# Distributed Data Aggregation Technology for Real-Time DDoS Attacks Detection

**Yu Chen**
*State University of New York – Binghamton, USA*

**Wei-Shinn Ku**
*Auburn University, USA*

## INTRODUCTION

The information technology has revolutionized almost every facet of our lives. Government, commercial, and educational organizations depend on computers and Internet to such an extent that day-to-day operations are significantly hindered when the networks are "down" (Gordon, Loeb, Lucyshyn & Richardson, 2005). The prosperity of the Internet also attracted abusers and attackers motivated for personal, financial, or even political reasons. What attackers aim at currently is beyond obtaining unauthorized network accesses or stealing private information, there have been attacks on Internet infrastructures (Chakrabarti & Manimaran, 2002; Moore, Voelker & Savage, 2001; Naoumov & Ross, 2006).

*Distributed Denial of Services* (DDoS) attacks is one of such attacks that can lead to enormous destruction, as different infrastructure components of the Internet have implicit trust relationship with each other (Mirkovic & Reiher, 2004; Specht & Lee, 2004). The DDoS attacker often exploits the huge resource asymmetry between the Internet and the victim systems (Chen, Hwang & Ku, 2007; Douligeris & Mitrokosta, 2003).

A comprehensive solution to DDoS attacks requires covering global effects over a wide area of *autonomous system* (AS) domains on the Internet (Mirkovic & Reiher, 2005). Timely detection of the ongoing attacks is the prerequisite of any effective defense scheme (Carl, Kesidis, Brooks & Rai, 2006). It is highly desirable to detect DDoS attacks at very early stage, instead of waiting for the flood to become widespread. It is mandatory for the detection systems to collect real time traffic data from widely deployed traffic monitors and construct the spatiotemporal pattern of anomaly propagation inside the network.

This chapter will introduce a novel distributed real time data aggregation technique named *Change Aggregation Tree* (CAT). The CAT system adopts a hierarchical architecture to simplify the alert correlation and global detection procedures. At intra-domain level, each individual router, which plays the role of traffic monitor, periodically report the local traffic status to the CAT server in the AS. At the inter-domain layer, CAT servers share local detected anomaly patterns with peers located in other ASes, where the potential attack victim is located.

## BACKGROUND

To monitor the traffic fluctuations in a real time manner, network devices often play the role of distributed sensor system that collects local data individually. However, as a large scale distributed system without a central administrator, it is challenging to create a spatiotemporal picture covering wide area cross multiple ISP networks. Unfortunately, such a big picture is essential to detect the anomalies embedded in the traffic flows (Chen, Hwang & Ku, 2007; Papadopoulos, Lindell, Mehringer, Hussain & Govindan, 2003). For this reason, efficient distributed data aggregation techniques have become a hot topic in research community. Due to the limited space, here we only provide a brief survey of reported works which are closely relevant to our work.

A couple of overlay based data aggregation techniques have been proposed to monitor local network traffic and detect anomalies and attacks collaboratively (Feinstein, Schnackenberg, Balupari & Kindred, 2003). In WormShield (Cai, Hwang, Pan & Papadopoulos, 2007), a balanced distributed data aggregation tree (DAT) was proposed, which is capable of collecting and aggregating the fingerprint of Internet worms generated locally. Comparing to the original overlay based data aggregation such as Chord (Stoica, Morris, Karger, Kaashoek & Balakrishnan, 2001), DAT

can compute global fingerprint statistics in a scalable and load-balanced fashion. Several data aggregation systems use advanced statistical algorithms to predict lost values (Zhao, Govindan & Estrin, 2003; Madden, Franklin, Hellerstein & Hong, 2002) and try to reduce the sensitivity of large scale data aggregation networks to the loss of data (Huang, Zhao, Joseph & Kubiatowicz, 2006).

## MAIN FOCUS

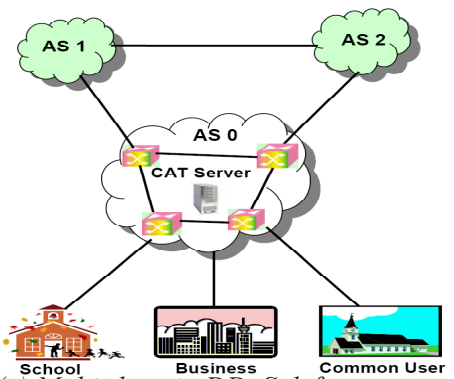
Efficient distributed data aggregation technique is critical to monitor the spatiotemporal fluctuations in Internet traffic status. Chen, Hwang and Ku (2007) have proposed a new distributed aggregation scheme based on change-point detection across multiple network domains. In order to establish an early warning system for DDoS defense across multiple domains, this system adopts a new mechanism called *change aggregation tree* (CAT), which adopts a hierarchical architecture and simplifies the alert correlation and global detection procedures implemented in ISP core networks.

### Distributed Change Point Detection

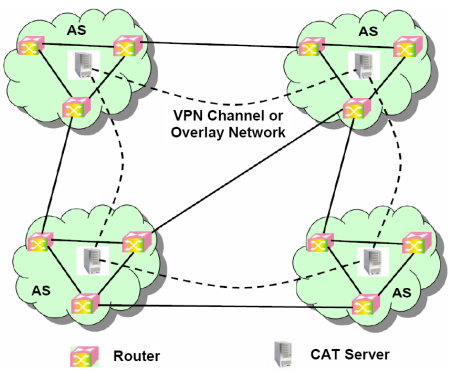
The *Distributed Change-point Detection* (DCD) scheme detects DDoS flooding attacks by monitoring the propagation patterns of abrupt traffic changes at distributed network points. Once a sufficiently large CAT tree is constructed to exceed a preset threshold, an attack is declared. Figure 1 presents the system architecture of the DCD scheme. The system is deployed over multiple AS domains. There is a central CAT server in each domain. The system detects traffic changes, checks flow propagation patterns, aggregates suspicious alerts, and merge CAT subtrees from collaborative servers into a global CAT tree. The root of the global CAT tree is at the victim end. Each tree node corresponds to an *attack-transit routers* (ATR). Each tree edge corresponds to a link between the attack-transit routers.

The DCD system has hierarchical detection architecture. There are three layers in this architecture as shown in Fig. 2. At the lowest layer, individual router functions as a sensor to monitor local traffic fluctuations. Considering the directionality and homing effects in a DDoS flooding attack, routers check how the wavefront changes. A router raises an alert and reports an anomalous traffic pattern to the CAT server. The second layer is at each network domain level. The CAT server constructs a CAT subtree that displays a spatiotemporal pattern of the attack flow in the domain. At the highest layer, the CAT servers at different domains form an overlay network or communicate with each other through *virtual private network* (VPN) channels. All CAT servers send the locally-generated CAT subtrees to

*Figure 1. Distributed change detection of DDoS attacks over multiple AS domains*




*(a) Multi-domain DDoS defense system*




*(b) Inter-domain communication via VPN tunnels or overlay network atop the CAT servers in 4 domains*