CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

Distributed Denial of Service attacks are widely regarded as a major threat to the Internet. Flooding-based and protocol based DDoS attack are very common methods to attack a victim machine by sending large volume of malicious traffic. Although a number of techniques have been proposed to defeat DDoS attacks, it is still hard to detect and respond to flooding-based DDoS attacks due to a large number of attacking machines, the use of source-address spoofing and the similarities between legitimate and attack traffic.

The related works on DDoS defense are categorized based on DDoS detection, response and defense framework. The detection techniques mainly include IP attributes-based DDoS detection or traffic volume-based DDoS detection. Response techniques primarily involve either packet filtering or rate limiting. Defense techniques can be categorized into three types based on the location of the defense system in the network: victim end defense, source end defense and distributed defense. The following chapter discusses some of the key techniques proposed by researchers to defend against DDoS attacks.
2.2 OVERLAY ARCHITECTURE

Secure Overlay Services (SOS) Architecture proposed by Keromytis et al. (2004) is a proactive approach to mitigate DDoS attacks. The architecture comprises of three layers: Secure Overlay Access Point nodes, (which receives the traffic destined for the target server), Beacon nodes, (which receives the packets from the Secure Overlay Access Points and forwards the packets to the Secret Servlets) and the Secure Servlet nodes, each layer comprising of multiple nodes. The nodes in each layer communicate only with nodes in the adjacent layer. The architecture aims to achieve communication between confirmed users and target.

When a server experiences a DDoS attack it redirects all traffic to the overlay network and all further transmissions to the protected target are validated at entry points of the overlay. A source that wants to communicate with the target contacts a Secure Overlay Access Point. After authenticating and authorizing the request, the Secure Overlay Access Point securely routes all traffic from the source to the target via one of the Beacon nodes which verifies the validity of the received information and forward traffic to the Secure Servlet which in turn forward validated traffic through the filtering routers to the target.

Xuan et al. (2005) performed an exhaustive analysis of the resilience of Keromytis’s three layered SOS architecture under intelligent DDoS attacks and proposed a generalized multi-layered SOS architecture, which aims to provide reliable communication between clients and a target under DDoS attack. It employs a set of overlay nodes arranged in hierarchical layers that control access to the target by traffic authentication and routing. The Secure Overlay Access Point layer is the first point of contact for a client which requires communication with the target system. The Secure Servlet layer is the last layer and forwards the packets through the filtering router to
the target. Multiple layers are present between the Secure Overlay Access Point and Servlet layers and they perform the functionality of Beacons, i.e. nodes in the layer n+1 will forward traffic that arrive from a node at layer n only.

Wang et al. (2006) performed an extensive study on the overlay systems proposed by Keromytis and Xuan and proposed the Secure Overlay Forwarding Systems (SOFS) – another generalized model of the original SOS architecture. The standard architectural features of the system included the layering (the number of layers between the client and server), mapping degree (number of next layer neighbors a node can communicate with) and node distribution (number of nodes per layer). A formal mathematical model of intelligent DDoS attacks was developed and analytical approaches and simulations were used to study the system performance and sensitivity of design features on performance under the attack models. The performance metric was the probability that a client can find a path to communicate with the server under ongoing attacks.

In et al. (2005) identified a major vulnerability in the SOS architecture – if any node in the SOS architecture (Secure Overlay Access Point / Beacon / Secure Servlet) is attacked the node will exit the overlay. If the attacker has the ability / resources to cover a large number of nodes at the same time and can identify a group of SOS nodes, an attack launched by them simultaneously leaves the target / victim vulnerable.

To overcome this vulnerability, In et al. proposed a Partaking-based Self-reform Algorithm to sustain availability against such intelligent attacks. Information about all packets is collected within certain time slice. Using clustering, the destination IP address is grouped into three categories – remote group, related group and neighboring group. When a node in the SOS architecture detects that it is under attack, it broadcasts this information over
the channel to its neighbor nodes before it quits the SOS system making the architecture more adaptive and resilient against intelligent DDoS attacks.

Touch et al. (2003) devised a dynamic parallel overlay architecture to provide dynamic defense against DDoS attacks among private groups of networked systems. Multiple layered Internet overlays are used to apply encryption, routing and configuration diversity, providing multiple alternate networks over which traffic is automatically routed. The result is a dynamic backbone (a DynaBone) that provides DDoS resistance with increased performance and the ability to react to attacks.

Chen and Chow (2004) designed a random peer-to-peer (RP2P) network that connects the registered client networks with the registered servers to provide an Anti-DDoS service (AID). The AID service is implemented as a distributed overlay system consisting of geographically dispersed AID stations for service registration and anti-DDoS operations and communicates among themselves via secure communication channels. The registration establishes an IPSec tunnel with the AID station. All tunnels together form an exclusive overlay network between the registered clients and the registered servers. This overlay network will be activated when a registered server is under attack ensuring that the server is open to all registered clients. The service is scalable, efficient and light-weighted and is used only when a DDoS attack occurs.

Jing et al. (2006) proposed a hierarchical domain-aware Overlay-based coOperative Defense Network (O2-DN) to construct a built-in distributed rate limit framework to defeat – DDoS attacks. The architecture has three major components: Attack Detection Agent which raises an alert on attack detection and sends a defense request to the Defense Service Provider, specifying the IP address if the victim, the attack signature (port and protocol used), requester’s certificate and digital signature of the requester.. The
Defense Service Provider verifies the authenticity and integrity of the request and performs rate limit decision making to determine the real time rate of incoming traffic flow towards the victim and the rate limit to be applied. It then dispatches the rate limit commands to the rate limiters in the ISP. Each Defense Service Provider has direct access to the rate limiter in the ISP within its management domain only. Rate Limiters performs the actual rate limiting of the specific malicious flow and ensures that the aggregated incoming traffic does not exceed the capacity of the victim’s resources.

Hamano et al. (2005) proposed a DDoS defense scheme against flood type attacks based on traffic redirection in which edge and border routers divert suspicious packets to a Central Defense Node(s) (C–DNs) using secure tunneling. The C–DN analyzes the received traffic and filter out the attack packets. They then decapsulate the remaining packets and forward them to the egress router through normal IP–destination–address–based forwarding, thus ensuring that a packet passes through a C–DN only once.

When the C–DN detects the termination of attack on the given IP address, it reports the event to the manager node which reconfigures the policy based routing rules of the edge / border routers, so that the packets are not forwarded to the C–DN but are forwarded to the destination through normal IP–destination–address–based forwarding.

2.3 ACTIVE NETWORKS, GATEWAYS AND ROUTING PROTOCOLS

Hess et al. (2003) proposed a Flexible Intrusion Detection and Response Framework for Active Networks (FIDRAN), a framework for flexible intrusion detection and response on an underlying active networking environment. The five major components of FIDRAN are management module, operational module, control module, security policy and response
mechanism. Management and control modules deal with configuration and administration issues. Operational module involves attack signature detection, anomaly detection, bandwidth monitoring and response triggering and security policy provides information needed by management and control module. The response mechanism involves simple packet discard, firewall reconfiguration, traffic redirection and invocation of additional active services like traceback mechanisms.

Kim et al. (2003) proposed an ACtive edge-Tagging (ACT) scheme to identify and isolate intruders in an Active Networks environment. The scheme identifies the physical location of attackers using the first-hop router (Monitor) capability to identify the actual IP source address of each incoming packet from host machines and isolates the attackers from potential victims by discarding the attack packets before the victim. The salient feature of the scheme is its effectiveness when multiple packets using spoofed IP source addresses are transmitted over an extended period of time.

Demir and Ghose (2005) suggested the use of Active Gateways for real-time protection against DDoS attacks. The solution uses a dynamic packet filtering on dual-ported active NIC based gateways to dynamically filter attacking packets based on locally measured request rates and information from the server (such as server loading, number of incomplete connections). This information is periodically sent from the servers or can be received on explicit requests sent from the gateway’s host to a server. The gateway incorporates actual loading and status information from the server into its filtering and forwarding rules and loads these rules onto the memory of the active NIC gateways. The gateway performs dynamic packet filtering to allow the servers to provide a guaranteed level of service.

Wang et al. (2003) proposed a Transport-Aware IP Router architecture that provides fine-grained service differentiation and resource
isolation among different classes of traffic aggregates. Transport layer information required for packet classification and resource management is inferred from the IP headers. The architecture consists of a fine-grained Quality-of-Service (QoS) classifier and an adaptive weight-based resource manager. A two-stage packet-classification mechanism is used to decouple the fine-grained QoS lookup from the usual routing lookup at core routers. The fine-grained service differentiation and resource isolation is achieved using separate queues and adaptive-weighted bandwidth allocation.

2.4 FILTERING AND RATE LIMITING

Zhang et al. (2006) proposed an expert system model to acquire prior knowledge from the ordinary network information and then use the prior knowledge to defend against DDoS attacks. The three components of the expert model are knowledge construction phase, detecting phase and filtering phase.

The knowledge construction phase predefines the prior knowledge such as detecting rules, state transition rules and access control information to reach the goal of detecting and filtering approach. In detecting phase, the detecting rules from knowledge base are used to monitor the network traffic to check if the attack traffic is launched. When switching to attacked state, the defending system generates new filtering policy to drop the garbage traffic, provide the critical service continuously and switch to survival state.

Kim et al. (2003) proposed a Statistics-based distributed detection and automated on-line attack characterization for overload control against DDoS attacks. The proposed scheme consists of three phases: Detection, Differentiate and Discard.
Detection process uses a Bloom Filter / leaky bucket arrays (BFLBA) to monitor key traffic statistics of each protected target, while keeping minimal per-target state information. Differentiation between legitimate and attacking packets destined towards the victim is based on a readily computed, Bayesian-theoretic metric of each packet called the Conditional Legitimate Probability (CLP), which estimates the legitimacy of a suspicious packet. The discard threshold is adjusted according to (1) the distribution of CLP of all suspicious packets and (2) the congestion level of the victim. Packets are discarded selectively by comparing the CLP of each packet with a dynamic threshold.

Yau et al. (2005) proposed a proactive approach to DDoS defense by installing a router throttle at selected upstream routers specifically to regulate the contributing packet rate to a moderate level by either dropping or rerouting that traffic to an alternate server. The aim of the control algorithm is to keep the server load within a maximum and minimum bound when a throttle is in effect. The defense is initiated by activating the rate throttle at a subset of its upstream routers. The adaptive throttle algorithm can effectively protect a server from resource overload and increase the ability of good user traffic to arrive at the intended server.

Minority First (MF) is a traffic metering and control scheme, proposed by Ahn et al. (2003) for defense against DDoS attacks. Control of malicious traffic is done based on Source–Traffic–Trunk (STT) based metering and queue mapping. STT is the aggregate of flow that comes from the same source network. An STT is assigned a Permission Load, which is the maximum load that the STT using a high priority queue is allowed to carry. MF scheme uses a differentiated service instead of a blocking policy and can protect legitimate traffic irrespective of the strength of the DDoS attacks.
Xiaofeng et al. (2005) proposed a mechanism to defend against SYN flooding attack based on Network Measurement System (NMS) that monitors the network traffic at real time based on passive measurement technologies, focusing on analyses and evaluations to network performance. NMS consists of a register server, monitors, probes and policy servers. The register server primarily provides common services such as the component information service involving naming, membership, address, permission and status and component registration and management. Monitors are responsible for component configuration and maintenance, policy edit and management, event monitoring, initiating measurement request to probes, collecting measurement results, result analysis and display.

Probe components arrange and schedule the network measurement utilities, collect, cache and retrieve the measurement results and send the results back to the monitor. Two components – DetSYNAtk and TraceSYNAtk – are located on the probe to detect and trace SYN flooding. Policy servers define the explanation and storage of the policy, the actions the system expects to achieve or the states the system expects to attain. The defense mechanism includes attack-detection, service-protection and attack-trace.

Fan et al. (2005) proposed a proactive Source Router Preferential Dropping (SRPD) scheme for detecting and controlling DDoS attacks at source networks. High sending rates associated with persistent long response times imply a potential attack. After identifying the high rate flows, SRPD measures each flows arrival rate and computes its average response time. If the average response time exceeds a predefined threshold, an ICMP control message is sent to the victim router to collect queue length usage information and the each flows status is updated. When the source router determines the presence of a DDoS attack it computes the drop probability of each flow and
malicious packets are dropped at the edge routers close to the attacking sources, thus reducing the congestion at the victim network. This minimizes the collateral damage to the legitimate traffic.

Chow et al. (2005) designed and implemented the Secure COLlective Defense (SCOLD) system against DDoS attacks. The key idea of SCOLD is to provide alternate routes between clients and target server via a set of proxy servers and alternate gateways when the normal route is unavailable or unstable due to network failures, congestion, or DDoS attacks. In SCOLD, the IP addresses of the alternate gateways and the SCOLD coordinator(s) are revealed only to the trustworthy proxy servers to protect them from being attacked by malicious clients. The traffic between clients and target server is transported over Internet through the indirect routes during an ongoing DDoS attacks.

When the server experiences a DDoS attack, the IDS on the server raises an alert and notifies the coordinator. The coordinator identifies the network from which the malicious/legitimate packets are generated. The coordinator then activates a set of proxy servers which updates the DNS servers of the networks from which legitimate packets are generated. DNS servers of networks from which the malicious traffic pattern was identified is not updated, allowing all the malicious traffic to continue along the same congested route to the target server. The legitimate clients use the indirect route to the target server, via the proxy server setup by the SCOLD system.

Kim et al. (2003) proposed a High-Speed Router Filter for Blocking TCP Flooding under DDoS Attack. The high-speed filter is composed of three modules. The packet screener determines whether a packet should be processed at the filter, connection verifier detects legitimate flows and assigns a flow ID and the filter array actually drops the attack packets. The filter can be inserted either to a server side port or to a network side port.
Wu et al. (2009) proposed an adaptive control mechanism using Shewhart's control charts based on network connection to aid in handling DDoS attacks. The mechanism is designed to prevent incoming traffic from exceeding a given threshold, while allowing as much incoming legitimate traffic as possible.

Li et al. (2009) proposed the use of Probability Metrics to distinguish DDoS Attacks from Flash Crowds. Flash crowds are the results of the legitimate users respond to special events such as breaking news or popular products (movies, music and software) release. The DDoS attack detection system includes flow anomaly detector, flow distribution estimator, hybrid probability metric calculator and decision device. The probability metrics used include the total variation metric which measures the difference of two discrete probability distributions and the Bhattacharyya metric which measures the similarity of two discrete probability distributions. The hybrid probability metric of total variation and similarity coefficient is used to distinguish DDoS attacks flow from Flash crowd flow and DDoS attacks flow from Normal network flow and even also can distinguish the anomaly being DDoS attacks flow or Flash crowd flow from Normal network flow very well.

Tupakula et al. (2005) proposed an Automated Model for Counteracting DDoS Attacks. The architecture involves a Controller-Agent model. In each ISP domain, a trusted entity (dedicated host) is nominated as a controller. The agents are implemented on all edge routers. During the time of an attack, the victim requests the controller in its domain to prevent the attack. A session is established between the victim and the controller after proper authentication of the victim. Depending on the number of agents present within its domain, the controller will generate and issue the controller Identity (ID) and unique agent ID to each agent and commands its agents to mark the
traffic to the victim. The agents filter the traffic that are destined to the victim and mark the traffic with controller ID and agent ID in the fragment ID field.

Tian et al. (2006) proposed an auction-based dynamic bandwidth arbitration method for defending against Flash Crowds and malicious Traffic Attacks. In auction mechanism the users submit bids and request a share of the total resource and a price they are willing to pay for it and the auctioneer allocates shares of the resource to the users based on their bids. Congestion evaluation component detects congestion on the network and rate estimator will then perform rate limiting on the amount of traffic destined for the victim Web server as estimated by the bandwidth arbitrator. An adaptive throttle algorithm controls the client’s sending rate and traffic that exceeds the rate limit will be dropped. Thus maximum throughput of Web server can be guaranteed and the total utility of users can be maximized too.

Chen and Park (2005) combined the concepts of Pushback and Packet Marking and proposed a novel attack mitigation scheme called Attack Diagnosis which is a reactive defense that is activated by a victim host after an attack has been detected. The execution involves four steps: (1) An Intrusion Detection System installed at the victim (or at its firewall) detects an attack; (2) The victim instructs the upstream routers to start marking packets with traceback information; (3) Based on the traceback information extracted from collected packets, the victim reconstructs the attack paths; and (4) The victim instructs the appropriate upstream routers to filter attack packets.

Karrer (2006) proposed an architecture – Edge-based Capabilities – that prevents DDoS attacks and malware spread at the source edge and thus prevent service disruption. Karrer introduced a novel network element termed gate – a traffic filter deployed at the network edge. The gate allows authenticated traffic to pass while it blocks or delays unknown and unauthenticated traffic.
Chen et al. (2005) proposed an adaptive packet dropping algorithm called MAlicious Flow Identification and Cutoff (MAFIC) to accurately identify and push back DDoS Attacks. MAFIC addresses two major challenges involved in DDoS defense – identifying source IP spoofing and multiple distributed zombies utilization. On receiving the notification of DDoS attack from the victim router, MAFIC records the variances of arrival rate of suspicious flows. By monitoring the response to packet loss from the flow source, malicious attacking flows are accurately identified and all their packets are then dropped before reaching the victim.

Yaar et al. (2004) proposed a Stateless Internet Flow Filter (SIFF), which allows an end-host to selectively stop individual flows from reaching its network. Network traffic is classified as privileged (prioritized packets subject to recipient control) or unprivileged (legacy traffic). Privileged packets are always given priority over non-privileged ones when contending for bandwidth.

Sung and Xu (2003) proposed a protocol-independent DDoS defense scheme that is able to dramatically improve the throughput of legitimate traffic during a DDoS attack by performing “smart filtering” – dropping DDoS traffic with high probability while allowing most of the legitimate traffic to go through. To make this distinction, it leverages on and extends IP traceback techniques to infer whether or not a network edge is on the path from an attacker. The scheme involves three modules – Enhanced Probabilistic Marking module which inscribes either a signaling mark (information for IP traceback) or data mark (information for filtering) into packet header with a certain probability, Attack Mitigation Decision-making (AMD) module which reconstructs attack paths and determines the probability with which each type of packets (signaling or data marked) should be forwarded and Preferential Packet Filtering module which filter packets
differentially based on the instructions issued to them from the AMD module, once an attack is detected.

Zheng et al. (2006) proposed a scheduling algorithm to Compare and Control Unresponsive flows to determine if a packet originates from an unresponsive flow or not by comparing the packet with three candidate packets selected randomly from the current queue. If there are two or more candidate packets having the same address as the incoming packet, the flow is considered as a potential unresponsive flow. A counter is enabled to record the number of times two or more candidate packets have the same address as the incoming packet. If this exceeds a threshold, the flow is confirmed as an unresponsive flow and its sending rate of data packets will be limited immediately.

Hu et al. (2002) proposed a time-window based filtering mechanism followed by a queue management policy. The Window-Based Packet Filtering (WBPF) is turned on only if the aggregated arriving flow rate is larger than the available bandwidth of the output link and the queue experiences packet drops, i.e., the WFPF is on during the congestion period.

When the WBPF is on, a new incoming packet is dropped if a packet from the same flow has arrived in the previous consecutive window since it should be from a high rate flow. The time-window size is dynamically calculated based on several parameters observed during the current window period. A basic idea is that the next window size is increased if too many packet are still being dropped at the queue (i.e., not enough packets are filtered at the filter) and is decreased if the link utilization is too little while packets are being dropped at the filter (i.e., too many packet are filtered).

DDoS attacks can either exhibit static characteristic or can dynamically change their behavior. Li et al. (2009) proposed a statistic based
approach to classify normal and abnormal behavior of packets in both situations and the use of filters at the perimeter of a victim network to drop abnormal packets.

Given a mixture of actual attack traffic and a legitimate traffic the statistical algorithm considers the network and transport layer header attributes such as Time To Live (TTL), IP address, packet size and port number and calculates the attack packet distribution perceived by the filter. Based on this estimation the filter devices a policy to forward or drop the packet.

Gupta et al. (2008) proposed an ISP Level Solution to Combat DDoS Attacks using Combined Statistical Based Approach. The proposed framework deals with the detection of variety of DDoS attacks by monitoring propagation of abrupt traffic changes inside ISP Domain and then characterizes flows that carry attack traffic. Two statistical metrics namely, Volume and Flow are used as parameters to detect DDoS attacks. Attack flows are identified and tagged in real time and filtering or rate throttling is applied to malicious traffic according to strength of attacks. Accuracy of threshold value settings and hence the effectiveness of an anomaly based detection and characterization system is increased by the use of Six-Sigma methods.

Chan et al. (2004) proposed the use of an Intrusion Detection Router (IDR) to provide technologies to detect and harden existing Internet infrastructure from large-scale network intrusions and DDoS attacks. IDR detects DDoS attack passing through a router by employing Bloom Filter to locate the exceptional heavy volume of packets going to the same destination. Once suspicious DDoS attack packets are identified, the DDoS attack detection module will pass them to the packet classification module for analysis which will identify highly suspicious packets to be discarded and
pass less suspicious packets to the traffic control module to rate limit the bandwidth of these packets, thus protecting the legitimate packets.

The perimeter-based defense against DDoS proposed by Chen and Song (2005) performs two major tasks. It first identifies the attack aggregates and then identifies the flooding sources and installs appropriate rate-limit filters on the edge routers connecting to the flooding sources. A rate-limit filter is a tuple \( <a, r> \) where \( a \) is a traffic aggregate and \( r \) is the rate limit for \( a \).

He et al. (2005) proposed a defense system requiring limited fixed length memory and low computation overhead to encourage active participation of Internet Service Providers in DDoS defense. Bloom Filter, a space efficient hash data structure, is modified and used to record the behavior of each packet. When a DDoS attack is launched, an exceptional heavy volume of SYN request packets are sent towards the victim IP. If the value of the count exceeds a pre-defined threshold value it is regarded as suspicious. If there is at least one count in each table containing suspicious value, the DDoS attack alarm is launched.

Mirkovic et al. (2005) proposed a proactive Source End DDoS netWork Attack Recognition and Defense (DWARD). DWARD is one of the most promising DDoS defense techniques proposed. It uses a novel traffic profiling techniques and adaptive response to achieve autonomous attack detection and accurate response. DWARD controls the outgoing traffic to the victim, while ensuring that all legitimate traffic to the victim is detected and forwarded.
2.5 DDoS AND CORRELATION

Correlation is a statistical measure, which describes the degree of association or strength of relationship between two or more variables or sets of data that has been observed as a function of the time separation between them. The following section discusses some of the schemes proposed by researchers to detect DDoS attacks using correlation techniques.

Li et al. (2009) proposed an abnormal correlation analysis method for DDoS attack detection based on the directional property of DDoS attack flow. Every origin destination (OD) flow observation is compared with predefined predictions to determine the abnormality of a flow. Spatial correlation is then estimated for all OD flow with the same destination. An abrupt change of spatial correlation indicates the presence of a DDoS attack.

Yu et al. (2010) proposed the use of correlation coefficient to discriminate DDoS attacks from flash crowds. Flash crowd is a network phenomenon, where a network or host suddenly receives a lot of traffic.

Khatoun et al. (2008) proposed an alert correlation approach for DDoS intrusion detection using a decentralized architecture involving an overlay network.

Wang and Wang (2011) proposed the analysis of the distribution of network traffic for DDoS attack detection. The detection method is based on the analysis of the correlation of the IP addresses.

Zargar and Joshi (2010) proposed a collaborative approach to DDoS detection by which each router in the Autonomous System (AS) captures the correlation between flows destined for a specific destination. The sampling of network traffic, detection of DDoS attack and response
responsibilities is distributed among all routers in the AS, such that each router can individually detect and respond to a DDoS attack.

A sophisticated DDoS attacker can discover the statistical filtering rules applied by a defense system and then control the zombies to generate flooding traffic according to these discovered rules. Sun et al. (2008) proposed a principal components analysis-based robust DDoS defense system which extracts nominal traffic characteristics by analyzing intrinsic dependency across multiple attribute values. The PCA-based scheme differentiates attacking packets from legitimate ones by checking if the current traffic volume of the associated attribute value violates the intrinsic dependency of nominal traffic. The correlation among different attributes hinders the attacker from accurately discovering the statistic filtering rules and enables the defense system to handle more sophisticated attacks.

Jin and Yeung (2004) devised a covariance analysis model for detection of SYN flooding attacks. The proposed scheme is a statistics-based on-line detection method that uses the flags in the control field of TCP header as raw data for detecting DDoS attacks of different intensities.

Li and Lee (2003) proposed the use of Wavelet analysis to capture complex temporal correlation across multiple time scales to detect DDoS attack. DDoS attack traffic would cause significant energy distribution deviation in short time period. The proposed scheme utilizes the spikes in the energy distribution in the wavelet analysis to detect the presence of DDoS attacks.

Thapngam et al. (2012) proposed a supervised learning system based on Linear Discriminant Analysis (LDA) to discriminate legitimate traffic from DDoS attack traffic. Pearson's correlation coefficient and Shannon's entropy are deployed for extracting dependency and predictability.
of traffic data respectively. LDA is then used to train and classify legitimate and attack traffic flows.

Zi et al. (2010) proposed a novel, adaptive clustering method combined with feature ranking mechanism for DDoS attacks detection. Preliminary variables are selected based on the analysis of network traffic and Modified Global K-means algorithm is used to identify the cluster structure of the target data. Linear correlation coefficient is then used for feature ranking. The feature ranking result is used to inform and recalculate the clusters. The method is effective and adaptive in detecting the separate phases of DDoS attacks.

Xiang et al. (2005) proposed a correlation approach based on sequential pattern mining techniques. By mining the alert sequences and iteratively consolidating the matching sequential alert patterns, the scheme greatly eliminated the related redundant alerts and quickly identified the DDoS alert sequence.

Malliga et al. (2008) proposed a flow based scheme to detect the DDoS attacks that adapts itself to the changing trends of the current traffic. Information entropy, a measure to find correlation among traffic flows, is used to deduce the current state of the dynamic network and identify the suspicious traffic at the source end.

2.6 OTHER TECHNIQUES

Xu (2003) proposed a two-step approach using efficient cryptographic techniques to recognize traffic that needs to be protected during a high scale DDoS attack. The first step is to distinguish packets that contain genuine source IP addresses from those that contain spoofed addresses. The second step is to prevent such attackers from consuming too much system
resource by allocating adequate resources (e.g., bandwidth) to legitimate traffic separated by this process, thus ensuring efficient service to a large percentage of clients during DDoS attacks. The proposed two step approach is an effective and practical countermeasure that allows a victim system or network to sustain high availability during DDoS attacks.

Blackert et al. (2003) performed a Distributed Denial of Service Defense Attack Tradeoff Analysis (DDOS-DATA) to analyze Distributed Denial of Service (DDOS) attacks and mitigation technologies. Three mitigation technologies were analyzed: (1) Rate-limiting: limits the flow of predefined packet types by limiting certain resources. Example – allocate only a subset of the available bandwidth per traffic type. (2) Active monitoring: Active Monitors observe and classify traffic and then free resources consumed by attacks. (3) Proof of Work: Proof of Work (or client puzzle) technologies require client payment in the form of CPU cycles in return for using the server.

The analysis provided a detailed understanding of the interaction between DDoS mitigation technologies and attackers and to develop an understanding of how well mitigation technologies perform and how they can be combined to limit the potential attack space. This type of understanding is critical to breaking the reactionary cycle that currently exists between attackers and defenders.

Chen et al. (2007) proposed a Collaborative Detection strategy to detect DDoS flooding Attacks over Multiple Network Domains at the traffic-flow level. At the early stage of a DDoS attack, some traffic fluctuations are detectable at Internet routers or at the gateways of edge networks. This abrupt traffic changes is detected across multiple network domains at the earliest time. The defense system uses a Distributed Change-point Detection (DCD) architecture using Change Aggregation Trees (CAT) and is implemented over
the core networks operated by Internet service providers. The system is built over attack-transit routers, which work together cooperatively. Each ISP domain has a CAT server to aggregate the flooding alerts reported by the routers. CAT domain servers collaborate among themselves to make the final decision.

Lim and Uddin (2005) proposed an innovative and practical defense for built on network processor for source based defense against SYN flooding. The embedded architecture resides in upstream border routers and detects wide-range of attacks and blocks large portion of attack traffic before flooding into core network.

The architecture has a sampling, detection, and response module. The sampling and response components reside at data plane, where high-speed packet classification and processing tasks are performed. The detection and rate-limiting components involve complex computational jobs and is located at control-plane. Change-point detection algorithm is employed to detect occurrence of SYN flooding attack.

Results are passed to the fuzzy-based adaptive rate-limiting controller, whereby central alarm is immediately triggered once an IP address is found suspicious to be an attacker. Upon rising of alarm, the traffic policing component takes over the responsibility to inspect traffic flows of the suspicious IP address. It communicates with the fuzzy rate-limiting controller, which then enforces rate limits and ensures forwarding of legitimate packets. Under the per-flow mitigation scheme, while the attacker is penalized with limited outgoing connection, the legitimate clients in the same subnet are free from collateral damage.

Lee and Thing (2004) proposed a new technique called Port Hopping to detect and filter malicious traffic and improve the reliability of
good traffic flow. In this technique the server’s port numbers change dynamically as a function of time. Different port numbers are used in different slots for the same service. A cryptographic key is shared between the server and the client. When the client needs to communicate with the server, it determines the server’s current port number using the shared secret key and time slot number.

The server filters illegitimate packets by inspecting the port number contained in the UDP/TCP headers. The technique is compatible with the UDP and TCP protocols and can be implemented by using the socket communication for the UDP protocol and for setting the TCP communications. The strength of the mechanism lies in the simplification of both the detection and filtering of malicious attack packets and it also does not require any modifications to the existing protocols.

You et al. (2007) proposed a two distance–based DDoS detection techniques – average distance estimation and distance based traffic separation, which detect attacks by analyzing distance values and traffic rates. The distance information of a packet can be inferred from the Time To Live (TTL) value of the IP header. In the average distance estimation DDoS detection technique, the prediction of mean distance value is used to define normality. The prediction of traffic arrival rates from different distances is used in the distance–based traffic separation DDoS detection technique. The mean absolute deviation based deviation model is used to separate the normality from the abnormality for both the techniques and detect the attacks effectively.

Lai et al. (2008) proposed an adaptive bandwidth allocation approach to defend against DDoS attacks by monitoring the traffic pattern. The bandwidth allocation policy assigns normal users to a high priority queue and suspected attackers to a low priority queue. A traffic factor is tuned to
control the rate of the malicious flow and adjust the allocations of bandwidth according to the provider’s system resources. The proposed priority queue-based scheme, improves the effectiveness in network services and resources by effectively blocking the attack traffic while maintaining constant flows for legitimate traffic.

Xiang and Zhou (2006) proposed a Mark-Aided Distributed Filtering scheme using Neural Network for DDoS Defense. Neural network is applied to find network anomalies and are deployed at distributed routers to identify the attack packets and filter them. The system has two parts, Offline Training System which collects traffic characteristics and trains the neural network without influencing the normal operation of the network and Online Filtering System that provides the fast decision making function to find the attack signals. When the attack is confirmed, those packets with the same address digest bits are filtered out.

Schneider and Calmet (2006) proposed a hybrid intelligent DDoS defense called ‘Fibered Guard’ which uses logical fibering architecture and data mining approach. Information about each connection is saved in three tables – Global, Local and History. Connection specific data and general health state of the system are stored in the global structure. A data mining plugin is used for comparison and result analysis. In the first step attributes such as source IP, destination IP, port number and connection time and consequences such as bandwidth, CPU load and number of open connections are considered. A set of cases consisting of attributes and consequence is defined. In the second step irrelevant attributes in the above attribute–consequence cases are eliminated. Repetitions in the remaining raw data are analyzed and a simple set of rules is determined.

Potential attacks are detected by the consequence they have in the system – primarily a sudden scarcity of system resources. An analysis of the
data may reveal a large volume of similar (and probably faulty) connection, all established over a short period of time. The connections are classified as malicious and the system starts random dropping of some of these connections to free up the system resources.

2.7 DDoS NETWORK ATTACK RECOGNITION AND DEFENSE

DWARD is installed at the source router that serves as a gateway between the deploying (source) network and the rest of the Internet. DWARD is configured with a set of local addresses whose outgoing traffic it polices – its Police Address Set. This set identifies all machines in the stub network or all customers of an ISP. DWARD consists of observation, rate-limiting and traffic – policing components.

2.7.1. Observation Component

The Observation Component monitors all packets passing through the source router and gathers statistics on two-way communications between the police address set and the rest of the Internet. The statistics are gathered at the aggregate flow – agflow (the aggregate traffic between the police address set and one foreign IP address) and connection granularity (the aggregate traffic between a pair of IP addresses and port numbers, where one address belongs to the police address set and the other is a foreign address).

Periodically, statistics are compared to legitimate traffic models and agflows and connections are classified as attack or legitimate. The observation component passes the information on agflow classification and behavior to the rate-limiting component which decides to impose, modify or remove the rate limit based on the agflow’s sending rate.
2.7.2. **Rate Limiting Component**

The Rate Limiting Component uses a list of classified legitimate connections to selectively enforce these rules on suspicious traffic. DWARD observes how the agflow statistics react to defense actions and detect and recover from false positives very quickly.

DWARD’s rate limit strategy applies modified TCP congestion control for fast recovery from false positives. A fast exponential decrease of the sending rate is performed when the attack is detected to quickly relieve the victim of high-volume traffic.

Once the attack subsides, DWARD performs a slow recovery of rate-limited agflows, linearly increasing the sending rate. This is done to probe the state of the receiver and to reevaluate its ability to handle traffic. After a while, if the attack is not repeated, DWARD performs a fast recovery of rate limited agflows, increasing the sending rate exponentially.

2.7.3. **Traffic Policing Component**

The Traffic Policing Component periodically receives rate-limited agflow information from the rate-limiting component and connection classification information from the observation component. It uses this information to reach a decision whether to forward or drop each outgoing packet. Packets from nonlimited agflows and good connections are always forwarded. TCP packets from transient connections on limited agflows, whose sequence number matches the predicted value, are forwarded if the Early Packet Rate Limit for the agflow is not exhausted. Other transient-connection packets are forwarded if the agflow’s rate limit is not exhausted.
2.8 ORGANIZATION OF THE CHAPTERS IN THE THESIS

The thesis is organized as follows.

Distributed Denial of Service, its attack methods, tools and defense are discussed in chapter 1. The literature review in the chapter 2 analyzes the various approaches proposed by researchers to detect and mitigate DDoS attacks.

The location of the defense system in the network and the implementation of the defense system as a Snort preprocessor plugin module are discussed in the chapter 3. The architecture of Dynamic DDoS Defense with an adaptive Spin Lock Rate control mechanism (D3SLR), the DDoS flood attack detection strategy and implementation of the detection component as Snort Flood Detection Preprocessor is discussed in the chapter 4.

Alert correlation strategy proposed to prioritize the alerts and identify the critical alerts in the alert stream is discussed in chapter 5. The use of an adaptive Spin Lock Rate control mechanism and the performance evaluation of D3SLR using the DDoS attack traces from University of Southern California / Information Sciences Institute (USC / ISI) Analysis of Network Traffic (ANT) Dataset is the subject of analysis in chapter 6.

Conclusions and discussion followed by selective references are in the last part of the thesis.

2.9 CONCLUSION

An analysis of the DDoS defense methods proposed by various researchers shows the use of IP traceback as the most popular approach to DDoS defense. DDoS defense is done by tracing the origin of attack and...
taking the host system out of action. This is a slow process during which the victim site can do nothing to restore its service to legitimate clients. Also many of the proposed solutions require some information to be embedded in the packet header, based on which packets are forwarded or dropped. While these solutions show a high degree of efficiency in simulation environments, they cannot be implemented in real networks without requiring extensive changes and cooperation from ISP and intermediate routers, which is not practically possible.

Secure Overlay Services (SOS) proposed by Keromytis et al. is a novel architecture which proactively prevents DDoS attacks. Access requests will be authenticated and routed via an overlay network to one of the Servlets, which then forward the requests to the target site. The overlay architectures require a certificate for accessing a protected server to be issued to each authorized client. The robustness of these overlay architecture against a DDoS attack comes from the fact that, only authenticated traffic can enter the overlay network, the locations of the Servlets are not predictable and the target site only accepts packets from the Servlets.

The primary mitigation strategy used in DDoS defense systems are packet filtering and rate limiting. Filtering techniques discard packets that do not match specific conditions specified at the router. When properly configured and supported by network operators, these approaches can effectively prevent DDoS attack. However this approach is dependent on the cooperation and implementation by network operators and Internet Service Providers. All ISPs usually do not have strong incentive to implement the filtering mechanisms into their routers since it increases the overhead but has no direct benefit to their own clients.
2.10 AIM OF THE THESIS

To overcome these limitations generally found in detection techniques, this thesis proposes a solution which does not require any information to be embedded in the IP header and does not require any modification at the router to mark the packets. The proposed solution does not require any redirection of packets and can be effectively used against a general purpose server.

Dropping all packets from suspicious source(s), while effective in creating an immediate relief at the victim may also cause a great deal of legitimate traffic to be dropped. The most crucial feature of a flood mitigation mechanism not its ability to stop attack traffic but to protect the legitimate traffic.

Hence this thesis proposes rate limiting as a more effective alternative to packet filtering approach. Here instead of dropping all malicious packets, the number of identified malicious packets allowed to pass through a router is limited to a certain threshold. This serves a twofold purpose. If a legitimate traffic flow is wrongly identified as a malicious flow, a part of its traffic still reaches the destination. If the flow is malicious, the packets reaching the target can be logged and later used for analysis.

Correlation techniques have been used by many researchers for DDoS attack detection based on the traffic flow characteristics. However this thesis proposes the use of autocorrelation technique to determine the strength of each individual malicious flow and based on the time duration during which the attack persists, an adaptive rate limit is applied on the malicious traffic flow towards the victim. The proposed architecture is discussed in the following chapter.