Efficient Technique for Fast IP Traceback

Byungryong Kim

DongBang Data Technology Co., Ltd.
Department of Computer Science and Engineering, Inha University
doolyn@gmail.com

Abstract. This paper suggests techniques to restrain the convergence time and the combinatorial explosion. IP traceback technique allows a victim to trace the routing path that an attacker has followed to reach his system. It has an effect of deterring future attackers as well as capturing the current one. FMS (Fragment Marking Scheme) is an efficient implementation of IP traceback. Every router participating in FMS leaves its IP information on the passing-through packets, partially and with some probability. The victim, then, can collect the packets and analyze them to reconstruct the attacking path. FMS and similar schemes, however, suffer a long convergence time to build the path when the attack path is lengthy. Also they suffer a combinatorial explosion problem when there are multiple attack paths. The convergence time is reduced considerably by insuring all routers have close-to-equal chance of sending their IP fragments through a distance-weighted sampling technique. The combinatorial explosion is avoided by tagging each IP fragment with the corresponding router's hashed identifier.

1 Introduction

DOS (Denial Of Service) attack typically is performed by sending a large number of packets to the victim system[4,10]. To hide the attacker's location, it often uses a spoofed IP address[3,11]. There are several IP traceback techniques that can trace the attack path in spite of IP spoofing[1,2,9,14]. Most of them, however, requires a heavy traffic or log analysis in the intermediate routers and are impractical without the management support at the routers.

Recently, efficient IP traceback techniques based on IP marking have been suggested[16,19]. In these schemes, the routes in the packet-traveling path mark their IP on the passing-through packets. If all routers write their IP's on all packets, the packet length will be increased out of control. To avoid that, the routers mark their IP's only probabilistically; that is, they mark only if a randomly generated number (between 0 and 1) is less than a certain sampling probability. Since a full IP address is still too large, they split it into a number of fragments (e.g. 8 fragments) and write only one of them. To let the victim know which fragment it is, they include the offset of that particular fragment in the packet. Also, since there could be several routers on the attack path, the routers need to convey information about the distance or hop counts from the victim. For this purpose, a distance field is included in the packet, too. Finally, to facilitate error-checking process, a hashed value of the original IP is included (in fragmented form).

Y. Luo (Ed.): CDVE 2006, LNCS 4101, pp. 211 – 218, 2006.

These solutions are based on an observation that a DOS attack typically involves a large number of packets, and even though the intermediate routers mark their IP's occasionally, the victim can construct the full attack path by collecting all the IP-marked packets. Since the attack path can be constructed automatically without further helps from routers, these solutions are superior to previous IP traceback techniques. However, they still have a number of drawbacks. First the constant sampling probability chosen by FMS, one of the prominent IP marking techniques, tends to penalize distant routers from the victim. The further is the router from the victim, the harder for it to deliver its IP fragments to the victim. Because of this, the time to collect all necessary IP fragments to reconstruct the full attack path (the convergence time, in short) becomes unnecessarily longer. Secondly, they are weak to multiple attack paths[5,7,8]. With multiple attack paths, the victim receives multiple IP fragments with the same offset and distance. Since there are multiple routers at the same distance, the victim has no way of knowing which IP fragment belongs to which router. In FMS, the victim has to try all possible combinations of IP fragments to recover the original IP's.[19] suggests a technique to avoid this, but it requires the victim to maintain a map of upstream routers.

In this paper, we suggest techniques to solve above two problems. The convergence time can be minimized by insuring all routers have equal chance of sending their IP fragments. Since all routers have an equal chance, there are no particular routers that are unduly penalized, and the victim wouldn't have to wait for the slowest router sitting idle. Giving a fair chance is possible through a distance-weighted sampling technique. Each router samples packets based on the values of the distance the packet has traveled so far. Our algorithm encourages the routers to choose short-traveled IP fragments over long-traveled ones for IP marking. This strategy tends to equalize the arrival rates of IP fragments from different routers. We analyze what would be the optimal sampling function that would make all routers have the same chance of IP marking.

The combinatorial explosion is avoided by tagging each IP fragment with the corresponding router's hashed identifier. Since IP fragments are tagged with a particular router's identifier, the victim can easily extract IP fragments belonging to the same router. However, this hashed identifier could collide, and we give an explanation about how often this collision could happen and what can be done when that happens. The rest of the paper is structured as follows. Section 2 surveys related researches. Section 3 explains the distance-weighted sampling technique. Section 4 explains the Tagged Fragment Marking Scheme. Section 5 gives the experimental results. And finally, Section 6 draws a conclusion.

2 Related Researches

Reconstructing an attack path during or after an attack is not an easy problem[6,14]. There are several techniques that are varied in cost and performance. Input Debugging technique[18] utilizes the input debugging feature provided in most routers. Using this feature, a router operator can associate a packet's egress port to the corresponding ingress port. Starting from the closest router, the victim can construct an attack path by identifying the upstream link at each router. This technique, however, requires the