INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X

(Volume2, Issue6)

Available online at: www.Ijariit.com

DETECTION AND MITIGATION OF BOTNET THROUGH MACHINE LEARNING IN MANETS

Mariya Ameer*
Research scholar dept of ECE,PEC,Mouli Kurukshetra University.
Mariyaameer1431@gmail.com

Aashish Gagneja
HOD dept of ECE, PEC, Mouli, Kurukshetra University
Gagneja2000@gmail.com

Navjot Kaur
Assistant Professor dept of ECE, PEC, Mouli, Kurukshetra University
Rosett_jot@yahoo.com

Abstract - Botnets are the networks of remotely controlled computer systems infected with a malicious program that allow cybercrimes to control the infected computers or machines without the user's knowledge. Botnets are the most ever-growing much interested evolved in the design of mobile adhoc networks (MANET).A botnet in mobile network is defined as a collection of nodes containing a malware called mobile malware which are able to bring the different elements into harmonious activities. Unlike Internet botnets, mobile botnets do not need to propagate using centralized structure. With the advent of internet and ecommerce application data security is the most critical issue in transferring the information throughout the internet. Botnets are emerging as the most significant threat facing computing assets and online ecosystems. The sharing of information through internet has been the main driver behind the Elite hacker into criminal activities. Their main target is to steal the vulnerable information from the individuals or from the organizations .In other words we can say their purpose include the distribution of spam emails, coordination of distributed denial of service (DDoS) and automatic identity theft. The proposed method is a classified model in which a Hill Cipher Algorithm and a Support Vector Machine are combined. A MANET environment with real time datasets is simulated for testing this model; the packet data of network flow was also collected. The proposed method was used to identify the critical features that determine the pattern of botnet. The experimental results indicated that the method can be used for identifying the essential botnet features and that the performance of the proposed method was superior to that of Artificial Fish Swarm Algorithm.

Keywords— : Botnet, Bot, Botmaster, Hill Cipher, MANET, SVM, Mitigation and detection.

I. INTRODUCTION

A Mobile Ad-hoc Network (MANET) can be characterized as gathering of portable hubs. It doesn't depend on any settled foundation. Since it is a framework less system, the versatile hubs in the system powerfully setup ways among themselves to transmit parcels from the source to destination and it is a self-arranging system. MANET can be utilized as a part of various applications, for example, combat zone correspondence, crisis help situation, hazardous situations and open air exercises, and so forth. The way of MANET is a progressively evolving process, because of its powerfully changing procedure; it is defenceless for extensive variety of assault. A MANET is a multi-bounce remote system that is shaped powerfully from versatile hubs without the help of a unified facilitator.

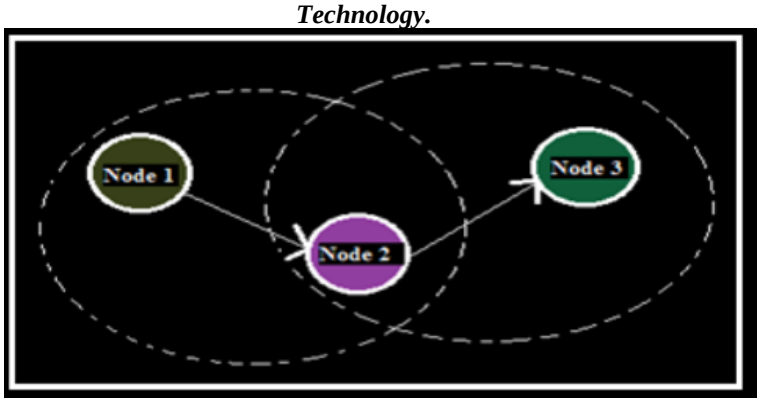

Figure 1.1

With late execution headway in remote innovation, convenient processing gadgets have become essential gadgets of our everyday life. The utilization of a compact gadget is compelled by its vitality and making power preservation. A Mobile Ad hoc Network is a gathering of autonomous portable hubs that can convey to each other by means of radio waves. The versatile hubs that are in radio scope of each other can straightforwardly impart, though others require the guide of middle of the road hubs to course their bundles. Each of the hubs has a remote interface to speak with each other. These systems are completely circulated, and can work at wherever without the assistance of any altered base as access focuses or base stations.

In view of the increasing demand for the wireless information and data services, providing faster and reliable mode access is becoming an important concern. A Mobile ad hoc network (MANET) is a network called spontaneous network that can be established without any infrastructure or any kind of topology. It means that all nodes behave as a router and participate in its discovery and maintenance of the routes. Its routing protocol has to be able to manage new difficulties that an ad hoc network creates such as node mobility, limited power, quality of service, and bandwidth issues. The challenges defined above create set new requirements on MANET network with routing protocols and make them more exposed to attacks [2]. In MANETS, all the participating nodes are involved in the routing process. Since conventional routing protocols are designed for predefined infrastructure networks, which cannot be used in mobile ad hoc networks so new routing protocols were designed to accomplish the requirement of less infrastructure ad hoc network. Figure 1.1 shows a straightforward ad-hoc network with three nodes. Node one and node three aren't at intervals vary of every other; but the node a pair of may be accustomed forward packets between node one and node a pair of. The node 2 will act as a router associate degreed these 3 nodes along type an ad-hoc network.

We see network security applications often require analyzing volumes of data to identify abnormal patterns or we can say activities .Today a novel system called Bot Graph [7] can be used to detect a new type of botnet spamming attacks targeting major web email providers. Botnets are the prevailing mechanism for the felicitation of the distributed denial of service (DDoS) attacks on the computer networks [8].

Because botnet viruses are ever-changing, in each pattern and attack ways, protection against these viruses became very tough. Most botnet investigation studies [1] have applied basic net virus detection ways like Honey internet and anomaly-based, signature primarily based, or machine-learning techniques. The anomaly primarily based and signature-based ways are ordinarily used. Within the anomaly-based methodology, once the detection system observes that the traffic within the user network exhibits uncommon actions, it determines that the user may be the victim of a botnet virus. The advantage of exploitation the anomaly primarily based [10] methodology is that unknown botnets may be detected; the disadvantage is that the speed of misjudgment may be high.in the signature-based method, associate degree uncommon packet info is often engineered, and once the system detects that the web packets of a user adapt to the info, the user may be infected by a botnet virus. The advantage of exploitation this methodology could be a high detection rate; but, the information should be updated. As a result of this method possess disadvantages, so they weren't employed in this research instead the Machine-learning methodology was adopted for police investigation botnet viruses. Methodology which will be used to discover unknown botnet viruses and features a high detection rate was developed by exploitation feature choice [4]. Feature choice is employed for distinction of crucial options of a quantity of multidimensional information and later on exploiting these options for analysis. Recently, SVM has become progressively common as a it can do high classification with little coaching sets . The most purpose of the SVM is to ascertain associate degree best hyper plane to classify information and build a classification model. The results indicated that the hill cipher exhibited wonderful performance in operate optimization, and therefore the potential of applying the Hill Cipher in optimization issues were also disclosed. During this study, a classified model was projected combining associate degree Hill Cipher formula and SVM. The projected methodology determined the crucial options defining the pattern of a botnet and thus botnet detection and mitigation can be done easily.

Here Section 2 introduces the SVM, AFSA, and Hill Cipher. Section 3 introduces the botnet detection methodology exploiting the SVM and also the Hill Cipher. Section 4 presents the experimental results and Section 5 provides a conclusion and suggestions for future work.

II.BACKGROUND

2.1 Support Vector Machine

The SVM was proposed by Cortes and Vapnik . It is a managed learning model taking into account basic danger minimization and the Vapnik–Chervonenkis measurement. A SVM is normally connected in machine learning [3] and for taking care of arrangement or relapse issue. In this manner, the primary reason for a SVM is distinguishing the ideal hyper plane to break down different order information. The ideal hyper plane has the maximal edge connected with the different order information, as appeared in Figure 2.1. Two dark focuses and three white focuses exist on the maximal edge line, which speak to two sorts of grouping information; these focuses are called vectors. These vectors can be utilized for grouping new information. At the point when the information is not directly distinguishable, the bit capacity must be utilized to outline information into the Vapnik-Chervonenkis dimensional space. Three types of kernel function (φ) exist: radial basis functions (RBFs), polynomials, and sigmoid. Using the appropriate kernel function for transforming the data is imperative for increasing the classification speed.

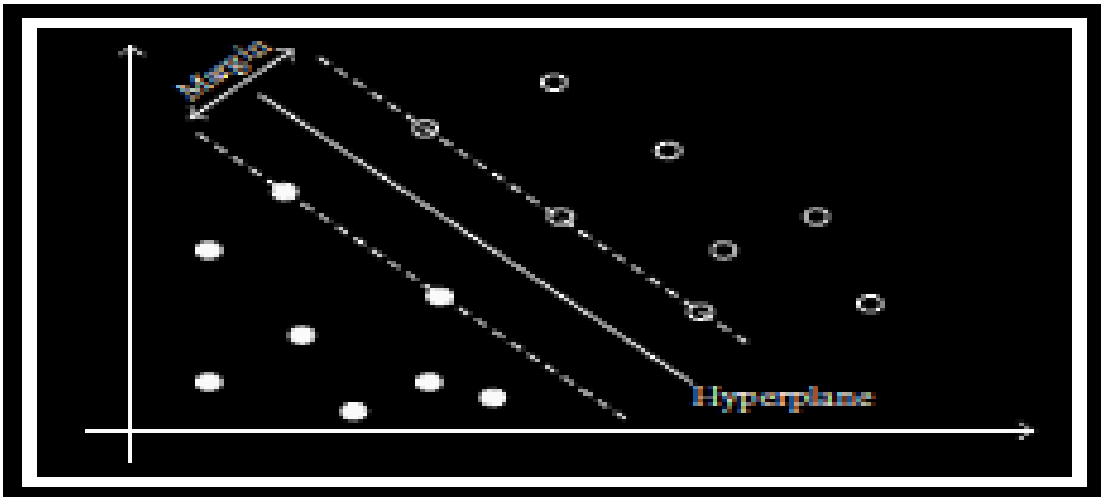

Figure - 2.1: The optimal hyper plane

2.2 Hill Cipher Algorithm

Hill Cipher Encryption is a technique used to scramble messages utilizing the network as a key. In this key, there are nine pieces utilized arbitrary whole numbers that set a grid of 3x3. Every number will take up with each other to create the figure content, yet we can't for all time utilize these numbers to reestablish the first messages. The numbers must have the definite estimation of the determinant. Prior to the numbers could be utilized, we ought to test these numbers whether we meet the right determinant. In the event , if the outcome isn't right, the whole numbers must be done once more. Producing keys on Hill Cipher calculation by consolidating Genetic calculation is relied upon to accelerate the quest of the reasonable key for the Hill Cipher encryption. Hill Cipher is the module number method in cryptography. Hill Cipher utilizes the symmetric key as the secret key to change over plaintext to figure content. The Symmetric key is one of the cryptography frameworks that utilization the same sort of keys in encryption and unscrambling. The key used to encode is really unique in relation to decode the messages, however they utilize the same recipe as the vessel of data trade either on encryption or unscrambling part. The general hypothesis of grid utilized as a part of Hill Cipher is the increase between the network and the converse of the framework. Without getting the right key, the procedure of encryption and unscrambling can't be performed.

2.3 Artificial Fish Swarm Algorithm

The AFSA is employed to see the optimum or most satisfactory resolution in an exceedingly restricted time by regularly looking for attainable solutions employing a met heuristic within the AFSA[3], the position of each fish is taken into account an answer, and each resolution contains a fitness price that's evaluated.

III. PROPOSED METHOD

Both the AFSA and Hill Cipher are met heuristic algorithms; however, they employ distinct optimization mechanisms. The AFSA has demonstrated success in numerous applications [3], but we found that Hill Cipher yields superior optimization performance. In this study, the SVM was employed as the classifier, using the Hill Cipher and AFSA to perform feature selection. Classifiers can establish a classified model and use it to assign data to the correct categories. Method involves these steps, which comprised the cross- key that encrypts in really totally different way to decipher the messages; however they use a similar formula. We must inverse the key before it is used to decipher the validation process. For example, the first portion of the data was used as the test data and the remaining data were used as training data; whereas in the next round, the second portion of the data was used as the test data and the remaining data were used as training data. The steps involved in detecting and mitigating botnet in MANET include:

Step 1: Generate MANET Scenario using MATLAB
Step 2: Implementation of AODV protocol
Step 3: Generate Load on Each Node
Step 4: To implement of Hill Cipher and SVM algorithm.
Step 5: To establish a system by hybrid algorithm
Step 6: To improve efficiency and execution time comparative to previous one

At the initial steps of the Hill Cipher, the algorithm assigns a random feature subset to every ciphered text, and the SVM is used to obtain the classification accuracy based on the fitness of every cipher.

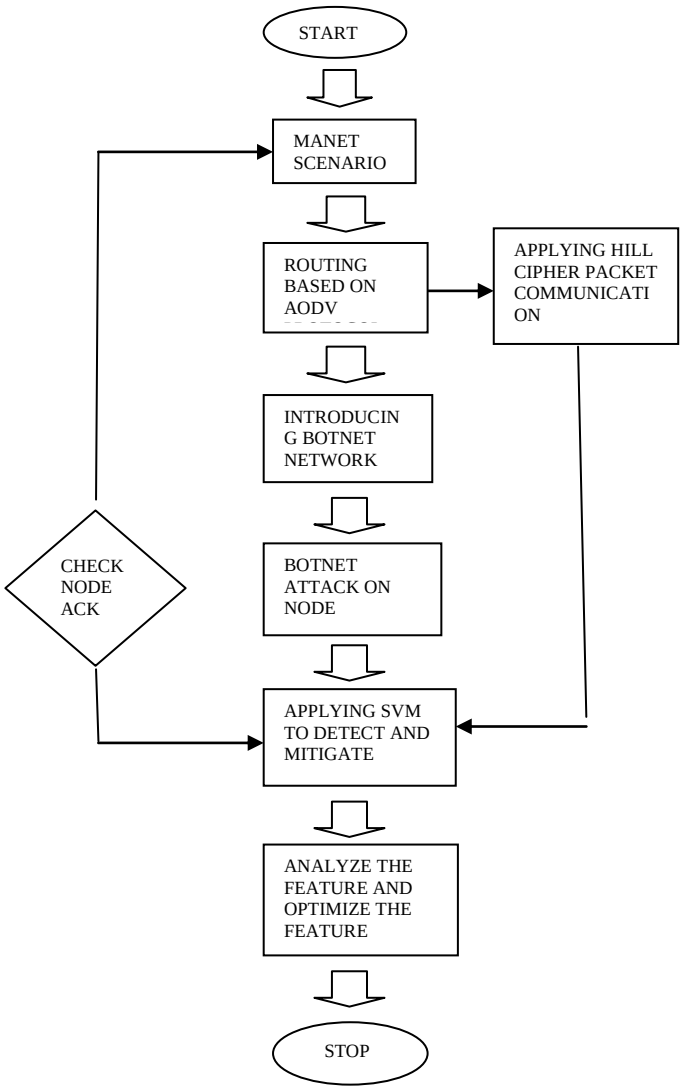

Figure 3.1 Proposed Flowcharts

IV. EXPERIMENTAL RESULTS

In this section, the proposed algorithm is evaluated via computer simulation using MATLAB simulator. All simulation results are obtained by using two algorithms Hill Cipher and Support Vector Machine. To estimate the performance of feature selection using the Hill Cipher combined with SVM, the performance of Hill Cipher was compared with that of AFSA, including the classification accuracy, the number of features of the optimal solution and the time spent applying each algorithm to perform calculation. The parameters used in this study are presented as follows. The number of nodes were 100, the number of cluster head were 5, max x-lim for network land was 100,max y-lim for network land was 100,max z-lim for network land was 50, number of receivers were 6,time taken was 15 sec, gain of transmitting antenna was 12 and gain of receiving antenna was 13.

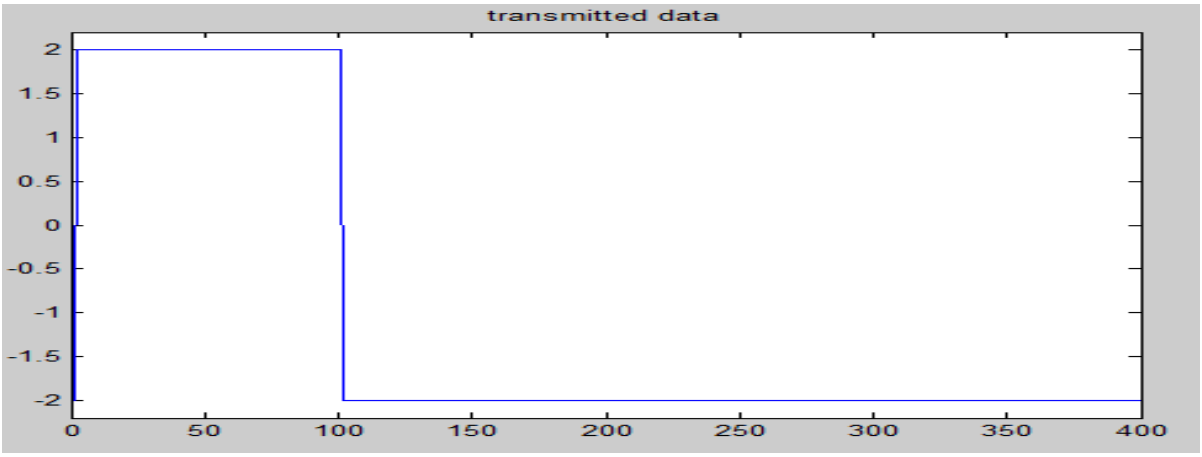

Figure 4.1: Data Transmit Between Sources to Destination

Fig 4.1 represents the data transmission from source node to destination node for 100 nodes, the number of cluster head were 5, max x-lim for network land was 100, max y-lim for network land was 100,max z-lim for network land was 50, number of receivers were 6,time taken was 15 sec, gain of transmitting antenna was 12 and gain of receiving antenna was 13.

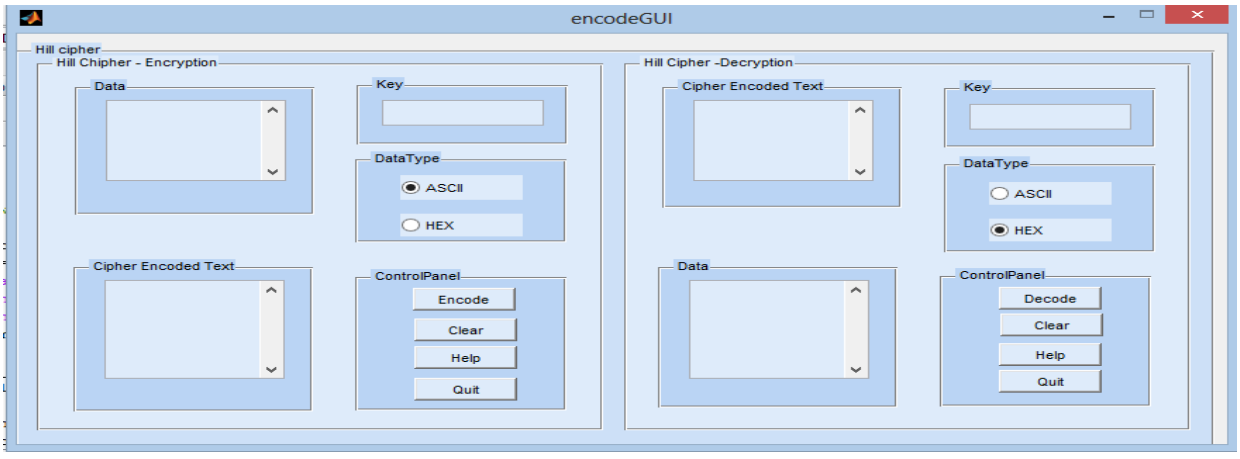

Figure 4.2: Front End Simulation using Hill Cipher.

Fig 4.2 shows the front end simulation of Hill Cipher Algorithm for 100 nodes, the number of cluster head were 5, max x-lim for network land was 100,max y-lim for network land was 100,max z-lim for network land was 50, number of receivers were 6,time taken was 15 sec, gain of transmitting antenna was 12 and gain of receiving antenna was 13.

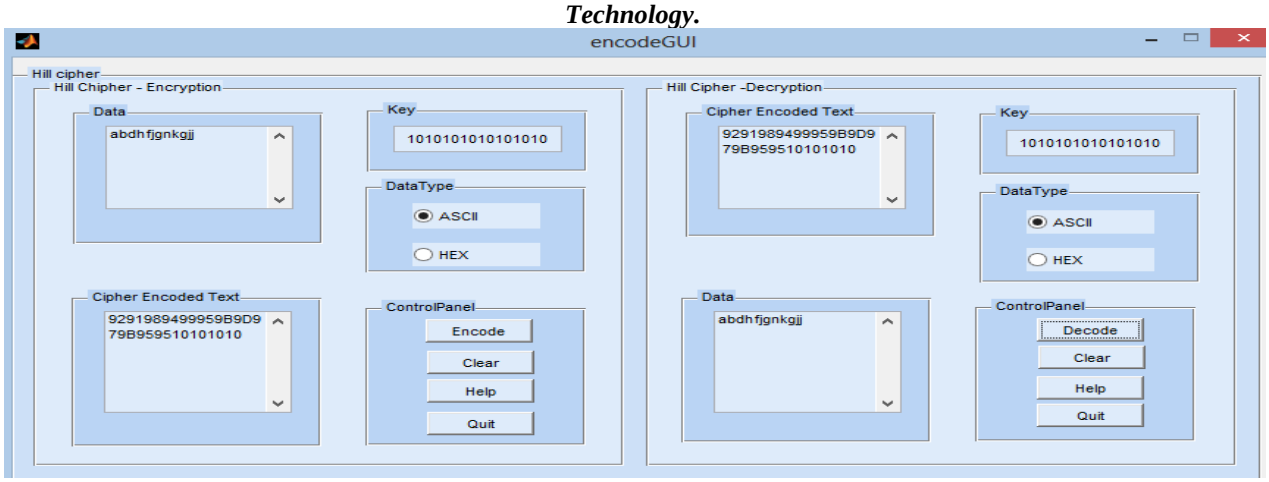

Figure 4.3 Encrypt and decrypt data using hill cipher

Fig 4.3 shows the encryption and decryption of data by using Hill Cipher Algorithm for 100 nodes, the number of cluster head were 5, max x-lim for network land was 100,max y-lim for network land was 100,max z-lim for network land was 50, number of receivers were 6,time taken was 15 sec, gain of transmitting antenna was 12 and gain of receiving antenna was 13.

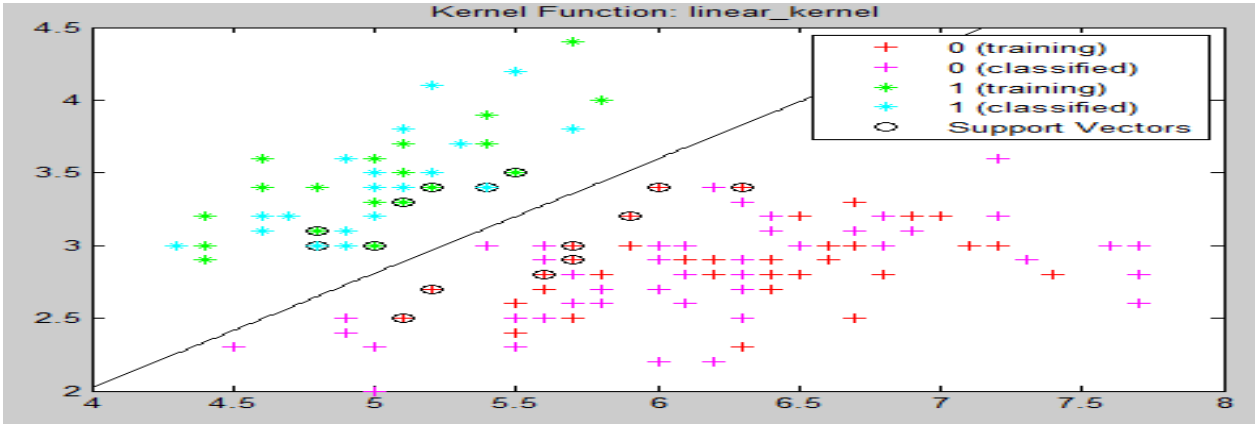

Figure 4.4 Total optimized feature

Fig 4.4 shows the number of optimized features by making use of SVM for 100 nodes, the number of cluster head were 5, max x-lim for network land was 100,max y-lim for network land was 100,max z-lim for network land was 50, number of receivers were 6,time taken was 15 sec, gain of transmitting antenna was 12 and gain of receiving antenna was 13

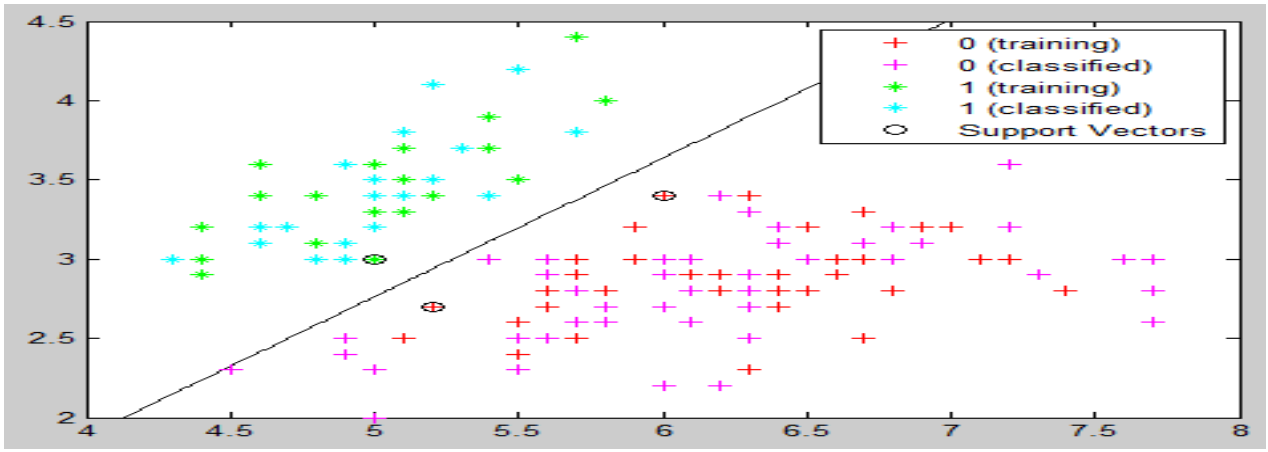

Figure 4.5 Selected Feature out of Optimized Features

Fig 4.5 shows total number of selected features by making use of SVM for 100 nodes, the number of cluster head were 5, max x-lim for network land was 100,max y-lim for network land was 100,max z-lim for network land was 50, number of receivers were 6,time taken was 15 sec, gain of transmitting antenna was 12 and gain of receiving antenna was 13.

Real time dataset were collected table 4.6 shows the experimental results for dataset classified using Hill Cipher and SVM. The average classification accuracy, number of selected features of the optimal solution subset, and the total time between the ASFA and Hill Cipher were compared. The numbers of selected features of the Hill Cipher were less than the ASFA thus reducing the detection time

Table 4.6

Datasets	**AFSA-SVM**				**Hill-Cipher SVM**		
	Number of Selected Features	**Average Accuracy rate (%)**	**Execution Time (sec)**		**Number of Selected Features**	**Average Accuracy rate (%)**	**Execution Time (sec)**
Botnet	6	97.76	198.43		3	100	44.27

In this study, a feature selection method for detecting botnet viruses is proposed, which is the Hill Cipher- SVM method. The experimental results showed that the Hill-Cipher yielded higher classification accuracies than exploitation the AFSA-SVM; however less time was spent to get a lesser range of feature subsets. In sensible applications, classification accuracy is usually the primary priority, however insure processes, like botnet virus detection [5], detection speed is as crucial as accuracy. To get the required detection speed, the info needed for process should be reduced underneath the premise that the accuracy level is that the same; therefore, during this situation, the Hill Cipher-SVM technique is superior. The result additionally shows that both Hill Cipher and AFSA will still be applied for distinctive the crucial options of botnet, filtering surplus options, and exploitation these algorithms in varied applications simply.

V.CONCLUSION AND FUTURE WORK

For any extent of this work, rather than planned rule a brand new AI may be utilized for higher improvement. Therefore, in future studies, the planned technique should be tested for detective work P2P protocols or alternative sorts of botnet viruses. Finally, a feature-selection based mostly detection system for detective work botnet viruses will hopefully be made within the future.

REFERENCES

*[*1] Maryam Feily,Alireza Shahrestani and Sureswaran Ramadas"A Survey of Botnet and Botnet Detection ,National Advanced IPv6 Center of Excellence (NAv6)IMPACT Research Team University Saints Malaysia (USM)Penang, Malaysia, Faculty of Computer Science and Information Technology University of Malaya (UM), NAv6 2009 Third International Conference on Emerging Security Information, Systems and Technologies.

[2]HemalathaU,Mohana.P,Manjubargavi.A,SanjanaViswanathan and I.Varalakshmi Murugan "Enhanced MANET Data Security againist BOTNET attacks using Hybrid IDS Model".International Journal of Emerging Technology in Computer science and Electronics(IJETCSE).ISSN:0976-1353 volume 12 Issue 2-January 2015.

[3]Kaun-cheng Lin,Sih-Yang Chen,and Jason C.Hung"Botnet Detection using Support Vector Machines with Artificial Fish Swarm Algorithm".Hindawi Publishing Corporation Journal of Applied Mathematics Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 4, Issue 1, January 2014).

[4] Sagar A. Yeshwantrao1, Prof. Vilas J. Jadhav Lecturer, 2Asst. Prof., Department of Computer Engineering,MGM's College of Engineering & Technology Kamothe, Navi Mumbai, State- Maharashtra, Country- India. Threats of Botnet to Internet Security and Respective Defense Strategies. International Journal of Emerging Kyushu University.

[5] Kamaldeep Singh, Sharath Chandra GuntukuAbhishek Thakur, and Chittaranjan Hota Big Data Analytics framework for "Peer-to-Peer Botnet detection using Random Forests". A Department of Computer Science, BITS – Pilani, Hyderabad Campus, AP 500078,

India. Information science 2014.School of Computer Engineering, Nanyang Technological University, 50 Nanyang Drive, Singapore 639798, Singapore India,. Information Science 2014.

[6] Pijush Barthakur, Manoj Dalal, Mrinal and Kanti Ghose Department of Computer Science and Engineering Sikkim Manipal Institute of Technology Sikkim, India. "A Framework for P2P Botnet Detection Using SVM (2012)". International Conference on Cyber-Enabled Distributed Computing and Knowledge Discover.

[7] Yao Zhaoy, Yinglian Xie, Fang Yu, Qifa Ke, Yuan Yu, Yan Cheny, and Eliot Gillumz BotGraph: Large Scale Spamming Botnet Detection. Northwestern University Microsoft Research Silicon Valley Microsoft Corporation.

[8]Esraa Alomari, Selvakumar Manickam, B. Gupta,Shankar Karuppayah and Rafeeq Alfaris 3University of New Brunswick, Canada 4RSCOE, University of Pune, India National Advanced IPv6 Centre (NAV6), University Saints Malaysia, Malaysia. "Botnet-based Distributed Denial of Service (DDOS) Attacks on Web Servers".Classification and Art.) International Journal of Computer Applications (0975 – 8887) Volume 49– No.7, July 2012.

[9] Jingyu Hua and Kouichi Sakurai "A SMS-Based Mobile Botnet Using Flooding Algorithm" Department of Informatics, 2014, Article ID 986428, 9 pages.

[10] James R. Binkley and Suresh Singh," An Algorithm for Anomaly-based Botnet Detection "Computer Science Dept. Portland State University.

[11] J. R. Quinlan, C4.5: Programs forMachine Learning,TheMorgan Kaufmann Series in Machine Learning, Morgan Kaufmann Publishers, San Mateo, Calif, USA, 1993.

[12] T. S. Furey, N. Cristianini, N. Duffy, D. W. Bednarski, M.Schummer, and D. Haussler, "Support vector machine classification and validation of cancer tissue samples using microarray expression data," Bioinformatics, vol. 16, no. 10, pp. 906–914,2000.

[13] M. Abdel Fattah, "The use of MSVM and HMM for sentence alignment," Journal of Information Processing Systems, vol. 8, no.2, 2012.

[14] G. P. Zhang, "Neural networks for classification: a survey," IEEE Transactions on Systems, Man and Cybernetics C: Applications and Reviews, vol. 30, no. 4, pp. 451–462, 2000.

[15] K. Sarkar, M. Nasipuri, and S. Ghose, "Machine learning based keyphrase extraction: comparing decision trees and artificial neural networks," Journal of Information Processing Systems, vol. 8, no. 4, pp. 693–712, 2012.

