A Scalable DDoS Detection Framework with Victim Pinpoint Capability

Haiqin Liu, Yan Sun and Min Sik Kim
School of Electrical Engineering and Computer Science
Washington State University, Pullman, USA
Email: {hliu, ysun, msk}@eecs.wsu.edu

***Abstract*—In recent years, various intrusion detection and prevention systems have been proposed to detect DDoS attacks and mitigate the caused damage. However, many existing IDS systems still keep per-flow state to detect anomaly, and thus do not scale with link speeds in multi-gigabit networks. In this paper, we present a two-level approach for scalable and accurate DDoS attack detection by exploiting the asymmetry in the attack traffic. In the coarse level, we use a modified count-min sketch (MCS) for fast detection, and in the fine level, we propose a bidirectional count sketch (BCS) to achieve better accuracy. At both detection levels, sketch structures are utilized to ensure the scalability of our scheme. The main advantage of our approach is that it can track the victims of attacks without recording every IP address found in the traffic. Such feature is significant for the detection in the high-speed environment. We also propose a SRAM-based parallel architecture to achieve high-speed process. Furthermore, we analyze accuracy estimation issues to provide hints for practical deployment with constraint memory. We finally demonstrate how to extend our original scheme to a collaborative detection framework. Experimental results using the real Internet traffic show that our approach is able to quickly detect anomaly events and track those victims with a high level of accuracy while it can save over 90% key storage compared with previous sketch-based approaches.**

***Index Terms*—intrusion detection, distribute denial of service, victim pinpoint capability, two-level scheme, asymmetry traffic, count min sketch, bidirectional count sketch**

I. Introduction

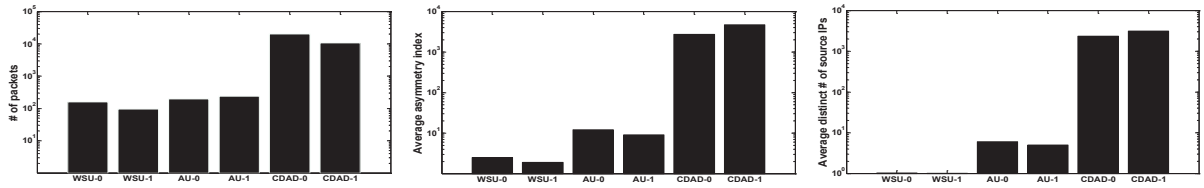
In recent years, distributed denial of service (DDoS) attacks have posed one of the most serious security threats to the Internet [2]. DDoS attacks can do a great damage to the network service by exhausting resources of a server or an ingress network near a server. Since Internet-based services increases every day, the damage that DDoS caused is more severe than before. How to effective defend against such attacks still attract attention from both academia and industry. The very first step for approaching this goal is to effectively detect the attacks, which is difficult because distinguishing attack traffic embedded in a huge amount of background traffic from legitimate traffic is a hard work.

Manuscript received August 15, 2011;accepted October 5, 2011.
This paper is based on "Fine-Grained DDoS Detection Scheme Based on Bidirectional Count Sketch," by H. Liu, Y. Sun, and M. S. Kim, which appeared in the Proceedings of IEEE International Conference on Computer Communication Networks (ICCCN), Hawaii, USA, August 2011.

Over the past decades, many intrusion detection systems (IDSs) have been proposed to fight against DDoS attacks. However, those existing schemes usually present a tradeoff between scalability and accuracy. That is to say, finer grained traffic monitoring can ensure the accuracy of detection while does not scale well. For example, two most popular open-source IDSs, Snort and Bro [3], [4], keep per-flow state to detect anomalies, which makes both of them not scale well in a high-speed network. Since the volume of the Internet traffic doubles every year, how to monitor a large amount of traffic in real-time is becoming more crucial in anomaly detection. Although dimensionality reduction proposed in [5], [6] may be effective in dealing with such a large data, it usually requires complex operations, and thus is impractical in real-time detection.

Recently, a series of sketch-based approaches have been proposed for anomaly detection [7]–[10]. Sketch [11] is a data structure to store a summary of a large data set for space efficiency. Kompella et al. presented Partial Completion Filters (PCF) by utilizing multiple hash tables for scalable attack detection in high-speed networks [7]. As we will see later, such scheme is essentially an simplified version of the original sketch [11]. The main drawback of their scheme is that it can only tell when an attack happens without providing any hint on where the anomaly occurs; the latter is critical in mitigating the attack at an early stage. Identifying victims is also useful in responding to attacks. For instance, an IDS can generate packet classification rules automatically based on the victim information, so as to minimize future damage. In order to provide sketches with the capability that can tell those keys with heavy change, a reversible sketch framework is proposed by [8], [9]. Such feature can be used to provide the victim pinpoint capability in DDoS detection. An improved reversible sketch is proposed by [10] in DDoS detection context. They proposed a flooding attack detection method using a count-min sketch (CMS) with multi-channel nonparametric CUSUM (MNP-CUSUM) [10]. The CUSUM [12] is a change-point detection technique that can accumulate those small offsets during the process to amplify a varying statistical feature so as to improve the detection sensitivity. Although their improved scheme can detect flooding events effectively, it suffers from the following shortcomings which render it still insufficient to detect general DDoS attacks effectively. First of all,

doi:10.4304/jcm.6.9.660-670



(a) Average # of incoming packets per DIP (b) Average AI per DIP (c) Average distinct # of source IPs

Fig. 1. Measurement on Real Traces

their scheme only takes the high frequency of packets in a flow as the evidence of an anomaly event. However, this feature of traffic alone is not enough to detect an anomaly. For example, a Flash Crowd event, which is caused by a large number of legitimate users simultaneously accessing the same server during historical events, can also result in an outburst of the traffic. Their scheme will lead to a large false positive in such a case. Moreover, their scheme suffers from the scalability problem. Because of the key recovery issue in the original sketch scheme they exploited, their method has to record every incoming destination IP (DIP) for the key recovery later, which makes it unscalable to a large amount of traffic due to the huge memory consumption. Finally, it applies a multi-channel CUSUM algorithm to every bucket in the sketch, which requires heavy computations in high speed networks.

In this paper, we propose a two-level approach for DDoS detection. We are motivated by the fact that a typical DDoS attack traffic possesses three characteristics: high frequency of incoming packets, asymmetry in interaction patterns, and high diversity of source IP (SIP) addresses. Our modified count-min sketch (MCS), bidirectional count sketch (BCS), and distinct IP addresses estimator are designed precisely for detecting these three characteristics. Although sketch is also used in our work to achieve high scalability, the differences between the previous sketch-based detection approaches and ours lie in the following ways:

a) Memory consumption: Our scheme outperforms previous works in terms of the space requirement. By utilizing the two-level model, most of benign traffic information, which may be considered as a redundancy for IDS systems, does not need to be recorded in the system. Moreover, the traditional key recovery process in sketch [9] requires sketches to record every input key, which will consume much space, especially when a large number of DIPs are involved in high-speed networks. By taking advantage of the high diversity of source IP addresses feature, our approach can reveal the victim set without recording every destination IP as previous works do.

b) Searching time: Our scheme also can achieve faster detection than previous sketch-based approaches in two aspects. Firstly, since most of traffic is benign, the adopted two-level scheme can greatly reduce the search space while the sketch adopted in [10] processes different kinds of traffic equally. Secondly, rather than applying CUSUM to multiple channels of each bucket in the high frequency anomaly detection phase, we utilize a light-weight exponentially weighted moving average (EWMA) technique to achieve the same goal while introducing much less computing overhead.

c) Accuracy: By taking the asymmetry feature into account, our approach can greatly reduce the false positives by distinguishing between Flash Crowds and DDoS attacks, and thus can improve the overall accuracy.

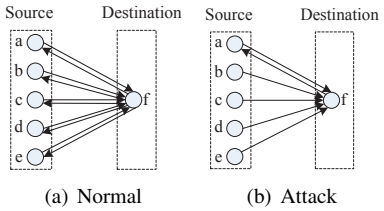
The remainder of this paper is organized as follows. Based on the real Internet measurements, Section II describes the data structures in our scheme and overviews the architecture. Section III presents the design and implementation in detail. Section IV analyzes the accuracy estimation and demonstrates a collaborative detection framework. Experimental results are presented in Section V, and we conclude in Section VI.

II. System Description

A. Measurement on Real Traces

The effectiveness of our scheme is based on the assumption that a typical DDoS attack traffic possesses three characteristics: high frequency of incoming packets, asymmetry in the interaction patterns, and high diversity of source IP addresses.

Fig. 2 demonstrates the "asymmetry" in typical DDoS attacks for web services. We pay attention to the fundamental difference in flow patterns between DDoS attacks and normal traffic. In order to exhaust resources at the server side, an attacker (or more likely a large number of "puppets" in his or her botnet) tries to generate as many requests as possible. When the number of requests exceeds the capacity of the server, we observe fewer responses from the server than requests it receives. We notice that it is not always true that an IP address will serve as both source and destination of traffic, especially when the UDP or ICMP protocol is involved. Thus, in this paper we only refer to the TCP flooding attack by default. Let $N_{\text{forward}}(i)$ denote the number of flows from other nodes to a server i, and $N_{\text{backward}}(i)$ the number of those flows that originate from the server i. By "flow," we mean a group of packets with the same pair of source IP (SIP) and destination IP (DIP) addresses. We do not consider the port information in the flow because we only consider the node-level interactions in our scheme. We define the asymmetry index $AI(i) = |N_{\text{forward}}(i) - N_{\text{backward}}(i)|$



(a) Normal (b) Attack

Fig. 2. Flow patterns of normal traffic and DDoS attacks

for the server i. We expect that the server under a DDoS attack will exhibit a higher AI value than the one experiencing no attack. For example, Fig. 2(a) shows normal interactions between clients and a server, and $AI(f)$ is 0. On the other hand, in an attack scenario depicted in Fig. 2(b), $AI(f)$ is 4.

In order to support the above assumption, we sought to evaluate a set of public traces. We derived Internet traffic from Auckland University (AU) [13] and Washington State University (WSU) as the normal traffic and the attack traffic from "The CAIDA DDoS Attack 2007 Dataset" (CDAD) [14]. The traces are labeled AU-0, AU-1, WSU-0, WSU-1, CDAD-0, CDAD-1, where the numbers with each label are corresponding to different time periods from each trace. The default time interval for the measurements is 10 minutes.

We firstly measure the average number of the incoming packets per DIP in each trace for comparison. Fig. 1(a) shows the results. We notice that in Fig. 1(a) the average numbers of the incoming packets per DIP of CDAD-0 and CDAD-1 are both much larger than that of other traces (nearly 100 times larger). We also measure the average number of flows per DIP and the average traffic volume per DIP over every trace, and we get the similar results that the numbers measured from CDAD are much higher than that of other traces. Thus, rather than using all the three metrics (packet, flow, volume), we only utilize the high frequency feature of incoming packets as the anomaly indicator during the coarse-level detection since these three metrics are highly correlated with each other in our measurement. We notice that some attacks might not possess such correlations. For example, alpha flows [15] usually result in high volume of traffic while only have a small number of flows per DIP or even low-rate DDoS attacks [16] can have low value of all the three metrics. In this paper, we do not aim at the development of a panacea, which is very hard if not impossible, for all kinds of attacks. Instead, we only focus on the detection of the general flooding attack in high speed networks. Other types of attacks need to be further filtered by extra works. Regarding the asymmetry features, we measure the asymmetry index of DIPs of each traces, the result of which is shown in Fig. 1(b). We can see that the AI value obtained from CDAD traces is greatly larger than that of other traces. Similar results can be seen in Fig. 1(c) after we measure the number of distinct IPs that are associated with each DIP in all the traces.

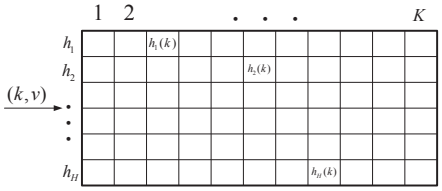


Fig. 3. Illustration of sketch data structure

B. Data Structure

K-ary sketch is a data structure to efficiently and accurately estimate the original signals by aggregating high dimensional data streams into fewer dimensions. As shown in Fig. 3, it consists of H hash tables of size K. A hash function for each row is selected independently and randomly from a set of hash functions. Each data item contains a key k_i and an associated value v_i. When a new item $s_i = (k_i, v_i)$ arrives, its value v_i is added to those buckets corresponding to the key k_i. The $CMS\_Query(key)$ function can return the minimum value among all the buckets corresponding to a specific key. In case of hash collisions, the colliding keys will be listed in the bucket for the key recovery purpose later. The key recovery process [9] can reveal those keys with high frequency in the sketch by looking into the intersection set of high value buckets across the whole sketch. The recovery process is crucial in tracking victims, which will greatly benefit in responding to attacks. Our proposed approach makes two important changes to this original sketch structure: modified count-min sketch (MCS) and bidirectional count sketch (BCS), which will be introduced in the Section II-C.

C. System Architecture

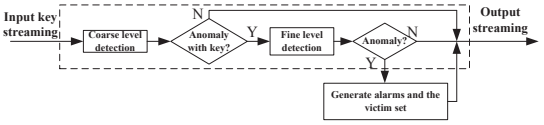


Fig. 4. High-level view of detection process

The overall framework of our detection system is shown in Fig. 4. During each detection period, the related information of every incoming packet is inserted into MCS for the coarse-level detection. We use DIP as the key, and the number of packets that are destined for that IP address as the associated value. MCS only maintains counters for input IP addresses. Compared with the original CMS structure, our MCS structure utilizes space more efficiently because no information on IP addresses themselves is stored in this structure. Besides, unlike CMS, MCS does not rely on CUSUM; MCS is used only for coarse-grained filtering and a light-weight EWMA technique is applied to each bucket to determine whether to generate an alarm for this bucket or not. Whenever an incoming packet satisfies the condition that every bucket it hashed into has an alarm signal, it will trigger the second

stage for finer-grained detection, where a new structure called BCS is used.

In BCS, those suspicious flows detected in the coarse-level detection in both directions are mapped into buckets. While the general sketch structure in which a value in each bucket can increase only, a bucket value in BCS may increase or decrease. We demonstrate how we apply BCS to exploiting the asymmetry of the attack traffic in Section III-B. Once an attack is detected, we use a light-weight distinct IP addresses estimator to pinpoint victims that have the most number of distinct sources, which is a strong indication of DDoS attacks.

III. Scheme Design and Implementation

In this section, we describe the detection processes of both coarse and fine level in details and then explain how the proposed distinct sources estimator works and why it can help to indicate DDoS attacks. Finally, a SRAM-based parallel architecture is proposed to achieve high-speed process.

A. Coarse-level Detection

In our scheme, each bucket in the MCS contains four values $(v_t, v_{t-\Delta t}, v_{\text{backup}}, Flag)$. v_t is the number of packets that are accumulated from $t-\Delta t$ to t, $v_{t-\Delta t}$ is the previous value of v_t, and v_{backup} is the value of v_t right before the alarm occurs (or null if there has been no alarm). $Flag$ is set to 1 whenever the alarm condition is satisfied; otherwise it is set to 0. Here, the definition of alarm conditions depends on the practical deployment, and we will further explain it when we describe Algorithm 1 below. For each incoming record, we update the sketch with $(k_i, 1)$ where k_i is the DIP and 1 represents the number of this incoming record. In our MCS, rather than returning the minimum value of v_t as the original sketch does, the CMS_Query function in MCS returns the minimum value of $Flag$ among all the buckets corresponding to a specific DIP to indicate whether it is under flood attacks. As we can see, the sketch adopted here only requires $O(H \times K)$ cells, which is constant.

The main purpose of MCS is to detect items with abnormal frequency at the coarse level. It is the first stage in the system, which every packet must go through. When a new packet arrives, hash values of H hash functions are computed, and the corresponding buckets are updated; the value in each bucket is incremented by 1. This accumulation process repeats every Δt seconds. The alarm condition is tested for all $H \times K$ buckets periodically. If the alarm condition is satisfied, then the alarm flag associated with the bucket is set to 1. Whenever there is an alarm, the previous v value of the bucket is recorded in the v_{backup} for determining whether the raised alarm is terminated or not.

We use an EWMA technique to decide whether there is an anomaly in each bucket, as shown in Algorithm 1. For each bucket, if the bucket status is normal, then we estimate v_t with an EWMA parameter α. Whenever $v_t \geq (1+\theta)\overline{v_{t-\Delta t}}$, which is considered as the satisfaction of the alarm condition, an alarm is raised. θ is the parameter that represents the percentage above the estimated value that can be considered to be an indication of anomalous pattern. The procedure is different after an alarm was raised. In order to estimate when the generated alarm should be terminated, we need to compare the current value with the specific value right before the time that the alarm happened. Such specific value is recorded in v_{backup} before the alarm is generated. Also, rather than using the previous value $\overline{v_{t-\Delta t}}$, we estimate the $\overline{v_t}$ by v_{backup} in order to eliminate the impact of the anomaly on the next following $\overline{v_t}$ series.

Algorithm 1: Adjustment procedure of $Flag$

```
for k = 1, h = 1 to K, H do
    if Flag = 0 then
        v̄_t ← (1 − α)v̄_{t−Δt} + αv_t ;
        if v_t ≥ (1 + θ)v̄_{t−Δt} then
            Flag ← 1 ;
            v_backup ← v̄_{t−Δt} ;
        end
    else
        v̄_t ← (1 − α)v_backup + αv_t ;
        if v_t < (1 + θ)v_backup then
            Flag ← 0 ;
        end
    end
end
```

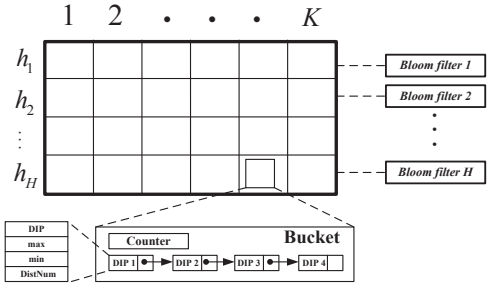


Fig. 5. Illustration of BCS data structure

We can do the coarse-level detection by querying the minimal value of alarm flag for a specific key. If $CMS\_Query(key) = 1$, then there may be an anomaly associated with the key. However, the coarse-level detection would yield a certain number of false positives. There are two possible reasons for false positives. The first possibility is hash collisions, which can be reduced by carefully selecting hash functions or enlarging the size of the sketch. The second possibility is flash crowds. They can also yield many items with high frequencies in the sketch. Thus, we need to examine traffic further to separate these possibilities from true attacks, which is the goal of our next technique, BCS, which detects anomalies at the finer level.

B. Fine-level Detection

The objective of the fine-level detection is to find out those DIPs exhibiting high asymmetric communication

Algorithm 2: BCS update procedure in the forward direction

```
for h = 1 to H do
    k ← BCS[h].hash(DIP) ;
    if DIP is not in BCS[h][k].list then
        insert DIP into BCS[h][k].list ;
        update BCS[h].BF by DIP|SIP ;
        BCS[h][k].counter ++ ;
    else
        if DIP|SIP is not in BCS[h].BF then
            update BCS[h].BF by DIP|SIP ;
            BCS[h][k].counter ++ ;
        end
    end
    DistinctSourcesEstimator(SIP) ;
end
```

Algorithm 3: BCS update procedure of the backward direction

```
for h = 1 to H do
    k ← BCS[h].hash(SIP) ;
    if SIP is in BCS[h][k].list then
        if SIP|DIP is in BCS[h].BF then
            BCS[h][k].counter − − ;
        end
    end
end
```

patterns. A successful DDoS attack employs a large number of zombies to exhaust resources of the target side. However, they are usually unaware of the exact capacity of the server. Therefore, to guarantee to overwhelm the server, an attacker sends as much traffic as allowed, exceeding the server's capacity. This results in highly asymmetric communication patterns between clients and the server, as shown in Fig. 2(b). There are two reasons for such an asymmetry pattern. First, the capacity of the server, namely the victim, to respond is limited while the attacker can keep launching new connections. Second, the SIP addresses of attack traffic are forged, and thus the server has to abort communications. During some time interval ΔT, for a specific IP address, we can detect whether this address serves as both source and destination or not. If yes, then the corresponding flow (the source and destination IP pair containing this IP address) can be considered as a normal flow. Motivated by this observation, we propose the BCS structure to monitor such kind of anomaly with fine granularity.

During one time interval, we use DIP as the key in updating the BCS structure. An illustration of BCS sketch is shown in Fig. 5. All the keys with hash collisions will be stored as a list in the corresponding bucket. Rather than incrementing corresponding H counters by 1 every time a new packet arrives, the counters increase only when the DIP belongs to a new flow. For example, a flow (s_i, d_i), where s_i denotes the SIP address of node i and d_i is the DIP address of node i, will contribute to the corresponding H buckets only once during a single period. On the other hand, (s_j, d_i), which is another flow with the same destination d_i, will contribute another 1 to the buckets that the key d_i is hashed into. Since we do not need to record SIP addresses in the sketch, we employ H bloom filters (BF) with m bits and k_{bf} hash functions as an ancillary structure to estimate whether a specific flow that new packets belong to has been inserted to the BCS structure or not. Algorithm 2 presents the details of how BCS works on the forward direction of traffic. We use $|$ as the string concatenation operator.

The procedure for the backward traffic is shown in Algorithm 3. Whenever we find a backward flow that can be paired with an existing forward flow in the BCS structure, the corresponding counter decreases. In this way, the counters with anomalous high values indicate an anomaly event caused by asymmetric communication patterns for a specific victim.

C. Distinct Sources Estimator

In order to avoid being detected, attackers may employ a large number of SIP addresses. In such cases, those DIPs that are associated with the largest distinct SIP addresses should be a good candidate for a victim under attack. Thus, how to find the number of distinct SIPs for a victim is crucial in the DDoS defense. Without recording the SIP addresses in the system, which requires too much memory, we need to find a way to estimate this number. For each DIP that is hashed into BCS, we pick a hash function $h\colon \mathcal{N} \rightarrow [0,1]$ which maps every number into $[0,1]$, and then we apply $h(\cdot)$ to all the SIP addresses that are associated with this DIP, and maintains the maximal value max and minimal value min, and then the number of distinct IP addresses, $DistNum$, can be estimated as $\frac{1}{2}\cdot\left(\frac{1}{min}+\frac{1}{1-max}\right)$. If the hash function that we choose is sufficiently random, then the above formula is a sufficiently good estimator for our purpose. In this way, each DIP which has been hashed into the BCS will be associated with a number: $DistNum$. For a specific DIP, this number $DistNum$ can be used as an indicator on how diverse the corresponding SIPs are.

D. Victims Identification

At the end of each time interval, for each row in the BCS, we compute the average counter value $\overline{C[h]}$ and the corresponding mean square deviation $D[h]$. For a specific bucket, whenever its counter value $BCS[h][k].counter$ satisfies the following condition, then it raises an alarm for an anomaly:

$$BCS[h][k].counter - \overline{C[h]} \geq \beta \cdot D[h] \qquad (1)$$

where β is an adjustment factor that should be empirically determined. Then, we merge those DIPs that correspond to those anomalous buckets together, and sort them by their $DistNum$. In addition, we eliminate those DIPs that satisfy the condition $BCS\_Query(DIP) < TH_{counter}$ in the merged set, where BCS_Query is similar to the original CMS_Query, which returns the minimum counter value through all the hashed buckets in the sketch

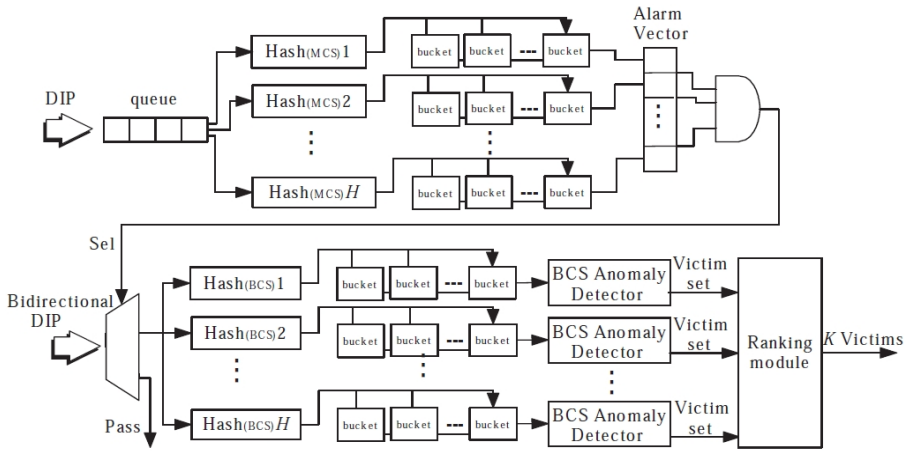


Fig. 6. Hardware architecture for the proposed scheme

BCS. $TH_{counter}$ is a threshold which can be empirically determined. Finally, those victims can be chosen from the merged DIP set by picking the top few DIPs with the largest $DistNum$ value. Or, we can set a threshold $TH_{DistNum}$ to select those victims that can satisfy $DistNum \geq TH_{DistNum}$ to process the selection of victims.

E. Hardware Architecture

Our proposed scheme can be implemented by hardware to achieve high speed process. Since the field programmable gate array (FPGA) technology has widely been utilized for real-time packet processing due to its capability of reconfigure and parallelism, we propose a SRAM-based parallel architecture as shown in Fig. 6. For each input DIP of incoming packets, we perform the hash computations over the H_{MCS} branches in parallel. The Carter-Wegman H3 hash function [17] can be utilized in our hardware-based scheme, since the H3 hash function mostly consists of XOR gates proportional to the number of output bits, which can make it easily implemented in hardware. The bit values in the vector are initially set to 0 and they will be periodically reset to 0 at the end of the detection interval. For each row at the first stage, whenever an alarm signal is generated, the corresponding bit in the vector will be set to 1. After doing the AND operation among all the bit values, it can decide whether to trigger the second stage detection or not. Similarly, the H_{BCS} branches during the finer level detection phase can also be executed in parallel. The BCS anomaly detector and ranking module can be implemented by FPGAs according to the flow logic discussed in the previous sections. The overall search process can be divided into several independent parts and it can be pipelined by assigning each part to a separate memory block to accelerate the overall processing speed. For example, the hash computation of the current incoming DIP and the anomaly detection of the previous DIP are independent with each other and thus can be mapped into two different stages.

IV. Analysis and Discussion

In this section, we firstly analyze the space requirement and then estimate the accuracy of our scheme. We further demonstrate how to extend the current scheme to a collaborative detection framework.

A. Space Requirements

The primary space consumption of the system is due to the two sketches (MCS and BCS) and H bloom filters that are employed at the finer detection stage. Let L_{mc} denote the length (in bytes) of each bucket in the sketch MCS and L_{bc} represent the same in the sketch BCS. Moreover, each bloom filter will occupy L_{bf} space. Suppose the MCS and BCS have the size $H_{mc} \cdot K_{mc}$ and $H_{bc} \cdot K_{bc}$, respectively, the total memory requirement will be:

$$H_{mc} \cdot K_{mc} \cdot L_{mc} + H_{bc} \cdot K_{bc} \cdot L_{bc} + H_{bc} \cdot L_{bf} \quad (2)$$

According to our method proposed above, the length L_{bc} will not be a constant value because the length of the linked list varies. In the practical deployment, we are able to limit the maximal number of the nodes in the link list. For example, for a specific link list, we can only keep those top few DIPs that are associated with the highest $DistNum$ value in the list.

B. Accuracy Estimation Analysis

We further sought to quantify the impact of the size of sketches to the overall accuracy of our framework. We estimate the accuracy of our system in terms of "false positive rate" (FPR) and "false negative rate" (FNR). In order to simplify the problem, we conduct our analysis on an assumption that the "false negative rate" of both MCS FNR_{mc} and BCS FNR_{bc} are negligible and we will demonstrate why this assumption holds in our scheme below. At the presence of certain number of malicious flows, the overall "FPR" depend on the accuracy performance of individual modules (MCS and BCS). Thus, we firstly define false positive rate of MCS FPR_{mc} and BCS FPR_{bc}, and then demonstrate how they contribute to the overall false positive rate.

Since MCS and BCS are both proposed based on sketch, we firstly conduct a general analysis of the sketch structure. Let us assume there are m different malicious keys and n buckets in each row of a sketch. Since the probability that a specific bucket is not hashed by a malicious key is $1-\frac{1}{n}$, the probability that a specific bucket is not hashed by every malicious key is $\left(1-\frac{1}{n}\right)^m$. Therefore, the probability that a bucket in a row is hashed by at least one malicious key is $1-\left(1-\frac{1}{n}\right)^m$ and the expectation of the number of buckets to which these m malicious keys hash is $n\left(1-\left(1-\frac{1}{n}\right)^m\right)$. When m is much less than n, we have:

$$\begin{aligned} n\left(1-\left(1-\frac{1}{n}\right)^m\right) &= n\left(1-\left(1-\frac{1}{n}\right)^{-n\cdot-\frac{m}{n}}\right) \\ &\approx n\left(1-e^{-\frac{m}{n}}\right) \text{ when } n >> 1 \\ &= m\cdot\frac{1-e^{-\frac{m}{n}}}{\frac{m}{n}} \\ &= m\cdot\frac{1-[1+\left(-\frac{m}{n}\right)+\frac{\left(-\frac{m}{n}\right)^2}{2!}+\cdots]}{\frac{m}{n}} \\ &\approx m \quad (when\ n >> m \geq 1) \end{aligned} \tag{3}$$

We define those buckets that are hashed by malicious keys as malicious buckets. Therefore, when $n >> m$, the number of the malicious keys can be used to estimate the expected number of malicious buckets in a row. For a key that is hashed into a malicious bucket in each row of sketch, whether it is benign or not, our scheme will judge it as a malicious key, which is the main cause of false positives of sketch scheme. We assume that there are totally N distinct incoming keys, which contains $N\cdot P_{normal}$ normal keys, $N\cdot P_{flashCrowd}$ keys associated with "Flash Crowd" events and $N\cdot P_{DDoS}$ keys with DDoS events. P_{normal}, $P_{flashCrowd}$ and P_{DDoS} are the proportion of normal keys, keys with "Flash Crowd" and keys with DDoS to the total number of different keys, respectively. Based on the above definition, we have $P_{normal}+P_{flashCrowd}+P_{DDoS}=1$. For the analysis, we call those keys that are related to "Flash Crowd" events as "Flashcrowd" keys and keys associated with DDoS events as DDoS keys.

In the coarse level detection, both "Flashcrowd" and DDoS keys are considered to be positive (malicious) instances. Therefore, based on the Eq. 3, the probability that a key is hashed into one of these malicious buckets in one row is given by $\frac{N\cdot(1-P_{normal})}{K_{mc}}$. For H_{mc} rows, the probability is:

$$FPR_{mc}=\left(\frac{N\cdot(1-P_{normal})}{K_{mc}}\right)^{H_{mc}} \tag{4}$$

For the fine level detection, only DDoS keys are considered to be positive (malicious) instances. Similarly, the probability that a key is hashed into one of these malicious buckets for each row in BCS is given by:

$$FPR_{bc}=\left(\frac{N\cdot P_{DDoS}}{K_{bc}}\right)^{H_{bc}} \tag{5}$$

We define the overall false positive rate by the definition:

$$FPR_{overall}=\frac{Total\ \#\ of\ false\ positive\ instances}{Total\ \#\ of\ negative\ instances} \tag{6}$$

For the overall scheme, negative instances contains normal keys and "Flashcrowd" keys and false positives mean those negative instances that are wrongly judged as DDoS keys. Thus, we have:

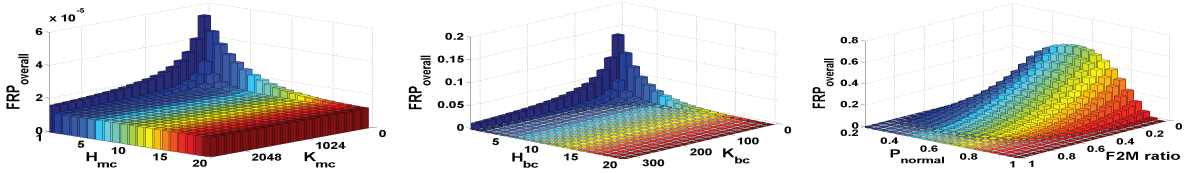
$$\begin{aligned} &FPR_{overall} \\ &= \frac{Total\ \#\ of\ false\ positive\ instances}{Total\ \#\ of\ negative\ instances} \\ &= \frac{NP_{normal}FPR_{mc}FPR_{bc}+NP_{flashCrowd}FPR_{bc}}{N\left(P_{flashCrowd}+P_{normal}\right)} \\ &= \frac{P_{normal}FPR_{mc}FPR_{bc}+P_{flashCrowd}FPR_{bc}}{P_{flashCrowd}+P_{normal}} \end{aligned} \tag{7}$$

From the Eq. 7, we can see that the overall false positive rate depends on the distribution of traffic and false positive rate of each individual module. We can estimate the overall false positive rate $FPR_{overall}$ by Eq. 7. For example, suppose the total number of distinct DIPs is 1000, and there are 180 "Flashcrowd" keys and 20 "DDoS" keys. Let us assume that $K_{mc}=1024$, $H_{mc}=10$ for MCS and $K_{bc}=128$, $H_{bc}=5$ for BCS. According to the Eq. 4 and Eq. 5, we have $FPR_{mc}=\left((180+20)/1024\right)^{10}\approx 8.08\times10^{-8}$ and $FPR_{bc}=(20/128)^5\approx 9.31\times10^{-5}$. Therefore, $FPR_{overall}$ can be estimated as:$(0.8\times8.08\times10^{-8}\times9.31\times10^{-5}+0.18\times9.31\times10^{-5})/(0.18+0.8)\approx 1.72\times10^{-5}$.

In order to further demonstrate the impact of various factors on $FPR_{overall}$, we vary different parameters and draw figures according to Eq. 7 as shown in Fig. 7. The default settings for the parameters we adopted are $H_{mc}=10$, $K_{mc}=1024$ for the sketch size of MCS, $H_{bc}=5$, $K_{bc}=128$ for the size of BCS and $P_{normal}=0.8$, $P_{flashCrowd}=0.18$ for the traffic distribution.

Fig. 7(a) and Fig. 7(b) show the impact of sizes of MCS and BCS on the overall false positive rate, respectively. From both of these two figures, we can see that by enlarging the size of sketches (either H or K), we can greatly reduce the overall false positives. Although keeping the size of sketches large will benefit the accuracy performance, it will also consume much memory space, which will be unaffordable for practical deployment. In practice, we should carefully design the size of sketches employed in the scheme based on space requirements. In Fig. 7(c), we show the impact of traffic distribution on $FPR_{overall}$, where F2M is defined as the ratio of the number of "Flashcrowd" keys to malicious keys (including "Flashcrowd" keys and DDoS keys). We can see that the $FPR_{overall}$ decreases as the proportion of "Flashcrowd" keys increases when we fix P_{normal}. This is because the probability that a normal key is hashed into a malicious bucket decreases as the number of malicious buckets diminishes.

Regarding the false negative rate, we first consider the false negative rate of MCS FNR_{mc} and BCS FNR_{bc}.



(a) Size of MCS VS. $FPR_{overall}$ (b) Size of BCS VS. $FPR_{overall}$ (c) Traffic distribution VS. $FPR_{overall}$

Fig. 7. The relation between various factors and the estimated overall false positive rate

Since we define those buckets that are hashed by the malicious keys as malicious buckets and we consider those keys that are hashed to malicious buckets in every row of sketches as malicious keys, there will be no false negative ideally. In our MCS stage, it is possible that some false negatives can be caused by the improved EWMA technique. For example, some of buckets, which should be considered as malicious buckets, still do not generate the corresponding alarm signal for detection. However, this case is much less often, because the malicious keys will always have much higher incoming frequency than the normal keys. Similar things happen at the BCS stage. Although there might be some false negatives in BCS due to the inherent false positive issue of bloom filters employed, such case rarely happens for the much lower false positive rate of bloom filters compared with sketches. From the perspective of the overall framework, positive instances only consist of DDoS keys and false negatives are those DDoS keys which are classified as normal or "Flashcrowd" keys by mistake. Thus, according to the definition of false negative rate, we have:

$$\begin{aligned} & FNR_{overall} \\ & = \frac{Total\ \#\ of\ false\ negative\ instances}{Total\ \#\ of\ positive\ instances} \\ & = \frac{NP_{DDoS}FNR_{mc} + NP_{DDoS}TPR_{mc}FNR_{bc}}{NP_{DDoS}} \\ & = FNR_{mc} + TPR_{mc}FNR_{bc} \\ & = FNR_{mc} + FNR_{bc} \quad (Since\ TPR_{mc} = 1) \end{aligned} \tag{8}$$

Where TPR_{mc} is the true positive rate of MCS. The true positive rate for MCS is approximately equal to 1 based on the definition of the positive instances in MCS. Since both FNR_{mc} and FNR_{bc} are negligible, the overall false negative rate also can be neglected. Moreover, from Eq. 8, we can see that the false negative rate is irrelevant to the distribution of the incoming traffic.

C. Collaborative Detection Scheme

Till now, our proposed two-level framework can be categorized as a host-based system, which can be deployed at an ingress router near the victim side. The nearer the detection module from victims is, the larger amount of attack traffic we can observe. Thus, in order to reduce the difficulty of detection, one possible solution is to deploy the proposed detection module at the targeted server. However, this preliminary solution is a bad idea for two reasons. First, one deployment can only protect one victim which render it not scale well. Secondly, it cannot even well protect the victim it supposed to protect. Because the ingress bandwidth resources near the victim server can be exhausted as well by the attack traffic, which will result in the same effect to the legitimate users since they cannot visit the victim server. Thus, a deployment that is a little far from a victim server might be a good choice.

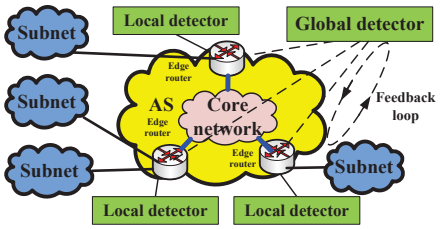


Fig. 8. Illustration of a collaborative framework

However, a single host-based system is inherently not robust enough no matter where it is deployed. It is entirely possible that some unaware or intentional internet behaviors can damage its effectiveness. For instance, due to network device failure problems or a specific routing protocol designed for congestion avoidance, a backward traffic associated with an original forward flow might be routed by a totally different path. As a result, the traffic asymmetry feature no longer can be observed by a single router. Furthermore, such scheme can be easily fooled by a sophisticate attacker, which can be considered as an intentional internet event. Since attackers always employ a large number of zombie machines around the world to launch attacks, traffic that comes from every corner of the world can be routed by different edge routers inside an AS. Thus, if we only take a single router into accounts, the volume of attack traffic might not be aggregated at a detectable level for a detection module while the final gather of attack traffic will still cause severe damage to victim servers. Therefore, a collaborative detection approach which can comprehensively consider the global circumstance will be an attractive solution. Fortunately, our proposed approach can be easily extended to a collaborative detection scheme, which will greatly reinforce our original work. Fig. 8 illustrates the overall collaborative framework. The edge routers are responsible for connecting subnets (it can be customer networks or other ASes) with the core network. Our collaborative detection framework contains multiple local detectors, one

global detector and a feedback loop between them. The functionality of each component is described as below.

a) Local Detector: A local detector can be deployed at an edge router, and it is responsible for:

- Summarizing traffic statistics from partial or all packets from both of two directional links
- Report the summarized traffic statistics to the global detector periodically
- Receive feedback instructions from the global detector and adjust the local information collection manner based on the feedback instructions
- Timely react to those DDoS events that can be detected at the local side

To be specific, a local detector maintains two main threads. The first thread is called "update thread", which keeps scanning every incoming packet and updates the traffic profile. The second thread, which we called as "report thread", periodically sends the built traffic profile to the global detector and finally refreshes the profile after reporting. A hash table with linked lists can be utilized for the traffic profile building. Each entry in the profile hash table contains six values $(DIP, num, suspFlag, min, max, sipPtr)$ with DIP as key value. num accumulates the number of those incoming packets associated with DIP during one period. $suspFlag$, which is set based on the feedback instructions from the global detector, can be used to decide whether to update the remaining values in one entry or not. Whenever a DIP is suspected by the global detector due to its high packet frequency, the $suspFlag$ will be set to 1. When $suspFlag = 1$, the "update thread" will keep updating the following values $(min, max, sipPtr)$ in an entry. min and max maintain the minimal and maximal of hash value by mapping all SIPs associated with the DIP into range $(0, 1)$ as we did in the distinct sources estimator. At the same time, $sipPtr$, which is a head pointer of a linked list, will be updated by inserting those SIPs into the list. As we can see, by reporting the built traffic profile, the global detector can obtain all the necessary information for further anomaly analysis.

b) Global Detector: The responsibility of a global detector contains:

- Receive those statistics reports from local detectors
- Perform anomaly detection based on packet frequency at coarse-level detection phase
- Perform anomaly detection based on both the distinct number of SIPs and asymmetry feature associate with each DIP at fine-level detection phase
- Send feedback instructions to local detectors based on anomaly detection results

The global detector also maintains two threads. The first thread, which we called as "MCS thread", is responsible for updating MCS and sending feedback to local detectors. The MCS update process is similar as we described in Section III. The total incoming frequency associated with a DIP can be obtained by: $\sum_{k=1}^{M} num_k$, $(1 \leq k \leq M)$, where M is the total number of local detectors that report their local frequency num of this DIP to the global detector. When a key is detected as suspicious key with high packet frequency during MCS detection phase in the global detector, those hash entries associated with this key at local detectors will be marked as suspicious by setting $suspFlag$ to 1. We called the second thread as "BCS thread". The "BCS thread" also does similar works as we have demonstrated in Section III. The min and max value associated with certain DIP in BCS can be obtained by: $min = MIN\,(min_1, min_2, \cdots, min_M)$ and $max = MAX\,(max_1, max_2, \cdots, max_M)$. As we can see, our original scheme can be extended in a distributed-executing way quite smoothly. Besides those advantages we pointed out before, one great benefit by running in a distributed way is that the workload of the central global detector can be largely reduced. As a result, the scalability performance can be further improved.

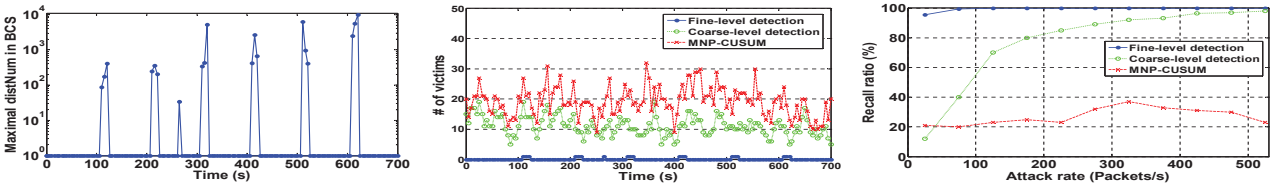
We notice that the number of one packet will be counted twice in a typical AS infrastructure. One count is at the ingress router and the other one is at the egress router. Similar thing happens when we measure the count for traffic asymmetry. However, it will not impact the overall performance, because both malicious and benign traffic will be amplified by the same proportion when we measure the frequency feature. Regarding the asymmetry feature, both forward and backward traffic will be counted twice, the effect of which will be offset to each other when we measure the asymmetry feature in BCS.

V. Evaluation

We evaluate the performance of the proposed scheme via simulations. We use the trace data from AU [13] as the background traffic. It contains packet traces captured from the link connecting Auckland University and the Internet. This background traffic, which contains both forward and reverse directions, has an average rate of 523 packets per second. We consider the accuracy of victim identification and the amount of memory consumption as two main performance metrics. Unless otherwise noted, the default settings for the parameters we adopted in our experiment are $\Delta t = 5s$ for the periodical sketch construction, $H_{mc} = 32$, $K_{mc} = 1024$ for the sketch size of MCS, $H_{bc} = 5$, $K_{bc} = 128$ for the size of BCS, $L_{bf} = 10000Bytes$ for bloom filters, $\alpha = 0.4$, $\theta = 0.5$ for the coarse level detection, $\Delta T = 5s$, $\beta = 2$ and $TH_{counter} = 10$ for the fine-level detection.

A. Detection Accuracy Evaluation

We generate the flooding traffic using attack tools we developed. The attack rates vary from 25 to 500 packets per second (25, 50, 75, 100, 200, and 500) and the duration of each the attack is 20 seconds. Those attacks are injected at the offset of every 100 seconds. Our goal is to try to gauge the detection sensitivity of our scheme under a large range of attack rates. Fig. 9(a) shows the maximal $DistNum$ series among all the detected victims in the sketch BCS as the time goes. The six spikes

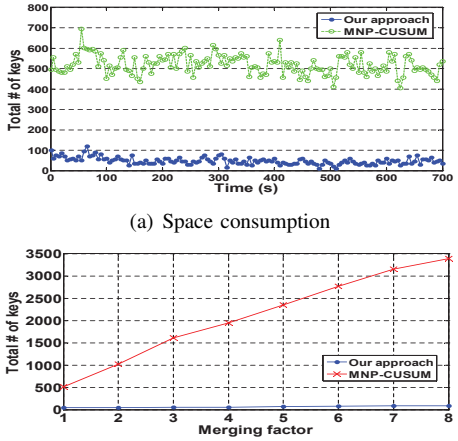


(a) Maximal DistNum value in BCS (b) # of detected victims (c) Recall ratio

Fig. 9. Evaluation of accuracy

(excluding the smallest one) indicate all six DDoS attacks we injected. Even when the attack rate is as low as 25 packets per second, which happens at the offset of 100 seconds, our scheme is still able to identify such low rate attacks while maintaining high accuracy. The maximal $DistNum$ values well reflect the rates of corresponding attacks. After we manually inspected the background traffic, we found that the remaining spike with the lowest value in Fig. 9(a) represents a low rate flooding attack in the original trace. Fig. 9(b) demonstrates the number of victims that are identified by the coarse-level and fine-level detection. On average, the coarse-level detection identify 12 victims per interval. All of those victims experience high rates of requests, which may be caused by flash crowds or DDoS attacks. However, after we further filter those potential victims using the fine-level detection, at most one victim per interval remains, which is the actual attack contained in the traffic. Moreover, the average victim number detected by the MNP-CUSUM approach [10] is around 21, which is even higher than the coarse-level detection of our approach. This is because the original CUSUM technique does not take care of the quick termination after the alarm happens, which results in that too many buckets in the sketch remains high value for a long time. Therefore, it usually causes many false positives. After we modified the original CUSUM techniques by a method for quickly terminating alarm as proposed in [19], the average number is significantly reduced to around 12, which can be due to flash crowds.

We also measure the recall ratio under different attack rates. A recall ratio is the fraction of the true victims in the estimated victims returned by our scheme. The estimated victims identified by the coarse-level detection is the set of all DIPs which satisfy $CMS\_Query(DIP) = 1$. In Fig. 9(c), we can see that the recall ratio of the fine-level detection is very stable; nearly 100% of victims are accurately identified. Even when the attack rate is as low as 25 packets per second, the recall ratio is still over 95%. However, with the coarse-level detection only, the ratio is much lower. It requires more than 350 packets per second (about 66% of the background traffic rate) to achieve the ratio over 95%. Again, due to the alarm termination problem, the MNP-CUSUM technique performs poorly here. Its recall ratio is around 23% on average.



(a) Space consumption

(b) Storage scalability

Fig. 10. Evaluation of space consumption

B. Space Consumption

We also sought to measure the memory consumption. Basically, the overall space consumption of sketch-based approaches consists of two different parts. The first part, which can be attributable to the sketch structure itself, takes constant size of small space while the other part, which serves for assisting functions such as the key storage, occupies dynamic size. Since the scalability performance of sketch-based approaches greatly depends on the dynamic part, we compare our approach against [10] by measuring the number of keys that should be stored. The results are shown in Fig. 10(a). During one interval, there are 47 keys that are needed to be stored in our scheme on average while the average number of the keys of MNP-CUSUM approach is around 519. Our approach can save up to 90% keys, which translates to less memory consumption and searching space, when comparing with the previous approach. In order to evaluate the storage scalability, we shift time stamps of different periods of traces from AU and then merge them together in order to enhance the traffic intensity. We define "Merging factor" as the number of different periods, which can also reflect the intensity of the traffic. Then, we measure the required key storage over various approaches as shown in Fig. 10(b). Our method nearly keeps constant number of keys when the merging factor increases, while the MNP-CUSUM holds a linear-like trend in the same case. That is because our method only record those suspicious DIPs rather than storing every DIP.

From the total number of keys in the sketches and the default parameter settings, the total memory consumption of our scheme can be estimated using the Eq. 2. The average memory cost is around 563.6 KB, which we consider can be easily accommodated in modern routers.

VI. Conclusion

In this paper, we present a fine-grained DDoS detection scheme based on the BCS structure to counter the threat of DDoS attacks. Our approach employs a two-level model to reduce both the size of the search space and time, and further make identification of specific victims possible in the high-speed network environment. We adopt the MCS structure in coarse-level detection to achieve fast detection, and the BCS structure in the fine-level to further guarantee the accuracy. We believe that this approach can accurately identify victims of DDoS attacks with a low memory footprint and give a timely response. We also propose a SRAM-based parallel architecture to achieve high-speed process. We finally analyze accuracy estimation issue and demonstrate a collaborative detection scheme based on the original single-host detection scheme. Experimental results show that our scheme outperforms previous sketch-based methods with respect to both storage scalability and detection accuracy.

Our future work will focus on designing a collaborative defense framework against DDoS attacks. Our proposed detection scheme can be used to facilitate defense against DDoS attacks in the following way. Since all the victims can be accurately detected by our collaborative scheme, an automatical rules generator can be developed to reinforce firewall and IDS systems in real-time.

Acknowledgment

We thank the reviewers for their detailed comments.

References

[1] H. Liu, Y. Sun, and M. S. Kim, "Fine-grained DDoS detection scheme based on bidirectional count sketch," in *Proceedings of IEEE International Conference on Computer Communication Networks*, Hawaii, Aug. 2011.

[2] J. Mirkovic and P. Reiher, "A taxonomy of DDoS attack and DDoS defense mechanisms," *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 2, pp. 39–53, Apr. 2004.

[3] M. Roesch, "Snort - lightweight intrusion detection for networks," in *Proceedings of the 13th USENIX Conference on System Administration*, Nov. 1999.

[4] V. Paxson, "Bro: A system for detecting network intruders in real-time," *Computer Networks*, vol. 31, no. 23–24, pp. 2435–2463, Dec. 1999.

[5] A. Lakhina, M. Crovella, and C. Diot, "Diagnosing network-wide traffic anomalies," in *Proceedings of ACM SIGCOMM*, Aug. 2004.

[6] ——, "Mining anomalies using traffic feature distributions," in *Proceedings of ACM SIGCOMM*, Aug. 2005.

[7] R. R. Kompella, S. Singh, and G. Varghese, "On scalable attack detection in the network," *IEEE/ACM Transactions on Networking*, vol. 15, no. 1, pp. 14–25, Feb. 2007.

[8] R. Schweller, A. Gupta, E. Parsons, and Y. Chen, "Reversible sketches for efficient and accurate change detection over network data streams," in *Proceedings of the ACM SIGCOMM Internet Measurement Conference*, Taormina, Sicily, Italy, Oct. 2004, pp. 207–212.

[9] R. Schweller, Z. Li, Y. Chen, Y. Gao, A. Gupta, Y. Zhang, P. Dinda, M.-Y. Kao, and G. Memik, "Reverse hashing for high-speed network monitoring: Algorithms, evaluation, and applications," in *Proceedings of IEEE INFOCOM*, Apr. 2006.

[10] O. Salem, S. Vaton, and A. Gravey, "A scalable, efficient and informative approach for anomaly-based intrusion detection systems: theory and practice," *International Journal of Network Management*, vol. 20, pp. 271–293, Sept. 2010.

[11] A. C. Gilbert, Y. Kotidis, S. Muthukrishnan, and M. J. Strauss, "Quicksand: Quick summary and analysis of network data," DIMACS, Tech. Rep. 2011-43, 2001.

[12] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Prentice Hall, 1993.

[13] "Auckland-IV trace data," 2001, http://wand.cs.waikato.ac.nz/wand/wits/auck/4/.

[14] P. Hick, E. Aben, K. Claffy, and J. Polterock, "The CAIDA DDoS Attack 2007 Dataset," http://www.caida.org/data/passive/ddos-20070804_dataset.xml (accessed on 2010-02-28).

[15] S. Sarvotham, R. Riedi, and R. Baraniuk, "Network traffic analysis and modeling at the connection level," in *Proceedings of Internet Measurement Workshop*, San Francisco, Nov. 2001.

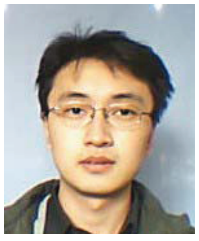
[16] A. Kuzmanovic and E. W. Knightly, "Low-rate TCP-targeted denial of service attacks and counter strategies," *IEEE/ACM Transactions on Networking*, vol. 14, pp. 683–696, Aug. 2006.

[17] J. L. Carter and M. N. Wegman, "Universal classes of hash functions," *Journal of Computer and System Sciences*, vol. 18, no. 2, pp. 143–154, 1979.

[18] H. Liu, Y. Sun, V. C. Valgenti, and M. S. Kim, "Trustguard: A flow-level reputation-based DDoS defense system," in *Proceedings of the 5th IEEE International Workshop on Personalized Networks*, Las Vegas, Jan. 2011.

[19] H. Liu and M. S. Kim, "Real-time detection of stealthy DDoS attacks using time-series decomposition," in *Proceedings of IEEE International Conference on Communications 2010*, May 2010.

Haiqin Liu received the B.S. degree in electrical engineering from Harbin Institute of Technology, Harbin, China, in 2005, the M.S. degree in network communication system and control from the University of Science and Technology of China, Hefei, China, in 2008. He is currently a Ph.D. candidate in computer science at Washington State University, Pullman. His research interests include network security, content distribution network and peer-to-peer network.



Yan Sun received the B.S. degree in applied physics in 2005 from University of Science and Technology, Beijing, China, the M.S. degree in 2008 in microelectronics from the University of Science and Technology of China, Hefei, China, and is currently a Ph.D. candidate in Computer Science, Washington State University. His research interests include network security, high-performance VLSI systems and computer architectures.

Min Sik Kim received the B.S. degree in computer engineering from Seoul National University, Seoul, Korea, in 1996, and the Ph.D. degree in computer science from the University of Texas at Austin in 2005. At present, he is an Assistant Professor of computer science with the School of Electrical Engineering and Computer Science, Washington State University, Pullman. Prior to joining Washington State University, he cofounded and served as Chief Technology Officer of Infnis, Inc., from 2002 to 2004. His research interests include overlay networks, network monitoring, and network traffic analysis.