MALICIOUS IP TRACKBACK: REVEALING THE LOCATIONS OF IP SPOOFERS

Mohammed Aquib. A. Hulkoti[1], M S Patel[2]

[1]*Student, Department of CSE, T. JIT, Bangalore, Karnataka, India*
[2]*Associate Professor, Department of CSE, T. JIT, Bangalore, Karnataka, India*

Abstract

It has been for long period that, counterfeit source IP location has been utilized to disguise the genuine area of attackers. Various unique IP trackback techniques have been utilized to trap the IP spoofers. Notwithstanding, because the need of physical devices implementation, there has been definitely not a generally embraced IP trackback outcome in any event at the Internet level. Therefore, yet there has been no disappearance of fog on the areas of spoofers. Malicious IP trackback (MIT) is being proposed to sidestep the implementation difficulties of existing techniques. MIT researches ICMP error messages (tagged as "Path Backscatter") activated by spoofed packets, and spoofers are trapped taking into account open accessible data (e.g., topology). Along this, MIT without any prerequisite implementation of any special device can track attackers. This article represents the reasons, gathering, and the simulation results for "Path Backscatter", exhibits the techniques and viability of MIT, and with the implementation of MIT on the data set of path backscatter. These outcomes can encourage uncover the attackers that uses counterfeit addresses, these attackers has been observed very long, but difficulties in implementation lead to problem in understanding it thoroughly. In spite of all, MIT cannot be guaranteed in all the attacks, it might be the most valuable system to track attackers before a network level trackback system is conveyed in real-time. Simulation and evaluation results in MIT can be used to get more effective algorithms and mechanisms to defend against these attacks. An IP trackback technique is important to overcome the problem of distributed denial of service attack.

Keywords: *Internet protocol (IP) spoofers, path backscatter, Internet Security and Internet control message protocol (ICMP).*

----------***----------

1. INTRODUCTION

Attackers propelling attacks with counterfeit IP address in an IP spoofing attack, which have been a prolonged security issue in Internet. To hide their real location, attackers use unused address or address yet to be assigned. Many of the denial of service attacks depend on "IP spoofing" attack. Dos attacks which depend on IP spoofing includes SYN flooding, Smurf, DNS amplification etc [12]. To seize the source of malicious traffic is of incredible significance. IP spoofers cannot be blocked, till their original locations are not disclosed. Indeed, even simply drawing closer the spoofers, for instance, deciding the autonomous systems or network they hide themselves, attackers can be detected in a smaller region, and filters can be implemented nearer to the attackers before normal traffic and affected traffic get combined. At last, recognizing the sources of spoofing traffic can develop a notoriety framework for autonomous systems, where the service providers can be forced to verify the source of that address.

In any case, it is prickly to track the source of a malicious traffic at the Internet level. There are two major challenges, in constructing an IP trackback technique. As the first challenge would be the expense in supporting a trackback technique in an routing environment. Presently operating normal routers do not support existing trackback techniques and extensively overload the routers, especially in case of networks with heavy load of traffic. Second challenge is, getting all service providers together in one roof. Since the attackers would be at each corner in the world, but forcing each of internet service provider at that corner to implement their own trackback technique would be pointless. IP trackback implementation, no doubt would be advantage to track the attacker, but would create significant load on the network.

Because of deployment difficulties in existing trackback techniques, this paper attempt to reveal the area of attackers using the routes created by routers with basic supported functionalities, when the attack takes place.

Rather than proposing another IP trackback technique with enhanced tracing ability, this proposes a novel result, named Malicious IP Trackback (MIT), to sidestep the difficulties in implementation. There would be many reasons for a router to avoid the transmission of malicious traffic, such as TTL limit. Routers in such situation, triggers an ICMP message (as path backscatter) at the spoofed address of a node. Since spoofers are settled near routers to conduct the attacks, hence path backscatter could conceivably uncover the areas of attackers.

2. RELATED WORK

Although all the existing IP trackback techniques perform IP trackback, but the introduced MIT technique is different from all the existing techniques. MIT is enlivened by various IP spoofing perception practices. As existing techniques are proposed to reveal the source of malicious traffic or to trace the route of attackers. Existing techniques, proposes five primary classifications: ICMP trackback, packet marking, link testing, logs on the router, overlay system and hybrid tracing.

A huge amount of ICMP messages are created to an authority or the end node of the route in the ICMP trackback technique [10]. The route, on which attack has been conducted, can be rebuilt by the utilization of ICMP messages. If routers are equipped with "iTrace", then destination would get the test results of ICMP messages with specific possibility. Extensive extra packets will be created to expend the effectively focused resource, transmission capacity as the major disadvantage in ICMP trackback technique. This extra amount of packets will make the attack more harmful, in the case of attacker willing to block the resources such as bandwidth. Thus, the extra traffic created due to ICMP production will decrease the performance of a processor in a router.

The technique of packet marking requires each successor router alters the details of predecessor router in the packet header and decides further transmission of a packet. Route of a packet can be then rebuilt by the destination node, through the packets transmitted by the sender node. Packet marking techniques are for the most part thought to be lightweight since they don't cost transmission capacity on routers and the transmission capacity of link. Not each of the routers in a network supports packet marking technique

This technique is used to identify the source of attacking traffic by testing the links between routers in a network. This technique tries to determine the link which carries the malicious traffic, by testing the incoming routes from a router to victim. This technique is operated on the predecessor nodes again and again, until attacking node is caught. The major drawback of this technique is, it works until the attack is in progress.

Log on the router can reproduce the attacking path from the router, as the router keeps a record whenever a packet is sent [7]. Storage capacity on the router can be utilized by using bloom filter, which reduces the quantity of bits to store a packet. In spite of all, to accomplish a sufficiently low collision possibility in current rapid networks, the capacity expense is still too vast for commercial routers.

CenterTrack [3] proposes utilization of extraordinary routers with special tracing functionality, so whenever there is malicious traffic other than normal traffic, it is shifted to this special router from normal operating router using an overlay system [12]. This system would explicitly add an overhead to the administration of internet and tunnels, as the fact such a technique can diminish the necessity of normal functioning routers. It has been observed that, attackers can be precisely detected in autonomous system, if overlay system is able to join some of the autonomous systems in its system. So, the main purpose in this technique is making autonomous system's participate. Hence, this technique requires normal router to be upgraded to trace the spoofers.

Although there are many advantageous existing techniques, but all these techniques are far away to be implemented and seems like impossible to deploy on the level of wide area network. At present, tracing the spoofers still seems like not possible or difficult.

3. STUDY OF PROPOSED WORK

Not each one of the packets is successful to accomplish their target. Packet drop in a network may occur, because of several reasons and thus may generate an ICMP message at node by the router. In an event, that source location is counterfeited, the real owner of the spoofed address would get the message. As the above scenario illustrates a reflection attack and locations of the nodes utilized by attackers. Figure 1 illustrates the above situation.

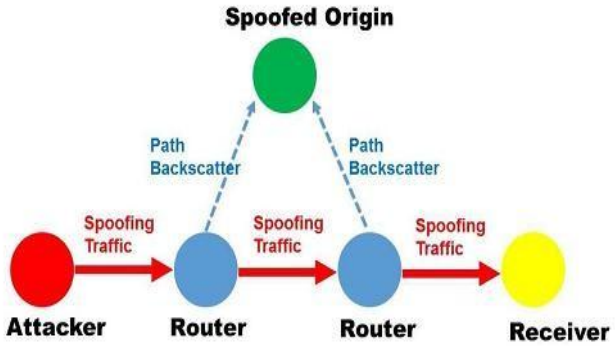


Fig -1: Generation and collection of path backscatter messages in an attack.

Figure 2. illustrates the fields in path backscatter message format. Each ICMP message includes the original source node address and that packet's IP header. In this way, each ICMP path backscatter message gives two types of addresses, where first address includes the victim of reflection attack's address residing in between the route of spoofer and the destination of a malicious traffic, and second address includes the destination node's address. Whereas header of an IP packet includes many advantageous details, such as TTL time consumed by a malicious packet. As it has been observed that, because of some nodes in a network alter the address (e.g., Network Address Translation), but the original addresses assigned to sender and receiver machine will be different.

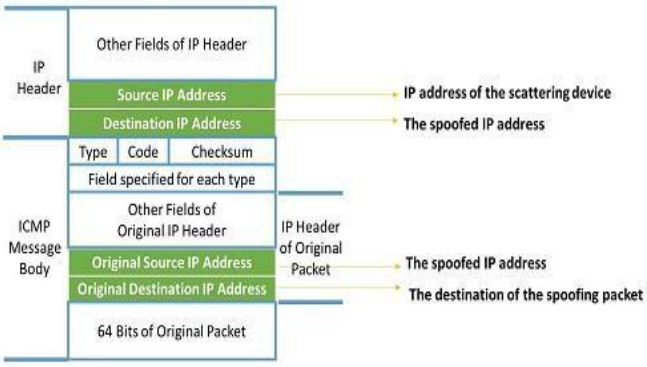


Fig -2: Path backscatter message format.

3.1 Classification and Reasons for Path Backscatter

This paper tries to illustrate the reason for the classification of ICMP "Path Backscatter" messages as in Table I. Especially, this paper try to make out causes that, messages are generated near the attacking nodes. It ought to be noticed that, path backscatter messages are triggered in a large size close to victim node. Be that as it may, considering a large amount of malicious messages, if just a little proportion of these messages trigger path backscatter messages close to the attacker, it would be great advantageous from the aggregated dataset of path backscatter message. Although, these ICMP messages are not created near by the attackers, areas in which they are generated are nearer to the spoofers than the victims. Therefore, they can be utilized as a part of the first pace for trackback.

Table -1: Classification of Path Backscatter classes

Type	Class
Time Exceeded	TIMXCEED_INTRANS
Destination Unreachable	UNREACH_FILTER_PROHIB, UNREACH_NET_PROHIB, UNREACH_HOST_PROHIB, UNREACH_HOST, UNREACH_NET, UNREACH_NEEDFRAG
Source Quench	SOURCEQUENCH
Redirect	REDIRECT_HOST, REDIRECT_NET
Parameter Problem	PARAMPROB

Data units pursuing "0" TTL value triggers "TIMXCEED_TRANS" message. ICMP messages recognize all the above messages in Table I , as included in path backscatter. In spite of fact that spoofers can enable the underlying value of TTL to be sufficiently substantial to abstain from activating all these messages, they may purposefully transmit the packets starting with smaller values of TTL, as routers on the way of spoofers triggers the messages of "TTL Exceeding" to devour the router resources.

In most cases, filtering techniques implemented in the area in between the attacker and victim, triggers "UNREACH_FILTER_PROHIB","UNREACH_NET_PRO -HIB" and "UNREACH_HOST_PROHIB" messages. If the packet is unable to reach a specified destination, then "UNREACH_HOST" and "UNREACH_NET" messages get triggered. If the attacker, propel attack against a private address or address yet to be allocated, the above messages are generated. Whenever a "Don't Fragment" flag is on and length of the spoofing traffic is more than maximum transmission unit (MTU) on the path of a node, it would trigger "UNREACH_NEEDFRAG" message. So, when the routers are under attack these messages get triggered.

When the buffer size of a router is exhausted and has no space for the first packet, then "SOURCEQUENCH" message is triggered. It can be come about because of inability of a router to send such a huge amount of malicious traffic.

If the attacking source pursues more than two gateways, then "REDIRECT_HOST" and "REDIRECT_NET" messages get triggered. Else in the first packet, if the router comes out with an issue in the header parameter then "PARAMPROB" message is triggered.

3.2 Problems in Security with Path Backscatter Messages

Practically, it can be observed that path backscatter messages are as simple to falsify a malicious data packet creation. Accordingly, the gatherer should deploy a filter to remove all the false ICMP messages from the normal messages.

Attackers use all conceivable TTL values for sending counterfeit ICMP messages to the victim, yet the victim has an option to overcome this problem, by verifying whether those messages are with different values of TTL from a node or such can be distinguished by node calculation.

4. PROPOSED WORK

This paper introduces IP trackback mechanism in light of exploring ICMP messages through Malicious IP Trackback (MIT). MIT is developed by observing all the existing techniques. The fundamental technique, which depends on topology and routing knowledge. Notwithstanding, by and large the routing data is difficult to accomplish. The technique work on the off chance that the routing details is obscure. Attackers can be traced without availability of routing and topological details, in some exceptionally extraordinary circumstances

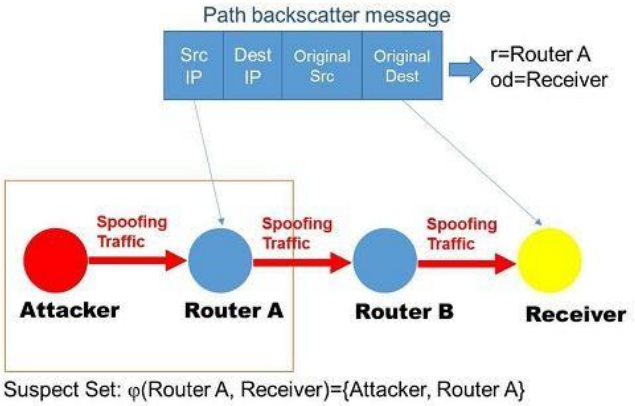


Fig -3: The doubtful set identified by a path backscatter.

This paper utilizes the details of a path to trace the region in which the attacker resides. To illustrate the list of nodes on any one of the route from p to q can be shown by path(p, q) and to indicate all the routes from p to q can be shown by PATH(p, q). List of all nodes from which a packet can detour a reflector to reach it's original destination can be denoted by ϕ(r, od). So the main intenition behind MIT is, when an ICMP message is trapped whose sender is the router as a reflector (r) and real destination (od), the packet should detour a reflecto while coming from attacker and moving towards to real destination.

$$\varphi(r, od) = \{v | r \in path(v, od), path(v, od) \in PATH(v, od)\}.$$

The smallest set in which attacker resides is decided by ϕ(r, od). The set of malicious nodes gives the outcome set ϕ(r, od). As outlined in figure 3, on the off chance that every path is free of loop, {Attacker, Router A} is the malicious node from the set determined by the Path Backscatter. On the off chance that the paths and physical arrangement of the network details are available, this technique can be utilized to viably decide the malicious node set.

4.1 Availability of Routing Details

Topological details of a network can be obtained in some trackback situations. For instance, trace route can be used to get details of topology at router-level, and topological details at the Autonomous System-level can be obtained from the border gateway routing protocol. In addition, various ASes make open to all their topologies [13]. Be that as it may, the courses of a system are constantly regarded as business confidential and are not open to everyone.

Based on routing, two assumptions can be proposed:

1. Loop-Free Assumption [12]: There should be no loop in the path of a packet, that"s the condition of this assumption. The assumption would get satisfied unless there is no convergence of malfunctioning or the routing.

```
1: function GETSUSPECTSET_LOOPFREE(G,r,od)
2:     SuspectSet ← ∅
3:     c ← null
4:     P ←shortest path from r to od
5:     for Vertex v in P do
6:         if v == r then
7:             Continue
8:         end if
9:         G' ← G.remove(v)
10:        if r and od are disconnected in G' then
11:            c ← v
12:            break
13:        end if
14:    end for
15:    SG ← G.remove(c)
16:    for Vertex v in SG do
17:        if v and r are connected in SG then
18:            SuspectSet ← SuspectSet + v
19:        end if
20:    end for
21:    return SuspectSet
22: end function
```

Fig -4: The algorithm depicts the loop-free assumption to identify the malicious nodes in a set.

To discover all the gratifying verticals through collection of all the set of nodes is almost impractical on for wide area networks. Figure 4 illustrates an algorithm based on the above assumption. A shortest path from destination to source is identified from the above algorithm. The algorithm checks whether there is a path break between receiver and original destination, if the link between them is removed, that is along the second the link. At whatever point such a link to a node c is found, expelling that route from gateway (G), and the set containing every one of the links which are still associated with receiver is only the malicious node set.

2. Valley-Free Assumption [12]: There should be no valley in the path of a packet at the autonomous system level, to hold the assumption well. In spite of the fact that the high intricacy of AS relationship has decreased the comprehensiveness of this assumption Due to the fact of decreased coordination between autonomous systems is highly complicated, it is still the most widely recognized prototype at autonomous system level routing.

The above assumption cannot be used in different topologies other than AS-level topology. Considering the size of Internet topology at AS-level, for a Path Backscatter message, it is exorbitant to discover every one of the autonomous systems.

```
1: function GETSUSPECTSET_VALLEYFREE(G,r,od)
2:     if od ∈ Cone(r) then
3:         return G.nodes()
4:     else
5:         return Cone(r)
6:     end if
7: end function
```

Fig -5: Algorithm depicts Valley-Free assumption, to identify malicious nodes in a set.

4.2 Unavailability of Routing and Topological Details

This section overcomes the limitation of pursuing details of topology of a network and routing. This is possible, using three special types of ICMP messages:

1. The "Path Backscatter" messages whose real node distance check is 1 or 0. Such messages are triggered 1 or 2 nodes from the attackers. These messages are generated possibly near the gateway of the attackers.
2. ICMP messages which are triggered near the gateway of the attacker are of type "Redirect".
3. The ICMP messages triggered pursuing private address or address yet to be allocated. These messages are commonly triggered in between the attacker and the actual receiver node by the default free zone router

5. OUTCOMES OF PROPOSED WORK

MIT is altogether different from any current trackback technique. There is no specific possibility of path backscatter message generation, which is the primary

contrast. Despite the fact that the era of path backscatter cannot be avoided, the aggregate amount of path backscatter is very outstanding. Table 1 outlines the datasets utilized as a part of this assessment. The messages that were gathered by CAIDA around 37 days in 2008, which also includes path backscatter messages from freshest. Despite the fact that there are colossal measures of path backscatter messages created, their era does not have a specific likelihood. Accordingly, it is difficult to assess MIT comparable as the existing IP trackback techniques.

Table -2: Datasets.

Dataset	Source
Path backscatter dataset	CAIDA 2008 Backscatter dataset
AS Level Internet topology	RouteView BGP data
AS relationship	Inferred AS relationship from CAIDA
IP-to-AS Mapping	RouteView BGP data
AS topologies	Topology zoo [34]
IP geolocation	IPInfoDB [37]

Consequently, it cannot be guaranteed how good MIT will work in each attack. To avoid the unsure characteristics of path backscatter message era, after getting the path backscatter (r, od) tuple we assess the likelihood of finding the attacker. To accomplish this, this paper carryout some assumptions based on attacks irrespective of IP spoofing and ICMP messages i.e. path backscatter.

1. Attackers utilize random areas for attacking.
2. Attackers pick random destinations for IP spoofing attack.
3. Path backscatter messages triggered in between the attacker and original destination.

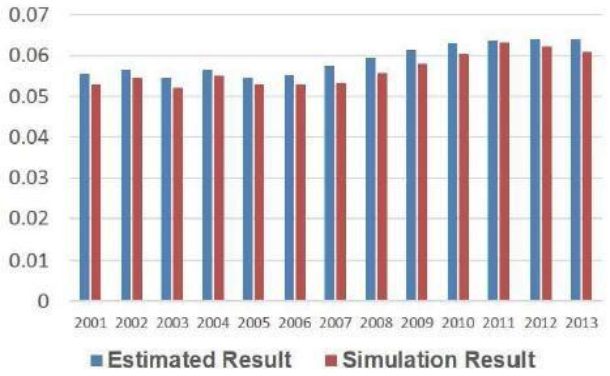


Fig -5: The evaluated possibility of exact detection using topology at Autonomous system level topology taking into account the loop free assumption with the result of simulation.

Without inclusion of intermittent characteristics of path backscatter message era, the adequacies of the techniques are really identified by the network structure. In spite of the fact that exceptionally constrained data can be utilized as a part in tracing the spoofers, the attacker tracing is observed to be successful generally because of the structure of power law in networks.

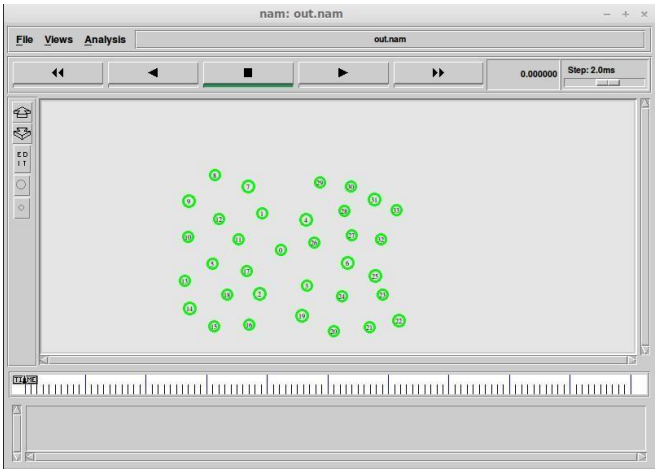


Fig -6: Deployment of nodes in a network

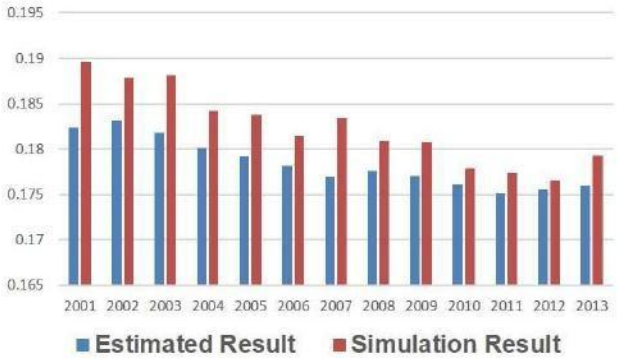


Fig -7: The evaluated possibility of exact position in topology at autonomous level taking into account the valley-free assumption with the result of simulation.

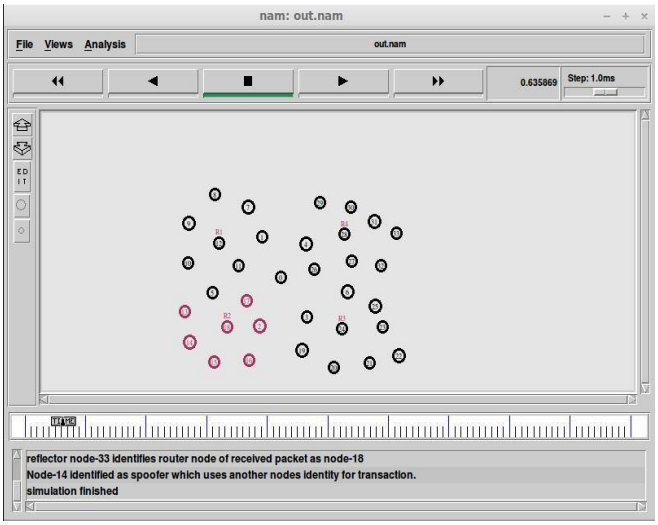


Fig -8: Malicious nodes set in a network, where spoofer resides, with the implementation of MIT on the network.

6. CONCLUSION

In this article we have exhibited a new mechanism based on, "Path Backscatter," for assessing IP spoofing attack in the network. Utilizing this strategy, we have watched across the board IP spoofing attacks in the network, disseminated among a wide range of areas and service providers. The

quantity of the attacks we evaluated are of large amount, with a very few amount of enormous attacks becoming a critical portion of the general attack. In addition, this paper shows a extraordinary amount of attacks we see a shocking number of attacks coordinated at a couple of outside nations, at own devices, and towards specific network operations. This paper attempt to vanish the fog on the areas of spoofers in view of exploring the path backscatter messages. This paper proposes Malicious IP Trackback (MIT) which tracks attackers in view of Path Backscatter messages with open accessible data.

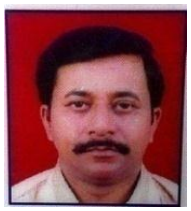
This paper indicates implementation of MIT when there is no information of routing and topology, else the routing is obscure, or none of this information is available. We displayed two successful assumptions to apply MIT in wide area networks and proved they are flawless. This article demonstrates that, the viability of MIT in light of deduction and simulation. It also demonstrates the implementation of MIT on the dataset of path backscatter, where spoofers are trapped.

REFERENCES

[1]. S. M. Bellovin, "Security problems in the TCP/IP protocol suite, "ACM SIGCOMM Comput. Commun. Rev., vol. 19, no. 2, pp. 32–48, Apr. 1989.
[2]. R. Stone, "CenterTrack: An IP overlay network for tracking DoS floods, "in Proc. 9th USENIX Secur. Symp., vol. 9. 2000, pp. 199–212.
[3]. S. Savage, D. Wetherall, A. Karlin, and T. Anderson, "Practical network support for IP trackback," in Proc. Conf. Appl., Technol., Archit., Protocols Comput. Commun. (SIGCOMM), 2000, pp. 295–306.
[4]. L. Gao, "On inferring autonomous system relationships in the internet,"IEEE/ACM Trans. Netw., vol. 9, no. 6, pp. 733–745, Dec. 2001.
[5]. A. C. Snoeren et al., "Hash-based IP trackback," SIGCOMM Comput. Commun. Rev., vol. 31, no. 4, pp. 3–14, Aug. 2001.
[6]. A. Mankin, D. Massey, C.-L. Wu, S. F. Wu, and L. Zhang, "On design and evaluation of „intention-driven" ICMP trackback," in Proc. 10th Int. Conf. Comput. Commun. Netw., Oct. 2001, pp. 159–165.
[7]. H. C. J. Lee, V. L. L. Thing, Y. Xu, and M. Ma, "ICMP trackback with cumulative path, an efficient solution for IP trackback," in Information and Communications Security. Berlin, Germany: Springer-Verlag, 2003, pp. 124–135.
[8]. S. Bellovin. ICMP Trackback Messages. [Online]. Available:http://tools.ietf.org/html/draft-ietf-itrace-04, accessed Feb. 2003.
[9]. D. Moore, C. Shannon, D. J. Brown, G. M. Voelker, and S. Savage, "Inferring internet denial-of-service activity," ACM Trans. Comput. Syst., vol. 24, no. 2, pp. 115–139, May 2006.
[10]. J. Liu, Z.-J. Lee, and Y.-C. Chung, "Dynamic probabilistic packet marking for efficient IP trackback," Comput. Netw., vol. 51, no. 3, pp. 866–882, 2007.
[11].R. P. Laufer et al., "Towards stateless single-packet IP trackback," in Proc. 32nd IEEE Conf. Local Comput. Netw. (LCN), Oct. 2007, pp. 548–555. [Online]. Available: http://dx.doi.org/10.1109/ LCN.2007.160
[12]. Guang Yao, Jun Bi, Senior Member, IEEE, and Athanasios V. Vasilakos, Senior Member, IEEE, "Passive IP Traceback: Disclosing the Locations of IP Spoofers From Path Backscatter", vol. 10, no. 3, march 2015,pp 471.
[13]. Y. Xiang, W. Zhou, and M. Guo, "Flexible deterministic packet marking: An IP traceback system to find the real source of attacks," IEEE Trans. Parallel Distrib. Syst., vol. 20, no. 4, pp. 567–580, Apr. 2009.

BIOGRAPHIES

Mr. Mohammed Aquib. A. Hulkoti. Completed B. E computer science in SIET College Bijapur, Under VTU Belgaum. Presently pursuing M. tech Computer Networking in TJIT college Bangalore.



Mr. M S Patel, Currently working as, Associate Professor in T. JIT, Bangalore. Pursuing more than 13 years of Teaching and 2 years of industrial experience. Interested Areas: Computer Networks and Cloud Computing.