Enhanced probabilistic packet marking traceback mechanism

Tan, Wei Peng

2006

Tan, W. P. (2006). Enhanced probabilistic packet marking traceback mechanism. Master's thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/2464>

<https://doi.org/10.32657/10356/2464>

Nanyang Technological University

Downloaded on 26 Apr 2023 19:22:22 SGT

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**ENHANCED PROBABILISTIC PACKET MARKING
TRACEBACK MECHANISM**

**TAN WEI PENG
SCHOOL OF COMPUTER ENGINEERING
2006**

Enhanced Probabilistic Packet Marking Traceback Mechanism

Tan Wei Peng

School Of Computer Engineering

A thesis submitted to the Nanyang Technological University
in fulfilment of the requirement for the degree of
Master of Engineering

2006

CONTENTS

	PAGE
CONTENTS	I
LIST OF FIGURES	VI
LIST OF TABLES	IX
LIST OF ABBREVIATIONS	X
ABSTRACT	XI
ACKNOWLEDGEMENTS	XIII
<u>CHAPTER ONE</u>	<u>1</u>
OVERVIEW OF DENIAL OF SERVICE AND DISTRIBUTED DENIAL OF SERVICE ATTACKS	1
1.1 Background	1
1.2 Denial of Service (DoS)	3
1.2.1 TCP SYN Flood Attack	4
1.2.2 UDP Flood Attack	6
1.2.3 Smurf Attack	6
1.2.4 Teardrop Attack	7
1.2.5 Bonk Attack	7
1.2.6 Land Attack	7
1.2.7 Ping of Death Attack	8
1.3 Distributed Denial of Service (DDoS)	8
1.3.1 DDoS Attack Architecture	10

1.3.2	Tribble Flood Network (TFN)	12
1.3.3	Trinoo	13
1.3.4	Stacheldraht	14
1.3.5	TFN2k	14
1.4	High Bandwidth DDoS Flooding Attacks	15
1.4.1	Direct Attack	15
1.4.2	Reflector Attack	16
1.5	Research Objective	18
1.6	Report Organization	18
 CHAPTER TWO		19
OVERVIEW OF DDOS DEFENSE MECHANISMS		19
2.1	Attack Detection and Response Mechanisms	19
2.1.1	Source-End Network Defense Mechanism	22
2.1.2	Intermediate Network Defense Mechanism	23
2.1.3	Victim-End Network Defense Mechanism	25
2.1.4	A Comparison of Various Attack Detection And Response Mechanisms	26
2.2	IP Traceback	27
2.2.1	Controlled Flooding	30
2.2.2	ICMP Traceback	31
2.2.3	Hash-Based IP Traceback	33
2.2.4	Overlay Network	36
2.2.5	Probabilistic Packet Marking (PPM)	37

2.2.6	Adjusted Probabilistic Packet Marking	39
2.2.7	Dynamic Probabilistic Packet Marking	40
2.2.8	Comparison of the Traceback Schemes	41
 CHAPTER THREE		42
ENTROPY-MINIMIZATION CLUSTERING TECHNIQUE FOR PROBABILISTIC PACKET MARKING SCHEME		42
3.1	Overview of Probabilistic Packet Marking (PPM) Scheme	42
3.1.1	High Combination Overhead	47
3.1.2	Large Number of False Positives	48
3.2	PPM with Entropy-Minimization Clustering Technique	49
3.3	Simulation Studies	55
3.3.1	Simulation Model	55
3.3.2	Network Scenario 1	56
3.3.3	Network Scenario 2	57
3.4	Results and Performance Analysis	58
3.4.1	PMF Analysis for Network Scenario 1 using PPM	59
3.4.2	PMF Analysis for Network Scenario 1 using APPM	60
3.4.3	Combination Overhead Analysis for Network Scenario1	62
3.4.4	PMF Analysis for Network Scenario 2 using PPM	63
3.4.5	PMF Analysis for Network Scenario 2 using APPM	64
3.4.6	Combination Overhead Analysis for Network Scenario2	65
3.5	Summary of Results	67

CHAPTER FOUR	68
LEGITIMACY INVESTIGATION AND INTENTION-BASED PROBABILISTIC PACKET MARKING SCHEME	69
4.1 Effectiveness of Probabilistic Packet Marking (PPM) and its Variant Schemes	69
4.2 Legitimacy Testing	71
4.3 Intention-Driven iTrace	74
4.4 Legitimacy Investigation and Intention-Based Probabilistic Packet Marking Scheme (LI-IPPM)	74
4.4.1 Proposed Traceback Architecture	77
4.4.2 Architecture of LI-IPPM	80
4.4.3 Intention-Based Probabilistic Packet Marking Module	81
4.4.4 Investigation Agent Module	85
4.4.5 Query Processing Agent and Authentication Test Module	88
4.4.6 Attack Graph Generation Module	89
4.5 Analysis	90
4.6 Analysis and Simulations Studies	94
4.6.1 Combination Overhead Analysis	95
4.6.2 Path Reconstruction Time Analysis	96
4.7 Summary of Results	98
CHAPTER FIVE	100
CONCLUSION	100
5.1 Future Work	101

Nanyang Technological University

LIST OF PUBLICATIONS	102
LIST OF REFERENCES	103

LIST OF FIGURES

	PAGE
Figure 1-1: Number of attacks reported to the CERT/CC	1
Figure 1-2: DDoS attacks on the Rise	2
Figure 1-3: Total Losses caused by Different Security Incidents	3
Figure 1-4: Denial of Service Attack	4
Figure 1-5: Standard TCP Three-Way Handshake	4
Figure 1-6: TCP SYN Flood Attack	5
Figure 1-7: Smurf Attack	6
Figure 1-8: Distributed Denial of Service Attack	8
Figure 1-9: Attack Sophistication vs. Required Intruder Knowledge	9
Figure 1-10: Three-Tier DDoS Attack Architecture	11
Figure 1-11: Tribble Flood Network (TFN)	13
Figure 1-12: Trinoo	14
Figure 1-13: Direct Attack	16
Figure 1-14: Reflector Attack	17
Figure 2-1: Deployment Location of Attack Detection and Response Mechanisms	20
Figure 2-2: Effectiveness of Attack Detection and Response Mechanisms	21
Figure 2-3: Ingress Filter	22
Figure 2-4: Egress Filter	23
Figure 2-5: Illustration of Route-Based Packet Filtering (RPF)	24
Figure 2-6: Local Attack Detection Mechanism	25
Figure 2-7: Controlled flooding Scheme	30
Figure 2-8: ICMP Traceback Scheme	32

Figure 2-9: Fields shaded are masked out before digesting	33
Figure 2-10: Hash-Based IP Traceback Scheme	34
Figure 2-11: CenterTrack Scheme	36
Figure 2-12: PPM Traceback Scheme	38
Figure 3-1: Probabilistic Packet Marking (PPM) Scheme	43
Figure 3-2: Encoding Edge-ID Fragments in IP Identification Field	43
Figure 3-3: Encoding Scheme	44
Figure 3-4: Encoding of edge-id fragments	45
Figure 3-5: Path Reconstruction	45
Figure 3-6: Address Verification Procedure	46
Figure 3-7: Packets leave the Bottleneck Router Equally Spaced	49
Figure 3-8: Simple Network Illustration	50
Figure 3-9: Inter-Packet Spacing of Different Cluster	51
Figure 3-10: PMF of the Inter-Packet Spacing	52
Figure 3-11: Network Scenario 1	57
Figure 3-12: Network Scenario 2	58
Figure 3-13: PMF of the inter-packet spacing for the correct cluster using PPM Scheme	59
Figure 3-14: PMF of the inter-packet spacing for the correct cluster using APPM Scheme	61
Figure 3-15: Combination overhead using PPM/APPM	62
Figure 3-16: Combination overhead using PPM/APPM with Entropy-Minimization Clustering Technique	62
Figure 3-17: PMF of the inter-packet spacing for the correct cluster using PPM	

Scheme	64
Figure 3-18: PMF of the inter-packet spacing for the correct cluster using APPM	
Scheme	65
Figure 3-19: Combination overhead using PPM/APPM	65
Figure 3-20: Combination overhead using PPM/APPM with Entropy-Minimization	
Clustering technique	66
Figure 4-1: Spoofing Additional Traceback Paths	69
Figure 4-2: SYN Cookies Legitimacy Testing	72
Figure 4-3: ICMP Echo Message Legitimacy Testing	73
Figure 4-4: Proposed Traceback Architecture	79
Figure 4-5: Intention-Based Probabilistic Packet Marking Module	82
Figure 4-6: Set of Fields Used For Marking	84
Figure 4-7: Architecture of Legitimacy Investigation Module	87
Figure 4-8: Computational Overhead	93
Figure 4-9: Tree Topology	94
Figure 4-10: Combination Overhead	95
Figure 4-11: Number of packets required for reconstruction	97

LIST OF TABLES

	PAGE
Table 1: A Summary of Reflector Attack Methods	17
Table 2: A Comparison of Various Attack Detection And Response Mechanisms	27
Table 3: Comparison of the Traceback Schemes	41

LIST OF ABBREVIATIONS

ABBREVIATIONS	FULL EXPRESSIONS
AGGM	Attack Graph Generation Module
APPM	Adjusted Probabilistic Packet Marking
CERT/CC	Computer Emergency Response Team Coordination Center
DDoS	Distributed Denial of Service
DGA	Data Generation Agent
DoS	Denial Of Service
DPPM	Dynamic Probabilistic Packet Marking
FNR	False Negative Ratio
FPR	False Positive Ratio
HMAC	Hash Message Authentication Code
ICMP	Internet Control Message Protocol
ID-Tag	Identity Tag
IDMEF	Intrusion Detection Message Exchange Format
IETF	Internet Engineering Task Force
IP	Internet Protocol
IPPM	Intention-Based Probabilistic Packet Marking Module
IPv6	Internet Protocol Version 6
ISP	Internet Service Provider
ITUA	Intrusion Tolerance by Unpredictability and Adaptation
LAD	Local Attack Detection
LI-IPPM	Legitimacy Investigation and Intention-Based Probabilistic Packet Marking Scheme
NPSR	Normal Packet Survival Ratio
NSA	National Security Agency
PMF	Probability Mass Function
PPM	Probabilistic Packet Marking
QoS	Quality of Service
QPAM	Query Processing Agent Module
RPF	Route-Based Packet filtering
SCAR	SPIE Collection and Reduction Agent
STM	SPIE Traceback Manager
TCP	Transmission Control Protocol
TFN	Tribble Flood Network
TR	Tracking Router
TTL	Time-To-Live
UDP	User Datagram Protocol

Abstract

ABSTRACT

Distributed Denial of Service (DDoS) attacks on the Internet have become a pressing issue following several high-profile attacks on e-commerce enterprises like Yahoo, Amazon and Ebay [2] [3]. New occurrences of DDoS attack continue to be reported and threats have continued to increase. A recent DDoS attack was aimed at 13 root servers that offer the primary roadmap for approximately all Internet infrastructures [60]. It was the largest and most complex DDoS attack ever and only four or five of the root server systems were able to endure the attack and stay available to legitimate users throughout the attack.

It is difficult to defend against DDoS because of the lack of security features in TCP/IP specifications. Due to the stateless nature of the Internet, it is difficult to determine the true sources of spoofed IP packets. This thesis will first explore the various issues involved in an IP Traceback scheme, Probabilistic Packet Marking (PPM). In PPM, routers will probabilistically mark packets with partial path information. Based on this information in the marked packets, the victim tries to reconstruct the full path even though the IP addresses of the packets are spoofed. However, PPM suffers from high combination overhead and large number of false positives during path reconstruction. Attackers can also introduce uncertainty by inserting fake path information in the attack packets. Another disadvantage is that it is incapable of performing effective traceback for a large scale DDoS attack.

This thesis introduces two new schemes: Entropy-Minimization Clustering Technique for Probabilistic Packet Marking Scheme [102] and Legitimacy Investigation and Intention-Based Probabilistic Packet Marking Scheme [103] to improve the performance of PPM. The first scheme, Entropy-Minimization Clustering Technique for Probabilistic Packet Marking Scheme is developed to provide a more effective traceback mechanism. The new technique divides the attack traffic into clusters and processes them in parallel. This

Abstract

method of dividing the path reconstruction into smaller clusters significantly reduces the total number of combinations that need to be checked and will in turn minimize the probability of reconstructing a false positive. Our simulation results show that the combination overhead can be reduced by an average of N^9 times, where N is the number of clusters. Our new approach has the same advantage as PPM scheme because it is entirely passive and does not generate any probe traffic into the network. In contrast to the previous work, the new technique is much more efficient and effective during path reconstruction under large-scale DDoS attacks.

The Legitimacy Investigation and Intention-Based Probabilistic Packet Marking Scheme devises a new countermeasure that puts the attacker into a dilemma if they should choose to use spoofed IP addresses during their attack. It also addresses the uncertainty problem under injection of fake path information by the attacker. The origin of the attack can be traced in real time with minimum uncertainty and traceback process is optimized such that those paths that do not forward attack traffic to the victim will be excluded. Path reconstruction can be limited to trace only infected edges thus reducing combination overhead.

In addition, unlike PPM, which is an after-the-fact reaction scheme to an attack, our new scheme allows the victim to maintain high availability to legitimate clients during DDoS attack and traceback process. It is particularly effective in identifying and isolating attack traffic that contained spoofed source address thus improving the overall throughput of the legitimate traffic. The introduction of a novel concept “Identity Tag” provides an inexpensive authentication process to identify false marking as well as resolve the high combination overhead problem in PPM. The results from our simulation also demonstrate that the new approach significantly improves the effectiveness and robustness of PPM by reducing combination overhead and the number of packets required for successful traceback.

Acknowledgements

ACKNOWLEDGEMENTS

I would like to express my deepest and warmest gratitude to my academic supervisors, Professor Lee Bu Sung and Dr. Lee Chee Jwai for their invaluable advice and guidance throughout this research. I am especially grateful to them for the supports and encouragement given throughout the project.

I would also like to extend my sincere appreciation to all the research students and staff members working in Center for Multimedia and Network (CeMNeT) and Institute for Infocomm Research (I2R) for their great assistance. Acknowledgement also goes to Hg Han Sun and For Wei Khing for their kind assistance and encouragement.

CHAPTER ONE

OVERVIEW OF DENIAL OF SERVICE AND DISTRIBUTED DENIAL OF SERVICE ATTACKS

1.1 Background

Denial Of Service (DoS) and Distributed Denial of Service (DDoS) attacks on the Internet have become a pressing issue following several high-profile attacks during recent times [1]. Since early 2000, a number of e-commerce enterprises like Yahoo, Amazon and Ebay [2] [3] were brought down worldwide by the DDoS flood attacks. These attacks have resulted in huge financial losses due to an intolerable level of disruption in Internet services and the threat has continued to ascend. In recent years, there has been an increasing trend whereby the attackers are targeting international corporations with attacks designed not only to disrupt business but as a form of extortion as well [92]. This form of extortion is becoming more common and most of the companies are willing to pay rather than endure the attacks. Some of them have even modernized into a DDoS-for-hire scheme [93] that provides service for clients to launch attacks against business competitors.

Figure 1-1: Number of attacks reported to the CERT/CC

Moore, Voelker and Savage's work [4] demonstrated that DDoS attacks are prevalent in the Internet. They observed 12,805 attacks against 5000 distinct Internet hosts belonging to approximately 2000 distinct organizations during a three-week period. The Computer Emergency Response Team Coordination Center (CERT/CC) [94] has also been keeping overall statistics on Internet attacks from 1993 through 2003. Figure 1-1 shows that the number of attacks reported to the CERT/CC has raised steeply over the past decade due to the widespread use of automated attack tools. On top of that, there is no indication of any sort to expect a drop in the frequency of DDoS attacks in the near future (figure 1-2).

Figure 1-2: DDoS attacks on the Rise [97]

The Computer Crime and Security Survey 2004 [95] is conducted by CSI with the participation of the San Francisco Federal Bureau of Investigation's Computer Intrusion Squad. Figure 1-3 shows the losses caused by different types of computer security incident. The virus and denial of service categories emerged as the two major type of security breaches that generates the largest total losses. This is due to the rise in the degree to which virus attacks were intertwined with denial of service attacks (e.g. numerous variation of the MyDoom worm carried a time-triggered denial of service attack program in its payload).

Figure 1-3: Total Losses caused by Different Security Incidents

1.2 Denial of Service (DoS)

A Denial of Service (DoS) [13] is characterized by an attacker's attempt to thwart legitimate users from using that service. A DoS attack is a type of security breach that does not result in the theft of information or other security loss. It typically causes momentary loss of network availability and services. In the worst-case scenario, for example, an important Internet service accessed by millions of users can be brought down resulting in high monetary loss.

DoS attacks come in many forms and are most commonly executed against scarce resources such as network bandwidth and system resources. Some classic types of DoS attacks are described in the following sections.

Figure 1-4: Denial of Service Attack

1.2.1 TCP SYN Flood Attack

One of the most popular DoS attacks is the *TCP SYN flood attack*, in which the attacker exploits the TCP's three-way handshake mechanism [36] and its limitation in maintaining half-open connections.

Figure 1-5: Standard TCP Three-Way Handshake

A TCP session is established by using a three-way handshake mechanism, which is shown in figure 1-5. The initial step in the procedure is a SYN packet that is sent from the client to the server to establish the connection. The server replies a SYN/ACK, to synchronize and continue the connection establishment. After that the client responds with an ACK message to the server and a connection will be established.

Figure 1-6: TCP SYN Flood Attack

In a TCP SYN flood attack [47] as shown in figure 1-6; the attacker launches a large number of SYN packets using forged source IP address to create many half-open connections that are never completed. The server will never receive the final ACK packet to complete the three-way handshake and the spoofed SYN requests will ultimately exhaust the server's backlog queue, resulting in all incoming SYN requests to be dropped.

1.2.2 UDP Flood Attack

UDP is a connectionless protocol that does not utilize a handshake mechanism for connection establishment and this makes it fairly straightforward for it to be exploited in flood attacks. Pepsi attack [66] is an example of UDP flood where the attacker sends a large number of forged UDP packets aimed at random diagnostic ports on the victim's machine. The scarce resources such as CPU time and memory required to process all these packets would force the victim to cease operation.

1.2.3 Smurf Attack

Figure 1-7 illustrates a Smurf attack in progress. An attacker may also use ICMP Echo Request (PING) packets to bring down the victim's network. Smurf [48] [66] is a classic attack where the attacker sends a large amount of ICMP echo request packets to a broadcast address using the victim's IP address as the source IP.

Figure 1-7: Smurf Attack

The ICMP Echo Reply packets from all the devices will eventually flood all the available bandwidth at the victim's network. In such an attack, the attacker can disable a high-bandwidth network easily because the total amount of attack traffic generated is amplified by a factor equivalent to all devices that respond to the ICMP echo packets. The Fraggle [49] attack is a simple rewrite of the Smurf attack and its only difference lies in sending UDP echo packets instead of ICMP Echo packets.

1.2.4 Teardrop Attack

Teardrop attack [66] is an attack tool that exploits the vulnerability in the TCP/IP fragmentation re-assembly code. When data is sent across a TCP/IP network, it is fragmented into small fragments with an offset field in their TCP header that specifies the starting and ending points of certain data. Teardrop attack takes advantage of this vulnerability [67] by using fragments with invalid overlapping values in the offset field to crash a vulnerable device when it tries to reassemble the data.

1.2.5 Bonk Attack

The Bonk attack [68] is modified from the Teardrop attack where the offset of the second fragmented packet is set to be larger than the length of the IP header. Similar to the Teardrop attack, this may cause the device to crash.

1.2.6 Land Attack

During a Land attack [68], the attacker sends a SYN packet in which the source address and port are the same as the destination. Some implementations of TCP/IP are vulnerable to packets that are crafted in this particular manner and it may cause the device to crash if not handled properly.

1.2.7 Ping of Death Attack

Another well-known DoS attack is the Ping of Death [66] where the attacker sends an IP packet of a size greater than 65,535 bytes to the targeted device. This will result in the buffer overflow problem during reassemble process and may cause the device to crash or reboot.

1.3 Distributed Denial of Service (DDoS)

A DDoS attack [13] deploys numerous machines to exhaust the victim's resources. The essential difference between DDoS and DoS attack lies in the distributed nature, which makes it extremely difficult to trace and stop the attack. It has proven to be a troublesome issue and produced new challenges for the Internet network security community because the cooperative power of a diverse group of attacking machines can easily generate attack volume that is significant enough to make any network inoperable. Legitimate traffic will not be able to compete with the malicious flood and has little chance of obtaining service.

Figure 1-8: Distributed Denial of Service Attack

DDoS attacks have become a major threat to the stability of the Internet [69] because DDoS tools have grown more user-friendly and they are maturing to the point in which even an unsophisticated attacker can do severe damage (Figure 1-9).

Figure 1-9: Attack Sophistication vs. Required Intruder Knowledge [98]

A detail account of the evaluation of DDoS tools is provided by Dave Dittrich at the University of Washington [50] and in the next section we will describe a few of the more popular DDoS tools.

- Trinoo
- Tribe Flood Network (TFN)
- TFN2K
- Stacheldraht

1.3.1 DDoS Attack Architecture

The terminology used in DDoS analysis is often confusing. For clarity, we shall use the following definitions [70]:

- ***DDoS Attack Architecture***: Hackers sometimes setup their own community by sharing resources like compromised machines and information on how to access them. Gradually, all the compromised machines will be organized and connected together forming the overall DDoS attack architecture.
- ***Attacker***: An attacker is the one who utilizes the DDoS attack architecture to launch an attack against a victim.
- ***Master***: Masters are compromised machines with “DDoS master software” installed. The master receives attack signals and acts like a middleman to coordinate and dispatch attack commands to a group of daemons. Master is also known as handler.
- ***Daemons***: Daemons are compromised machines with “DDoS daemons software” installed. The daemons also know as slaves or zombies take on the role as attacking agents by sending out numerous attack packets to the victim.

Figure 1-10: Three-Tier DDoS Attack Architecture

During the initial phase, the attacker first sets up a DDoS attack architecture by recruiting multiple compromised machines by exploiting their security vulnerabilities. This process is normally performed automatically through scanning of remote machines and infecting them with the attack codes such as Stacheldraht or TFN. The infected slaves can then be used for recruitment of new agents. Among all the infected machines, a few of them will be defined as masters while the rest will play the role of attacking agents. The master acts like a middleman to coordinate all the daemons and issue attack requests. A typical description of the entire process of building a Three-Tier DDoS attack Architecture [71], consisting of attacker, masters and daemons is listed:

1. A stolen account is set up as a depository for attack tools, root kits, daemon and master programs, lists of vulnerable machines, compromised hosts identity, keys to access them and etc.
2. A scan is performed to find other vulnerable machines with exploitable security vulnerabilities, such as wu-ftpd, RPC services and etc.
3. A script is then created on the vulnerable machines to perform and complete the exploits.
4. Pre-compiled binaries of the daemon are then created and hidden among the compromised machines.
5. For maximum multitasking and automation of the installation process, another script is secretly created in the background without the owner's knowledge.
6. Usually the master is installed with a "root kit" to hide the presence of the program files and communication activities.

1.3.2 Tribble Flood Network (TFN)

The TFN [72] exhibits a two-tier architecture, consisting of only attacker and daemons. The attacker controls and sends commands to all daemons using a number of connection methods like remote shell bound to a TCP port, UDP based client/server remote shells, ICMP based client/server shells or normal "telnet" TCP terminal sessions. After which communication from the attacker to daemons is accomplished via ICMP_ECHOREPLY packets. The TFN is proficient in waging a SYN flood, ICMP flood, UDP flood, and Smurf attacks.

Figure 1-11: Tribble Flood Network (TFN)

1.3.3 Trinoo

The Trinoo [71] is the first widely known DDoS tool used to launch coordinated UDP flooding attack from numerous sources. It exhibits a three-tier architecture, consisting of attacker, masters and daemons. The additional tier adds an extra layer to the communication, thus making it harder to detect and trace back. The attacker controls and sends commands to one or more master servers via TCP connection to port 27665, which in turn communicate with the daemons via UDP packets on port 27444 to coordinate an attack. To prevent other attackers from controlling the masters and daemons, the attack commands are usually password protected.

Figure 1-12: Trinoo

1.3.4 Stacheldraht

Stacheldraht [74] merges features from Trinoo and TFN by adding encrypted communication, automated update methods and three-tier architecture. The attacker uses an encrypting telnet program to communicate with the master and communication from the master to daemon is done via ICMP or TCP. Similar to TFN, Stacheldraht provides ICMP flood, SYN flood, UDP flood, and Smurf style attacks.

1.3.5 TFN2k

TFN2k [73] is the heir of TFN; it is an improved version of TFN with additional unidirectional and encrypted communication between its two-tier architecture. It uses TCP, UDP, and ICMP packets that are randomized and encrypted making it more

difficult to detect and filter. Decoy packets are also used to complicate attempts to trace the compromised machines.

1.4 DDoS Flooding Attacks

Generally speaking, there are two types of DDoS flooding attacks [16], direct attack (figure 1-13) and reflector attack (figure 1-14). High bandwidth attack is the most popular DDoS attack because it is difficult to stop and has destructive outcome. The magnitude of the combined traffic is significant enough to overwhelm most victim Internet connection (bandwidth exhaustion) or other packet-processing resources (system resource exhaustion).

1.4.1 Direct Attack

In a direct attack [96], an attacker launches a DDoS attack by dispatching commands with the victim's address, attack duration, attack methods to the masters. The masters will then coordinate and issue instructions to all the daemons to execute the DDoS attack. Attack methods can be TCP, ICMP, UDP or a mixture of them. DDoS direct attack is most often a bandwidth starvation attack where a sufficient number of daemons are amassed to generate a flood of malicious traffic aimed at the victim network.

A desire for large throughput had led to the design of high bandwidth pathways in the intermediate network while the end network invested in only as much bandwidth as they might need to use. Such a scenario aggregates DDoS and figure 1-13 illustrates how the massive flood is combined with numerous distributed daemons through the Internet route to congest the victim's network. Under such an attack, the legitimate traffic will not be able to compete with the malicious flood and has little chance to obtain service.

Figure 1-13: Direct Attack

1.4.2 Reflector Attack

A reflector attack [96] is an indirect attack in which reflectors are innocently used as attack launchers. A reflector is any device such as Web servers, DNS servers or routers that will return a response. The attacker sends a large number of packets using the victim's IP address as the source IP to the reflectors and the reply traffic will ultimately flood the victim's network. When TCP attack packets are used, the reflector may respond with either SYN-ACK or RST packets. Smurf is a classic reflector attack that is triggered by sending a large amount of ICMP echo request to a broadcast address using the victim's IP address as the source IP. In addition, attacker can also trigger ICMP error message like ICMP port unreachable or ICMP time exceeded messages.

As illustrated in figure 1-14, a reflector attack utilizes a set of predetermined reflectors and the magnitude of the attack depends on factors such as the number of the reflectors, transmission frequency and size of the reflected packets. Since there are many such reflectors available in the Internet, the attacker can easily diffuse the attack traffic using

many reflectors and the total reflected traffic volume could easily overwhelm the victim's network. Various types of reflector attacks [16] are summarized in table 1.

Figure 1-14: Reflector Attack

Attack Methods	Packets sent by an attacker to a reflector (with a victim's address as the source address)	Packets sent by the reflector to the victim in response
Smurf	ICMP ECHO queries to a subnet-directed broadcast address	ICMP ECHO replies
SYN flooding	TCP SYN packets to public TCP servers	TCP SYN-ACK packets
RST flooding	TCP packets to non-listening ports	TCP RST packets
ICMP flooding	ICMP queries (usually echo queries) UDP packets to non-listening UDP ports IP packets with low TTL values	ICMP replies (usually echo replies) ICMP port unreachable messages ICMP time exceeded messages
DNS reply flooding	DNS (recursive) queries to DNS servers	DNS replies (usually much larger than DNS queries)

Table 1: A Summary of Reflector Attack Methods

1.5 Research Objective

Recent occurrences of DDoS attacks, which employed forged source addresses have proven to be a troublesome issue and produced new challenges for the Internet Security Community. Due to the stateless nature of the Internet, it is difficult to trace the actual source of such spoofed attack without relying on the source IP address. Tracing the true attacking source can be very useful because it is most effective to perform filtering at the source network so that damage to legitimate traffic at the victim network can be minimized. If the attack traffic is capped early, it will confine the overall damage thus preserving Internet resources. Tracing is also helpful in identifying the attacking machines as well as collecting evidences for law enforcement.

1.6 Report Organization

This report is organized into a total of five chapters. Chapter 2 introduces the existing DDoS defense Mechanisms. Chapter 3 describes the proposed enhancement to PPM with Entropy-Minimization Clustering technique. Chapter 4 gives an elaboration on the work on Legitimacy Investigation and Intention-Based Probabilistic Packet Marking Scheme (LI-IPPM). Finally, Chapter 5 contains the conclusion and future work.

CHAPTER TWO

OVERVIEW OF DDOS DEFENSE MECHANISMS

The increasing threat of DDoS problem has attracted much attention from the research community and over the years it has led to the advent of numerous DDoS defense proposals. In general, DDoS defense research can be categorized into two main areas:

- Attack Detection and Response Mechanisms
- IP Traceback

2.1 Attack Detection and Response Mechanisms

Attack detection schemes are responsible for identifying DDoS attacks packets. The scheme should be as accurate as possible because false positive results can lead to inappropriate responses that can cause denial of service to legitimate users. Attack detection scheme [13] can be classified into mechanisms that deploy:

1. Pattern Detection: This scheme requires storing the signature of known attacks in a database. Some of the disadvantages of this scheme are that it requires regular new attack signature update and it is ineffective against new or old attack with slight modification. One example that deploys pattern detection is NetRanger [100].
2. Anomaly Detection: This scheme uses normal traffic dynamics and traffic model to detect anomalies. The current state of the network traffic is periodically compared with the traffic model. The advantage of this technique is that new attacks may be identified. Some examples of mechanisms that deploy anomaly

detection are XenoService [51], Dynabone [52], Dynamically Provisioned Monitoring [53], PeakFlow [54], QuO Middleware [55], Intrusion Tolerance by Unpredictability and Adaptation (ITUA) [56].

3. Hybrid Detection: This scheme combines the features of pattern and anomaly detection scheme to identify new attacks and devise new attack signatures using data from attack discovered to update the database. Some examples of mechanisms that deploy Hybrid detection are CISCO IPS Sensor Software Version 5.0 [99], RealSecure® Network Sensor (version 7.0) and RealSecure® Guard [101].

Figure 2-1: Deployment Location of Attack Detection and Response Mechanisms

The goal of the attack detection scheme is to detect DDoS attack, as early as possible and after identifying the attack characteristic, the response mechanism (e.g. filtering) is responsible for alleviating the impact and improve the overall throughput of the victim. The attack detection and response mechanism (e.g. figure 2-1) can be performed at different locations along the attack path and with regards to the deployment location, they can be differentiated into:

- Source-End Defense Mechanism
- Intermediate Network Defense Mechanism
- Victim-End Network Defense Mechanism

Figure 2-2: Effectiveness of Attack Detection and Response Mechanisms [16]

Figure 2-2 shows a DDoS attack generated from many distributed sources, like the top of a funnel. All the individual streams of attack traffic will be aggregated at the end of the funnel to form a massive flood towards the victim. Detecting the DDoS attack will be easy at the victim-end network since it can monitor all the attack packets. In contrast, it will be much more difficult to detect attacks at the source-end network since most attacking sources are located in a dispersed area. Opposite to the case of attack detection, it is more effective if the response mechanisms are located closer to the attacking sources

because damage to legitimate traffic can be minimized if the attack is filtered close to the source.

2.1.1 Source-End Network Defense Mechanism

The goal of source-end network defense mechanism [13] [16] is desirable because DDoS attack should be stopped as close to the source as possible in order to minimize wastage of resources. But a DDoS defense system located at the source network faces tough detection challenges because it can only observe a small portion of the attack and there is usually no distinct characteristic between the attack and legitimate traffic.

Ingress/Egress Filtering [78] is one of the most renowned source-end network defense mechanism against IP spoofing attacks. It restricts the flow of traffic to a known valid address range by discarding packets whose addresses do not belong to the stub network and limits the impact of attacks using random forged source address. A point worth noting is that Ingress/Egress filtering is not always effective because attackers can still employ subnet spoofing within the permitted prefix range to bypass the filter.

Figure 2-3: Ingress Filter

Figure 2-4: Egress Filter

Other examples of source-end network defense mechanism include Multops [14] and D-ward [15].

2.1.2 Intermediate Network Defense Mechanism

The DDoS defense mechanism is positioned in the intermediate network [13] [16] (e.g. Internet core) to provide defense service to a large number of Internet hosts. It employs a number of distributed detection systems and filtering mechanisms to detect DDoS attack traffic in the Internet core so that they can be suppressed before reaching the victim network.

Route-Based Packet filtering (RPF) [16] [17] is an example of intermediate network defense mechanism proposed by Park and Lee. This technique basically extends the Ingress Filtering function to the intermediate network (e.g. Internet core) and it uses the BGP routing information to filter attack packets that comes from an incorrect link.

Figure 2-5: Illustration of Route-Based Packet Filtering (RPF)

Figure 2-5 shows an Internet Autonomous Systems (ASs) graph. Assume an attacker in AS 7 is launching a DDoS attack against a victim in AS 4 by using spoofed IP address belonging to AS 2. Using the routing information, AS 6 at the peering point with AS 7 will know that packets from AS 2 to AS 4 do not enter through link (7,6) and will therefore discard the attack packets.

Currently the number of ASs is more than 10,000 and for this scheme to be effective, RPF filters need to be placed in at least 1800 ASs. The RPF scheme is ineffective against flooding attacks that use valid IP addresses and it also requires modification to the BGP protocol. Inclusion of the source address in the BGP message will significantly increase the message size as well as the processing time.

2.1.3 Victim-End Network Defense Mechanism

Majority of the DDoS defense mechanisms are positioned in the victim network. The greatest motivating factor for such a deployment is because it suffers the greatest impact during the attack.

Local Attack Detection (LAD) [16] is an example of victim-end network defense mechanism. The deployment and technical complexity for LAD are significantly lower compared to RPF because the detection activities are mainly centralized in the victim network. Attack detection is relatively easy in the victim network since it can observe all the attack packets but the effectiveness of filtering is low because differentiating the legitimate traffic from attack flow is difficult at this point. The filtering mechanism may not help to improve the victim's overall throughput because both attack and legitimate packets would also be dropped.

Figure 2-6: Local Attack Detection Mechanism

Other examples of victim-end network defense mechanism are Resource Accounting [19][20], Protocol Security Mechanisms [21][22] and Netbouncer [23].

2.1.4 A Comparison of Various Attack Detection And Response Mechanisms

Following are some essential metrics [16] used in Table 2 for comparing the various Attack Detection and Response Mechanisms:

False Positive Ratio (FPR): FPR is the number of normal packets that are falsely classified as attack packets divided by the total number of confirmed normal packets.

False Negative Ratio (FNR): FNR is the number of attack packets that are falsely classified as normal packets divided by the total number of confirmed attack packets.

Normal Packet Survival Ratio (NPSR): There is usually no distinguishable characteristic between the attack and legitimate traffic and packet filtering usually drops attack packets as well as normal packets. NPSR gives the percentage of normal packets that can reach the victim during a DDoS attack.

	Ingress Filtering	Route-Based Packet filtering (RPF)	Local Attack detection (LAD)
Detection locations	Source Network	Intermediate Network (e.g. Internet core)	Victim Network
Filtering locations	Same as the detection locations	Same as the detection locations	Same as the detection locations and further upstream ISP networks if backpressure is used
Attack signatures	Spoofed source IP addresses	Spoofed source IP addresses according to the BGP routing information	Traffic anomalies and misuses detected by local intrusion detection systems
False positive ratio (FPR)	= 0	= 0 if the BGP routes are correct	≥ 0 (= 1 in a sufficiently large scale DDoS attack)
False negative ratio (FNR)	≥ 0 (= 0 if all attack packets use spoofed addresses)	≥ 0 (small if most attack packets use spoofed addresses)	≥ 0 (= 0 in a sufficiently large-scale DDoS attack)
Normal packet survival ratio (NPSR)	≥ 0 (= 1 if all attack packets use spoofed addresses)	≥ 0 (large if most attack packets use spoofed addresses and the number of the AS nodes involved in the packet filtering is sufficiently large)	≥ 0 (= 0 in a sufficiently large-scale DDoS attack)
New communication protocols	Not required	Modifications to BGP protocols	Attack alert protocols between victims and their upstream ISP networks if backpressure is used
Computation requirement	Low	Moderate	Low
Technical complexity	Low	High	Moderate without backpressure mechanisms

Table 2: A Comparison of Various Attack Detection And Response Mechanisms [16]

2.2 IP Traceback

DDoS attack is difficult to defend because attacker uses spoofed IP address in the attack packets. In the Internet protocol (IP) [75], there is a header field that contains the source IP address to identify the packet's origin. However, the lack of security features in TCP/IP specifications allows attacker to generate attack packets using spoofed source address. Due to the stateless nature of the Internet, it is difficult to determine the true sources of these spoofed IP packets in a straightforward way.

This section provides an overview on various proposed IP traceback schemes to identify the true sources of the attacker. Some metrics [61] used for comparing various IP traceback schemes are as follow:

1. **ISP Involvement:** Low ISP involvement is a desirable quality because ISPs do not experience a direct benefit from the traceback scheme.
2. **Number of Attacking Packets Needed for Traceback:** This metric affects the computational time required to reconstruct the full attack path. Ideally, the traceback scheme should be able to trace the attacker with as few packets as possible.
3. **The Effect of Partial Deployment:** This metric determines how practical a traceback system is because given the size and complex topologies of today's Internet, not all ISPs will implement this system at the same time. Therefore it is essential for the traceback scheme to produce meaningful results even when deployed partially on the Internet.
4. **Processing Overhead:** This metric refers to the additional processing that may occur on the ISPs network devices and at the victim site. An ideal traceback scheme should incur minimum processing overhead at both the ISPs and victim network.
5. **Bandwidth Overhead:** Additional traffic generated by the traceback scheme is undesirable as it may require the ISP to upgrade existing equipments and invest in additional capacity.

6. **Memory Requirements:** This metric determines the amount of additional memory required by the traceback scheme. Large memory requirement on routers is highly undesirable since it will result in incurring additional costs for the ISP.
7. **Ease of Evasion:** This metric determines the effectiveness of the traceback scheme against a carefully orchestrated attack. This quality should be as low as possible so that the traceback scheme cannot be easily evaded.
8. **Protection:** Protection refers to the ability of a traceback scheme to provide reliable trace information even when a number of devices used in the traceback scheme have been compromised. If protection is low, false positives/negatives [9] will arise making traceback a major challenge. False positive refers to router that appears in the reconstructed attack graph but not in the real attack graph. Similarly false negative refers to router that is left out in the reconstructed graph but exists in the real attack graph.
9. **Scalability:** Scalability is a measure on how scalable the IP traceback scheme is as network size increases. One of the important considerations is that the additional configuration involved in newly added devices should be minimum and independent of each other.
10. **Ability to Handle Major DDoS Attacks:** This is a necessary and important quality that determines how effective the traceback scheme operates under highly distributed and severe DoS attack.
11. **Ability to Trace Transformed Packets:** This is an important metric for a traceback scheme to handle packet transformation such as Network Address Translation and

tunneling. A good traceback scheme would be able to handle any attacks that use packet transformation as a decoy.

2.2.1 Controlled Flooding

Controlled flooding [61] [76] works by generating short burst of traffic to load the upstream network links and by observing the changes in the incoming rate of attack packets, one can identify the links that will lead to the attacking source.

Figure 2-7: Controlled flooding Scheme

This technique is illustrated in figure 2-7 with an attack path R1-R2-R3-R4. The chargen service on the routers is enabled and the controlled flooding equipment loads the upstream network links using short burst of traffic. If there is no change in the attack's intensity, these links and all further upstream paths that may utilize it are excluded from further testing. If a drop in the rate of attack packets is observed, the loading process will be recursively repeated until the source of the attack is identified.

Nowadays, the chargen service on most routers is disabled by default and this technique by itself serves as a form of DoS attack that may disrupt service and cause inappropriate response. This scheme also requires an accurate map of the network topology and the attack must last long enough for a successful trace. It is not able to trace large-scale DDoS attacks because it relies on the fact that during the attack, the links of the attack path will be heavily loaded. This assumption maybe valid for a single-source DoS attack but it may not hold under a DDoS attack. Extremely high bandwidth overhead is also incurred during the traceback process.

2.2.2 ICMP Traceback

ICMP traceback was originally introduced in [62]. Each ICMP Traceback enabled router is configured to statistically produce ICMP traceback message or *iTrace* (e.g. 1 in every 20,000 packets) directed to the same destination as the selected packet.

Figure 2-8 illustrates the ICMP traceback scheme with an attack path R1-R2-R3-R4. Each ICMP Traceback enabled router (e.g. R2) will emit an *iTrace* message which contains the previous and next hop information, a timestamp, an initial TTL value that is set to 255 and part of the payload in the processed IP packet. The victim will eventually receive all the addresses of the routers on the attack path. By using the value in the TTL fields, these addresses can be sorted to reconstruct the attack path.

The ISP involvement for ICMP traceback scheme is low and packet transformation is not an issue. Additional ICMP Traceback enabled routers can be added independently indicating good scalability. It is compatible with existing protocols and allows post-attack analysis. The ability to handle large-scale DDoS attacks is poor because the chances of receiving a useful *iTrace* are very small.

Figure 2-8: ICMP Traceback Scheme

The overhead incur at each ICMP Traceback enabled router is minimal (e.g. 1 in every 20,000 packets) and does not require major upgrades to router hardware. This scheme is compatible with existing routers and network infrastructure. A negligible amount of additional memory is necessary on all the routers and major processing overhead will only occur at the victim site during path reconstruction. A point worth noting is that unless there is a security mechanism for ICMP traceback, attackers can easily inject false iTrace messages to confuse the victim.

2.2.3 Hash-Based IP Traceback

Alex Snoeren introduces Hash-Based IP Traceback [12] [61]. In hash-based traceback, the router computes a 32 bits hash digest using partial information from the packet and stores it in an efficient memory system called a bloom filter. The hash digest is computed using the first 8 bytes of the payload and the non-shaded portions as shown below.

Version	Header Length	Type of Service	Total Length			
Identification				D F	M F	Fragment Offset
TTL		Protocol	Checksum			
Source Address						
Destination Address						
Options						
Payload						

Figure 2-9: Fields shaded are masked out before digesting

Figure 2-10: Hash-Based IP Traceback Scheme

Figure 2-10 illustrates the Hash-Based IP Traceback scheme with an attack path R1-R2-R3-R4. Data Generation Agents (DGAs) functionality is installed on these routers and the whole network is divided into regions. The DGA produces hash digest of each departing IP packet and stores them in a memory system. The SPIE Traceback Manager (STM) is a central management unit that controls the SPIE Collection and Reduction Agents (SCARs) and DGAs within each region. When an attack is detected, the SCARs produce individual attack graph for its region and the STM gathers all this information to complete the full traceback process.

The steps to reconstruct the attack path through the various regions are as follows:

1. The STM is notified when a DDoS attack is detected.
2. The STM communicates and dispatch instructions to the SCARs.
3. The SCARs queries the DGAs and check the memory system (e.g. digest table) stored locally.
4. After analyzing the information stored in the memory system, the SCARs will identify the routers that have seen the attack packet digests. The SCAR can then reconstruct the attack path through its network. This information is then transferred to the STM to build the full attack graph.

ISP involvement in this scheme is high because additional equipment has to be installed to support the function of DGAs, SCARs and STM. Inter-ISP tracing is possible provided that all participating ISPs form an agreement for cooperation. It is compatible with existing protocols and supports incremental implementation. There is no additional processing and bandwidth overhead incurred at the victim site but large amount of additional memory is required in the network.

This scheme is able to handle DoS attack and it requires only a single packet to trace each attacker. The limitation of this scheme is that the computational and storage overhead is

very high since it needs to compute and store a bloom filter digest for every packet. Although it supports post-attack analysis, traceback must be performed within a very short period of time before copies of the digests are overwritten. It is less suitable for large scale DDoS attacks because it is very resource intensive in terms of storage and processing requirements.

2.2.4 Overlay Network

CenterTrack [61] [77] is an overlay network, consisting of IP tunnels that are used to reroute packets from edge routers to Tracking Routers (TRs). The TRs monitor all traffic that is routed through the network.

This scheme is illustrated in figure 2-11 where the TR is installed in the center of the star topology and IP tunnels are used to reroute packets directly from edge routers. The TR examines all the packets and makes decision whether to drop or forward them.

Figure 2-11: CenterTrack Scheme [61]

There are several limitations in this scheme like ISP involvement is large because the ISP has to monitor, detect and identify the attack. The scalability is poor because additional configuration is required if new edge router is added to the region. Another limitation is that this scheme requires great infrastructure change because in order to trace through multiple ISPs, it would be necessary to connect all the TRs into a single system.

In this scheme, the additional traffic generated for every packet is about 20 bytes, and for an attack composed of really small packets, this can result in significant bandwidth overhead. Handling packet transformations is not an issue for this scheme but the protection is rather low because it may not produce valid traces if a router is subverted.

2.2.5 Probabilistic Packet Marking (PPM)

Savage et al proposed one promising scheme known as Probabilistic Packet Marking (PPM) [8] [61]. In this scheme, routers will mark the packets with partial path information using a fixed probability $p=1/25$. Based on the information in the marked packets, the victim can reconstruct the full path even though the IP addresses of the packets are spoofed.

Figure 2-12 illustrates the PPM traceback scheme with an attack path R1-R2-R3-R4. Each router will probabilistically mark its partial address information into the packet headers. Eventually the victim will receive enough information to reconstruct the addresses of all routers along the attack path using the path reconstruction process.

Figure 2-12: PPM Traceback Scheme

The PPM scheme has a low overhead for routers and supports partial or incremental deployment. It also does not require interactive operational support from Internet Service Providers (ISP), nor major upgrade to the current router hardware. ISP involvement is therefore low and additional PPM-enabled routers can be deployed independently, which also indicates good scalability. The overhead incur at each PPM-enabled router is minimal, and major processing overhead is incurred at the victim site during reconstruction. Bandwidth overhead for this scheme is zero since all information is scrambled in the IP packet header. The PPM scheme can handle packet transformation but fragmented traffic will be corrupted by the scheme resulting in backward compatibility issues. It also allows for post attack analysis and traceback.

PPM scheme has a number of drawbacks. Among them is the high combination overhead and large number of false positives during path reconstruction. Attackers can also introduce uncertainty during path reconstruction by inserting fake path information in the attack packets. Another disadvantage is that it is incapable of performing effective traceback for a large scale DDoS attack.

2.2.6 Adjusted Probabilistic Packet Marking

Tao Peng, Christopher and Kotairi introduce a new approach Adjusted Probabilistic Packet Marking (APPM) [11]. In the APPM scheme, instead of using a fixed probability p , each router uses a different marking probability to mark each packet. They propose to use a decreasing marking probability of routers as the packets traverse towards the victim. Each router will mark the packets with respect to the distance of the router from the destination using probability distribution:

$$p = \frac{1}{c - d_v}$$

c is a constant and d_v represents the distance of the router from the destination.

2.2.7 Dynamic Probabilistic Packet Marking

Jenshiuh, Zhi Jian and Yeh Ching also proposed a Dynamic Probabilistic Packet Marking (DPPM) [80] to improve the effectiveness of PPM. The only difference between PPM and DPPM scheme is the determination of marking probability. Instead of using a fixed marking probability, they proposed to use the Time-to live (TTL) value in the IP header to deduce how far a packet has traveled and then choose the marking probability as an inverse function of hop count traveled.

2.2.8 Comparison of the Traceback Schemes

		PPM	ICMP Traceback	Hash-Based IP Traceback	Controlled flooding	Overlay Network
ISP Involvement		Low	Low	Fair	None	Large
Scalability		High	High	Fair	N/A	Poor
Number of attack packets required for traceback		High	High	Low	High	Low
Is partial deployment possible?		Yes	Yes	Yes	N/A	No
Network processing overhead	Packet transit	Low	Low	Low	None	Low
	During traceback	None	None	Low	None	High
Victim processing overhead	Every packet	None	None	None	None	None
	During traceback	High	High	None	Low	None
Bandwidth overhead	Every packet	None	Low	None	None	High
	During traceback	None	None	Low	High	None
Memory requirements	Network	None	Low	High	None	Low
	Victim	High	High	None	Low	None
Ease of evasion		Low	High	Low	N/A	Low
Protection		Fair	Fair	Fair	N/A	Fair
Ability to handle packet Transformations		Good	Good	Good	Good	Good
Ability to handle major DDoS attacks		Poor	Poor	Fair	Unable	Good
Limitations		DoS and DDoS attack only.	DoS and DDoS attack only.	Strict timing constrains on traceback. Single point of failure	DoS only. Manual. Unsafe. Inconsistent. Traceback is possible only while attack is in progress.	Single ISP. Single point of failure

Table 3: Comparison of the Traceback Schemes [61]

CHAPTER THREE

ENTROPY-MINIMIZATION CLUSTERING TECHNIQUE FOR PROBABILISTIC PACKET MARKING SCHEME

Probabilistic Packet Marking (PPM) and Adjusted Probabilistic Packet Marking (APPM) have been proposed for the identification of the true sources of spoofed IP packets typically used in DoS attacks. However, PPM/APPM suffers from high combination overhead and large false positives under large-scale DDoS attacks. In this chapter, we present an Entropy-Minimization Clustering Technique to solve the limitations in PPM scheme. The new approach has the same advantage as PPM/APPM scheme but it is much more efficient and effective in the path reconstruction under DDoS attacks. The technique divides the attack traffic into clusters based on shared bottleneck. The attack graph for each cluster can then be reconstructed separately and grafted together to form the complete attack graph. This method of dividing the path reconstruction into smaller clusters significantly reduces the total number of combinations that need to be checked and will in turn minimize the probability of reconstructing a false positive. The technique also preserves the advantages of the PPM scheme, as it works with any type of traffic (TCP, UDP, etc). It does not generate any new network traffic and it utilizes only the information at the IP layer. Theoretical analysis and simulation studies using ns-2 software are carried out to evaluate the proposed technique. The results demonstrate that the new approach gives significantly higher precision and lower combination overhead for attack paths reconstruction under large scale DDoS.

3.1 Overview of Probabilistic Packet Marking (PPM) Scheme

PPM scheme (figure 3-1) is based on the idea that when a router decides to mark a packet, it uses edge-sampling algorithm to write one of the edge-id fragments at random in the IP packet's 16-bit Identification field (figure 3-2).

Savage et. al. [8] proposed to use the edge-sampling algorithm to write edge information that represents the link between itself and the previous router into the packet. In theory, this algorithm requires two 32 bits address-sized fields, start and end, and an 8 bits (theoretical maximum number of hops allowed) distance field in each IP packet. Each router will probabilistically mark the packet by writing its own address into the start field and writes zero into the distance field. If a subsequent router sees that the distance field is already zero, it will assume that the previous router has marked the packet and it will write its own address into the end field and increment the distance field by one. If a router decides not to mark the packet, it will always increment the distance field. Edgesampling algorithm requires 64 bits of space for the two static address-sized fields and another 8 bits for the distance field in every IP packets.

Figure 3-3: Encoding Scheme

In order to reduce the storage requirements in each packet, Savage et. al. modified the edge-sampling algorithm using an edge-id fragment encoding scheme to overload the 16-bit identification field used for fragmentation. The encoding scheme (figure 3-3) calculates a uniform 32-bit hash of the router's IP address and this hash is interleaved

with the original IP address by inserting the hash and original address into the odd and even bits respectively. The resulting quantity is then broken into 8 fragments and each fragment is selected randomly and stored with an offset.

Figure 3-4: Encoding of edge-id fragments

Figure 3-5: Path Reconstruction

Downstream routers will use XOR to combine fragments at the same offset to form an edge-id fragment (figure 3-4). Over time the victim will receive enough edge-id fragments (e.g. d , $c \oplus d$, $b \oplus c$ and $a \oplus b$ edge data). The victim then uses the edge sampled in these packets to reconstruct the attack graph. The original path can be reconstructed from XORing all the received edge-id fragments as shown in figure 3-5.

Figure 3-6: Address Verification Procedure

In figure 3-6, the victim produces a 64 bits string by combining any 8 edge fragments and this string is then de-interleaved by separating the address and hash portion. The address is accepted as valid if the resulting hash over this address portion is the same as the hash portion extracted.

3.1.1 High Combination Overhead

We denote a router i at distance d as $R_d(i)$ and each edge fragment element in $R_d(i)$ as $EF_d(i,f)$, where $f \in [0,7]$ represents the offset of each fragment. Each packet will probabilistically contain one of the eight edge fragments from the router $R_d(i)$ that marked the packet.

$$R_d(i) = \{EF_d(i,0), EF_d(i,1), EF_d(i,2), EF_d(i,3), EF_d(i,4), EF_d(i,5), EF_d(i,6), EF_d(i,7)\}$$

There are eight edge fragment elements for each router $R_d(i)$ and if there are multiple path attacks, the victim will not be able to distinguish which eight fragments are encoded from the same router. In order to reconstruct the attack graph, the victim needs to consider all possible combination of the eight elements and the result is accepted as valid if the hash value of the odd bits matches the even bits. Let N_d represents the total number of distinct routers at distance d .

We denote C_d as the total number of combinations that needs to be considered at distance d .

$$C_d = (N_d)^8$$

The use of the XOR function further complicates the reconstruction because all resulting edge-id at distance $d+1$ has to be XOR with the previous edge-id and all combination of XOR values must be checked as the attack paths diverge.

We denote Z_{d+1} as the total number of combinations that has to be XOR.

$$Z_{d+1} = (N_d) \times (N_{d+1})^8$$

So the total number of combinations Z_{total} to be checked for all the distances is

$$Z_{total} = \sum_{0 \leq d \leq \max} \{(N_d) \times (N_{d+1})^8\}$$

The combination overhead is very high when there are multiple attacks paths during DDoS attack.

3.1.2 Large Number of False Positives

For large distributed attacks, the PPM scheme has serious limitations due to the difficulty in distinguishing which eight fragments are from the encoding of the same router during multiple path attacks. In order to reconstruct the attack path, the victim needs to consider all possible combination of the eight elements and the result is accepted as valid if the hash value of the odd bits matches the even bits. Consequently, the probability of misattributing an edge, as well as the amount of state needed, increases very quickly with the fan-out of an attack graph.

The probability of accepting an arbitrarily constructed candidate edge-id is $1/2^{32}$ if a random hash length $h=32$ is used. The probability of reconstructing an incorrect edge for

an event where there are m attackers located in m distinct routers at distance d away is at most: $1-(1-1/2^{32})^{m^8}$ [8].

3.2 PPM with Entropy-Minimization Clustering Technique

We propose an entropy-minimization clustering technique [24] at the victim end network that will help to improve the attack path reconstruction by reducing the high combination overhead and number of false positives for PPM scheme. During a DDoS attack, the victim will have to reconstruct the whole attack graph by testing all combination and the task will be especially difficult when multiple attacks paths diverge. With the clustering technique, the victim only needs to check the combination within each cluster in order to establish the attack path. It will be able to reconstruct the attack graphs within each cluster separately and graft all resulting graphs to form a complete attack graph.

Figure 3-7: Packets leave the Bottleneck Router Equally Spaced

The entropy-minimization clustering technique allows the victim to cluster the flows that share the same bottleneck path. The scheme relies exclusively on the packets' arrival timing at the receiver to group flows that run through the same bottleneck router. It works with any type of traffic (TCP, UDP, etc) and it does not generate any new network traffic. The fundamental idea behind the entropy-minimization clustering technique is as follows,

if the queue at the bottleneck router is congested, then the output link is always used to the maximum capacity and packets leave the bottleneck equally spaced (figure 3-7). The transmission time of one packet (T) is equivalent to the time difference between the first bits of two consecutive packets.

$$T = \text{packet-size} / \text{bottleneck-bandwidth}$$

Figure 3-8 shows a simple network where the victim receives all incoming traffic over the same link.

Figure 3-8: Simple Network Illustration

Figure 3-9 illustrates the inter-packet spacing at different locations in the above network. In figure 3-8, A1 and A2 are behind the same bottleneck router B1 and figure 3-9(a) shows the inter-packet spacing (A1 and A2) at the output of B1. Similarly A3 and A4 share the same bottleneck router B2 and figure 3-9(b) shows the inter-packet spacing (A3 and A4) at the output of B2. Figure 3-9(c) shows the inter-packet spacing at the victim which is the superimpose of the output of B1 and B2.

Figure 3-9: Inter-Packet Spacing of Different Cluster

Figure 3-9(c) shows that the uniform inter-packet spacing will be lost at the victim as packets blend with traffic coming from B1 and B2. The victim will have to make use of some clustering techniques to group the flows that share the same bottleneck router into the correct cluster. If the victim mistakenly groups A2 and A3 together, it will result in more random inter-packet spacing as shown in figure 3-9(d).

Figure 3-10: PMF of the Inter-Packet Spacing

Figure 3-10 shows the probability mass function (PMF) for the inter-packet spacing of the correct cluster in Figure 3-9(a) {A1 and A2}, and the incorrect cluster in figure 3-9(d) {A2 and A3}. The PMF in figure 3-10(a) shows one long pulse because the inter-packet spacing of the correct cluster has a uniform spacing. Conversely, the grouping of A2 and A3 results in more random inter-packet spacing and the resulting PMF in figure 3-10(b) shows many small pulses at various locations.

Since most DDoS attacks are only effective so long as they occupy the resources of the victim. Consequently, most attacks comprised thousands to millions of packets with the intention of depleting the network bandwidth connecting the server to the rest of the Internet. If this attack succeeds, the network bandwidth will be sufficiently depleted and legitimate users will experience severe or complete service degradation. Therefore we can make a fair assumption that attacking sources send enough packets to keep the bottleneck router's queue occupied. Consequently, packets leave the bottleneck evenly spaced and the victim can cluster all the flows according to their shared bottleneck using information from the probability mass function (PMF) of the inter-packet spacing. Since

the PMF of the inter-packet spacing of the correct clustering will have more structure than that of an incorrect clustering, the victim can use a clustering technique to minimize the entropy using the PMF of the inter-packet spacing.

The discussion of well-know clustering techniques is widespread in literature. Authors of [24] further discussed in more in-depth details about the performance of various clustering techniques on different topologies with different traffic patterns and showed that 2D Cluster-Weighted KMeans has the best performance. With such a clustering technique, we compute using the cost function below:

$$Cost = \frac{1}{N} \sum_{c=1}^{c=N} N_c H_c$$

The 2D Cluster-Weighted KMeans technique uses iterative procedures that start with some initial random clustering for the flows. During each iteration, the flow is moved among clusters to obtain an incremental reduction in the average entropy. The cost function is a weighted average of the entropies of the cluster, where the weighting factor is the number of flows N_c in the cluster, N is the number of clusters and H_c is the entropy of cluster c . We provide the following definitions:

- Entropy: a measure of the uncertainty in a random variable [25]. The entropy $H(x)$ of a discrete random variable x (inter-packet spacing), with probability mass function $p(x)$ is defined by the following expression. The entropy (H_c) is computed using a random vector where the first and second component is the current and previous inter-packet spacing respectively.

$$H(x) = -\sum_x p(x) \log_2 p(x)$$

- Flow (N_c): the set of packets with identical marking value in the identification field. We assume that all packets in a flow follow the same route and consequently share the same bottleneck [24] [26].
- Cluster (N): the set of flows ordered by their arrival time. A correct cluster consists of flows that run through similar bottleneck router.

The 2D Cluster-Weighted KMeans technique is used to identify and group the attack flows into smaller clusters before the victim performs the path reconstruction. Within each smaller cluster we can reconstruct the attack graph for each cluster separately and graft all resulting graphs to form the complete attack graph. This method of dividing the path reconstruction into smaller cluster will significantly reduce the total number of combination that needs to be checked and this will in turn minimize the probability of accidentally reconstructing a false positive. With this clustering technique we can also perform the path reconstruction simultaneously for each cluster to provide a faster construction of the attack graph.

The clustering technique may not be 100% accurate in grouping all the flows into the correct cluster but this will not affect the accuracy of the reconstruction process because the address verification is protected with an error detection procedure in the PPM scheme. The victim produces a 64 bits string by combining any 8 edge fragments and this string is then de-interleaved by separating the address and hash portion. The address is accepted as valid if the resulting hash over this address portion is the same as the hash portion extracted. When the reconstruction process for each particular cluster is completed, all the incorrect edge-fragment-ids from different paths will be separated. A final round of path reconstruction process will then be carried out for all the incorrect edge-fragment-ids collected from every cluster.

3.3 Simulation Studies

Simulation studies are used to evaluate the performance of Entropy-Minimization Clustering technique for PPM and APPM scheme, in terms of probability mass function (PMF) for the inter-packet spacing of the correct cluster. We also compare the effectiveness of the PPM and Entropy-Minimization Clustering technique with PPM, in terms of combination overhead for path reconstruction.

3.3.1 Simulation Model

NS-2 network simulation software is used to implement the PPM/APPM scheme and construct agents for the attackers, the intermediate routers and the victim. For the PPM scheme, each router will mark the packets with a fixed probability $p=1/25$ as proposed in [8]. Under the APPM scheme [11], each router will probabilistically mark packets based on the distance of the router from the destination:

$$p = \frac{1}{c - d_v}$$

$c=25$ is selected because few paths exceed this length [27] [28] and d_v represent the distance of the router from the destination.

We evaluate the PPM and APPM scheme using the topology given in Figure 3-11 and the receiver (R) will create a data record by logging the arrival time and the value in the identification field of each packet. After generating the data record using ns-2, 2D Cluster-Weighted KMeans technique is used to identify and group the attack flows

according to their common bottleneck before path reconstruction. Matlab software is used to implement the clustering technique and to analyze the PMF of the inter-packet spacing to determine the correct clusters.

3.3.2 Network Scenario 1

Using the topology shown in figure 3-11, we generate a total of 15 attacking sources situated at node 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 19, 20, 21, 22 and 24 respectively. Each attacking source uses a TCP-based attack (e.g. TCP SYN packets) and we select different permutation for the attackers' start time. We also vary each attacking source's rate such that enough attacking packets are sent to keep the queues at the 3 bottleneck routers B1, B2 and B3 congested. We also experiment using different packet dropping mechanisms such as Random Early Detection and droptail.

The upstream topology in figure 3-11 was designed for the worst case with multiple distributed attack paths giving rise to high combination overhead during the reconstruction procedure. We select a higher link bandwidth for the intermediate router so that all the individual streams of traffic from the many separate attacking sources will be aggregated to form a single massive flood towards the victim network.

Figure 3-11: Network Scenario 1

3.3.3 Network Scenario 2

For comparison purposes we perform another set of simulation studies using the topology in figure 3-12 with 50 attackers situated at distance 8 hops away from the victim. In the simulation scenarios, the attack traffic arriving at the victim causes 2 bottleneck routers that will result in 2 correct clusters.

NS-2 network simulation software is used to generate 50 distributed attacking sources situated at distance of 8 hops away from the victim. All the 50 attacking sources are

evenly distributed within each of the 2 clusters and the topology was simulated for the worst case with one attacker situated at each of the 50 end nodes, giving rise to multiple distributed attack paths. In the topology, a total of 108 intermediate routers is constructed and we select a slightly higher link bandwidth for all intermediate routers between the victim and attacker sites. Each attacking source uses a TCP-based attack (e.g. TCP SYN packets) and we select different permutation for the attackers' start time and use different dropping mechanisms such as Random Early Detection and Droptail. We also vary each attacking TCP source rate such that the bottleneck router $B1$ and $B2$ queues are congested.

Figure 3-12: Network Scenario 2

3.4 Results and Performance Analysis

In this section, we compare the performance of Entropy-Minimization Clustering technique for PPM and APPM scheme, in terms of probability mass function (PMF) for the inter-packet spacing of the correct cluster. We also compare the performance of the two schemes in terms of combinational overhead to establish the full attack graphs.

3.4.1 PMF Analysis for Network Scenario 1 using PPM

(a) Cluster 1

(b) Cluster 2

(c) Cluster 3

Figure 3-13: PMF of the inter-packet spacing for the correct cluster using PPM scheme

Figure 3-13 shows the PMF for the inter-packet spacing of the correct cluster using PPM scheme. The figure shows many small pulses that follow a first long spike because under the PPM scheme each router will mark the packet with a fixed probability $p=1/25$ as proposed in [8]. There will be a probability of $\alpha = (1-p)^d$ that a packet will reach the destination without any marking, where d is the average attack path length. The inter-packet spacing within the correct cluster will be more randomized since there will be a probability α that the packets will not be marked by the intermediate routers. They cause the PMF to show many tiny spikes that follow the long one, which represent the bottleneck spacing. Therefore the PMF of the 3 correct clusters in figure 3-13 is more random compared to figure 3-10 resulting in a higher average entropy.

3.4.2 PMF Analysis for Network Scenario 1 using APPM

(a) Cluster 1

(b) Cluster 2

(c) Cluster 3

Figure 3-14: PMF of the inter-packet spacing for the correct cluster using APPM Scheme

For comparison purpose, we perform a similar simulation using APPM scheme [11] where each router computes a different marking probability $p = \frac{1}{c - d_v}$ to mark packets.

The advantage of this scheme is that a router chooses a high marking probability if the packet is far away from the victim and decrease the marking probability as the packet travels along the path. By increasing the marking probability at the edge routers, majority of the spoofed packets can be overwritten by router information and simultaneously the number of unmarked packets is reduced.

The figure 3-14 shows that the first long spike has a higher value compared to those in Figure 3-13. This will help to achieve a better performance for the number of flows to be correctly classified because the PMF of the correct clusters will have lower entropy due to more uniform inter-packet spacing.

3.4.3 Combination Overhead Analysis for Network Scenario1

The combination overhead for the reconstruction of various hops of the attack path for the PPM/APPM and PPM/APPM with Entropy-Minimization Clustering technique scheme was obtained.

Figure 3-15: Combination overhead using PPM/APPM

Figure 3-16: Combination overhead using PPM/APPM with Entropy-Minimization Clustering Technique

In figure 3-15 we show that PPM/APPM scheme does not scale well under DDoS attacks in the sense that with more attacking machines used, the greater is the effort needed by the victim during the reconstruction procedure. The total number of combinations for the topology given in figure 3-11 can be computed using

$$Z_{total} = \sum_{0 \leq d < 3} \{ (N_d) \times (N_{d+1})^8 \}$$

Using PPM/APPMM with Entropy-Minimization Clustering technique we are able to reduce the high combination overhead for path reconstruction significantly. In figure 3-11, the total number of router at any distance d is evenly distributed among the 3 clusters and we can therefore compute the total number of combinations within each cluster to be

$$Z_{total} = \sum_{0 \leq d < 3} \left\{ \left(\frac{N_d}{N} \right) \times \left(\frac{N_{d+1}}{N} \right)^8 \right\}$$

where N represents the total number of cluster.

The result in figure 3-16 shows that through identification and clustering of the attack paths, we can group the packets from different paths into 3 smaller clusters and within each cluster we are able to reduce the combination overhead by an average of 3^8 times.

3.4.4 PMF Analysis for Network Scenario 2 using PPM

(a) Cluster 1

(b) Cluster 2

Figure 3-17: PMF of the inter-packet spacing for the correct cluster using PPM scheme

The figure 3-17 shows that the long spike, which represents the bottleneck spacing, has a higher value compared to those in Figure 3-13 and 3-14. It is clear that the unmarked probability $\alpha = (1-p)^d$ of an attack packet reaching the victim is a decreasing function of d . As the average attack path length d is increased, the number of unmarked packets is reduced and this results in a less randomized inter-packet spacing within the correct cluster. Therefore we observe that the PMF of the 2 correct clusters are less random resulting in lower entropy.

3.4.5 PMF Analysis for Network Scenario 2 using APPM

(a) Cluster 1

(b) Cluster 2

Figure 3-18: PMF of the inter-packet spacing for the correct cluster using APPM scheme

Figure 3-18 shows that the long spike, which represents the bottleneck spacing, has a higher value compared to those in Figure 3-17. This also illustrates that the PMF of the correct cluster using APPM scheme performs much better than PPM scheme, since the entropy is lower due to more uniform inter-packet spacing within the correct cluster and this will help to achieve a better performance for the number of flows to be correctly classified.

3.4.6 Combination Overhead Analysis for Network Scenario2

Figure 3-19: Combination overhead using PPM/APPM

Figure 3-20: Combination overhead using PPM/APPm with Entropy-Minimization
Clustering technique

We compare the performances of the two schemes in terms of combinational overhead to establish the full attack graphs. Similarly, the total number of combinations for the topology given in figure 3-12 can be computed using

$$Z_{total} = \sum_{0 \leq d < 3} \left\{ (N_d) \times (N_{d+1})^8 \right\}$$

In figure 3-20, the total number of router at any distance d is evenly distributed between the 2 clusters and we can compute the total number of combinations within each cluster using

$$Z_{total} = \sum_{0 \leq d < 3} \left\{ \left(\frac{N_d}{N} \right) \times \left(\frac{N_{d+1}}{N} \right)^8 \right\}$$

The result in figure 3-20 shows that through identification and clustering of the attack paths into 2 smaller clusters, the combination overhead can be reduced by an average of 2^9 times.

The results from both network scenario 1 and 2 show that we can achieve an average combination reduction ratio of N^9 for each individual cluster, where N is the number of clusters. Therefore the combination reduction ratio will increase with higher N .

3.5 *Summary of Results*

We present PPM/APPm with Entropy-Minimization Clustering Technique that reduces the combination overhead and enhances the accuracy of the attack path reconstruction procedure. Our new approach has the same advantage as PPM/APPm scheme because the entropy-minimization clustering technique is entirely passive and does not generate any probe traffic into the network. In contrast to previous work, our technique is much more efficient and effective during path reconstruction under large-scale DDoS attacks. The technique is used to divide the attack traffic into clusters based on shared bottleneck and the attack graph for each smaller cluster are then reconstructed separately and grafted together to form the complete attack graph. This method of dividing the path reconstruction into smaller clusters will significantly reduce the total number of combinations that needs to be checked and will in turn minimize the probability of reconstructing a false positive.

Theoretical analysis and simulation studies using ns-2 software are carried out to evaluate the proposed technique. The results demonstrate that the new approach reduces the combination overhead and enhances the accuracy of the path reconstruction procedure. The results from network scenario 1 and 2 show that through clustering of the attack paths we are able to achieve an average combination reduction ratio of N^9 for each individual cluster, where N is the number of clusters. Therefore the performance of our technique will be better with higher N . With this clustering technique we can also perform the path reconstruction simultaneously for each cluster to provide a faster construction of the full attack graph and achieve a lower false positive rate.

CHAPTER FOUR

LEGITIMACY INVESTIGATION AND INTENTION-BASED PROBABILISTIC PACKET MARKING SCHEME

Although IP spoofing plays an important role in attacks, it is also to the attacker's advantage to inject spoofed marking value into the packets. The spoofing of the marking field in the attack packets allows the attacker to introduce uncertainty and reduce the effectiveness of PPM. These form of attacks have proven to be a troublesome issue and produced new challenges for the Internet network security community.

In this chapter, we introduce a new countermeasure against IP spoofing called Legitimacy Investigation and Intention-Based Probabilistic Packet Marking Scheme (LI-IPPM) to optimize traceback process. Since IP spoofing plays an important role in DDoS attack, LI-IPPM uses this to its advantage and put the attacker into a predicament because they either choose to expose their true location by using real IP address or risk being rapidly traced by using spoofed IP addresses. The origin of the attack can be traced in real time with minimum uncertainty and traceback process is optimized such that those paths that do not forward attack traffic to the victim will be excluded. Path reconstruction can be limited to trace only infected edges thus reducing combination overhead.

In addition, unlike PPM, which is an after-the-fact reaction scheme to an attack, LI-IPPM allows a victim network to maintain high availability to legitimate clients during DDoS attack and traceback process. It is particularly effective in identifying and isolating attack traffic that contained spoofed source address thus improving the overall throughput of the legitimate traffic. We also introduce the novel concept of "Identity Tag" (ID-Tag) for lightweight authentication test as well as to reduce the high combination overhead during path reconstruction. LI-IPPM preserves the advantage of PPM and improves its effectiveness by reducing uncertainty, path reconstruction time and combinational

overhead. Theoretical analysis and simulation studies are carried out using ns-2 software to evaluate the LI-IPPM in terms of combination overhead and path reconstruction time. Furthermore we also analyze tradeoff between the marking probability and performance metrics.

4.1 Effectiveness of Probabilistic Packet Marking (PPM) and its Variant Schemes

Kihong and Heejo [29] studied the efficiency of PPM scheme under injection of corrupted marking by the attacker. They showed that the attacker could introduce uncertainty factor m and reduce the effectiveness of PPM by inserting fake path information in the attack packets. Traceback is extremely challenging because there is very little that can be trusted. When a victim receives a marking, it has no way of telling whether an upstream router has marked the packet or the attacker has forged the information. If any routers along the attack path do not overwrite the spoofed packet, it can impede the victim's ability to identify the true attack sources. An attacker can always send false signals to create the illusion of additional routers in the traceback graph.

Figure 4-1: Spoofing Additional Traceback Paths

Although IP spoofing plays an important role in attacks, it is also to the attacker's advantage to inject spoofed marking value into the packets. An attacker can spoof the marking field with edge (X_i, Y_i) where $i=1,2,3\dots k$ as shown in figure 4-1. If downstream routers do not overwrite the spoofed marking, the traceback will give x_1 to x_k number of false sources of attack in addition to the real one.

They also showed that PPM is effective against single source DoS attack and it has limitation in the case of large scale DDoS attack. This is because given a desired attack volume N to achieve denial of service; attacker can use concurrent and small volume of attack traffic from separate sources. Even in the absence of spoof marking, by increasing the number of attacking sources (M), the number of packets N/M from each source will decrease and may even fall below the minimum constraint for the victim to receive sufficient number of packets marking from the furthest routers. PPM scheme relies on the assumption that attacker sends numerous packets because routers mark each packet with only a small piece of path information and the victim must receive many such packets in order to reconstruct the full attack path. This assumption may be valid under a single-source DoS attack but it may not hold under a DDoS attack. It is also shown that the larger the number of attacking sources M , the higher the amplification of uncertainty m . With spoofed markings in DDoS attacks, PPM will even be less effective and uncertainty will be greatly amplified during traceback.

The spoofing of the marking field in the attack packets makes traceback by PPM a major challenge. To address this problem, Tao Peng, Christopher and Kotairi introduce Adjusted Probabilistic Packet Marking (APPM) [11]. In the APPM scheme, instead of using a fixed probability p , each router uses a different marking probability to mark each packet. Although this scheme reduces the unmarked probability for any packet, it is still vulnerable to spoofing of the marking field. According to [79], their results show that APPM cannot eliminate spoofing of the marking field because the constant $c=30$

proposed in [11] is too high and large percentage of unmarked packets will reach the victim for small path length.

Jenshiuh, Zhi Jian and Yeh Ching also proposed a Dynamic Probabilistic Packet Marking (DPPM) [80] to improve the effectiveness of PPM. Instead of using a fixed marking probability, they proposed to use the Time-to live (TTL) value in the IP header to deduce the marking probability. Although DPPM may outperform PPM, it still suffers an uncertainty of z under spoofed TTL attack, where z is the difference between initial TTL value and the spoofed packet's TTL value.

4.2 Legitimacy Testing

In [23], they presented an approach for legitimacy testing using an adaptation of the idea of SYN cookies proposed by [35] and ICMP echo message proposed by [37].

SYN cookies technology allows the defense mechanism to handle TCP connection request in a stateless manner as connection establishment is performed only when the host returns the legitimate ACK. In figure 4-2, when the ACK arrives at the defense mechanism and if the cookie has been verified then the original SYN along with client's sequence number is recreated based on the information in the ACK packet and the stored cookie. This SYN packet will be forwarded to the victim's server, the defense mechanism will complete the 3-way handshake. It will also handle the transformation of the sequence number on each packet on the TCP connection to and from the server.

Figure 4-2: SYN Cookies Legitimacy Testing

Figure 4-3: ICMP Echo Message Legitimacy Testing

When dealing with packets of type ICMP or UDP, the defense mechanism as shown in figure 4-3 employs the use of ICMP echo message [37] to determine if the incoming request belongs to a live host.

In figure 4-3, if an ICMP echo reply is received and the Hash Message Authentication Code (HMAC) can be verified, the request packet will be forwarded to the server. A HMAC is computed using a key hash function and taking the following input: a packet

tag, the incoming request packet, source IP address, payload-length, expiry-time, and a nonce. In order to authenticate and verify the ICMP echo replies, this HMAC is included into the outgoing ICMP echo request.

4.3 *Intention-Driven iTrace*

Intention-Driven iTrace [70] was proposed to enhance ICMP traceback (iTrace). It was shown that iTrace suffered a serious statistic problem such that the chance for useful and valuable iTrace messages could be extremely small against a major DDoS attack with the goal of minimizing the probability of generating useful iTrace. To address this limitation, they introduce the concepts of “usefulness” and “value” and propose an Intention-Driven iTrace scheme to improve the value of each generated iTrace messages.

They propose that when a network is under attack, the intrusion detection system will use the BGP routing protocol as the vehicle to distribute the intention signal to all routers. Each iTrace enabled router utilizes this information to decide whether a new iTrace message will be sent to the victim.

4.4 *Legitimacy Investigation and Intention-Based Probabilistic Packet Marking Scheme (LI-IPPM)*

Our Legitimacy Investigation and Intention-Based Probabilistic Packet Marking Scheme (LI-IPPM) acquires inspiration from the Legitimacy Testing and Intention-Driven iTrace to improve the performance of PPM. Since IP spoofing plays an important role in DDoS attack, our scheme uses the concept of Legitimacy Testing to its advantage and devise a countermeasure that put the attacker in a dilemma because they either opt to expose their true location by using real IP address or risk being quickly traced via spoofed IP address. Many IP traceback schemes including PPM focus solely on tracking the location of the

attacking sources and it is usually an after-the-fact reaction to a DDoS attack. During this period of time when the attack is raging on, the victim can do nothing to alleviate the damage or provide better services to the legitimate clients. Our proposed scheme addresses this limitation and it allows a victim network to maintain high availability to legitimate clients during DDoS attacks and traceback process. It is particularly effective in identifying and isolating majority of the DDoS traffic that contained spoofed source address thus improving the overall throughput of the legitimate traffic. Our scheme extends PPM technique by improving its effectiveness and robustness under DDoS attacks by eliminating uncertainty factor m and minimizing the number of packets required for successful traceback. It uses a countermeasure that employs source IP spoofing characteristic to trace the origin of the attack in real time with minimum uncertainty. During traceback process, it is able to differentiate packets that contain genuine source IP addresses from spoofed addresses and attack packets are filtered before reaching the victim's premises. Our new scheme preserves the advantages of the PPM and improves its effectiveness and performance in four aspects:

- 1) The total number of false positives and false negatives is an important factor in determining the accuracy of the traceback scheme. We denote false positive [9] as router that appears in the reconstructed attack graph but not in the real attack graph. Similarly false negative [9] refers to router that is left out in the reconstructed attack graph but exists in the real attack graph. One contributing factor for false positive is due to a fundamental shortcoming in the PPM scheme where attackers can always send false signals to invent additional routers in the traceback graph. Let $\alpha_o(p)$ be the probability that a spoofed packet is not overwritten by any router:

$$\alpha_o(p) = (1 - p)^d$$

where d is the total path length.

Since PPM is based on marking packets using a fixed probability $p=1/25$, a large number of attack packets will reach the victim unmarked. For example, an attacking source at a distance $d = 10$ hops away, 67% of the attack packets will reach the victim unmarked. Another contributing factor that results in false negative is that packet markings are not authenticated and a compromised router can simply forge the markings of any other routers to prevent the traceback scheme from identifying it. The intrinsic ability of attackers to spoof the marking field in the attack packets will result in amplification of uncertainty factor m and makes traceback by PPM a major challenge.

- 2) The combination overhead plays a critical role in affecting the efficiency of real time traceback. The PPM scheme experiences extremely high combination overhead under large-scale attacks because it is not able to distinguish which eight fragments are encoded from the same router.
- 3) Due to the fixed probability marking feature in PPM scheme, routers further away from the victim will contribute relatively less marking because downstream routers will overwrite majority of the marking. Let $\alpha_i(p)$ be the probability that a packet is marked at router i :

$$\alpha_i(p) = p(1-p)^{d-i}$$

where d is the total path length.

The probability that the arriving packet is last marked by the furthest router will be

$$\alpha_1(p) = p(1-p)^{d-1}$$

Which is a decreasing function of distance d . This will increase the computational time required to reconstruct the full attack path because the victim needs to receive numerous packets before it can obtain sufficient path information from further upstream router.

- 4) Unlike PPM, our scheme LI-IPPM allows a victim network to sustain high availability to legitimate clients while the attack is raging on. It also maximizes the victim's throughput by protecting the traffic of established connections as well as the rate at which new connection are opened during traceback process.

4.4.1 Proposed Traceback Architecture

The size and multiple administrative networks comprising of different access and security policies in the Internet makes it impossible to centrally manage the entire traceback process. We therefore propose our traceback architecture based on distributed management approach [82] that controls tracing through a number of interconnected network domains. Our traceback architecture is designed for efficiency and scalability in large-scale network like the Internet.

In order for our traceback architecture to function well across multiple domains, it would be necessary to build a cooperating community where all participating domains form an agreement for cooperation. The detail description of the agreement or certification procedure has been active for the past decade; we therefore propose to adopt a similar style of operation like CERT [86], CSIRT [83] [85] and FIRST [84]. In their approach,

all managerial details (e.g. access policy, encryption keys exchange, security matters, etc) are arranged and set up before the architecture is deployed [82]. Each participating network domain has an area of responsibility and all interconnected domains make an effort to cooperate forming our traceback architecture based on distributed management approach.

Figure 4-4 illustrates our traceback architecture and the interaction among the various functional components in a multiple domains network. The function operation of Investigation Agent, Query Processing Agent, Authentication Test Module, Attack Graph Generation Module and Intention-Based Probabilistic Packet Marking Module are dispersed among separate components in the cooperating community. The Investigation Agent, Query Processing Agent, Authentication Test Module and Attack Graph Generation Module are installed at a set of strategic perimeter routers typically belonging to a local ISP. The Intention-Based Probabilistic Packet Marking Module is an extension to the current PPM scheme and is installed on every Internet router.

A manager device is assigned to each participating network domain and controls a group of routers installed with Intention-Based Probabilistic Packet Marking Module. Its main responsibility is to respond to the Investigation Agent's alert request and sends commands to activate the intention-based module located within its domain. Each master device is inactive initially and is triggered when it receives an activation alert message from the Investigation Agent during an ongoing attack. We assume that DDoS attack generally will not significantly impact unidirectional alert requests from Investigation Agent to Manager Device, although performance in the other direction can be affected. This is usually true in today's routers that use switched architecture and full-duplex link.

Figure 4-4: Proposed Traceback Architecture

When a request is presented to the manager, it cryptographically verifies its authentication and integrity. Any Investigation Agent that wishes to employ its service must be properly authenticated in order to prevent any damage by attackers. The request

authentication is indispensable because attacker can intentionally generate false request and the response from such devices may provide a tool for them to exploit. Once a session is established between the Manager and the Investigation Agent, the Manager will act like a proxy to translate information and command to activate the respective Intention-Based Probabilistic Packet Marking Module. The message exchanged between all devices must be protected for integrity and confidentiality. But as our interest here is to focus on the mechanisms used to identify the attacking source, we therefore make the assumption that there is already an existing secure key management and authentication scheme and the managers from different network domain are willing to cooperate in tracing the DDoS attack. We also proposed to adopt the Intrusion Detection Message Exchange Format (IDMEF) [87] [88] [89] as the common language for communication and transporting alert messages among devices. This protocol is currently in draft state and is developed by the Internet Engineering Task Force (IETF) [89].

Although our LI-IPPM solves several problems in PPM scheme, these improvements do not come for free. In our scheme, additional bandwidth overhead will be introduced but it will only occur during traceback process after an attack has been identified. At any given time, we expect at most only a minor portion of the Internet under attack, since attacker usually target sites with heavy customer utilization in order to cause maximal disruption. To achieve this aim, they will also be inclined to focus the combined attack power of multiple machines onto such sites. Therefore bandwidth overhead will be infrequent and acceptable in volume since we use a victim initiated negotiation approach for activation and traceback process. If additional bandwidth overhead were to incur at all, it would be preferable to occur only during traceback process.

4.4.2 Architecture of LI-IPPM

Figure 4-5 and figure 4-6 depict the four modules in our proposed scheme. In the following section, we describe and illustrate the functionality for each of the module:

- Intention-Based Probabilistic Packet Marking Module
- Investigation Agent Module
- Query Processing Agent and Authentication Test Module
- Attack Graph Generation Module

4.4.3 Intention-Based Probabilistic Packet Marking Module

Intention-Based Probabilistic Packet Marking Module (IPPM) as shown in figure 4-5 is an extension to the current probabilistic marking module and it is run on each Internet router. Due to the fixed marking probability featured in PPM scheme, there is a very low chance that the victim will receive valuable marking information from routers near to the attackers. Routers far away from the victim will contribute relatively less marking because majority of them will be overwritten by downstream routers. As a result a significant amount of packets needs to be collected and this increases the computational time required for path reconstruction. Furthermore, fixed marking probability feature allows the attackers to introduce uncertainty and reduce the effectiveness of PPM by inserting fake path information in the attack packets.

Figure 4-5: Intention-Based Probabilistic Packet Marking Module

In order to address the limitations in fixed probability packet marking, we introduce the concept of IPPM, which takes into account the intention of the receiver. Since there are so many possible network addresses in the Internet and by merely marking packet with a fixed probability will not be desirable. It is clear that it will be more beneficial to generate more markings to victim who desperately need it than just randomly mark and send this information to a receiver who is not under attack or does not care about tracing. Efficient path reconstruction depends a lot on the coordination of upstream routers because it affects how fast the useful information from the furthest upstream routers can be received. To improve the efficiency of traceback, it is important to design an effective, secure and accurate scheme that can allow a router to vary the marking probability based

on the intention of a victim. The IPPM is compatible with the PPM scheme since it is just an extension to it. It preserves all the advantages of the PPM as it will only generate more marking for packets destined to someone who is interested to use that information and it will not affect the performance of the marking ratio to other network.

This module divides the marking probability into Intention-Based or fixed probability marking and it uses the edge-id fragment encoding scheme [8] to create an edge-id table. The table contains the 64-bits hash results of all links between itself and its adjacent routers. The proposed idea utilizes the fact that routers are able to determine the physical network interface port packets. This ability is already been used in input debugging [30] [31], CentreTrack [32], link testing [33] and Huffman Codes marking scheme [34].

The IPPM uses 25 bits for marking and this set of fields is shown below:

- 1) TOS Field: The type of service field is an 8-bit field in the IP header that is currently used to allow hosts a way to give hints to routers as to what kind of route is important for particular packets. Authors of [10] show that that setting this field arbitrarily makes no measurable difference in packet delivery.
- 2) The ID Field: The ID field is a 16-bit field used by IP to permit reconstruction of fragments and authors of [8] show that less than 0.25% of all Internet traffic is fragments.
- 3) The Unused Fragment Flag: There is an unused bit in the fragment flags field that current Internet standards require to be zero. Authors of [10] have found that setting this bit to one has no effect on current implementations; with the exception that when receiving the packet, some systems will think it is a fragment. The

packet is still successfully delivered, however, because it looks to those systems as though it were fragment 1 of 1.

Figure 4-6: Set of Fields Used For Marking

4.4.4 Investigation Agent Module

Investigation Agent Module as shown in figure 4-7 is installed at a set of strategic perimeter routers typically belonging to a local ISP. It is triggered using a victim-initiated negotiation approach when the network suffers from severe congestion. During traceback process, it is also able to differentiate attack packets from legitimate traffic as well as providing high availability to legitimate clients while the attack is raging on.

Attack packet types can be TCP, ICMP, UDP or a mixture of them. It has been shown by Moore et. al. [4] that more than 94% of the DDoS attacks use TCP, followed by UDP (2%) and ICMP (2%) packets. In the TCP case, SYN flooding is the most commonly used. During a DDoS flood attack; the Investigation Agent Module will maximize the victim's throughput by protecting the traffic of established connections against arbitrary packet attacks. It is able to differentiate legitimate clients who wish to open new connection and it will not throttle the rate at which new connections are opened. The objective of the Investigation Agent is to isolate and protect legitimate traffic from huge volume of DDoS attack flow. It will differentiate packets that contain genuine source IP addresses from spoofed addresses. The Investigation Agent will not rate limit the arrival of SYN packets at which new connection are opened.

When faced with a SYN flood, the agent will use the SYN cookies technology to intercept SYN packets. SYN cookies technology allows the defense mechanism to handle TCP connection request in a stateless manner as connection establishment is performed only when the host returns the legitimate ACK. If the original SYN packet uses a spoofed address that is not in used then no ACK will be returned. If the original spoofed address belongs to another host who did not send the request then a reset (RST) packets will be generated. In both cases, the original SYN request will be effectively ignored. State is only reserved when the host returns the ACK. Once the connection is completed, the agent will contact the server to establish the connection and keep acting as a proxy to

translate the flow between the client and server. The IP source and destination address and TCP source and destination port will be updated in the address information table. After which the agent will only allow packets with established connection to reach the victim. It will also prevent illegitimate packets send to arbitrary ports to flow through and the victim will only receive packets at ports at which it is listening or as part of an established connection.

When dealing with DDoS attacks that use ICMP or UDP packets, the Investigation Agent employs the use of ICMP echo message to determine if the incoming request belongs to a live host. If the source address of the incoming packet is not in the address information table, the Investigation Agent will challenge the source using ICMP echo request.

All attack packets that contain spoofed addresses source IP addresses will then be forwarded to the Query Processing Agent and Authentication Test Module for real time traceback processing. Since IP spoofing plays an important role in DDoS attack, the Investigation Agent uses this to its advantage to devise a countermeasure that places the attacker in a dilemma because they either opt to expose their true location by using real IP address or risk being promptly traced via spoofed IP address.

In view of the fact that the Investigation Agent Module has the ability to generate outgoing messages in response to incoming requests, which can be exploited by attacker to launch reflector attacks. A crafty attacker can send attack packets to the Investigation Agent Module using other victim's address as the source IP to launch an indirect attack on other network. To address this limitation, we propose a Reflector Prevention-Caching device to monitor the source IP address of incoming requests. This device controls the amount of outgoing requests to each IP address and monitors for failed connections. If a certain IP address with a large number of failed connections is observed within a short interval of time, protection measures (e.g. blacklisting or outgoing request control) can be employed.

Figure 4-7: Architecture of Legitimacy Investigation Module

4.4.5 Query Processing Agent and Authentication Test Module

Query Processing Agent and Authentication Test Module as shown in figure 4-7 is installed together with the Investigation Agent Module. The Query Processing Agent Module (QPAM) makes decisions by using information from Investigation Agent Module and Attack Graph Generation Module to determine which upstream routers should receive the Intention-Based marking signal.

A serious vulnerability in PPM is that it is unable to identify false marking inscribed by the attacker, which will result in false positives or negatives during path reconstruction. In order to address this limitation, the Authentication Test Module is used to identify fake marking and eliminate uncertainty.

To authenticate the packet marking, we proposed a novel concept of “Identity Tag” that makes use of the 9 bits in TOS and Unused Fragment Flag field. The QPAM first generates a sequence of 9-bits Identity Tag (ID-Tag) and maintains a database for them. Beginning at the network domain nearest to the victim, the QPAM issues an alert request and creates a session with the upstream manager. As mention earlier, all alert requests must be properly authenticated and only when verification is successful, the manager will issue the Intention-Based signal to activate each IPPM for a short period of time. In response to the request from the QPAM, each respective router will also be assigned a unique ID-Tag. This ID-Tag varies each time whenever the QPAM issues a request and it will remain valid for a sufficient period of time before expiring to prevent replay attacks. The QPAM identifies each upstream router using the assigned ID-Tag and the Authentication Test Module is used to check the validity of the received packets before path reconstruction.

The Authentication Test Module does not require expensive cryptographic operations and the authentication process for ID-Tag is easy to verify. Inexpensive authentication is very important because expensive public-key operations will create new opportunities for attacker. It is always vital to ensure our scheme does not leave new ways that can be exploited by attacker.

4.4.6 Attack Graph Generation Module

Attack Graph Generation Module (AGGM) as shown in figure 4-7 is installed together with the Investigation Agent and Query Processing Agent and Authentication Test Module. The AGGM performs intelligent decisions using information from Query Processing Agent and Authentication Test Module to identify infected edge and reconstruct accurate attack paths.

The combination overhead plays a critical role in affecting the efficiency of real time traceback. For large distributed attacks, the PPM scheme has serious limitation due to high combination overhead during attack path reconstruction [8] [9]. In order to resolve the high combination overhead problem, the ID-Tag provides an additional key contribution besides authentication testing.

The AGGM uses the ID-Tag together with the distance field value to group and link the fragments. Since each upstream router is assigned a unique ID-Tag, the Path Reconstruction Agent can use it to identify which eight fragments are from the encoding of the same router. During path reconstruction, the victim just needs to consider the eight elements with the same ID-Tag and distance field to find the combination overhead:

$$Z_{total} = \sum_{0 \leq d \leq \max} \left\{ (2.15)^8 \times (N_d) \right\}$$

where N_d represents the total number of distinct routers at distance d . The above equation utilizes the fact that in a router-level Internet map, the average degree (number of neighboring routers) is 3.15 [34] [57]; hence the average number of upstream routers will be 2.15.

Based on the information given by QPAM and Investigation Agent Module, the AGGM is able to identify the infected edge and path reconstruction can be limited to trace this set of paths that leads back to the attacking sources. During path reconstruction, the AGGM will send feedback signal to the QPAM indicating which upstream routers should receive the Intention-Based marking signal. This is an improvement compared to Intention-Driven iTrace because we optimize our traceback process such that routers that do not forward attack traffic or not involved in the attack will not be activated. This will also ensure that additional bandwidth overhead introduced is infrequent and acceptable in volume.

4.5 Theoretical Analysis

In this section, we will analyze the issues related to the tradeoff between the choice of marking probability and the performance metrics. The choice of marking probability directly influences the accuracy and efficiency of LI-IPPM. The performance metrics we will use are as followed:

1. **Path Reconstruction Time:** An important criterion for real time traceback is the ability to identify the attacking source based on as few packets as possible. In our proposed scheme, less number of packets is required to get the job done but at the expense of some additional computational overhead and bandwidth overhead. The added communication among the various devices allows the IPPM to vary the

marking probability based on the intention of a victim. This improves the efficiency of traceback because useful information from the furthest upstream routers can be received faster. Additional bandwidth overhead will be infrequent and acceptable in volume since we use a victim-initiated approach and it will only occur during traceback process after an attack has been identified.

2. **Uncertainty Factor (m)** [29]: This is an important factor in determining the accuracy of the traceback scheme. Fixed marking probability feature makes traceback by PPM a major challenge because it allows the attacker to introduce uncertainty. In our proposed scheme, the marking probability is increased during traceback process to ensure that all false markings injected by the attacker are overwritten.
3. **Computational Overhead ($C_{Overhead}$)**: Each marking by the IPPM poses some computational overhead on the router. For our analysis, we choose to use the number of marking performed by each router as the individual $C_{Overhead}$ experienced. Consider an attack with N_{total} number of packets:

$$N_{total} = N_{Attack} + N_{Good} + N_{Other}$$

where N_{Attack} and N_{Good} denote the number of attack and good packets sent to the victim. N_{Other} represents the cross-traffic (number of packets sent to other destination).

In PPM, all routers have the same individual $C_{Overhead}$ incurred because each router uses $p_{ppm} = 1/25$ to mark packets.

$$C_{Overhead} = N_{Total} \times p_{ppm}$$

On the other hand, in APPM, each router uses $p = \frac{1}{c - d_v}$ to mark packets.

Therefore, the individual $C_{Overhead}$ is

$$C_{Overhead} = \frac{N_{Total}}{c - d_v}$$

where $c=25$ is selected since few Internet paths exceed this length [8].

Similarly in DPPM, each router r_d ($1 \leq d \leq D$) uses $p=1/d$ to mark packets and the individual $C_{Overhead}$ is

$$C_{Overhead} = \frac{N_{Total}}{d}$$

Under the LI-IPPM scheme, the individual $C_{Overhead}$ is

$$C_{Overhead} = (N_{Other} \times p_{ppm}) + (N_{Attack} + N_{Good}) \times p_{LI-IPPM}$$

Figure 4-8: Computational Overhead

For simplicity, we assume that the attacker is 24 hops away from the victim, $p_{LI-IPPM}=1$, $N_{Total}=100,000$, $N_{Attack}=30,000$, $N_{Good}=20,000$ and $N_{other}=50,000$. From Figure 4-8, it can be seen that all routers in PPM have constant $C_{Overhead}$ while the first few routers in DPPM and APPM suffer very high overhead.

For LI-IPPM, some interesting points deserve our attention. The additional computational overhead $(N_{Total}+N_{Good}) \cdot p_{LI-IPPM}$ for each individual router along the attack path only occurs for small duration Δt . In the next section, we will show that Δt is very small because it only takes approximately 8 packets to identify an infected router. Furthermore, additional $C_{Overhead}$ is incurred only during traceback, which is relatively infrequent. Since DDoS attacks are the exception rather than the norm, the overhead incurred by LI-IPPM in the long run is approximately $(N_{other}+N_{Good}) \cdot p_{ppm}$ which is practically the same as PPM. Another point worth noting is that Internet routers are already capable of executing real time manipulations such as recalculating the IP header checksum and decrementing the TTL value of every packet they forward. Therefore the overall performance of the router need not suffer if the LI-IPPM marking can be performed in parallel with the above manipulation. In addition, our scheme optimizes the marking process by generating information only for packets destined to victim.

We also observe that the path reconstruction time and uncertainty factor m will decrease with higher marking probability. In a real Internet topology, the $C_{Overhead}$ incur in each IPPM router varies according to its relative distance from the attacking source. In a DDoS attack with multiple distributed sources, all the individual streams of attack traffic will ultimately aggregate together forming a massive flood towards the victim. It is therefore clear that routers closer to the victim will suffer slightly higher $C_{Overhead}$ compared to further upstream routers. However due to the high volume of attack traffic, IPPM nearer to the victim will get its job done much faster and minimize the duration of overload. On the other hand, since the attacking sources are sufficiently distributed, each IPPM near the source will only observe a small portion of the attack traffic and the $C_{Overhead}$ incurred will also be lower.

4.6 Analysis and Simulation Results

We use NS-2 network simulation software to implement the LI-IPPM. We construct a tree topology and generate agents for 1024 attackers situated at a distance of 10 hops away from the victim. The tree topology as shown in figure 4-9 is one in which each interior node has a fan-out of exactly 2 children.

Figure 4-9: Tree Topology

We will compare and evaluate the effectiveness of our scheme in terms of combinational overhead for path reconstruction as well as the time taken to establish the full attack graph.

4.6.1 Combination Overhead Analysis

The combination overhead plays a critical role in affecting the efficiency of real time traceback and from figure 4-10, we illustrate that PPM, APPM and DPPM does not cope well under major DDoS attack. It suffers from extremely high combination overhead because it is not able to distinguish which eight fragments are encoded from the same router.

Figure 4-10: Combination Overhead

In order to reconstruct the attack graph, the victim needs to consider all possible combination of the eight fragments from multiple attack paths, which will result in

$$Z_{total} = \sum_{0 \leq d \leq 10} \{(N_d) \times (N_{d+1})^8\}$$

where N_d represents the total number of distinct routers at distance d [8] [9].

Using LI-IPPM, we are able to reduce the high combination overhead significantly. The total number of combinations to be checked during path reconstruction is

$$Z_{total} = \sum_{0 \leq d \leq 10} \{(2)^8 \times (N_d)\}$$

where N_d represents the total number of distinct routers at distance d . The above equation utilizes the fact that every interior node in our tree topology has exactly 2 upstream routers. The number of neighboring routers in our simulation is 3, which is approximately the same as a real router-level Internet map with 3.15 neighboring routers [34] [57]. The results from figure 4-10 show that with LI-IPPM, the total combination overhead can be reduced by an average of 10^{21} times.

4.6.2 Path Reconstruction Time Analysis

The time required to reconstruct the full attack path using PPM, APPM, DPPM and LI-IPPM was obtained. We use the largest value for marking probability $p=1$ to illustrate the maximum efficiency of LI-IPPM and $p=1/25$ for PPM as proposed in [8]. Also worthy of mention is that $p=1$ is not applicable for PPM because with this choice of p , the last hop router will overwrite all the markings from upstream routers.

Figure 4-11: Number of packets required for reconstruction

Due to the fixed p in PPM, routers far away from the victim will contribute relatively less marking. This will increase the time taken to reconstruct the full attack path because the victim needs to receive numerous packets before it can obtain sufficient path information from further upstream router. In PPM, the average number of attack packets required to reconstruct each independent attack path of length d is

$$E(n) < \frac{k \ln(kd)}{p(1-p)^{d-1}}$$

where k is the number of edge fragments (e.g. $k=8$) [8].

Although APPM performs better than PPM when the attack path is large, it cannot achieve the optimum performance for small path length. This is due to the fact that constant $c=25$ is still too high for small path length and large percentage of packets will not be marked. A point worth noting is that APPM relies on the Internet routing protocol to obtain the d_v value thus making it more difficult for attacker to tamper with.

For the DPPM scheme, we used an unified initial TTL value $T_{ini}=32$ as proposed in [80] and each router will deduce the marking probability based on the packet's TTL value. DPPM performs far better than APPM and PPM because the first router will always mark with $p=1$ when it sees any packets with TTL equals T_{ini} . However, the DPPM results in figure 4-11 assume that the TTL value has not been tampered with. Attackers can easily beat the DPPM by sending packets with spoofed TTL value and the first router will view it as originated at $1+z$ hops away. If a well-informed attacker uses $z=20$, router r_d ($1 \leq d \leq 10$) will mark these packets with

$$p = \frac{1}{d+z}$$

and it will take the victim approximately 1000 packets to reconstruct the full attack path.

Using LI-IPPM with $p=1$, we are able to reduce the total path reconstruction time significantly because only an average of 8 packets is needed to reconstruct each edge regardless of its distance from the victim. The average number of attack packets required to reconstruct each independent attack path of length d is approximately

$$E(n) = 8 \times d$$

The results from figure 4-11 show that LI-IPPM significantly outperforms all the 3 schemes. Furthermore, using the maximum p value ensures zero uncertainty since all false markings injected by the attacker will be overwritten.

4.7 Summary of Results

We present LI-IPPM using the concept of Legitimacy Testing to devise a countermeasure that put the attacker into a dilemma if they should choose to use spoofed IP addresses during their attack. It uses a countermeasure that exploits source IP spoofing

characteristic to trace the origin of the attack in real time with minimum uncertainty. Traceback process is optimized such that those paths that do not forward attack traffic to the victim will be excluded and path reconstruction can be limited to edges that lead back to the attacking sources. Our scheme improves the effectiveness and robustness of PPM under DDoS attacks by reducing combination overhead and the number of packets required for successful traceback. It also allows a victim network to improve the overall throughput of the legitimate traffic and sustain high availability during the traceback process. The concept of IPPM and Authentication Test Module is introduced into PPM to address its statistic problem and to identify false marking inscribed by the attacker. The introduction of the novel concept ID-Tag provides an inexpensive authentication process to identify false marking as well as resolve the high combination overhead problem in PPM.

Next, we identified some essential performance metrics and analyze the tradeoff between marking probability and computational overhead incurred by the LI-IPPM. Theoretical analysis and simulation studies using ns-2 software are also carried out to evaluate the LI-IPPM. The results from our simulation demonstrate that LI-IPPM is able to reduce the combination overhead significantly by 10^{21} times. In addition, the reconstruction time is also reduced by approximately 16 times compared to PPM.

CHAPTER FIVE

CONCLUSION

A detail investigation of issues related with Denial of Service Attack, Distributed Denial of Service Attack, DDOS Defense Mechanisms and IP Traceback Schemes has been done. Problems associated with Probabilistic Packet Marking (PPM) has also been identified.

Two new schemes, Entropy-Minimization Clustering Technique for Probabilistic Packet Marking Scheme [102] and Legitimacy Investigation and Intention-Based Probabilistic Packet Marking Scheme [103] have been proposed to improve the performance of PPM, APPM and DPPM.

Entropy-Minimization Clustering Technique for Probabilistic Packet Marking Scheme has been proposed to solve the high combination overhead and large false positives problems in PPM and APPM schemes. The Entropy-Minimization Clustering Technique divides the attack traffic into clusters based on shared bottleneck. The attack graph for each cluster can then be reconstructed separately and grafted together to form the complete attack graph. This method of dividing the path reconstruction into smaller clusters significantly reduces the total number of combinations that need to be checked and will in turn minimize the probability of reconstructing a false positive. The results from our simulation demonstrated that through clustering of the attack paths, we are able to achieve an average combination reduction ratio of N^{θ} for each individual cluster, where N is the number of clusters. Our proposed scheme has the same advantage as PPM/APPM scheme because the entropy-minimization clustering technique is entirely passive and does not generate any probe traffic into the network. It utilizes only the information at the IP layer and works with any type of traffic (TCP, UDP, etc).

Legitimacy Investigation and Intention-Based Probabilistic Packet Marking Scheme (LI-IPPM) uses a countermeasure that employs IP spoofing characteristic to optimize traceback process. It also addresses the uncertainty problem under injection of false marking value by the attacker. LI-IPPM preserves the advantage of PPM and improves its effectiveness by reducing uncertainty, path reconstruction time and combinational overhead. The origin of the attack can be traced in real time with minimum uncertainty and path reconstruction will be limited to trace those edges that forward attack traffic. It also allows a victim network to improve the overall throughput of the legitimate traffic during DDoS attack and traceback process. In addition, we introduce the novel concept of Identify Tag for lightweight authentication test as well as to reduce the high combination overhead during path reconstruction. The results from our simulation demonstrated that combination overhead could be reduced significantly by 10^{21} times compare to PPM, APPM and DPPM. In addition, the reconstruction time is also reduced by approximately 16 times compare to PPM.

5.1 FUTURE WORK

As for ongoing and future work, we plan to study the performance of our proposed schemes on real network traffic. We are also investigating several issues associated with IPv6, the proposed successor to IPv4. PPM scheme does not address implementation in IPv6 and there could be challenges because IPv6 does not have an identification field and the IP address is 128 bits in length. We believe a feasible solution could be worked out before IPv6 become widely deploy and one possible answer is to overload the 24-bit flow label field in IPv6 [104].

LIST OF PUBLICATIONS

- Wei-Peng Tan, Bu-Sung Lee, Henry C. J. Lee, “Entropy-Minimization Clustering Technique For Probabilistic Packet Marking Scheme,” accepted by 12th IEEE International Conference On Networks 16-19 Nov. 2004, Pg 292-298.
- Wei-Peng Tan, Bu-Sung Lee, Henry C. J. Lee, “Legitimacy Investigation And Intention-Based Probabilistic Packet Marking Scheme,” accepted by IEEE, 6th Malaysia International Conference on Communications (MICC) in conjunction with 13th IEEE International Conference on Networks (ICON), 17-18 Nov. 2005.

LIST OF REFERENCES

- [1] L. Garber. “Denial-of-service attacks rip the Internet”. *Computer*, pages 12–17, Apr. 2000.
- [2] CNN Headline News, “ Distributed DoS Attacks on Yahoo, Buy.com, eBay, Amazon, Datek, E*Trade”, Feb 7-11, 2000
- [3] CNN Headline News, “Distributed DoS Attack Block Microsoft Web Sites”, Jan. 26, 2001.
- [4] D. Moore, G. Voelker, and S. Savage.” Inferring Internet Denial of Service Activity”. Proc. of USENIX Security Symposium’ 2001, August 2001.
- [5] K. Houle and G. Weaver, “Trends in Denial of Service Attack Technology,” <http://www.cert.org/>, October 2001.
- [6] Sven Dietrich, Neil Long, and David Dittrich. “Analyzing distributed denial of service attack tools: The Shaft case”. In 14th Systems Administration Conference, LISA 2000.
- [7] Dave Dittrich. “Distributed Denial of Service (DDoS) attacks/tools resource page”. <http://staff.washington.edu/dittrich/misc/ddos/>, 2003.
- [8] Savage,S.; Wetherall, D.; Karlin, A.; Anderson, T.; “Practical network support for IP Traceback,” *Networking, IEEE/ACM Transactions on*, Volume: 9, Issue: 3, June 2001 Pages:226 – 237
- [9] Dawn X Song and Adrian Perrig, “Advanced and Authenticated Marking Schemes for IP traceback”, In Proc. Of IEEE INFOCOM 2001, volume 2, pp.878-886, Anchorage, AK
- [10] Drew Dean, Franklin Matt, Adam Stubblefield, “An Algebraic Approach to IP Traceback”, *ACM Transactions on Information and systems Security*, May 2002, volume 5, issue 2, pp.119-137.

List Of References

- [11] T. Peng, C. Leckie and R. Kotagiri, "Adjusted probabilistic packet marking for IP Traceback", Networking 2002, Pisa, Italy, May 2002
- [12] A. C. Snoeren et al., "Single-Packet IP Traceback," IEEE/ACM Trans. Net, vol. 10, no. 6, Dec. 2002, pp. 721–34.
- [13] Jelena Mirkovic, Peter Reiher, "A taxonomy of DDoS attack and DDoS defense mechanisms," ACM SIGCOMM Computer Communication Review, April 2004, Volume 34 Issue 2
- [14] T. M. Gil and M. Poletto. "MULTOPS: a data-structure for bandwidth attack detection," In Proceedings of 10th Usenix Security Symposium, August 2001.
- [15] J. Mirkovic, G. Prier, and P. Reiher., "Attacking DDoS at the Source," In Proceedings of the ICNP 2002, November 2002.
- [16] Chang, R.K.C., "Defending against flooding-based distributed denial-of-service attacks: a tutorial," Communications Magazine, IEEE, Volume: 40, Issue: 10, Oct. 2002 Pages: 42 – 51
- [17] K. Park and H. Lee, "On the Effectiveness of Route-Based Packet Filtering for Distributed DoS Attack Prevention in Power-Law Internets," Proc. ACM SIGCOMM, Aug. 2001, pp. 15–26.
- [18] R. Mahajan, S. M. Bellovin, S. Floyd, J. Ioannidis, V. Paxson, S. Shenker, "Controlling high bandwidth aggregates in the network," July 2002, ACM SIGCOMM Computer Communication Review, Volume 32 Issue 3
- [19] A. Juels and J. Brainard. "Client puzzles: A cryptographic countermeasure against connection depletion attacks," In Proceedings of the 1999 Networks and distributed system security symposium, March 1999.
- [20] Y. L. Zheng and J. Leiwo. "A Method to Implement a Denial of Service Protection Base," In Information Security and Privacy, volume 1270 of LNCS, pages 90-101, 1997.
- [21] J. Leiwo, P. Nikander, and T. Aura. "Towards network denial of service resistant protocols," In Proceedings of the 15th International Information Security Conference, August 2000.

List Of References

- [22] T. Aura, P. Nikander, and J. Leiwo. DOS-Resistant Authentication with Client Puzzles. Lecture Notes in Computer Science, 2133, 2001.
- [23] Thomas, R.; Mark, B.; Johnson, T.; Croall, J,” NetBouncer: client-legitimacy-based high-performance DDoS filtering,” DARPA Information Survivability Conference and Exposition, 2003. Proceedings, Volume: 1, 22-24 April 2003 Pages:14 - 25 vol.1
- [24] Katabi D. Bazzi I. Xiaowei Yang, “A passive approach for detecting shared bottlenecks” Computer Communications and Networks, 2001. Tenth International Conference on, 15-17 Oct. 2001 Pg: 174 – 181
- [25] C. E. Shannon, “A Mathematical Theory of Communication”, Bell Sys. Tech Journal, 1948
- [26] V. Paxson, “Measurements and Analysis of End-to-end Internet Dynamics”, PhD thesis, Computer Science Division, University of California, Berkeley, April 1997
- [27] R.L. Carter and M.E. Crovella, “Dynamic server selection using dynamic path characterization in wide-area networks”, in Proc. IEEE INFOCOM, vol. 3, Apr. 1997, pp. 1014-1021
- [28] W. Theilmann and K. Rothermel, “Dynamic distance maps of the internet”, in Proc. IEEE INFOCOM, vol. 1, 2000, pp. 275-284.
- [29] K. Park and H. Lee, “On the effectiveness of Probabilistic Packet Marking for IP Traceback under Denial of Service Attack,” 2001, IEEE INFOCOM Conference, June 2001.
- [30] Admin. “Characterizing and Tracing Packet Floods Using Cisco Routers”. [Online] Available:
http://www.seconf.net/firewalls_and_VPN/Characterizing_and_Tracing_Packet_Floods_Using_Cisco_Routers.html, Oct 16, 2002
- [31] J. Glave. “Smurfing cripples ISPs.Wired Technology News.” [Online]. Available: <http://www.wired.com/news/news/technology/story/9506.html>, 1998
- [32] R. Stone, “CenterTrack: An IP overlay network for tracking DoS floods,” in Proc. 2000 USENIX Security Symp., July 2000, pp. 199–212.

List Of References

- [33] H. Burch and B. Cheswick, "Tracing anonymous packets to their approximate source," in Proc. 2000 USENIX LISA Conf., Dec. 2000, pp. 319-327.
- [34] Choi, K.H.; Dai, H.K."A marking scheme using Huffman codes for IP traceback," Parallel Architectures, Algorithms and Networks, 2004. Proceedings. 7th International Symposium on, 10-12 May 2004 Pages:421 - 428
- [35] D. J. Bernstein, SYN Cookies, <http://cr.yp.to/syncookies.html>
- [36] J. Lemon, "Resisting SYN Flood DoS Attacks with a SYN Cache," Proceedings of the BSDCon 2002, Pages: 89 - 97
- [37] J. Postel, "RFC0792: Internet Control Message Protocol," Year of Publication: 1981
- [38] Ratul Mahajan, Sally Floyd and David Wetherall, "Controlling High-Bandwidth Flows at the Congested Router," 9th International Conference on Network Protocols (ICNP) November 2001
- [39] John Ioannidis and Steven M. Bellovin, "Implementing Pushback: router-based defense against DDoS attacks", In Proc. Network and Distributed System Security Symposium (NDSS), San Diego, CA, February 2002
- [40] T. Peng, C. Leckie, and K. Ramamohanarao. "Defending against distributed denial of service attack using selective pushback." In Proceedings of the Ninth IEEE International Conference on Telecommunications (ICT 2002), 2002.
- [41] S. Floyd and V. Jacobon, "Random Early Detection Gateways for Congestion Avoidance," IEEE/ACM Trans. Networking, vol. 1, no. 4, 1993, pp. 397-413.
- [42] D. Lin and R. Morris, "Dynamics of Random Early Detection," in Proceeding of ACM Sigcomm, ACM, New York, 1997, pp. 127-137.
- [43] T. Ott, T. Lakshman, and L. Wong, "SRED: Stabilized RED," in Proceeding of IEEE Infocom, IEEE Press, Piscataway, N.J., 1999, pp 1346-1355.
- [44] P. McKenny, "Stochastic Fairness Queuing," in Proceeding of IEEE Infocom, IEEE Press, Piscataway, N.J., 1990, pp. 733-740.
- [45] A. Manin and K. Ramakrishnan, "Gateway Congestion Control Survey," IETF RFC

List Of References

- (Informational) 1254, Aug. 1991, <http://www.rfc-editor.org/rfc.html>.
- [46] CERT Coordination Center. "Denial of Service Attacks."
http://www.cert.org/tech_tips/denial_of_service.html
- [47] CERT Advisory CA-1996-21. "TCP SYN Flooding and IP Spoofing Attacks."
<http://www.cert.org/advisories/CA-1996-21.html>
- [48] CERT Advisory CA-1998-01. "Smurf IP Denial-of-Service Attacks."
<http://www.cert.org/advisories/CA-1998-01.html>
- [49] IBM Global Services, "Denial-of-service attacks: Understanding network vulnerabilities", White paper (2001)
- [50] David Dittrich, "Distributed Denial of Service (DDoS) Attacks/tools"
<http://staff.washington.edu/dittrich/>
- [51] J. Yan, S. Early, R. Anderson, "The XenoService – A distributed defeat for distributed denial of service", In Proceedings of ISW 2000, October 2000.
- [52] Information Sciences Institute, "Dynabone," <http://www.isi.edu/dynabone/>
- [53] Mazu Networks, "Dynamically Provisioned Monitoring,"
http://www.mazunetworks.com/white_papers/provmon-toc.html
- [54] Arbor Networks, "PeakFlow DoS for Hosting Providers Datasheet,"
http://www.arbornetworks.com/up_media/up_files/PFDoS_ServProv_1.6.pdf
- [55] BBN Technologies, "Applications that participate in their own defense,"
<http://www.bbn.com/infosec/apod.html>
- [56] BBN Technologies, "Intrusion tolerance by unpredictability and adaptation,"
<http://www.bbn.com/infosec/itua.html>
- [57] C. R. Palmer, G. Siganos, M. Faloutsos, C. Faloutsos, and P. B. Gibbons. "The connectivity and fault-tolerance of the Internet topology." The 2001 Workshop on Network-Related Data Management; in cooperation with ACM Special Interest Group on Management of Data / Principles of Database Systems. May 2001.

List Of References

- [58] Paruchuri, V.; Durresi, A.; Kannan, R.; Iyengar, S.S.; “Authenticated autonomous system traceback,” Advanced Information Networking and Applications, 2004. AINA 2004. 18th International Conference on, Volume: 1, 2004 Pages: 406 - 413 Vol.1
- [59] Tian Hai-tao; Huang Liu-sheng; Lei Yun-fei; Chen Guo-liang; “A new scheme for IP traceback under DoS attack,” Parallel and Distributed Computing, Applications and Technologies, 2003. PDCAT' 2003. Proceedings of the Fourth International Conference on , 27-29 Aug. 2003 Pages:189 - 193
- [60] D. McGuire and B. Krebs,” Attack On Internet Called Largest Ever,” <http://www.securityfocus.com/news/1413>, Washington Post, Oct 23 2002
- [61] Belenky, A., Ansari, N. “On IP traceback,” Communications Magazine, IEEE, Volume: 41, Issue: 7, July 2003 Pages: 142 - 153
- [62] S. M. Bellovin, “ICMP Traceback Messages,” IETF draft, 2000; <http://www.research.att.com/~smb/papers/draft-bellovin-itrace-00.txt>
- [63] H. Burch and B. Cheswick, “Tracing Anonymous Packets to Their Approximate Source,” Proc. USENIX LISA, 2000, pp. 319-27.
- [64] Computer Emergency Response Team. CERT Advisory CA-2000-01 “Denial-of-Service Developments.” <http://www.cert.org/advisories/CA-2000-01.html>, Jan. 2000.
- [65] C. Perkins. “IP Mobility Support.” RFC 2002, Oct. 1996.
- [66] TechExams.Net. “Security+ TechNotes - DoS Attacks,” <http://www.techexams.net/technotes/securityplus/attacks-DDOS.shtml>
- [67] CERT, “Advisory CA-1997-28 “IP Denial-of-Service Attacks”, <http://www.cert.org/advisories/CA-1997-28.html>
- [68] Norman, “Attacks against weaknesses in the TCP/IP protocol: About Smurfs and other vermin,” http://www.norman.com/documents/wp_smurf.shtml
- [69] CERT, “Results of the Distributed-Systems Intruder Tools Workshop,” http://www.cert.org/reports/dsit_workshop-final.html

List Of References

- [70] Mankin, A.; Massey, D.; Chien-Lung Wu; Wu, S.F.; Lixia Zhang, "On design and evaluation of intention-driven ICMP traceback" Computer Communications and Networks, 2001. Proceedings. Tenth International Conference on, 15-17 Oct. 2001. Pages: 159 – 165
- [71] David Dittrich, "The DoS Project's trinoo distributed denial of service attack tool," <http://staff.washington.edu/dittrich/>
- [72] David Dittrich, "The "Tribe Flood Network" distributed denial of service attack tool," <http://staff.washington.edu/dittrich/>
- [73] Jason Barlow and Woody Thrower AXENT Security Team, "TFN2K - An Analysis (Revision:1.3)," http://securityresponse.symantec.com/avcenter/security/Content/2000_02_10_a.html
- [74] David Dittrich, "The "stacheldraht" distributed denial of service attack tool," <http://staff.washington.edu/dittrich/>
- [75] Aljifri, H.; "IP traceback: a new denial-of-service deterrent," Security & Privacy Magazine, IEEE , Volume: 1 , Issue: 3 , May-June 2003 Pages:24 – 31
- [76] H. Burch and B. Cheswick, "Tracing Anonymous Packets to Their Approximate Source," Proc. USENIX LISA, 2000, pp. 319-27.
- [77] R. Stone, "Centertrack: An IP Overlay Network for Tracking DoS Floods," Proc. 9th USENIX Sec. Symp, 2000, pp. 199-212.
- [78] P. Ferguson and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks Which Employ IP Source Address Spoofing," RFC 2827, May 2000.
- [79] Rizvi, B. Fernandez-Gaucherand, E. "Analysis of adjusted probabilistic packet marking," IP Operations and Management, 2003, (IPOM 2003), 3rd IEEE Workshop on, 1-3 Oct. 2003 Pages: 9 – 13
- [80] Jenshiub Liu, Zhi-Jian Lee, Yeh-Ching Chun, "Efficient dynamic probabilistic packet marking for IP traceback," Networks, 2003. ICON2003. The 11th IEEE International Conference on 28 Sept.-1 Oct. 2003 Pages: 475 – 480

List Of References

- [81] T. Baba and S. Matsuda, “Tracing Network Attacks to Their Sources”, IEEE Internet Computing March, April 2002, pp 20-26.
- [82] G. Koutepas, F. Stamatelopoulos, B. Maglaris, “[Distributed Management Architecture for Cooperative Detection and Reaction to DDoS Attacks](#)”, March 2004 Journal of Network and Systems Management, Volume 12 Issue 1
- [83] Computer Security Incident Response Team (CSIRT), <http://www.cert.org/csirts/>
- [84] Forum of Incident Response and security team, <http://www.first.org/>
- [85] Trusted Introducer for CSIRTs in Europe, <http://www.ti.terena.nl/>
- [86] CERT Coordination Center (CERT/CC), <http://www.cert.org/>
- [87] David Curry, Herve Debar, “Intrusion Detection Exchange Format”, Internet Drafts, 19/07/2004, group: idwg, <http://www.potaroo.net/ietf/ids-wg-idwg.html>
- [88] Intrusion Detection Message Exchange Format, <http://xml.coverpages.org/idmef.html>
- [89] The Internet Engineering Task Force, “Intrusion Detection Exchange Format (idwg)”, Last Modified: 2004-11-11, <http://www.ietf.org/html.charters/idwg-charter.html>
- [90] Yau, D.K.Y. Lui, J.C.S. Feng Liang, “Defending against distributed denial-of-service attacks with max-min fair server-centric router throttles”, 2002. Tenth IEEE International Workshop on Quality of Service, 2002, page 35- 44
- [91] Tao Peng, Chris Leckie, Ramamohanarao Kotagiri, “Defending Against Denial of Service Attacks and other High Bandwidth Traffic Aggregates,” Ph.D Conversion Report, Jan. 2002
- [92] Denise Pappalardo and Ellen Messmer “[Extortion via DDoS on the rise](#)” Computerworld May 16, 2005
- [93] Kevin Poulsen, “FBI busts alleged DDoS Mafia” SecurityFocus 2004-08-26
- [94] Computer Emergency Response Team Coordination Center (CERT/CC) Statistics, http://www.cert.org/stats/cert_stats.html

List Of References

- [95] CSI/FBI Computer Crime and Security Survey 2004, <http://www.only4gurus.com/v3/preview.asp?ID=6769>
- [96] Dubendorfer, T.; Bossardt, M.; Plattner, B. "Adaptive Distributed Traffic Control Service for DDoS Attack Mitigation" Parallel and Distributed Processing Symposium, 2005. Proceedings. 19th IEEE April 2005
- [97] Christos Douligieris, Aikaterini Mitrokotsa, Security Event "Trust and Confidence in a Fast and Mobile Environment", July 2004
- [98] Dave McCurdy, "Internet Security Alliance", ANSI Annual Conference, Oct 2003
- [99] Cisco IPS 4200 Series Sensors, "Cisco IPS Sensor Software Version 5.0" http://www.cisco.com/en/us/products/hw/vpndevc/ps4077/products_data_sheet0900aecd801e6a45.html
- [100] Cisco Secure IDS/NetRanger: <http://www.cisco.com/en/us/products/index.html>
- [101] Network Sensor (version 7.0) and RealSecure® Guard: <http://www.iss.net/issEn/delivery/prdetail.jsp?type=ISS&oid=19302>
- [102] Wei-Peng Tan, Bu-Sung Lee, Henry C. J. Lee, "Entropy-Minimization Clustering Technique For Probabilistic Packet Marking Scheme," accepted by 12th IEEE International Conference On Networks 16-19 Nov. 2004, Pg 292-298.
- [103] Wei-Peng Tan, Bu-Sung Lee, Henry C. J. Lee, "Legitimacy Investigation And Intention-Based Probabilistic Packet Marking Scheme," accepted by IEEE, 6th Malaysia International Conference on Communications (MICC) in conjunction with 13th IEEE International Conference on Networks (ICON), 17-18 Nov. 2005.
- [104] S. Deering, Internet Protocol, Version 6 (IPv6) RFC 2460, Dec. 1998