International Journal of Cyber-Security and Digital Forensics (IJCSDF) 9(1): 26-33
The Society of Digital Information and Wireless Communications (SDIWC), 2020 ISSN: 2305-0011

Integrated Intrusion Detection Scheme using Agents

Kajal Rai*, Ajay Guleria, M. Syamala Devi

Kajal Rai*, Department of Computer Science and Applications, Panjab University, Chandigarh, India. Email: kajalrai.pu@gmail.com
Ajay Guleria, System Manager, Indian Institute of Science, New Delhi, India. Email: guleria.ajay@gmail.com
M. Syamala Devi, Department of Computer Science and Applications, Panjab University, Chandigarh, India. Email: syamala@pu.ac.in

ABSTRACT

Misuse based Intrusion Detection System (IDS) employs various signature matching techniques against stored databases to find intrusions in a network. Anomaly-based Intrusion Detection System does behavior modeling of network traffic and classifies it as normal behavior or attacks. A behavior which deviates from normal is the indication of an attack. The proposed integrated IDS combine the benefits of both approaches. An agent is a software program that is capable of doing independent actions on behalf of user. In the proposed system is a multi-agent based IDS three agents are used (Interface agent, training agent, and detector agent). In the proposed integrated approach, first the incoming packets are classified using a misuse approach based on a decision tree by agents, and then the packets are passed to the anomaly phase where these packets are again classified via agents using anomaly approach based where payload frequencies are categorized. The final decision is made by taking the output from both approaches. The primary benefit of an integrated strategy is that it can uncover both known and novel attacks in the network. The proposed integrated intrusion detection system is tested on the standard data set, i.e., DARPA and collected data set of Panjab University, Chandigarh. It is observed from the results that integrated approach achieve higher accuracy than the individual approaches.

KEYWORDS

Anomaly Detection, Integrated Scheme, Intrusion Detection, Misuse Detection, Multi-Agent System

1. INTRODUCTION

Misuse based IDS uses various procedures that correspond in pattern to determine similarity among system events and well-known attack signatures warehoused in the signature database. Anomaly detectors generate normal profiles of operators, computer machines, and all the network links by utilizing their recognized behaviors. After the profiles are generated, the system observes new occurrence of data, match the new data with previously warehoused profiles, and attempts to find deviations from usual behavior. The data with significant deviations are anomalies. Benefits of misuse based IDS are that they are a perfect and effective technique for the uncovering of known attacks and have high, revealing speeds as it devotes a smaller amount of time for fixing false positives.

Benefits of anomaly-based IDS are they are efficient to sense novel and unforeseen attacks. Limitations of misuse based IDS are that they are ineffective to identify unfamiliar attacks and variation of well-known attacks. Limitation to anomaly-based IDS is that it can make a misclassification in the uncovering of attacks because of intrusive training data, and it is quite challenging to generate warnings in real-time [1]. Hence, there is a need for integrated IDS in which both the approaches are incorporated and produce more efficient results in terms of accuracy, less false alarms, and less delay.

Hybrid IDS takes advantages of both signature-based detection and anomaly detection approaches and is proficient of detecting both known and novel attacks with low false positive rates. Hybrid IDS can be constructed using either sequential approach or parallel approach. In sequential approach either misuse detection is followed by anomaly detection or vice-verse. In parallel approach both the approaches work simultaneously. Most of the researchers are using sequential approach for their work and the reason for using it is that most of the packets get classified in the first attempt and only remaining ones are passed to the second phase and lots of computing overhead can be saved.

An agent is an entity that observes its environment through sensors, performs actions in that environment towards achieving its goals, and is capable of acting independently with negligible support from other entities and interacts with other agents or humans. A multi-agent system (MAS) comprises of several collaborative agents with every single one performing a specific task such as data analysis, feature extraction, etc. Agent-based IDS comes in the category of distributed IDS. In this, there is no central station and henceforth no particular point of failure. Agents used in these systems gain knowledge from their experience, communicate and work together to accurately detect network intrusions. The proposed work use agents to detect attacks in the network and to prevent human intervention.

MAISNID consists of three agents: Interface agent (IA), Training Agent (TA) and Detector Agent (DA). IA captures network packets and extracts useful features from them and sends attack packets information to the administrator for further legal action. TA accomplishes training for misuse-based and anomaly-based intrusion detection using decision tree and n-gram approach respectively. DA is a mobile agent which collects packets from different machines in the network by migrating itself along with its state and code. Then DA predicts the class of coming packets by moving to the machine where the database resides and sends related information about abnormal packets to the administrator. Thus the integrated scheme combines both misuse-based and anomaly-based approach. Multi-Agent System is more beneficial for this purpose as it can gain knowledge from their experience, communicate, and work together to detect network intrusions accurately. The main benefit of the integrated scheme is that it can discover both known and novel attacks in the network. This system assists in the analysis of network packets and detects affected hosts in the network.

The proposed integrated intrusion detection system is tested on standard data-set, i.e., DARPA and collected data-set of Panjab University. On DARPA data-set, the testing is done on three ports, namely, port 21 (FTP), port 23 (TELNET), and port 80 (HTTP). On university data-set, the testing is done on port 80 (HTTP) and port 443 (HTTPS).

2. RELATED WORK

Yu et al. [2] present the co-operative architecture for multiple IDSs which was intended to identify real-time attacks in the network. Co-operative agents and several detection sensors were used to discover attacks in the network. Various IDS products were combined to discover intrusions in a dynamic environment. Communication between agents is accomplished by exchanging messages and agents take decisions in a co-operative manner. Agents used in this system were IDMEF (Intrusion Detection Message Exchange Format) Agent, Merging Agent, Clustering Agent, Host Agents, and Coordinator Agents. The framework was designed for post-detection alert examination and adopting defensive safety actions.

DIDS (Distributed Intrusion Detection System) was implemented by employing co-operative intelligent agents distributed across the networks [3]. Fuzzy rule-based classifiers are used to identify attacks in a network. Distributed Soft Computing-based IDS (D-SCIDS) was developed as a mixture of several classifiers like a neuro-fuzzy and genetic-fuzzy classifier. By hybridizing fuzzy classifiers, robustness and flexibility are achieved. The main disadvantage of this system was the extreme usage of system resources and the correlation of alarms. Zhang et al. [4] designed agent-based distributed IDS, which was known as HIDSFCN. The authors used sequential hybrid IDS in which misuse detection was followed by anomaly detection. It uses a Radial Basis Function (RBF) of SVM to construct the classifier. In this system, there is no direct communication between agents; instead, they only intercommunicate with local sensors and local transreceivers. The proprietary component of this system was the Information Collect Center, which was used to gather data from the transceivers and store them in a built-in SRAM-the rule database.

MObile-VIsualization Hybrid IDS (MOVIH-IDS), was a hybrid artificial intelligent IDS proposed by Herrero et al. [5]. It incorporates an unsupervised connectionist IDS to provide an efficient network security architecture. The agents that were used in this system: Sniffer, Preprocessor, Analyser, ConfigurationManager, Coordinator, and Visualizer have been designed and implemented. Sniffer agent captures the network traffic, and the captured network packets were passed to other

agents for analysis. Preprocessor agent does the preprocessing of the network packets captured by sniffer agent. Preprocessing involves the feature selection from the packets. Analyzer agent then analyzed the processed packets. This agent incorporates two behaviors, namely, learning and exploitation. Configuration Manager agent manages the preprocessing and analysis of network traffic depending upon the values of parameters such as segment length, features to extract, etc. Coordinator agent coordinates the agents and balances the workload among them. Visualizer agent provides information about the system and analyzed data to other agents in the system in a graphical form.

An agent-based IDS was suggested by Ganapathy et al. [6] for MANETs. Techniques were used for feature selection, outlier detection, and improved multiclass SVM classification. Outlier detection and SVM were used as preprocessing techniques. Two new algorithms were suggested for identifying the intruders in a distributed database environment. Trust management and coordination in the processing of transactions were taken care of by Intelligent Agents. The system consists of six agents, namely, data collection agent, Preprocessing data agent, outlier detection agent, weight assignment agent, and classification agent, and selection agent. Network data was collected by the data collection agent. The data preprocessing agent used attribute selection algorithm and selects only the useful attributes from the dataset by means of projection. Outlier Detection Agent used weighted-distance-based outlier detection technique in which the agent used an outlier factor for finding the outlier point in the attribute set. Weight Assignment Agent allocates the most appropriate weight to all the features of the dataset. Classification Agent used an enhanced SVM algorithm where the agent uses the appropriate distance measurement formula to classify the data efficiently. Selection Agent chooses the distance measure to use for classification. From the experiments, it has been observed that the classification accuracy for DoS is 99.77%, Probe is 99.70%, and other attacks is 79.72%.

Okba et al. [7] proposed a mobile-agent based distributed IDS. Four types of agents were developed in this system by using the Aglets platform. Collector agent was used for gathering information; Analyzer agent and Redirector agent were used for analyzing the captured data, states, and behavioral analysis, respectively. Last, the generator agent launches these agents and manages the messages got from Analyzer and Redirector agents. Dasgupta et al. [8] developed an IDS which is an agent-based IDS that can simultaneously monitor several network activities at different levels. It can monitor the activities at system level, packet level, user level, etc. The identification of anomalous events and response to those events are managed not at the central system instead the work is distributed by agents. The system integrates artificial intelligence methods such as fuzzy logic and neural network for detection of network attacks. Elbasiony et al. proposed a sequential hybrid IDS in [9] in which misuse detection phase is followed by anomaly detection phase. Random Forest was used in misuse detection phase to generate signature patterns automatically from the training data-set. In anomaly detection, network connections records were clustered by weighted k-means clustering technique. But when high number of clusters is used for detecting numerous attacks the system generates high false positive rates because of low alert correlation between clusters.

It is observed from the review that there are many issues in intrusion detection systems including high resource consumption, large number of false alarms, a lot of time to train the model, etc. However hybrid systems can address a number of these issues. Hybrid IDS can be constructed either by running misuse and anomaly module simultaneously in parallel or it can be constructed by running one module after another in sequential way. In either case the main issue is the updating of attack database to find the most recent attack in the network. In our proposed integrated scheme, the IDS not only detect the attacks but also the signature database is also updated time to time to detect the latest attacks in the network.

3. AGENT-BASED INTEGRATED SCHEME

3.1 Design of Intrusion Detection System using Integrated Scheme

Design of proposed Multi-agent based Integrated Scheme for Network Intrusion Detection (MAISNID) consists of three modules, namely,

misuse detection module, anomaly detection module and integrated module. All these modules are implemented using three software agents. These three agents work in tandem with each other for deployment of the proposed integrated scheme.

3.2 Decision Tree-based Misuse Detection

Misuse-based IDS is also acknowledged as signature-based IDS. This approach is provided with a database that has a number of attack signatures. The test data gathered by IDS is matched against stored attack signatures using various pattern-matching techniques, and if a match exists, an alert is generated. The instances that do not equalize with any attack signatures are seen as a part of genuine activities. Two critical steps of misuse detection are extracting useful features and classifying based on those features.

For our research work, we used information gain ratio for feature selection. For selecting features to build a classifier model, the gain ratio is chosen because it gives equal importance to all the values in the attribute domain so that there is no bias towards multi-valued attributes, and it can work for both supervised and unsupervised classifiers.

The classification algorithm is applied to build a model from the labeled data set. Every data element is well-defined by the values of the features. We developed a classifier model for misuse detection using decision tree because of its capability to efficiently manage missing values in training data. It can also delete the branches that do not play an important role in data classification. Decision trees are developed by searching recursive splits on learning which attempts to find the best partition for predicting the output class. Each split is done according to the values of an attribute. At the first step, the root node is selected by finding the best split among all attributes. Then this process is repeated at every new node till certain ending criterion is reached [10]. The most important concern in creating a decision tree is the selection of a node that partitions the data-set into two or more classes. We have proposed an algorithm for constructing a decision tree which is based on C4.5 decision tree algorithm, which is explained in [11].

3.3 N-gram based Anomaly Detection

To create normal profiles for our research work, statistical modeling of experimented data streams using n-gram analysis for byte frequency distribution is performed. In context with network packets, an n-gram is the arrangement of n-adjacent bytes in a payload element. A sliding window with size n is passed over the entire payload and existence of each one n-gram is calculated. Further, the payload bytes are converted into ASCII characters, and the frequency vector is computed for each length [12]. Frequency vector is the count of each ASCII character in the payload of the packet in which the data is in the form of bytes. In order to model the payload, we fragment this stream of bytes into sets based on port number, IP addresses and payload length. Within one port, there is a deviation in the payload length, for example, for TCP packets on port 80, the payload length ranges from 0-1460.

We constructed the model on a specific port for each payload length from each connection. Before model construction, feature extraction is performed. Anomaly detection model is based on the frequency distribution of n-grams in the payload. As there are 256 characters in the ASCII character set, the payload can be represented by 256^n dimensional feature space. With higher n, a higher quantity of information can be taken out from payload. However, as n goes higher, the feature space dimension grows exponentially. Also, with a given payload P of length m, the frequency of n-grams can be figured out in O(n) time. We choose n=1 for model construction which reduces feature space to 256 dimensions only and also it computes the frequency model in O(1) time.

To reduce the number of models and save storage space, the length of payload for each port coming from specific source and to particular destination are clustered using lengthwise clustering where the length is divided into groups of 10 each from 1-1460. The details of how length-wise clustering is done are given in [13]. After doing length-wise clustering of payload data, payload bytes are converted into ASCII characters and frequency vector is computed for each length group. Frequency vector is the count of each ASCII character in the payload of the packet. Then the mean and standard deviation of all frequency

vectors for each payload length is computed. Mean and variance are calculated and are used as the model for the given length group. This procedure is repeated for all other IPs in the set. The algorithm for this is proposed in [13] which describe the steps for generating normal profiles.

3.4 Integrated Scheme

The integrated scheme includes two-layers. At the first layer, misuse detection model is constructed then anomaly detection is done at the second layer. Coming packets are analyzed in misuse module and if the packet-attributes are matched with stored-attack signatures, it is tagged as an attack packet; otherwise it is passed to the anomaly module. In anomaly module, payload of the packet is converted to ASCII frequency vector using the algorithm proposed in [13]. Then Cosine Similarity is used to find distance between packet frequency vector and normal profiles stored in the database. If distance is greater than the predefined threshold, then packet is labeled as anomalous. Equation 3.1 gives the mathematical formula for calculating Cosine Similarity. Cosine Similarity is used as it is highly efficient for high-dimensional spaces by taking less time in computation.

$$\text{Cosine Similarity} = \cos(\theta)$$

$$= \frac{X.Y}{||X||.||Y||} \qquad (3.1)$$

$$\text{Cosine Distance} = 1 - \text{Cosine Similarity}$$

where, X and Y are vectors.

3.5 Multi-Agent Design of Integrated Scheme

Figure 1 shows the overall design of the model, in which three agents are working in a multi-agent environment. The three agents used in the proposed scheme are Interface Agent (IA), Training Agent (TA) and Detector Agent (DA).

Communication between agents takes place to send and receive messages, analyze messages, and comprehend them. Administrator interacts with IA, and in turn, IA interacts with TA and DA. IA, on the request of the administrator, captures packets from the organization network, and extracts the useful features from packets for further processing. It also collects log files of the same time duration at which network packets are captured. Then the extracted features from packets and information about these features from log files are passed by IA to TA. TA then labels these packets by analyzing log files information and creating a database of attack signatures.

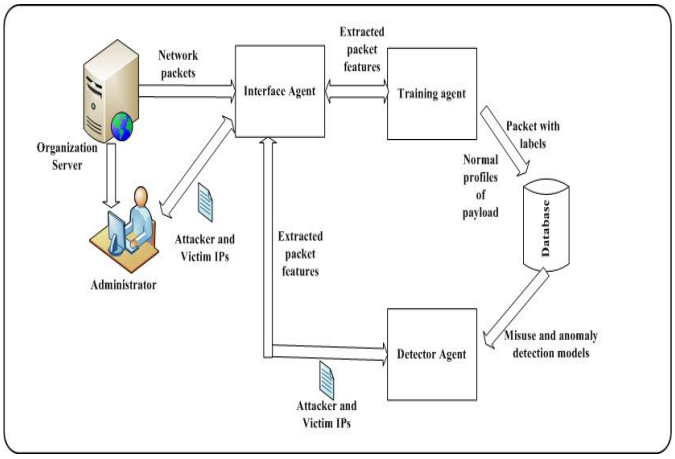

Figure 1: Multi-Agent based Intrusion Detection Model

TA next uses this database, which contains attack as well as normal connection packet information and builds the misuse detection model using decision tree algorithm. TA also builds normal profiles from the captured packets by computing the frequency vector of payload data using the n-gram approach. These normal profiles serve as anomaly detection models and are stored as the database for anomaly detection. After making detection models for both misuse-based module and anomaly-based module, TA informs IA that training has been done and the models are ready to detect attacks in the network. IA then asks the DA to test the newly coming network packets. IA does not only capture these packets, but features are also extracted by IA and passed to DA. DA first uses the misuse detection model to classify packets and passes the obtained results to IA. The packets, classified as normal by misuse detection model, are passed to the anomaly detection module and are classified by it. The information of detected attack packets as a result of the integrated module is kept by DA, and then this information is passed to IA and then to the administrator for further action.

After testing the information of all those packets which are tagged as ‘Anomalous' in any module is kept in a file, and this information is later passed

to the administrator for further action. Also, the Internet Protocols (IPs) which are tagged anomalous again and again within a certain period are considered as attack IPs and based on that the signature database is also updated. This updated database of attack signatures is then used to rebuild the decision tree model so that the newly coming attacks in the network are found at an early stage. Detector Agent by applying the integrated approach tests the data and displays the anomalous packets information such as IP, port number, etc.

Role of each agent is important for the functionality of the system. Interface Agent is designed to capture network packets for training misuse and anomaly detection model and provide useful information about observed attack packets to the administrator. Training Agent is designed to train the model for both modules. Detector Agent is used to find attacks in the network. Specific tasks performed by each agent are given in Figure 2.

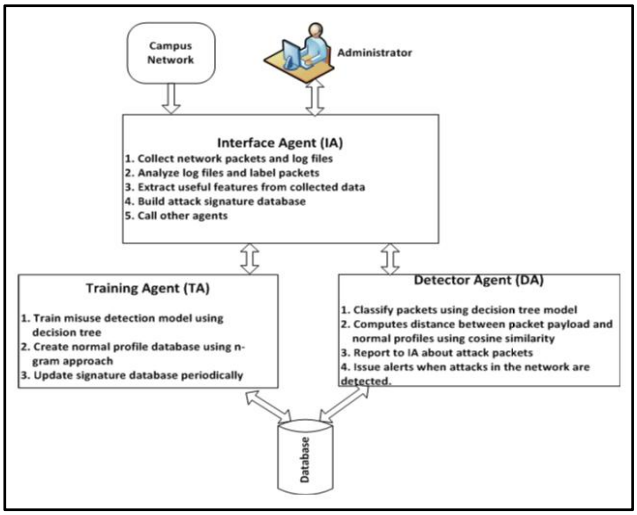

Figure 2: Tasks of Agents

4. IMPLEMENTATION OF INTEGRATED SCHEME

4.1 Data-sets Used

DARPA'99 Data-set: DARPA data-set is a well-known standard data-set for intrusion detection. It consists of different types of data such as TCPDUMP data, Basic Security Module (BSM), and audit data. For our experiments, DARPA 1999 TCPDUMP data is used [14]. As most of the Internet applications use TCP Transmission Control protocol), we only examined the inbound TCP traffic from the hosts 172.16.X.X for ports 21, 23, and 80. In this research work, network packets have been read, and useful features are extracted using information gain. The features selected for misuse detection are source IP, source port, destination IP, destination port, payload length, and TCP flags. The values for source port range from 23 to 49724; payload length ranges from 0-1460. The built-in MATLAB function grp2idx() is used to create groups of data categorized by IPs. There are 38 unique source IPs and 11 unique destination IPs in DARPA'99 training set.

Panjab University Data-set: In Panjab University, a Unified Threat Management System (UTM) is placed between campus network and the Internet. This UTM has a firewall and other features embedded in it. It inspects the total incoming and outgoing traffic. The UTM is further connected to the campus network through a layer core switch. The 10Gbps port on which this Layer 3 switch is connected is mirrored to another port on the same switch. The mirrored port on core switch is used to collect whole the inward and outward traffic of campus to and from the Internet and is dumped to a server connected on this port. University data-set is collected to model normal and anomalous events and also for testing the proposed model. For this purpose, 3 hours of Panjab University campus network data is collected at the egress router and analyzed using Wireshark. Log files are also collected of these 3 hours which are maintained by the server. From the research point of view, we have extracted only web traffic of port 80 and port 443 (i.e., HTTP and HTTPS only) on which we train the IDS. The different parameters like protocol, source IP, destination IP, source port, destination port, payload length, flags associated to the packets, etc. are extracted from each network packet. To label these packets the fields in log files associated with these packets where UTM shows that these packets are blocked or allowed are used. This database acts as input to build a misuse detection model.

4.2 Tools and Techniques used in implementing the proposed model

Several tools, libraries, and techniques/technologies are used in different phases for implementation of proposed integrated scheme. The details are shown in Table 1.

Table 1: Tools and Libraries Used

S. No.	Tool / Library	Purpose
1	Java Development Kit	For the development of Model
2	Java Swings and JFrames	To design interface
3	Java Agent Development Environment (JADE)	For multi-agent system development
4	MySQL	For database creation and management
5	JDBC drivers	The database is connected with agents using JDBC drivers.
6	Wireshark	For capturing network packets
7	jNetPcap	To extract useful information from network packets.
8	MATLAB	For mathematical computation, algorithm implementation, and interact with functions written in Java.
9	MATLAB Control Library	To integrate Java programs with MATLAB.

5. TESTING AND RESULTS

The integrated approach is tested on DARPA as well as on collected University data-set. The results of testing of the integrated approach are given in Table 2. The performance is based on accuracy, detection rate (DR), and false positive rate (FPR).

This table shows the DR, FPR and accuracy of all three modules i.e., misuse detection module, anomaly detection module, and integrated module on both data-sets.

Approach	Data-set	DR (in %)	FPR (in %)	Accuracy (in %)
Misuse Detection Approach	DARPA'99	70.08	19.84	72.04
Anomaly Detection Approach	DARPA'99	53.67	0.29	57.2
Integrated Approach	**DARPA'99**	**80.45**	**12.02**	**81.36**
Misuse Detection Approach	Panjab University	85.7	14.54	85.68
Anomaly Detection Approach	Panjab University	71.2	1.23	71.5
Integrated Approach	**Panjab University**	**88.69**	**10.25**	**88.71**

From the table, it is observed that the integrated approach gives better results than the approaches used individually for misuse detection and anomaly detection. The detection rate for both data-sets is less when using misuse or anomaly approach individually but it increased when integrated approach is used. Also the accuracy increases with proposed integrated approach for both data-sets as not only the attacks in the database get detected but also the novel attacks on which the classifier was not trained were discovered.

6. CONCLUSION AND SCOPE FOR FUTURE WORK

The proposed integrated scheme is developed and tested by utilizing multi-agent technology. This integrated scheme uses three agents, namely, Interface Agent, Training Agent, and Detector Agent. The time and effort involved in the evaluation of network packets are optimized by the use of agents as they are capable of working as independent software modules and can easily communicate with each other. This makes the system autonomous, and the administrator only needs to take future actions about the suspected machines. Updating the attack signature database is also done by the training agent whenever new intrusions are found in the network. By updating the database new intrusions coming in the network are easily found out at early stages of detection.

Integrated scheme used decision tree approach in misuse detection module to train the system. In constructing the decision tree, the split value on which the tree is divided into two parts is of foremost concern. This system used a gain ratio to compute the split value of a node. Proposed scheme used statistical approach, i.e., 1-gram approach for training the system and testing packets payload in the anomaly detection module that takes the linear time to scan the payload. This can be extended further for 2 or 3 grams to construct more efficient models of normal profiles by providing more computational resources with respect to processor, memory, and storage. Instead of the statistical approach, machine learning

techniques such as SVM and Neural Network can also be used. There is also scope for combining multi-agent based intrusion prevention system which can take decisions on its own to protect the system and data without the need for any human intervention.

In our research, for anomaly detection, we are dealing with complete packets. The proposed approach is more suitable for offline or sand-boxing. For online analysis, an alternate approach based on flows can be employed. In flow-based approach, only header part of packets is considered for analysis. Based on the type of network devices deployed in the network NetFlow, sflow, jflow, or other flow tools can be used.

7. REFERENCES

[1] H.J. Liao, C.H.R. Lin, Y.C. Lin, and K.Y. Tun: Intrusion Detection System: A Comprehensive Review: Journal of Network and Computer Applications. vol. 36, no. 1, pp. 16-24, 2013.

[2] J. Yu, Y.V. Ramana Reddy, S. Selliah, S. Reddy, V. Bharadwaj, and S. Kankanahalli: TRINETR: An Architecture for Collaborative Intrusion Detection and Knowledge-based Alert Evaluation: Advanced Engineering Informatics, Elsevier. vol. 19, no. 2, pp. 93-101, 2005.

[3] A. Abraham, R. Jain, J. Thomas, and S.Y. Han: D-scids: Distributed Soft Computing Intrusion Detection System: Journal of Network and Computer Applications. vol. 30, no. 1, pp. 81-98, 2007.

[4] B. Zhang, X. Pan, and J. Wang: Hybrid Intrusion Detection System for Complicated Network: In Fourth International Conference on Fuzzy Systems and Knowledge Discovery. (FSKD 2007), vol. 4, pp. 251-255, Aug 2007.

[5] A. Herrero, E. Corchado, M.A. Pellicer, and A. Abraham: MOVIHIDS: A MObile-Visualization Hybrid Intrusion Detection System: Neurocomputing. vol. 72, no. 13, pp. 2775-2784, 2009.

[6] S. Ganpathy, P. Yogesh, and A. Kannan, Intelligent Agent-based Intrusion Detection System using Enhanced Multiclass SVM: Computational Intelligence and Neuroscience. vol. 20, pp. 1-10, 2012.

[7]B. Djemaa, and K. Okba: Intrusion Detection System: Hybrid Approach based Mobile Agent: In International Conference on Education and e-Learning Innovations. pp. 1-6, July 2012.

[8]D. Dasgupta, and H. Brian: Mobile Security Agents for Network Traffic Analysis: In Proceedings of DARPA Information Survivability Conference amp; Exposition II, DISCEX'01, vol. 2, pp. 332-340, 2001.

[9] R.M. Elbasiony, E.A. Sallam, T.E. Eltobely, and M.M. Fahmy: A Hybrid Network Intrusion Detection Framework based on Random Forests and Weighted k-means: Ain Shams Engineering Journal. vol. 4, no. 4, pp. 753-762, 2013.

[10] A. Alazab, M. Hobbs, J. Abawajy, and M. Alazab: Using Feature Selection for Intrusion Detection System: In International Symposium on Communications and Information Technologies. (ISCIT), IEEE, pp. 296-301, 2012.

[11]Kajal Rai, M. Syamala Devi, and Ajay Guleria: Decision Tree based Algorithm for Intrusion Detection: International Journal of Advanced Networking Applications. (IJANA), vol. 7, no. 4, pp. 2828-2834, 2016.

[12] K. Wang, and S.J. Stolfo: Anomalous Payload-based Network Intrusion Detection: International Workshop on Recent Advances in Intrusion Detection (RAID). Springer, Berlin, Heidelberg, vol. 4, pp. 203-222, 2004.

[13] Kajal Rai, M. Syamala Devi, and Ajay Guleria: Packet-based Anomaly Detection using n-gram Approach: International Journal of Computer Science and Engineering. (IJCSE), vol.6, no. 5, pp. 366-372, 2018.

[14]https://www.ll.mit.edu/r-d/datasets/1999-darpa-intrusion-detection-evaluation-dataset