SYMBOLIC REPRESENTATION OF INTERNET TRAFFIC DATA USING MULTIPLE KERNEL FUZZY C-MEANS

N. Manju[1] and B.S. Harish[2]
[1]Department of Information Science and Engineering, Sri Jayachamarajendra College of Engineering, India
[2]Department of Information Science and Engineering, JSS Science and Technology University, India

Abstract

Network traffic classification is a core part of the network traffic management. Network management is a critical task since the various new applications are emerging every moment and increase in the number of users of an internet. Due to this problem, there is a need of internet traffic classification for smooth management of an internet by the internet service providers (ISP). Network traffic can be classified based on port, payload and statistical approach. In the proposed work, a novel method to represent internet traffic data based on clustering of feature vector using Multiple Kernel Fuzzy C-Means (MKFCM) is proposed. Further, feature vector of each cluster is used to build an interval valued representation (symbolic) using mean and standard deviation. In addition, this interval valued features are stored in knowledge base as a representative of the cluster. Further, to classify the symbolic interval data, we used symbolic classifier. To validate the effectiveness of the proposed model, experimentation is conducted on standard Cambridge University internet traffic dataset. Further, the proposed symbolic classifier compared with other existing classifiers such as Naïve Bayes, KNN and SVM classifier. The experiment outcome infers that; the proposed symbolic representation classifier performs better than other classifiers.

Keywords:

Internet Traffic, Representation, Symbolic Feature, Classification

1. INTRODUCTION

Internet traffic classification is very much essential to provide Quality of Service (QoS) [1]. Intrusion detection [2]-[4] and other various types of internet traffic detection [5], [6] is a very important task in today's network management and security. The main reason is with the demand of cloud computing [7], [8] where the deployment of number of applications are quickly increasing via the internet. There are various approaches to identify the internet traffic. Well known and fundamental approaches are based on the port number from a given TCP or UDP packet header of internet traffic assigned by the Internet Assigned Number Authority (IANA) [9]. However, the traffic allocates port numbers dynamically due to increase in the number of applications. Therefore, even though the port based approach is simple and faster, there are no effective methods to classify the IP traffic [10], [11]. Alternative approaches are based on the payload signatures to identify the internet traffic [12]. This method has several issues such as difficult to manage large database, cannot handle encrypted internet traffic and lawful inspection. Therefore, to overcome the drawback of port based and payload based approach, an alternative approach called flow based statistical methods are used at present day research [13].

2. LITERATURE SURVEY

Identification and classification of internet traffic data plays a vital role in internet traffic management. The task poses various challenges. The challenges include missing values in the dataset, high dimension of the feature set, class imbalance and timely classification in case of real time application and lack of ability in accurate classification. Data cleaning and data preprocessing is the initial step where it contributes in cleaning the artifacts present in the internet data. The method of cleaning data using Variation of Edited Nearest Neighbor (VENN) is presented in [14] and result shows improved accuracy of traffic classification. Various feature selection methods are applied to reduce the dimensionality which in turn increases the accuracy. Further it takes less time to classify internet traffic. Balanced Feature Selection (BFS) method is proposed in [15] to reduce the features and chooses optimal feature subset. Rough set method is proposed in [16] which not only reduces the dimensions, also improves the accuracy and computing performance of the classifier. Weighted Symmetrical Uncertainty Area Under roc Curve (WSU_AUC) is proposed to select stable and robust features in [17]. To study the merits and demerits of proposed feature selection methods, different metrics such as goodness, stability and similarity are proposed in [18]. The presented methods used for dimensionality reduction which addresses the multi class imbalance problem. Analysis of multiclass imbalance problem and possible solutions are investigated in [19]. Various classification methods are introduced using machine learning approaches including Data Gravitation Classification (DGC) [20] - [22]. There are other real time applications which require fast and timely classification of internet traffic. To demonstrate the effectiveness of classification, Naïve Bayes and Decision Tree based C4.5 algorithms are used in [23] over VOIP traffic and online games.

To handle the issues related to internet traffic, various machine learning approaches are studied in the literature. To mention a few, supervised classification methods such as Bayes Net, Naïve Bayes, Decision Tree and Neural Network [24]. The unsupervised machine learning approaches such as K-Means, Expectation Maximization [25], [26], Autoclass and DBSCAN [27] are widely used in the literature.

In [28], the authors proposed a symbolic representation model using interval valued symbolic features for text documents. The enhanced work is proposed by applying adaptive Fuzzy C-Means (FCM) in [29]. The proposed approach preserves intra-class variation. However, this method is suitable only for linear data. The usage of kernel with FCM (KFCM) and Multiple Kernel FCM (MKFCM) are the two essential variants of FCM. These variants of kernels are basically used to cluster nonlinear data [30]. The outcome of KFCM relies on selection of right kernel function. At the same time, selection of kernels for many applications is not an easy task. To handle this issue, MKFCM is used in [31]. MKFCM is based on the KFCM. However, it uses combination of kernel functions. MKFCM is flexible in choosing the kernel functions. During text classification, Multiple Kernel Fuzzy Clustering is applied using Euclidean distance, Cosine similarity, Jaccard coefficient, Pearson correlation coefficient to determine pair wise distance between the documents.

In the proposed work, we are recommending to use a new method to efficiently express intra-class variation using MKFCM clustering methods. After clustering, we are representing each of the clusters with an interval valued of feature vector (symbolic representation) using mean and standard deviation. Further, the representation will be stored in knowledge base which is a representative of each cluster. In internet traffic classification, the expert system is used to assign the class label depending upon the degree of belongingness.

The remaining part of the paper is structured as follows: the proposed method is presented in section 3, detailed experimentation is depicted in section 4 and section 5 presents the concluding remarks along with future work.

3. PROPOSED METHOD

The proposed method has two steps: i) Representation based on Multiple Kernel FCM (MKFCM) and ii) Internet Traffic Classification.

3.1 REPRESENTATION BASED ON MKFCM

In the proposed system, initially the internet traffic data is represented by feature vector matrix. To reduce dimensionality of feature vector matrix, we adopt Regularized Locality Preserving Indexing (RLPI) method [32]. Unfortunately, RLPI has noticeable intra-class variations. Thus, to handle this variations, we propose cluster based Multiple Kernel Fuzzy C-Means (MKFCM) method. In this method, we are considering intra-class variations via MKFCM clustering approach. We represent each cluster by an interval valued feature vector.

Let d_1, d_2, d_3,…, d_N be a set of N data points and $F_k = f_{k1}, f_{k2}, f_{k3},\ldots,f_{km}$ be a set of m features. The objective function of MKFCM is given by,

$$J\left(w,U,V\right)=\sum_{i=1}^{N}\sum_{c=1}^{C}u_{ic}^{m}\left\|\Phi\left(d_i\right)-\Phi\left(v_i\right)\right\|^2 \tag{1}$$

where w is the weights, U is the membership and V is the cluster center. C is the number of cluster and N is the number of features. The u_{ic} is the membership value of the i^{th} data point to c^{th} cluster.

The v_c is c^{th} cluster center. Φ is an implicit nonlinear map and

$$\left\|\Phi\left(d_i\right)-\Phi\left(v_c\right)\right\|^2=K_L\left(d_i,d_i\right)+K_L\left(v_c,v_c\right)-2K_L\left(x_i,v_c\right) \tag{2}$$

where K_L is the inner product of kernel function. i.e. $K_L\left(x,y\right)=\Phi\left(x\right)^T\Phi\left(y\right)$, given by,

$$\begin{aligned}K_L\left(d_i,v_c\right)=w_1K_1\left(d_i,v_c\right)+w_2K_2\left(d_i,v_c\right)+w_3K_3\left(d_i,v_c\right)\\+\ldots+w_lK_l\left(d_i,v_c\right)\end{aligned} \tag{3}$$

where $w=(w_1,w_2,w_3,...,w_l)$ is a vector which consists of weights and the sum of weights equal to 1. i.e., $w_1+w_2+w_3+,...,+w_l=1$ and value of each weight should be greater than or equal to zero. i.e. $w_l \geq 0, \forall_l$.

In the proposed method, we use four kernels to calculate the pair wise distance between each data point and the cluster. Following are the kernels used: Euclidean distance, Cosine similarity, Jaccard co-efficient and Pearson Correlation Coefficient.

The main purpose of MKFCM is to find the combination of weights w, membership U and cluster center V which minimizes the objective function present in Eq.(1). To obtain the membership value (u_{ic}), Eq.(1) is used with Lagrange multiplier. The membership function transforms to:

$$u_{ic}=\frac{1}{\sum_{c=1}^{C}\left(\frac{D_{ic}^2}{D_{ic}^2}\right)^{\frac{1}{m-1}}} \tag{4}$$

where $D_{ic}^2 = \| \phi\left(d_i\right)-\phi\left(v_c\right)\|^2$

The weight w is obtained by solving Eq.(1). Further, using Lagrange multiplier l, weight w becomes:

$$wl=\frac{\frac{1}{\beta_l}}{\frac{1}{\beta_1}+\frac{1}{\beta_2}+\ldots+\frac{1}{\beta_L+}} \tag{5}$$

where the coefficient β_l is given by,

$$\beta_l=\sum_{i=1}^{N}\sum_{c=1}^{C}u_{ic}^{m}\alpha_{icl} \tag{6}$$

where the coefficient α_{icl} is given by,

$$\alpha_{icl}=K_l\left(d_i,d_i\right)-2\sum_{j=1}^{N}u_{jc}K_l\left(d_i,d_j\right)+\sum_{j=1}^{N}\sum_{k=1}^{N}u_{jc}u_{kc}K_l\left(d_j,d_k\right) \tag{7}$$

MKFCM method is used to cluster the training data points. The intra-class variance of each feature is captured in the form of interval value. i.e., $\left[f_{ck}^-, f_{ck}^+\right]$, where $f_{ck}^-=\mu_{ck}-\sigma_{ck}$ and $f_{ck}^+=\mu_{ck}$ $+\sigma_{ck}$. The μ_{ck} is the mean of k^{th} feature of data points present in the c^{th} cluster. The σ_{ck} is the standard deviation of the k^{th} feature of vector present in the c^{th} cluster. The interval value serve as upper and lower limits of the feature vector in a cluster. The reference data point for class C_c is formed using the representation of each feature in the form of interval value. i.e.

$$RF_c=[f_{c1}^-, f_{c1}^+],[f_{c2}^-, f_{c2}^+],...,[f_{cm}^-, f_{cm}^+] \tag{8}$$

The interval values are used to represent the c^{th} cluster. Therefore, we will have N number of symbolic vectors representing clusters corresponding to a class.

Algorithm 1: MKFCM method

```
Data: RLPI feature vector with $F$ features, set of kernel function
      $K$ number of cluster $C$, fuzzification degree $m$ and
      convergence criteria $\varepsilon$
Result: Symbolic feature vector RF,
Step 1: Initialize membership matrix $U$
```

Step 2: Repeat: Calculate normalized membership value

$$u_{ic} = \frac{u_{ic}^{m}}{\sum_{i=1}^{N} u_{ic}}$$

Step 3: Calculate the α_{icl} coefficient using Eq.(7)

Step 4: Calculate the β_l coefficient using Eq.(6)

Step 5: Update weights using Eq.(5)

Step 6: Calculate distance as $D_{ic}^2 = \sum_{i=1}^{L} \alpha_{icl} wl^2$

Step 7: Update membership value using Eq.(4)

Step 8: until: $\{U(t) - U(t-1) < \varepsilon\}$

Step 9: Calculate μ_{ck} and σ_{ck} for each cluster C_c

Step 10: Represent each cluster using symbolic vector RF as in Eq.(8)

3.2 INTERNET TRAFFIC CLASSIFICATION

To classify the internet traffic, in the proposed work, we considered a test data point which has *m* features containing crisp values. The features of test data point are compared with corresponding interval valued features. The number of features which falls within the interval decides the belongingness. The class label of the test data point is obtained by calculating the degree of belongingness B_c. The degree of belongingness B_c is calculated as,

$$B_c = \sum_{k=1}^{m} c(f_{tk},[f_{cm}^{-}, f_{cm}^{+}]) \tag{9}$$

where,

$$C(f_{tk},[f_{cm}^{-}, f_{cm}^{+}]) = \begin{cases} 1 & f(f_{tm} \geq f_{cm}^{-} \; and \; f_{tm} \leq f_{cm}^{+} \\ 0 & otherwise \end{cases} \tag{10}$$

The features of the test data point falling within the respective feature interval of the reference class contributes as value 1 towards B_c. The value of B_c of all remaining clusters are computed and class label is assigned based on the test data point for which the class has highest B_c.

4. EXPERIMENTAL SETUP

4.1 DATASET

The authors in [33] have used the internet traffic data obtained from Cambridge University. Internet traffic data flows associated with various categories of classes such as www, mail, bulk, attack, P2P, database and services. The dataset which are used consists 324,277 data flows. The flows are split into 7 classes and uses 12 attribute values to determine the classes. The preferred attributes and its definition is shown in [34]. The 1000 random flows for each class are equally chosen for our experimentation to avoid class imbalance issue. A total number of 7000 flows were used for the experimentation.

4.2 PRE-PROCESSING

Preprocessing the data is the first step before start of experimentation to get the effective result. During preprocessing, raw data will be converted into meaningful format. To pre-process the data, we have applied normalized function as shown in the Eq.(11) to convert the data into 0-1 range.

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{11}$$

where, *x* = current value of the feature set, x_{min} = minimum value of the feature set and x_{max} = maximum value of the feature set.

4.3 EXPERIMENTATION

This section presents the results of the experiment which are conducted and discussed the effectiveness of the method proposed on Moore's dataset. In this experiment we use multiple kernels to compute the pair wise distance between two data points. The choice of the kernels is: Euclidean distance (KFCed), Cosine similarity (KFCcs), Jaccard Coefficient (KFCjc) and Pearson Correlation Coefficient (KFCpcc). The fuzzification degree (m) is set to 2 and the Convergence Criteria (ε) is set as 0.0001 by empirical evaluation to build symbolic representation model.

The obtained classification accuracy of the proposed symbolic classifier is compared with Naive Bayes classifier (NB), K-Nearest Neighbour (KNN) and Support Vector Machine (SVM) classifier. During the experimentation, data is divided into two sets having (a) 50% training and 50% testing and (b) in the second set of experiment we have set 60% of training 40% testing. Both a and b sets are run separately for 5 times to measure the accuracy which is tabulated in Table.1. Classification accuracy is chosen as the metric to compute and compare the accuracy across different classifiers. The traffic data for training are chosen randomly for each class to create symbolic feature vectors.

From Table.1, it is observed that, the proposed method based on symbolic representation obtained a better result. This is because, for a given internet traffic dataset, we cannot predict which kernel performs well. However, when we combine kernels, multiple kernels contribute more to the cluster. Thus, there is an improvement in the result.

Table.1. Comparative analysis of the Proposed Method

Classifiers	**Training vs. Testing**	**Kernels**				
		KFC$_{ed}$	***KFC$_{cs}$***	***KFC$_{jc}$***	***KFC$_{pcc}$***	***MKFCM***
NB	50:50	71.35	74.50	70.85	73.25	77.95
	60:40	73.15	75.00	71.20	74.50	79.55
KNN	50:50	70.25	70.85	70.30	71.25	73.25
	60:40	71.55	72.65	73.55	73.95	74.90
SVM	50:50	76.45	77.15	78.90	79.10	80.25
	60:40	78.50	79.80	79.10	81.25	83.15
Symbolic Classifier	50:50	78.65	79.20	80.25	81.55	82.95
	60:40	80.45	81.55	82.90	83.75	84.60

The Fig.1 depicts the performance of various kernels used viz., Euclidean distance (KFCed), Cosine similarity (KFCcs), Jaccard Coefficient (KFCjc) and Pearson Correlation Coefficient

(KFCpcc) and the proposed ultiple Kernel Fuzzy C-Means (MKFCM).

Fuzzy C-means is a well-known soft clustering approach. However, identifying the right kernel in FCM and implementing kernel function is not easy in practical. Reason behind this is to map nonlinear data relationships to relevant feature space. To address this problem, MKFCM is proposed. MKFCM is an extension of Fuzzy C-Means with multiple kernels. Combining many kernels and adjusting the kernel weights automatically gives the better result when compared to other methods such as Naive Bayes classifier (NB), K-Nearest Neighbour (KNN) and Support Vector Machine (SVM) classifier.

The proposed MKFCM method achieves highest accuracy of 82.95 and 84.60 for training and testing ratio of 50:50 and 60:40 split respectively.

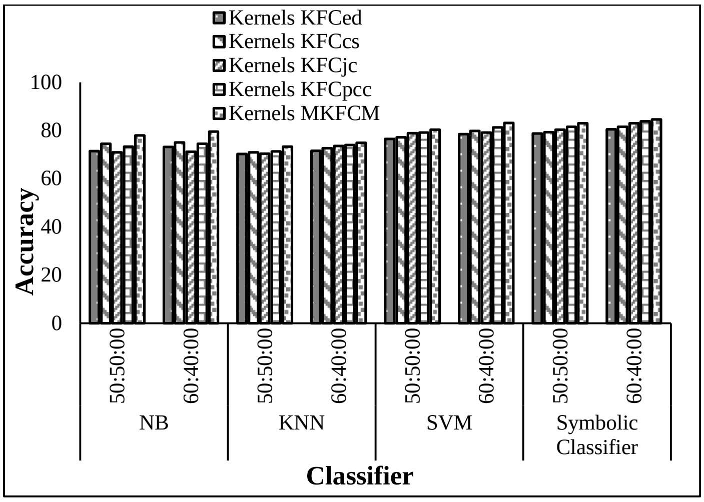

Fig.1. Comparative analysis of various kernels with the proposed method

5. CONCLUSIONS

Internet traffic classification plays a vital role in internet traffic management by the internet service providers. This is due to exponential growth in the number of applications and increase in the number of users who depend more on internet. Classifying those applications or categories of an application to provide Quality of Service to the user is fundamental necessity in the current trend. Hence, in this paper, a new representation for internet traffic data is presented. Internet traffic data is represented using symbolic features. This is a new representation model for the internet traffic data using clustering based Multiple Kernel Fuzzy C-Means (MKFCM) algorithm. The kernels used for the experimentation are the Euclidean distance (KFCed), Cosine similarity (KFCcs), Jaccard Coefficient (KFCjc) and Pearson Correlation Coefficient (KFCpcc). The proposed symbolic representation will give interval valued representation. To evaluate the efficacy of the proposed model, experimentations are conducted on the standard Cambridge University internet traffic dataset. Further, results are compared with the Naive Bayes classifier (NB), K-Nearest Neighbour (KNN) and Support Vector Machine (SVM) classifier. The experimentation outcome reveals that, the symbolic representation using MKFCM clustering techniques obtain better classification accuracy using symbolic representation approaches compared to NB, KNN and SVM classifier. In future, the proposed work will extend to use symbolic feature selection methods. This will intern reduce the higher dimensions to lower dimensions. Further, it is also planned to explore other features which will capture intra-class variations effectively.

REFERENCES

[1] M. Roughan, S. Sen, O. Spatscheck and N. Duffield, “Class-of-Service Mapping for QoS: A Statistical Signature-based Approach to IP Traffic Classification”, *Proceedings of 4th ACM Conference on Internet Measurement*, pp. 135-148, 2004.

[2] T. Karagiannis, K. Papagiannaki and M. Faloutsos, “BLINC: Multilevel Traffic Classification in the Dark”, *Proceedings of ACM International Conference on Computer Communication Review*, pp. 229-240, 2005.

[3] T.T. Nguyen and G.J. Armitage, “A Survey of Techniques for Internet Traffic Classification using Machine Learning”, *IEEE Communications Surveys and Tutorials*, Vol. 20, No. 1, pp. 56-76, 2008.

[4] A. Finamore, M. Mellia, M. Meo and D. Rossi, “Kiss: Stochastic Packet Inspection Classifier for UDP Traffic”, *IEEE/ACM Transactions on Networking*, Vol. 18, No. 5, pp. 1505-1515, 2010.

[5] Y. Xiang, W. Zhou and M. Guo, “Flexible Deterministic Packet Marking: An IP Traceback System to Find the Real Source of Attacks”, *IEEE Transactions on Parallel and Distributed Systems*, Vol. 20, No. 4, pp. 567-580, 2009.

[6] Z.M. Fadlullah, T. Taleb, A.V. Vasilakos, M. Guizani and N. Kato, “DTRAB: Combating Against Attacks on Encrypted Protocols through Traffic-Feature Analysis”, *IEEE/ACM Transactions on Networking*, Vol. 18, No. 4, pp. 1234-1247, 2010.

[7] R. Buyya, C.S. Yeo, S. Venugopal, J. Broberg and I. Brandic, “Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility”, *Future Generation Computer Systems*, Vol. 25, No. 6, pp. 599-616, 2009.

[8] M. Armbrust, A. Fox and R. Griffith, “A View of Cloud Computing”, *Communications of the ACM*, Vol. 53, No. 4, pp. 50-58, 2010.

[9] H. Dreger, A. Feldmann, M. Mai, V. Paxson and R. Sommer, “Dynamic Application-Layer Protocol Analysis for Network Intrusion Detection”, *Proceedings of 15th International USENIX Association Symposium on Security*, pp. 257-272, 2006.

[10] A.W. Moore and K. Papagiannaki, “Toward the Accurate Identification of Network Applications”, *Proceedings of International Workshop on Passive and Active Network Measurement*, pp. 41-54, 2005.

[11] S. Sen, O. Spatscheck and D. Wang, “Accurate, Scalable in-Network Identification of P2P Traffic using Application Signatures”, *Proceedings of 13th International Conference on World Wide Web*, pp. 512-521, 2004.

[12] R.C. Jaiswal and S.D. Lokhande, “Machine Learning based Internet Traffic Recognition with Statistical Approach”, *Proceedings of Annual IEEE International Conference on Networking and Security*, pp. 1-6, 2013.

[13] R.Y. Wang, L.I.U. Zhen and Z. Ling, "Method of Data Cleaning for Network Traffic Classification", *The Journal of China Universities of Posts and Telecommunications*, Vol. 21, No. 3, pp. 35-45, 2014.

[14] L. Zhen and L. Qiong, "A New Feature Selection Method for Internet Traffic Classification using ML", *Physics Procedia*, Vol. 33, pp. 1338-1345, 2012.

[15] M. Sun, J. Chen, Y. Zhang and S. Shi, "A New Method of Feature Selection for Flow Classification", *Physics Procedia*, Vol. 24, pp. 1729-1736, 2012.

[16] H. Zhang, G.Lu, M.T. Qassrawi, Y. Zhang and X. Yu, "Feature Selection for Optimizing Traffic Classification", *Computer Communications*, Vol. 35, No. 12, pp. 1457-1471, 2012.

[17] A. Fahad, Z. Tari, I. Khalil and I. Habib, "Toward an Efficient and Scalable Feature Selection Approach for Internet Traffic Classification", *Computer Networks*, Vol. 57, No. 9, pp. 2040-2057, 2013.

[18] S. Wang and X. Yao, "Multiclass Imbalance Problems: Analysis and Potential Solutions", *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, Vol. 42, No. 4, pp. 1119-1130, 2012.

[19] L. Peng, Lizhi, B. Yang, Y. Chen and X. Zhou, "An Under-Sampling Imbalanced Learning of Data Gravitation Based Classification", *Proceedings of 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery*, pp. 1-6, 2016.

[20] L. Peng, H. Zhang, B. Yang and Y. Chen, "A New Approach for Imbalanced Data Classification based on Data Gravitation", *Information Sciences*, Vol. 288, pp. 347-373, 2014.

[21] L. Peng, H. Zhang, B. Yang and Y. Chen, "Imbalanced Traffic Identification using an Imbalanced Data Gravitation-based Classification Model", *Computer Communications*, Vol. 102, pp. 177-189, 2017.

[22] T.T.T. Nguyen, G. Armitage, P. Branch and S. Zander, "Timely and Continuous Machine-Learning-based Classification for Interactive IP Traffic", *IEEE/ACM Transactions on Networking*, Vol. 20, No. 6, pp. 1880-1894, 2012.

[23] Neeraj Namdev, Shikha Agrawal and Sanjay Silkari, "Recent Advancement in Machine Learning based Internet Traffic Classification", *Procedia Computer Science*, Vol. 60, pp. 784-791, 2015.

[24] Y. Wang, Y. Xiang, J. Zhang, W. Zhou, G. Wei and L.T. Yang, "Internet Traffic Classification using Constrained Clustering", *IEEE Transactions on Parallel and Distributed Systems*, Vol. 25, No. 11, pp. 2932-2943, 2014.

[25] Hardeep Singh, "Performance Analysis of Unsupervised Machine Learning Techniques for Network Traffic Classification", *Proceedings of 5th IEEE International Conference on Advanced Computing and Communication Technologies*, pp. 1-6, 2015.

[26] J. Erman, M. Arlitt and A. Mahanti, "Traffic Classification using Clustering Algorithms", *Proceedings of SIGCOMM workshop on Mining Network Data*, pp. 1-5, 2006.

[27] D.S. Guru, B.S. Harish and S. Manjunath. "Symbolic Representation of Text Documents", *Proceedings of 3rd Annual ACM Bangalore Conference*, pp. 1-5, 2010.

[28] B.S. Harish, B. Prasad and B. Udayasri, "Classification of Text Documents using Adaptive Fuzzy C-Means Clustering", *Proceedings of IEEE International Conference on Recent Advances in Intelligent Informatics*, pp. 205-214, 2014.

[29] K.R. Muller, S. Mika, G. Ratsch, K. Tsuda and B. Scholkopf, "An Introduction to Kernel-Based Learning Algorithms", *IEEE Transactions on Neural Networks*, Vol. 12, No. 2, pp. 181-200, 2001.

[30] Hsin Chien Huang, Yung-Yu Chuang and Chu-Song Chen, "Multiple Kernel Fuzzy Clustering", *IEEE Transactions on Fuzzy Systems*, Vol. 20, No. 1, pp. 120-134, 2012.

[31] D. Cai, X. He, W.V. Zhang and J. Han, "Regularized Locality Preserving Indexing via Spectral Regression", *Proceedings of 16th ACM International Conference on Information and Knowledge* Management, pp. 741-750, 2007.

[32] A. Moore, Z. Denis Zuev and M. Crogan, "Discriminators for Use in Flow-Based Classification", Technical Report, Department of Computer Science, Queen Mary University of London, pp. 1-19, 2013.

[33] F. Ertam and A. Engin, "A New Approach for Internet Traffic Classification: GA-WK-ELM", *Measurement*, Vol. 95, pp. 135-142, 2017.