Available online at www.sciencedirect.com

ScienceDirect

Computer Communications 30 (2007) 3193–3202

computer communications

www.elsevier.com/locate/comcom

Performance analysis of probabilistic packet marking in IPv6

Xuan-Hien Dang [a,*], Emil Albright [a], Abdullah A. Abonamah [b]

[a] Department of Computer Science, University of Akron, Akron, OH 44325-4003, USA
[b] Institute for Technological Innovation, Zayed University, P.O. Box 500320, Dubai, United Arab Emirates

Available online 29 June 2007

Abstract

Probabilistic packet marking (PPM) has received considerable attention as an IP traceback approach against distributed Denial-of-Service attack, which is one of the most challenging security threat in the Internet. PPM is a technique that seeks to identify the source of such attacks by marking individual packets with portion of the attack path, and then relies on the volume of attack traffic generated to ensure that the whole path can be reconstructed. However, modifying the identification field in the IPv4 packet header to mark packet incurs backward incompatibility for IP fragmented packets. In this paper, we address this issue and analyze the viability of PPM under the next-generation Internet Protocol, IPv6. In doing so, we consider the flaws inherent to IPv4 implementations that limit their backward compatibility, and demonstrate how these shortcomings can be avoided in IPv6. We show that the Flow Label field in the IPv6 datagram header can be safely and effectively overloaded to implement PPM schemes, and present simulation results verifying the applicability and efficiency of this approach.

Keywords: IP traceback; IPv6; Probabilistic packet marking; Security; Performance evaluation

1. Introduction

Denial-of-Service (DoS) attacks constitute a major problem for modern computer networks as they direct massive floods of packets to a victim network. Processing these packets consumes much if not all of the victim's resources such as bandwidth, and thus denies access to legitimate users. To hold the initiators of an attack responsible, they must first be identified. However, the IP design with its stateless packet routing process does not support reliable identification of the originator. The task of tracking and identifying an IP packet origin is known as the IP traceback problem, and to date there exists no clear solution.

A variety of IP traceback techniques have been proposed and assessed [1], each with its own advantages and disadvantages. They can be roughly categorized as either infrastructure logging or end-host storage approaches. The infrastructure logging approach involves a centralized management with logging of packet information as packets travel through routers in the network. The main limitations of this approach are clearly its complexity and significant storage requirements. The end-host storage approach relies on a marking method, which specifies the attack path by having routers store the information required for IP traceback in the packet header. The end-host may then reconstruct the network path by combining information contained in the volume of marked packets received. An example of this technique is the probabilistic packet marking (PPM) scheme. Because of the limited space in the packet header, PPM proposes that routers probabilistically mark IP packets with partial information identifying the marker in the identification (ID) field.

This technique appears promising for several important reasons: small administrative and computational overhead requirements, no extra traffic generated, post-mortem, and compatibility with the existing IP protocol [2]. However, the overloading of the ID field was assumed to negligibly affect packet fragmentation as supported by an empirical

* Corresponding author. Tel.: +1 330 972 7191; fax: +1 330 374 8630.
E-mail addresses: dangx@uakron.edu (X.-H. Dang), emil_albright@yahoo.com (E. Albright), Abdullah.Abonamah@zu.ac.ae (A.A. Abonamah).

0140-3664/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2007.05.060

traffic analysis showing that less than 0.5 % of all packets in the Internet are fragmented [3]. Nevertheless, the fragmentation issue was addressed in [4] by introducing additional memory requirements and processing overhead on the routers.

In this paper, we analyze and discuss the issues in IPv6 implementation of PPM as no ID field exists in IPv6 packets, and also demonstrate how the overloading of the flow label field can be used to overcome the problem. To date, most of the research with PPM has been intended for the Internet Protocol version 4, which is still plagued by many problems [5]. However, IP version 6 was developed to accommodate the explosive growth of the Internet and to gradually replace IPv4. Although IPv6 deployment is still in the early stages, most of the operating systems and network devices (e.g., routers) are already IPv6-capable and many applications have been ported to IPv6. However, some of the limitations of IPv4 are still unresolved. Hence, in this paper, we address one of the security issues of IPv4 in IPv6.

In this paper, we consider PPM traceback techniques, and propose to port one implementation to IPv6. Section 2 presents history and work related to IP traceback techniques. We also consider the compatibility problems arising from IPv4 implementations. In Section 3, we discuss design issues and feasibility of PPM in IPv6 using the Flow Label field. Section 4 presents an implementation of PPM for IPv6 to verify the validity of our proposed approach. An analysis of its performance follows in Section 5. Finally, Section 6 offers conclusions and future work.

2. Background and related work

2.1. IP traceback

IP traceback is a technique aimed at specifying the transmission path, or the routers in the path, of the attack packets with forged sources. The stateless nature of the Internet makes it very difficult to ascertain the origin of an IP packet. As was noted as early as 1985 [6], the IP protocol provides no real means of authentication for packet origins, and thus essentially operates entirely on trust when dealing with inter-network traffic. Since a software user can easily modify the IP header fields, it was observed that solving the IP traceback problem would need to involve hardware. A variety of methodologies have been proposed to this effect and can be divided into two main groups: logging and marking.

2.1.1. Packet logging

Packet logging calls for routers to store identifying information about the IP packets they process to provide a record of through traffic. This is potentially a powerful authentication technique, with applications beyond DoS deterrence, in so far as a single packet can be traced back to its source. However, packet logging requires prohibitive overhead in terms of processing and storage in the face of reasonable traffic volume. Some highly efficient techniques have been proposed such as hash-based IP traceback [7], but even these have functional limits in terms of log retention duration [8] arising from the sheer volume of data involved. While logging shows potential, it has definite limits that can call its practicality into question.

2.1.2. Packet marking

Packet marking calls for some or all routers along an attack path to send packets with some identifying information, either by generating extra ICMP-based edge marking packets [9] or by encoding it directly in the packet header [2,10,11].

With ICMP traceback, traffic passing through nodes would occasionally generate a small ICMP message to be sent along with the data packet to the destination. This packet would identify the packet's path through the node, as well as when it was received. The victim would take all such messages and reconstruct the attack path from them. While ICMP traceback allows for post-attack traceback, it necessarily increases the overall volume of traffic across the network, and also increases the computational burden of routers.

In packet header marking, as the overhead of adding full address path information is excessive, each router generally adds only its own address, or the edge between itself and its downstream neighbor, to the packet. In order to minimize space, there is only one such marking allowed per packet, and each router probabilistically decides whether to overwrite it. A DoS victim reconstructs the attack tree from the marked packets it has received. Instances of header packet marking include PPM, Algebraic-Based Traceback Approach (ATA) [12] and Deterministic Packet Marking (DPM) [1]. Fig. 1 shows a basic overview of the PPM procedure, with a constant marking probability $p = 3\%$ for each router.

2.2. Probabilistic packet marking

2.2.1. Fragment edge marking

PPM is a technique based on the use of edge-sampling, which is to write edge information instead of node informa-

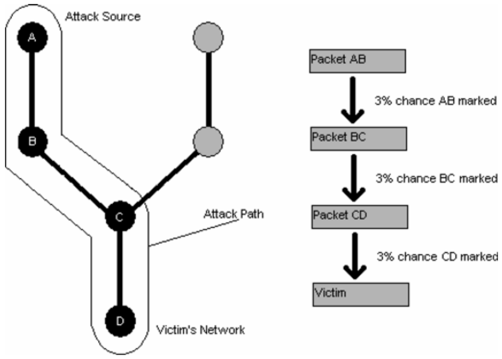

Fig. 1. Probabilistic packet marking.