FORECASTING DISTRIBUTED DENIAL OF SERVICE ATTACK USING HIDDEN MARKOV MODEL

[1]Afolorunso, A. A., Abass O. Longe, H. O. D, Adewole, A. P.

[1]School of Science and Technology, National Open University of Nigeria, Lagos, Nigeria
[1,2,3,4]Department of Computer Sciences, University of Lagos, Lagos, Nigeria

ABSTRACT
Distributed denial of service (DDoS) attack bombards the network with loads of packets and requests that consumes the system resources in terms of time, memory, and processors. This paper presents a proposed method for forecasting DDoS in networks. The proposed model employs hidden Markov model (HMM) to forecast DDoS attacks. The method uses the inherent characteristic features of DDoS to determine the observable states of the system. To avoid intractable computations, Kullback-Leibler divergence algorithm was employed to reduce the number of observable states to three. The proposed model is formulated and trained through experiments using DARPA 2000 data set and the preliminary results shows that the characteristic features of the DDoS and the entropy concept can be used to formulate an HMM to predict DDoS.

***Keywords*:** Distributed denial of service, Forecasting, Hidden Markov model, Kullback-Leibler divergence.

1.0 INTRODUCTION

Globalinter connectivity has made private networks accessible to outsiders. This has created the issue of security especially in the competitive industries where malicious rivals may gain access into other organisations in order to cause damage. It has therefore become imperative to find ways of securing such networks from malicious attacks. Intrusion Forecasting Systems (IFS) evolved as a proactive method of stopping damage to systems. Whereas IDS alerts about intrusion after damage might have been done, IFS will alert before the attack takes place. There are different kinds of intrusion into systems such as DDoS attacks, Worms, Domain Name System (DNS) attacks, router attacks among many others Debar H., Dacier M., Wespi A. (1999), Kim, G., Lee, S. and Kim, S. (2014)). As stated earlier, there are different forms of attacks into systems and each has its peculiar patterns and signatures(Debar, et al. (1999). In DDoS attacks, unlike other types of attack, the attacker does not use the holes in the system security but rather he/she launches attack against its availability. Here an attacker compromises a large number of connected systems by exploitation of network software vulnerabilities (Vinchurkar, D. P., and Reshamwala, A. 2012). After attack software is installed on these systems through secured channels, the compromised hosts on which attack software is installed simultaneously bombards the victim with unwanted (use another word) packets. The amount of malicious traffic generated by such hosts becomes so much that a victim's system resources are used up and it cannot attend to legitimate requests and hence becomes paralyzed.

DDoS progresses in stages and can therefore be said to have different phases. At each phase there are some observable events that occur and these events can be used to predict the state of the system and what could happen in the system in the foreseeable future. According to the experiments run by the MIT Lincoln Lab (MIT Lincoln Lab, 2000) in which a DDoS attack was run by a novice attacker over multiple networks and audit sessions, DDoS attack session can be grouped into five phases as follows:Lee, K., Kim, J., Kwon, K. H., Han, Y., Kim, S.(2008)

1) IPsweep to the DMZ(demilitarized zone) hosts from a remote site.
2) Probe of live IP's to look for the sadmind daemon running on Solaris hosts.

3) Breaks-in via the sadmind vulnerability, both successful and unsuccessful on those hosts.
4) Installation of the Trojan mstream DDoS software on three hosts in the DMZ.
5) Launching the DDoS.

During each phase, there are certain features that could be observed in the network traffic. During the first phase, for instance, the attacker sends a lot of spurious Internet Control Message Protocol (ICMP)echo trying to get security vulnerable hosts that would act as agents and handlers for the attack. It is, therefore, possible to predict DDoS attack by studying the traffic features of the network at each stage.

In this study, the entropy concept is employed to analyse the network traffic at each phase of the attack. The network features at each phase are then used as the observable parameters in the proposed forecasting method. The rest of the paper is organised as follows: Section 2 presents previous research works related to this study; section 3 presents the proposed method; simulation experiments are presented in Section 4; Section 5 is the results, while conclusion and future work are presented in section 6 and 7 of the paper respectively.

2.0 RELATED WORKS

Jemili, F., Zaghdoud, M. and Ahmed, M. B. (2009) proposed an intrusion detection and prediction system which recognizes an upcoming intrusion and predicts the attacker's attack plan and intentions. Graph techniques were applied based on Bayesian reasoning for learning. Also, inference was applied to recognize the attack type and predict upcoming attacks. The inference process is based on hybrid propagation, which takes into consideration both the uncertain and imprecise character of information. While the system demonstrates high performance in detecting intrusions, correlating and predicting attacks, there were still some challenges in attack plan recognition and, of course, the system worked in a static environment.

Pontes, E. and Guelfi A. E. (2009) showed a collaborative architecture of IDS with prediction approaches, covering the gaps of the current forecasting techniques. A proof of concept of the architecture was presented, which allows conclusion about the improvement in forecasts for IDS to cope with Unwanted Internet Traffic (UIT).

Gao, Bo, Hui-Ye Ma, and Yu-Hang Yang (2002) developed an Hidden Markov Model (HMM) to predict attacks in the application layer, an approach, which they claimed could be extended for network layer.

Arnes, A., Valeur, F., Vigna, G. and Kemmerer, R. A. (2006) used to represent the likelihood of transitions between security states. The models parameters were manually estimated, which was more tedious and prone to error than using of a training algorithm.

Haslum, K., Abraham, A, and Knapskog, S (2009) used an HMM that models only integrity and confidentiality, and make no attempts to model availability. They believe that availability is best modeled separately. Preliminary experimental result from the system indicates that their proposed framework is efficient for real time distributed intrusion monitoring and prevention.

Lee et al (2008) propose a method for proactive detection of DDoS attack by exploiting the characteristic nature of its architecture comprising of handlers and agents selection, communication and compromise, and finally, attack. Cluster analysis was performed for proactive detection of the attack. The method demonstrated good performance in detecting precursors of DDoS attack as well as the attack itself. However, there is need to extend the work to predict different types of DDoS attack and datasets.

Kim, S., Shin, S., Kim, H., Kwon, K. and Han, Y. (2010) worked on a hybrid approach that combined time-series analysis, probabilistic modeling (Markov Chain model) and data mining method to forecast intrusion. Experimental results showed that among the three hybrid methods, the combination of the time-series analysis and the Markov chain method shows the best performance in reducing the false alarm rate. They concluded that there is need to develop new intrusion forecasting methods that provide improved accuracy of predictions with a lower false alarm rate as well as developing an alert correlation algorithm between each forecasting method in order to achieve earlier and more precise forecasting of attacks.

An, X. and Jutla, D. (2006) proposed a model that applied DBN to privacy intrusion detection in temporal environments. The approach's objective was to handle general internal attacks on databases to steal large volume of private data. However, the model was not formulated from real data.

Shin, S., Lee, S., Kim, H. and Kim, S. (2013) proposed a probabilistic approach to effectively forecast and detect network intrusions. The approach uses a Markov chain for probabilistic modeling of abnormal events in network systems. K-means clustering was performed to define the network states, with the introduction of the concept of an outlier factor. The degree of abnormality of the incoming data was stochastically measured, in real-time, based on the defined states. The performance of

the approach was evaluated by experiments using the DARPA 2000 data set. Although in the work, the various features of the DDoS attack and the concept of entropy were mentioned, it was not used in the formulation of the model. Also, the need to improve the accuracy of prediction of the model was mentioned.

Flores, J. J., Antolino, A. and Garcia, J. M. (2010) proposed a system that performs network anomaly detection through the use of Hidden Markov Models (HMMs). Genetic Algorithms (GAs) were used in designing and training the HMMs used in detecting anomalies. This helps in the automation of the use of HMM because with it, users do not have need for statistical knowledge necessary for software that trains HMMs from data. The GA determines the necessary parameters such as, number of states, connections and weights, and probability distributions of states. Comparing the results obtained with those from Baum-Welch algorithmproved that, in all tested cases, GA outperforms Baum-Welch (Flores, et al (2010)). The best of the resulting HMMs was used to perform anomaly detection in network traffic activity with real data.

Considering the previous schemes, there is commonlytradeoff between prediction efficiency and cost. Increasing the prediction accuracy often leads to increase in false alarm rate or increase in computational overheads or memoryoverheads. While accurate prediction is very important for preparing defense measures in DDoS attacks, most of the previous researches have been focused on the traffic generated by agents to extract prediction parameters. It is valuable to study and analyze the traffic generated during attack preparation phases as well as that generated during attack phases for optimal attack prediction. Hence, the need to develop a methodto predict the possibility of intrusion in time for steps to be taken to avert it without adding too much overhead in terms of resources consumption which might adversely affect the performance of the system.

3.0 THE PROPOSED METHOD

In this section, the Hidden Markov Model and its major properties are described briefly. After which the problem to be solved in this paper is clearly reviewed.

Prediction of network intrusion, of which DDoS attack is a type, is generally viewed as a pattern recognition problem, and as such, it can be solved using one of two broad approaches: structural and empirical (Kumar and Ravi, 2007). The former derives the probability of a network for default based on its characteristics and dynamics, while the latter approach relies on previous knowledge and relationships in the area under study, learns from existing data or experience, and deploys statistical or intelligent methods to predict intrusion. The HMM that is employed in this work falls under intelligent methods.

Just as several methods had been applied to forecasting DDoS attacks, HMM had been applied to forecasting network intrusion, the difference had been in the methods used in determining the observable states of the system.

3.1 Hidden Markov models (HMMs)

A HMM is a temporal probabilistic model with two embedded stochastic processes: an unobservable (hidden) process S, which can be observed only through another (visible/observable) stochastic process O. Each state in Q (the set S of hidden states)has state-transition probabilities (which are not visible) and a probability distribution over the possible values of O. The key assumption is that the current hidden state of the system is affected only by its previous state.

HMM is usually defined as a 5-tuple (S, O, P, Φ, π), (Ibe, O. C. (2009)) where:

$S = \{s_1, s_2, \ldots, s_N\}$ is a finite set of N states

$O = \{o_1, o_2, \ldots, o_M\}$ is a finite set of M possible symbols

$P = \{P_{ij}\}$ is the set of state-transition probabilities where p_{ij} is the probability that the system goes from state s_i to s_j

$\Phi = \{\varphi_i(o_k)\}$ are the observation probabilities, where $\varphi_i(o_k)$ is the probability that the symbol o_k is emitted when the system is in state s_i.

$\pi = \{\pi_i\}$ are the initial probabilities; i.e. π_i is the probability that the system starts in state s_i.

Since the states and output sequence are understood, it is customary to denote the parameters of an HMM by $\lambda = (P, \Phi, \pi)$.

HMM can be formally represented as in Figure 1 below where S_i are the hidden states that we would like to estimate and the O_i are the observation random variables from which the S_i are to be estimated. The letters B and E indicate the *Beginning* and *End* of the sequence of states.

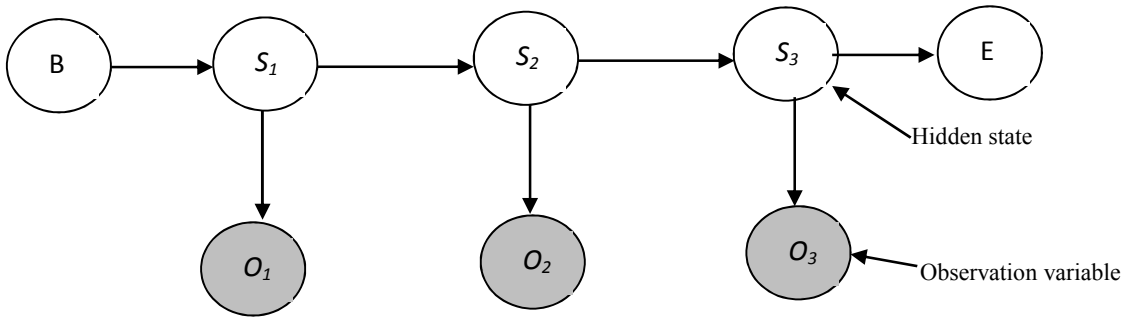

Fig. 1: General Structure of an HMM

There are three basic problems to be solved in HMM and each of these problems has specific solution methods/algorithm similar to those of Bayesian Networks (BNs) and Dynamic Bayesian Networks (DBNs).

- The Classification/Evaluation problem: this deals with how to efficiently compute the probability that a particular model $\lambda = (P, \Phi, \pi)$ generated a given observation sequence $V = v_1, v_2, \ldots, v_T$ of length T where $v_i \varepsilon O$.
- The Inference/Decoding problem: determining the most likely sequence of hidden states that could have generated a given observation sequence given a model $\lambda = (P, \Phi, \pi)$
- The Learning/Estimation problem: finding the model λ that best explains a given observation sequences. i.e. to estimate the most likely HMM parameters for a given observation sequence.

The first problem is usually solved by the application of the Forward-Backward algorithm (Ibe, O. C. (2009)).

The second HMM problem is solved by the application of a specific type of the forward algorithm named Viterbi algorithm. Though Viterbi algorithm (Ibe, O. C. (2009)), was originally designed for decoding convolutional codes, it is applied in many other areas. In HMM, it is used to find the most likely state sequence $Q^* = \{q_1^*, q_2^*, \ldots, q_T^*\}$ for a given observation sequence $O = o_1, o_2, \ldots, o_T$.

Let *function:*

$$\arg\max_y\{z\}$$

denotes the argument y that corresponds to the maximum of the expression z. The Viterbi algorithm simultaneously maximises both the joint probability $P[q, O]$ and the conditional probability $P[q|O]$. It defines the variable $\delta_t(i)$ as below:

$$\delta_t(i) = \max_{q1,q2,qt-1} P[q_1, q_2, \ldots, q_{t-1}, q_t = s_i, o_1, o_2, o_{t-1}, \ldots, o_t|\lambda]$$

This means $\delta_t(i)$ is the largest probability along a single path that accounts for the first t observations and ends in state s_i. It is therefore the probability of the most likely state path for the partial observation sequence. A variable c stores the node of the incoming edge that leads to this most probable path. i.e.

$$\psi_1(i) = \arg\max_{1\le i\le N}\{\delta_{t-1}(i)p_{ij}\}$$

The full algorithm is as in Ibe, O. C. (2009)

The third HMM problem, the learning problem, deals with how to adjust the HMM parameters so that the given set of observations, usually called the training set, is represented by the model in the best way for the intended application. It is an optimization problem that seeks to find the parameters of the HMM that maximise the probability of a given observation sequence. An iterative method called the Baum-Welch algorithm (sometimes called forward-backward algorithm) which is a special case of the Expectation Maximization (EM) method is employed to solve this. The algorithm is as in Ibe, O. C. (2009).

3.2 Selection of parameters

In order to effectively predict the DDoS attack, there is need to take cognisance of the DDoS attack right from the preparation stage. As written in the introduction, there are certain traffic features that could be observed at each stage of the attack. These traffic features could be used to predict the possibility of a DDoS attack. First, the traffic features throughout the entire phases of the DDoS is studied with a view of using them as the observable states of the HMM. Some of these features changes abnormally at each phase of the attack. In order to launch a DDoS attack, the attacker needs to select agents and handlers and this is usually done by sending ICMP Echo Request packets to find handlers and agents that could help attack (IPsweep). During this stage, a lot of ICMP traffic is transmitted to several host located on the Internet. Therefore, at this stage not only is there abnormally high rate of ICMP traffic occurrence compared to normal network traffic

but also randomly distributed destination IP address in the network flow. Next come the communication and compromise stage when increased volume of occurrence of a traffic type such as ICMP, UDP, and TCP SYN can be observed since any of these can be used in message exchange between the handlers and agents. Finally, comes the attack stage. At this stage, the reverse of the distribution of destination IP address and source IP address during the IPsweep stage is observed. i.e. the attack packets have randomly distributed source IP address focused on the destination IP address of the victim.

Lee et al (2008) employed the concept of entropy to measure the degree of divergence of these features. Entropy H for an information source with n independent symbols each with choice probability P_i is given as:

$$H = -\sum_{i=1}^{n} P(i) \log_2 P(i) \quad (1)$$

Entropy can be computed on a sample of consecutive packets. Comparison of entropy value of some sample of packet header fields to that of other samples of packet header fields provides a mechanism for perceiving changes in the randomness and can therefore be used as prediction parameter.

When entropy value is used, the value of source IP address becomes small and that of destination IP address becomes large in the IPsweep phase. On other hand, in the DDoS attack period, the entropy value of source IP address increases and that of destination IP address converges to a very small value.

The entropy values of source and destination port numbers can also be applied to predict DDoS attacks because some types of DDoS attacks use random port numbers in the attack period (Criscuolo, P. J. (2000)). In addition, the entropy value of packet type is worth observing because DDoS attacks use specific packet type such as ICMP flood attack and UDP flood attack (Houle, K. J., and Weaver, G. M. (2001), Criscuolo, P. J. (2000)). Convergence of packet type entropy to a small value is an indication that an attack is underway. As earlier mentioned, during the attack stage, the agents send huge amount of packets to the victim thereby jamming the network. Therefore the number of packets is a definite indication of an attack in progress.

From the aforementioned analysis, the parameters for DDoS attack prediction can be represented as follows:

- Entropy of source IP address
- Entropy of source port number
- Entropy of destination IP address
- Entropy of destination port number
- Entropy of packet type
- Occurrence rate of ICMP
- Occurrence rate of UDP
- Occurrence rate of TCP-SYN, and
- Number of packets

4.0 THE EXPERIMENTS

To implement the model, the **DARPA 2000 Intrusion Scenario Specific Datasets** (MIT Lincoln Lab. 2000) was used. Specifically, LLDOS 2.0 - Scenario Two which presents attack scenario data sets created for DARPA that includes a DDoS attack.

The downloaded data is the first attack scenario data set to be created for DARPA. It includes a distributed denial of service attack run by a novice attacker. It was reported that future versions of this and other example scenarios will contain more stealthy attack versions.(MIT Lab. 2000)

This attack scenario is carried out over multiple networks and audit sessions. The sessions have been grouped into five attack phases, over the course of which the attacker probes the network, breaks in to a host by exploiting the Solaris sadmind vulnerability, installs trojan mstream DDoS software, and launches a DDoS attack at an offsite server from the compromised host.

The proposed model was developed through a training process using both the DARPA 1999 attack-free dataset and the DARPA 2000 dataset with attack. From the data, the parameters listed in section 3.2above were estimated and the results are as presented in Table 1and 2 below.

Table 1: Estimates of DDoS Features

S/N	Variables	Attack free data (DARPA 1999)	Intrusion Specific data (DARPA 2000)
1.	Entropy of source IP address	1.59	0.52
2.	Entropy of source port number	1.61	0.41
3.	Entropy of destination IP address	1.58	5.12
4.	Entropy of destination port number	1.50	0.43
5.	Entropy of packet type	1.12	0.55
6.	Number of packets	37.0	42.3
7.	Occurrence rate of ICMP	0.0	0.87
8.	Occurrence rate of UDP	0.0	0.99
9.	Occurrence rate of TCP-SYN	0.02	0.02

Table 2: **Estimates of DDoS Features at each Phase**

Variables	Phase1	Phase 2	Phase 3	Phase 4	Phase 5
Entropy of source IP address	0.52	0.07	0.06	0.07	0.03
Entropy of source port number	0.41	0.13	0.12	0.12	13.2
Entropy of destination IP address	5.12	0.07	0.08	0.07	12.8
Entropy of destination port number	0.43	0.14	0.11	0.12	12.9
Entropy of packet type	0.55	0.05	0.06	0.04	0.03
Number of packets	42.3	1.21	1.32	1.21	6235
Occurrence rate of ICMP	0.87	0.01	0.01	0.01	0.01
Occurrence rate of UDP	0.01	0.98	0.01	0.0	0.01
Occurrence rate of TCP-SYN	0.02	0.02	0.02	0.01	0.92

In Table 1, for the DARPA 2000 data set, the parameters were estimated at the first phase of the attack scenario. Table 2 presents the estimates of the parameters at each phase of the attack. From Table 2 it can be observed that there was no observable change in the network traffic for phases 4 and 5. This is understandable because those are the phases in which the attacker intrudes the hosts of agents and installs the attack software.

To formulate the HMM, the transition matrix, the initial probability matrix, and the emission matrix have to be determined. However, a detection model must not have too many parameters else the computation will become intractable and the operation time will increase. For this reason, the transition matrix is formulated to be a 2 x 2 matrix (it is assumed that the system can only be in one of two states: normal (N) and compromised (R) which correspond to the hidden states). It should be noted that for this work phase 1 of the DDoS phases was taken as the compromised state. This is logical since the objective of the work is to predict attack. We want to use the precursors to the DDoS attack to predict the possibility of an attack.

The observable states in this work have been reduced to three and the states are drawn from the nine parameters listed above. To reduce the space of observable variables from nine to three, Kullback-Leibler divergence (KLD) (Aczél, J. and Daróezy, Z. (1975)) algorithm is applied.

4.1 Definition of Kullback-Liebler divergence

For discrete probability distributions P and Q, the KLD of Q from P is defined to be:

$$(P||Q) = \sum_i^{\mathrm{n}} P(i) \ln \frac{P(i)}{Q(i)} \qquad (2)$$

It can be described as the expectation of the logarithmic difference between the probabilities P and Q, where the expectation is taken using the probabilities P. The KLD is only defined if $Q(i) = 0 \Rightarrow P(i) = 0$, for all i (absolute continuity). If the quantity 0ln0 appears in the formula, it is interpreted as zero since $\lim_{x\to 0} x\ln(x) = 0$.

From the results got from the implementation of the KLD, the model was formulated using Entropy of Source IP address (denoted by *S*), Entropy of Destination IP address (denoted by *D*), and occurrence rate of packet type (denoted by *C*) as the observable states.

The entropies of each of the selected variables were then computed using equation (1)

The dataset is very large, so for the model training, it was partitioned into several subsets with each subset having hundreds of records with different sizes and below are the results obtained:

5.0 RESULTS

The model comprising of a 2 x 2 transition matrix, a 2 x 3 output matrix and initial probability matrix was formulated using the aforementioned. To estimate the parameters for normal network traffic without attack, the DARPA 1999 week 1 attack free dataset was used. The HMM model was formulated and trained using the DARPA 2000 data set. The dataset was divided into two disjoint sets: the training set and the test set. The learning algorithm was applied to the training set, generating the HMM model after which the percentage data in the test set that were correctly classified by the model was measured. These steps were repeated for different sizes of training set and different randomly selected training sets of various sizes.

The results obtained are as below:

The HMM [$\lambda = (P, \Phi, \pi)$]

The initial probability matrix (π) = $0.72 \quad 0.28$

The Transition Matrix (P) = $\begin{matrix} 0.82 & 0.18 \\ 0.98 & 0.02 \end{matrix}$

The Observation Matrix (Φ) = $\begin{matrix} 0.45 & 0.37 & 0.18 \\ 0.72 & 0.08 & 0.20 \end{matrix}$

The model is graphically represented as below:

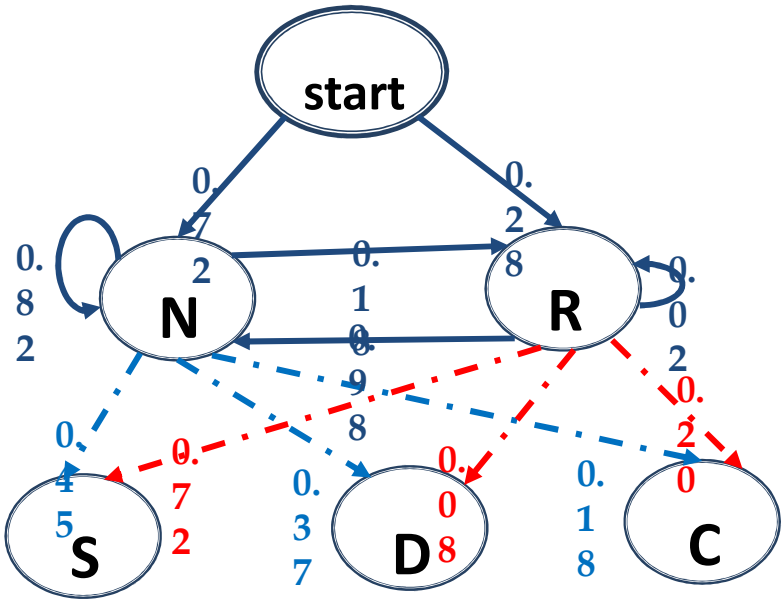

Table 2: Model Training Results:

No. of Iterations	P		Φ		
1.	0.8422	0.1578	0.4206	0.4241	0.1553
	0.9927	0.0073	0.6788	0.1095	0.2117
2.	0.8066	0.1934	0.4489	0.3810	0.1701
	0.9758	0.0242	0.7500	0.0645	0.1855
3.	0.8190	0.1810	0.4757	0.3529	0.1714
	0.9713	0.0287	0.7452	0.0892	0.1656
4.	0.8141	0.1859	0.4501	0.3588	0.1911
	0.9874	0.0126	0.7197	0.0884	0.1919
5.	0.8341	0.1659	0.4394	0.3764	0.1843
	0.9571	0.0429	0.7387	0.0878	0.1734

6.0 CONCLUSION

In this paper, we present a forecasting model based on the precursors to DDoS attack. The concept of entropy was used in determining the observable states of the HMM. In order to avoid computational intractability, KLD was used to reduce the number of parameters after which the HMM was formulated as a state-space representation. Based on the defined states, the initial probability matrix, the transition matrix and the emission matrixwas built. After which the model was employed in real-time determination of the level of abnormality of the incoming data.

The performance of the proposed method was empirically measured using the DARPA 2000 data set. The research has shown that features of the DDoS attack can be used as observable states of an HMM to predict attack in a network.

7.0 FUTURE WORK

More experiments would be performed and the system performance would be compared with the existing work. In order to do this, the HMM's hidden states would be increased so also the observable states.

References:

Aczél, J. and Daróezy, Z. (1975), On Measures of Information and Their Characterizations, New York-San Francisco-London. Academic Press. XII, 234 S., (Mathematics in Science and Engineering 115)

An, X. and Jutla, D. (2006). Privacy Intrusion Detection Using Dynamic Bayesian Networks. ICEC '06 Proceedings of the 8th international conference on Electronic commerce

Arnes, A., Valeur, F., Vigna, G. and Kemmerer, R. A. (2006). Using Hidden Markov Models to Evaluate the Risks of Intrusions: System Architecture and Model Validation

Bengio, Y. and Frasconi, P. (1995). Diffusion of context and credit information in Markovian models. *J. of AI Research*, 3:249–270.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer. ISBN 0-387-031073-8.

Cappe, O., Moulines, E. and Ryde´n, T. (2005), Inference in Hidden Markov Models. Springer Series in Statistics

Criscuolo, P. J. (2000). Distributed denial of service Trin00, Tribe Flood Network, Tribe Flood Network 2000, and Stacheldraht CIAC-2319. Department of Energy Computer Incident Advisory (CIAC), UCRLID-136939, Rev. 1, Lawrence Livermore National Laboratory.

Debar, H., Dacier, M., &Wespi, A. (1999). Towards a taxonomy of intrusion-detection systems. Computer Networks, 31(8), 805–822.

Durbin, J. and Koopman, S. J. (2000). Time series analysis of non-Gaussian observations based on state space models from both classical and Bayesian perspectives (with discussion). *J. Royal Stat. Soc. B.*

Durbin, J. and Koopman, S. J. (2001). *Time Series Analysis by State Space Methods*. Oxford University Press.

Gao, Bo, Hui-Ye Ma, and Yu-Hang Yang. (2002) "Hmms (hidden Markov models) based on anomaly intrusion detection method." Machine Learning and Cybernetics, 2002. *Proceedings. 2002 International Conference on*. Vol. 1. IEEE,.

Faltin, F. and Kenett R. (2007), Encyclopedia of Statistics in Quality & Reliability, Wiley & Sons

Flores, J. J., Antolino, A. and Garcia, J. M. (2010). Evolving Hidden Markov Models For Network Anomaly Detection. 10.1109/ICNS.2010.44 Conference: Networking and Services (ICNS), 2010 Sixth International Conference on

Ghahramani, Z. (2001) An Introduction to Hidden Markov Models and Bayesian Networks. International Journal of Pattern Recognition and Artificial Intelligence

Govindu, S. K. (2005). Intrusion forecasting system. Available on http://www.securitydocs.com/library/3110

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models, and the Kalman Filter*. Cambridge University Press,

Haslum, K., Abraham, A,and Knapskog, S (2008) "Fuzzy Online Risk Assessment for Distributed Intrusion Prediction and Prevention Systems", IEEE ICCMS.

Houle, K. J., & Weaver, G. M. (2001). Trends in denial of service attack technology. CERT and CERT Coordination Center, Carnegie Mellon University.

Ibe, O. C. (2009) Markov Processes for Stochastic Modelling, Elsevier Academic Press.

James, G., Witten, D., Hastie, T., Tibshirani, R. and (2013). *An Introduction to Statistical Learning*. Springer.

Jemili, F., Zaghdoud, M. and Ahmed, M. B. (2009) Hybrid Intrusion Detection and Prediction multiAgent System, HIDPAS, *(IJCSIS) International Journal of Computer Science and Information Security,Vol. 5, No.1,*

Jordan, M. I. (1999). *Learning in Graphical Models*, MIT Press, Cambridge.

Julisch, K. (2002) Data mining for intrusion detection: A critical review. IBM Research, Zurich ResearchLaboratory.

Kim, G., Lee, S. and Kim, S. (2014). A novel hybrid intrusion detection method integrating anomalydetection with misuse detection. Expert Systems with Applications. Vol. 41

Kim, S., Shin, S., Kim, H.,Kwon, K. and Han, Y. (2010), Hybrid Intrusion Forecasting Framework for Early Warning System, IEICE TRANS. INF. & SYST., VOL.E91–D, NO.5

Kumar, P. and Ravi, V. (2007). Bankruptcy Prediction in Banks and Firms via Statistical and Intelligent Techniques, European Journal of Operational Research, 180 (1)

Korb, K. B. and Nicholson A. E. (2004) *Bayesian Artificial Intelligence*, Chapman & Hall/CRC

Lee, K., Kim, J., Kwon, K. H., Han, Y., Kim, S (2008). DDoS attack detection method using cluster analysis. Expert Systems with Applications 34

Leu, F. and Li, Z. (2009), Detecting DoS and DDoS Attacks by Using an Intrusion Detection and Remote Prevention System.Information Assurance and Security, 2009. IAS '09. Fifth International Conference on (Volume:2)

Li, M. (2011) *Application of Data Mining Techniques in Intrusion Detection*, An Yang. Institute of Technology

Lin, S. C. and Tseng S. S. (2004). Constructing detection knowledge for DDoS intrusion tolerance. Expert Systems with Applications, 27

MIT Lincoln Lab (2000). DARPA intrusion detection scenario specific datasets. <http://www.ll.mit.edu/IST/ideval/data/2000/2000_data_index.html>.

Murphy, K. P. (2002). "Dynamic Bayesian Networks: Representation, Inference and Learning," Ph.D. Thesis, University of California, Berkeley

Pavlovic, V. (1999) "Dynamic Bayesian Networks for Information Fusion with Application to Human-Computer Interfaces," Ph.D. Thesis, University of Illinoisat Urbana-Champaign

Pearl, J. (2000). *Causality: Models, Reasoning and Inference*. Cambridge Univ. Press.

Pontes, E. (2012), "Distributed Multiagent Intrusion Forecasting System (DMIFS) Based Prediction Models for Cyber Attacks Detection System," In *Proceedings of the 7th European Symposium on Research in Computer Security*, Zurich, Switzerland, pp. 264-280

Pontes, E. and GuelfiA. E. (2009), IFS - Intrusion Forecasting System Based on Collaborative Architecture, Digital Information Management, ICDIM 2009. Fourth International Conference on

Reddy E. K. (2013) Neural Networks for Intrusion Detection and Its Applications. Proceedings of the World Congress on Engineering Vol II

Russell, S. and Norvig, P. (2003), *Artificial Intelligence: A Modern Approach,* Prentice-Hall, Inc.

Shin, S., Lee, S., Kim, H. and Kim, S. (2013). Advanced probabilistic approach for network intrusion forecasting and detection. Expert Systems with Applications. Vol. 40

Stich, T. (2004). Bayesian networks and structure learning, Diploma Thesis, Computer Science and Engineering, University of Mannheim, available at: http://66.102.1.104/scholar?hl=en&lr=&q=cache:j36KPn-8hWroJ:www.timostich.de/resources/thesis.pdf.

Tran, T. P., Cao, L., and Tran, D. (2009) Novel Intrusion Detection using Probabilistic Neural Network and Adaptive Boosting. International Journal of Computer Science and Information Security, Vol. 6, No. 1

Ye,N., Li, X., Chen, Q., Emran, S. M. and Xu, M. (2001) Probabilistic Techniques for Intrusion DetectionBased on Computer Audit Data, IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems And Humans, Vol. 31, No. 4

Vinchurkar, D. P., and Reshamwala, A. (2012). A Review of Intrusion Detection System Using Neural Network and Machine Learning Technique. International Journal of Engineering Science and Innovative Technology (IJESIT) Volume 1, Issue 2