International Conference of Science, Engineering & Environmental Technology (ICONSEET), 2(15): 113-126, 2017
ISSN 0794-9650

www.repcomseet.com

Development of Threats Detection Model for Cyber Situation Awareness

Adenusi Dauda A.[a]*, Ayeleso E. C.[b], Kawonise A. K.[b], Ekuewa J.B.[b], Adebayo A. A.[a]

[a]The Polytechnic Imesi-Ile, P.M.B. 001, Imesi-Ile, Osun State.
[b]Federal Polytechnic Ede, P.M.B.231, Ede, Osun State.

Abstract: This study development of a threats detection model for gaining experience of cyberspace condition. This was with a view to timely detecting anomalous activities and taking proactive decision to safeguard the cyberspace. The situation awareness model was modeled using Artificial Intelligence (AI) technique. The cyber situation perception sub-model of the situation awareness model was modeled using Artificial Neural networks (ANN), simulated in MATLAB R2015a using standard intrusion dataset of NSL-KDD'99 and evaluated for threats detection accuracy using precision, recall and overall accuracy metrics. The comprehension and projection sub-models of the situation awareness model were modeled using Rule-Based Reasoning (RBR) techniques. The simulation result obtained for the performance metrics showed that the cyber-situation sub-model of the cyber-situation model performs better with increase in number of training data records.

***Keywords: cyberspace, cyber-situation, cybersecurity, Artificial Intelligence, Awareness, Rule-based, Intrusion*.**

1. INTRODUCTION

Situation awareness has been recognized as one of the important, yet unsolved, issues in many different domains, including human controlled and monitored mobile communication networks, social networks, physical and cyber security systems, disaster monitoring and recovery, epidemic monitoring and control, intelligent transportation systems, financial and investment services, and tactical and operational battle field command and control. Creating cyber awareness is critical to network security. An effective Situation Awareness (SA) system will assist computer networks administrator to understand network conditions which will enable him take cybersecurity precautionary measures. The common feature of all SA systems is that they need to react to a dynamic environment that changes its state independently of whether the human or computer agents act on it. The agents want to act on the environment so that its evolution, at least in the area of interest to the agents, leads to the satisfaction of their goals. Towards this end, the agents need to collect information about the environment (usually from many different sources), make decisions based on the collected information and their knowledge, act according to their decisions, collect feedback from the environment in response to the actions, and update their knowledge (learn) to make better decisions in the future.

Endsley defined Situation Awareness (SA) as the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future [1]. Klein [2] ties the notions of goals, cue salience, expectations, and identification of typical actions to SA, and believes that it is central to the decision-making process. Rasmussen [3] describes a hierarchically organized system for SA consisting of three levels: skill-based, rule-based, and knowledge-based. Pew [4] proposes that SA should integrate the environment, goals, system, available physical and human resources, and other actors Rousseau, et.al [5] state that most SA researchers and practitioners agree that SA represents a body of knowledge with a set of processes (or functions) that serve to develop and update that knowledge. SA can be broken down into several different components. The first component is that of being aware of the current environment. The second component is that of ascertaining the significance of certain events and aspects of the current environment. Third, one must be able to tie the awareness to timely and appropriate responses. In many situations, “appropriate” responses are determined by the degree of their success in accomplishing a particular goal. The goals can be a work goal, such as “navigate the channel to deliver the goods” or “hold the defensive position”. Sometimes it is simply necessary to understand or interpret the environment and report what is seen or noted as anomalies. Researchers have used different techniques of components made up of cyber situation awareness model, (perception, comprehension and projection). Several authors have used soft computing techniques ranging from Artificial Neural Network (ANN), Fuzzy Logic (FL), agent-based, genetic algorithm, Bayessian networks to machine learning techniques for modeling these components [6] [7]; [8].

Creating cyber awareness is critical to effective cybersecurity. The emerging threats are sophisticated, complex and highly dynamic. The actions of cyber threats could be disastrous and inimical to cybersecurity. Organisation networks need to be secure, government information and trade secrets must be prevented from cyber-attack. One of the ways of ensuring this is to develop a cyber situation awareness model. With this, network administrators and other concerned stakeholders will be aware of network situations from time to time and will be able to take appropriate measure to ensure the safety of their networks. Research showed that the study of situation awareness in domain of cybersecurity is still at infant stage and the field is emerging. However, some efforts have been made in this direction but more still have to be done. It is based on this premise that a model to create cyber situation awareness for effective network security, is proposed in this study.

2. RELATED WORKS

Endsley [1] defines SA as the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future. Klein [2] ties the notions of goals, cue salience, expectations, and identification of typical actions to SA, and believes that it is central to the decision-making process. Rasmussen [3] describes a hierarchically organized system for SA consisting of three levels: skill-based, rule-based, and knowledge-based. Pew [4] proposes that SA should integrate the environment, goals, system, available physical and human resources, and other actors. Rousseau, et. al., [5] state that most SA researchers and practitioners agree that SA represents a body of knowledge with a set of processes (or functions) that serve to develop and update that knowledge.

Hayden [9] emphasizes the importance of communication and association across interfaces in a complex system. The paper stated the means by which communication occurs and the speed with which it occurs. At the same time, associations between parts of the system establishes a network of interactions. Together, the amount of communication and association among systems partially determines the behaviour of the whole system.” McGuinness and Foy [10] extended Endsley’s Model by adding a fourth level, which they called resolution. This level tries to identify the best path to follow to achieve the desired state change to the current situation. Resolution results from drawing a single course of action from a subset of available actions. McGuinness and Foy [10] believe that for any fusion system to be successful, it must be resilient and dynamic. It must also address the entire process from data acquisition to awareness, Prediction and the ability to request elaboration (drill-down) for additional data and finishing with an appropriate action. McGuiness and Foy put Endsley’s model and their model into perspective with an excellent analogy. They state that Perception is the attempt to answer the question “What are the current facts?” Comprehension asks, “What is actually going on?” Projection asks, “What is most likely to happen if...?” and Resolution asks, “What exactly shall I do?” The answer to the resolution question is not to tell a decision maker what specific action to perform or what specific decision to make but instead provides options of end actions and how they affect the environment.

Krishna et *al*., [11] proposed model for detecting unknown or novel attack in computer networks. The intrusion detection model which is anomaly in nature was developed using hybrid ANN approach. KDD CUP ’99 dataset and real dataset set were used for evaluating the model. The proposed hybrid ANN model performed better than simple model in terms of detection rate, attack classification, training time and response time. Govindarajan [12] in his own paper addressed problem of effectively classifying intrusion in computer networks, especially in the face of increasing cyber-attacks. The author presented two classification methods involving multilayer perception and radial basis function and an ensemble of multilayer perceptron and radial basis function. The analysis of results shows that the performance of the proposed ensemble method is superior to that of single usage of existing classification methods such as multilayer perceptron.

Adelina *et al*., [13] solved the problem of alert identification. i.e ability to identify attack from normal packet among large set of data log in the system. A hybrid clustering algorithm (Extensive Leader Farthest Linkage (ELFL)) is proposed. It is applied to alert dataset to cluster alert. The clustering rate of this algorithm is high. Therefore, it reduces false alarm. Kamaruzaman*et al*., [6] argued that human intervention is still much need in traditional IDS. Traditional IDS can detect intrusion but cannot respond toward it. The paper therefore, analyzed the evolution of IDS and proposed how mobile Agent could increase the integrity of the traditional IDS without human intervention. The implementation of the intelligent mobile Agent is expected to increase integrity of IDS. Nabil [8] argued that the work of an administrator increases with large volume of data to be analyzed. He therefore developed an intrusion detection system using mobile agent and artificial neural network techniques. The technique increased efficiency and accuracy of intrusion detection.

Shanmugavadivu [14] realized that traditional intrusion detection relies too much on the extensive knowledge of security experts, in particular on their familiarity with the computer system to be protected. This is a problem of dependency. The author therefore designed a Fuzzy logic-based system for effectively identifying the intrusion activities within a network. Automated strategy was used for fuzzy rules generations which are obtained from the definite rules using frequent items. The experiment and evaluation of the proposed model was carried out using NSL-KDD Cup 99 dataset. The experiment performed showed that the proposed Fuzzy logic-based model achieved higher precision in identifying whether records are normal or attack. Mahbod et *al*. [15] in their paper titled “A Detail Analysis of the NSL-KDD Cup 99 Dataset” addressedthe problem which affects performance evaluation of system and which results in poor evaluation of IDS through the use of NSL-KDD Cup 99 dataset. The authors conducted statistical analysis on the NSL-KDD Cup 99 dataset to find the deficiency in term of redundant data record. Based on the deficiency of the NSL-KDD 99 dataset, new NSL-KDD dataset has been proposed. This dataset has overcome some redundant disadvantage of the NSL-KDD 99 dataset.

3. MOTIVATION

The modern day Cyberspace is full of uncertainty in form of sophisticated and devastating cyber threats which are smart enough to spoofed security policies of computer networks, thereby causing havoc which could be debilitating to normal functioning of a Computer Network and adversely affect security of confidentiality, integrity and availability of information in government or private organisation networks. It will be a good effort in the right direction to develop a model that can capture the condition of the cyberspace, comprehend it and project the near future at any point in time for the cyberspace, to assist network administrator or concerned individual to take proactive Cybersecurity actions when the need arises. Therefore, this research proposes to develop cyber situation awareness model which will always assist user to understand Cyberspace condition from time to time.

Having the understanding of impending or potential attacks in Computer Network is crucial to Network security. The cyber threats are complex and could be detrimental to normal network operations. Cyber threats are devastating to the extent that if proper measure is not taken to counteract them, they have adverse effect on socio-economic and security of organisations and could affect global security and economy. Therefore, this research is pertinent as it proposes a solution which will create situation awareness of the cyberspace for effective cybersecurity.

4. Model Design Overview

Understanding the situation in the cyberspace is critical to effective cybersecurity. Situation awareness is a proactive means of dealing with uncertainty. Network administrators need aid in gaining full awareness of their networks in order to assist them in effective cybersecurity decision making. In this study a cyber situation awareness model was developed with aim of improving cybersecurity strategy and proactive decision making.

The overall design objective of the model proposed in this study is to develop a cyber situation awareness model which can be used to proactively predict cyberspace condition and based on facts gathered, reasonable cybersecurity decision could be made. The model was designed using Artificial Intelligence (AI) techniques and Rule Base Reasoning (RBR). The conceptual overview and the detail architecture of the model are shown in Figures 1 and 2 respectively. The proposed model has three modules: perception module, comprehension and projection modules. The perception module was modelled as intrusion detection system (IDS) using Artificial Neural Network (ANN) technique. The comprehension module was modelled as Rule Base Reasoning (RBR). The projection module of the system was also modelled as Rule Base Reasoning (RBR). The Perception component of the proposed model has a dimensionality reduction sub-module. The dimensionality reduction component was implemented using Principal Component Analysis (PCA).

The research data used in this study was the standard and conventionally accepted dataset of NSL-KDD'99 cup used for evaluating intrusion detection system. It works in such a way that perception module analyses the network traffic for abnormal behaviour. If the perception module which is also the IDS detects any abnormality, it will immediately send out signal to comprehensioner module for interpretation of the situation. The comprehensioner having interpreted the situation will send to the projection module for prediction of what is happening and what may likely happen in the cyberspace.

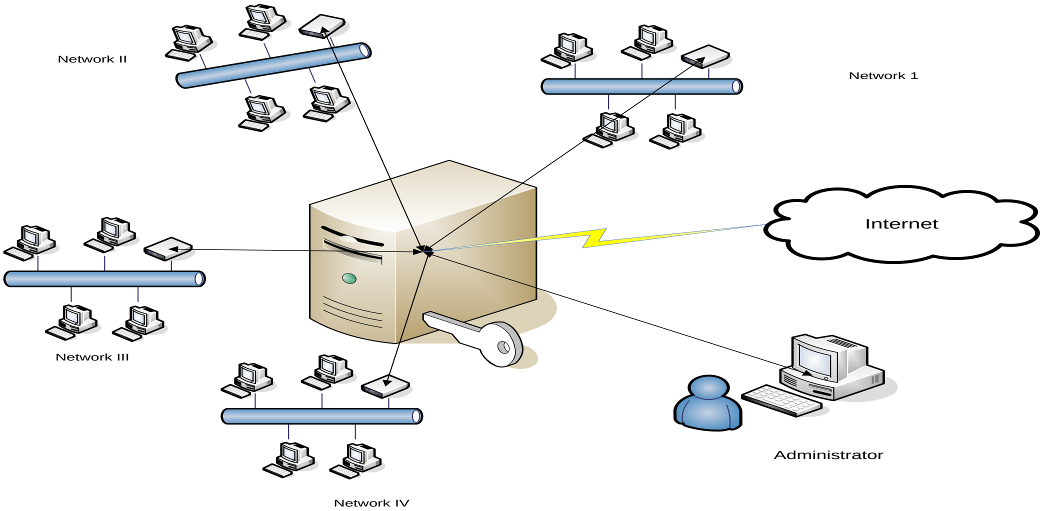

Figure 1: Conceptual View of the Cyber Situation Awareness Model

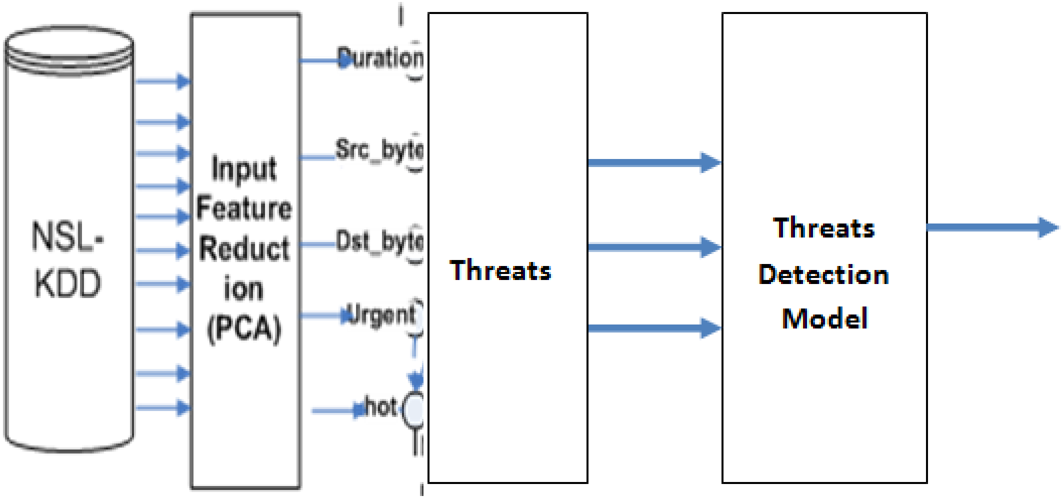

Figure 2: Proposed Model

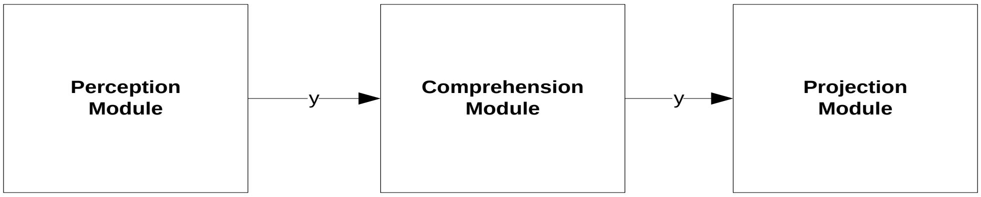

Figure 3: Situation Awareness Model Architecture

Perception Component

This component of the model is used to sense the condition and activities going on in the cyberspace. The cyberspace condition could be normal or malicious at a particular point in time. Therefore, there is need to monitor what goes on in the cyberspace. The perception component performs this function and it is core to this model because the accuracy and effectiveness of the proposed cybersecurity awareness model proposed in this study depends on this component. This component is made up of dimensionality reduction and Intrusion Detection System (IDS) sub-components.

A. Dimensionality Reduction

Effective input attributes selection from intrusion detection datasets thereby reducing the input data dimension is one of the important research challenges for constructing high performance IDS. Irrelevant and redundant attributes of intrusion detection dataset may lead to complex intrusion detection model as well as reduce detection accuracy. The attribute selection methods of data mining algorithms identify some of the important attributes for detecting anomalous network connections. The main advantage of doing dimensionality reduction is that it will reduce the memory requirement and increases the speed of execution thereby increases the overall performance. The dimensionality reduction here is done by using Principal Component Analysis (PCA).

(i) Principal Component Analysis

Principal Component Analysis is one of the most widely used dimensionality reduction techniques for data analysis and compression. It is based on transforming a relatively large number of variables into a smaller number of uncorrelated variables by finding a few orthogonal linear combinations of the original variables with the largest variance. The first principal component of the transformation is the linear combination of the original variables with the largest variance; the second principal component is the linear combination of the original variables with the second largest variance and orthogonal to the first principal component and so on. In many data sets, the first several principal components contribute most of the variance in the original data set, so that the rest can be disregarded with minimal loss of the variance for dimension reduction of the data [15]. PCA reduces the amount of dimensions required to classify new data and produces a set of principal components, which are orthonormal eigenvalue/eigenvector pairs. The steps for doing principal component analysis are given in the algorithm as follow:

1. Get input data
2. Subtract the mean
3. Calculate the covariance matrix
4. Calculate the eigenvectors and eigenvalues of the covariance matrix.
5. Sort the Eigen values in descending order.
6. Calculate the feature vectors.[i]

B. Intrusion Detection System (IDS)

Intrusion Detection Systems can be classified based on their monitoring activity, detection technique and their response to the attack. This paper explores two methods which are: monitoring activity and detection technique.

(a). Intrusion Detection based on monitoring activity

The Intrusion Detection Systems can be broadly divided into two types based on the fact that whether they monitor the whole network or a particular host. Accordingly, they are termed Host based Intrusion Detection System (HIDS) and Network based Intrusion Detection System (NIDS).

i. Host based Intrusion Detection System

The Host Based Intrusion Detection System resides in the host and traffic data is analyzed individually in each host. The IDS monitors the various file systems, network events and system calls to detect any possible threat to the system.

ii. Network based Intrusion Detection System (NIDS)

The NIDS monitors the packets passing through the entire network and analyses the packets. Network Based Intrusion Detection System is particularly useful for monitoring traffic of many systems all at once.

(b). Intrusion Detection based on detection technique

Intrusion Detection System is also classified on the basis of the technique used by the IDS to look up for vulnerabilities. They are mainly classified into Signature Based Detection and Anomaly Detection.

i. Signature Based Detection

The Signature Based Detection compares a possible threat with the attack type already stored in the IDS. The limitation of this type of detection technique is that if any new type of threat comes which is not already known to the IDS, the system becomes vulnerable to that attack.

ii. Anomaly Based Detection

The Anomaly Based Detection is a detection technique by which the IDS looks for vulnerabilities based on rules set forth by the user and not on the basis of signatures already stored in the IDS. This type of detection usually uses Artificial Intelligence to distinguish between normal traffic and anomalous traffic.

Mathematical model of the MLP based IDS
A = X_1 = ‘duration’
B = X_2 = ‘src_bytes’
C = X3 = ‘dst_bytes’
D = X4 = ‘land’
E = X_5 = ‘wrong fragment’
F = X_6 = ‘urgent’
G = X_7 = ‘hot’
I = X_8= ‘num_failed_login
AK = X_9= ‘dst_host_rerror_rate’
AL = X_n = ‘dst_host_srv_rerror_rate output’

$$A_{out(1)}, A_{out(2)}, A_{out(3)}, \dots\dots\dots\dots\dots\dots\dots, A_{out(k)} \qquad (1)$$

Where,

k is the number of output defined for the detector (IDS) model.

Comprehension Component.

This component includes the process of combining, interpretation, storing and retaining information. Comprehension of information means that multiple pieces of information (received from multiple sources) must be integrated and relevance of the information must be evaluated within the scope of the current situation and the goals of the network administrator.

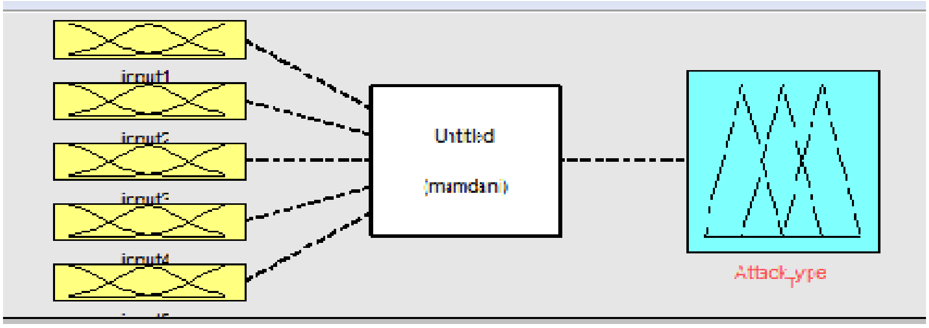

Figure 4: Fuzzy Logic-Based Comprehensionner

The output of perception module is fed into comprehension module for interpretation of the cyber situation. This module was modelled using rule base reasoning (RBR) approach. Some of the rules written to convey the meaning of cyber situation at a particular point in time are as written below.

Rule1: $IF A_1^c = 1 \ AND A_2^c = 0 \ AND A_3^c = 0 \ AND A_4^c = 0 \ AND A_5^c = 0 \ AND A_6^c = 0 \ THEN \text{"NORMAL"}$

Rule 2: $IF A_1^c = 0 \ AND A_2^c = 1 \ AND A_3^c = 0 \ AND A_4^c = 0 \ AND A_5^c = 0 \ AND A_6^c = 0 \ THEN \text{"DOS"}$

Rule 3: $IF A_1^c = 0 \ AND A_2^c = 0 \ AND A_3^c = 1 \ AND A_4^c = 0 \ AND A_5^c = 0 \ AND A_6^c = 0 \ THEN \text{"R2L"}$

Rule 4: $IF A_1^c = 0 \ AND A_2^c = 0 \ AND A_3^c = 0 \ AND A_4^c = 1 \ AND A_5^c = 0 \ AND A_6^c = 0 \ THEN \text{" PROBE"}$

Rule 5: $IF A_1^c = 0 \ AND A_2^c = 0 \ AND A_3^c = 0 \ AND A_4^c = 0 \ AND A_5^c = 1 \ AND A_6^c = 0 \ THEN \text{"U2R"}$

Rule6:$IF A_1^c = 0 \ AND A_2^c = 0 \ AND A_3^c = 0 \ AND A_4^c = 0 \ AND A_5^c = 0 \ AND A_6^c = 1 \ THEN \text{"EMERGENT"}$

Projection Component.

This module is used to predict future events from the available past and present data, which experienced operators do on a regular basis. The analysis and interpretation made by the comprehension module is used for prediction of the cyberspace situation. This module was also modeled using Rule Based Reasoning (RBR).

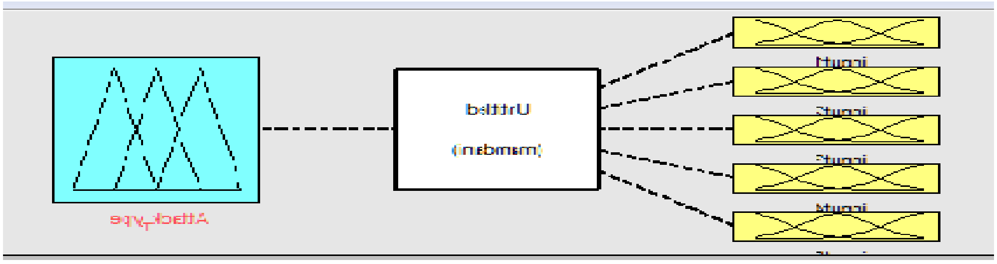

Figure 5: Fuzzy logic-based Projectionner

If malicious activity is suspected in the cyberspace, the comprehension module will interpret it and reveal the type of threats looming in the cyberspace. It is based on this interpretation that the projection component will make a projection on what is like to happen and the solution to the looming cyber situation. A sample rule written for this module is written as follow

IF Situation = “Normal” THEN “No threat”

IF Situation = “DOS” THEN “Legitimate access may be denied”
IF Situation = “R2L” THEN “Information stealing”
IF Situation = “Probe” THEN “Searching for system weakness”
IF Situation = “U2R” THEN “Weakness problem”

RESEARCH DATASET DESCRIPTION.

In 1998, DARPA in concert with Lincoln Laboratory at MIT launched the DARPA 1998 dataset for evaluating IDS. The DARPA 1998 dataset contains seven weeks of training and also two weeks of testing data. In total, there are 38 attacks in training data as well as in testing data. The refined version of DARPA dataset which contains only network data (i.e. Tcpdump data) is termed as KDD dataset [15].

A variety of attacks incorporated in the dataset fall into following four major categories:

1. Denial of Service Attacks: A denial of service attack is an attack where the attacker constructs some computing or memory resource fully occupied or unavailable to manage legitimate requirements, or reject legitimate users right to use a machine.
2. User to Root Attacks: User to Root exploits are a category of exploits where the attacker initiate by accessing a normal user account on the system (possibly achieved by tracking down the passwords, a dictionary attack, or social engineering) and take advantage of some susceptibility to achieve root access to the system.
3. Remote to User Attacks: A Remote to User attack takes place when an attacker who has the capability to send packets to a machine over a network but does not have an account on that machine, makes use of some vulnerability to achieve local access as a user of that machine.
4. Probing: Probing is a category of attacks where an attacker examines a network to collect information or discover well-known vulnerabilities.

These network investigations are reasonably valuable for an attacker who is staging an attack in future. An attacker who has a record, of which machines and services are accessible on a given network, can make use of this information to look for fragile points. Table I shows different attack type included in the KDD’99 dataset for evaluating the performance of IDS. It is very difficult to execute the proposed system on the KDD cup 99 dataset since it is a large scale.

Table 1: Attack Category described in KDD’99 dataset used

S/N	**Attack Group**	**Different Attacks**
1.	Denial of Service attacks	Back, Land, Neptune, smurf, teardrop
2.	Remote to Local Attacks	Ftp_write, guess_passwd, imap, multihop, phf, spy, warezclient
3.	Probes Attacks	Satan, ipsweep, nmap, portsweep
4.	User to Root Attacks	Buffer_overflow, Loadmodule, perl, rootkit

Therefore, a subset of 10% of KDD Cup 99 dataset for training and testing was used. The number of records taken for testing and training phase is given Table 2 and Table 3. This selection is based on previous research in intrusion detection system [14].

Table 2: Training Dataset		Table 3 : Test dataset	
Training Dataset		**Testing Dataset**	
Normal	25,000	Normal	26,000
DOS	25,000	DOS	26,000
Probe	4107	Probe	4107
RLA	77	RLA	77
URA	42	URA	42

5. PERFORMANCE EVALUATION MODELLING OF THE IDS MODULE

Accuracy deals with the proper detection of attacks and the absence of false alarms. Inaccuracy occurs when an intrusion-detection system suspects a legitimate action in the environment as anomalous or intrusive. This metric is used independently to evaluate the performance of the Intrusion Detection System component (IDS) developed for this model. This component is expected to be installed on all the client networks that will use this proposed cybersecurity platform, as it is the component of the system that will sense the cyber situation. For this component to be suitable for the purpose of generating reliable information suitable for sensing cyber situation, it must have high detection accuracy couple with capability to detect zero day threats. To accomplish this, ANN was chosen and experimented with. The accuracy of the techniques was carried out using precision, recall and overall accuracy as described below.

Precision of accuracy metric is calculated as

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

Recall of the detection accuracy is calculated as

$$\text{Recall} = \quad \mathrm{Re}\,call = \frac{TP}{TP + FN} \tag{3}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \tag{4}$$

Where,

TP = True positive
TN = True negative
FN = False negative
FP = False positive

6. Experimental Design for the IDS Module

In the study, ANN-based model was designed for the intrusion detection model. The dimensionality reduction process conducted using principal component analysis was applied to the model in order to get the required input for the model. By so doing, redundancies and irrelevant attributes that do not contribute any significance value to the dataset classification were removed. In this study a multi class problem was solved. Here, five classes case were described. An output layer with five neurons (output states) was used: [1 0 0 0 0] for normal conditions, [0 1 0 0 0] for DoS, [0 0 1 0 0] for R2L, [0 0 0 1 0] for probe and [0 0 0 0 1] for U2R. The desired output vectors used in training and testing phases are simply as mentioned above.

Model Simulation Result and Analysis

In this section the mathematical models formulated for the perception component of the cyber situation model was simulated. The threat detection accuracy simulation for perception component was carried out using ANN technique. MATLAB R2015a software was used in this research, the dataset was divided into both training and test set for the simulation. The training dataset was used to train the ANN-based perception component so that it could learn.

In the testing phase, randomly selected threat records from the training dataset were presented to the trained ANN-based perception model. The results were obtained and recorded. Secondly, threat data records were also randomly selected from test dataset to ascertain detection capability of the perception component of the cyber situation model. The essence of testing with test dataset was to test for novelty detection capability of the developed perception sub-model of the cyber situation awareness model. The test dataset is new to the perception model because it has not been exposed to such dataset before. The test dataset has a lot of interesting features which assisted us in analyzing the performance of the developed model. Dataset were randomly selected from different categories and were presented to the models for classification. The obtained results were then used to compute overall accuracy of the developed cyber situation awareness models. The overall accuracy of the developed model was computed based on the definitions, namely; precision, recall and accuracy chosen as performance metrics. These parameters are normally used to estimate the class prediction. Table 4 – Table 9 show different simulation results both at training and test phases using different sizes of dataset figure 6 – Figure 11 also show simulation results both at training and test periods.

Table 4: Performance of ANN-based perceptioner during training (5,000 threats record)

S/N	**Packet Type**	**Actual Data**	**Predicted Data**				**Precision**	**Recall**	**Accuracy**
			Correct Classification		**Wrong Classification**				
			TP	**TN**	**FP**	**FN**			
1.	Normal	100	0	74	26	0	1.00	0.00	0.74
2.	DOS	100	72	0	0	28	1.00	0.72	0.72
3.	Probe	100	68	0	0	32	1.00	0.68	0.68
4.	R2L	20	13	0	0	7	1.00	0.65	0.65
5.	U2R	20	12	0	0	8	1.00	0.60	0.69

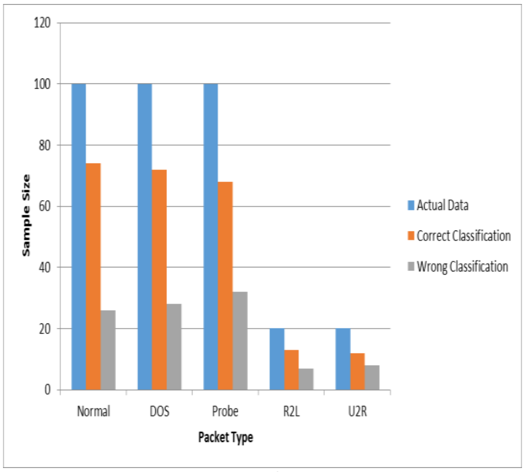

Figure 6: Detection Accuracy of ANN-based Perceptioner on Actual and Predicted Dataset During Training (5,000 Records).

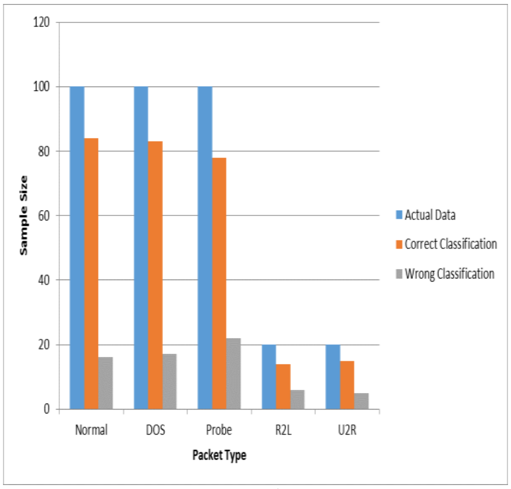

Figure 7: Detection Accuracy of ANN-based Perceptioner on Actual and Predicted Dataset During Training (5,000 Records)

Table 5: Performance of ANN-based perceptioner during training (20,000 threats record)

S/N	**Packet Type**	**Actual Data**	**Predicted Data**				**Precision**	**Recall**	**Accuracy**
			Correct Classification		**Wrong Classification**				
			TP	**TN**	**FP**	**FN**			
1.	Normal	100	0	84	16	0	1.00	0.00	0.84
2.	DOS	100	83	0	0	17	1.00	0.83	0.83
3.	Probe	100	78	0	0	22	1.00	0.78	0.78
4.	R2L	20	14	0	0	6	1.00	0.70	0.70
5.	U2R	20	15	0	0	5	1.00	0.75	0.75

Table 6 : Performance of ANN-based perceptioner during training (60,000 threats record)

S/N	**Packet Type**	**Actual Data**	**Predicted Data**				**Precision**	**Recall**	**Accuracy**
			Correct Classification		**Wrong Classification**				
			TP	**TN**	**FP**	**FN**			
1.	Normal	100	0	99	1	0	1.00	0.00	0.99
2.	DOS	100	97	0	0	3	1.00	0.97	0.97
3.	Probe	100	95	0	0	5	1.00	0.95	0.95
4.	R2L	20	18	0	0	2	1.00	0.90	0.90
5.	U2R	20	17	0	0	3	1.00	0.85	0.85

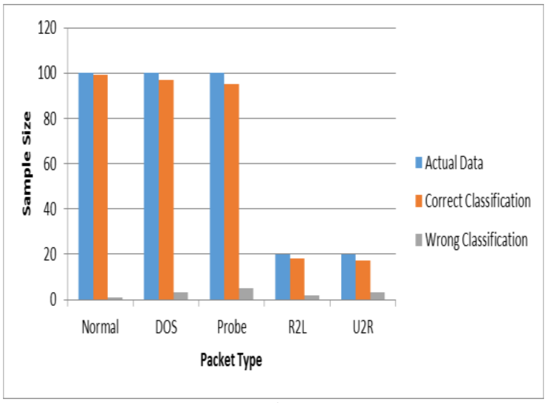

Figure 8: Detection Accuracy of ANN-based Perceptioner on Actual and Predicted Dataset During Training (60,000 Records).

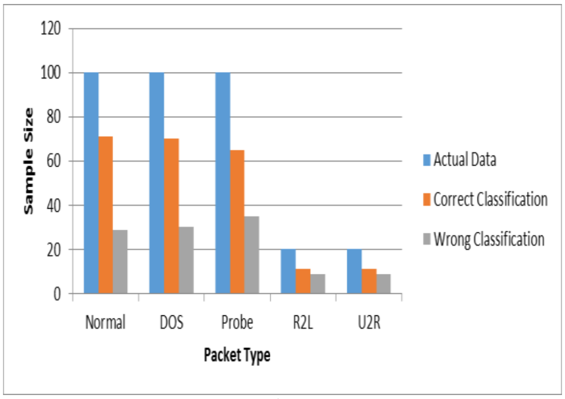

Figure 9: Detection Accuracy of ANN-based Perceptioner on Actual and Predicted Dataset During Training (60,000 Records

Table 7: Performance of ANN-based perceptioner during test (5,000 threats record)

S/N	**Packet Type**	**Actual Data**	**Predicted Data**				**Precision**	**Recall**	**Accuracy**
			Correct Classification		**Wrong Classification**				
			TP	**TN**	**FP**	**FN**			
1.	Normal	100	0	71	29	0	1.00	0.00	0.71
2.	DOS	100	70	0	0	30	1.00	0.70	0.70
3.	Probe	100	65	0	0	35	1.00	0.65	0.65
4.	R2L	20	11	0	0	9	1.00	0.55	0.55
5.	U2R	20	11	0	0	9	1.00	0.55	0.55

Table 8 : Performance of ANN-based perceptioner during test (20,000 threats record)

S/N	**Packet Type**	**Actual Data**	**Predicted Data**				**Precision**	**Recall**	**Accuracy**
			Correct Classification		**Wrong Classification**				
			TP	**TN**	**FP**	**FN**			
1.	Normal	100	0	81	19	0	0.00	0.00	0.81
2.	DOS	100	79	0	0	21	1.00	0.79	0.79
3.	Probe	100	72	0	0	28	1.00	0.72	0.72
4.	R2L	20	12	0	0	8	1.00	0.60	0.60
5.	U2R	20	13	0	0	7	1.00	0.65	0.65

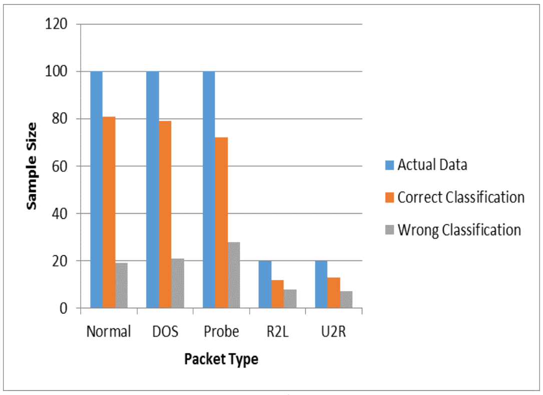

Figure 10: Detection Accuracy of ANN-based Perceptioner on Actual and Predicted Dataset During Test (20,000 Records).

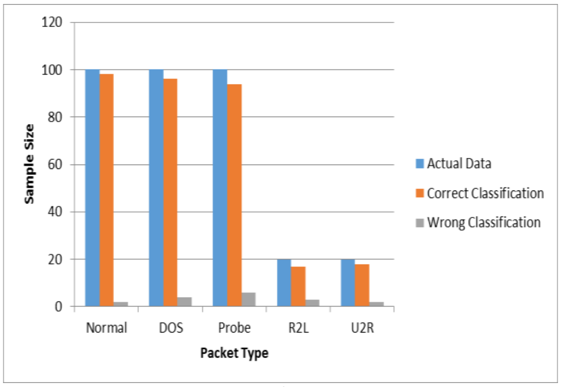

Figure 11: Detection Accuracy of ANN-based Perceptioner on Actual and Predicted Dataset During Test (60,000 Records)

Table 9: Performance of ANN-based perceptioner during test (60,000 threats record)

S/N	Packet Type	Actual Data	Predicted Data				Precision	Recall	Accuracy
			Correct Classification		Wrong Classification				
			TP	TN	FP	FN			
1.	Normal	100	0	98	2	0	0.00	0.00	0.98
2.	DOS	100	96	0	0	4	1.00	0.98	0.98
3.	Probe	100	94	0	0	6	1.00	0.94	0.94
4.	R2L	20	17	0	0	3	1.00	0.85	0.85
5.	U2R	20	18	0	0	2	1.00	0.90	0.90

7. SUMMARY

In this research, efforts have been channeled towards presenting a development of a cyber-situation awareness model for gaining knowledge of the cyberspace condition. The model has the capability for threats detection (perception), analysis (comprehension) and cyberspace condition projection. Figure 6-11, clearly shows the performance of the study for different threat records on actual and predicted dataset during training and test phase.

The simulation results showed that the situation awareness model during training with 5000 threats record classified records with 0.74 accuracy for normal records, 0.72 accuracy for DOS, 0.68 accuracy for Probe, 0.65 for R2L and 0.69 accuracy for U2R threat record. The perception and classification accuracy of the model increased with increase in the training data. The model had highest accuracy when the training data was increased to 60,000 threat records. The model had accuracy of 0.99 for normal data records, 0.97 for DOS, 0.95 for Probe, 0.90 for R2L and 0.85 accuracy for U2R. During test period when the model was trained with 5000 threat records, 0.71 accuracy was recorded for normal threat records, 0.70 accuracy for DOS, 0.65 for Probe, 0.55 for R2L and 0.55 accuracy was recorded for U2R threats record. Likewise, when the situation awareness model was test during testing with 60,000 threats records, 0.98 accuracy was obtained for normal data records, 0.98 for DOS, 0.94 for Probe, 0.85 for R2L and 0.90 accuracy for U2R. In table 4, for normal threats of 100, no predicted data were true to be normal threat, 74 were predicted to be true as not normal threat and 26 were wrongly classified as false threats. 72, 68, 13 and 12 were classified correctly for DOS, probe, R2L and U2R threats respectively.

The cyber situation model designed was able to meet its overall goal of assisting network administrators to gain experience of cyberspace condition. The model was capable of sensing the cyberspace condition, perform analysis based on the sensed condition and predicting the near future condition of the cyberspace.

Various issues associated with intelligent computing and computer data communication and network with emphasis on Cybersecurity were examined. Other issues such as intrusion detection system (IDS) and machine learning techniques were also examined. Relevant literatures were reviewed to examine modern cyber threats and current cybersecurity trends. The Study proposed a cyber-situation awareness model having three major components: the perception component provides information about the status, attributes and dynamics of relevant elements within the cyber environment. The comprehension component encompasses how people combine, interpret, store, and retain information. The projection represents a prediction of the elements of the environment. The cyber-situational awareness model was designed using artificial intelligence technique. In other to accomplish cyberspace threat perception, an artificial neural network (ANN) used to model it. The comprehension component was achieved by applying Rule Base Reasoning (RBR). This was achieved by creating patterns of normal network packets using IF……THEN programming construct. The projection component was also achieved using RBR.

In order to evaluate the performance of the situation awareness model, mathematical models were adopted for the perception module of the cyber situation awareness model. The model was simulated using MATLAB R2015a. The cyber situation perception sub-model was evaluated by obtaining the performances on both training and test datasets of KDD'99. Precision, recall and overall accuracy were used as performance metrics. The simulation result obtained for the performance metrics showed that the cyber-situation sub-model of the cyber-situation model better with increase in number of training data records.

8. Conclusion

In conclusion, the challenge of effectively safeguarding the cyberspace from complexity and sophistication of emerging cyber threats motivated this research. The study therefore presented a proactive approach to secure the cyberspace. The goal is to accurately and timely gain knowledge of the cyberspace for quick analysis, projection and proactive decision making has been achieved with the use of ANN-based approach presented in this study.

From the results, it was clear that for the model to provide accurate awareness, the more number of threats used in training the model the more accurate the model will get. This is true as artificial neural network performs better with large dataset.

The effectiveness of the methodology enhanced the security property of confidentiality, integrity and availability of computer networks in this era of emerging sophisticated threats. The designed model will greatly have enhanced the cyberspace availability, quality of service and users' confidence. The model developed in this study can be used for cyberspace monitoring for effective cybersecurity strategy. The acceptability and implementation of this model will be a useful tool for network administrator to effectively manage their network.

FUTURE WORK

From the results of this research, the following are recommended:

(i) For effective operation of the developed cybersecurity software, government should provide a national backbone. A backbone is a very fast with high bandwidth connection. This will facilitate effective operation of the new solution developed in this research. Because effective operation of the cybersecurity software requires high speed link connections and bandwidth.

(ii) Securing the cyberspace is not an easy task, it requires huge capital. It is therefore recommended that government make special budget for cybersecurity research and should help in training cybersecurity professional.

(iii) Government should enact strong cyber laws, reviews these laws from time-to-time and, if necessary, drafts new criminal laws, procedures, and policy to deter, respond to and prosecute cybercriminals. This is because technical solution may not be enough to address the complexity of the problems.

(iv) Individual, private organization and government establishment should cooperate with this effort to secure the cyberspace by securing their network using this new platform when fully developed. This is because the proposed cybersecurity solution has a promising future to cope with the complexity of the modern cyber threats.

REFERENCES

[1] Endsley, M.R. (1995). “Toward a Theory of Situation Awareness in Dynamic Systems”. Human Factors. 37(1), 32-64, 1995.

[2] Klein, G. (1997). “The Recognition-Primed Decision (RPD) Model”: Looking Back, Looking Forward. in C.E. Zsambock and G. Klein (eds), Naturalistic Decision Making, Lawrence Erlbaum Associates, Mahwah, pp. 285-292, 1997.

[3] Rasmussen, J. (1983). “Skills, Rules, and Knowledge”: Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models”. IEEE Transactions on Systems, Man and Cybernetics. 13 (3), 257-266.

[4] Pew R. W, (2000). The state of situation awareness measurement: heading toward the next century. In M. R. endsley and D. J. Garland, eds. Situation awareness analysis and measurement. Mahwah, N. J: Lawrence Erlbaum Associates, 33 – 50.

[5] Rousseau, R., Tremblay, S., and Breton, R., (2004) Defining and modeling situation awareness: a critical review, In: S. Banbury and S. Tremblay, eds. A cognitive approach to situation awareness: theory and application. Aldershot: Ashgate.

[6] Kamaruzaman M., Mohd A., Mohd S., Mohammad A.K. and Mohd R. M. (2011). Mobile Agents in Intrusion Detection System: Review and Analysis”. Modern and Applied Science. 5(6), 218 – 231.

[7] SueraBoran and KerimGoztepe (2010). Development of a Fuzzy decision support system for com modity Acquisition using Fuzzy Analytic Network process.

[8] Nabil E. K., Karim H., andNahla E. Z. (2012). “A Mobile Agent and Artificial Neural Networks for Intrusion”Detection.Journal of Software. 7(1), 156 – 160.

[9] Hayden, N.K. (2006). “The Complexity of Terrorism: Social and Behavioural Understanding, rendsfor the Future”. Draft Document, to be published by Routledge Press.

[10] McGuinness B. and Foy. J. L. (2000) “A subjective measure of SA: The crew awareness ratingscale (cars)”. In Proceedings of the first human performance, situation awareness, and automation conference, Savannah, Georgia, USA.

[11] Krishna Champaneria, Bhavin Shah Krunal Panchal, (2014). “Survey of Adaptive Resonance Theory Techniques in IDS”, International Journal of Emerging Technology and Advanced Engineering, Volume 4, Issue 12.

[12] Govindarajan M. andChandrasekaran R.M. (2010). “Intrusion Detection using Neural Based Hybrid Classification Methods”. Elsevier, Computer Networks Journal. 55(2011), 1662 – 1671.

[13] Adelina J. D., Anushiadevi R. And Lakshminarayanan T. R. (2012).” An Efficient Algorithm for Clustering Intrusion Alert”. Journal of Theoretical and Applied Information Technology. 37 (2), 234 – 240.

[14] Shanmugavadivu R. and Nagarajan N. (2012). “Learning of Intrusion Detector in Conceptual Approach of Fuzzy Towards Intrusion Methodology”. International Journal of AdvancedResearch in Computer Science and software Engineering. 2(5) 246 – 250.

[15] MahbodTavallaee, EbrahimBagheri, Wei Lu, and Ali A. Ghorbani, (2009). “A Detailed Analysis of the KDD CUP 99 Data Set”. Proceeding of the 2009 IEEE Symposium on Computational Intelligence in security and defence Application (CISDA).