

# An Efficient Nonnegative Matrix Factorization & Game Theoretic Framework Based Data Clustering

Dr. V. Umadevi<sup>1</sup>, R.Sumithra<sup>2</sup>

Research Guide<sup>1</sup>, Department of Computer Science, Jairams Arts & Science College, Attamparappu, Kakkavadi(PO), Karur, Tamilnadu, India-639003

MPhil Scholar<sup>2</sup>, Department of Computer Science, Jairams Arts & Science College, Attamparappu, Kakkavadi(PO), Karur, Tamilnadu, India-639003

**Abstract**—This Mostly, factorization of matrices is not unique, Non-negative Matrix Factorization (NMF) changes from the Principal Component Analysis, Singular Value Decomposition, Nystrom Method, and it imposes the controls that the factors must be non-negative. The proposed method utilizes a most powerful tool derivative from evolutionary game theory, which permits re-organizing the clustering attained with NMF method, making it consistent with the structure of the data set. The new propose a method to filter the clustering results obtained with the nonnegative matrix factorization (NMF) technique, imposing consistency constraints on the final labeling of the data set. The research community focused its effort on the initialization and on the optimization part of this method, without paying concentration to the final cluster assignments. The propose a game theoretic framework in which each object to be clustered is symbolized as a player, which has to choose its cluster membership. The detailed obtained with NMF method is used to initialize the approach space of the players and a weighted graph is worn to model the interactions among all the players. These connections allow the players to choose a cluster which is coherent with the clusters chosen by similar players, a property which is not guaranteed by NMF, since it produces a soft clustering of the data. The proposed results on common benchmarks show that our model is able to progress the performances of many NMF formulations..

**Keywords**— Enter Data Cluster; Nonnegative Matrix Factorization; Weighted Graph ; Game Theoretic.

## I.INTRODUCTION

This Clustering has received a significant amount of attention as an important problem with many applications, and a number of different algorithms have emerged over the years. Recently, the use of Non-Negative Matrix Factorization (NMF) for partitioned clustering has attracted much interest. The popularity of NMF has significantly increased; the authors proposed multiplicative NMF algorithms

which they applied to image data. At present, NMF and its variants have already found a wide spectrum of applications in several areas such as pattern recognition and feature extraction dimensionality reduction, segmentation and clustering, text mining and neurobiology. The concept of matrix factorization is used in a wide range of important applications and each matrix factorization relies on an assumption about its components and its underlying structures, it is an essential process in each application domain. Very often, the data sets to be analysed are non-negative, and sometimes they also have a sparse representation. In machine learning approach, sparseness is strongly related to feature selection and certain generalizations in learning algorithms, while non-negativity relates to possibility distributions.

Clustering [6] is an unsupervised learning mechanism which locates unknown groups of related dataset. In the areas of information retrieval, bioinformatics, and digital image processing, it has been a very imperative problem where many algorithms have been residential using various objective functions. K-means clustering approach is a well-known scheme that tries to minimize the sum of squared distances between own cluster center and each data point. K-means has been generally applied thanks to its comparative simplicity. However, it is well known that K-means is flat to discover only a local minimum values and, therefore, strongly depends upon K-means initial conditions. A common approach in this condition is to run K-means with many different initial conditions and choose the best solution. Modifications to algorithms were also created by refining initial guesses or bountiful variations at the expense of more processing overhead time. The reformulated the minimization problem as a hint maximization problem and suggested an algorithm using a matrix factorization. Regarding the same objective function, the sum of squared errors, recently proposed a method called Affinity Propagation that performs clustering by passing messages between data points.

Clustering, which partitions a data set into different groups unsupervised, is one of the most essential topic in statistical learning approach. Most

established clustering mechanisms are designed for one-side data clustering, i.e. cluster either data points or features. Still, in many real-world applications, the clustering based analysis is interested in two side data clustering results, i.e. group the data points and features concurrently, e.g., “words” and “documents” in document analysis, “items” and “users” in collaborative filtering technique, “genes” and “samples” in microarray data analysis, etc. Classically, instead of being independent, the different clustering tasks on dataset and features are closely connected, and it is challenging for traditional clustering algorithms to exploit the data and features interdependence more efficiently. Consequently, co- culturing procedures, which aim to cluster both features and data simultaneously by limit the interrelations between them, have been proposed in recent researches [7].

However, the techniques are mentioned above focus on one-side data clustering, i.e. clustering the data side based on the related along the aspect side. Motivated by the duality between features (e.g. documents) and data points (e.g. words), i.e. data points can be group based on their sharing on features, while features can be group based on their sharing on the dataset points, several co-clustering techniques have been projected in the past decade and shown to be superior to conventional one-side data clustering. For instance, proposed a bipartite spectral graph partition method to co-cluster documents and words. Still, it requires that each document cluster is related with a word cluster, which is a very strong limit. The existing algorithm an information theoretic co-clustering algorithm, which can be seen as the conservatory of information blockage method to two-side clustering. The existing projected an orthogonal nonnegative matrix tri-factorization (ONMTF) to co-cluster documents and words, which owns an stylish mathematical form and heartening performance.

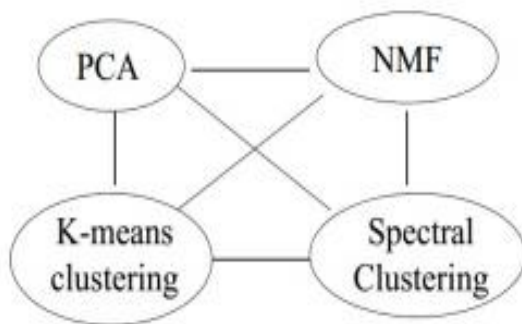


Fig 1:Proposed NMF method

## II.RELATED WORK

The document clustering, a typical process is Latent Semantic Analysis (LSA) which involves a Singular Value Decomposition (SVD) of the

document-term matrix. The explore the relationship between NMF and Probabilistic Latent Semantic Analysis (PLSA), ultimate that the hybrid connections among them give the best results. The established links between PLSA and NMF, and they claim that PLSA solves NMF with KL I-divergence, and for this cost purpose, PLSA provides a better reliability. A comparison of several NMF techniques with various databases was performed. They completed that the NMF techniques generally give better performance than k-means. In fact, the NMF method is rather equivalent to soft PLSA, and k-means typically also gives the equivalent results as NMF. More lately, explore the relationships between K-means/spectral clustering and Nonnegative Matrix Factorization (NMF), and proposed method to use Nonnegative Matrix Trifactorization (NMTF) to co-cluster words and documents at the same time. Due to its mathematical elegance and hopeful experiential results, NMTF method has been further residential to address diverse aspects of co-clustering. However, a notorious bottleneck of NMTF based co-clustering techniques is the measured computational speed because of intensive matrix multiplications concerned in each iteration step of the resolution algorithms, which makes these techniques tough to be apply to large scale data in actual world applications [14].

Bi-clustering (co-clustering) of gene expression data set and advocated the import of such concurrent clustering of genes and conditions for learning more coherent and significant data clusters. They formulated their problem of bi-clustering by proposing a mean squared residue score for measuring cluster quality. One of the initial bi-clustering formulations, block clustering was introduced by Hartigan who called it direct clustering. He proposed various bi-clustering value measures and models including the partition techniques are old in this paper. However, only gives a greedy technique for a hierarchical co-clustering model. This method begins with the entire data set in a single block and then at each stage find the column or row divide of every block into two pieces, choosing the one that produces largest reduction in the total within block variance. The splitting is continued till the reduction of within block discrepancy due to additional splitting is less than a known threshold [4].

The existing proposed non-overlapping and overlapping two-mode partitioning technique, of which the non-overlapping two-mode technique attempting to minimize the same objective function bi-clustering method for gene expression data set using mean squared scum as the measure of the rationality of the conditions and genes. The algorithm constructs one bi-cluster at a time a low mean squared remains plus a large variation from the steady gives a good criterion for spotting a bi-cluster. A sequence of node (i.e. column or row) removals and additions is practical to the gene condition matrix, while the mean squared remains of

the bi-cluster is kept beneath a agreed threshold. They nearby an algorithm called FLOC (FLexible Overlapped bi-clustering) that concurrently produces  $k$  bi-clusters whose mean remains are all less than a pre-defined constant  $r$ . FLOC incrementally shifts a column or row out of or into a bi-cluster depending on whether the column or row is already included in that bi-cluster or not, which is called an exploit [20].

The spectral bi-clustering approaches similar to the one proposed method on gene expression data set to produce checkerboard based structure. The largest several right and left singular vectors of the normalized gene expression data set matrix are subtracted and then a final grouping step using  $k$ -means and regularized cuts is practical to the data predictable to the topmost curious vectors. Different normalizations of conditions and genes. The information-theoretic bi-clustering algorithms that sights a non-negative matrix as an experiential joint probability sharing of two discrete random variables and shams the bi-clustering problem as an optimization problem in sequence theory: the optimal bi-clustering maximizes the joint in sequence between the clustered random variables topic to constraints on the number of column and row clusters [5].

Principal component analysis (PCA) is a generally used statistical technique for unsupervised learning dimension reduction.  $K$ -means clustering method is a mostly used data bi-clustering for unsupervised learning tasks. Here we show that principal components (PCA) are the incessant solutions to the separate cluster membership pointers for  $K$ -means clustering approach. To address this main problem, in this proposed paper, we suggest a Dual Regularized Co-Clustering (DRCC) based on semi-nonnegative matrix tri-factorization method. To address this problem, in this proposed paper, we propose a Dual Regularized Co-Clustering (DRCC) method based on semi non-negative matrix tri-factorization method, which inherits the pros of ONMT. We think that not only the data points but also the skins are discrete samplings method from some manifolds, namely data manifold and characteristic manifold respectively [21].

### **III. PROPOSED APPROACH**

The proposed approach employs game theory method, which lets to reorder the clustering gets with NMF, creating it reliable with the structure of the data set. With our advance, it inflicts that the cluster association has to be re-negotiated for all the objects. To this end, we use a dynamical system viewpoint, in which it is compulsory that similar objects have to fit in to similar clusters, so that the last clustering will be reliable with the structure of the data set. This viewpoint has demonstrated its efficiency in dissimilar semantic categorization situations, which engage a high numeral of interrelated categories and necessitate the employ of contextual and comparison in sequence. The Game

Theoretic Nonnegative Matrix Factorization (GTNMF), our advance to NMF clustering refinement method. We remove the mark vectors of each object in a dataset then, depending on the NMF algorithm worn, it furnish as input to NMF the attribute vectors or a parallel matrix. GTNMF gets as input the matrix  $W$  obtained with NMF and the resemblance graph  $A$  of the dataset to create a reliable clustering of the data.

These limits can be conquering using the relational in order of the data and the stage a consistent labelling. For this cause, in this proposed paper we employ a more powerful tool resulting from evolutionary game theory, which permits to reorder the clustering getter with NMF, making it reliable with the structure of the data set. With our advance, we inflict that the cluster membership has to be re-negotiated for all the objects. To this end, we use a dynamical system viewpoint, in which it is compulsory that alike objects have to belong to alike clusters, so that the final clustering will be reliable with the structure of the data set. This viewpoint has established its efficiency in dissimilar semantic classification scenarios, which engage a high figure of interrelated group and need the use of background and resemblance information.

#### **A. NMF Clustering**

NMF is employed as clustering algorithm in different applications. It has been successfully applied in parts-of-whole decomposition, object clustering, face recognition, multimedia analysis, and DNA gene expression grouping. It is an appealing method because it can be used to perform together objects and feature clustering. The generation of the factorized matrices starts from the assumption that the objects of a given dataset belong to  $k$  clusters and that these clusters can be represented by the features of the matrix  $W$ , which denotes the relevance that each cluster has for each object. This description is very useful in soft clustering applications because an object can contain information about different clusters in different measure. For example, a text about the launch of a new car model into the marked can contain information about economy, automotive or life-style, in different proportions. Hard clustering applications require to choose just one of these topics to partition the data and this can be done considering not only the information about the single text, but also the information of the other texts in the texts collection, in order to divide the data in coherent groups. In many algorithms, the initialization of the matrices  $W$  and  $H$  is done randomly and have the drawback to always lead to different clustering results. In fact, NMF converges to local minima and for this reason has to be run several times in order to select the solution that approximates better the initial matrix. To overcome this limitation there were proposed different approaches to find the best initializations based on feature clustering and SVD techniques. These initializations allow NMF to converge always to the same solution. The spherical  $k$ -means to

partition the columns of  $X$  into  $k$  clusters and selects the centroid of each cluster to initialize the corresponding column of  $W$ . Nonnegative Double Singular Value Decomposition (NDSVD) computes the  $k$  singular triplets of  $X$ , forms the unit rank matrices using the singular vector pairs, extracts from them their positive section and singular triplets and with this information initializes  $W$  and  $H$ . This approach has been shown to be almost as good as that obtained with random initialization.

A different formulation of NMF as clustering algorithm (SymNMF). The main difference with classical NMF approaches is that SymNMF takes a square nonnegative similarity matrix as input instead of a  $n \times m$  data matrix. It starts from the assumption that NMF was conceived as a dimension reduction technique and that this task is different from clustering. In fact, dimension reduction aims at finding a few basis vectors that approximate the data matrix and clustering aims at partitioning the data points where similarity is high among the elements of a cluster and low among the elements of different clusters. In this formulation, a basis vector strictly represents a cluster.

**B. Game Theory**

Game theory is a mathematical framework able to model the essentials of decision making in interactive situations. In its normal-form representation, it consists of a finite set of players, a set of pure strategies for each player, and a utility function, which associates strategies to payoffs. Each player can adopt a strategy in order to play a game and the utility function depends on the combination of strategies played at the same time by the players involved in the game, not just on the strategy chosen by a single player. An important assumption in game theory is that the players are rational and try to maximize the value of  $u$ . Furthermore, in non-cooperative games the players choose their strategies independently, considering what other players can play and try to find the best strategy profile to employ in a game.

Nash equilibria represent the key concept of game theory and can be defined as those strategy profiles in which each strategy is a best response to the strategy of the co-player and no player has the incentive to unilaterally deviate from his decision, because there is no way to do better. The players can also play mixed strategies, which are probability distributions over pure strategies. A mixed strategy profile can be defined as a vector where  $m$  is the number of pure strategies and each component denotes the probability that the player chooses its pure strategy. Each mixed strategy corresponds to a point on the simplex and its corners correspond to pure strategies.

The initialization of the matrices  $W$  and  $H$ , is crucial and can lead to different matrix decompositions, since it is performed randomly in many algorithms. To the contrary, the step involving

the final clustering assignment received less attention by the research community. In fact, once  $W$  and  $H$  are computed, soft clustering approaches interpret each value in  $W$  as the strength of association among objects and clusters and hard clustering approaches assign each object  $j$  to the cluster  $C_k$ , where:

$$k = \arg \max(W_{j1}, W_{j2}, \dots, W_{jk}) \quad (1)$$

In a two-player game, a strategy profile can be defined as a pair. The expected payoff for this strategy profile is computed as:

$$u_i(p, q) = p \cdot A_{iq}, u_j(p, q) = q \cdot A_{jp} \quad (2)$$

where and are the payoff matrices of player  $i$  and  $j$  respectively.

In evolutionary game theory, we have a population of agents which play games repeatedly with their neighbours and update their beliefs on the state of the system choosing their strategy according to what has been effective and what has not in previous games, until the system converges. The strategy space of each player  $i$  is defined as a mixed strategy profile  $x_i$ , as defined above. The payoff corresponding to a single strategy can be computed as the average payoff is:

$$u_i(e_i^h) = \sum_{j=1}^n (A_{ij} x_j) h \quad (3)$$

where  $n$  is the number of players with whom the games are played and is the payoff matrix among player  $i$  and  $j$ .

$$u_i(x) = \sum_{j=1}^n x_j^T A_{ij} x_j \quad (4)$$

It must be in 24 pt Regular font. Author name must be in 11 pt Regular font. Author affiliation must be in 10 pt Italic. Email address must be in 9 pt Courier Regular font.

**C. Game Dynamics**

The replicator dynamic equation is use in order to find those states, which correspond to the Nash equilibria of the games,

$$x^h(t + 1) = x^h(t) \frac{u(e^h, x)}{u(x, x)} \quad \forall h \in S \quad (5)$$

This equation allows better than average strategies to grow at each iteration and we can consider each iteration of the dynamics as an inductive learning process, in which the players learn from the others how to play their best strategy in a determined context.



**D. Game Theoretic Nonnegative Matrix Factorization**

Each data point, in our formulation, is represented as a player that has to choose its cluster membership. The weighted graph  $A$  measures the influence that each player has on the others. The matrix  $W$  is used to initialize the strategy space  $S$  of the players. to constrain the strategy space of each player to lie on the standard simplex, as required in a game theoretic framework. The dynamics are not started on the center of the  $K$ -dimensional simplex, as it is commonly done in unsupervised learning tasks, but on a different interior point, which corresponds to the solution point of NMF and do not compromise the dynamics to converge to Nash equilibria. Now that we have the topology of the data  $A$  and the strategy space of the game  $S$  we can compute the Nash equilibria of the games according to equation (5). In each iteration of the system each player plays a game with its neighbor's  $N_i$  according to the similarity graph  $A$  and the payoffs are calculated as follows:

$$u_i(e^h, s) = \sum_{j \in N_i} (a_{ij} s_j) h \quad (6)$$

And

$$u_i(s) = \sum_{j \in N_i} x_i^T (a_{ij} s_j) \quad (7)$$

assume that the payoff of player  $i$  depends on the similarity that it has with player  $j$ ,  $a_{ij}$ , and its preferences,  $x_i$ . During each phase of the dynamics a process of selection allows strategies with higher payoff to emerge and at the end of the process each player chooses its cluster according to these constraints. Since Equation 5 models a dynamical system it requires some criteria to stop. In the experimental part of this work we used as stopping criteria the maximum number of iterations = 100 and  $\|s_t - s_{t+1}\| < \epsilon$ , where  $\| \cdot \|$  is the Euclidean norm between the strategy space at time  $t$  and at time  $t + 1$ .

**E. Algorithm Implementation**

The generation of the factorized matrices starts from the assumption that the objects of a given dataset belong to  $k$  clusters and that these clusters can be represented by the features of the matrix  $W$ , which denotes the relevance that each cluster has for each object.

Step 1: In its normal-form representation, it consists of a finite set of players  $I$

Step 2: Calculate utility function  $S_i$

Step 3: Each player can adopt a strategy in order to play a game and the utility function depends on the combination of strategies played at the same time by the players involved in the game, not just on the strategy chosen by a single player.

Step 4: In a two-player game, a strategy profile pair  $(p, q)$

Step 5: The strategy space of each player  $i$  is defined as a mixed strategy profile  $x_i$ ,

Step 6: The replicator dynamic equation is used in order to find those states, which correspond to the Nash equilibria of the games.

Step 7: consider each iteration of the dynamics as an inductive learning process.

**IV. EXPERIMENTAL RESULTS**

The evaluation of GTNMF has been conducted on datasets with different characteristics. The used textual (Reuters, RCV1, NIPS) and image (COIL-20, ORL, Extended YaleB and PIE-Expr) datasets. The objects belonging to small clusters in order to make the dataset more balanced, simplifying the task. We tested our method using this approach and also keeping the datasets as they are (without reduction), which lead to situations in which it is possible to have in the same dataset clusters with thousands of objects and clusters with just one object (e.g. RCV1). The proposed approach has been validated using two different measures, accuracy (AC) and normalized mutual information (NMI).

**Table 1: Performance of GTNMF Compared to several Semi-NMF-PCA approaches. Normalized Mutual Information**

Dataset	NMF	Semi-NMF-PCA	SymNMF+GT
Reuters	0.451 (±0:026)	0.502 (±0:014)	0.51 (±0:016)
RCV1	0.51 (±0:007)	0.406 (±0:007)	0.422 (±0:007)
PIE-Expr	0.939 (±0:008)	0.95 (±0:004)	0.968 (±0:004)
ORL	0.691 (±0:015)	0.888 (±0:006)	0.921 (±0:006)
COIL-20	0.619 (±0:017)	0.871 (±0:009)	0.875 (±0:012)
ExtYaleB	0.356 (±0:006)	0.308 (±0:005)	0.313 (±0:005)

**Table 2: Performance of GTNMF Compared to several Semi-NMF-PCA approaches. Accuracy**

Dataset	NMF	Semi-NMF-PCA	SymNMF+GT
Reuters	0.322 ( $\bar{\sigma}$ 0:024)	0.517 ( $\bar{\sigma}$ 0:044)	0.528 ( $\bar{\sigma}$ 0:043)
RCV1	0.383 ( $\bar{\sigma}$ 0:009)	0.289 ( $\bar{\sigma}$ 0:014)	0.292 ( $\bar{\sigma}$ 0:015)
PIE-Expr	0.783 ( $\bar{\sigma}$ 0:023)	0.81 ( $\bar{\sigma}$ 0:021)	0.85 ( $\bar{\sigma}$ 0:019)
ORL	0.465 ( $\bar{\sigma}$ 0:019)	0.776 ( $\bar{\sigma}$ 0:017)	0.811 ( $\bar{\sigma}$ 0:018)
COIL-20	0.478 ( $\bar{\sigma}$ 0:023)	0.727 ( $\bar{\sigma}$ 0:036)	0.729 ( $\bar{\sigma}$ 0:037)
ExtYaleB	0.194 ( $\bar{\sigma}$ 0:007)	0.228 ( $\bar{\sigma}$ 0:007)	0.235 ( $\bar{\sigma}$ 0:008)

In particular, it can notice that despite the different settings (textual/image datasets) the proposed algorithm is able improve the NMI performance in 33/36 cases with a maximum gain of 15:3% (which is quite impressive) and a maximum loss of 0:2%. Constant gain in the NMI means, in practice, that the algorithm is able to partition better the dataset, making the final clustering closer to the ground truth.

**V.CONCLUSION**

The proposed system presented GTNMF, a game theoretic model to improve the clustering results obtained with NMF going beyond the classical technique used to make the final clustering assignments. The W matrix obtained with NMF can have a high entropy which make the choice of a cluster very difficult in many cases. This approach we try to reduce the uncertainty in the matrix W using evolutionary dynamics and taking into account contextual information to perform a consistent labelling of the data. In fact, with our method similar objects are assigned to similar clusters, taking into account the initial solution obtained with NMF. The proposed system conducted an extensive analysis of the performances of our method and compared it with different NMF formulations and on datasets with different features and of different kind. The results of the evaluation demonstrated that approach is almost always able to improve the results of NMF and that when it has negative results those results are practically non-significant. The algorithm is quite

general thanks to the adaptive auto-tuning of the payoff matrix and can deal with balanced and completely unbalanced datasets. As future work are planning to use different initialization of the strategy space, to use new similarity functions to construct the games graph, to apply this method to different problems and to different clustering algorithms.

**REFERENCES**

- [1] A. Banerjee, S. Merugu, I. S. Dhillon, and J. Ghosh, "Clustering with bregman divergences." Journal of Machine Learning Research, vol. 6, pp. 1705–1749, 2005.
- [2] M. Belkin and P. Niyogi. "Laplacian eigenmaps and spectral techniques for embedding and clustering." In NIPS, 2001.
- [3] P. Carmona-Saez, R. D. Pascual-Marqui and A. Pascual-Montano, "Biclustering of gene expression data by non-smooth nonnegative matrix factorization." BMC Bioinformatics, vol. 7(78), pp. 1–18, 2006.
- [4] H. Cho, I. Dhillon, Y. Guan, and S. Sra, "Minimum sum squared residue based co-clustering of gene expression data." SIAM SDM, pp. 114–125, 2004.
- [5] I. S. Dhillon. "Co-clustering documents and words using bipartite spectral graph partitioning." ACM SIGKDD, pages 269–274, 2001.
- [6] N.Mahendran, " Collaborative Location Based Sleep Scheduling with Load Balancing in Sensor-Cloud "International Journal of Computer Science and Information Security (IJCSIS), ISSN: 1947-5500, volume 14, Special Issue, October 2016, PP: 20-27.
- [7] N.Mahendran, "Sleep Scheduling Schemes Based on Location of Mobile User in Sensor-Cloud", International Journal of Computer, Electrical, Automation, Control and Information Engineering Volume 10, No: 3, 2016 PP: 615-620.
- [8] C. HQ, Ding, T. Li and M. Jordan. I. "Convex and semi nonnegative matrix factorizations." TPAMI, 32(1), pp. 45–55, 2010.
- [9] Q. Gu and J. Zhou, "Co-clustering on manifolds." ACM SIGKDD, 2009.
- [10] J. Kim and H. Park., "Sparse nonnegative matrix factorization for clustering." Technical report, Georgia Institute of Technology, 2008.
- [11] T. Li and C. Ding., "The relationships among various nonnegative matrix factorization methods for clustering." IEEE ICDM, pp. 362— 371, 2006.
- [12] F. Shahnaz, M. Berry, P. Pauca, and R. Plemmons, "Document clustering using non-negative matrix factorization." Information Processing and Management, vol. 42, pp. 373–386, 2006.
- [13] A. Strehl and J. Ghosh, "Cluster ensembles - a knowledge reuse framework for combining multiple partitions." Machine Learning Research, pp. 583–617, 2002.
- [14] H. Wang, F. Nie, H. Huang and F. Makedon, "Fast nonnegative matrix tri-factorization for large-scale data co-clustering." IJCAI, 2011.
- [15] R. Zass and A. Shashua, "A unifying approach to hard and probabilistic clustering." IEEE ICCV, pp. 294–301, 2005.
- [16] Z. Zhang and Z. Zha. "Principal manifolds and nonlinear dimensionality reduction via tangent space alignment." SIAM Scientific Computing, Vol 26, pp. 313–338, 2004.
- [17] Z. Y. Zhirong Yang and J. Laaksonen, "Projective nonnegative matrix factorization with applications to facial image processing." Pattern Recognition and Artificial Intelligence, vol. 21(8), pp. 1353– 1362, 2007.
- [18] D. Cai and X. He and J. Han, "SRDA: An efficient algorithm for large-scale discriminant analysis." TKDE, vol. 20, pp. 1–12, 2008.
- [19] L. Lov'asz and M.D. Plummer, "Matching Theory." AMS Chelsea Publishing Series, American Mathematical Soc., 2009.
- [20] R Gomathi, N Mahendran, "An efficient data packet scheduling schemes in wireless sensor networks", in

- Proceeding 2015 IEEE international Conference on Electronics and Communication Systems (ICECS'15), ISBN: 978-1-4799-7225-8, PP:542-547, 2015
- [21] S Vanithamani, N Mahendran, "Performance analysis of queue based scheduling schemes in wireless sensor networks", in proceeding 2014 IEEE international Conference on Electronics and Communication Systems (ICECS'14), ISBN: 978-1-4799-2320-5, PP: 1-6, 2014.
- [22] P Kalaiselvi, N Mahendran, "An efficient resource sharing and multicast scheduling for video over wireless networks", in proceeding 2013 IEEE international Conference on Emerging Trends in Computing, Communication and Nanotechnology (ICECCN'13), ISBN: 978-1-4673-5036-5, PP:378-383, 2013