Source-End DDoS Defense[*]

Jelena Mirković Gregory Prier Peter Reiher
University of California Los Angeles
Computer Science Department
3564 Boelter Hall
Los Angeles, CA 90095, USA
{sunshine, greg, reiher}@cs.ucla.edu

Abstract

A successful source-end DDoS defense enables early suppression of the attack and minimizes collateral damage. However, such an approach faces many challenges: (a) distributing the attack hinders detection; (b) defense systems must guarantee good service to legitimate traffic during the attack; and (c) deployment costs and false alarm levels must be sufficiently small and effectiveness must be high to provide deployment incentive. We discuss each of the challenges and describe one successful design of a source-end DDoS defense system — the D-WARD system. D-WARD was implemented in a Linux router. We include experimental results to illustrate D-WARD's performance.

1 Introduction

Distributed denial-of-service (DDoS) attacks misuse network resources of numerous subverted machines to generate packet streams targeting a single host or a set of hosts. The attack can be performed through a few malformed packets that exploit vulnerabilities in applications and protocols installed at the victim (*protocol attacks*), or through a vast number of seemingly legitimate packets that overwhelm the victim's resources (*flooding attacks)*. While a victim can defend against protocol attacks by applying protocol patches and keeping its applications up to date, it is usually helpless against flooding attacks and must request aid from upstream networks for attack suppression.

Ideally, flooding DDoS attacks should be stopped as close to the sources as possible. Source-end DDoS defense has several advantages over intermediate-network and victim-end defense approaches:

- **Congestion avoidance.** Restraining attack streams near the source preserves Internet resources that are usually overwhelmed by the attack traffic. This reduces overall congestion and increases resources available to legitimate users.
- **Small collateral damage.** Many DDoS defense systems respond to the attack by filtering or rate-limiting all traffic to the victim. Legitimate traffic thus suffers collateral damage. Moving DDoS defense closer to the sources reduces the range of legitimate traffic adversely affected by the response, as the traffic from uncompromised source networks proceeds to the victim unhampered.
- **Easier traceback.** Being close to the source facilitates easier attack traceback and investigation.
- **Sophisticated detection strategies.** Routers closer to the sources are likely to relay less traffic than intermediate routers and can dedicate more of their resources to DDoS defense. This facilitates use of more complex detection strategies.

However, source-end defense faces hard challenges with regard to:

- **Defense effectiveness.** There are no common characteristics of the streams comprising the attack that can facilitate early detection and filtering. Additionally, if the attack is sufficiently distributed, each participating machine or source-end network does not observe any higher outgoing load than usual.
- **Source-end response must be selective.** Networks hosting source-end defense do not experience a direct benefit from the defense system's operation and are thus poorly motivated to deploy it. In order to provide deployment incentive, a system must not only successfully detect and restrain many attacks but also: (a) provide good service to legitimate traffic between the de-

[*] This work is funded by DARPA under contract number N66001-01-1-8937.

ploying network and the victim; (b) have a low level of false positives; and (c) have a low deployment cost.

We discuss here each of the challenges of source-end defense and describe the D-WARD system as one successful design of a source-end defense system. D-WARD performs autonomous detection and suppression of DDoS attacks originating from the deploying network while guaranteeing good service to existing legitimate connections to the victim of the attack. The attack detection is achieved through constant monitoring of two-way traffic between the deploying network and the rest of the Internet. Online traffic statistics are periodically examined, looking for anomalies such as low number of responses, delayed response packets or presence of spoofing. If these anomalies are coupled with high and aggressive outgoing traffic to the victim, an attack will be detected. As a response, D-WARD installs a selective rate limit, differentiating between legitimate packets that are always forwarded to the victim, and the attack traffic which is severely constrained. D-WARD was first proposed in [18]. In this paper we present the design rationale behind the system and show how it meets the source-end defense challenges.

Section 2 discusses the challenge of source-end DDoS attack detection, and Section 3 discusses the challenge of an appropriate source-end response to the attack. Section 4 focuses on the deployment incentive for a source-end defense. Section 5 offers a brief description of the D-WARD system and discusses how it meets outlined challenges. Section 6 presents D-WARD performance results and deployment cost. Section 7 gives an overview of related work, and Section 8 concludes the paper.

2 Source-Based DDoS Attack Detection

One approach to DDoS defense would be to detect the attack at one point (i.e., at the victim or the congested core router), then respond to the attack further upstream. While this approach is appealing, it requires a secure, reliable and scalable communication mechanism between detection and response points, which is still an open problem. We thus limit our discussion to autonomous systems that perform unaided attack detection and response.

Attack detection is easiest at the victim network: high-volume of incoming traffic or disturbed operation can be readily used as a sign of DDoS attack. Effective response, however, depends on the attack volume and victim network resources. No victim-end defense is possible against sufficiently high-volume attacks — they overwhelm network resources even before they reach defense system, leaving legitimate clients outside. Additionally, high level of traffic aggregation hinders differentiation between legitimate and attack flows, leading to non-selective response. Thus, while protecting the victim, the response penalizes some legitimate traffic, still leading to denial-of-service.

The selectiveness and effectiveness of response improve as defense system is moved from the victim closer to the sources of the attack, but the detection accuracy deteriorates. Response is most effective at the source-end network, as attack streams can be stopped before they enter the Internet. Also, sophisticated profiling can be done to facilitate selectiveness of the response, since the attack traffic at the source is not highly aggregated. However, attack machines can be distributed among many source networks, thus each source network only observes a small amount of attack traffic that may appear legitimate, hindering detection.

Following observations can be used to design effective source-end detection:

- **Source-end firewall.** Firewalls perform attack detection using known attack signatures. This detection has a low level of false positives and can successfully be used to filter out malformed packets that are used for protocol attacks.

- **Threshold anomaly detection.** Each source-end network can define a set of thresholds for various traffic types, describing expected values for a set of parameters, such as average packet rate per connection, average number of outgoing UDP packets per destination, outgoing packet size distribution given time of the day, etc. These threshold values can be obtained through extensive training, guaranteeing low levels of false positives, and can detect a wide range of attacks. However, traffic patterns change and models need to be retrained. If the retraining is automatic (i.e. system adjusts its models to slow-changing traffic trends), attackers can misuse this to avoid detection.

- **Two-way traffic dynamics.** Since outgoing attack streams appear legitimate at the source network, just observation of the outgoing traffic alone cannot provide sufficient information to detect anomalous behavior and raise the alarm. However, it is generally the case that attack streams are nonresponsive to congestion signals—i.e., the outgoing attack stream to the victim will not reduce its rate if notified of congestion through reduction in the number of peer responses. Source-end system can use this observation to detect DDoS attacks misusing inherently responsive transport protocols such as TCP.

- **Spoofing detection.** Attack packets frequently spoof source IP address to avoid detection and accountability. Therefore, occurrence of aggressive spoofing in the outgoing flow can be a sure sign of DDoS attack. In order to detect aggressive spoofing, the system can: (a) enforce ingress filtering on all outgoing packets, thus preventing random spoofing; and (b) restrict number of outgoing connections per a single destination. These steps help detect those attacks that are performed through inherently non-responsive transport protocols, such as UDP, in cases where they deploy aggressive spoofing.

- **Connection semantics.** Sophisticated attacks can be performed while preserving correct two-way traffic dynamics and avoiding aggressive spoofing. It is much more difficult to preserve correct connection semantics, since this requires an attacker to store state per connection, thus putting more burden on the attacking machine. If system can afford to monitor, or even sample, per connection state, this enhances its ability to detect subtle attacks that preserve two-way dynamics, such as degrading attacks, slowly increasing rate attacks, small rate attacks and pulsing attacks.

3 Source-End DDoS Response

The source-end response to a detected DDoS attack must take into account the complexity of source-end detection which results in a low confidence of the attack signal. Since detection is unreliable, response must be **liberal**, to minimize damage to legitimate traffic inflicted by false detections. Response must also be **selective**, i.e., the system must be able to detect and preserve legitimate traffic to the alleged victim. The selectiveness of the response plays a crucial role in defining deployment incentive. If the system were not selective in constraining outgoing traffic to the victim, legitimate packets from a deploying network would be regularly dropped whenever the attack was detected. In this scenario, a network would be worse off deploying a source-end DDoS defense, because its legitimate packets, that otherwise might get some responses, would be invariantly dropped. This negative feature, along with the fact that the victim, not the source network deploying the defense system, harvests the benefit of DDoS defense, would be a strong argument against the source-end defense. On the other hand, the selective response that favors legitimate traffic will provide better service to legitimate clients of the source network during the attack, as their packets will not compete with the attack packets for the limited bandwidth. Instead, the legitimate packets receive preferential treatment and are sent promptly to the victim.

4 Deployment Incentive

A source-end defense system faces the hard deployment incentive challenge. Its cost is sustained by the deploying network, which does not receive substantial benefit from its operation. Furthermore, a victim cannot be asked to compensate for this cost, as being payed to stop participating in the attack would create opportunity for extortion. On the other hand, source-end defenses are necessary to stop Internet misuse as all other points of defense (intermediate and victim networks) do not prevent attack traffic from congesting shared Internet resources. Similar problems in the history have been solved by legislation. It is possible that in the future a contracted or legislative action will hold those who do not take reasonable steps to secure their system liable for damages inflicted by attackers misusing their machines. In that case, a source-end defense system would become part of an established security practice, and therefore a network deploying the system could not be held liable if its machines are misused for an attack.

Many people have concluded that stopping attacks completely is impossible, since there is a vast number of machines whose owners are unaware of security holes or are unwilling to fix them. A single source-end defense system installed at the network's exit router would prevent DDoS attacks originating from anywhere in the network, in spite of unsecured machines within.

Good performance characteristics, such as low number of false alarms, high effectiveness of attack response, a service guarantee to legitimate traffic, and low deployment cost would further strengthen the motivation for deployment of source-end DDoS defenses.

5 D-WARD

The D-WARD system is a source-end DDoS defense system whose goal is to detect and constrain outgoing attacks from the deploying network, while inflicting minimal damage to the legitimate traffic. The system is installed at the *source router* that serves as a gateway between the deploying network and the rest of the Internet. It monitors the traffic passing through the router in both directions and correlates these observation to detect anomalies that can be a sign of DDoS attack. Upon detection, it selectively imposes a rate limit on the outgoing flow to the victim, attempting to detect and forward legitimate packets regardless of the limit. The rate limit is dynamically adjusted based on subsequent observations and on the behavior of the limited flow.

D-WARD is a self-regulating reverse-feedback system. It consists of observation and throttling components that can be part of the source router itself, or can belong to a separate unit that interacts with the source router to obtain traffic statistics and install rate-limiting rules. The observation component gathers two-way traffic statistics and detects attacks. The throttling component then adjusts rate limit rules and communicates them to the source router. The imposed rate limits modify associated traffic flows and thus affect future observations, closing the feedback loop.

5.1 Attack Detection

D-WARD's observation component monitors two-way traffic at *flow* (the aggregate traffic between the source network and a foreign host) granularity in order to detect difficulties in communication that could be a sign of a DDoS attack, such as reduction in number of responses from the foreign peer. Additionally, it monitors two-way traffic at *connection* (the aggregate traffic between a local and a foreign IP addresses and port numbers) granularity, attempting to identify legitimate connections that should receive good service in case the associated flow becomes rate-limited.

D-WARD uses observations listed in Section 2 to design its attack detection method. Two-way traffic dynamics is used to devise normal flow and connection models for TCP and ICMP flows. During a TCP session, the data flow from the source to the destination is controlled by the constant flow of acknowledgments in the reverse direction. Under a TCP attack, the victim cannot generate sufficient number of replies for the received packets. A legitimate TCP flow backs off when such a reduction in responses is detected and lowers its sending rate, whereas the attack flow does not. D-WARD uses this to define a normal TCP flow model as TCP_{rto}—the maximum allowed ratio of the number of packets sent and received in the aggregate TCP flow to the peer. The ICMP protocol specifies many different message types. During normal operation the "timestamp," "information request," and "echo" messages should be paired with the corresponding reply. Under ICMP request attack (using "timestamp," "information request," and "echo" packets), the affected victim will not be able to generate sufficient amount of replies for received packets. Using this observation, the normal ICMP flow model defines $ICMP_{rto}$—the maximum allowed ratio of the number of echo, timestamp, and information request and reply packets sent and received in the aggregate flow to the peer. This method will not be effective under "reflection" attack, where source network receives a number of fake ICMP requests and bombards the victim by a flood of replies. It is likely that a fixed rate limit applied to incoming ICMP requests on a flow would be sufficient to handle this type of attack. Same ratios are used, along with more detailed connection semantics, to detect legitimate TCP and ICMP connections.

Spoofing detection is used to detect those attacks that misuse one-way communication, such as UDP attacks. D-WARD deploys ingress filtering and defines the normal UDP flow model as a set of thresholds: n_{conn}—an upper bound on the number of allowed connections per destination, and p_{conn}—a lower bound on the number of allowed packets per connection The model classifies a flow as an attack when at least one of these thresholds has been breached.

Threshold anomaly detection is deployed to detect attacks through ICMP "destination unreachable," "source quench," and "redirect," messages whose frequency is expected to be small. It is also used to define UDP_{rate}—a maximum allowed sending rate per UDP connection. This threshold enables detection of the attack through non-spoofed aggressive UDP connections.

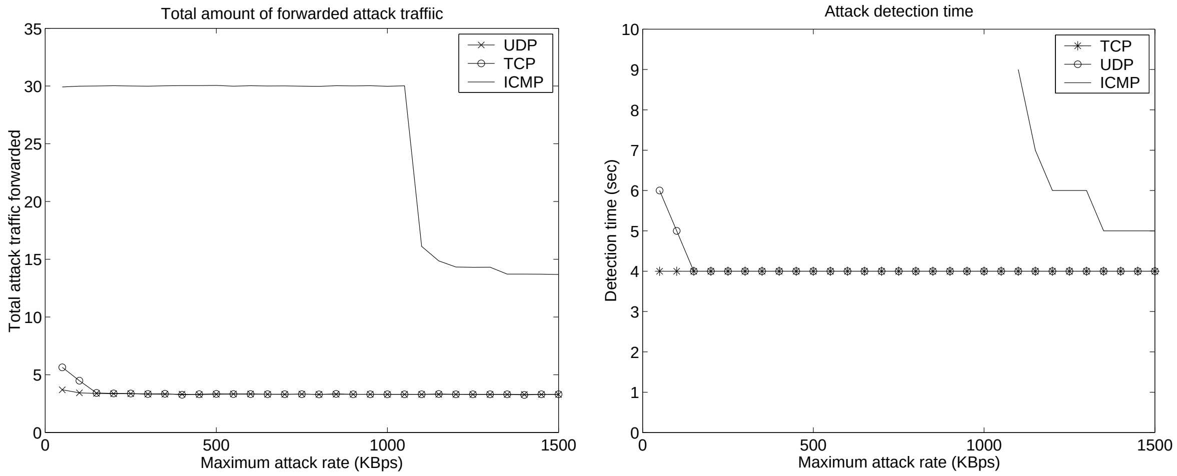
Connection semantics is used to detect subtle attacks, such as increasing rate and pulsing attacks, and attacks violating TCP connection semantics. Along with flow classification the observation component performs connection classification to detect legitimate connections that should receive good service. It monitors the amount of the outgoing traffic not belonging to legitimate connections—the *non-validated traffic*. Under normal system operation the amount of non-validated traffic is zero, since all connections are classified as legitimate. When new connections are initiated to the host, they create short-lived peaks of non-validated traffic that are soon brought to zero in subsequent classification steps. Under normal load these peaks are wide apart. Under an attack, the amount of non-validated traffic will exhibit prolonged bursts. The bursts are detected by calculating the minimum amount nv_{min} of observed non-validated traffic in previous N_{bursts} classification intervals. The attack detection is triggered if nv_{min} is greater than zero. Periodic sampling of the amount of non-validated traffic is used to detect pulsing attacks. Collected samples are stored in a first-in-first-out (FIFO) queue of size N_q. Attack detection is triggered if, at any observation interval, the minimum element in this queue has a non-zero value. The size of the queue N_q determines the probability that the attack will be detected; a shorter queue increases the chances of detection, but also may increase the number of false positives. Sampling probability defines the speed of filling the queue and thus affects the detection.

A source-firewall is currently not a part of D-WARD but it could be easily added as a stand-alone component and would nicely complement D-WARD's operation.

5.2 Attack Response

To meet the requirement for liberal response, D-WARD applies rate-limiting, rather than filtering, to the attack flow. The problem of regulating the sending rate of a one-way flow to the level manageable by the receiver (or the route to the receiver) has been recognized and addressed by the TCP congestion control mechanism. In process of defining the best value for the rate limit, D-WARD strives to solve a similar problem at a more aggregated scale. It controls the total flow to the peer, and it infers the peer's state from its response packets and the attack detection signal.

If the attack is detected within a flow, its rate limit is initially set to be a portion of its sending rate at the time of the detection. Upon subsequent attack detections, the rate limit



(a) Forwarded attack traffic, scaled by the maximum attack rate

(b) Attack detection time

Figure 1. Constant rate attack.

is decreased exponentially, providing the fast relief for the victim. The rate cannot be decreased more than $MinRate$ This guarantees each flow a chance to recover from false alarms. Once the attack is aborted, the imposed rate limit is linearly increased for a certain predefined time interval. This prevents oscillations and recurring attacks. After this slow recovery, the rate is increased exponentially and finally removed.

Along with attack detection, the rate limit depends also on the flow's behavior, i.e., its compliance to the imposed limit. This compliance is measured through the *compliance factor*, a ratio of outgoing bandwidth from this flow before and after it passes through the throttling component. Rate limit is inversely proportional to the compliance factor. If the flow is compliant the ratio is close to 1, and the rate limit will be decreased slower and increased faster. If the flow is aggressive, the ratio is close to 0, and this restricts the rate limit severely.

To meet requirement for the selective response, D-WARD stores information on those connections that are classified as legitimate. If an outgoing packet belongs to one of these connections, it is forwarded to the victim regardless of the rate limit.

6 Test Results and Analysis

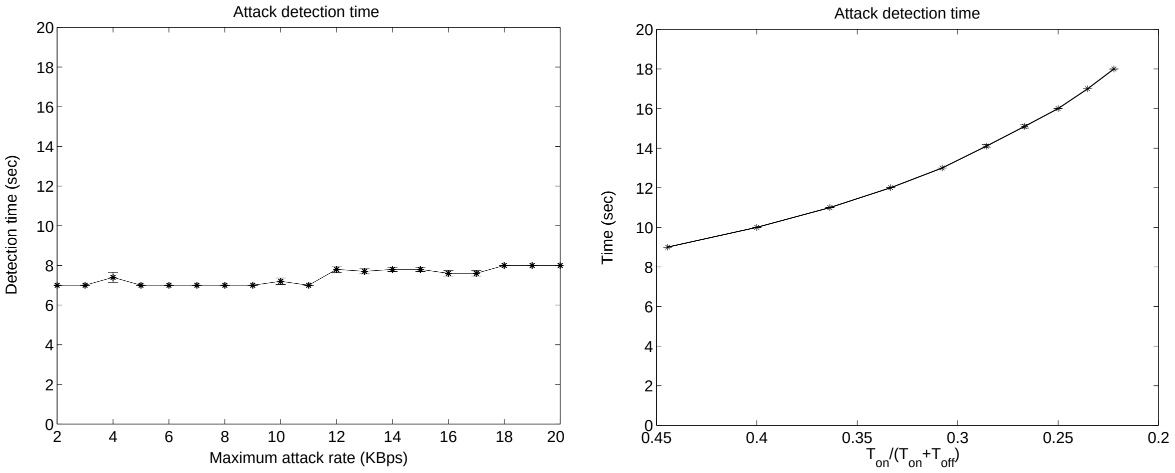
D-WARD was implemented in a Linux router, partly at the application level and partly as a kernel module. The application gathers traffic statistics, detects the attacks and calculates the appropriate rate limits. The kernel module then enforces the limits on outgoing flows.

To illustrate D-WARD's performance we provide some experimental results. We tested D-WARD in the simple test network. It consists of a source router deploying D-WARD, the attacker and the legitimate client who both belong to the source network, and a foreign host playing the role of victim. The simplicity of the test network does not affect the validity of the test results. Since D-WARD operates autonomously and analyzes only its incoming and outgoing traffic, multiple attacking domains could only speed up detection of the attack, because the victim would feel the denial-of-service sooner. So testing with only one source network does not give advantage to D-WARD. The effect of deploying multiple attack and legitimate client machines in the source network is mimicked by using only two machines that generate high traffic loads. Since D-WARD analyzes all incoming and outgoing traffic regardless of the links it comes on, the number of machines generating the traffic is transparent to the system. Furthermore, one attacking machine can fill the pipe to the router and thus more machines would not be able to increase incoming load into the router.

In order to test different attack scenarios we developed a customizable DDoS attack tool. It uses a master-slave architecture to coordinate attacks among multiple slaves. Attack traffic mixture (relative ratio of TCP SYN, ICMP_ECHO and UDP packets), packet size, attack rate, target ports, spoofing techniques and attack dynamics can be customized.

6.1 Constant-Rate Attack

To test detection of constant rate attacks, we generated TCP SYN, ICMP_ECHO, and UDP attacks with constant rate, varying the maximum rate from 100 KBps to 2 MBps. Short-lived FTP transfers were generated throughout the test to account for legitimate traffic. We measured the cu-



(a) Attack detection time (increasing rate attack)

(b) Attack detection time (pulsing attack)

Figure 2. Variable rate attacks.

mulative attack and good traffic that was delivered to the victim during the test. Figure 1(a) gives results of these tests. The cumulative attack traffic is scaled by the maximum attack rate. Since the D-WARD response eventually brings the forwarded attack traffic to a very low level, the scaled cumulative attack traffic shown on Y axes represents the time (in seconds) until the attack is constrained. A line with a cross mark represents the traffic passed to the victim in the case of a UDP attack, a line with a circle mark represents the case of a TCP attack, and a solid line represents the case of an ICMP attack. All attacks use a fixed packet size of 1KB.

Constant rate attacks pass similar amounts of the attack traffic for both UDP and TCP cases. This is due to the sudden onset of the attack, which creates a sufficient disturbance in the network to be quickly detected and controlled. ICMP attacks can pass undiscovered until a very high maximum attack rate is reached (1.1MBps). At higher attack rates, they are detected with an efficiency similar to UDP and TCP attacks, and quickly constrained. TCP and UDP attacks are constrained within 4-5 seconds. The ICMP attacks (when discovered) are constrained within 15 seconds. Figure 1(b) gives the time needed for detection of the attack. This time is measured from the start of the attack. In all experiments no good traffic was dropped, which proves that D-WARD is selective in its response.

6.2 Increasing-Rate and Pulsing Attacks

To test detection of more subtle attacks, we generated slowly increasing rate and pulsing attacks. In the first experiment, we initiated frequent and intensive file transfers between the legitimate client and the victim to achieve a high volume of legitimate traffic passing through D-WARD. We then generated a low-volume, increasing-rate TCP SYN attack. We used $N_{bursts} = 5$.

Figure 2(a) shows the detection time, when the maximum attack rate is varied over tests. The detection time is given from the beginning of the attack. It can be observed that the detection time does not depend on the maximum attack rate and is between N_{bursts} and $2 * N_{bursts}$ observation intervals. The total legitimate traffic dropped in these experiments is under 0.05%.

In the second experiment, we generated the same legitimate traffic as in the previous test, but attacked the victim with pulsing attacks. The duration of ON periods, T_{on}, was fixed at $N_{bursts} - 1$, i.e., 4 seconds, while the duration of OFF periods, T_{off}, was varied from 4 to 14, effectively changing the ratio $\frac{T_{on}}{T_{on}+T_{off}}$ from 0.5 to 0.2. Sampling probability was fixed at 0.6, and N_q was set to 4.

Figure 2(b) shows the detection time for pulsing attacks. The detection time depends exponentially on the ON/OFF period ratio but is reasonably small even for very small ratios (e.g., for the attack whose inactive intervals last 5 times its active intervals, the detection time was 18 seconds). This means that the attacker who would like to avoid detection would have to perform small-rate, small-duration, repetitive attacks with large inactive periods, which forces him to subvert a much larger number of machines than before. The total legitimate traffic drop in these runs was smaller than 0.01%.

6.3 False Alarms

In order to test the rate of false alarms with realistic traffic, we modified the system to read packet header data from a tcpdump-generated trace file instead of sniffing it from the network. We used packet traces gathered from UCLA Computer Science network during August 2001. The network has approximately 800 machines and experiences an average of 5.5 Mbps of outgoing traffic and 5.8 Mbps of incoming traffic. We assume that no attack has occurred during the trace-gathering process.

We determine the level of false positives by measuring the number of flow and connection misclassifications (the number of times that any flow was misclassified as attack or suspicious, and the number of times that any connection was misclassified as bad). We report this measure relative to the total number of flow and connection classifications performed during the trace. In all measurements D-WARD had a very low level of false positives: less than 2% for flow misclassifications and less than 0.1% for connection misclassifications. As collateral damage will be inflicted only in the case when both the flow is classified as attack or suspicious and the legitimate connection is misclassified as bad, we are confident that D-WARD's operation would not have any noticeable impact on legitimate traffic in the real network.

7 Related Work

There are many approaches to solving the serious problem of DDoS. Space permits only a brief review of those approaches most related to source-end defense and to D-WARD.

Intrusion detection systems (IDS) deploy signature and anomaly-based detection of intrusion attempts. While they apply very sophisticated methods for detection, IDS systems generally do not take automated action to stop malicious behavior. D-WARD on the other hand has a set of fairly simple flow and connection models that it uses for attack detection, but provides efficient response to detected attacks. Combining signature and anomaly-based detection methods of intrusion-detection systems (such as deployed in NetRanger [3], NID [4], SecureNet PRO [17], RealSecure [11], [21] and NFR-NID [19]) with D-WARD would likely enhance detection accuracy.

Several DDoS defense systems [16] and [1] perform anomaly detection (usually at the victim network) by observing numerous traffic parameters and defining a range of allowed values based on the analysis of packet trace data. The attack response is to impose a non-selective fixed rate limit to offending streams, thus likely damaging legitimate traffic. Instead of fixed legitimate traffic models that D-WARD uses, [16] and [1] train their models based on the observed traffic in the network. Once underlying traffic patterns change, models need to be retrained to avoid false positives. Like all learning approaches, these detection models can be mistrained by the attacker to regard attack traffic as legitimate. On the other hand, models tailored especially to traffic characteristics of deploying network are likely to yield better detection accuracy than those deployed by D-WARD. D-WARD's selective response guarantees good service to legitimate traffic and is likely to inflict less collateral damage than fixed rate-limiting deployed by [16] and [1].

MANAnet [5] is a reverse firewall system that prevents DDoS attacks by limiting the rate of "unexpected" packets at a network's exit router. This is a commercial product and because information concerning its detection and response mechanisms is not publicly available we cannot provide a detailed comparison with D-WARD.

MULTOPS [10] proposes a heuristic and a data-structure that network devices can use to detect both outgoing and incoming DDoS attacks. The proposed data structure is a multi-level tree, storing certain traffic characteristics in nodes corresponding to subnet prefixes. The attack is detected by abnormal packet ratio values in the node statistics. MULTOPS is likely to detect smaller range of attacks than D-WARD, as it uses simpler traffic models. MULTOPS attack response is fixed and non-selective rate-limiting which yields higher collateral damage than D-WARD's selective rate-limiting.

In [8] Floyd et al. propose intermediate-network defense — aggregate congestion control (ACC). Routers detect and control flows that create congestion (frequently a sign of DDoS attacks), by deploying Random Early Detection [9]. Congested router applies non-selective rate limit at the aggregate traffic, thus inflicting collateral damage [12]. Unlike D-WARD, which is an autonomous system, ACC-enabled routers can achieve cooperative defense by propagating rate-limit requests to their adjacent upstream neighbors. ACC has to be deployed contiguously in order to facilitate rate-limit propagation. ACC routers will detect only high-bandwidth attacks that create congestion, D-WARD deploys more sophisticated traffic models and successfully detects small-rate attacks.

Secure Overlay Services (SOS) [13] prevent denial-of-service attacks on critical servers by routing requests from previously authenticated clients to those servers via an overlay network. All other requests are filtered by the overlay. SOS is a distributed system that offers excellent protection to the specified target at the cost of modifying client systems, thus it is not suitable for protection of public servers.

Several traceback mechanisms have been proposed to locate attacking nodes ([23], [6], [24], [2] [22]). These systems provide information about the identity of attacking machines, but do not stop DDoS attacks. The complexity of a traceback mechanism is large if the attack is distributed, and some mechanisms are prone to packet marking attacks.

Several filtering mechanisms have been proposed to prevent spoofing source addresses in IP packets ([7], [14], [20]). While IP spoofing is not necessary in DDoS attacks, it helps attackers hide the identity of attacking machines so they can reuse them for future attacks. D-WARD and many other DDoS prevention mechanisms would benefit from more reliable packet source addresses.

Protocol and application scrubbing [15] have been proposed to remove ambiguities from transport and application protocols. Scrubbing can eliminate many vulnerability attacks that use protocol ambiguities to bypass intrusion detection systems. A protocol scrubber could complement D-WARD by preventing outgoing vulnerability attacks.

8 Conclusion

Source-end DDoS defense can be an effective approach to constraining DDoS attacks close to their sources. It faces many challenges in order to be widely deployed. Source-end detection is difficult and unreliable. Source-end response has to be flexible and selective to compensate for poor detection and to offer deployment incentive. In this paper we discussed each of these challenges and defined requirements for a successful DDoS defense system. We then discussed one possible design - a D-WARD system.

The D-WARD system is highly successful at blocking the kinds of DDoS attacks that are commonly perpetrated today. But as systems like D-WARD make such simple attacks infeasible, attackers will develop more sophisticated attacks that might slip by basic defense mechanisms. This paper also demonstrates an approach to making D-WARD resilient to a wider class of attacks, including sophisticated attacks that slowly ramp up and attacks geared toward wasting resources rather than entirely crippling the target. Once detected, the attacks can be throttled to limit damage to the victim.

Our experiments with D-WARD show promise for source-end DDoS defense. While it is not a complete solution to DDoS attacks, we believe that source-end defense is one of the crucial building stones of the complete solution and essential for promoting Internet security.

References

[1] Arbor Networks. The Peakflow Platform. http://www.arbornetworks.com.

[2] S. Bellovin, M. Leech, and T. Taylor. ICMP Traceback Messages. *Internet draft, work in progress*, October 2001.

[3] Cisco. NetRanger Overview. http://www.cisco.com/univercd/cc/td/doc/product/iaabu/csids/csids1/csidsug/overview.htm.

[4] Computer Incident Advisory Capability. Network Intrusion Detector Overview. http://ciac.llnl.gov/cstc/nid/intro.html.

[5] Cs3, Inc. MANAnet DDoS White Papers. http://www.cs3-inc.com/mananet.html.

[6] D. Dean, M. Franklin, and A. Stubblefield. An algebraic approach to IP Traceback. In *Proceedings of the 2001 Network and Distributed System Security Symposium*, February 2001.

[7] P. Ferguson and D.Senie. Network Ingress Filtering: Defeating Denial of Service Attacks which Employ IP Source Address Spoofing. *RFC 2827*.

[8] S. Floyd, S. M. Bellovin, J. Ioannidis, K. Kompella, R. Mahajan, and V. Paxson. Pushback Messages for Controlling Aggregates in the Network. *Internet draft, work in progress*, July 2001.

[9] S. Floyd and V. Jacobson. Random Early Detection Gateways for Congestion Avoidance. *IEEE/ACM Transactions on Networking*, August 1993.

[10] T. M. Gil and M. Poletto. MULTOPS: A Data-Structure for Bandwidth Attack Detection. In *Proceedings of 10th Usenix Security Symposium*, August 2001.

[11] Internet Security Systems. Intrusion Detection Security Products. http://www.iss.net/securing_e-business/security_products/intrusion_detection/index.php.

[12] J. Ioannidis and S. M. Bellovin. Pushback: Router-Based Defense Against DDoS Attacks. In *Proceedings of NDSS*, February 2002.

[13] A. D. Keromytis, V. Misra, and D. Rubenstein. SOS: Secure Overlay Services. In *Proceedings of SIGCOMM 2002*, 2002.

[14] J. Li, J. Mirkovic, M. Wang, P. Reiher, and L. Zhang. SAVE: Source Address Validity Enforcement Protocol. In *Proceedings of INFOCOM 2002*, June 2002.

[15] G. R. Malan, D. Watson, F. Jahanian, and P. Howell. Transport and application protocol scrubbing. In *Proceedings of INFOCOM 2000*, pages 1381–1390, 2000.

[16] Mazu Networks. Mazu Technical White Papers. http://www.mazunetworks.com/white_papers/.

[17] MimeStar.com. SecureNet PRO Feature List. http://www.mimestar.com/products/.

[18] J. Mirkovic, G. Prier, and P. Reiher. Attacking DDoS at the Source. In *Proceedings of the ICNP*, November 2002.

[19] NFR Security. NFR Network Intrusion Detection. http://www.nfr.com/products/NID/.

[20] K. Park and H. Lee. On the Effectiveness of Route-Based Packet Filtering for Distributed DoS Attack Prevention in Power-Law Internets. In *Proceedings of ACM SIGCOMM 2001*, August 2001.

[21] P. Porras and P. Neumann. EMERALD: Event Monitoring Enabling Responses to Anomalous Live Disturbances. In *Proceedings of the Nineteenth National Computer Security Conference*, October 1997.

[22] S. Savage, D. Wetherall, A. Karlin, and T. Anderson. Practical Network Support for IP Traceback. In *Proceedings of ACM SIGCOMM 2000*, August 2000.

[23] A. C. Snoeren, C. Partridge, L. A. Sanchez, C. E. Jones, F. Tchakountio, S. T. Kent, and W. T. Strayer. Hash-Based IP Traceback. In *Proceedings of ACM SIGCOMM 2001*, August 2001.

[24] D. X. Song and A. Perrig. Advanced and authenticated marking schemes for IP Traceback. In *Proceedings of IEEE Infocom 2001*, 2001.