A review of detection approaches for distributed denial of service attacks

Parneet Kaur, Manish Kumar and Abhinav Bhandari

Department of Computer Engineering, Punjabi University, Patiala, India

ABSTRACT

Distributed Denial of Service (DDoS) attacks are the intimidation trials on the Internet that depletes the network bandwidth or exhausts the victim's resources. Researchers have introduced various defense mechanisms (such as attack prevention, traceback, reaction, detection, and characterization) against DDoS attacks, but such attacks are still growing year by year, and the ideal solutions of this problem are eluded so far. In the past, various signature-based and anomaly-based approaches were introduced for the detection of DDoS attacks, but only a few of them have focused on the nature of anomalies. Most of the detection approaches do not provide efficient real-time detection with high detection rate and low faux pas. In this paper, a classification of detection approaches against DDoS attacks has been presented with an aim to go deep insight into the DDoS problem for the beginners in this research area. The detection approaches have been explained along with their pluses and minuses. Further, this review paper includes the different functional classes to which the detection approaches belong to. In the end, a comparison of signature-based, anomaly-based and hybrid detection approaches is depicted in tabular form.

1. Introduction

In the present era, the services like banking, electronic commerce, social networking (chat rooms) and newsgroups are directed through the Internet (Zhou, Leckie, & Karunasekera, 2010). Denial of Service (DoS) attacks may impede the rise and continuity of these Internet-based applications. It disrupts or degrades the network services (by depleting the network bandwidth or router processing capacity) or victim resources (by exhausting disk or database bandwidth, file descriptors, buffers, sockets, CPU cycles, memory) and stops the legitimate user from accessing a specific Internet service (Saman & Tipper, 2013). Such attacks hog the victim's resources so that it cannot respond to the services requested by an authenticated user.

Distributed Denial of Service (DDoS) attacks are the global attacks and have become a severe problem of today's Internet. DDoS attacks are adroit in nature that follows the same techniques as regular DoS attacks, but performs the attack on a much larger scale through botnets (Douligeris & Mitrokotsa, 2004) as shown in Figure 1. A botnet is a wide chain of hundreds or thousands of remotely controlled compromised hosts (zombies or bots or slave agents) under the control of one or more intruders to attack a particular victim.

Every computer connected to the Internet is an attractive target for attackers for making bots or zombies, even if the user does not know about it. Zombies are enrolled through the use of worms, backdoors or Trojan horses by sending an e-mail content, a captivating link, or a trust-inspiring sender address to the vulnerable machines (Prasad, Mohan, & Rao, 2014; Saman & Tipper, 2013). Sometimes, the data originates from a single bot is very small, but the cumulative traffic from a sufficient number of bots emerging at the end user's system is enormous that exhaust its resources. Therefore, Low-rate DDoS (LDDoS) attacks are devastating and harder to expose as the traffic appears to be normal that a particular link can control (Zhang, Cai, Chen, Luo, & Yin, 2012). On the other hand, High-rate DDoS (HDDoS) attacks are quickly recognized with the prevailing detection methods. Nowadays, DDoS attacks are conducted in the form of packet flooding and link flooding attacks. Such attacks have increased on the Internet because the attacker knows what information can be obtained where and how. Due to the presence of vulnerabilities in Internet protocols, web applications, and operating systems, it becomes easy for the attacker to launch such attacks. Such attacks are performed with the motives like hactivism (to generate media attention), to gain profit through extortion.
(like blackmailing), personal reasons (like revenge or disputes), economical reasons (nastiness) and due to political reasons (Prasad et al., 2014). The most common targets for these dreadful assaults are the gaming, media, web applications and software industries.

The rest of this paper is organized as follows: Section 2 describes the background of DDoS attacks and need of detection against DDoS attacks. Section 3 classifies detection approaches according to their functionalities. Section 4 provides the comparison of detection approaches. Section 5 presents the various issues of existing detection approaches. Finally, Section 6 concludes the paper and Section 7 presents the future directions in this research area.

2. Background and motivation

This section presents the history as well as the basic strategy of DDoS attacks. It also depicts the need of defense mechanisms against such attacks. DDoS attacks are not new offenses against web applications (Li, Kao, Zhang, Chuang, & Yen, 2015). Initially, DDoS attacks were launched in August, 1999 against different organizations and continued attacking the various websites like Yahoo, Amazon, Buy.com, CNN and eBay since then (Bhuyan, Kashyap, Bhattacharyya, & Kalita, 2014; Burago-hain, Kalita, Singh, & Bhattacharyya, 2015). In 2009, a DDoS attack was launched that disrupted the network services of most popular websites like Live Journal, Facebook, Amazon, and Twitter (Acohido & Swartz, 2009). In 2010 and 2011, more than 75,000 computer systems in 2500 organizations and 4 million computers in 100 countries were affected by DDoS attacks respectively (Li et al., 2015). Each day, more than 7000 DDoS attacks are launched by the attackers (Mousavi, 2014). The average attack volume reached at 48.25Gbps in the first quarter of 2013, which is 718% more as compared to the last quarter of 2012 (Mousavi, 2014). In recent times, DDoS attacks have become smaller in duration. According to (Questions to ask your DNS host about DDoS), the largest recorded DDoS attacks have grown 1000% since 2008, from 40 Gbps to 400+ Gbps in 2013 and such attacks happen at an average rate of 3000 times a day. According to a survey by Verisign, there is an increase of 111% in DDoS attacks every year (Verisign). Verisign mitigated 85% more attacks in the fourth quarter of 2015 as compared to fourth quarter of 2014 (Bisson). In 2015, the largest attack was about 500 Gbps that disrupted an entire ISP’s network of the country of Kenya (Baraniuk). DDoS attack was conducted against the BBC website to 602 Gbps in the first quarter of 2016 (Khandelwal). According to the records in (Woolf, 2016), the largest DDoS attacks in the history was orchestrated in October,2016 using a new Mirai botnet against the servers of an American company named as Dyn, that steer much of the Internet’s Domain Name System (DNS) infrastructure. Mirai was the primary source of pernicious attack traffic. Unlike other botnets, Mirai botnets had used the Internet of Things (IoT) devices such as digital cameras and DVR players to bring down the websites (including Twitter, Netflix, the Guardian, CNN, Reddit, and many others) in Europe and US. According to the estimates of Dyn, the attack had prodigious attack strength of 1.2 terabits (1200 gigabytes) per second and had intricate ‘100,000 malicious agents’. As stated in (Bhandari, Sangal, & Kumar, 2015), the average strength of DDoS attacks has been shown in Figure 2.
Basically, DDoS attacks are of two types namely flooding attacks and vulnerability attacks (Nadiammai & Hemalatha, 2014; Prasad et al., 2014) as described in Figure 3. In flooding attacks, the attacker sets a zombies army to send junk or attack packets to the destination in order to raise the traffic to a level that a victim cannot handle and the victim system brings down or crashes (Prasad et al., 2014). On the basis of attack mechanism, (Aggarwal & Gupta, 2015) categorize the flooding attacks into direct and indirect (through reflectors) DDoS attacks. On the basis of protocol level that is targeted, flooding attacks are grouped into Network/Transport level (Net-DDoS attacks) and Application level DDoS flooding attacks (App-DDoS attacks) (Saman & Tipper, 2013). The attacks like TCP, UDP, and ICMP flooding comes under the category of Net-DDoS flooding attacks while HTTP flooding comes under App-DDoS flooding attacks. In (Xie & Yu, 2009a, 2009b), the authors introduce the App-DDoS attacks and discuss the incapability of network-level detection methods for catching the App-DDoS attacks. These attacks are growing rapidly, harder to detect and cause severe problems in accessing a particular online service (or web server) as compared to the Net-DDoS attacks. In vulnerability attacks, the attacker browses for unprotected openings in the software implementation and exploits them to bring the system down or to recruit zombies for further attacks. These attacks use the exacted performance of different protocols (such as TCP and HTTP) to ravage the resources of the victim server and prevent it from processing events or requests from the authorized users.

2.1. Defense mechanisms

DDoS defense system consists of four phases: Attack prevention, Attack detection and characterization, Traceback and Attack reaction (Douligeris & Mitrokotsa, 2004). Attack prevention tries to suspend attacks before they actually occur. Attack prevention schemes fix security holes in Internet hosts like weak protocols, inadequate authentication plans, unprotected computer systems and operating systems that are used to originate a DDoS attack (Gupta, Misra, & Joshi, 2012). Prevention schemes are the best solutions to secure the weak systems as well as the protocols and are needed to raise the global security level (Gupta et al., 2012). Attack detection schemes expose the attacks after they actually happen and characterization departs the attack packets from the normal traffic, for example, signature-based and anomaly-based detection techniques. Traceback schemes attempt to hinder the malicious packets at their origin and find the attack source despite the spoofed source IPs before or after the attack occurrence. Attack reaction schemes try to minimize the loss caused by DDoS attack when it is underway. This phase decreases the attack influence and maximizes the quality of services delivering to the legitimate users under attack.

2.2. Detection

When a system is under DDoS attack, unexpected fluctuations in the network traffic are noticed. Detection systems are software or hardware products that automatically monitor the abrupt changes in the network and analyze
the detection process (Gyanchandani, 2012). In the detection phase, DDoS attacks are detected and legitimate packets are distinguished from attack packets. Detection methods recognize DDoS attacks with the directory of known (or familiar) attack patterns or by identifying irregularities in standard network behaviour (Douligeris & Mitrokotsa, 2004). Due to the lack of clear DDoS attack profile or signature, the detection schemes observe unexpected shifts in IP packet traits or traffic volume to catch these appalling attacks. Attack detection methods erect a model or profile by observing the regular functioning of the interface, validate the incoming flux against the paradigm and discover oddities with the perpetual shifts in the network. The detection approaches can be implemented locally, to protect a particular victim or remotely, to expose propagating attacks in the core network. Early detection and the detection accuracy of DDoS attacks have become the critical measures for the realization of a defense system. So, every detection technique should outline normal traffic intelligently, accurately, and recognize aberrations with high normal packet survival ratio, low false positive and false negative ratios and it should be cost effective in terms of resource consumption and per packet computations.

3. DDoS attacks detection: approaches, functional classes and metrics

This section introduces the classification of different detection approaches and their functional classes. A brief introduction to different metrics has been depicted in order to compare the detection methods.

3.1. Approaches

This sub-section begins with the review of existing approaches to DDoS attack detection. Though a diversity of detection approaches has been proposed in the research in preceding years, but the security tools with detection capabilities have several important obstacles that remain to be solved. The choice of detection approach totally depends on the various factors such as the type of anomalies, processing data type and behaviour, working environment of the organization, computational cost, and the required security level (Raut & Singh, 2014). Moreover, the performance of the detection schemes depends on how well it is operated and tested on all network protocols. Nowadays, Soft Computing or Artificial Intelligence-based methods are applied extensively for the attack detection (Singh, Hans, Kumar, & Singh, 2015). On the basis of analysis methods, detection approaches are classified into Signature-based, Anomaly-based and Hybrid detection (Agarwal & Mittal, 2012; Wu & Yen, 2009) as described in Figure 4.

3.1.1. Signature-based detection

Signature-based detection is also known as Misuse detection, Pattern detection, Knowledge-based or Rule-based detection. This approach captures the required behaviour.
from the available datasets (such as protocol stipulations, network traffic occurrences) and also collects facts about various attacks and system exposures. It uses the acquired expertise to identify the occurrence of anomalous events and produces an alarm if an attack is identified (García-Teodoro, Díaz-Verdejo, Macia-Fernández, & Vazquez, 2009). It uses an index of ‘signatures or patterns’ of public attacks and matches the incoming traffic with the stored patterns to identify the attack instances (Agarwal & Mittal, 2012). This approach is effective only in case of known attacks because new attacks or slightly modified old and known attacks go unrecognized as they don’t have signatures in the database (Chauhan, Mishra, & Kumar, 2012; Lee & Xiang, 2001). The signatures and patterns of various attacks consist of several fields of IP packets (source and target IPs, ports and the keywords in the payload of a packet) (Xia, Qu, Hariri, & Yousi, 2005). Such systems are too slow and depend on the fault conditions (or behaviour) of the victim system. Fault conditions may arise due to a large number of open TCP connections, excessive utilization of bandwidth and due to the exceeded total throughput (Thottan & Ji, 2003). SNORT (Gupta et al., 2012), BRO (Gupta et al., 2012), IDES (Xia et al., 2005), and INBOUNDS (Xia et al., 2005) techniques are based on signature-based detection approach. Nowadays, this approach is followed only by the network administrator (Nadiammai & Hemalatha, 2014). Wu, Tseng, Yang, and Jan (2011) present a detection method consists of classification trees and a traffic pattern matching algorithm. The classification trees separate the DDoS traffic from normal traffic after analyzing the incoming and outgoing packet rate, transmission rate, TCP, SYN and ACK flag rate. A pattern matching algorithm is used to detect the traffic flow that is identical to attack flow and looks for the origin of the attack. In (Limwiwatkul & Rungsawang, 2004), the authors analyze the TCP/IP packets against some well-marked rules and conditions to distinguish the attack and regular traffic. In (Thapngam, Yu, Zhou, & Beliaakov, 2011), the authors use the transmission rate to recognize the attack traffic. This study shows that transmission rate of attack traffic is high as compared to real network traffic because the slave agents under the command of their masters generate the attack traffic in a very short time frame while the regular traffic waits for the server’s response thereby extend the time period. Such methods cannot detect efficiently because the attackers can easily send the mimic attack traffic towards the victim using flash events. Thomas, Mark, Johnson, and Croall (2003) describe detection scheme called NetBouncer that prepares a database of legitimate users. If the incoming packets do not belong to a legitimate client, then the packets need to prove their legitimacy through a series of legitimacy tests. The new clients are added to the database if they pass the test and send a sufficient number of packets until a session window expire. Various methods are used to construct attack signatures with the help of State Transition Analysis, Expert Systems, Petri Nets, Description Languages and Adept Systems.

3.1.1. State transition analysis. This approach represents the attack patterns with state transition diagrams (Wu & Yen, 2009). In this approach, an attack is viewed as a chain of actions from an initial state of the system to the target compromised state (Gyanchandani, 2012). It uses a list of key actions which are required for launching a favourable attack (Raut & Singh, 2014). The complexity of the system rises with the increase in a number of states and parameters that do not remain constant over the network. It is an offline-detection mechanism against DDoS attacks. Wang, Phan, Whitley, and Parish (2010) propose an Augmented Attack Tree (AAT) based anti-DDoS model that captures the network behaviour from the victim server and transforms into state transitions to detect different types of attacks.

3.1.1.2. Expert systems. This approach builds a set of rules to specify the well-known attacks and draw some conclusions from the rules and facts. Then the incoming traffic instances are matched against the rules to check whether any rule is satisfied and detect the inconsistent behaviour of the system (Thottan & Ji, 2003). But this approach requires re-building the rule frequently to assist the newly discovered vulnerabilities (Gyanchandani, 2012).

3.1.1.3. Petri nets. In this approach, complex attacks signatures are written manually by the system administrators with IDIOT tool (Gyanchandani, 2012). The approach is conceptually simple to build attack signatures and the signatures are represented in graphical form. It is computationally very expensive to match the complex signatures with new traffic instances.

3.1.1.4. Description scripts. Various scripting languages are used to describe attack signatures on the system and networks. Scripting languages identify the series of distinct events that are representative of various attacks (Rama, 2011).

3.1.1.5. Adept systems. This approach uses the human expertise to solve the problem of uncertainties in attack signatures. Adept systems are prepared based on comprehensive knowledge of signatures linked with popular
attacks that are presented by specialists from their past experience (Rama, 2011).

### 3.1.2. Anomaly-based detection

Anomaly-based detection approach (also known as novelty detection, outlier detection, behaviour based or one-class learning scheme) is capable of detecting new, unknown and novel (unidentified) attacks. This approach mirrors the standard network behaviour and compares it with the incoming data instances (Alenezi & Reed, 2012). When the divergence between an observed and expected behaviour surpasses a predefined threshold, the detection system generates an anomaly alarm; hence an attack is disclosed (Garcia-Teodor et al., 2009; Xiang, Li, & Zhou, 2011). Anomaly-based schemes produce a lot of false signals due to the varying nature of a system or network behaviour and uncertainties present in the acquired data. The input to a detection approach can be in the form of individual data instances (such as an object, vector, point, observation (Chandola, Banerjee, & Kumar, 2009) or a collection of data instances. Data instances may or may not relate to each other. Each input instance has a set of attributes and each attribute can be discrete, categorical or continuous in nature. Most of the detection schemes deal with the individual input instances in which there is no relationship among the different instances (Chandola et al., 2009). On the basis of nature of anomalies, detection approaches are sub-categorized into Point anomaly, Contextual anomaly, and Collective anomaly-based detection. Collective anomaly detection approach has become the most challenging research field as compared to point and contextual anomaly detection.

#### 3.1.2.1. Point anomaly-based detection

If a single data instance is considered as an anomaly as compared to the remaining dataset, then the approach is known as Point anomaly-based detection. Nowadays, it is the most significant and interesting field of the research on anomaly-based detection. Various approaches have been adopted to recognize point anomalies in the network traffic namely Statistical Methods, Data Mining, Artificial Intelligence (Al Based, Information Theoretic Based, Nearest Neighbour Based Detection) which are described below:

- **Statistical methods**: Statistical methods used in anomaly detection systems prepare a model (Chandola et al., 2009) or normal profile (Esposito, Mazzariello, Oliviero, Romano, & Sansone, 2007; Prasad et al., 2014) to represent the assumption behaviour of a system (or network) and continuously monitors the bi-directional traffic runs between the victim network and the rest of the Internet in on-line as well as off-line detection mode (Xie & Yu, 2006; Xie & Yu, 2009). This is basically done by measuring statistical properties (i.e. means and variances) of various parameters of normal traffic (like activity measures i.e. login and logout time for each session, traffic rate, CPU time used, packet rate for each protocol and the number of different IP addresses). A statistical inferring analysis (like $\chi^2$-test) is practiced to decide if an unseen case refers to the statistical basis or not. The test examples having a low probability (i.e. certain thresholds or baselines are not met) are declared as anomalies (Chandola et al., 2009; Garcia-Teodor et al., 2009; Gyan-chandani, 2012; Lazarevic, 2016). Moreover, the detection method assigns a score to each anomalous activity. If the anomaly score exceeds the baseline, the system generates an anomaly alarm. Statistical methods are deployed at any network (source-end, victim-end, and core-end network) for the discovery of Net-DDoS attacks (Mirkovic, Prier, & Reiher, 2002; Nguyen & Choi, 2010; Prasad et al., 2014) as well as App-DDoS attacks (Jin & Yeung, 2004; Thottan & Ji, 2003; Xie & Yu, 2009). SSM (Prasad et al., 2014), CAT-DCP (Chen, Hwang, & Ku, 2007), ARIMA model (Zhang, Jiang, Wei, & Guan, 2009) are widely used statistically based detection technique. A batch detection method has been proposed in (Blazek, Kim, Rozovskii, & Tartakovsky, 2001), to identify attack instances by analyzing the statistical changes. D-WARD (Bhuyan et al., 2014; Mirkovic et al., 2002) and MULTOPS (Gil & Poletto, 2001) techniques offer the features of filtering and rate-limiting on incoming traffic at the source-end. On the other hand, COSSACK (Papadopoulos, Lindell, Mehringer, Hussain, & Govindan, 2003) and DefCOM (Mirkovic & Reiher, 2005) techniques detect the flooding attacks at the victim-end and inform the filters or the rate limiters installed at the source-end. In (Chen & Song, 2005), the authors introduce a perimeter-based method for ISP located on boundary routers to detect the attack generator. CUSUM scheme discussed in (Alenezi & Reed, 2012; Carl, Kesidis, Brooks, & Rai, 2006) observes the unstable fluctuations in the traffic from the long-term network performance. When the aggregate difference exceeds the threshold, the system generates an anomaly alarm. In (Peng, Leckie, & Ramamohanarao, 2004; Wang, Zhang, & Shin, 2004), the authors suggest the source IP based detection methods that monitor the changes in network traffic behaviour at the gateway level. Statistical methods are categorized into Parametric and Non-parametric detection.

- **Parametric detection**: Parametric methods assume that the system has experience of latent distribution and assess the statistical specifications from the given data (Garcia-Teodor et al., 2009). The techniques like Statistical Moments, Operational (or Threshold Based) Model, Gaussian Model, Regression Model and Spectral Analysis comes under the category of parametric detection (Chandola et al., 2009).
(a) Statistical moments

In this scheme, a specified confidence range or an interval is set based on statistical properties (correlations or moments) like statistical mean, standard deviation. If any event drops outward the set interim i.e. above or below the moment, it is declared as anomalous. This scheme offers more flexibility as compared to operational model because the confidence range depends on the observations that can vary from user to user (Rama, 2011). Therefore, it gives higher weights to the recent activities.

(b) Operational (or Threshold based) model

In this detection scheme, we compare the given observation (or an event) with a predefined limit (upper limit is $n$ and the lower limit is $m$ or 0). When the count of events that occur during a particular period is more than $n$ or less than $m$ then the detection system generates an alarm and detects the anomalies. For example, when the count of password failure exceeds the threshold, the system results in failed log-in (Gyanchandani, 2012; Rama, 2011; Raut & Singh, 2014). Moreover, the threshold is based on the mean of various parameters or metrics. This scheme is effective only if there are not any intermittent variations in standard data behaviour and the tolerant level of an appropriate event needs to be recognized in advance (Gyanchandani, 2012; Islam & Jamil, 2005). If the malicious activities have more than one event or the threshold limits are not significant, the scheme cannot detect anomalies efficiently.

(c) Spectral analysis

High-dimensional datasets are used for the detection purpose. It is very difficult to store, process, transmit and understand the huge amount of multidimensional datasets. It makes the detection technique more complex and expensive. In order to handle the multidimensional datasets effectively, spectral analysis approach is used (Alenezi & Reed, 2012; Patcha & Park, 2007). This approach converts a collection of attributes of correlated variables into a reduced set of linearly uncorrelated principal components and outlines the variations in the collected data (Chandola et al., 2009; Xie & Yu, 2009). It transforms a high-dimensional space into lower dimensional subspaces (such as embeddings or projections), where normal and abnormal behaviour is represented differently and anomalies are easily identified. This approach is also termed as Signal processing based detection or Wavelet analysis (Purwanto, Kuspriyanto, Hendrawan, Rahardjo, 2014). Principal Component Analysis (PCA) technique detects peculiarities in multivariate time-series data (Munz & Carle, 2007). Cheng, Kung, and Tan (2002) introduce a spectral analysis based detection method to identify the attack flows. An anomaly detection procedure on the basis of wavelet transmutation and the probability assumption has been introduced in (Li & Li, 2009). A two-phase automated detection method has been presented in (Dainotti, Pescape, & Ventre, 2009) that consist of the change point detection technique and consecutive wavelet alterations to identify the exceptional profiles in the network traffic. In (Bhuyan et al., 2014), the authors introduce an energy distribution based wavelet analysis for the exposure of DDoS attacks.

Non-parametric detection: In this approach, the system has no knowledge of potential distribution and it prepares a model from the given observations (Chandola et al., 2009). Various approaches like Markov Model, Histogram-Based and Time Series Model come under the category of non-parametric detection.

(a) Markov model

This approach uses an event counter metric to ascertain the consistency of particular events on the basis of the previous event and uses a state transition matrix to define the likelihood of an appropriate event (Gyanchandani, 2012; Raut & Singh, 2014). This scheme works by scrutinizing the system at regular interludes and sustains an account of its state. Each observation in the system is treated as a state and whenever an event occurs the system changes its state (Rama, 2011). When the system’s state changes, if the computed probability of the occurrence of that particular state is miniature at a given instant then that situation is thought as unusual (Rama, 2011). In this model, computed state change possibilities are the isolated parameters to identify anomalies and system state is directly perceptible. It looks for the transitions between certain activities or commands where strings of activities are particularly significant. It is ineffective for real-time services where a considerable large traffic and event rates occur in high-speed networks (Patcha & Park, 2007).

(b) Histogram-based detection

Frequency based and counting based histograms are used to formulate a profile of normal data. A test instance is considered as normal if it comes in any of the containers (or bins) of the histogram, otherwise, it is treated as anomalous. An anomaly score is ascribed to a test case based on the frequency or height of a particular bin to which the instance belongs to. If the test instances slip outside or in the empty region in case of small size bins, it
results in high false alarm rate. If the anomalous instances drop in the bins of the histogram in the case of large size bins, it results in high false negative rate. Therefore, optimal size bins should be preferred to construct the histograms in order to overcome the high false alarm ratio and false negative degree. Other variants of histogram based non-parametric statistical detection are IS Statistics, Packet Header Anomaly Detection (PHAD), Application Layer Anomaly detection (ALAD) (Chandola et al., 2009). This scheme is used for system call based and web-based anomaly detection.

(c) Time series model

This scheme includes an interlude timer, along with an event counter (or resource measure) as in the case of threshold scheme and a statistical database is prepared that reckons the order, the inter-arrival moments as well as the values of observations (Garcia-Teodoro et al., 2009; Gyanchandani, 2012; Rama, 2011). The observations (observed traffic) with low possibilities of their occurrence are viewed as anomalous. In this method, anomalies are the data points straying from normal patterns. The detection system measures the network behaviour over time and detects various shifts in behaviour. Therefore, when the attacks are performed in the form of series, they are easily detected. When swift changes in the common network behaviour occur due to anomalous conditions, the scheme cannot detect effectively (Gyanchandani, 2012). Cabrera (Cabrera et al., 2001) introduce a detection method using time series analysis that consists of a correlation process and statistical tools like Granger Causality Test (GCT) and Auto-Regressive Model for the identification of DDoS attacks.

Data mining: Data mining approach is based on ‘pattern finding’ (Garg & Chawla, 2011). It uses the statistical model to extract the useful, previously ignored (or hidden) patterns and their relationships, from the large data stores or entire domain (Aggarwal & Gupta, 2015). It decreases the amount of data that must be focused to detect (or uncover) the real attacks (Dickerson & Dickerson, 2000; Gyanchandani, 2012; Narayana, Prasad, Srividhya, & Ranga, 2011; Raut & Singh, 2014). This approach offers high detection accuracy when combines with Artificial Intelligence or Machine Learning methods. Data mining methods are categorized into Clustering, Classification and Associative rule mining based detection.

Classification: This approach tries to divine the class of a new, previously unseen data instances on the basis of class-labelled training dataset and a decision tree (or classification tree) (Ektefa, Memar, Sidi, & Affendey, 2010) is adopted to analyze each instance as normal or malicious (Gyanchandani, 2012; Raut & Singh, 2014). Top-down and bottom-up approaches are used to build a decision tree. This method rests on the availability of correct labels for standard class instances, which is not feasible. Therefore, the label assigned to each test instance shows some difficulty in assigning an anomaly score to each test case (Gyanchandani, 2012). Multi-class techniques use persuasive algorithms to separate the instances referring to distinct classes. In (Lee, Stolfo, & Mok, 1999), the authors develop a classifier that extracts the system features to represent the programme and user behaviour in order to recognize the anomalies in the network traffic.

Clustering: In clustering data mining techniques, the system finds the hidden patterns from unlabelled data with different proportions (number of attributes). It is based on natural grouping of similar data instances. The entities or records that do not belong to any of the clusters (i.e. by-product of clusters) are treated as an unusual activity or an attack (Barbara, Wu, & Jajodia, 2001). Complex data types are handled properly by the supporting clustering algorithms (Gyanchandani, 2012). As stated in (Chandola et al., 2009), several clustering algorithms like DBSCAN, ROCK, SNN, FindOut, WaveCluster algorithms have been introduced to detect the normal clusters and the residual instances are treated as anomalous. Besides these, various clustering-based methods namely Grid-based, Model-based, Density-based, Partitioning and Hierarchical clustering techniques were discussed in (Pei, Upadhyaya, Farooq, & Govindaraju, 2004). It has been mentioned in (Jin & Yeung, 2004) that clustering methods are applied to extract the HTTP-based flash passes from the App-DDoS attacks traffic. But, this approach is ineffective in case the new instances (anomalies) form their own clusters among themselves (Chandola et al., 2009; Gyanchandani, 2012).

Associative rule mining: Nowadays, this approach is not much popular and it is being displaced by other data mining methods. This approach discovers the anomalies by analyzing the correlation between different attributes (Barbara, Couto, Jajodia, & Wu, 2001; Chauhan et al., 2012). It is based on Boolean association rules and finds the regularities between the attributes. In case the system manipulates a large number of attributes, the detection process becomes slow in execution time (Tajbakhsh, Rahmati, & Mirzaei, 2009). Moreover, processing an abundance of rules is a challenging task for this approach.

Artificial intelligence-based detection: In this approach, the detection system can change its execution procedure on the basis of recently collected data (Patcha & Park, 2007). The system can improve its performance on certain test cases on the basis of prior results. This approach coincides with data mining methods or statistical methods which focus on getting the rules that generate the new data (Garcia-Teodoro et al., 2009; Patcha & Park, 2007). It
are discussed below:

Analysis

rithms, Support Vector Machines, Bayesian Networks, Fuzzy Logic Approach, Genetic Algorithms, Classification Trees, Local Models, Hidden Markov Models and Reinforcement Learning (Wu & Yen, 2009). Different AI-based detection approaches namely Neural Networks, Bayesian Networks, Fuzzy Logic Approach, Genetic Algorithms, Support Vector Machines and System Call Sequence Analysis are discussed below:

**Neural networks:** Neural networks are introduced as an alternative to statistical methods that divine the subsequent command on the basis of a series of previous commands from a particular user. Neural networks are well trained, purely feed forward and back propagation networks that give better results as compared to basic signature matching methods (Gyanchandani, 2012). This approach does not expect any explicit user model to predict the user’s behaviour (Raut & Singh, 2014). In this approach, neurons are trained with the data collected from the audit logs of various users for a particular period. This is basically done to represent the characteristic patterns of normal traffic. Whenever the incoming network traffic is served to the prepared neurons if its range exceeds a preset threshold then the system generates a signal; hence an anomaly is detected (Chandola et al., 2009; Patcha & Park, 2007). The reconstruction error (i.e. actual output minus desired output) is undeviatingly accepted as an anomaly score for detecting anomalies. *Hopfield Networks, Radial Basis Function (RBF) Based (Karimazad & Faraahi, 2011), Replicator Neural Networks (RNN), Linear Vector Quantization Artificial Neural Networks (LVQ-ANN), Back Propagation Neural Networks (BP-ANN), Resilient Back Propagation (RBP) Neural Networks and Time Delay Neural Networks (TDNN)* are used for anomaly-based DDoS attacks detection (Prasad et al., 2014). Neural networks detect anomalies from limited, noisy, imprecise or uncertain information and recognize the future unseen patterns along with previously observed attack patterns. Neural networks are deployed at victim-end networks and operate in supervised (Buragohain et al., 2015) as well as unsupervised mode (Jalili, Imani-Mehr, Amini, & Shahriari, 2005; Prasad et al., 2014). But it is very costly and time-consuming process as it needs extra time for collecting and analyzing the training data (or neurons). Jalili et al. (2005) propose a detection method called SPUN-NID that consists of a statistical pre-processor to extract the traffic features and unsupervised neural networks to differentiate the attack and regular traffic. RBF based neural networks discussed in (Karimazad & Faraahi, 2011), offers the features of robustness, parallelism, and tolerance of imprecision, faults, and uncertainty (Prasad et al., 2014). Machine Learning and Soft Computing methods are the sub-areas under the Artificial Intelligence-based (Al-based) approach. Machine learning includes the technologies like Bayesian Decision Theory, Linear Discrimination, Multivariate Methods, Multilayer Perceptrons, Clustering, Classification Trees, Local Models, Hidden Markov Models and Reinforcement Learning (Wu & Yen, 2009). Different AI-based detection approaches namely Neural Networks, Bayesian Networks, Fuzzy Logic Approach, Genetic Algorithms, Support Vector Machines and System Call Sequence Analysis are discussed below:

**Bayesian networks:** Bayesian approach for DDoS attack detection is used as a combination of Bayesian networks along with statistical methods (Kruegel, Mutz, Robertson, & Valeur, 2003). Bayesian networks are used to predict outcomes or discover cause-effect relationships when the system has uncertain or incomplete knowledge of network traffic. It develops a graphical model that encodes the probabilistic correlations (or conditional interdependencies) between distinct variables and predicting events (Chauhan et al., 2012; Wu & Yen, 2009). It captures all the existing knowledge of network traffic and represents the uncertain knowledge in expert systems. This new technique is still evolving and mostly used for solving data-analysis problems. Bayesian networks are also known as belief networks, Bayesian belief networks, and causal probabilistic networks. This method is induced to both univariate as well as multivariate datasets (Chandola et al., 2009), enhances the capacity to expose new attacks and lessens the false alarms to the possible extent. It is an efficient and principled approach to combine both prior knowledge and data as well as avoids the over-fitting of data (Gyanchandani, 2012).

**Fuzzy logic approach:** The concept of fuzziness is used along with data mining methods for highlighting anomalies or network attacks (Garg & Chawla, 2011). It is based on fuzzy set hypothesis under which reasoning is estimated rather than accurately procured from classical predicate logic (Chauhan et al., 2012; Harjinder, 2013). It uses fuzzy sets and fuzzy rules to handle a large number of input parameters (CPU usage time, activity rate, connection interval) that can be hazy in nature and incomplete datasets (Dickerson & Dickerson, 2000). Fuzzy systems effectively combine the inputs from various sources and construct if-then rules to describe security attacks (Eskin, Arnold, Prerau, Portnoy, & Stolfo, 2002; Raut & Singh, 2014). In (Shiaeles, Katos, Karakos, & Papadopoulos, 2012), the authors introduced the Fuzzy estimators that detect DDoS attacks using mean packet inter-arrival times and also find the offending IP addresses in real-time with high detection accuracy. It is an effective approach against port scans and probes. It relies on attacking explicit rules for detection rather than building a model for depicting the current status of the system. A fuzzy reasoning based approach along with statistical analysis using wavelet transformation and Schwarz information criterion has been introduced in (Xia, Lu, Li, & Tang, 2010) that can detect the DDoS traffic accurately and effectively.
Genetic algorithms: Genetic algorithms are heuristic quest algorithms based on the development thoughts of natural selection and eugenics tools to find the approximate solutions or to determine the optimization enigmas. It uses the evolutionary algorithm techniques like selection, crossover (mating or recombination), mutation, inheritance and elitism (Li, Guo, Tian, & Lu, 2008). It follows the postulate of ‘survival of the fittest’ so whenever many users attempt to obtain the scanty resources, the fittest users dominate the weaker one. A series of iterations are performed to replace the low fitness users with the help of a fitness function. Genetic algorithms are proficient in acquiring classification rules with the knowledge collected from incoming traffic and select optimal parameters for detection process to differentiate the attack passes from normal data (Harjinder, 2013). It selects the fine test cases as the instructing dataset and minimizes the false positive rates when human input is used in a feedback loop. It is a flexible and robust approach because, in the presence of noise or changing inputs, it is not easily affected. The measures like detection rate, false positives and the ratio of reduced training dataset are combined in a fitness function. So, the system should take care to raise the fitness function defined (i.e. increase the detection rate and decrease the false positives and instances in training dataset) (Li et al., 2008; Patcha & Park, 2007). This approach involves an assemblage of agents to monitor the network parameters so there is a need of intra-agent communication and it has a long training procedure. In (Lee, Kim, Lee, & Park, 2012), the authors propose an early detection method that consists of a traffic matrix using genetic algorithms and a packet based window size for the detection of DDoS attacks.

Support vector machines: This approach maps the training data obtained from the primary input space into a higher dimensional characteristic space using kernels and acquires the favourable isolating hyper-plane or a decision boundary in the form of support vectors (Chauhan et al., 2012; Gyanchandani, 2012; Rama, 2011; Wu & Yen, 2009). Then the new incoming instances are mapped into the same space and their regions are estimated to which they belong to. If the new instances do not belong to a particular region, they are treated as anomalous (Chandola et al., 2009). This is basically done to mould a linearly non-separable problem into a linearly separable one (Rama, 2011). Decision boundary is extremely robust to outliers (Gyanchandani, 2012). This method is superior to neural networks and clustering methods in case of accuracy and speed. It offers high detection accuracy with low faux pas and handles the unseen data and over-fitting problems effectively (Nadiammai & Hemalatha, 2014). One-Class SVM (OCSVM) gives better results as compared to One-Class Bayesian Networks (Heller, Svore, Keromytis, & Stolfo, 2003). An IP Address Interaction (IAI) based SVM classifier has been developed in (Cheng, Yin, Liu, Cai, & Wu, 2009) to identify the DDoS attack flows in a troop of regular network flows with high detection accuracy and low false alerts.

System call sequence analysis: System calls are used as the functional interface between the programme and the operating system kernel. By analyzing the sequence of system calls, we can detect whether the system is under attack or not. In this technique, normal system calls are divided into several short sequences, which can be considered as data items of the training set (Zu & Hu, 2016). It applies an algorithm to formulate a normal outline of the system on the basis of inter-associations in fixed time series of system calls. When a system call sequence deviates from normal behaviour sequence profile, it can be treated as anomalous. It maintains a database that collects the normal behaviour of each and every programme of the system. The database monitors the programme’s behaviour and whenever the sequence of system calls for a particular programme is not found in the database, the system indicates anomalies (Patcha & Park, 2007). This approach observes each and every system call so it results in high computational overheads and performance debasement of the monitoring system (Chandola et al., 2009). Moreover, the irregularity of system call leads to increase in false positive ratio and makes the distinction of unusual system calls more difficult.

Information theoretic-based detection: This approach detects anomalies by auditing the erudition constituents of a normal dataset with various conjectural estimates like Entropy, Multiscale entropy, Dominant state analysis, Hellinger distance, Mahalanobis distance, Relative uncertainty distribution, Kullback-Leibler divergence distance, Chi-square and Mutual information (Purwanto et al., 2014). It finds the irregularities that are presented in the information content of the normal dataset. It deals with different data types like compressed, categorical, sequential, spatial and graphical data in which the data instances are naturally ordered and reduces the complexity of datasets. Information entropy is defined as a measure that computes the incertitude corresponds to a test instance (Bhandari et al., 2015; Ray, 2004). More randomness associates with the instances results in higher entropy (Androulidakis, Chatzigiannakis, & Papavassiliou, 2009; Bhandari, Sangal, & Kumar, 2016). It identifies the anomalous network behaviour but results in high false signals due to the low rifts in unusual and standard traffic. Entropy can be calculated on the basis of different parameters such as the change in packet size distribution statistics (Agarwal & Mittal, 2012), transmission rate, IP address rate (Bhatia, Mohay, Tickle, & Ahmed, 2011) and URL accessed (Li, Zhou, Li, Bai, & Liu, 2009). As per the
study in (Park, Li, Gao, Lee, & Deng, 2008), the authors presents an FDD mechanism using randomness checks to predict the source IP addresses at the server from the previous connection requests. Therefore, this method is used to distinguish the flash events from DDoS traffic because source IP addresses of the attack traffic are not foreseeable and act as arbitrary locations on the victim system. In (Zhou, Jia, Wen, Xiang, & Zhou, 2014), the authors propose a detection unit that reveals the App-DDoS attacks by obtaining the ratio of the entropy of source IPs and the URL analyzed. The study shows that the above ratio will be large for App-DDoS attacks as compared to normal flash events. In (Sachdeva & Kumar, 2014), the authors introduce a cluster entropy concept to differentiate the flash events from DDoS traffic in which cluster are obtained from the users that previously access the web-service or belong to the same administration network and the entropy of different clusters is calculated. The study concludes that the value of entropy will be small for the predicted flash event from the same network and large for the DDoS attack traffic due to increase in the count of new networks. Information entropy-based anomaly detection methods are used in Software Defined Networks (SDN) and Cloud Computing environment against DDoS attacks (Mousavi, 2014; Navaz, Sangeetha, & Prabhadevi, 2013). Entropy-based metrics with PCA algorithm detect the anomalies that are not detected by volume based (HDDOS attacks) detection methods (Nychis, Sekar, Andersen, Kim, & Zhang, 2008). Features of this method include low false positive rate, high detection accuracy, on-line detection and the early detection of LDDoS attacks (Agarwal & Mittal, 2012; Xiang et al., 2011). It is deployed at core-end and victim-end networks (Prasad et al., 2014) for the detection of App-DDoS and LDDoS attacks (Bhandari et al., 2016; Bhuyan et al., 2014; Xiang et al., 2011).

Nearest Neighbour-based detection: This approach detects the anomalies that are far from the dense (or close) neighbour of the normal instances and uses the distance or density based measures to find the similarities (or distance) between the two or more data instances. An anomaly score is estimated for an observation either on the basis of its distance to its Kth nearest neighbour (K-NN classifier) (Nguyen & Choi, 2010; Oo & Phyuf, 2014) or the relative density of test instance. This approach deals with both continuous and categorical data types. It offers high detection accuracy, early detection, easy implementation and less computation time (Nguyen & Choi, 2010). In (Eskin et al., 2002; Zhang & Wang, 2006), anomaly score the new instance is reckoned as the aggregation of its distance from its k nearest neighbours. A proactive detection method that divides the DDoS attack into different phases and analyzes the network status in each phase using K-NN classifiers has been developed in (Nguyen & Choi, 2010). PAD algorithm discussed in (Heller et al., 2003) is based on density function and comparable to One-Class Support Vector Machine (OCSVM) to find unusual events in Windows registry. Density-based detection is not an effective approach in case the dataset of the normal instances has the varying density regions. To overcome this problem, concept of relative density to the neighbour’s density was proposed in (Breunig, Kriegel, Ng, & Sander, 2000). This approach is mostly deployed at core-end networks (Bhuyan et al., 2014) and detects LDDoS attacks (Prasad et al., 2014; Xiang et al., 2011). Table 1 presents the pros and cons of different point anomaly-based detection approaches.

3.1.2.2. Contextual anomaly-based detection. If an incoming event is abnormal in a well-defined context or situation, then it is recognized as the contextual anomaly or conditional anomaly (Song, Wu, Jermaine, & Ranka, 2007). Every data instance is defined by its contextual and behavioural attributes. Contextual attributes help in defining the contextual (or neighbour) characteristics and behavioural attributes help in defining the non-contextual characteristics for a particular instance (Chandola et al., 2009). Various approaches are used to detect contextual anomaly to find the deviations in the neighbour of an instance (i.e. find deviations from average) by using the values of behavioural attributes. Moreover, contextual attributes are expressed in the form of spatial, graphs, sequential and profile attributes (Chandola et al., 2009). Contextual anomalies are similar to point anomalies and they are anomalous within a particular context. Therefore, various point anomaly-based detection techniques (for example Information theoretic based detection) are used to identify contextual anomalies in an appropriate circumstance. The contextual anomaly-based detection system identifies a context for the incoming traffic instance with the help of contextual attributes and estimates a freak score for malicious instances with the help of a point anomaly-based detection approach. This approach helps in detecting the real world anomalies where data instances within a context tend to be similar. It recognizes the peculiarities that are not exposed by point anomaly-based detection methods. In some cases, specifying a context is not easy, so using contextual anomaly-based detection does not make any sense.

3.1.2.3. Collective anomaly-based detection. If the data instances are related to each other and the assemblage of data instances is unusual to the residual dataset, then it is known as the collective anomaly. Note that, a single instance in a collective anomaly may or may not be anomalous but the collection of such
Artificial Intelligence-based
Nearest neighbour based
Information theoretic based
sequences can be univariate and multivariate in nature
sequence datasets (Chan & Mahoney, 2005). Note that,
example time-series data, system call sequence and event
as abnormal with respect to the regular behaviour. For
ordered and identifies the subsequences that behave
sequential data in which the data instances are linearly

Table 1. Pros and cons of point anomaly-based detection approaches.

| Approaches | Pros | Cons |
|------------|------|------|
| Statistical methods based anomaly detection | - Offer a good detection rate for the ‘zero days’ or an extremely new attack if there exist statistical characteristics for the network traffic. - Controls the rate limits if an attack is confirmed. - Provides accurate results for long-term malicious activities (i.e., ‘low and slow’ attacks). - Prior knowledge about normal activity, security flaws and the attacks themselves, is not required as the expected system behaviour is prepared from observations. Therefore, such methods are simpler to manage and there is no need to refresh signatures. - These systems look for individual elements of a particular activity and generate an alarm when an attack is detected without waiting for the completion of that activity. | - Difficult to set parameters (or metrics) and unrealistic assumptions of a quasi-stationary process that may affect the threshold level, false positives, and false negatives. But such hypotheses do not exist for high-dimensional real datasets. - Such systems need accurate statistical distributions, but only a few normal profiles are modelled using purely statistical methods. - Although it provides accurate and effective results but it is a time-consuming process (i.e. takes days or weeks for results). |
| Data mining based anomaly detection | - Analyzes the lengthy, continuous patterns (i.e. different IPs, same activity). - Allow experts to concentrate on actual attacks. - Classifies the false alarm dynamics and ‘bad’ sensor indications. | - It produces very high false positive rate - Fail to be applied in real-time detection environment. |
| Information theoretic based anomaly detection | - Detects anomalies that are present in a huge amount of the information content of normal datasets. - Reduce the complexity of dataset. | - The optimal size of the substructures (like subsequences and sub-graphs) should be preferred to detect anomalies. - Performance depends on the selected information theoretic measures. - Difficult to assign an anomaly score to the test instances. - Performance decreases as the number of attributes increases. - Performance depends on the selected distance measure, which is a challenging task to be computed for complex data structures like graphs, sequences. - Very small size of samples may adversely affect the anomaly score computations. - Performance depends on the input parameters from the system or user’s point of view. - Fail to provide accurate results due to lack of sufficient data and learnable functions. - It needs high resource consumption and complex computations for detecting anomalies. |
| Nearest neighbour based anomaly detection | - Purely a data-driven approach and operates in unsupervised mode. - The possibility of an anomaly drops to set a dense neighbour if the approach operates in semi-supervised mode. | |
| Artificial Intelligence-based detection | - Facility to adjust its execution tactics on the basis of recently collected data. - Offers high detection accuracy, but more expensive as compared to another detection approaches. - Deployed at victim-end network and supervised in nature. | |

instances appears as anomalous. In sequence data, graph
data and spatial data representation, the test cases are
related to each other and the relationship among the
data instances becomes the basis of detecting collective anomalies. Various approaches have been used for
collective anomaly-based detection such as Sequential anomaly, Graph anomaly, and Spatial anomaly detection approaches. Various approaches have been used for
collective anomaly-based detection such as Sequential anomaly, Graph anomaly, and Spatial anomaly detection approaches.

Sequential anomaly detection: This approach deals with sequential data in which the data instances are linearly ordered and identifies the subsequences that behave as abnormal with respect to the regular behaviour. For example time-series data, system call sequence and event sequence datasets (Chan & Mahoney, 2005). Note that, sequences can be univariate and multivariate in nature and the sequential anomalies can be reduced to point anomalies that are easy to handle. There are a number of ways for handling anomalous sequential data such as:

- Detects anomalous sequence from a set of sequences (Budalakoti, Srivastava, Akella, & Turkov, 2006; Chan & Mahoney, 2005) and models the sequences with the help of various modelling tools like Probabilistic Suffix Tree (PST) and Sparse Markov Trees (SMT) (Chandola et al., 2009). It works in semi-supervised and unsupervised mode.
- Detects anomalous subsequence (or discords) that is anomalous in a long sequence (Bu et al., 2007) and works in unsupervised mode.
- Determine the frequency of a query pattern that is anomalous as compared to expected frequency in a particular sequence (Chandola et al., 2009).
The patterns of perversions are filed in a directory or database. Useful in recognizing only defined and well-known attacks in real-time and performs in supervised mode. Deployed at any network (source-end, victim-end or core-end network). Offers the features of robustness, scalability, and flexibility.

Identifies attack by labelling the activity as either abnormal or normal. Able to expose new, unusual or ‘zero days’ attacks and exceptional patterns if do not correspond to the presumed normal functionalities. A different profile of normal activities can be built for different systems, applications or networks to confuse the attacker what activities it can take without going identified.

It runs in supervised as well as unsupervised mode. Combined use of two or more detection approaches. While the hybrid methods still drop some sorts of attacks, but its low false notifications rate strengthens the plausibility of exploring most of the alerts.

Unable to identify the new, novel attacks. Prior knowledge of attack signatures is needed and it should be updated regularly. Moreover, Knowledge should be accumulated with finer details for better detection.

It is very complex and prolonged task as the subsistence of the knowledge base entails thoughtful and comprehensive dissection of each vulnerability. This approach faces the various generalization issues. Difficult to extract the network features so a training phase is needed to recognize normal activity profile. Setting threshold value is difficult to adjust the false positive and the false negative rates. Defined a rule set is difficult. High false positive rate leads to low detection efficiency. Offers low throughput and are computationally expensive because of the cost of retaining a record of, and perhaps refreshing several system profile metrics.

Exploits benefits of both signature and anomaly-based detection procedures so the resulting hybrid systems are not perpetually favourable. Leads to high complexity and implementation cost.

**Graph anomaly detection:** This approach deals with the graphical data in which the data instances are depicted as points (or vertices) and linked to other vertices through edges. It detects the sub-graphs that are unusual within the large graph (Noble & Cook, 2003). Various measures like entropy have been applied to the sub-graphs to determine its anomaly score.

**Spatial anomaly detection:** This approach deals with spatial data and detects the sub-regions (spatial anomalies) that are irregular to rest of the data (Chandola et al., 2009). It can be used to find contextual and collective anomalies.

### 3.1.3. Hybrid detection

Hybrid approach for DDoS detection combines the two or more of above detection strategies. The monitoring capabilities of a detection system can be improved by developing a hybrid model that is meant by analyzing the regular system behaviour and impertinent attacker behaviour. This approach attempts to familiar as well as anonymous attacks if it consists of both anomaly and signature-based detection techniques (Bhuyan & Kalita, 2012). When a signature-based technique is adopted along with the anomaly-based technique, then the hybrid system can detect the intruder who tries to change the attack patterns stored in the signature database (Patcha & Park, 2007). It offers the features of both anomaly-based and signature-based methods like high detection rate and low false signal rate (Wu & Yen, 2009). For example, EMERALD (Patcha & Park, 2007), NeGPAIM (Botha, Solms, Perry, Loubser, & Yamoyany, 2002), RST-SVM (Chen, Cheng, Chen, & Hsieh, 2009), NFBoost (Raj & Selvakumar, 2013) are the hybrid detection systems that are developed from the combination of above detection approaches. Moreover, signature-based SNORT method is combined with one or more anomaly-based methods (such as PHAD, NETAD, ALAD, LERAD) in order to develop a hybrid model for real-time detection (Nadiammai & Hemalatha, 2014). In (Asosheh & Ramezani, 2008), the authors propose a combined data mining procedure for the automated exposure of attacks that consists of a clustering (K-mean) and classification (K-nearest neighbour) algorithms. In (Agarwal & Mittal, 2012), the authors propose a hybrid approach that consists of Entropy-based and SVM based detection methods that remove the demerits of both the techniques and results in low false alert rate and high detection accuracy. Sometimes, the output of different classifiers such as Bayesian Networks, Neural Networks and Decision Tree (DT) are combined using multiple fusion techniques to improve the system’s consummations (Modi et al., 2013). Combining different approaches makes the detection system stronger but the detection results are not always very good. In fact, developing a hybrid detection system from different approaches that can interoperately effectively and efficiently is a challenging task (Patcha & Park, 2007). A summary of different detection approaches has been conferred in Table 2.

### 3.2. Functional classes

A detection approach belongs to one or more different functional classes. A review of different detection approaches and the functional classes to which they
belong to has been shown in Table 3 based on our current literature survey and the different functional classes are explained below:

### Table 3. Functional classes of different detection approaches.

| Functional classes                      | Anomaly-based detection | Data Mining | Artificial Intelligence based detection | Signature-based detection |
|----------------------------------------|-------------------------|-------------|------------------------------------------|---------------------------|
|                                        |                         | CM          | CBM                                      | ARM                       |
|                                        |                         |             | BN                                       | FL                        |
|                                        |                         |             | GA                                       | SVM                       |
|                                        |                         |             | SCSA                                     |                           |
| Core-end detection                     | ✓                       | ✓           | ✓                                        |                           |
| Source-end detection                   | ✓                       | ✓           | ✓                                        |                           |
| Victim-end detection                   | ✓                       | ✓           | ✓                                        |                           |
| Supervised mode                        | ✓                       | ✓           | ✓                                        |                           |
| Semi-supervised mode                   | ✓                       | ✓           | ✓                                        |                           |
| Unsupervised mode                      | ✓                       | ✓           | ✓                                        |                           |
| Net-DDoS detection                     | ✓                       | ✓           | ✓                                        |                           |
| App-DDoS detection                     | ✓                       | ✓           | ✓                                        |                           |
| HDDoS detection                        | ✓                       | ✓           | ✓                                        |                           |
| LDDoS detection                        | ✓                       | ✓           | ✓                                        |                           |
| Profiling based                        | ✓                       | ✓           | ✓                                        |                           |
| Model based                            | ✓                       | ✓           | ✓                                        |                           |
| One-class setting                      | ✓                       | ✓           | ✓                                        |                           |
| Multi-class setting                    | ✓                       | ✓           | ✓                                        |                           |

CM: Clustering methods; ANN: Artificial Neural Networks; GA: Genetic Algorithms; CBM: Classification based methods; BN: Bayesian Networks; SVM: Support Vector Machines; ARM: Associative rule mining; FL: Fuzzy Logic; SCSA: System Call Sequence Analysis.

3.2.1. **Source-end, victim-end and core-end detection**

The detection approaches depend on the nature of data that is available either from the end-users (source or victim) or the network. The end-user information contains the data from TCP and UDP packets and it is specific to a particular user application. Various detection approaches are implemented on either source-end or victim-end. Moreover, the detection approaches deployed at victim-end operates in on-line as well as off-line mode (Prasad et al., 2014). The network-based information contains the data from intermediate or core router’s physical interfaces and their forwarding engines (Thottan & Ji, 2003). The detection approaches implemented at core networks detect anomalies and inform the source to slow down the data traffic. In (Yu, Guo, & Stojmenovic, 2012), the authors propose a detection method that monitors the traffic at edge routers of ISP domains. Most of the Artificial Intelligence-based detection approaches are deployed at victim-end (Prasad et al., 2014).

3.2.2. **Supervised, semi-supervised and unsupervised mode**

DDoS detection approaches can operate in one of the following three modes: *Supervised*, *Semi-supervised* and *Unsupervised mode* (Buragohain et al., 2015). In *supervised mode*, the detection approach requires a trained dataset (or a classifier) to find the anomalies, where the trained dataset includes input variables and output classes. The trained dataset is used to extract the hidden functions and predict the class of input variables (incoming traffic instances). This mode is similar to a predictive model. For example, Classification techniques comes under the category of supervised data mining (Aggarwal & Gupta, 2015). In the *unsupervised mode*, the detection system identifies the hidden functions (or patterns) from a given dataset without having any trained dataset but it produces less detection accuracy (Nadiammai & Hemalatha, 2014). For example, Clustering and Associative rule mining comes under the category of unsupervised data mining (Aggarwal & Gupta, 2015). Approaches that work in the *semi-supervised mode* have incomplete training data i.e. training data is meant for only for normal class and some targets are missing for anomaly class (Nadiammai & Hemalatha, 2014). It produces high detection ratio and low false alerts. Therefore, it is more applicable as compared to supervised mode. But, it is very difficult to collect the entire anomalous behaviour in a training dataset.

3.2.3. **Profiling-based and modelling-based detection approaches**

In profiling based approaches, profiles of normal behaviour of various real computer system components (like different types of network traffic, programmes, users) are developed and certain patterns or activity deviations from the normal profile are identified to be considered as anomalous (Lazarevic, 2016). UNIX shell commands, audit events, system calls, keystroke and network packages are used as data sources for developing the attack profiles (Rama, 2011). In spite of building profiles, different models like replicator neural networks or unsupervised support vector machines are fabricated for the attack detection. In model-based approaches, the model
represents the normal functioning of the system and the anomalies are recognized as divergences from the standard behaviour.

3.2.4. Net-DDoS and App-DDoS attack detection approaches

Most of the approaches discern the Net-DDoS attacks, only a few of them can detect App-DDoS attacks (Xie & Yu, 2006). Swaddler anomaly-based detection technique (Cova, Balzarotti, Felmetsger, & Vigna, 2007), Fuzzy estimators (Shiaeles et al., 2012) and K-means clustering (Yu, Li, Chen, & Chen, 2007; Zhong & Yue, 2010) detects App-DDoS attacks in the real-time environment.

3.2.5. High-rate and low-rate DDoS detection approaches

Most of the detection approaches captures only high-rate traffic (HDDoS attacks or volume-based attacks), but some of them detect only low-rate traffic (LDDoS attacks) or both (Gupta et al., 2012). HDDoS detection strategy is unable to detect several types of anomalies, which can be analyzed by Entropy-based detection (Androulidakis et al., 2009; Lakhina, Crovella, & Diot, 2005; Nychis et al., 2008). Low rate traffic is difficult to detect because it behaves as normal throughout its journey but aggregates only at the victim network (Xiang et al., 2011). Congestion Participation Rate (CPR) and Cumulative Amplitude Spectrum (CAS) based approaches are effective in detecting LDDoS attacks (Zhang et al., 2012).

3.2.6. One-class and multi-class setting

In one-class based setting, the detection approaches prepare a single class that represents the usual network behaviour and the entities that do not belong to the class are treated as anomalous. But in multi-class based setting, the detection system prepares a set of multiple normal classes and the entities that do not belong to any of the multiple classes, are considered as anomalous.

3.3. Metrics

This section presents the various metrics that are crucial to compare and assess the performance of different detection approaches:

3.3.1. Complexity

This metrics defines the overall complexity of a particular detection approach related to time, memory storage, coordination among different components and the computations needed for the detection system.

3.3.2. False positive (FP)

If the test instance is innocuous but the detection system falsely reports it as anomalous, then it is considered as false positives (Estevez-Tapiador, Garcia-Teodoro, & Diaz-Verdejo, 2004) and the system generates an anomaly alarm of overabundance of false positives. False positive rate depends on the threshold value, where very low threshold value results in high false positive rate.

3.3.3. False negative (FN)

If the test instance is abnormal or malicious but it is labelled as innocuous, then it is considered as false negatives (Estevez-Tapiador et al., 2004) and the detection does not generate an alarm on necessary malicious traffic. The high false negative rate has been noticed if the threshold value is set too high.

3.3.4. False alarm rate (FAR)

False alarm rate is considered as a count of false positive or false negative. High false alarm rate initiated in anomaly detection makes it complicated to relate a distinct anomaly signal with the events that causes them. False alarm rate can also be affected if the attacker trains the detection system to accept the malicious traffic as normal.

3.3.5. Implementation cost

It consists of the total cost needed for implementing a particular detection technique on the source-end, victim-end or core-end (intermediate) network.

3.3.6. Reliability

Reliability is defined as how well a detection approach is performing its required functions in a particular time period under stated conditions or in the case of a failure (component or system failure).

3.3.7. Detection rate

Detection rate is measured as the number of malicious packets identified by the detection approach divided by the total number of malicious packets in the dataset.

3.3.8. Detection accuracy

Detection accuracy is defined as the degree to which the result of a detection approach conforms to the correct identification of malicious packets of DDoS traffic. Accuracy depends on the proportions of false positive and false negative rates against detection of DDoS attacks.

3.3.9. Type of monitoring

This parameter decides whether a particular approach monitors the high or low data rate traffic or identifies the malicious packets i.e. it decides whether the detection approach is traffic volume based or IP attribute based (Alenezi & Reed, 2012). In (Alenezi & Reed, 2012; Karimazad & Faraahi, 2011), it has been discussed that
Table 4. Comparison of detection approaches.

| Metrics                  | Anomaly-based detection | Signature-based detection | Hybrid detection |
|--------------------------|-------------------------|---------------------------|------------------|
| Complexity               | Medium                  | Low                       | High             |
| False positive rate      | High                    | Very low                  | Low              |
| False negative rate      | High                    | Medium                    | Low              |
| False alarm rate         | High                    | Low                       | Medium           |
| Implementation cost      | Medium                  | Low                       | High             |
| Reliability              | Yes                     | Yes                       | –                |
| Detection rate           | Medium                  | High (only for known attacks) | High             |
| Detection accuracy       | Medium                  | High                      | High             |
| Type of monitoring       | Traffic monitoring, Packet monitoring | Traffic monitoring, Packet monitoring | –               |
| Real-time detection      | Medium                  | High                      | –                |

anomaly-based detection analyzes both the traffic flow characteristic as well packets contents. TOPAS (Munz & Carle, 2007), a detection system that allows parallel deployment of different detection algorithms offers packet-based monitoring, on-line analysis, and real-time detection.

3.3.10. Real-time detection

The existing approaches must be deployed in real networks with a suitable level of detection rate, accuracy, and false alarm rate. Preference should be given to detection speed rather than accuracy in a real-time environment.

4. Comparison of detection approaches

In this section, a comparison of detection approaches (Alenezi & Reed, 2012; Modi et al., 2013; Ranju, 2014) has been presented on the basis of previously discussed parameters. Table 4 presents the comparison of different detection approaches based on our current literature review.

5. Issues in existing detection approaches

- Recent trends show that various detection techniques are presented in the theory, but only some of them run effectively on all protocols and works in the real environment. Developing and enacting an ideal and real-time detection system is indeed a hard task. In order to suffice the growing demands for detection and response, there are many issues faced by the researchers:
  - Detection schemes involve complex computations due to which time taken by the system is too long to find the anomalous conditions. Therefore, detection speed must be given preference over detection accuracy for the disclosure of attacks in real-time. It should be effective against a variety of attack tools available today. Therefore, it should not be exposed to attacks, producing an impending disruption of its services (Xiang et al., 2011).
  - The detection procedures should rest on a small fraction of input (traffic) parameters, and sturdy against future trials by the attacker. It should be capable of handling the masses and functions accurately in high-speed real networks.
  - Accurate isolation of HDDoS attack traffic and regular flash events (with the minimal support needed or low false alerts), real-time updating of network statistics and quick identification of spoofed IPs is the most challenging task in real-time detection environment.
  - Most of the detection methods analyze the packet contents and traffic flow characteristics for the attack exposures. But the attacker undoubtedly modifies the packet contents and traffic flow traits, thus the detection system fails. Moreover, it is very difficult to analyze the encrypted packets (Li et al., 2015). IP attributes (IP protocol type, packet-size) based detection techniques adversely affect the performance by increasing the computational complexity and false positive rate.
  - Nowadays, a combined approach of different detection approaches has become the utmost necessity for defending against unknown or novel attacks. A single router cannot identify that a particular network or a victim is under attack and adjust its traffic to decrease the impact of DDoS attacks (Reddy, Siva, & Malathi, 2013).
  - Most of the anomaly-based detection methods try to find the anomalies in network traffic as well packet contents. But none of the detection algorithms has focused on types of anomalies that it can detect. Type of the anomalies is associated with the types of botnets, flash crowds and the types of attacks (Purwanto et al., 2014).

6. Conclusion

In this review paper, we surveyed on several detection approaches against DDoS attacks. It is very complicated to discern which detection approach should be followed for a circumstantial dilemma. Signature-based detection
approach can disclose only known attacks and results in high detection accuracy with the low false notifications. But the attacker can quickly adjust the attack signatures or perform attacks with small variations. Therefore, the attacks remain unidentified by this approach. Nowadays, anomaly-based detection approach has been widely used for the detection of Net-DDoS as well as App-DDoS attacks. The key challenges for this approach are online analysis, manipulating a huge amount of data and the increasing false signal ratio due the presence of uncertainty in data. Supervised and semi-supervised techniques are fancied for controlling the huge amount of data but unsupervised techniques are adopted for catching unfamiliar attacks. Nevertheless, such schemes do not fit for the real-time detection. Therefore, implementing a mixed approach of supervised and unsupervised techniques that can recognize both unknown and known DDoS attacks in the real-time environment is a challenging task. From this review paper, we have concluded that the researchers have stated different defense mechanisms against the DDoS attacks. But due to lack of benchmarks against which the performance of defense tools may be compared, the best solutions for defending against such attacks are improbable.

7. Scope of future work

We strongly believe that a perfect comprehensive real-time defense framework could be the best and effective approach to battle DDoS attacks. Building a defense mechanism as close as to the attack source with an evitable participation of various service providers offering a source address validation and filtering features, we hope to find sooner in the near future.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Acohido, B., & Swartz, J. (2009). Hacker attack takes down Twitter, Facebook, LiveJournal. Hacker attack takes down Twitter, Facebook, Live Journal, ed.
Agarwal, B., & Mittal, N. (2012). Hybrid approach for detection of anomaly network traffic using data mining techniques. Procedia Technology, 6, 996–1003.
Aggarwal, A., & Gupta, A. (2015). Survey on data mining and IP traceback technique in DDoS attack. International Journal of Engineering and Computer Science ISSN:2319-7242, 4(6), 12595–12598.
Alenezi, M., & Reed, M. J. (2012). Methodologies for detecting DoS/DDoS attacks against network servers. Proceedings of the seventh international conference on systems and networks communications—ICSNC, November 18–23, Lisbon, Portugal. pp. 92–98. IARIA.
Androulidakis, G., Chatzigiannakis, V., & Papavassiliou, S. (2009). Network anomaly detection and classification via opportunistic sampling. IEEE Network, 23(1), 6–12.
Asosheh, A., & Ramezani, N. (2008). A comprehensive taxonomy of DDoS attacks and defense mechanism applying in a smart classification. WSEAS Transactions on Computers, 7(7), 281–290.
Baraniuk, C. Retrieved from http://www.bbc.com/news/technology-35376327
Barbara, D., Couto, J., Jajodia, S., & Wu, N. (2001). ADAM: A testbed for exploring the use of data mining in intrusion detection. ACM Sigmod Record, 30(4), 15–24.
Barbara, D., Wu, N., & Jajodia, S. (2001). Detecting novel network intrusions using Bayes estimators. SIAM, 1–17. doi:10.1137/1.9781611972719.28
Bhandari, A., Sangal, A., & Kumar, K. (2015). Destination address entropy based detection and traceback approach against distributed denial of service attacks. International Journal of Computer Network and Information Security, 7(8), 9–20.
Bhandari, A., Sangal, A. L., & Kumar, K. (2016). Characterizing flash events and distributed denial-of-service attacks: An empirical investigation. Security and Communication Networks, 9(13), 2222–2239.
Bhatia, S., Mohay, G., Tickle, A., & Ahmed, E. (2011). Parametric differences between a real-world distributed denial-of-service attack and a flash event. IEEE, 210–217. doi:10.1109/ARES.2011.39
Bhuyan, M. H., & Kalita, J. K. (2012). Network anomaly detection: methods, systems and tools.
Bhuyan, M. H., Kashyap, H. J., Bhattacharyya, D. K., & Kalita, J. K. (2014). Detecting distributed denial of service attacks: Methods, tools and future directions. The Computer Journal, 57(4), 537–556.
Bisson, D. Retrieved from http://www.tripwire.com/state-of-security/risk-based-security-for-executives/risk-management/report-ddos-attacks-grew-in-number-size-and-sophistication-in-q4-2015/
Blazek, R. B., Kim, H., Rozovskii, B., & Tartakovsky, A. (2001). A novel approach to detection of denial-of-service attacks via adaptive sequential and batch-sequential change-point detection methods. Citeseer, 220–226.
Botha, M., Solms, R. V., Perry, K., Loubser, E., & Yamoyanj, G. (2002). The utilization of artificial intelligence in a hybrid intrusion detection system. SAICSIT ’02 Proceedings of the 2002 annual research conference of the South African institute of computer scientists and information technologists on enablement through technology. pp. 149–155.
Breunig, M. M., Kriegel, H.-P., Ng, R. T., & Sander, J. O. R. (2000). LOF: Identifying density-based local outliers. ACM, 29, 93–104.
Bu, Y., Leung, O. T.-W., Fu, A. W.-C., Keogh, E. J., Pei, J., & Meshkin, S. (2007). WAV: Finding Top-K discord in time series database. SIAM, 449–454. doi:10.1137/1.9781611972771.43
Budalakoti, S., Srivastava, A. N., Akella, R., & Turkov, E. (2006). Anomaly detection in large sets of high-dimensional symbol sequences.
Buragohain, C., Kalita, M. J., Singh, S., & Bhattacharyya, D. K. (2015). Anomaly based DDoS attack detection. International Journal of Computer Applications, 123(17), 35–40.
Cabrera, J. B., Lewis, L., Qin, X., Lee, W., Prasath, R. K., Ravichandran, B., & Mehra, R. K. (2001). Proactive detection
of distributed denial of service attacks using mib traffic variables-a feasibility study. IEEE, 609–622.

Carl, G., Kesidis, G., Brooks, R. R., & Rai, S. (2006). Denial-of-service attack-detection techniques. Internet Computing, IEEE, 10(1), 82–89.

Chan, P. K., & Mahoney, M. V. (2005). Modeling multiple time series for anomaly detection. ICDM ’05 Proceedings of the Fifth IEEE International Conference on Data Mining, November 27–30. pp. 90–97. Washington, DC: IEEE Computer Society.

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys, 41(3), 1–58.

Chauhan, A., Mishra, G., & Kumar, G. (2012). Survey on data mining techniques in intrusion detection. Lap Lambert Academic Publ.

Chen, R.-C., Cheng, K.-F., Chen, Y.-H., & Hsieh, C.-F. (2009). Detection of DDoS attacks over multiple network domains. Parallel and Distributed Systems, IEEE Transactions, 18(12), 1649–1662.

Cheng, C.-M., Kung, H., & Tan, K.-S. (2002). Use of spectral analysis in defense against DoS attacks. IEEE, 3, 2143–2148.

Cheng, J., Yin, J., Liu, Y., Cai, Z., & Wu, C. (2009). DDoS attack detection using IP address feature interaction. International conference on intelligent networking and collaborative systems, November 4–6, Barcelona, Spain. IEEE.

Cova, M., Balzarotti, D., Felmetsger, V., & Vigna, G. (2007). Swadler: An approach for the anomaly-based detection of state violations in web applications. In C. Kruegel, R. Lippmann, & A. Clark (Eds.), Recent advances in intrusion detection. RAID 2007 (pp. 63–86). Lecture Notes in Computer Science, Vol. 4637. Berlin: Springer.

Dainotti, A., Pescape, A., & Ventre, G. (2009). A cascade architecture for DoS attacks detection based on the wavelet transform. Journal of Computer Security, 17(6), 945–968.

Dickerson, J. E., & Dickerson, J. A. (2000). Fuzzy network profiling for intrusion detection. 19th international conference of the North American fuzzy information processing society – NAFIPS, July 13–15, Atlanta, GA, USA. IEEE.

Douglis, C., & Mitrokotsa, A. (2004). DDoS attacks and defense mechanisms: Classification and state-of-the-art. Computer Networks, 44, 643–666.

Ektefa, M., Memar, S., Sidi, F., & Affendey, L. S. (2010). Intrusion detection using data mining techniques information retrieval & knowledge management. International conference on information retrieval & knowledge management (CAMP), March 17–18, Shah Alam, Selangor, Malaysia. IEEE.

Eskin, E., Arnold, A., Prerau, M., Portnoy, L., & Stolfo, S. (2002). A geometric framework for unsupervised anomaly detection. In D. Barbará & S. Jajodia (Eds.), Applications of data mining in computer security (pp. 77–101). Advances in Information Security, Vol. 6. Boston, MA: Springer.

Esposito, M., Mazzariello, C., Oliviero, F., Romano, S. P., & Sansone, C. (2007). Real time detection of novel attacks by means of data mining techniques. In C. S. Chen, J. Filipe, I. Seruca, & J. Cordeiro (Eds.), Enterprise information systems VII (pp. 197–204). Dordrecht: Springer.

Estevez-Tapiador, J. M., Garcia-Teodoro, P., & Diaz-Verdejo, J. E. (2004). Anomaly detection methods in wired networks: A survey and taxonomy. Computer Communications, 27(16), 1569–1584.

Garcia-Teodoro, P., Diaz-Verdejo, J., Macia-Fernandez, G., & Vazquez, E. (2009). Anomaly-based network intrusion detection: Techniques, systems and challenges. Computers & Security, 28(1), 18–28.

Garg, K., & Chawla, R. (2011). Detection of DDoS attacks using data mining. International Journal of Computing and Business Research (IJCBR), 2(1).

Gil, T. M., & Poletto, M. (2001). MULTOPS: A data-structure for bandwidth attack detection. Proceedings of the 10th conference on USENIX security symposium, August 13–17, Washington, DC, USA. IEEE.

Gupta, B., Misra, M., & Joshi, R. C. (2012). An ISP level solution to combat DDoS attacks using combined statistical based approach. arXiv preprint arXiv:1203.2400.

Gyanchandani, R. N. Y. M. (2012). Taxonomy of anomaly based intrusion detection system: A review. International Journal of Scientific and Research Publications, 2(12). ISSN 2250-3153.

Harjinder, J. M. (2013). A review of machine learning based anomaly detection techniques. International Journal of Computer Applications Technology and Research, 2(2), 185–187.

Heller, K., Svore, K., Keromytis, A. D., & Stolfo, S. (2003). One class support vector machines for detecting anomalous windows registry accesses, 2–9. doi:10.7916/D85M6CCF

Islam, M. H., & Jamil, M. (2005). Taxonomy of statistical based anomaly detection techniques for intrusion detection. International Conference on Emerging Technologies, Islamabad, pp. 270–276, September 17–18.

Jalili, R., Imani-Mehr, F., Amini, M., & Shahriari, H. R. (2005). Detection of distributed denial of service attacks using statistical pre-processor and unsupervised neural networks. In R. H. Deng, F. Bao, H. Pang, & J. Zhou (Eds.), Information security practice and experience (pp. 192–203). ISPEC 2005. Lecture Notes in Computer Science, Vol. 3439. Berlin: Springer.

Jin, S., & Yeung, D. S. (2004). A covariance analysis model for DDoS attack detection. IEEE, 4, 1882–1886.

Karimazad, R., & Faraahi, A. (2011). An anomaly-based method for DDoS attacks detection using RBF neural networks, 11. Khandelwal, S. Retrieved from http://thehackernews.com/2016/01/biggest-ddos-attack.html

Kruegel, C., Mutz, D., Robertson, W., & Valeur, F. (2003). Bayesian event classification for intrusion detection. Proceedings of the 19th annual computer security applications conference, December 8–12, Las Vegas, NV, USA. IEEE.

Lakhina, A., Crovella, M., & Diot, C. (2005). Mining anomalies using traffic feature distributions. ACM, 35, 217–228.

Lazarevic, A. (2016). Anomaly Detection/Outlier Detection in Security Applications.

Lee, S. M., Kim, D. S., Lee, J. H., & Park, J. S. (2012). Detection of DDoS attacks using optimized traffic matrix. Computers & Mathematics with Applications, 63(2), 501–510.

Lee, W., Stolfo, S. J., & Mok, K. W. (1999). A data mining framework for building intrusion detection models. Proceedings of the 1999 IEEE symposium on security and privacy, May 14, Oakland, CA, USA. IEEE.
Lee, W., & Xiang, D. (2001). Information-theoretic measures for anomaly detection. Proceedings 2001 IEEE symposium on security and privacy, May 14–16, Oakland, CA, USA. IEEE.

Li, K., Zhou, W., Li, P., Hai, J., & Liu, J. (2009). Distinguishing DDoS attacks from flash crowds using probability metrics. Third international conference on network and system security, October 19–21, Gold Coast, Queensland, Australia. IEEE.

Li, M., & Li, M. (2009). A new approach for detecting DDoS attacks based on wavelet analysis. IEEE, 1–5.

Li, S.-H., Kao, Y.-C., Zhang, Z.-C., Chuang, Y.-P., & Yen, D. C. (2015). A network behavior-based botnet detection mechanism using PSO and K-means. ACM Transactions on Management Information Systems (TMIS), 6(1), 3–00.

Li, Y., Guo, L., Tian, Z.-H., & Lu, T.-B. (2008). A lightweight web server anomaly detection method based on transductive scheme and genetic algorithms. Computer Communications, 31(17), 4018–4025.

Limwiwatkul, L., & Rungsawang, A. (2004). Distributed denial of service detection using TCP/IP header and traffic measurement analysis. IEEE, 1, 605–610.

Mirkovic, J., Prier, G., & Reiher, P. (2002). D-WARD: A source-end defense against flooding denial-of-service attacks. IEEE Transactions on Dependable and Secure Computing, 2(3), 216–232.

Moud, C., Patel, D., Borisanjia, B., Patel, H., Patel, A., & Rajarajan, M. (2013). A survey of intrusion detection techniques in cloud. Journal of Network and Computer Applications, 36(1), 42–57.

Mousavi, S. M. (2014). Early detection of DDoS attacks in software defined networks controller.

Munz, G., & Carle, G. (2007). Real-time analysis of flow data for network attack detection. 10th IFIP/IEEE international symposium on integrated network management, May 21–25, Munich, Germany. IEEE.

Nadiammai, G., & Hemalatha, M. (2014). Effective approach toward intrusion detection system using data mining techniques. Egyptian Informatics Journal, 15(1), 37–50.

Narayana, M. S., Prasad, B., Srividhya, A., & Ranga, K. P. (2011). Data mining machine learning techniques–A study on abnormal anomaly detection system. International Journal of Computer Science and Telecommunications, 2(6).

Navaz, A., Sangeetha, V., & Prabhadevi, C. (2013). Entropy based anomaly detection system to prevent DDoS attacks in cloud. arXiv preprint arXiv:1308.6745.

Nguyen, H.-V., & Choi, Y. (2010). Proactive detection of DDoS attacks utilizing k-NN classifier in an anti-DDoS framework. International Journal of Electrical, Computer, and Systems Engineering 4 (4): 247–252.

Noble, C. C., & Cook, D. J. (2003). Graph-based anomaly detection. In Proceedings of the 9th ACM SIGKDD international conference on knowledge discovery and data mining (pp. 631–636). New York, NY: ACM Press.

Nychis, G., Sekar, V., Andersen, D. G., Kim, H., & Zhang, H. (2008). An empirical evaluation of entropy-based traffic anomaly detection (pp. 151–156). Proceedings of the 8th ACM SIGCOMM conference on Internet measurement, October 20–22, Vouliagmeni, Greece. New York, NY: ACM.

Oo, T. T., & Phyu, T. (2014). Analysis of DDoS Detection System based on Anomaly Detection System.

Papadopoulos, C., Lindell, R., Mehringer, J., Hussain, A., & Govindan, R. (2003). COSSACK: Coordinated suppression of simultaneous attacks. Proceedings DARPA information survivability conference and exposition, April 22–24, Washington, DC, USA. IEEE.

Park, H., Li, P., Gao, D., Lee, H., & Deng, R. H. (2008). Distinguishing between fe and ddos using randomness check. In International conference on information security (pp. 131–145). Lecture Notes in Computer Science, Vol. 5222. Berlin: Springer.

Patcha, A., & Park, J.-M. (2007). An overview of anomaly detection techniques: Existing solutions and latest technological trends. Computer Networks, 51(12), 3448–3470.

Pei, J., Upadhyaya, S. J., Farooq, F., & Govindaraju, V. (2004). Data mining for intrusion detection: Techniques applications and Systems. 20th international conference on data engineering, April 2, Boston, MA, USA. IEEE.

Peng, T., Leckie, C., & Ramamohanarao, K. (2004). Proactively detecting distributed denial of service attacks using source IP address monitoring. In International conference in networking (pp. 771–782). Lecture Notes in Computer Science, Vol. 3042. Berlin: Springer.

Prasad, K. M., Mohan, A. R., & Rao, K. V. (2014). Dos and DDoS attacks: Defense, detection and traceback mechanisms–A survey. Global Journal of Computer Science and Technology, 14(7).

Purwanto, Y., Kuspiyanto, Hendrawan, Rahardjo, B. (2014). Traffic anomaly detection in DDoS flooding attack. 8th international conference on telecommunication systems services and applications (TSSA), October 23–24, Kota, Indonesia. IEEE.

Questions to ask your DNS host about DDoS. Retrieved from http://www.circleid.com/posts/20141016_3_questions_to_ask_your_dns_host_about_ddos/

Raj, P. A., & Selvakumar, S. (2011). Distributed denial of service attack detection using an ensemble of neural classifier. Computer Communications, 34(11), 1328–1341.

Raj, P. A., & Selvakumar, S. (2013). Detection of distributed denial of service attacks using an ensemble of adaptive and hybrid neuro-fuzzy systems. Computer Communications, 36(3), 303–319.

Rama, V. V. (2011). A review of anomaly based intrusion detection systems. International Journal of Computer Applications, 28(7), 26–35.

Ranju, N. M. (2014, April). Survey on DDOS attack prevention and detection techniques. International Journal for Advance Research in Engineering and Technology, 2.

Raut, A. S., & Singh, K. R. (2014). Anomaly based intrusion detection–A review. International Journal on Network Security, 5(3), 7–00.

Ray, A. (2004). Symbolic dynamic analysis of complex systems for anomaly detection. Signal Processing, 84(7), 1115–1130.

Reddy, P., Siva, R., & Malathi, C. (2013). Techniques to differentiate DDoS attacks from flash crowd. International Journal of Advanced Research in Computer Science and Software Engineering, 3(6), 295–299.

Sachdeva, M., & Kumar, K. (2014). A traffic cluster entropy based approach to distinguish DDoS Attacks from flash event using DETER testbed. ISRN Communications and Networking 2014.

Saman, J. J., & Tipper, D. (2013). A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks. IEEE Communications Surveys and Tutorials, 15(4), 2046–2069.
Shiaeles, S. N., Katos, V., Karakos, A. S., & Papadopoulos, B. K. (2012). Real time DDoS detection using fuzzy estimators. Computers & Security, 31(6), 782–790.

Singh, N., Hans, A., Kumar, K., & Singh, M. P. (2015). Comprehensive study of various techniques for detecting DDoS attacks in cloud environment. International Journal of Grid and Distributed Computing, 8(3), 119–126.

Song, X., Wu, M., Jermaine, C., & Ranka, S. (2007). Conditional anomaly detection. IEEE Transactions on Knowledge and Data Engineering, 19(5), 631–645.

Tajbakhsh, A., Rahmati, M., & Mirmehdi, A. (2009). Intrusion detection using fuzzy association rules. Applied Soft Computing, 9(2), 462–469.

Thapngam, T., Yu, S., Zhou, W., & Beliakov, G. (2011). Discriminating DDoS attack traffic from flash crowd through packet arrival patterns. IEEE conference on computer communications workshops, April 10–15, Shanghai, China. IEEE.

Thomas, R., Mark, B., Johnson, T., & Croall, J. (2003). Netbouncer: Client-legitimacy-based high-performance DDoS filtering. IEEE, 1, 14–25.

Thottan, M., & Ji, C. (2003). Anomaly detection in IP networks. IEEE Transactions on Signal Processing, 51(8), 2191–2204.

Verisign. Retrieved from https://www.verisign.com/en_IN/security-services/ddos-protection/ddos-report/index.xhtml

Wang, H., Zhang, D., & Shin, K. G. (2004). Change-point monitoring for the detection of DoS attacks. IEEE Transactions on Dependable and Secure Computing, 1(4), 193–208.

Wang, J., Phan, R. C.-W., Whitley, J. N., & Parish, D. J. (2010). Augmented attack tree modeling of distributed denial of services and tree based attack detection method. 10th IEEE international conference on computer and information technology, June 29–July 1, Bradford, UK. IEEE.

Woelfl, N. (2016, October 26). Retrieved from https://www.theguardian.com/technology/2016/oct/26/ddos-attack-dynmirai-botnet

Wu, S.-Y., & Yen, E. (2009). Data mining-based intrusion detectors. Expert Systems with Applications, 36(3), 5605–5612.

Wu, Y.-C., Tseng, H.-R., Yang, W., & Jan, R.-H. (2011). DDoS detection and traceback with decision tree and grey relational analysis. International Journal of Ad Hoc and Ubiquitous Computing, 7(2), 121–136.

Xia, Z., Lu, S., Li, J., & Tang, J. (2010). Enhancing DDoS flood attack detection via intelligent fuzzy logic. Informatica, 34(4), 497–507.

Xiang, Y., Li, K., & Zhou, W. (2011). Low-rate DDoS attacks detection and traceback by using new information metrics. Information Forensics and Security, IEEE Transactions, 6(2), 426–437.

Xie, Y., & Yu, S.-Z. (2006). A novel model for detecting application layer DDoS attacks. IEEE, 2, 56–63.

Xie, Y., & Yu, S.-Z. (2009a). A large-scale hidden semi-Markov model for anomaly detection on user browsing behaviors. IEEE/ACM Transactions on Networking, 17(1), 54–65.

Xie, Y., & Yu, S.-Z. (2009b). Monitoring the application-layer DDoS attacks for popular websites. IEEE/ACM Transactions on Networking, 17(1), 15–25.

Yu, J., Li, Z., Chen, H., & Chen, X. (2007). A detection and offense mechanism to defend against application layer DDoS attacks. Third international conference on networking and services, June 19–25, Athens, Greece. IEEE.

Yu, S., Guo, S., & Stojmenovic, I. (2012). Can we beat legitimate cyber behavior mimicking attacks from botnets? Proceedings IEEE INFOCOM, March 25–30, Orlando, FL, USA. IEEE.

Zhang, C., Cai, Z., Chen, W., Luo, X., & Yin, J. (2012). Flow level detection and filtering of low-rate DDoS. Computer Networks, 56(15), 3417–3431.

Zhang, G., Jiang, S., Wei, G., & Guan, Q. (2009). A prediction-based detection algorithm against distributed denial-of-service attacks. In Proceedings of the 2009 international conference on wireless communications and mobile computing: Connecting the World wirelessly (pp. 106–110). New York, NY: ACM.

Zhang, J., & Wang, H. (2006). Detecting outlying subspaces for high-dimensional data: The new task, algorithms, and performance. Knowledge and Information Systems, 10(3), 333–355.

Zhong, R., & Yue, G. (2010). DDoS detection system based on data mining (pp. 62–65). Proceedings of the second international symposium on networking and network security, April 2–4.

Zhou, C. V., Leckie, C., & Karunasekera, S. (2010). A survey of coordinated attacks and collaborative intrusion detection. Computers & Security, 29(1), 124–140.

Zhou, W., Jia, W., Wen, S., Xiang, Y., & Zhou, W. (2014). Detection and defense of application-layer DDoS attacks in backbone web traffic. Future Generation Computer Systems, 38, 36–46.

Zu, Q., & Hu, B. (2016). Human centered computing: Second international conference, HCC 2016, Colombo, Sri Lanka, January 7–9, 2016, Revised selected papers. Springer International Publishing.