Enhancing Sparse Adjacency Matrix for Community Detection in Large Networks

Ali Falah Yaqoob * and Basad Al-Sarray
Department of Computer Science, College of Science, University of Baghdad, Baghdad-Iraq.
* Corresponding author: ali.f.yaqoob94@gmail.com

Abstract

This paper presents a problem of community detection in sparse network. Graph represents the network with {0, 1} symmetric matrix, this matrix is defined to be sparse when most of its entries are zeros. The problem of community detection of this type of networks is non-deterministic polynomial-time hardness (NP-hard) problem. Here, we give a simple idea to regularize the sparse matrix by adding a heuristic parameter to the entries of the matrix. This work performs integrating Tabu Search via Fuzzy C-mean to compute variants of the modularity maximization. The results show the ability of the proposed method to define structure of the network by optimizing different types of the quality functions; the results show the global function gives the high value in must runs when apply it on a large sparse real networks.   [DOI: 10.22401/ANJS.22.4.10]

1. Introduction

Networks show up in various topics like, web based life, electrical power systems, correspondence systems, Politic, science and so on. All in all, the structures of the networks are founded by applying the numerical methods, these structures aren't extraordinary, accordingly the vast majority of time the advancement calculations received to get ideal provide for bold depiction of uncover appropriate patterns [1]. Chin et al, 2015 worked on SBM and community detection with sparse graphs, they used spectral algorithm with optimal rate of recovery to analyze a robust spectral algorithm for the SBM model [2]. Newman, 2013 proposed spectral methods based on the eigenvectors of matrices are widely used in the analysis of network data, particularly for community detection and graph partitioning. Standard methods based on the adjacency matrix and related matrices, however, break down for very sparse networks, which includes many networks of practical interest [3]. Qin et al, 2013 presented the regularized spectral under the degree corrected Stochastic Block Model (DCSBM), because of the important of spectral clustering in finding clusters in networks [4]. Rohe et al, 2011 worked on the spectral clustering with assumption of high dimensional Stochastic Block Model (SBM). They studied the asymptotic properties of clustering by spectral clustering under the SBM, where they found the "eigenvectors of the normalized Laplacian graph asymptotical convergence to the eigenvectors of the normalized Laplacian graph" [5]. Clustering strategies are utilized to infer the parcel of finite points into groups, with presumption the items in a similar cluster are more like each other than points in various clusters as indicated by some characterized criteria. It is an unsupervised critical procedure ordinarily uses to characterize the structures in information. The principal idea of grouping procedures is finding the clusters which can give important description of structure of the distribution of points in groups. The communities in networks define the structure in network, with the goal that the level of affiliation is viewed as solid for data lying in the same block of the segment and feeble for element in other blocks. Network density describes the strongest of relations among the elements of a network. Dense social networks for coupled with strong ties inside the group from others. A network is defined based on two sets: a set of nodes (vertices) that represent the selected elements, and a set that represent ties between limits. They define a graph, the lines can be directed, or undirected edge [6].

The rest of the paper, section.2 proposed regularization operator, section.3 presents

modularity optimization and introducing five function of the quality functions, section.4 proposed algorithm Tabu search based Fuzzy c-Means, section. 5 results and discussion of the most observations and analysis of metric that used to measure the ability of quality function to find the optimal partitions.

2. Regularization of sparse network Adjacency matrix.

Since the concentration of sparse random networks fails because the degree distribution is too irregular, so, the question is : Can the regularization of the network in some way solves the problem?.

One simple way to deal with very low-degree nodes, proposed by [7] the idea is to add the same small positive number τ to all entries of the adjacency matrix Λ. Where

$$\Lambda_\tau := \Lambda + \frac{\tau}{N} 11^T \ \text{...............................} (1)$$

by using the Laplacian of Λ_τ instead. Another way to manage low degree nodes, proposed by [8] and is to add a constant τ directly to the diagonal of D in the definition of the Laplacian. Chaudhuri 2012 defined the regularizing Laplacian graph as

$$L_\tau = D_\tau^{-\frac{1}{2}} A \ D_\tau^{-\frac{1}{2}} \in R^{N \times N} \ \text{....................} (2)$$

Where $D_\tau = D + \tau I$, $for \ \tau \geq 0$.

This technique divides the nodes into two random subsets and only uses the induced sub-graph on one of those random subsets to compute the spectral decomposition.

The steps of regularized spectral clustering algorithm are [13], [14].

Algorithm.1 Pseudocode of Laplacian graph regularization
Input: adjacency matrix Λ, number of clusters C, regularize parameter τ.
Output: $T_1, \dots, T_C$, *node i assigned to cluster r if the i'th row of V^* is assigned to T_r.*
Processes: Begin

1. Implement the adjacency matrix Λ, define number of clusters C, regularize parameter τ, Find the regularized graph based on τ.
2. Evaluate the eigenvectors $[v_1, \dots, v_C] \in R^{N \times N}$, contains the C largest eigenvalues of regularized graph, by embedding the $[v_1, \dots, v_C] \in R^{N \times N}$ into the columns.
3. Form the matrix $V^* \in R^{N \times C}$ from V by normalizing the rows of V by projecting each row of V onto the unit sphere, $V_{ij}^* = \frac{v_{ij}}{\Sigma_j (V_{ij}^2)^{\frac{1}{2}}}$.
4. Row of V^* in $R^{N \times C}$, and computing centers of the cluster.
   To create C non-overlapping sets $T\_1, \dots, T_C$ with union in T
5. The mean idea is: find the principal eigenspace of the (regularized) Laplacian graph; then determine the clusters in the low dimensional eigenspace.

End

There are two main kinds of social network data: Socio-matrix (Adjacency Matrix) use to represent directed or undirected ties between actors using a numerical matrix. There is one column for each actor and one row for each actor.

An Edge list is the other primary form of data storage for social network analysis.

This only captures data about existing ties so it needs to be supplemented with knowledge of the total number of actors in the network (even if they do not have any ties).

3. Proposed heuristic operator
The idea of community detection for sparse networks studied by Can M.Le, where he used the Laplacian graph concentration with regularized dataset under community detection [9].

To define community in sparse networks we introduce a heuristic operator as a regularization parameter by adding it adjacency matrix of the sparse network.

Let Ξ be a regularization parameter that provides high-probability bounds on the spectral norm. This operator defined by

$$\Xi: [0,1] \rightarrow [0,1], \ \text{s.t} \ \Xi = \xi \ (n-1) \ v \ \psi \ \text{.................................} (3)$$

Where ξ is between $\{0,1\}$, n is size of adjacency matrix, v is a random real number and ψ is an $n \times 1$ vector with ones. Ξ adding to the entries of the adjacency matrix Λ, thereby replacing by Λ_Ξ where,

$$\Lambda_{\Xi} = \Lambda + (\Xi / N)I^{T} \dots\dots\dots\dots\dots\dots\dots\dots (4)$$

This regularization will solve the problem of low-degree vertices.

4. Modularity maximization

The graph is used for representing the network, *cluster* used for *community*; *clustering* for defining community *detection*.

Let $\mathcal{G} = (\mathcal{T}, \mathcal{S})$ be a *graph* consists of a *node set* $\mathcal{T}$ and an *edge set* $\mathcal{S}$. The nodes contained in $\mathcal{T}$ will always be labelled as $1, 2, 3, \dots, \mathcal{N}$. the *external* edges of $\mathcal{T_C}$, which have one end in $\mathcal{T_C}$ and the other end in $\mathcal{T} - \mathcal{T_C}$, the set of nodes which do not belong to $\mathcal{T_C}$. Community detection of $\mathcal{G} = (\mathcal{T}, \mathcal{S})$ is the activity of graph clustering, finding a partition $\mathcal{T} = \{\mathcal{T}_1, \dots, \mathcal{T}_K\}$ of a graph into clusters; the nodes contained in each cluster must somehow be more related to each other than to nodes outside the cluster, thus forming a community.

A quality function is a function $\Upsilon(\mathcal{G}, \mathcal{T})$ (i.e., it depends on both the graph $\mathcal{G}$ and the partition $\mathcal{T}$) the value of which characterizes how good V is as a partition of $\mathcal{G}$.

Hence the best decomposition of $\mathcal{G}$ into communities is the partition $\mathcal{T}^* = \{\mathcal{T}_1^*, \dots, \mathcal{T}_k^*\}$ which maximizes Υ, i.e.

$$\mathcal{T}^* = \arg\max_{\mathcal{T}} \Upsilon(\mathcal{G}, \mathcal{T}) \dots\dots\dots\dots\dots\dots\dots (5)$$

And then good communities are the elements of $\mathcal{T}$ which achieves a high $\Upsilon(\mathcal{G}, \mathcal{T})$ score. The most popular quality function are: For $c \neq 1$, M_c defined to be the sum of node degrees in community c, and μ refer to the sum of all degree in the network. M_{cc} represents the twice the total number of edges within community c, and μ^2 is twice the number of edges in the network. [12].

Table (1)
Quality function for optimizing community detection in real networks.

Function	Mathematical description	description
Girvan-Newman	$\Upsilon_{GNM}(\mathcal{G}, \mathcal{T}) = \sum_{c=1}^{C}\left(\frac{S_c^i}{2m} - \left(\frac{m_c}{2m}\right)^2\right)$	$m_c = \sum_{x \in T_c}\sum_{y \in T} \Lambda_{xy}$ total degree of cluster T_c [6]
Weighted global density	$\Upsilon_{Wglob} = \sum_{c=1}^{C}\frac{1}{2}\cdot\left[\Upsilon^i(T_c, \mathcal{G}) + 1 - \Upsilon^e(T_c, \mathcal{G})\right]$	$\Upsilon^i(\mathcal{G}, \mathcal{T}) = \frac{\sum_{c=1}^{C}\sum_{i \in T_c}\sum_{j \in T_c}^{C}\Lambda_{ij}}{\sum_{c=1}^{C}|T_c|^2}$, global internal densities $\Upsilon^e(\mathcal{G}, \mathcal{T}) = \frac{\sum_{c=1}^{C}\sum_{i \in T_c}\sum_{j \in T - T_c}^{C}\Lambda_{ij}}{\sum_{c=1}^{C}|V_c||\mathcal{T} - \mathcal{T_c}|}$ global external densities [6]
Local density quality	$\Upsilon_{Wlocal}(\mathcal{G}, \mathcal{T}) = \sum_{c=1}^{C}\frac{|T_c|}{2|T|}\cdot\left[\Upsilon^i(T_c, \mathcal{G}) + 1 - \Upsilon^e(T_c, \mathcal{G})\right]$	$\Upsilon^i(\mathcal{G}, \mathcal{T}) = \frac{\sum_{i \in T_c}^{C}\sum_{j \in T_c}^{C}\Lambda_{ij}}{|T_c|^2}$, local inner densities $\Upsilon^e(\mathcal{G}, \mathcal{T}) = \frac{\sum_{i \in \mathcal{T_c}}\sum_{j \in \mathcal{T} - \mathcal{T_c}}\Lambda_{ij}}{|V_c||\mathcal{T} - \mathcal{T_c}|}$, local outer densities [6]
Distance based quality	$\Upsilon_{Dist}(\mathcal{G}, \mathcal{T}) = \frac{1}{|\mathcal{T}|^2}\|\Lambda_{\mathcal{G}} - \Lambda_{\mathcal{T}}\|$	$\|.\|$ is a matrix norm. $\Lambda_{\mathcal{G}}$ is the adjacency matrix of $\mathcal{G}$ [6]
Node membership quality	$\Upsilon_{NM}(\mathcal{G}, \mathcal{T}) = \frac{1}{2|\mathcal{T}|}\sum_{i \in V}[\varsigma(x - \mathcal{T}[x]) + 1 - \varsigma(i, \mathcal{T} - \mathcal{T}[i])]$	$[x]$ indicates the cluster to which $x \in \mathcal{T}$ and $\Lambda \subseteq \mathcal{T}$ [6]

5. Integrating Tabu Search vie Fuzzy C-Mean for community detection in sparse networks.

Modularity is a metric for measuring the best partition that used to define the structure of the network, where the modularity with high value refers to defining a good community structure of the network, the partition corresponding to high value of modularity is defined with highest quality. Discovering the communities is getting by maximizing the modularity over all possible partitions of a network. However, it is computationally prohibitively expensive to exhaustively search all such partitions for the optimal value of modularity since modularity optimization is known to be NP-hard. A lot of researches introduced heuristic methods to find high-modularity partitions in a reasonable time. Tabu Search (TS) is an effective heuristic algorithm, it was developed by Fred Glover in 1988. [10]

Fuzzy C-Mean algorithm is trying to divide a finite n elements $\sigma = \{ \sigma_1, \ldots, \sigma_n \}$ into a collection of c fuzzy clusters with respect to some given criterion [11]. Given a finite set of data, the algorithm output is a set of the K cluster centres $O = \{ o_1, \ldots, o_K \}$, the matrix of the partition which is denoted by $\Omega = \{ \omega_{i,j} \in [0,1], \ i = 1, \ldots, n, j = 1, \ldots, o \}$, here $\omega_{i,j}$, refers to the degree to which element, σ_i belongs to cluster c_j. FCM aims to minimize an objective function. Clustering is the process of grouping feature vectors into classes in the self-organizing mode. Let $\{ \sigma^{(\iota)} \colon \iota = 1, \ldots, \lambda \}$ be a set of λ feature vectors, where each feature vector $\sigma^{(\iota)} = \left(\sigma_1^{(\iota)}, \ldots, \sigma_N^{(\iota)} \right)$ has N components. The process of clustering is to assign the feature vectors into K clusters usually by the minimum distance assignment principle. FCM uses the weights that minimize the total weighted mean-squares errors:

$$I = \sum_i^{mp} \sum_j^{nc} |\sigma_i - o_j|^2 . w_{ij}^m \ \ldots\ldots\ldots\ldots\ldots (6)$$

The centroid of each cluster, the membership of a pattern to a cluster given in equations below

$$o_j = \frac{\sum_i^{mp} \sigma_i . \omega_{ij}^m}{\sum_i^{mp} \omega_{ij}^m}, \quad w_{ij} = \frac{\left(\frac{1}{|\sigma_i - o|^2} \right)^{\frac{1}{m-1}}}{\sum_k^{nc} \left(\frac{1}{|\sigma_i - o_j|^2} \right)^{\frac{1}{m-1}}}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7)$$

Algorithm. 2 Computing centers of the cluster.

Input: Number of clusters C
Output: Centroid for each cluster C, coefficients of each points of being in the clusters
Processes: Begin
Step.1. Determine C number of clusters.
Step.2. Assign randomly to each point coefficients for being in the clusters.
Step.3. Repeat until the maximum no. of iterations is reached, or when the getting the condition of convergences.
End

Tabu search requires the following basic elements to be defined
1. **Configuration**: generate the initial solution or an assignment of values to variables.
2. **Move**: a move is a finding a new trial solution is generated from the current one.
3. **Set of candidate moves**: is the set of all possible moves out of a current configuration. If this set is too large, one could operate with a subset of this set.
4. **Tabu restrictions**: these are certain conditions imposed on moves which make some of them forbidden. These forbidden moves are known as Tabu. It is done by forming a list of a certain size that records these forbidden moves. This is called the Tabu.list.
5. **Aspiration criteria**: these are rules that override Tabu restrictions, if a certain move is forbidden by Tabu restriction, then the aspiration criterion, when satisfied can make this move allowable.

Algorithm.3 Pseudocode of integrated Tabu search via Fuzzy C-Means clustering method

Initialization: Generate centroid of data, using the data as a parameters of fuzzy c-mean function by call **Algorithm.2** above.
Processes: Begin
Note*: [] represents as an empty array.

Step. 1. TabuList $=$ []
Step. 2. Insert regularization parameters on adjacency matrix
Step. 3. while (Iters $\leq$ MaxIter)do
Step. 4. Candidate solution List $=$ [   ]
Step. 5. For (Sol. best in Sol. Optim region

(neighborhood))

If $\big($there isn't any features in $($Sol. new, TabuList$)\big)$

Define Sol. new constraint Sol. new list

Step. 6. Sol. new get from Locate Best

   Candidate(Candidate List)
   Step. 7. If (Quality(S. new) $<$
            $=$ Quality(Sol. Optim))
      Sol. Optim $=$ Sol. new ,
 Tabu List Feature Differences(S. new, S. opt)
Step. 8. While (Tabu List $>$ Tabu List Size),
      Delete Feature (Tabu List)
End while
Step. 9. Return (Sol. Opt)
Step. 10. Iters $=$ Iters $+$ 1 ;
End for

Step Output: **Best Quality of tabu search with Fuzzy c − mean clustering**

End

6. Results
6.1 General Framework
   The exterminates are designed with some fixed parameters and some of other parameters are changing according to requirements and type of the algorithms.

1. **Data-pre-processing**: The first step in implementing the experiments are treating the data, there are many challenges of using complex networks data. These challenges are:
   a. Isolated nodes and the duplicate edges,
   b. duplicated edges by using degree of nodes.
   In order to fix these problems, we are working on removing these isolated nodes duplicated edges.
2. **Graph Generators**: Different graph generators require different graph parameters as input but there is one input which must be specified for every graph generator, we have to refer to whether the

adjacency matrix Λ of the graph will contain zeros or ones in the diagonal (these correspond to the absence or presence of self-loops for the graph nodes).
3. Graph Clustering: Every graph clustering function requires as minimum input a description of the graph to be clustered.
4. **Cluster Number Selection**: The cluster number selection criterion applied on Newman-Girvan modularity $Y_{GNM}(G, \mathcal{T})$ for $c = 1, 2, \dots, C$; the function returns as output the integer C iff $\mathcal{T}^C$ is the clustering with highest modularity. Number of cluster is necessary to study the behavior of the network, in this study we work with two kinds of assumption on number of clusters, $C = 2$, and C auto-select by defining range of C and the best optimal will be adopted as the best partition via number of metrics.
5. **Partition Similarity**: These functions are used to compute the similarity between two clustering (partitions) $\mathcal{T}_1$ and $\mathcal{T}_2$
6. **Quality function**: While partition similarity functions evaluate a partition by comparing it to the true partition, quality functions evaluate a partition $\mathcal{T}$ for which.
7. **Choosing the Evaluation Function**: This function will evaluate the Jaccard similarity between V (the clustering obtained by the graph clustering algorithm) and $\mathcal{T}_0$ which is the reference clustering, returned from the Graph Generator function.
   Here, the results using Newman's networks NMI is the normalized mutual information index of partition (clustering) similarity. It takes values in the interval (0, 1). The maximum value 1 indicates maximum similarity, i.e., identity. The sequential of the proposed algorithms are given respectively by the following sections.

6.2 Integrating TS for modularity maximization
   There are many of challenge that are appearing during the analysis of complex networks. The degree of nodes isolated or duplicated, the treatment for two cases are: remove the isolated nodes when $Degree\ (G) = 0$ and duplicated edges. This section deals with the problem of community detection in large networks by integrating TS with FCM.

First method is TS-FCM method: First step is generate the regularization of adjacency matrix and then it with the data to integrate Tabu search with FCM to find the best quality functions that give the optimal partitions of the graph which represents the network. The maximum values of modularity computing by applying Fuzzy c-mean to determine the initial centroid of the clusters.

Second method is FCM-TS method: First step is generate the centroids by FCM to integrate and run the regularization of adjacency matrix. Then take the regularized adjacency matrix with the centroids of FCM to integrate Tabu search with FCM. This section shows the results of applying the proposed algorithm on the real sparse network, the details of the data given in Table 2. There are many of challenge that are appearing during the analysis of complex networks. The degree of nodes isolated or duplicated, the treatment for two cases are: remove the isolated nodes when $Degree\ (\mathcal{G}) = 0$ and duplicated edges. After the step of pre-processing, is using integrating Tabu search via FCM with Euclidean distance to optimize the quality

functions modularity. The regularization of the network is performing by applying the heuristic operator on the adjacency matrix of the real networks given by Equation (4) , the number of clusters are auto selecting based on the maximum values of the modularity function(the objective functions).

a. Internet: 6473 nodes and 11284 edges

b. Political Blogs: 1490 nodes and 9537 edges

c. Twitter: 9900 nodes and 21000 edges

d. Protein : 2361 nodes and 9537 edges

e. Facebook: 4039 nodes and 84241edges

Fig.(1): The different networks representing by graph.

Table (2)

Networks Details.

Dataset	No. of nodes	No. of isolated Nodes	No. of nodes after deletion of isolated nodes
Facebook	4039	80	3958
Protein	2361	77	2284
Political	1490	382	1107
Internet	6473	29	6444
Twitter	9900	6728	2623

Table (3)
Comparison of Tabu search Algorithm with Fuzzy C-mean.

Dataset	Best-Quality of TS-FCM for 5 Quality functions	Best-Quality of FCM using TS-FCM	FCM obj.fun by TS	Best-Quality of FCM-Ts for 5 Quality functions
Facebook	0.8465	0.000143	15039982.662834	0.98350
Protein	0.5931	0.000024	616401.035692	0.99665
Political	0.5159	0.000065	470152.707395	0.97955
Internet	0.8445	0.000078	20887.637974	0.98945
Twitter	0.83526	0.000043	355626.54362	0.83528

a. FCM objective value   **b. FCM-TS Best quality**   **c. Quality functions comparison**
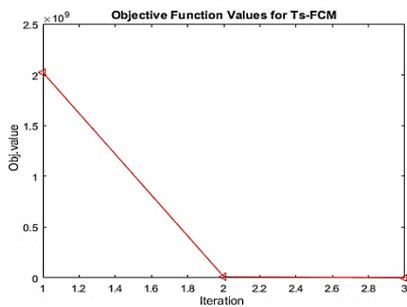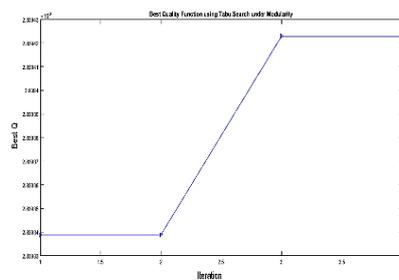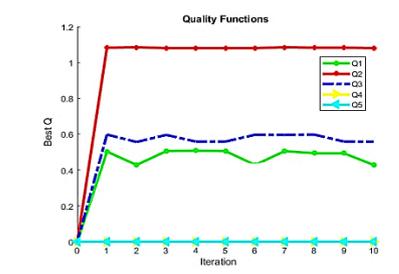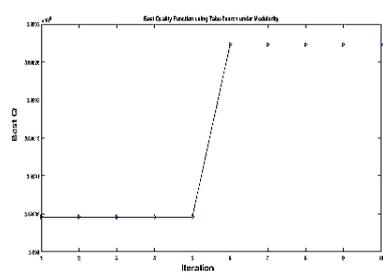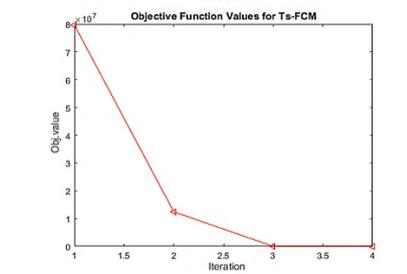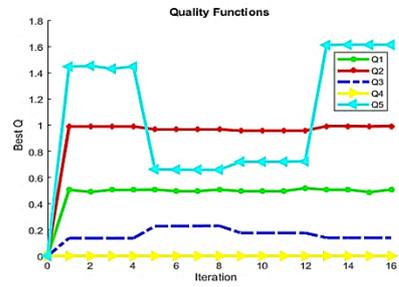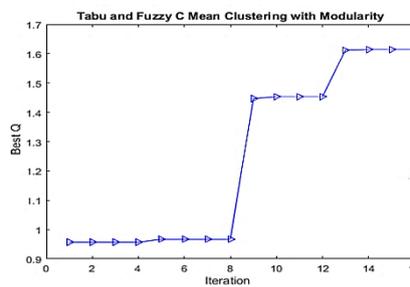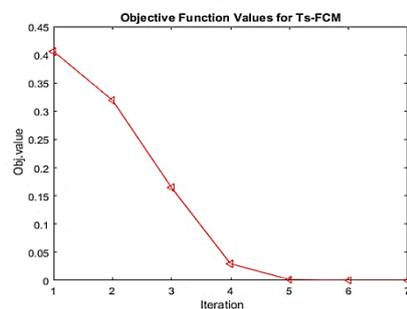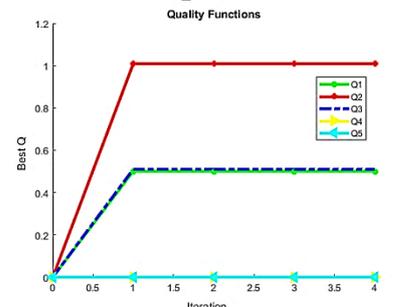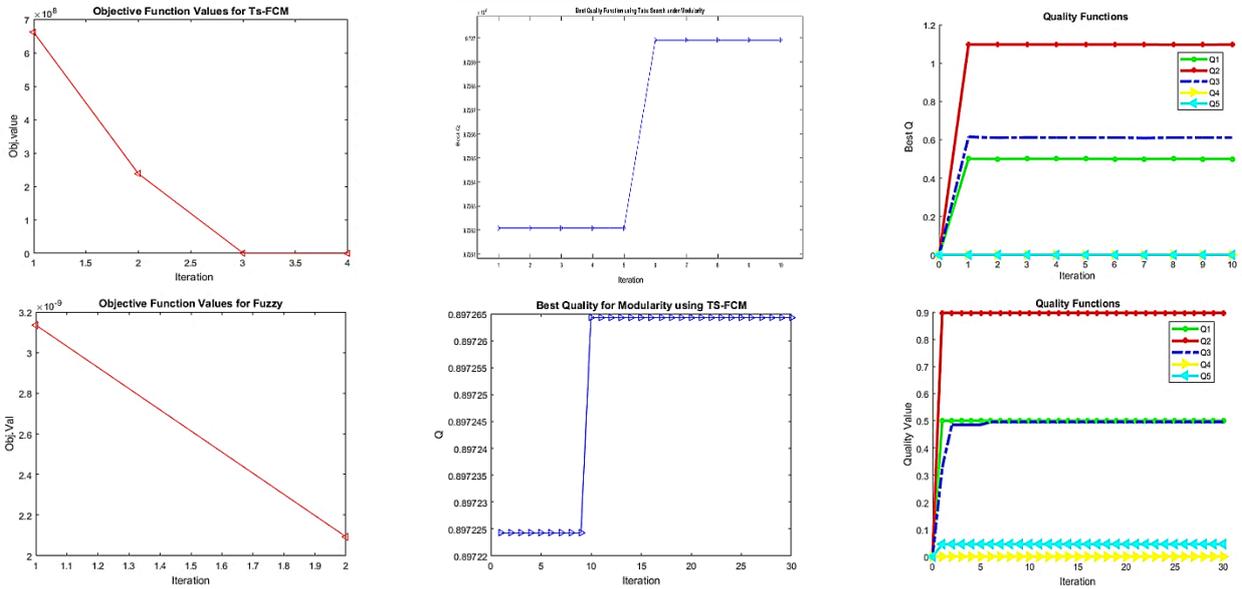
Fig.(2): ITS-FCM objective function with respect to number of iterations: Internet AS-Level Network, Facebook, Political Blogs Network, Protein , and Twitter Network.

a. **Facebook network**

b. **Internet AS-Level**

c. **Political Blogs network**

d. **Protein network.**

e. **Twitter network**
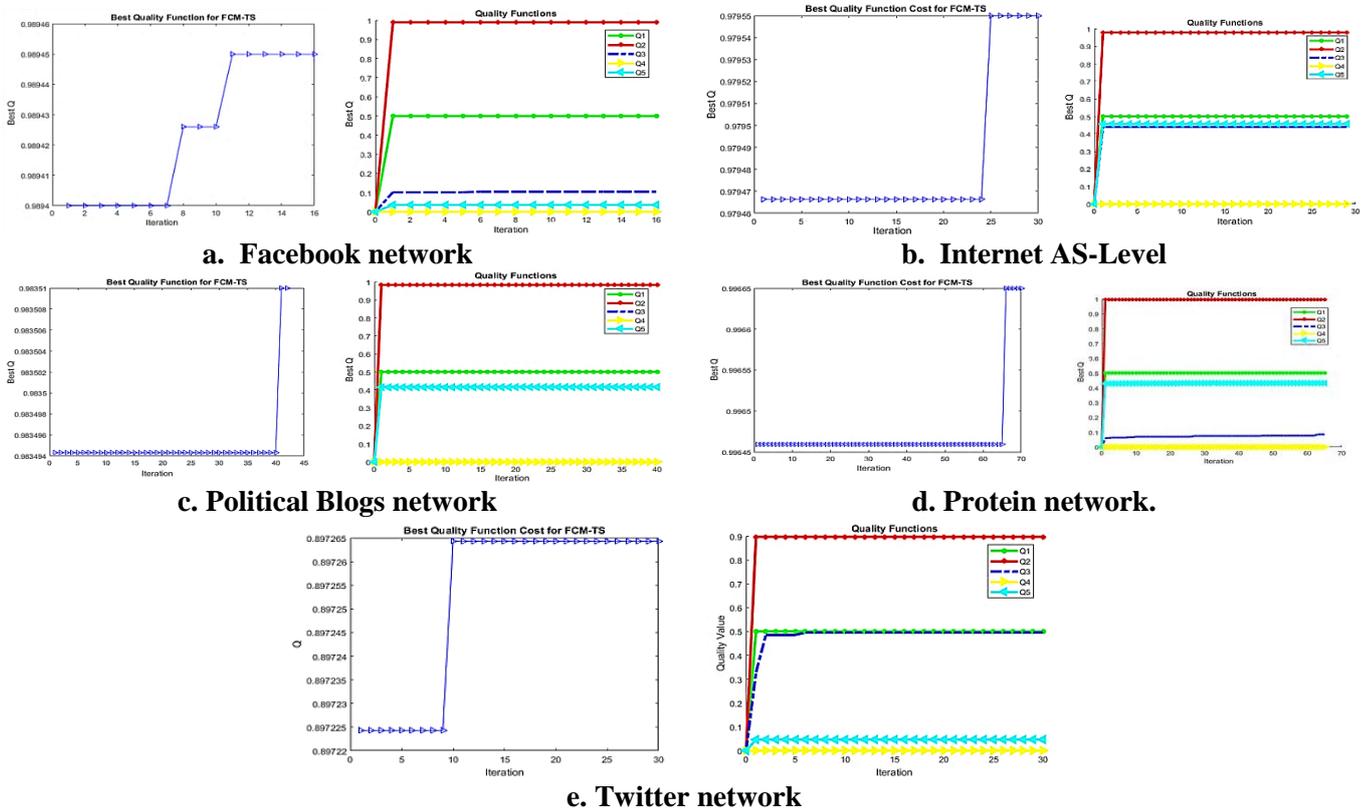
Fig.(3): ITS-FCM for objective function values.

6.3 Performance Comparison with Partition Selection Measurements

Partition selection measurements normally uses to evaluate the community detection methods, general steps are given by the following:

a. **Normalized Mutual Information :** normalized mutual information criteria is one of partition selection measurements on community detection. We use NMI measurement with all datasets that for finding the similarity between two partitions of all datasets.

b. **Jaccard Measurement** : We used Jaccard criteria for finding the similarity between two partitions of all Networks.

c. **Random Index Measurement** : We used Random Index criteria for finding the similarity between two partitions of all Networks.

d. Relative Number of Clustering Error Measurement : We use Relative Number of Clustering Error criteria for finding the similarity between two partitions of all networks. All networks give the same quality equal to 0.5, but $\Upsilon_{obj.1}$ gives zeros when V_1 has the same number of clusters as V_2.
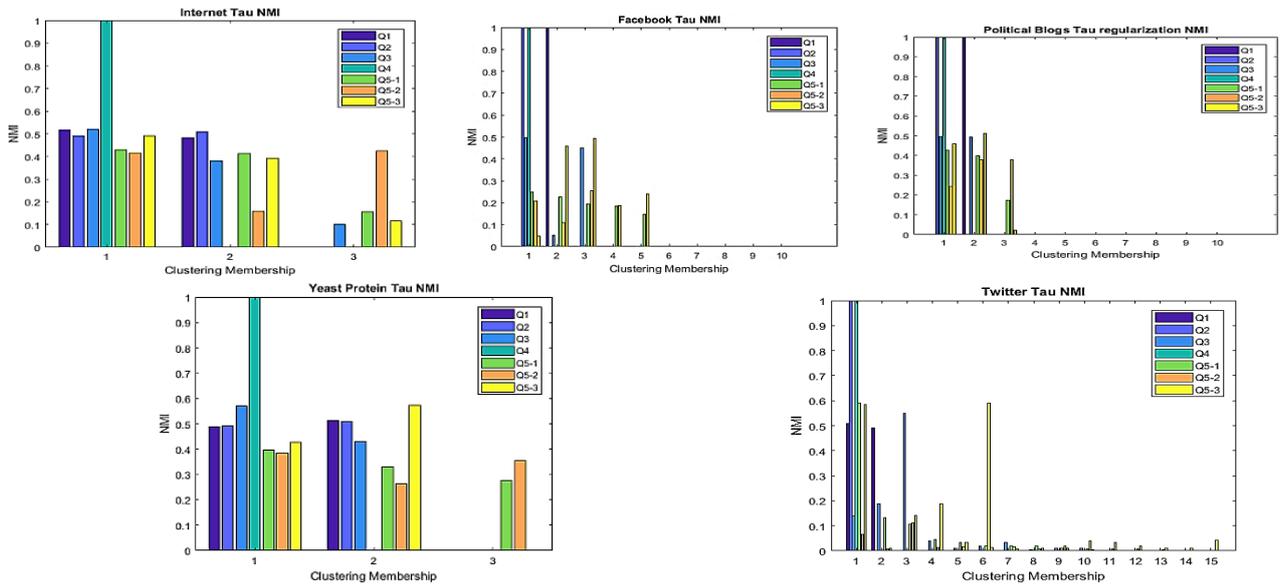
Fig.(4): NMI partition selection criteria for Internet , Facebook, Political Blogs, Protein, Twitter networks.

a. **Jaccard partition selection criteria**

b. **Random index partition selection criteria**

c. **Relative Cluster number Error partition selection**
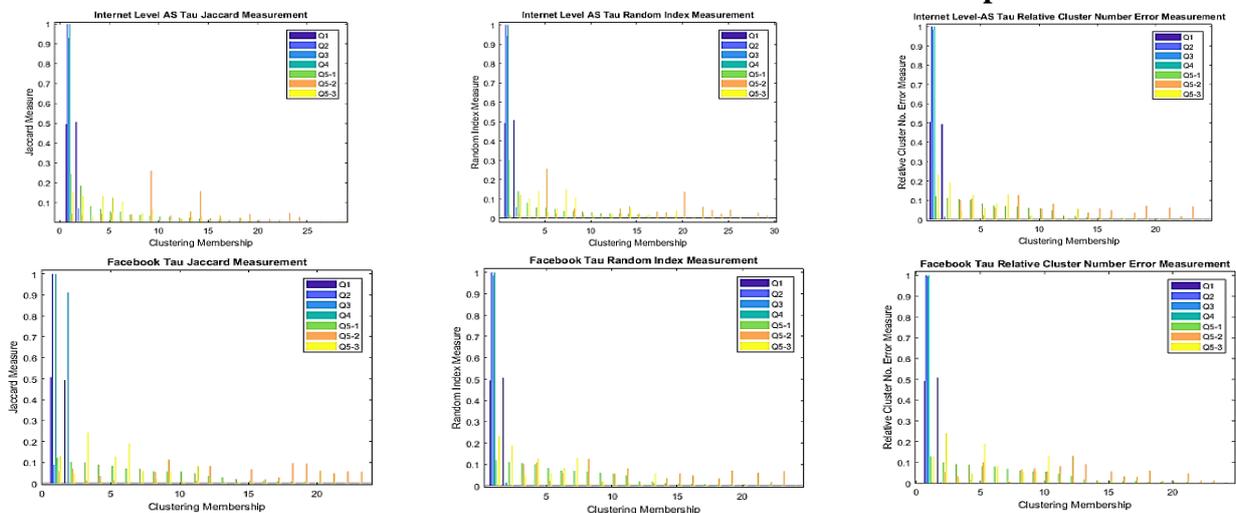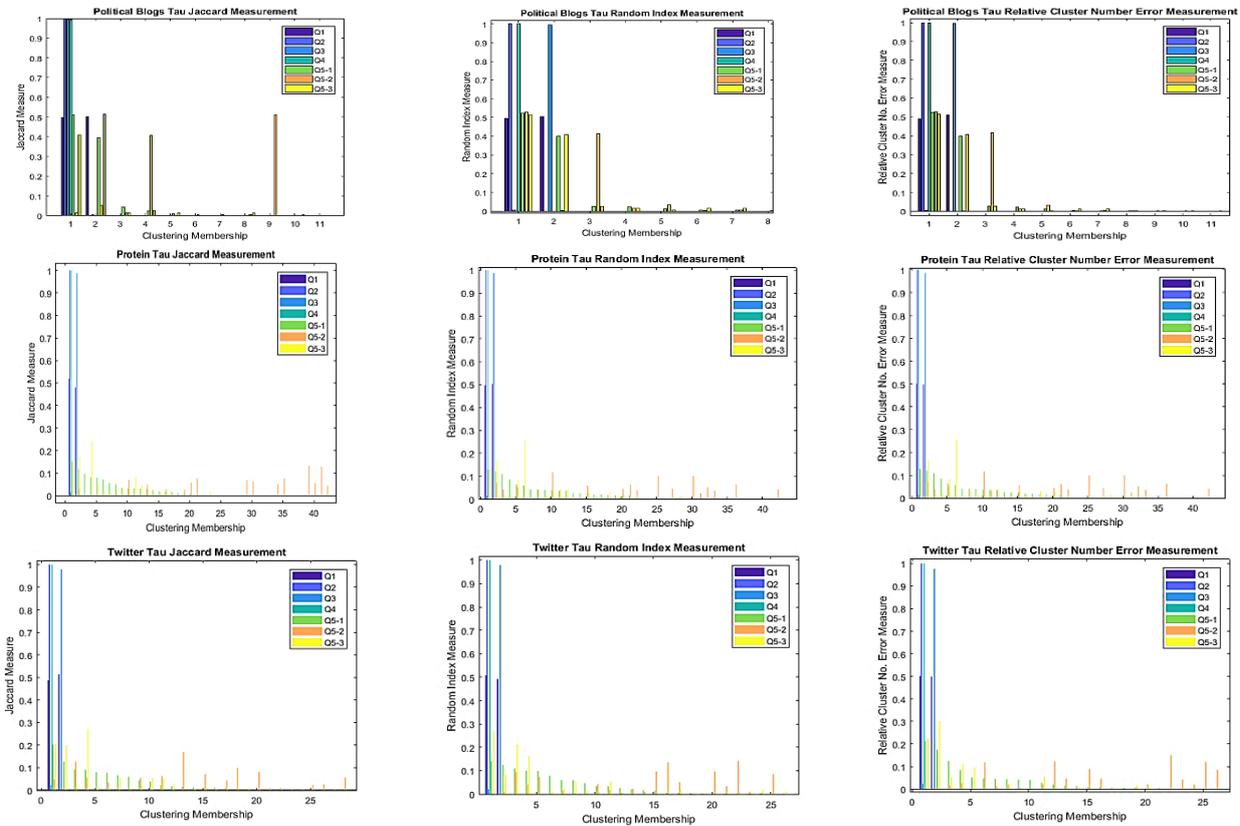
Fig.(5): Another methods of partition selection criteria for Internet, Facebook, Political Blogs, Protein, and Twitter networks.

6.4. Discussion

In this work, Fig.(2) represent the first method experiment of Integrating of Tabu search with Fuzzy c-mean after the regularization. The part (a) of Fig.(2) represents Fuzzy objective value, part (b) represents the best quality function of integrating Tabu with FCM, and part (c) represents the comparison of 5 quality objective functions of modularity for all networks that presented in Fig.(1). Fig.(3) represent second method experiment of Integrating TS-FCM. Fig.(4) represents an NMI partition selection criteria to find the similarity between two partitions of all networks and Fig.(5) represent the experiment of other partition selection methods to find the similarity between two partitions of all networks. Fig.(4) and (5) shows in the light lines in many networks because these networks contain a high quality of clustering membership for similarity of partitions.

7. Conclusion

Network classified into dense network and sparse networks, finding community in sparse networks is a challenge that defines the structure of the network, so we define heuristic operator called regularization operator of the sparse network, and apply it on special kinds of networks such as Protein, Facebook twitter, and political blogs. This paper deals with modularity for community detection in sparse real networks, we introduce a regularization parameter for regularize the data network and apply an integrating Tabu vie Fuzzy to optimize a five types of the quality function to find community detection of the sparse network. The results show the ability of the algorithm to find maximum value of the quality functions.

References
[1] Fortunato, Santo. "Community detection in graphs." Physics Reports 486.3: 75-174, 2010.
[2] Chin, Peter and Rao, Anup and Vu, Van. "Stochastic block model and community detection in sparse graphs: A spectral algorithm with optimal rate of recovery". Conference on Learning Theory, 391- 423, 2015.

[3] Newman, M. E. J. "Spectral community detection in sparse networks." arXiv preprint arXiv: 1308.6494, 2013.

[4] Qin, Tai and Rohe, Karl. Regularized spectral clustering under the degree-corrected stochastic block model Advances in Neural Information Processing Systems. 3120-3128, 2013.

[5] Rohe, Karl and Chatterjee, Sourav and Yu, Bin and others .Spectral clustering and the high-dimensional stochastic block model. The Annals of Statistics. Institute of Mathematical Statistics, 39 (4), 1878—1915, 2011.

[6] Newman, Mark EJ, and Michelle Girvan. "Finding and evaluating community structure in networks." Physical review E 69,(2) , 026113, 2004.

[7] Amini, Arash A., et al. "Pseudo-likelihood methods for community detection in large sparse networks." The Annals of Statistics, 41.4: 2097-2122, 2013.

[8] K. Chaudhuri, F. Chung, and A. Tsiatas. Spectral clustering of graphs with general degrees in the extended planted partition model. Journal of Machine Learning Research Workshop and Conference Proceedings, 23:35.1 – 35.23, 2012.

[9] Le, Can M. Estimating Community Structure in Networks by Spectral Methods. PhD thesis 2016.

[10] Glover, F., & Laguna, M. Tabu search". Springer US, 2013.

[11] Dunn, J. C. "A fuzzy relative of the isodata process and its use in detecting compact well separated clusters". Journal of Cybernetics, 3 (3), 32-57, 1973.

[12] Newman, Mark EJ. "Fast algorithm for detecting community structure in networks." Physical review E 69.6: 066133, 2004.

[13] T. Qin and K. Rohe. Regularized spectral clustering under the degree corrected stochastic block model. In Advances in Neural Information Processing Systems, 3120–3128, 2013.

[14] A. Joseph and B. Yu. Impact of regularization on spectral clustering. arXiv: 1312.1733, 2013.