EFFICIENT DETECTION OF BLACK HOLE ATTACKS IN MANETS USING MACHINE LEARNING TECHNIQUES AND NETWORK DYNAMISM

Mohammed Shabaz Hussain [1] and Khaleel Ur Rahman Khan[2]
[1] *Research Scholar, Rayalaseema University, Kurnool, AP*
[2]*Department of Computer Science Engineering, ACE, Hyderabad, India*
(E-mail: mshabazh@gmail.com,khaleelrkhan@gmail.com)

Abstract — It is implicit that the entire network nodes among each other to transmit data packets inside a multi-hop manner adhoc networks as the amount of cyber at tacks have enlarged, detecting the infringements inside networks becomes a very difficulty job. Intended for network s intrusion detection scheme, numerous data mining algorithms along with machine learning strategies are pursued .In this paper we propose a novel hybrid approach of two stage IDs system with combination of J48 decision tree and dual attack detection meant for black hole attack for manet's .The projected scheme selects intrusion detection system network node amid three added features such as Hop count, The Destination Sequence number and the route errors reported. We integrate J-48 classifiers into manet.Experimental results showed a significant improvement over the existing approaches

Keywords—*Mobile Adhoc Networks; Intrusion Detection System; Decision Tree Algorithms; Machine Learning Technique;, Ad Hoc On Demand Vector Routing; Black Hole Attack; J48 Classification;, C4.8 Algorithm.*

t; formatting; style; styling; insert (key words)

I. Introduction

In the focal point of current technological advancements involving new concepts such as cloud computing, big data analytics, social media networks, our communities produce a large amount of information. In the midst of this enormous data generated by these new technologies, the search for useful data became important for data scientists, marketers and even businesses. In the centre of this wide range of data exchanged through the Internet or networks, information security is a foremost concern, even though several intrusion deterrence techniques have been developed in the earlier period to get rid of probable threats inspite-of, intrusion attacks are still are increased and augment in complexity, that is why a system is required to detect any unusual or unauthorized traffic that could cause damage to a specific network.

This protection mechanism can be implemented using an IDS that can be defined as a set of software or hardware devices capable of gathering, analyzing and detecting any unwanted, suspicious or malicious traffic on a specific computer host or network node [1].To accomplish security issue, an IDS should utilize some factual or realistic data scheme to peruse and decipher the data it gathers and reports any malicious movement in the network system[2].There still exist one principle issue with respect to the interruption in network system that is the association of human cooperation with regards to mark the traffic between an anomaly node and normal node, another significant concern is the new challenge of "Large Data" and "Distributed computing. These two techniques produce a lot of information that must be gathered and investigated by the intrusion detection mechanism regularly. The IDS needs to manage multi-dimensional information created by these enormous amounts of information. It is important to consider the interruption dataset can be immense in size, the measure of perceptions developed, however the list of tracked anomaly nodes can likewise increment essentially and may create an impressively measure of bogus positives results as it can contain numerous excess or copy records [3].An artificial intelligence approach and information mining strategy which is the use of ML techniques in huge database are generally known and used to lessen intricacy.

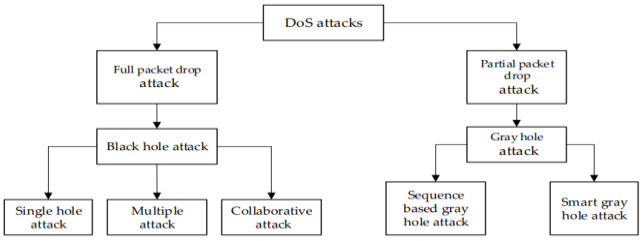

Figure 1. Types of Dos attacks in MANETs.

ML help towards optimizing performance criterion by example data information or precedent experience by means of program, schemas are defined by few attributes, towards optimizing the attributes of the model by making use of training data. The system could be prescient towards making predictions in the prospect, or expressive to increase information from data. To get an analytical or significant job, ML normally uses two important schemes: Classification and Clustering. Inside categorization, the program have to forecast the most feasible class or label for novel observation in one or more already defined classes/label though clustering, the classes may not be predefined throughout the learning progression. If we want to recognize the kind of clustering, intrusion may be further cooperative [5]. Unlikely, these schemes encompass unsuccessful towards ensuring a better performance of detection pace. Furthermore, those already existing intrusion detection system aims towards analysing the entire features which could result in a misclassifying of intrusion detection and a fair amount of time duration when developing the model, even with a little concern and critic concerning the model assessment of KDD datasets [8], researchers tranquil to test their model.

The rest of the paper is organized as follows. The next section reviews the related works. In Section 3, we compare the proposed algorithm performance results with other different approaches and continued till the sections number 4,5,6 and 7. The procedure Performance evaluation is explained provided in 8. Finally, the conclusions are briefed in Section 9.

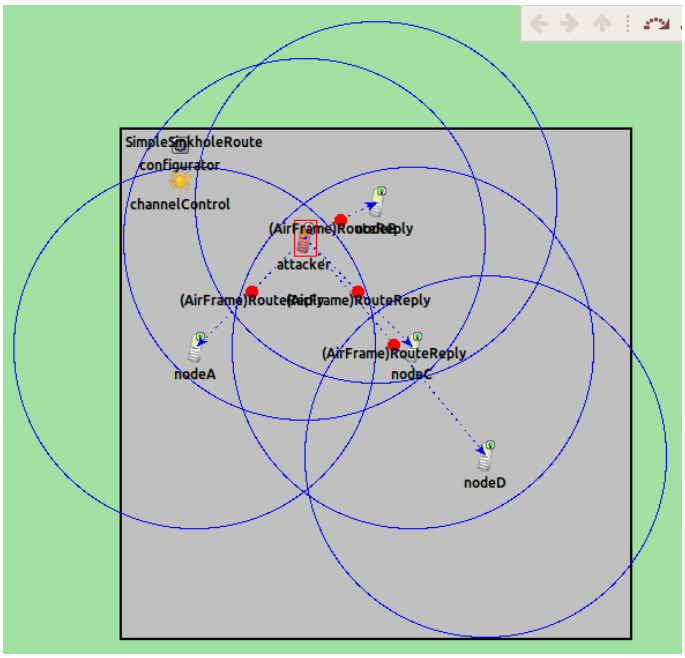

Figure 2. MANET nodes under Black-Hole attack.

II. RELATED WORK

Classification is defined as a procedure of creating model of classes through a collection of records that hold class labels. Decision Tree Algorithm is used to discover the method the attributes-vector behave for numerous instances. The algorithm generates the rules for the forecast of the objective variable. Through the assistance of tree classification scheme, the critical allocation of the information is easily comprehensible [5]. The data mining tool WEKA, open source J48 is a Java accomplishment of the C4.5 scheme. Tree pruning options are provided by the WEKA tool. Fitting pruning intended for précising could be utilized as an instrument in case of potential threat. Furthermore, in algorithms the categorization is done recursively until every single node or leaf is pure, to assist in the classification of the information must to be as better as feasible. This scheme generates the rules from where scrupulous recognition of that information is generated. The purpose is gradual simplification of a decision tree until it gains symmetry of accurateness and litheness.

Numerous techniques of data mining has been worn intended for intrusion detection system. James P. Anderson [3] categorized the coercion and introduces a model which will be able to identify anomalies in the user's behaviour. Afterward many researchers used various techniques i.e., RST (Rough Set Theory), Principal Component Analysis (PCA), SVM (Support Vector Machine), towards making an effective IDS, learning genetic network programming (GNP), Levenberg Marquardt (LM)etc., towards creating an effective IDS. In 2007, researchers Barton P. Miller and Shai Rubin proposed a novel scheme known as proto-matching which concatenate normalization, protocol analysis, as well as pattern matching hooked on a one phase [4]. In 2009, Shang Haikun, Meng Jianliang [5], use K-Means algorithm for IDS. Afterward in 2010, Sara, Fatimah, Mohammaderza, and Lilly [6] employ two proposal i.e., SVM and C4.5 to detect intrusion in network and uncover that C4.5 algorithm efficient than SVM with respect to network intrusion detection. A. Baig, Zubair [7], in his AODE-dependent NIDS, recommended as the Naive Bayes do not precisely find intrusion in network. In 2012, Jain [8], compare quad ML algorithms i.e., BayesNet, J48, NB and OneR , for IDS and outcome depicts to facilitate that J48 decision tree provides better accurateness in comparison with other three schemes. In the year 2012, R Rangaduari proposed a novel scheme Adaptive NIDS makes use of a Hybrid scheme which utilizes two stage approaches: in the early stage, makes use of a probabilistic classifier and in stage two, a traffic model based on HMM is used. V. Jaiganesh [10] utilized Levenberg Marquardt Learning in

combination Kernelized SVM for intrusion detection in network. Gholam Reza Zargar [11] describe the applying of carefully selected, semi-supervised learning, nonparametric schemes to the intrusion in network problem, in their research they compared the efficiency of various model types by making use of feature-based information extracted from operational networks.

Decision tree scheme is categorization technique. It is dependent on the principle of divide and conquer approach. A decision tree constitutes of decision nodes as well as leaf nodes, though decision node predicts assessment on one of the attributes as well as leaf node depicts class value [12]. Each path starting from the root node to leaf is a rule. Classification inaccuracy is performance key factor of decision trees. Classification error can be defined as the proportion of misclassified case [12]. In implementation, the sample training data sets will be usually large, that fallout in additional quantity of branches as well as layers in the created decision tree. Inside decision tree whilst the class classifications will be high, classification accuracy is considerably condensed. There are various schemes to generate decision tree i.e., J48, ID3, BFTree, LMT, FT, LMT and several more. For our research we make use of u J48 algorithm since it have superfluous accuracy rate [8]. Quinalan proposed J48 algorithm.

The major commitments of present work are as follows. Firstly, we present task offloading from multiple client vehicles to surrogate nodes in vehicular edge networks. Furthermore, the proposed task offloading scheme offers a viable solution for offloading tasks of real-time applications by utilizing a viable vehicular resource discovery strategy to find out the up-to-date neighbouring computational resources. Thirdly, it gives scalability by restricting the number of messages stormed amid request-response exchange between clients and surrogates in an Energy Efficient way.

III. Proposed algorithm for black hole detection in MANET

The proposed algorithm is composed of three phases namely resource discovery phase, fitness evaluation phase and surrogate selection phase. Significance of black hole attack identification

The MANET are a network of infrastructure-less nodes that communicate over wireless links. The nodes in MANET act as both routers and end devices. That means the nodes not only generate traffic but also forward it to other nodes of the network. For this reason protocols like AODV are used. When communicating over the wireless links, there always exists the threat of packets being dropped fully or partially. This comes from nodes within the network or outside the network. The most common types of attack are black hole and gray hole attacks. The malicious nodes have certain characteristics that are an anomaly to the regular behavior of the MANET nodes.

IV. Behaviour of malicious nodes

Nodes in MANET need a communication protocol like AODV to successfully communicate. The AODV protocol has three main types of messages RREQ, RREP and RERR messages. The regular behaviour of MANET nodes is to reply to the RREQ packets using RREP packets if a route exists to a particular destination node. But the malicious nodes will try to attract the senders by manipulating its behaviour. It does so by replying using RREP to the RREQ packets from route requesting nodes even if there is no route to the destination node in question. This is generally done as follows

- Replying to RREQ packets using a higher sequence number
- Replying to RREQ packets using a very small hop count value to the requested destination node.

V. Modelling of Black Hole attack

In the present technique to identify a malicious node which carries the black hole attack we first model its behaviour using NETA framework. NETA is a INET 2.1.0 based framework. INET provides realistic implementations of structures and protocols on each protocol layers, especially

in the higher ones: network, transport and application layers. For example, INET develops Mobile Ad-hoc Networks (MANETs) routing protocols as AODV, DSR, OLSR, etc., and several TCP implementations as well as several applications. Thanks to the versatile and extensible schema of NETA, a huge amount of attacks could be implemented in each protocol layer. It is a framework built on top of OMNET++ simulation framework that specifically can be used to model various attacks in MANET. It includes a network attack controller module that facilitates the injection variables that alter the parameters of the underlying routing protocol that transform the behavior characteristics of a normal node to a malicious node. In our case the underlying routing protocol is AODV.

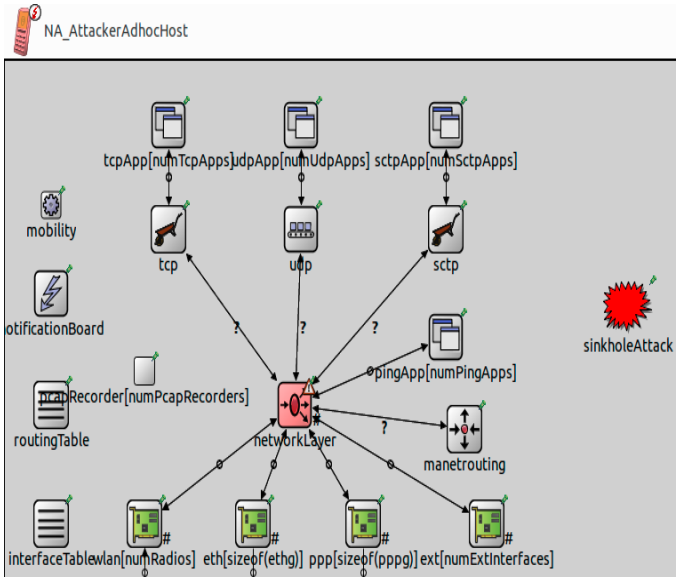

Figure 3. Node architecture for simulating Black Hole attack

VI. Identification of features to classify the BLACK HOLE attack

To accurately model the behaviour of the malicious node which carries the black hole attack we need to look from various perspectives. Apart from the two important parameters identified above as the main features of a malicious node carrying a black hole attack, we try to include more parameters to increase the efficiency of identifying mechanism. Those are as follows

- Evaluating the trust of the forwarding node
- Requesting a route for a non-existent destination node

In our simulations we evaluate the trust of the forwarding node to be eligible as a genuine forwarder. We suggest querying the joint neighbours of the forwarding node about the packets that were successfully forwarded to their requested destinations. This query can be implemented as part of the hello messages that are used for neighbour discovery in AODV protocol. The other way round to evaluate the trust is the number of times the malicious node reports the route error. It was found that the regular behaviour of a node carrying black hole attack is to avoid report a route error. In AODV protocol, when a destination node moves to a new location and becomes unreachable to the nodes forwarding the traffic to it, it needs to be reported to the originator node. The malicious node for black hole attack will not report any route error. Hence a regular node will have a non-negative value for the RERR messages it forwarded to the originator nodes. We suggest to query the history of the route error reporting of the malicious node.

Table 1. Sample values for features selected to detect Black-Hole attack

Sequence Number from neighbor	Hop Count	Number of RERRs reported	State
10001	5	3	Normal node
10066	1	0	Black hole node
10001	8	5	Normal node
10001	3	1	Normal node
10047	1	0	Black hole node
10001	4	1	Normal node
10001	9	1	Normal node
10001	7	1	Normal node
10001	8	0	Normal node
10021	5	0	Normal node
10001	1	0	Normal node

VII. Creating the dataset using the messages generated by AODV protocol

By using the data above for a large number of nodes that are part of the MANET which possess both normal behavior and malicious behavior we collect the above said features for nodes. This dataset is used to train the j 48 algorithm for classifying nodes into either malicious nodes or normal nodes

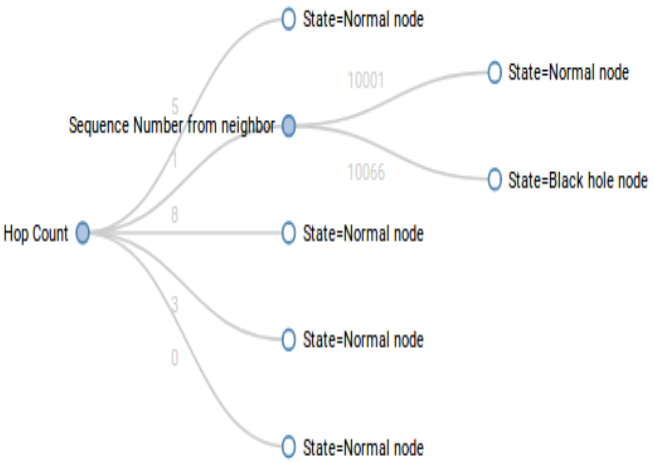

Figure 4. Decision tree using features selected for attack classification

```
void NS_CLASS handleMessageFromAttackController(cMessage *msg){

  // It is necessary to call Enter_Method for doing context switching (4.10 of User Manual)
  Enter_Method("NA_AODVUU: handle message from attack controller");

  LOG << "NA_AODVUU: Received message: "<< msg->getFullName() << "\n";

  // BEGIN NA_SINKHOLE - sancale
  // Activate sinkhole
  if (not strcmp(msg->getFullName(), "sinkholeActivate")) {
      NA_SinkholeMessage *dmsg;
      dmsg = check_and_cast<NA_SinkholeMessage *>(msg);
      LOG << "--> Activating module NA_AODVUU for Sinkhole...\n";
      LOG << "    Sinkhole Probability received: "<< dmsg->getSinkholeAttackProbability() << "\n";
      LOG << "    Sink only when route in table: " << dmsg->getSinkOnlyWhenRouteInTable() << "\n";
      //Now sinkhole attack is activated in this module
      sinkholeAttackIsActive = true;
      sinkholeAttackProbability = dmsg->getSinkholeAttackProbability();
      sinkOnlyWhenRouteInTable = dmsg->getSinkOnlyWhenRouteInTable();
      seqnoAdded = dmsg->getSeqnoAdded();
      numHops = dmsg->getNumHops();
      delete(msg);
```

Figure 5. The implementation of the Attack configuration module

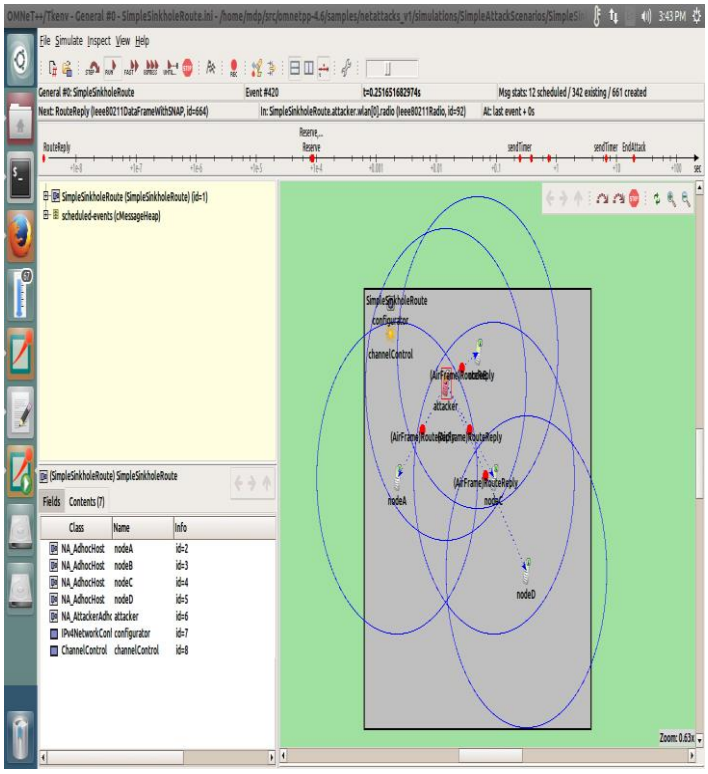

Figure 6. Simulation of Black Hole attack using OMNET++ simulation framework

VIII. Performance Evaluation

The performance of APEATOVC algorithm compared with that of Greedy Surrogate Selection and PEATOVC schemes. The details of these are given below.

```
##########################################################
#       Parameters for the Attack (Sinkhole No Route)    #
##########################################################
# SINKHOLE ATTACK
**.attacker.sinkholeAttack.active = true
**.attacker.sinkholeAttack.startTime = 0s
**.attacker.sinkholeAttack.endTime = 20s
**.attacker.sinkholeAttack.sinkOnlyWhenRouteInTable = true
**.attacker.sinkholeAttack.sinkholeAttackProbability = 1
**.attacker.sinkholeAttack.seqnoAdded = uniform(50, 60)
**.attacker.sinkholeAttack.numHops = 1
```

Figure 7. Simulation parameters used in the experiment

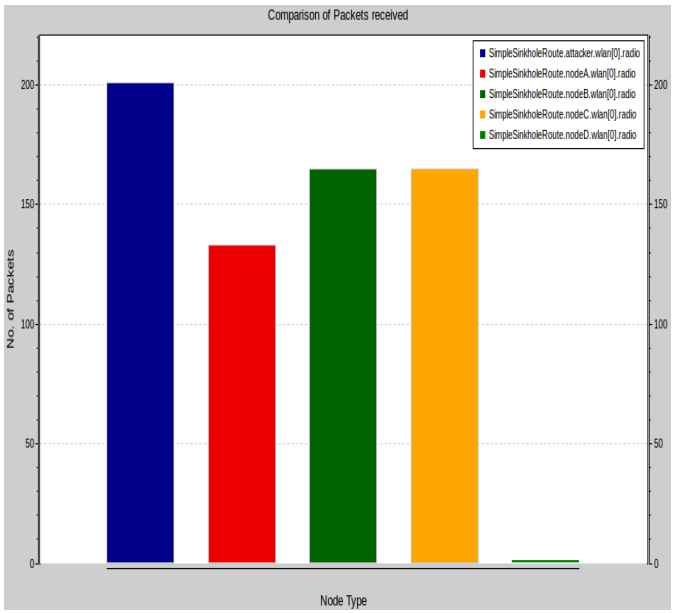

Figure 8. Number of Packets received by MANET nodes under Black-Hole attack

Table 2. Results of applying the J48 Classification algorithm to detect the Black hole attack

Correctly Classified Instances	456	99.13%
Incorrectly Classified Instances	4	0.87%
Kappa statistic		0.9703
Mean absolute error		0.0155
Root mean squared error		0.0927
Relative absolute error		5.21%

Root relative squared error		24.00%
Total Number of Instances	460	

Table 3. Comparison of algorithms used to detect the Black hole attack

Algorithm used for Back hole detection	Accuracy
Ad Hoc On-Demand Distance Vector (AODV)	87.63%
Local Intrusion Detection (LID)-ADOV	89.35%
Hybridization of Particle Swarm Optimization-Genetic Algorithm (HPSO-GA)	95.13%
DDBG	98.15%
NDJ48	99.13%

IX. Conclusion

An efficient technique which combines machine learning and network dynamism is presented in this work. It is evaluated using widely used simulation tool. Its effectiveness is evaluated by modelling the Black hole attack when using AODV a routing protocol in MANET. We envision to apply Support Vector Machine Technique to study the machine learning approach in MANET Intrusion Detection System.

REFERENCES

[1] P. Garcia-Teodoro, J. Diaz-Verdejo, G. MaciaFernandez, and E. Vazquez, "Anomaly-basednetwork intrusion detection: Techniques, systems and challenges," Computers and Security, vol. 28, no. 12, pp. 18-28, 2009.

[2] D. Hadiosmanovi, L. Simionato, D. Bolzoni, E.Zambon, and S. Etalle, "N-Gram against the machine: on the feasibility of the n-gram network analysis for binary protocols," In Research in Attacks, Intrusions, and Defenses, 2012, pp. 354-373.

[3] James P. Anderson, "Computer Security Threat Monitoring and Surveillance," Technical report, James P. Anderson Co., Fort Washington, Feb 1980.

[4] Shai Rubin, Somesh Jha, and Barton P. Miller, "Protomatching Network Traffic for High Throughput Network Intrusion Detection," In the Proceedings of the 13th ACM conference on Computer and Communications Security, pages 47-58. ACM, 2006.

[5] Meng Jianliang, and Shang Haikun, "The application on intrusion detection based on K-Means cluster algorithm," International Forum on Information Technology and Application, 2009.

[6] Mohammadreza Ektefa, Sara Memar, Fatimah Sidi, and Lilly Suriani Affendey, "Intrusion Detection Using Data Mining Techniques," In the proceedings of IEEE International Conference on Information Retrieval & Knowledge Management, Exploring Invisible World, CAMP10, 2010, pp. 200-203.

[7] Zubair A. Baig, Abdulrhman S. Shaheen, and Radwan AbdelAal, "An AODE-based Intrusion Detection System for Computer Networks," World Congress on Internet Security (WorldCIS), pp. 28-35, IEEE 2011.

[8] Yogendra Kumar Jain and Upendra, "An Efficient Intrusion Detection Based on Decision Tree Classifier Using Feature Reduction," International Journal of Scientific and Research Publications, vol. 2, issue 1, ISSN 2250-3153, Jan. 2012

[9] Rangadurai Karthick R., Hattiwale V.P., and Ravindran B., "Adaptive network intrusion detection system using a hybrid approach," 4th International Conference on Communication Systems and Networks (COMSNETS), vol.1, no. 7, pp. 3-7, Jan. 2012

[10] V Jaiganesh and P Sumathi, "Kernelized Extreme Learning Machine with Levenberg-Marquardt Learning Approach towards Intrusion Detection," International Journal of Computer Applications, vol. 54, pp. 38-44, September 2012.

[11] Gholam Reza Zargar, and Tania Baghaie, "Category-Based Intrusion Detection Using PCA," Journal of Information