DDoS Discrimination by Linear Discriminant Analysis (LDA)

Theerasak Thapngam, Shui Yu and Wanlei Zhou
School of Information Technology, Deakin University
Burwood, VIC 3125, Australia
Email: {tthap, syu, wanlei}@deakin.edu.au

Abstract— **In this paper, we propose an effective approach with a supervised learning system based on Linear Discriminant Analysis (LDA) to discriminate legitimate traffic from DDoS attack traffic. Currently there is a wide outbreak of DDoS attacks that remain risky for the entire Internet. Different attack methods and strategies are trying to challenge defence systems. Among the behaviours of attack sources, repeatable and predictable features differ from source of legitimate traffic. In addition, the DDoS defence systems lack the learning ability to fine-tune their accuracy. This paper analyses real trace traffic from publicly available datasets. Pearson's correlation coefficient and Shannon's entropy are deployed for extracting dependency and predictability of traffic data respectively. Then, LDA is used to train and classify legitimate and attack traffic flows. From the results of our experiment, we can confirm that the proposed discrimination system can differentiate DDoS attacks from legitimate traffic with a high rate of accuracy.**

Keywords; DDoS attacks; correlation coefficient; entropy; Linear Discriminant Analysis, traffic patterns; learning machine;

I. INTRODUCTION

Today, Distributed Denial of Service (DDoS) attacks are serious threats to computer hosts on the Internet. A recent report by [1] has revealed the largest attack size has doubled year after year, to more than 100 Gbps, which is a surprising 1000% increase in attack size since 2005. The attacks can be carried out by a large number of compromised hosts, called *zombie armies*. These hosts become the attack tools associated with performing DDoS attacks and are the reason why legitimate users experience a decline or absence of their service. Moreover, mimicking DDoS attacks [2] place more pressure on the defence system to the differentiate attacks as opposed to legitimate flows (*flash crowd*). On an individual attack source, the packet transmission is crafted in a random fashion. This method helps the attack traffic flies under the radar through the victim.

Research on DDoS detection has been able to identify DDoS attack packets and/or traffic flows. In the addition to detection of common attacks, DDoS detection uses statistical and heuristic analysis methods to identify attacks in progress. While, the key to any statistical-based detection system (SBDS) [3-5] is its ability to learn and distinguish *normal* from *anomalous* network activity, the heuristic-based detection system (HBDS) [2, 6-8] relies on optimisation of its *threshold* decision. Moreover, HBDS needs fine-tuning to produce stability, and improvement results of anomaly detection in network traffic and minimise false positives/negatives. This configuration process needs manual and semi-automatic adjustment for its threshold. Hence, creating an autonomous DDoS detection system with a full learning algorithm is still rarely implemented in this research area.

As we follow the assumption from our previous paper [8], the action of a DDoS attack source follows the instructions programmed by an attacker. Its function is repeatable to generate and transmit DDoS attack packets to the reflector/victim. This repeatable behaviour is apparently different from behaviour of a human user or an Internet proxy. Since we can extract the repeatable feature of an attack source, we can measure the degree of predictability in behaviour of packet arrival.

In this paper, we propose a solution to discriminate DDoS using the supervised learning model from the pattern behaviour of traffic sources by observing packet arrivals. This proposed technique is an effective method to discriminate traffic flows among DDoS attack sources and legitimate users. The packet arrival rate is detected as measurement data. We can success in the measurement of predictability features using Pearson's correlation coefficient and Shannon's entropy. We also classify those features into two groups (attack and legitimate) using Linear Discriminant Function as we explain later in the next section. Since we can measure the degree of pattern behaviour, the flows from the attack sources must be filter out, but legitimate flows must get through the server.

The contributions of this paper are listed as follows:

- *Reliability*: Our DDoS discrimination system maximises the accuracy in detecting and minimising a false positive rate (FPR) and a false negative rate (FNR) in the results. By using the triple check from Pearson's correlation, the Shannon's entropy and LDA, the statistical relationship of a traffic flow is measured with a high rate of accuracy.
- *Flexibility*: Our DDoS discrimination system could be detect any form of attack packets, such as malformed IP, TCP, UDP, ICMP, Application-based floods, etc. Our detection methods also work well with low-rate attack, flash-crowd attack, shrew attack, periodical attack, and pulsing attack.
- *Ability to Learn*: Our DDoS discrimination system may be enhanced to learn the classification based on its knowledge. Since we have measurement decision modules to provide a double check for the accuracy of results, the knowledge from feedback can be reuse by the training algorithm. Hence, the amendable knowledge is supervised to the classification module and maximises the accuracy.

The rest of this paper is organized as follows. Section 2 reviews the related work of our research. In Section 3, we summarise the mathematical tools related to our research. Section 4 discusses the problems of DDoS attacks and the specific challenges these raise. In the next section, we propose a solution with the system modelling. In Section 6, we provide the results of our experiment with publicly available datasets. In the final section, we provide a summary and discuss the direction of our future work.

978-1-4673-0009-4/12/$26.00 ©2012 IEEE

II. Related Work

Both a heuristic-based detection system (HBDS) and a statistical-based detection system (SBDS) have their limitations. While the key to any SBDS is its ability to learn and distinguish normal from anomalous network activity, the HBDS relies on optimisation of its threshold decision and requires fine-tuning to produce stability, improvement, or precise results of anomaly detection in network traffic and minimise the false positives/negatives.

An anomaly detection system [7] deploys a Support Vector Machine (SVM) and a Dynamically Growing Self-Organizing Tree (DGSOT) in order to create a learnable algorithm and deploy a clustering analysis respectively. The approach was compared with the Rocchio Bundling technique and random selection in terms of accuracy loss and training time gain. SVM + DGSOT achieve the learning system for DDoS detection, however, the results are very low in accuracy (69.8%), very long in training time (13.18 hours), very high in its false negative rate (FNR) (37.8%), and very high in its false positive rate (FPR) (29.8%).

Human-vs.-bot differentiation [6] by human behaviour modelling was proposed with an adjustable threshold. This approach derives three aspects of human behaviour: 1) request dynamics, 2) request semantics and 3) the ability to process visual cues. This heuristic approach discovered the flash-crowd attacks hidden in test traffic with a high accuracy rate of around 95-99%. The adjustable threshold of 0.05 gave a low FNR (0.00%) and low FPR (1.94%).

A DDoS detection based on Chaos Theory [2] deploys the theory of network self-similarity to differentiate DDoS flooding attack from legitimate self-similar traffic in the network. The authors developed a neural network detector for training by DDoS prediction algorithm. When the threshold of sensitivity had a range of 88% to 94%, FPR is varied with a range of 0.05% to 0.45%.

III. Mathematical Models

Based on data from arrival rates, we need mathematical models to identify the degree of prediction. Since we categorize data into predictable and unpredictable data, the mathematical models must be able to judge the data by using a threshold. These possible models are the tools for self-similarity analysis such as the correlation coefficient and distance matrix. In this paper, we deploy Pearson's correlation coefficient, Shannon Entropy and Linear Discriminant Analysis as detection and discrimination tools.

A. Pearson's Correlation Coefficient

In this paper, we select a feature that can measure the degree of dependence in data by using *Pearson's correlation coefficient* [9] (here after called *the correlation*), which is defined as:

$$\rho_{X,Y} = \frac{\Sigma(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} \tag{1}$$

The correlation is used to measure dependence between two quantities (variables) X and Y with expected values μ_X and μ_Y and standard deviations σ_X and σ_Y. Both value of the standard deviations are finite and nonzero ($0 < \sigma_X < \infty$ and $0 < \sigma_Y < \infty$). In this paper, we define the data that gives us this value of 1 ($|\rho_{X,Y}| = 1$) as predictable data with a linear form. The value of 0 ($|\rho_{X,Y}| = 0$) defines unpredictable data in *linear form.*

B. Shannon Entropy

We select another feature that can measure the degree of uncertainty in data by using *Shannon entropy* (here after called *the entropy*), which is defined as:

$$H(X) = -\sum_{i=1}^{N} p(x_i) \log p(x_i) \tag{2}$$

The entropy is a measure of uncertainty/unpredictability for a random variable $X = \{x_i : i = 1, 2, 3, \dots, N\}$ where $p(x_i)$ is the probability mass function of outcome x_i. The entropy value is between 0 and $\log N$ ($0 \le H(X) \le \log N$). The entropy value is equal 0 ($H(X) = 0$) represented by the high *predictability* of the random variable X. However, the entropy value could reach its maximum value ($H(X) = \log N$) if the random variable X is high in *uncertainty/unpredictability*. In this paper, we define the data that gives us an entropy value of 0 ($H(X) = 0$) as *predictable* data. On the contrary, the value of maximum entropy ($H(X) = \log N$) defines *unpredictable* data.

C. Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is a statistical approach in classifying objects (people, things, events, etc.) based on their set of features that can be placed in two or more characteristic groups. A feature must be defined as an observation, property, attribute, variable or measurement of an object. In the training process, groups are known or predetermined and do not have order (i.e. nominal scale). A feature set of those objects is then measured which helps to solve the classification problem.

For classification of g groups, the Bayes' rule [10] could minimise total errors by assigning the object to group i which expresses the highest conditional probability against group j:

$$P(i|x) > P(j|x)\,, \qquad for\, \forall\, j \neq i \tag{3}$$

However, we cannot achieve the probability $P(i|x)$ of the class directly from the given measurement of the object. We can obtain the measurement and compute the probability for each class by knowing $P(x|i)$ as given in the Bayes Theorem [11]:

$$P(i|x) = \frac{P(x|i)\cdot P(i)}{P(x)} = \frac{P(x|i)\cdot P(i)}{\sum_{\forall j} P(x|j)\cdot P(j)} \tag{4}$$

Thus, the Bayes' rule (Eq. 3) becomes:

$$\frac{P(x|i)\cdot P(i)}{\sum_{\forall k} P(x|k)\cdot P(k)} > \frac{P(x|j)\cdot P(j)}{\sum_{\forall k} P(x|k)\cdot P(k)}\,, \qquad for\, \forall\, j \neq i \tag{5}$$

We can simplify the Eq. 5 as follow:

$$P(x|i)\cdot P(i) > P(x|j)\cdot P(j), \qquad for\, \forall\, j \neq i \tag{6}$$

We assign the object to group i if the proposed formula (Eq. 6) is satisfied. If we have multiple classes and multiple dimensions of measurement, with each dimension having many values, the computation of conditional probability $P(x|i)$ requires a large amount of data. In the experiment, we assume that data comes from some theoretical distribution. In this paper, we assume that the data comes from the multivariate Gaussian distribution [11] in the following formula:

$$P(x|i) = \left(\frac{1}{(2\pi)^{\frac{k}{2}}|\Sigma_i|^{\frac{1}{2}}}\right) exp\left(-\frac{1}{2}(x-\mu_i)^T \Sigma_i^{-1}(x-\mu_i)\right) \tag{7}$$

where, μ_i is vector mean and Σ_i is the covariance matrix of group i. Replacing the distribution formula (Eq. 7) into Bayes' rule in Eq. 6 and simplifying by several steps (see [10] for more details) create a new formula:

$$\ln\big(P(i)\big) + \mu_i\Sigma^{-1}x^T - \frac{1}{2}\mu_i\Sigma^{-1}\mu_i^T > \ln\big(P(j)\big) + \mu_j\Sigma^{-1}x^T - \frac{1}{2}\mu_j\Sigma^{-1}\mu_j^T, \qquad i \neq j \quad (8)$$

Denote that:

$$f_i = \ln\big(P(i)\big) + \mu_i\Sigma^{-1}x_k^T - \frac{1}{2}\mu_i\Sigma^{-1}\mu_i^T, \qquad i \neq j \quad (9)$$

By deriving the Eq. 8 and 9, we then have the Linear Discriminant Function:

$$f_i > f_j,\ \ for\ \forall i \neq j \qquad (10)$$

We could assign object measurement x to group i if Eq. 10 is satisfied. In this paper, we select the two features (data dependence ($|\rho_{X,Y}|$) and data predictability ($H(X)$)) as measurement x. Then we use LDA to train and classify measurement x into two different groups based on its nature. We explain this in detail in Section 5.

IV. Problem Statement

We considered the situation when a server derives suspicious traffic flow via its router. As illustrated in Fig. 1, a server connects to the Internet and provides a service to public Internet users. In a normal situation, the server can handle legitimate users with limited resources such as CPU process and memory/buffer. However, a busy server may suffer a *flash crowd* (FC) event, which is observed as a sudden high demand in service requests from Internet users. This FC event forces the server to work hard with its limited resources. The FC flow may also overwhelm the server and create a Denial of Service (DoS) condition which results in either a delay of response or a complete crash. Hence, the FC flow must be monitored and considered as a suspicious flow when it arrives at the server.

Since the flow could be either legitimate traffic and/or DDoS attack traffic, we treat the flow as suspicious flow, which has the potential to harm the server. The process of investigation and mitigation begins by taking a sample of the arrival rate (λ_k) of individual traffic from an individual source. Then we measure the degree of dependency and predictability of an individual flow. If the flow can be classified as a dependent and predictable flow, then it is an attack. If the flow can be classified as an independent and unpredictable flow, then it is a legitimate flow of traffic. Otherwise, the flow remains unknown and is considered suspicious.

V. System Modelling

The goal of this paper is to classify a suspicious flow into either legitimate or DDoS attack traffic. To satisfy this investigation process, the following system model (as shown in Fig. 2) will be proposed:

A. *Dependency*

In the process of dependency measurement of the arrival rate (λ_k), we nominate the correlation using Eq. 1. We denote X as a sample set of arrival rates ($X = \{\lambda_k : k = 1, 2, \ldots, N\}$), and Y as a sample set of sequence numbers ($Y = \{k : k = 1, 2, \ldots, N\}$). Then we calculate the correlation value ($\rho_{X,Y}$) from the two variables (X and Y). The value that we expect is between -1 and 1 ($-1 \leq \rho_{X,Y} \leq 1$). Then we pass the dataset of absolute values of the correlation ($|\rho_{X,Y}| = \{|\rho_1|, |\rho_2|, \ldots, |\rho_k| : k = 1, 2, \ldots, N\}$) to the process of LDA.

B. *Predictability*

In the process of predictability measuring of arrival rate (λ_k), we nominate entropy using Eq. 2. We denote X as a sample set of arrival rates ($X = \{\lambda_k : k = 1, 2, \ldots, N\}$). Then we calculate the entropy value $H(X)$ from the variable (X). The value that we expect is between 0 and $\log N$ ($0 \leq H(X) \leq \log N$). By default, we initial the number (N) of sample arrival rate to 10. Then we can expect the entropy to be between 0 and 1 (derived from $\log N = \log 10 = 1$). Finally, the dataset of the entropy ($H(X) = \{H_1, H_2, \ldots, H_k : k = 1, 2, \ldots, N\}$) is passed to the process of LDA.

C. *Linear Discriminant Analysis (LDA)*

In the process of LDA, we derive two features of the arrival rate (λ_k) which are processed by the correlation ($|\rho_{X,Y}| = \{|\rho_1|, |\rho_2|, \ldots, |\rho_k| : k = 1, 2, \ldots, N\}$) and the entropy ($H(X) = \{H_1, H_2, \ldots, H_k : k = 1, 2, \ldots, N\}$) modules. Then these datasets are processed in one of the following phases:

1) Training phase: One of the LDA processes is to let the discrimination system learns the classified (known) object category. We provide some samples from the DDoS attack datasets and legitimate FC datasets to the system. This is an important process for selecting the right training data for our discrimination system. Then, the LDA will create its threshold for our system for the purposes of decision making.

2) Classification phase: Since the discrimination system supervises the nature of data, it will have a general knowledge to classify the test data that we need to investigate. By using the LDA threshold ($\mathcal{T}$), we can measure the accuracy of our system.

In the process of LDA, we calculated the two features using the Linear Discriminant Function (Eq. 9). If the coordinate results ($f(x) = \{(f_i, f_j) : \forall i \neq j\}$) are in a training phase, this will create the threshold ($\mathcal{T}$) for further evaluation measurement. If the coordinate results ($f(x) = \{(f_i, f_j) : \forall i \neq j\}$) are in classification phase, they will be used to compare and classify using the threshold.

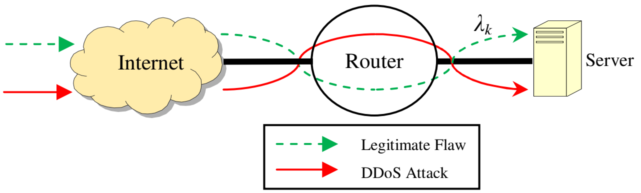

Figure 1: Arrival rate λ (packet/interval) from *k* IP address(es)

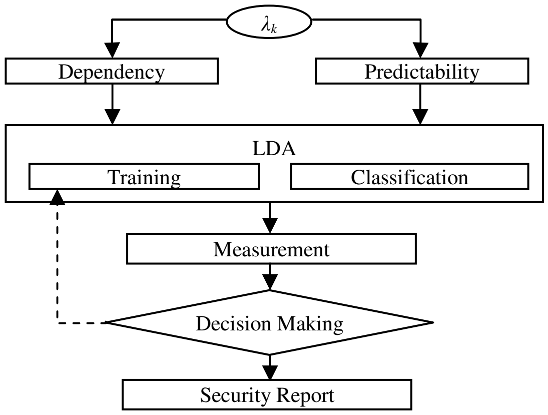

Figure 2: Architecture of a traffic discrimination system

D. *Measurement*

The measurement approach against threshold (τ) for the classification process is similar to the training process. However, the measurement data (x) is tested against the criteria of probability ($P(i)$) and the vector mean (μ_i) of group i plus the covariance matrix (Σ) of all groups. We also tested the measurement data (x) with group j as shown in Fig. 3. Then the results were evaluated for the accuracy as shown in Table 1.

E. *Decision Making*

The module of decision making is based on the threshold (τ) from the training datasets. As shown Table 1 and Fig. 3, if the measurement data (x) satisfies $f_i > f_j$ (Eq. 10), then it is classified into group i. Otherwise, it is in group j. In our paper, we classify group i and j as legitimate traffic and attack traffic respectively. Finally, we inform the next module on order to generate a report of the measurement result.

In the case of false detections increasing, this module gives us a double check from the previous and on-going behaviours for measurement data (x). We are looking forward to developing a learning discrimination system using its feedback as shown in Fig. 2. This advanced technology could provide us with an artificial intelligence (AI) of a smart DDoS detection system.

F. *Security Report*

The last part of the discrimination system is the module of security report. After a decision has been made by the decision making module, an action command/message is then generated to order/inform the security system. We are looking forward to customising this report generation and serving various types of intrusion detection systems (IDS) and intrusion prevention systems (IPS).

VI. Evaluations

In order to measure the discrimination performance, we have to test 2 datasets from real traces (the World Cup 98 website (WC) [12] and the MIT project of MStream attacks (MIT) [13]). We took each sample flow from legitimate and attack traffic from these datasets for the training of our system. Then we took another traffic flow from each dataset to test the processes of classification and measurement. The results of the discrimination processes are as follows:

A. *Results of Training*

The result of the training process derives from the measurement data (x) which is passed the correlation and entropy modules. We selected the WC200657 and WC555 as a legitimate flow (group i) and an attack flow (group j) respectively, as depicted in Fig. 4. Then, we trained the system to know their groups and create their own threshold as depicted in Fig. 5. As shown in Table 2, the training process can accurately discriminate between the legitimate and attack flows 100% of the time. This means that the calculated threshold is perfect for discriminating a legitimate flow from an attack flow in the early stage of the training process.

We can also confirm the high performance of discrimination by the statistics in Table 4. The threshold gives us the equal average distance (1.327) between the legitimate group and the threshold, and between the attack group and the threshold. Since we found a good threshold for our discrimination system, we expect this threshold to work well for the test datasets.

Table 1: List of accuracy measurement

$f_i > f_j$?	Result	Fact	Accuracy
True	Group i	Group i	True Positive
True	Group i	Group j	False Negative
False	Group j	Group i	False Positive
False	Group j	Group j	True Negative

Table 2: Results of training

Accuracy	Positive	Negative
True	100.0%	100.0%
False	0.0%	0.0%

Table 3: Results of classification

Accuracy	Positive	Negative
True	100.0%	92.3%
False	0.0%	7.7%

Table 4: List of LDA scores from Fig. 6 and 9

LDA Data	Average	Min.	Max.
Train Leg.	1.327	0.979	1.614
Train Att.	-1.327	-2.156	-0.750
Test Leg.	1.142	0.273	2.089
Test Att.	-1.829	-4.306	0.670

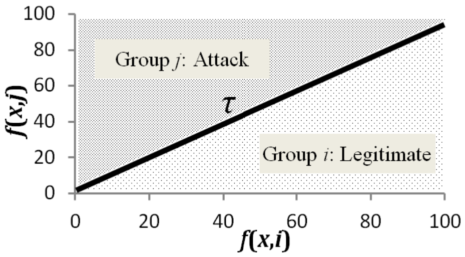

Figure 3: Threshold of Decision Making

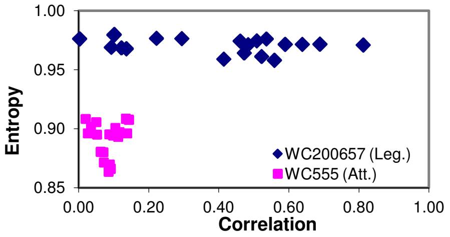

Figure 4: Scatter plot of the original data for training

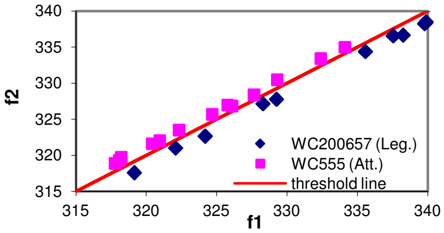

Figure 5: Zooming scatter plot with threshold from the LDA training process

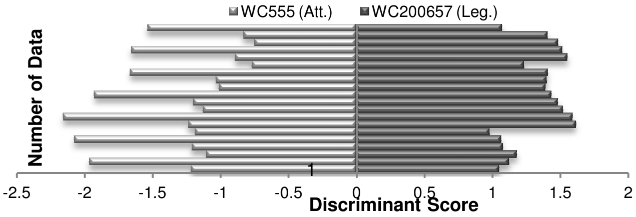

Figure 6: LDA score from the training process

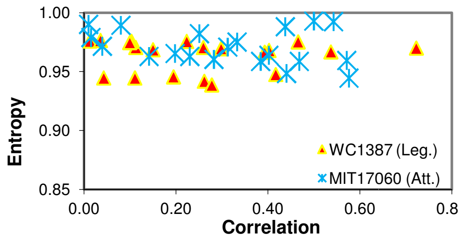

Figure 7: Scatter plot of the original data for classification

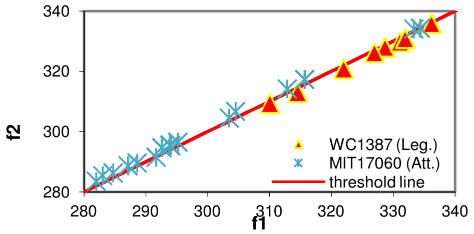

Figure 8: Zooming scatter plot with the threshold from the LDA classification process

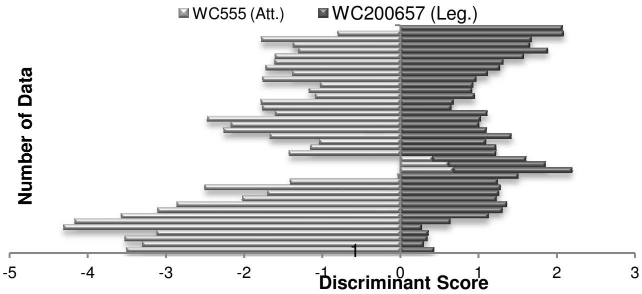

Figure 9: LDA score from the classification process

B. Results of Classification

The results of classification process derive from the measurement data (x) which is also passed by the correlation and entropy modules. We selected the WC1387 and MIT17060 as a legitimate flow (group i) and an attack flow (group j) respectively as depicted in Fig. 7. We measured the accuracy of the classification process against the trained threshold as depicted in Fig. 8. As shown in Table 3, the classification process can accurately discriminate the legitimate flow 100% of the time. However, the attack flow can only be detected with a rate of 92.3% accuracy and false negative of 7.7%. This means the legitimate flow has unique behaviour regarding unpredictable packet transmission as well as a high accuracy for detecting DDoS attack flow that can protect the system of the server.

We can also confirm the high performance of discrimination by the statistic from Table 4. The distance measurement gives us the absolute average LDA score (1.829) of an attack group that is higher than the absolute average LDA score (1.142) of the legitimate group. This means that the discrimination system has a high level of differentiation in the attack flow.

VII. Conclusion and Future Work

In this paper, we proposed an effective approach with a supervised learning system based on a Linear Discriminant Analysis (LDA) to discriminate legitimate traffics from DDoS attack traffic. We introduce a new LDA approach in this research area against DDoS attack. This solution helps reduce the complexity and configuration of the discrimination system. In addition, this approach also adopted a method of detecting human-based behaviour when browsing websites which is a huge difference from predictable behaviour of DDoS attack. This idea could detect DDoS attacks regardless of the types of attack packets and transmission methods.

Our discrimination system consists of 2 major stages (training and classification). The correlation and entropy extract the features of data dependency and predictability respectively. Then, we deliver these features into the LDA as its measurement data. We perform the experiment with famous trace datasets. The result of this process was good enough to create its own effective threshold. At the final stage, we use the threshold from the previous process to classify the groups as legitimate or attack flow. This classification process provides us with a total accuracy of 96.2%, with 100% accuracy in the detection of legitimate flows.

As future work, we are looking forward to improve the accuracy in discrimination with the double checks in decision making module. This may give us more accuracy but we may not implement this additional module in real time detection.

References

[1] Arbor Networks, "Worldwide Infrastructure Security Report: 2010 Report," Arbor Networks, 2011.

[2] A. Chonka, J. Singh, and W. Zhou, "Chaos Theory Based Detection against Network Mimicking DDoS Attacks," *IEEE Communications Letters,* vol. 13, pp. 717 - 719, 2009.

[3] Y. Xie and S. Z. Yu, "A Large-Scale Hidden Semi-Markov Model for Anomaly Detection on User Browsing Behaviors Networking," *IEEE/ACM Transactions on Networking,* vol. 17, p. 12, 2009.

[4] Y. Xie and S. Z. Yu, "Monitoring the Application-Layer DDoS Attacks for Popular Websites," *IEEE/ACM Transactions on Networking,* vol. 17, p. 11, 2009.

[5] F. Yi, S. Yu, W. Zhou, J. Hai, and A. Bonti, "Source-Based Filtering Scheme against DDOS Attacks," *International Journal of Database Theory and Applications,* vol. 1, pp. 9-22, 2008.

[6] G. Oikonomou and J. Mirkovic, "Modeling Human Behavior for Defense against Flash-Crowd Attacks," in *IEEE International Conference on Communications 2009 (ICC '09)*, 2009, pp. 1 - 6.

[7] L. Khan, M. Awad, and B. Thuraisingham, "A new intrusion detection system using support vector machines and hierarchical clustering," *the International Journal on Very Large Data Bases (The VLDB Journal),* vol. 16, pp. 507 – 521, 2007.

[8] T. Thapngam, S. Yu, W. Zhou, and G. Beliakov, "Discriminating DDoS Attack Traffic from Flash Crowd through Packet Arrival Patterns," in *the 30th Annual IEEE International Conference on Computer Communications (IEEE INFOCOM 2011)*, Shanghai, China, 2011, pp. 969 - 974.

[9] M. Natu and J. Mirkovic, "Fine-Grained Capabilities for Flooding DDoS Defense Using Client Reputations," in *the ACM SIGCOMM Large-Scale Attack and Defense Workshop (LSAD) in* 2007.

[10] Y. Chen, K. Hwang, and Y.-K. Kwok, "Filtering of Shrew DDoS Attacks in Frequency Domain," in *the First IEEE LCN Workshop on Network Security (WoNS 2005)*, Sydney, 2005.

[11] Y. Chen and K. Hwang, "Spectral Analysis of TCP Flows for Defense against Reduction-of-Quality Attacks," in *the 2007 IEEE International Conference on Communications (ICC'07)*, 2007, pp. 1203–1210.

[12] Y. Chen and K. Hwang, "Collaborative detection and filtering of shrew DDoS attacks using spectral analysis," *Journal of Parallel and Distributed Computing,* vol. 66, pp. 1137-1151, 2006.

[13] E. Kreyszig, *Advanced Engineering Mathematics*, nineth ed.: Wiley, 2006.

[14] K. Teknomo. (2006). *Discriminant Analysis Tutorial*. Available: http://people.revoledu.com/kardi/ tutorial/LDA/

[15] R. D. Yates and D. J. Goodman, *Probability and Stochastic Processes*, second ed. NJ, USA: John Wiley & Sons, 2005.

[16] M. Arlitt and T. Jin. (1998). *1998 World Cup Web Site Access Logs*. Available: http://www.acm.org/sigcomm/ITA/

[17] MIT Lincoln Laboratory. (1999). *Lincoln Laboratory Scenario (DDoS) 1.0*. Available: http://www.ll.mit.edu/mission/communications/ist/corpora/ideval/data/2000/LLS_DDOS_1.0.html.