding with allowable errors,” Communications of ACM 13, pp. 422–426, 1970.”

[7] C. Gong and K. Sarac, “Toward a practical packet marking approach for IP traceback,” Int. J. Network Security , vol. 8, no. 3, pp. 271–281, Mar. 2009.
[8] C. Gong and K. Sarac, “A more practical approach for single-packet IP traceback using packet logging and marking,” IEEE Trans. Parallel Distributed Syst. , vol. 19, no. 10, pp. 1310–1324, Oct. 2008.
[9] C. Gong and K. Sarac, “IP traceback based on packet marking and logging,” Proc. of IEEE International Conference on Communications, Seoul, Korea, May 2005.
[10] Glenn Carl and George Kesidis, Richard R. Brooks and Suresh Rai.” Denial-of-Service Attack - Detection Techniques”, IEEE Internet Computing, pp: 82-89, January • February 2006
[11] John Haggerty, Qi Shi and Madjid Merabti. “Early Detection and Prevention of Denial-of-Service Attacks: A Novel Mechanism With Propagated Traced-Back Attack Blocking”, IEEE Journal on Selected Areas in Communications, Vol. 23, No. 10, pp: 1994-2002, October 2005.
[12] Ming-Hour Yang and Ming-Chien Yang “RIHT: A Novel Hybrid IP Traceback Scheme” IEEE Transactions On Information Forensics And Security, Vol. 7, No. 2, April 2012.
[13] Masayuki Okada , Yasuharu Katsuno , Akira Kanaoka , Eiji Okamoto “32-bit AS Number Based IP Traceback” Fifth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, 2011.
[14] Ning Lu, Yulong Wang, Fangchun Yang, Maotong Xu “A Novel Approach for Single-Packet IP Traceback Based on Routing Path” 20th Euromicro International Conference on Parallel, Distributed and Network-based Processing, 2012.
[15] S. Malliga and A. Tamilarasi, “A hybrid scheme using packet marking and logging for IP traceback” Int. J. Internet Protocol Technol., vol. 5, no. 1/2, pp. 81–91, Apr. 2010.
[16] Shweta Vincent, J. Immanuel John Raja “A Survey of IP Traceback Mechanisms to overcome Denial-of-Service Attacks” RECENT ADVANCES in NETWORKING, VLSI and SIGNAL PROCESSING, pp.93-98, Feb.2010.
[17] S. Savage, D. Wetherall, A. Karlin, and T. Anderson, “Practical network support for IP traceback,” in Proc. ACM SIGCOMM2000,Stock- holm, Sweden, pp. 295–306 Aug. 2000.
[18] S.M.Bellovin,M.D.Leech,and T. Taylor, “ICMP traceback messages,” Internet Draft: Draft-Ietf-Itrace-04.Txt ,Feb.2003.
[19] X. J. Wang and Y. L. Xiao. “IP traceback based on deterministic packet marking and logging” in Proc. SCALCOM-EMBEDDEDCOM’09, Dalian, China, pp. 178–182, Sep. 2009.
[20] Yan Fen,,Zhu Hui , Chen Shuang-shuang , Yin Xin-chun “A Lightweight IP Traceback Scheme Depending on TTL” International Workshop on Information and Electronics Engineering (IWIEE), 2012.