❏   545

Intelligent Packet Delivery in Router Using Structure Optimized Neural Network

R. Deebalakshmi*[1], V.L. Jyothi[2]
[1]Sathyabama University, Chennai, India
[2]CSE, Jeppiaar Engineering College, Chennai, India

Article Info	ABSTRACT
	The Internet itself is a worldwide network connecting millions of computers and less significant networks. Computers communicated by routers. Crucial the role of a router is to our technique of communicating and computing. Routers are situated at gateways, the spaces where two or more networks connect, and are the decisive device that keeps data flow between networks and keeps the networks connected to the Internet. When data is sent between places on one network or from one network to a second network the data is always seen and intended for to the proper place by the router. The router carries out this by using headers and routing tables to establish the best path for routing the data packets. This trim down the effectiveness of edge router only when the path engaged, it will enhanced by classification method, predictable classification methods like port based ,deep packet inspection and statistical classification are give less precision. In this system structured optimized neural network is used for more precise organization. Classification output forwarded to router dynamically for intellectual packet delivery. The method will improve router competence by greater than before throughput and decreased latency.

Corresponding Author:

R. Deebalakshmi,
Research Scholar,
Sathyabama University, Jeppiar Nagar, Chennai, Tamilnadu, Indian-600119.
Email: deepar11@gmail.com

1. INTRODUCTION

Lakhina et al. [1] explained the distributions of packet features observed in flow a trace which reveals both the presence and the structure of a wide range of anomalies. Using entropy a summarization tool that the method expressed that the analysis of feature distributions leads to significant advances using feature distributions, anomalies naturally fall into distinct and meaningful clusters. Estan et al. [2] introduced scalable algorithms for identifying the large flows: sample and hold and multistage filters, which take a constant number of memory references per packet and use a small amount of memory. The technique described optimizations such as early removal and conservative update that further improve the accuracy of algorithms, as measured on real traffic traces, by an order of magnitude. Peng el al. [3] expressed Energy efficient clustering using energy prediction mechanism. They have used different energy level sensors with different monitoring objects. For cluster head selection residual energy of nodes is predicted instead of broadcasting it in every round, also they used communication cost.

Nguyen et al. [4] designed K-Means clustering algorithm for packet separations. Packets traced in computer networks. From traced packets flow is defined, and flows are grouped into clusters. Flows classified based on k-cluster centers. The cluster centers are defined by packet length and packets arrival time of application. Applications which are having same packet length or nearby packet length are compared with existing cluster, if it is matched combined to existing cluster otherwise new cluster group if formed. In machine learning SVM it uses supervised learning algorithm and analyze data used for classification and

regression. If no trained set that method is called unsupervised algorithm, here input flows are grouped as clusters.

There are two types of SVM.

1. Linear SVM and
2. Non- linear SVM.

Linear SVM: In linear SVM, it has linear n-dimensional vector and if there is a possibility of separating n-1 hyper plane vector it is called a linear classifier. The hyper plane is used to classify data and optimal hyper plane will give largest separation or, Bad selection of hyper plane will give noise data classification. Distance of data compared with hyper plane so maximum value stetted for it.

Non-Linear SVM: In non-linear SVM, the classifier is created by applying the kernel trick to maximum-margin hyper planes. They are effective in high dimensional spaces. They are still effective in cases where number of dimensions is greater than the number of samples. It uses a subset of training points in the decision function and so it is memory efficient. SVM has some disadvantages. Particularly, if the number of features is much greater than the number of samples, the method is likely to give poor performance. They do not directly provide probability estimates, these are calculated using an expensive five-fold cross-validation.

ONE-CLASS SVM: Another alternative of SVM is one-class SVM. It is an unsupervised algorithm that learns a decision function for novelty detection that it classifies new data as similar or different to the training set. Basically, it also separates all the data points from the origin and maximizes the distance from this hyperplane to the origin. The optimal hyperplane is the one that represents the largest separation, or margin, between the two classes. If the number of samples is more than the creation of optimal boundary is difficult and it performance is also affected. It uses only unsupervised algorithm. Hence, it cannot identify new traffic.

Moore et al. [5] designed machine learning Naïve Bayes technique to categorize internet traffic based on application. The traffic in the internet applications were classified into different categories, e.g. mail, web services, p2p, multimedia and games. The authors used accuracy as a classification metric to evaluate performance of classifier. This result depicted that naïve bayes techniques have 65 % accuracy in classification. Two refinements were performed for improving classification accuracy using naïve bayes kernel estimation and fast correlation based filter method. It gives 95% as the overall accuracy.

Luigi et al. [6] designed SeLeCT technique, a Self-Learning Classifier for Internet Traffic. It uses unsupervised algorithm to automatically groups the flow into homogeneous group based on application. It doesn't require prior knowledge of training set to identify the traffic flows. Automatically the algorithm groups the flows into homogeneous clusters using statistical features. It also simplifies the label assignment by assigning labels to the cluster based on application. Furthermore, it uses self seeding approach to process next batches of flows before assigning labels to previous cluster. The author evaluated the performance of SeLeCT using different traffic traces collected from ISP located in the different continents. The experiments showed that it achieves overall accuracy .The accuracy is achieved is nearly 98% and it discover new protocols and application in the traffic traces.

Bujlow et al. [7] discussed, C5.0 is the decision tree based algorithm and use the concept of machine learning algorithm. It is easier to use and memory efficient. It generates the decision tree based on set of training set .The trees are used to classify the set of test cases. The c5.0 classifier uses command line interface to generate the rules for decision tree and test the classifier. The experiment was executed many times using different set of training flows and attributes like packet size, packet length, number of flows. The attributes for flows are used to construct test cases. It produced 98% accuracy when accurate testing and training data used. Karagiannis et al. [8] introduced a more general classification system that groups the flows using host and port information then individual flows, this system works on payload details, flows are trained by payload, it gives less accuracy, if port and host information is unavailable.

Roughan et al. [9] classified the traffic into an undersized number of types apt for Qos applications. They use system such as clustering using nearest neighbor to give the necessary arrangement. Some degree of small set of features and an implicit assumption of the accuracy of the testing and training data-sets; the authors restrict themselves large sets of network-based applications. Afzal et al. [10] compared the performance of two variations of back propagation learning algorithm (Adaptive learning rate with momentum and Resilient). Both the algorithms are experimented on a variety of classification problems in order to assess the efficiency of these two learning approaches. Experimental results reveal that during testing and training Resilient propagation algorithm outperforms back propagation with Adaptive learning rate and momentum.

Hemalatha et al. [11] expressed the privacy-preserving methods was vulnerable to Sybil attacks, whereby a malicious user can act as to be multiple vehicles. It ensured the authenticity of the messages propagated in VANET; a straight-forward process is utilized public keys authorized by a certification authority (CA) to sign the messages. Fadli et al. [12] developed an Ethernet over SDH (synchronous digital hierarchy) and MPLS-TP (Multiprotocol Label Switching - Transport Profile) with ring security to

predictable network failure and computed performance of network. Aman et al. [13] integrated context transport and multicast quick reroute, and developed this integration to the standard network mobility management. Hashim et al. [14] discussed quantitative analysis and it computed Proxy Mobile Internet Protocol Version 6 (PMIPv6) multicast speedy reroute operations for data traffic.

Router routes the packets from one network to other network through the best path. To route the packet router must know the following details target address, neighbor router from which it can learn about remote networks possible routes to all remote networks. The best route is applied to each remote network and how to maintain and verify routing information. The router learns about remote networks from neighbor router or from an administrator, the router then builds a routing table that describes how to find the remote network, if a network is directly connected then the router already knows how to get to it. If a network isn't directly connected to the router, the router use two ways to reach remote network one is static routing and other is dynamic routing.

The Network traffic classification is required mainly for many network management tasks. Flow prioritization Traffic shaping/policing diagnostic monitoring. Also the network operators can know knowledge of what applications are present in a network, network-operators are superior to plan, budget and bill. Network-operators and users able to identify abnormality in network traffic to reduce the impact of malicious behavior. Beside classification of network traffics are those that define to understand the present and varying Internet: the modeling of traffic-mix, user-composition and so on. In neural network classification depend without knowing site detail, which machine used, without interaction of user and administrator and without packet payload information like IP address and port detail.

This paper summarize as follows: Section 2 discussed about proposed method. Section 3 explores about implemented result and Discussion. Section 4 concludes the overall work with future plan of research work.

2. PROPOSED SYSTEM

Neural network is a conceptual model designed as a computational model based on human brain to solve certain kind of problems, it's a network of many neurons and performs intelligent behavior, and neural network has ability to learn an adaptive system which means it can change network structure based on information flowing through it. Adaptation is done by adjusting weights of network.
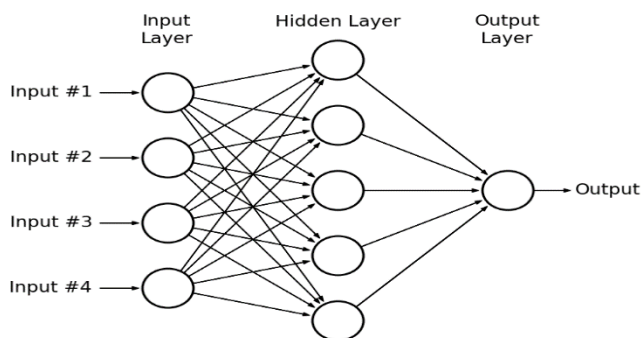

Figure 1. Sample Architecture of Neural Network

In Figure 1 circle represents the neurons and lines indicates connection between two neurons also it is a path for flow of information, each line has weight it's a number that control action between two neurons. If the neural network gives good output no need to adjust the weights, otherwise error generated by network. So the weights are adjusted to get good output. A perceptron contains one or more inputs or a processor and produce the single output, it follows the feed forward model which means inputs are sent to neurons are processed to get output.
Perceptron algorithm
Step 1: all inputs are multiplied by weights
Step 2: sum all the weighted inputs
Step 3: compute the output by sum passed through an activation function.

2.1. Back Propagation with Neural Network (BPNN)

The BPNN is a common method for training the artificial neural networks in conjunction with an optimization method, such as the gradient descent. The method calculates the gradient of a loss function with respects to all the weights in the network. The gradient is fed to the optimization method which in turn uses it to update the weights, in an attempt to minimize the loss function. Back propagation (BP) requires a known, desired output for each input value in order to calculate the loss function gradient. It is therefore usually considered to be a supervised learning method. Although, it is also used for some unsupervised network. It is a generalization of the delta rule to multi-layered feed forward networks, made possible by using the chain rule to iteratively compute gradients for each layer. The BP requires that the activation function used by the artificial neurons (or "nodes") be differentiable.

Given the number of nodes of the input layer, the hidden and the output network n, k, m, respectively, the total number of input samples is x_{pi}, which indicates that the P sample's the ith input value, v_{ki} indicates the ith node of input layer to the hidden layer of the kth node weight, ω_{jk} means the node weight from hidden layer ofthe k to the output layer of the j. For convenience, the threshold is included connection weights, and then the output of hidden layer node k is expressed in equation (1).

$$z_{pk} = f\left(net_{pk}\right) = f\left(\sum_{i=0}^{n} v_{ki} x_{pi}\right) \tag{1}$$

The output layer nodes for the node j: is expressed in equation (2).

$$y_{pj} = f\left(net_{pj}\right) = f\left(\sum_{i=0}^{n} w_{jk} z_{pk}\right) \tag{2}$$

where, standard sigmoid function is selected as incentive function are given equation(3).

$$f(x) = \frac{1}{e^{-x}} \tag{3}$$

The definition of global error functions can be expressed in equation (4).

$$E = \sum_{p=1}^{p} E_p = \frac{1}{2}\sum_{p=1}^{p}\sum_{j=1}^{m}(t_{pj} - y_{pj})^2 \tag{4}$$

where E_p, is error of the sample p, t_{pj} is the ideal output. The adjustment formulas of weights are as follows.
a. The weight adjustment formula of output layer neurons which mathematical expression are given in Equation(5)

$$\Delta\omega_{jk} = \eta\sum_{p=1}^{p}(t_{pj} - y_{pj}).y_{pj}(1 - y_{pj}).z_{pk} \tag{5}$$

where, η is the learning rate, whose general range is 0.1 - 0.3.

b. The weight adjustment formula of each hidden layer neurons are explained in equation(6)

$$\Delta v_{ki} = \eta\sum_{p=1}^{p}\left(\sum_{j=1}^{m}\delta_{pi}\omega_{jk}\right)z_{pk}(1 - z_{pk})x_{pi} \tag{6}$$

The basic idea of the BP algorithm is the learning process and can be divided into two stages: the first stage (forward propagation process), given input information through a layer by layer processing each hidden layer and calculate the actual output value of each unit of y_{pj}; the second stage (the reverse process), if the output layer fails to get the expected output value, then layer by layer recursively calculates the difference of error between the actual output and the expected output. The gradient descent method modifies the weights of the Δv_{ki}, $\Delta \omega_{jk}$, making the total error function minimum. Table 1 displays the input parameters of network traffic to evaluate the proposed techniques.

Table 1 Sample Input Parameters of Network Traffic

Classification	Example Application
DATABASE	postgres, sqlnet oracle, ingres
INTERACTIVE	ssh, klogin, rlogin, telnet
MAIL	imap, pop2/3, smtp
SERVICES	X11, dns, ident, ldap, ntp
ATTACK	Internet worm and virus attacks
GAMES	Microsoft Direct Play
P2P	KaZaA, BitTorrent, GnuTella
MULTIMEDIA	Windows Media Player, Real
BULK	ftp
WWW	www

2.2. Flow Features-Classification input

Flow metrics (duration, packet-count, total bytes) Packet inter-arrival time (mean, variance, 1st & 3rd quartiles, median, minimum, maximum. Size of TCP/IP control fields (mean, variance, 1st & 3rd quartiles, median, minimum, maximum,). Total packets (in each direction and total for flow) Payload size (mean, variance, 1st & 3rd quartiles, median, minimum, maximum,. . .). Effective Bandwidth Ranked list is top-ten fourier-transform components of packet inter-arrival times for each direction. Here, Table 2 explains the proposed classifier result of packet inter-arrival times in details.

Table 2 Proposeed Classifier Result of Packet Inter-Arrival Times

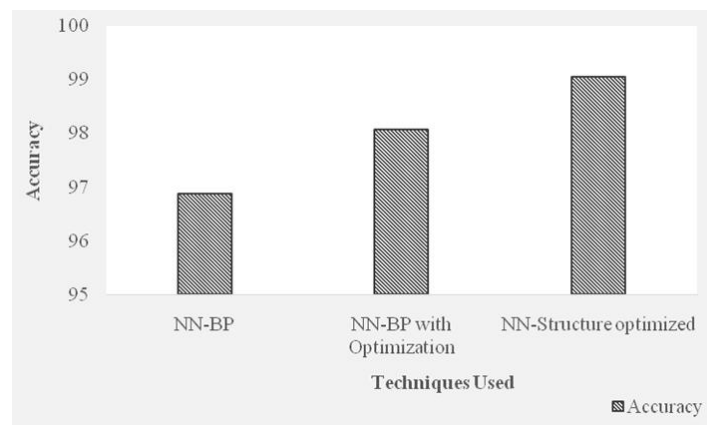| Algorithms	Accuracy
NN-BP	96.87
NN-BP with Optimization	98.07
NN-Structure optimized	99.05

Figure 2. Comparison Analysis of Neural Network Algorithms

Figure 2 shows the comparison of Neural Network Algorithms such as NN-BP, NN-BP with optimization and NN-structured optimization and it is clearly shown that the proposed mechanism NN-BP has higher accuracy of 99.05 estimated in terms of packet arrival times.

2.3. Important flow Routing

In general router route the packets from one device to other device which are in different network, for forwarding packets it refers the routing table, routing table contains details about nearby routers information. For important flow routing, some additional information is used to store in routing table. In routing table IP address along with packet count for particular application and packet time stamp is updated in routing table periodically. Flow measurement statistics report provides packet count for particular application. Router route packets intelligently using this classification output.
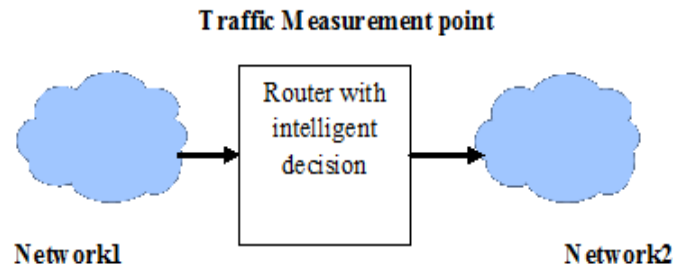
Traffic Measurement point

Figure 3. Important Flow Routing

In Figure 3, network1 wants to transfer the packet to network2. The router route the packet intelligently using traffic classification output. In router priority scheduling algorithm is used to schedule the packet, based on priority packets importance is given then it is forwarded to respective destination.
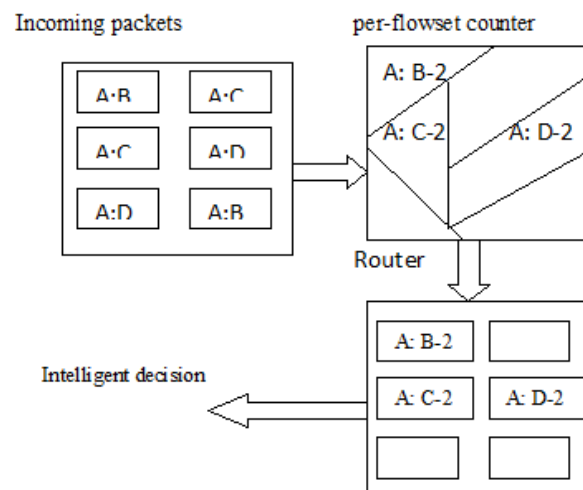

Figure 4. Routing Table Updatation

Router incoming packets are measured and classified based on categorization algorithm, flow are formed as flow-groups. This flow-groups information is maintained by counter, for each matching flow-groups counter is incremented; the application packet count along with the packet time stamp is updated in routing table periodically as shown in Figure 4. Router while route the packet check how to reach the destination address, and also knows the application detail, which packet belongs to which application and how many packets forwarded for particular application, based on the intelligent information route the packet intelligently. If some application packets counts are forwarded more and other application packets are just started for transmission, router follows the priority algorithm first priority is given for application which have more packets count, other application transmission is denied, this denial of service is given only three times for particular application using aging concept.

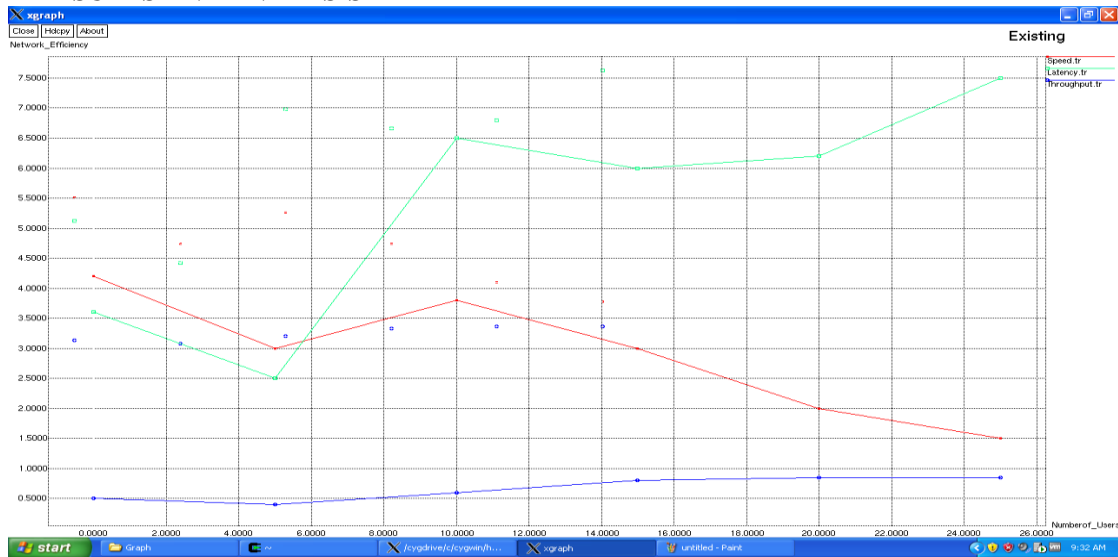## 3.   RESULTS AND ANALYSIS

Figure 5. Router performance without intelligence

Router performance without intelligence is shown in Figure 5. Here the parameters such as speed, latency and throughput are plotted for existing system with respect to their network efficiency.
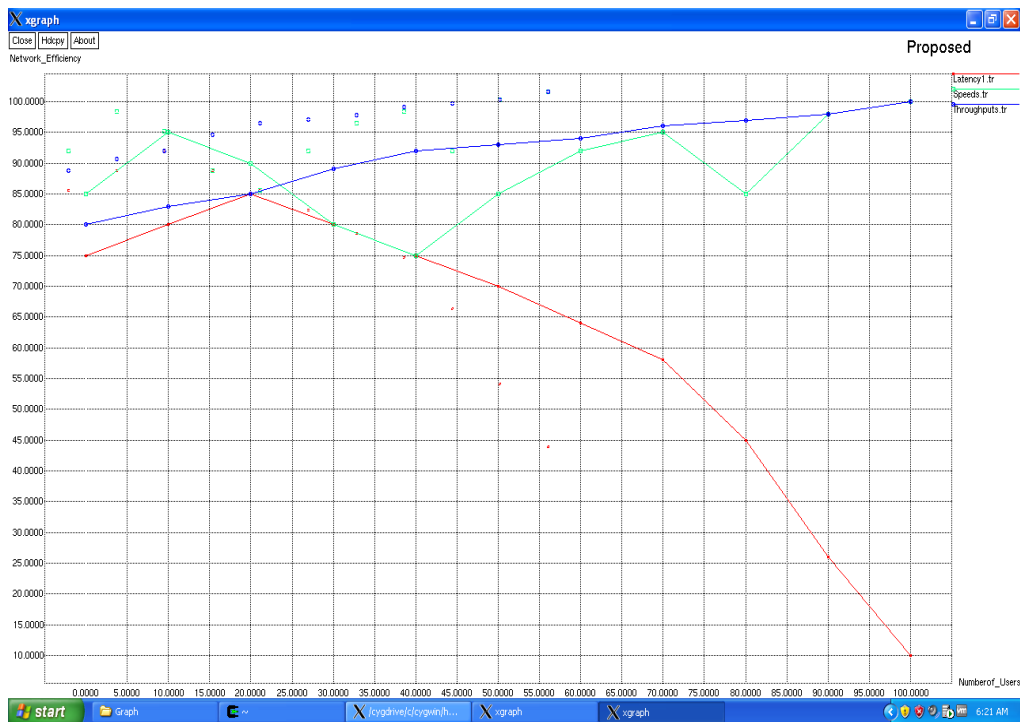

Figure 6. Router performance with intelligence

Router performance with intelligence is shown in Figure 6. Here the parameters such as speed, latency and throughput are plotted for proposed system with respect to their network efficiency. Achieving higher throughput shows that better performance of the network. Table 3 expresses the proposed algorithm performance result in parameters wise.

Intelligent Packet Delivery in Router Using Structure Optimized Neural Network (R. Deebalakshmi)

Table 3. Parameters used for comparing router performance

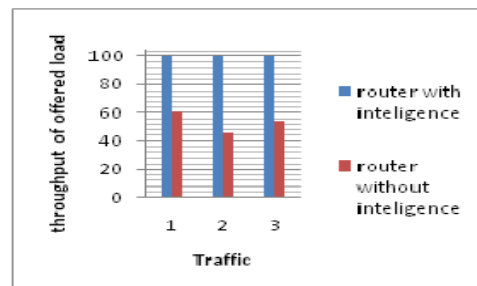| Parameters	Router With Intelligence	Router Without Intelligence
Throughput (percent)	96.20%	72.50%
Average Latency(microseconds)	53.2	152.3
Misordered Packets(Percent)	0%	2.15%

Figure 7. Performance Comparison of Router With Intelligence and Without Intelligence
Based on Load Offered

Figure 7 demonstrates the performance comparison of router with intelligence and router without intelligence on basis of load offered to the network. The throughput of offered load for router with intelligence is better compared to the router without intelligence.

4.   CONCLUSION

In router large volume of data can enter and leave, this huge amount of data is handled and classified by structure optimized neural network, the accuracy of proposed system is higher then traditional classification methods like port based ,deep packet inspection and statistical classification. This mechanism will improve router efficiency by increased throughput and decreased latency. Also proposed method is used by the administrator for taking good control decision about the network activities, classification of application usage and style, anomaly detection and accounting.

REFERENCES

[1]   Lakhina, Anukool, Mark Crovella, and Christophe Diot, "Mining anomalies using traffic feature distributions" , *ACM SIGCOMM Computer Communication Review*, vol. 35, no. 4, pp. 217-228, 2005.
[2]   C. Estan and G. Varghese, "New directions in traffic measurement and accounting: Focusing on the elephants, ignoring the mice", *ACM Transactions on Computer Systems*, vol. 21, no.3, pp.270-313, 2003.
[3]   Peng, J., Liu, T., Li, H., and Guo, B., "Energy-efficient prediction clustering algorithm for multilevel heterogeneous wireless sensor networks", *International Journal of Distributed Sensor Networks*, vol. 9, no. 2, pp. 1-8, 2013.
[4]   T.T. Nguyen and G. Armitage, "A survey of techniques for Internet traffic classification using machine learning" , *EEE Communications Surveys & Tutorials*, vol 10, no.4, p56-76, 2008.
[5]   A. Moore and D. Zuev, *"Internet traffic classification using Bayesian analysis techniques",* ACM SIGMETRICS Performance Evaluation Review, vol. 33, no.1, pp. 50-60,2005.
[6]   Grimaudo, L., Mellia, M., Baralis, E. and Keralapura, R., "Select: Self-learning classifier for internet traffic", *IEEE Transactions on Network and Service Management*, vol.11, no. 2, pp.144-157, 2014.
[7]   Bujlow, T., Riaz, T. and Pedersen, J.M., "*A method for classification of network traffic based on C5.0 Machine Learning Algorithm*, 2012 IEEE International Conference on Computing, Networking and Communications (ICNC), pp. 237-241,2012.
[8]   Karagiannis, T., Papagiannaki, K. and Faloutsos, M., "Blinc: multilevel traffic classification in the dark", *ACM SIGCOMM Computer Communication Review*, vol. 35, no.4, pp. 229–240, 2005.
[9]   Roughan, M., Sen, S., Spatscheck, O. and Duffield, N., "*Class-of-Service Mapping for QoS: A statistical signature-based approach to IP traffic classification*", Proceedings of the 4th ACM SIGCOMM conference on Internet measurement, Taormina, Sicily, Italy pp. 135-148, 2004.
[10]   Afzal, S., & Wani, M. A., "Comparative Study of Adaptive Learning Rate with Momentum and Resilient Back Propagation Algorithms for Neural Net Classifier Optimization", *International Journal of Distributed and Cloud Computing*, vol. 2, no.1 pp.1-6, 2014.

[11] T. Hemalatha, "SECURE Detecting Sybil Attack with a Scalable Protocol", *International Journal of MC Square Scientific Research*, vol. 4, no. 1, pp. 35-41, 2012.

[12] Fadli Sirait., "Capacity Improvement and Protection of LTE Network on Ethernet Based Technique", *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 16, no.1, pp. 1-10, 2018.

[13] Aman, A. H. M., Hashim, A. H. A., and Ramli, H. A. M., "Simulation Analysis for Multicast Context Delivery Network Mobility Management", *Indonesian Journal of Electrical Engineering and Informatics (IJEEI),* vol. 5, no. 4, pp. 390-394, 2017.

[14] Hashim, A. H. A., Aman, A. H. M., and Ramli, H. A. M., "Quantitative Evaluation for PMPIv6 Multicast Fast Reroute Operations", *Bulletin of Electrical Engineering and Informatics*, vol. 6, no. 4, pp. 371-376, 2017.