IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

Hybrid Network Intrusion Detection for DoS Attacks

S.Anu[1], K. Pradeep Mohan Kumar[2]

M.Tech Student, Department of CSE, Periyar Maniammai University, Thanjavur, India[1]

Research Scholar, Department of CSE, Periyar Maniammai University, Thanjavur, India[2]

Abstract: The growing use of computer networks, security has become a main challenging task for computer users. It is basically suffering from Denial of Service (DoS) attacks. DoS attack is an attack that sending large number of network traffic towards the Victim server of the organization that bring down the performance of networks. Existing techniques like firewall, access control and encryption mechanism is vulnerable to provide sufficient protection against Virus, warms and DoS attacks,etc.,. So we need new mechanism to monitoring computer systems and network traffic to identify any deviation of the original user behavior. This paper proposes a new hybrid based IDS model for DoS attacks and evaluate its performance based on Particle Swarm Optimization combine with Support Vector machine (PSO-SVM) to increase detection accuracy and reducing false alarms. Here, feature selection is one of the important processes to increase the classification accuracy of this model. So, Particle Swarm Optimization (PSO) is used for selecting necessary feature from the PMU 2015 dataset that will improve the performance of the proposed model. The proposed work was implemented in Mat lab7. 2. The result shows that the proposed hybrid IDS has high detection accuracy (99. 25%) and (0. 75%) of false alarms.

Keywords: Intrusion Detection System (IDS), Network Security, Denial of Service (DoS), PSO-SVM.

INTRODUCTION

Intrusion detection system (IDS) is a model that mainly monitors computer networks and systems for identifying malicious user activity or policy violations and generate report to a system administrator to take necessary action. IDS mainly classify in to two types:

Host Based IDS (HBIDS):
Host Intrusion Detection Systems (HIDS) run on individual hosts or devices on the network. A HIDS monitors the inbound and outbound packets from the device only and will alert the user or administrator if suspicious activity is detected. It takes a snapshot of existing system files and matches it to the previous snapshot. If the critical system files were modified or deleted, an alert is sent to the administrator to investigate.

Network Based IDS (NBIDS):
NIDS is strategically positioned at various points on the network to monitor traffic going to and from network devices. NIDS solutions offer sophisticated, real-time intrusion detection capabilities often consisting of an assembly of interoperating pieces: a standalone appliance, hardware sensors, and software components are typical components that make up an NIDS. Further NIDS is classified in to two types:

(i)Signature based IDS:
A signature based IDS having the well known malicious pattern of data packet. It mainly monitors network traffic for identifying suspicious patterns present in data packets or not. By using a database of well known intrusion types and their data patterns, a signature-based NIDS can quickly identify intrusions and initiate an appropriate course of action taken by system administrator.

(ii)Anomaly based IDS:
Anomaly-based IDS having well known normal user activity of the system. It monitors network traffic for identifying whether the user activity is normal or abnormal. Extracted pattern compare with set of normal user activity if any deviation occur it consider as a attack otherwise it consider as a normal user.

Denial-of-service (DoS) is an attack that makes a Victim resource unavailable to its respective users, such as to temporarily or indefinitely interrupt or suspend services of a system connected to the Internet. So, we need new model like Computational Intelligence is ready to encounter this kind of threads for building intelligent security system that detect abnormal activities of the users and discard it. [1].

Existing security systems are evaluated with DARPA 98 and KDDCup99 datasets that affects the performance of system, this dataset only having old pattern of the attacks that degrade the performance of security systems. So, new Hybrid based IDS is the more power full security system to detect the various types of new DoS attacks pattern with higher detection accuracy, reducing false alarms with the help of new dataset.

RELATED WORK

In the field of DoS and DDoS attack, many researchers have been done up to now. Many of them consisted of traditional approaches such as firewall and etc, which was not met all needs of a robust detection system. There upon, researchers attended to the artificial intelligence and data mining techniques. Iftikhar et al. at [4] introduced feature selection mechanism based on Principal Component

Analysis (PCA). However, since this method might ignore some sensitive features, a method was proposed based on Genetic Algorithm and multilayer perceptron (MLP) - The neural network algorithm for mapping input to appropriate output. KDD-cup was used for dataset. As a result, they selected 12 features among 44 features and claimed that accuracy has improved to 99%. As mentioned at [5, 6], PCA is not suitable for large dataset and this method is executable just for small dataset.

In [7] Singh and Silakari stated that PCA is not proper solution for non-linear dataset, therefore they presented an algorithm based on Generalized Discriminant Analysis (GDA), to generate small size of features and improve classification operation. They asserted that this method is premier than other classification method such as SelfOrganizing Map (SOM) and C4.5. KDDCup99 was used for dataset in that research, also 4 different attacks were reviewed: DoS attack, User to Root Attacks, Remote to Local (User) Attacks, and probing. Finally their method accuracy was about 0.98.

Most of the researches in the scope of intrusion detection attack, offer the model by analyzing raw packet data, and processing vast amount of data especially while occurring DoS/DDoS attack is the main challenge for researchers. For this reason, the idea of attack detection based on statistical data gained from network management protocol was raised. MAID [8] was an intrusion detection system that monitored 27 different SNMP MIB variables and compared the behavior of normal and attack packet. Normal behavior of packet was modeled using probability density function (PDF), and was kept as reference PDF. They compared five similarity metrics by examining algorithm on actual network data and attack. They stated that KST is able to detect more attacks in all situations even at low traffic intensities.

D.Dutta and K.Choudhury at [9] claimed that their research was the first intrusion detection system which was integrated Digital Signature of Network Segment (DSNS) with Particle Swarm Optimization (PSO). They also benefited SVM to optimize clustering operation and better centroids selection. PSO [10] is a Swarm Intelligence algorithm, which despite the high Efficiency has low computational complexity.

PROPOSED PSO-SVM IDS MODEL

The architecture of the proposed PSO-SVM model is shown in Fig.1. The architecture contains two phases (1) Training phase (ii) Testing phase. In training phase, the KDDCUP 99 and PMU 2015 datasets are used for data pre-processing, feature selection which is done using Particle Swarm Optimization (PSO) and classification using SVM is implemented by which DoS attack patterns are identified. Second stage is testing stage where the captured traffic is evaluated as in training stage, pattern is identified, matched with the stored DoS patterns in database and decisions will be taken accordingly. New patterns identified by analyzing the behavior of the traffic, if found against the legitimate traffic, that pattern is captured and updated in the database. The implementation of PSO-SVM based IDS has three phases that includes Preprocessing, Feature Selection, and Classifier.

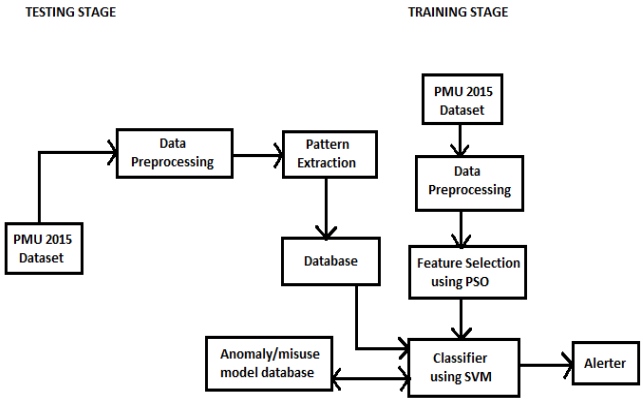

Fig. 1 Architecture of PSO-SVM based IDS

A. Data Preprocessing

Classification of malicious and normal data would not be effective if PMU2015 dataset are processed in its current format. Hence it is required to preprocess the data before SVM classification system is built.

The following activities are involved in data preprocessing:

- Mapping symbolic features to numeric value.
- Implementing scaling since the data has significantly varying resolution and ranges. The attribute data are scaled to fall within the range [-1, 1].
- Attack names were mapped to one of the five classes, 0 for Normal, 1 for DoS (Denial of Service), 2 for U2R (user-to-root: unauthorized access to root privileges), 3 for R2L (remote-to-local: unauthorized access to local from a remote machine) and 4 for Probe (probing: information gathering attacks) .

Data Normalization

When processing instances whose different features are on different scales, it causes bias towards some features over other ones. To solve this problem, these raw data sets need to be normalized PMU 2015 comprises of 113 features can be divided into 4 categories such as 'Boolean', 'String', 'Count', and 'Rate'. There are many methods for data normalization like min-max normalization, z-score normalization and normalization by decimal scaling. Here Min-max normalization performs a linear transformation on the original data shown in the eqn(1.1). Suppose that min_a and max_a are the minimum and the maximum values for attribute A. Min-max normalization maps a value v of A to v' in the range [new-min_a , new- max_a] by computing

$$v' = ((v - min_a) / (max_a - min_a)) * (new - max_a - new\ min_a) + new - min_a \quad ...Eqn\ (1.1)$$

where, 'v' indicates the original value, min_a represents minimum value in that column, max_a represents maximum value in that column.

B. Feature selection based on PSO

Particle swarm optimization (PSO) is a population-based stochastic optimization technique. PSO simulates the

DOI 10.17148/IJARCCE.2016.53212

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

social behavior of organisms, such as bird flocking and fish schooling, to describe an automatically evolving system. In PSO, each single candidate solution is "an individual bird of the flock", that is, a particle in the search space. Each particle makes use of its individual memory and knowledge gained by the swarm as a whole to find the best solution (Venter 2002). All of the particles have fitness values, which are evaluated by fitness function to be optimized and have velocities which direct the movement of the particles. During movement, each particle adjusts its position according to its own experience, as well as according to the experience of a neighboring particle and makes use of the best position encountered by itself and its neighbor. The particles move through the problem space by following a current of optimum particles.

The initial swarm is generally created in such a way that the population of the particles is distributed randomly over the search space. At every iteration, each particle is updated by following two "best" values, called p_{best} and g_{best}. Each particle keeps track of its coordinates in the problem space, which are associated with the best solution (fitness) the particle has achieved so far. This fitness value is stored and called p_{best}. When a particle takes the whole population as its topological neighbor, the best value is a global "best" value and is called g_{best}.

The pseudo code of the PSO procedure is given below.

Initialize population
While (number of generations, or the stopping criterion is not met)
For p = 1 to number of particles
If the fitness of X_p is greater than the fitness of p^{best}_p
then Update $p^{best}_p = X_p$
For k ∈ Neighborhood of X_p
If the fitness of X_k is greater than that of g^{best}
then Update $g^{best} = X_k$
Next k
For each dimension d
$V^{new}_{pd} = w \times v^{old}_{pd} + c_1 \times rand_1 \times (p^{best}_{pd} - \chi^{old}_{pd}) + c_2 \times rand_2 \times (g^{best}_{d} - \chi^{old}_{pd})$
if $v_{pd} \epsilon (v_{min}, v_{max})$ then
$V_{pd} = max (min(v_{max}, v_{pd}) v_{min})$
$\chi_{pd} = \chi_{pd} + v_{pd}$
Next d
Next p
Next generation until stopping criterion

V^{new}_{pd} and v^{old}_{pd} are the particle velocities, x^{old}_{pd} is the current particle position (solution) and x^{new}_{pd} is the updated particle position (solution). The values $pbest_{pd}$ and $gbest_d$ are defined as stated above.

The two factors $rand_1$ and $rand_2$ are random numbers between (0, 1), whereas c_1 and c_2 are acceleration factors, usually $c_1 = c_2 = 2$. Particle velocities of each dimension are tried to a maximum velocity V_{max}. If the sum of velocities causes the total velocity of that dimension to exceed V_{max}, then the velocity of that dimension is limited to V_{max}. V_{max} is a user-specified parameter.

Based on the rules of particle swarm optimization, we set the required particle number first, and then the initial coding alphabetic string for each particle is randomly produced. In our case we coded each particle to imitate a chromosome in a genetic algorithm; each particle was coded to a binary alphabetic string $S = F_1 F_{2.....} F_n$, n =1, 2, …, m ; the bit value {1} represents a selected feature, whereas the bit value {0} represents a non-selected feature. The adaptive functional values were data based on the particle features representing the feature dimension; this data was classified by a support vector machine (SVM) to obtain classification accuracy; the SVM serves as an evaluator of the PSO fitness function. For example, 113 feature of PMU2015 dataset (n=113 Sn = $F_1F_2F_3F_4F_5F_6F_{7.....}F_{111}F_{112}F_{113}$) is analyzed using PSO, 12 numbers of features are selected, Sn= ($F_1F_{13}F_{24}$ $F_{35}F_{47}F_{59}F_{60}F_{71}F_{89}F_{90}F_{107}F_{111}$).

The selected features from PMU2015 datasets are trained and tested with the proposed PSO-SVM model is shown in the Table I & II.

TABLE I: Selected feature of PMU2015 dataset

S. No	Selected Features	Description
1	Source IP	Source IP Address
2	Destination IP	Destination IP Address
3	Source byte	Sender byte
4	Destination byte	Destination byte
5	Protocol	http,telnet….
6	SYN count	Number of SYN segments observed (including rtx)
7	FIN count	Number of FIN segments observed (including rtx)
8	Count	No of connections to the same Destination
9	srv_count	Number of connections to the same service as the current connection in the past two seconds
10	same_srv_rate	Number of connections to the same service
11	diff_srv_rate	Number of connections to different services
12	dst_,host_same _src_port_rate	Number of connections to the current host having the same srcport

TABLE II: Rule structure of PMU2015 dataset

S. No	Attack Description	Attack Type
1	protocol = tcp, source Ip=172.20.62.33, Dest Ip=172.20.62.255, 178>src_byte<322, 10>Dst_byte<224, SYN count=1 or 2, 1>count<28, 1>srv_count<28, same_srv rate=1, diff_srv rate=0, 0>dst_,host_same_src_port_rate<1 FIN=1.	Normal
2	protocol=tcp, source Ip=172.20.62.33, Dest Ip=172.20.62.255, src_byte= 0, Dst_byte=0, SYN bit=3 to 160, 3>count<160, 3>srv_count<160, same_srv rate=0, 0>diff_srv rate<1,0>dst_, host_same_src_port_rate<1, FIN bit=0.	Neptune
3	protocol = UDP, source Ip=172.20.62.3, Dest Ip=172.20.62.25, src_byte= 221120, Dst_byte=0, SYN bit=238 to 512, 238>count<512, 238>srv_count<512, same_srv rate=1, 0>diff_srv rate<2, 0<dst_,host_same_src_port_rate<2 , FIN bit=0.	Smurf

DOI 10.17148/IJARCCE.2016.53212

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

C. Classification using PSO-SVM Approach

Support Vector Machines (SVMs) are a group of supervised learning methods that can be applied to classification or regression. Theoretically, SVM is a well-motivated algorithm that is based on the statistical learning proposed by Vapnik [87]. SVM has shown promising empirical good performance and successful application in many fields such as bioinformatics, text categorization, speaker verification, handwritten digit recognition, face detection, engineering and science, financial market evaluation, pattern recognition [88], image recognition [89] and many more. SVM has received overwhelming attention as a classification technique from diverse research communities due to its suitability, and works very well with high-dimensional data as well as avoiding the curse of the dimensionality problem .SVM was originally designed for binary classification in order to construct an optimal or maximal hyperplane so that the margin of separation between the negative and positive dataset will be maximised.

There are many hyperplanes that might classify the data. The best hyperplane is chosen based on the one that represents the largest separation, or margin, between the two classes. So we chose the hyperplane so that the distance from it to the nearest data point on each side is maximized. An SVM maps linear algorithms into non-linear space. It uses a feature called, kernel function, for this mapping. Kernel functions like polynomial, radial basis function are used to divide the feature space by constructing a hyperplane. The kernel functions can be used at the time of training of the classifiers which selects support vectors along the surface of this function. SVM classify data by using these support vectors that outline the hyperplane in the feature space. This process will involve a quadratic programming problem, and this will get a global optimal solution. Suppose we have N training data points {(x1, y1), (x2,y2), (x3, y3), ..., (xN , yN)}, where xi € Rd and yi €{+1,−1}. Consider a hyper-plane defined by (w, b), where w is a weight vector and b is a bias. The classification of a new object x is done with

$$f(x) = \text{sign}(w.x+b)=\text{sign}\left(\sum_i^N \alpha_i y_i (x_i, x) + b\right) \quad \ldots\ldots\ldots\ldots\ldots \text{Eqn (1.2)}$$

The training vectors x_i occurs only in the form of a dot product. For each training point, there is a Lagrangian multiplier α_i. The Lagrangian multiplier values α_i reflect the importance of each data point. When the maximal margin hyper-plane is found, only points that lie closest to the hyper-plane will have $\alpha_i > 0$ and these points are called support vectors. All other points will have $\alpha_i = 0$. That means only those points that lie closest to the hyper plane, give the representation of the hypothesis/classifier.

1) Create an initial enhanced population of chromosomes, i.e. a group of individuals with different chromosomes generated by PSO. Each individual chromosome consists of twelve parameters for PMU2015 dataset with different features namely Protocol, Source Ip, Dest Ip, src_byte,.dst_bytes, count (No of connections to the same Dest), SYNcount, srv_count, same_srvrate, diff_srvrate, dst_,host_same_src_port_rate, FINcount. The pattern weight of the initial population should be determined properly by user to include as many possible solutions as possible.

2) Calculate the fitness value of each individual in the initial population using Eqn (1.3) and rank them according to their fitness value.

$$F(X) = \sum_{i=1}^{n} Wi\, Xi + b \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{Eqn (1.3)}$$

Where *W* is the weight vector, *b* is a bias value, and *n* is the number of features. They ranked each feature depending on the value of its weight. The features with large weight values are considered to be the features of the greatest effect (important features) and are used for the detection process.

SIMULATION RESULTS AND DISCUSSIONS

We used a AMD Athlon™ 64 X2 Dual Core Processor 5000+ 2.59G MHz computer with 512MB RAM, and implemented on a Windows XP Professional operating system. The proposed PSO-SVM model was implemented in Mat Lab 7.2. During the evaluation,8.5 percent labeled data of PMU2015 datasets were used for training the proposed PSO-SVM. This PMU 2015 dataset contains three types of traffics and two types of DoS attacks about 410 MB and each traffic record has 113 features. It contains a total of 20,00,000 records constituting a size of 410 MB out of which 15,13,000 records are of Neptune type, 4,00,219 records are of Smurf type and the balance 86,781 records belong to the normal category. 42380 normal, 38523 Smurf, 90000 Neptune are the traffic records considered for training the proposed IDS. 35400 normal, 29549 Smurf, 79000 Neptune are the traffic records considered for testing the proposed IDS.

The detection accuracy of the proposed PSO - SVM model is measured by the precision, recall and F - measure; which is calculated based on the confusion matrix given in Table III. The criteria explained below holds good for both KDDCUP 99 and PMU 2015 dataset.

TABLE III : Confusion Matrix

	Predictor class	
	Normal	Attack
Normal	True Positives (TP)	False Positives (FP)
Attack	False Negatives (FN)	True Negatives (TN)

Where,

- True positives (TP): indicates the numbers of normal events are successfully labeled as normal.
- False positives (FP): refer to the number of normal events being predicted as attacks.
- False negatives (FN): number of attack events is incorrectly predicted as normal.
- True negatives (TN): numbers of attack events are correctly predicted as attack.

The performance metrics calculated from these are:

$$\text{Recall} = \frac{TP}{TP+FN} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{Eqn (1.4)}$$

DOI 10.17148/IJARCCE.2016.53212

$$\text{Precision} = \frac{TP}{TP+FP} \quad \text{.....Eqn (1.5)}$$

An IDS should achieve a high recall without loss of precision, where F-measure is a weighted mean that assesses the trade-off between them.

$$\text{F- Measure} = \frac{2*\text{Recall}*\text{Precision}}{\text{Recall}+\text{Precision}} \quad \text{.....Eqn (1.6)}$$

$$\text{Overall accuracy} = \frac{TP+TN}{TP+TN+FN+FP} \quad \text{.....Eqn (1.7)}$$

The number of records taken for testing and training phase along with the testing and training time with PMU 2015 dataset is given in Table IV.

TABLE IV: Training and Testing data from PMU2015 dataset

Data Type	Training Data	Training time (s)	Test data	Testing time (s)
Normal	42380	0.9	35400	0.78
Smurf	38523	0.73	28549	0.57
Neptune	90000	0.98	79000	0.87

TABLE V: Result Discussions when tested with PMU 2015 Dataset

METRICS (%)	DATA TYPE		
	Normal	Smurf	Neptune
Precision	99.17	99.1	99.36
Recall1	100.00	100.00	100.00
F-measure	99.81	99.87	99.76
Accuracy	99.17	99.15	99.44
Overall detection accuracy of PSO-SVM system when tested with PMU 2015 dataset			99.25%

By analyzing the results on simulation with PMU 2015 dataset as explained above, the overall performance of the proposed system is improved significantly and it achieves 99.25% accuracy for all types of attacks.

CONCLUSION

Intrusion Detection is a process of detecting intruders in a computer system in order to increase the security. Intrusion detection is an area in which more and more sensitive data are stored and processed in networked system. A series of experiments were conducted on the proposed PSO-SVM model with PMU2015 datasets to examine the effectiveness of our feature selection and its parameters in building effective IDS. We have used confusion matrices for evaluation of our proposed technique and the results are obtained on the basis of evaluation metrics namely precision, F-measure, recall and finally accuracy. The experiment results show that our approach is not only able to achieve the process of selecting important features but also to yield high detection rates for IDS. From the experiments, detection accuracy of PSO-SVM is 99.25 % and 0.75% false alarm rate against PMU 2015 Dataset. The results show that the proposed hybrid approach will outperform the existing individual approaches by reducing the number of features and increasing the detection rates.

REFERENCES

[1] J. Mirkovic and P. Reiher, "A taxonomy of DDoS attack and DDoS defense mechanisms,"ACM SIGCOMM Computer Communication Review, vol. 34, pp. 39-53, 2004.
[2] J. Ding, Advances in network management: CRC press, 2010.
[3] T. Pang-Ning, M. Steinbach, and V. Kumar,"Introduction to data mining," in Library of Congress, 2006.
[4] A. Iftikhar, A. Azween, A. Abdullah, A. Khalid, and H. Muhammad, "Intrusion detection using feature subset selection based on MLP," Scientific Research and Essays, vol.6, pp. 6804-6810, 2011.
[5] H. M. Imran, A. Abdullah, M. Hussain, S. Palaniappan, and I. Ahmad, "Intrusion Detection based on Optimum Features Subset and Efficient Dataset Selection," International Journal of Engineering and Innovative Technology (IJEIT), vol. 2, pp. 265-270, 2012.
[6] K. Delac, M. Grgic, and S. Grgic, "Independent comparative study of PCA, ICA, and LDA on the FERET data set," International Journal of Imaging Systems and Technology, vol. 15, pp. 252-260, 2005.
[7] S. Singh, S. Silakari, and R. Patel, "An efficient feature reduction technique for intrusion detection system," in Machine Learning and Computing, 2009. International Conference on, 2011.
[8] J. Li and C. Manikopoulos, "Early statistical anomaly intrusion detection of DOS attacks using MIB traffic parameters," in Information Assurance Workshop, 2003. IEEE Systems, Man and Cybernetics Society, 2003, pp. 53-59.
[9] D. Dutta and K. Choudhury, "Network Anomaly Detection using PSO-ANN,"International Journal of Computer Applications, vol. 77, 2013.
[10] J. Kennedy, "Particle swarm optimization," in Encyclopedia of Machine Learning, ed: Springer, 2010, pp. 760-766.
[11] J. Yu, H. Lee, M.-S. Kim, and D. Park, "Traffic flooding attack detection with SNMP MIB using SVM," Computer Communications, vol. 31, pp. 4212-4219, 2008.
[12] J. Yu, H. Kang, D. Park, H.-C. Bang, and D. W. Kang, "An in-depth analysis on traffic flooding attacks detection and system using data mining techniques," Journal of Systems Architecture, vol. 59, pp. 1005-1012, 2013.
[13] J. B. Cabrera, L. Lewis, X. Qin, W. Lee, and R. K. Mehra, "Proactive intrusion detection and distributed denial of service attacks—a case study in security management," Journal of Network and Systems Management, vol. 10, pp. 225-254, 2002.
[14] http://www.hping.org/hping3.html, July 2014
[15] http://www.cs.waikato.ac.nz/ml, July 2014

DOI 10.17148/IJARCCE.2016.53212