Iranian Journal of Mathematical Sciences and Informatics
Vol. 16, No. 1 (2021), pp 105-121
DOI: 10.29252/ijmsi.16.1.105

# Graph Clustering by Hierarchical Singular Value Decomposition with Selectable Range for Number of Clusters Members

Azam Sadeghian, Seyed Abolfazl Shahzadeh Fazeli*, Seyed Mehdi Karbassi

Department of Mathematics, Yazd University, Yazd, Iran.

E-mail: a_sadeghian@stu.yazd.ac.ir
E-mail: fazeli@yazd.ac.ir
E-mail: smkarbassi@yazd.ac.ir

Abstract. Graphs have so many applications in real world problems. When we deal with huge volume of data, analyzing data is difficult or sometimes impossible and clustering data is a useful tool for these data analysis. Singular value decomposition(SVD) is one of the best algorithms for clustering graph but we do not have any choice to select the number of clusters and the number of members in each cluster. In this paper, we use hierarchical SVD to cluster graphs to desirable number of clusters and the number of members in each cluster. In this algorithm, users can select a range for the number of members in each cluster and the algorithm hierarchically cluster each clusters to achieve desirable range . The results show in hierarchical SVD algorithm, clustering measurement parameters are more desirable and clusters are as dense as possible. In this paper, simple and bipartite graphs are studied.

**Keywords:** Graph Clustering, Singular Value Decomposition, Hierarchical Clustering, Selectable Clusters Number.

**2010 Mathematics subject classification:** 68W99

## 1. Introduction

Clustering data is one of the best preprocessing for data mining because in clustering data, data with same properties are grouped together and decision for

*Corresponding Author

Received 04 February 2018; Accepted 18 Jun 2019

these groups is made easier. In some applications, data are modeled as graphs. For graph clustering, there are some methods like hierarchical methods, partitioning methods and density based methods. One of which is by using SVD. In most existing clustering methods, users do not have any choice for the number of clusters and the number of members in each cluster. In some applications, the number of clusters should be fix or the number of members in each cluster should be almost equal. In this paper, by using SVD method and hierarchical methods, graphs can be clustered to desirable number of clusters and there is a range for the number of members in each cluster. In Section 2, graph clustering methods are described and some graph measurement parameters are defined. In this section, also clustering with SVD and its limitations are described. In Section 3, a new method by using SVD sign method merging with hierarchical clustering method is defined to cluster graphs. Complexity of this algorithm is also derived. In Section 4, several data for simple and bipartite graphs like Facebook data are clustered with SVD and hierarchical SVD methods. Some conclusion remarks are finally drawn in Section 5.

## 2. Graph Clustering

2.1. **Graph and Graph Clustering.** Graphs can be used to model so many practical problems such as social networks, bioinformatics data, recommendation systems and information systems[11]. For example in social networks, each vertex can be assigned to a member and relations between people are shown as edges. Formally, in graph $G(V, E)$, $V$ is a set of vertices and $E$ is a set of edges between the vertices in $V$ such that:

$$E \subseteq \{(u, v)|u, v \in V\}.$$

A graph is connected when there is at least one path between every pairs of $u, v \in V$ with edges in $E$.
A graph is dense if the number of its edges are close to maximum number of impossible edges. The maximum number of edges in a graph with vertices $V$ and edges $E$ is:

$$\frac{|V|(|V|-1)}{2} \tag{2.1}$$

and density parameter for this graph is defined as:

$$D = \frac{2|E|}{|V|(|V|-1)} \tag{2.2}$$

For a simple graph with vertex set $V$, the adjacency matrix $A$ is a square $|V| \times |V|$ matrix such that the elements $A_{ij}$ is either one when there is an edge from vertex $i$ to vertex $j$, or zero when there is no edge. So adjacency matrix for simple graph is symmetric.

A bipartite graph is a graph whose vertices can be divided into two disjoints and independent sets $V_1$ and $V_2$ such that every edge connects a vertex in $V_1$ to a vertex in $V_2$. When relations between two different classes of objects are needed, bipartite graphs are used. For example, relations between drugs and diseases or mines in countries.

The adjacency matrix $A$ for a bipartite graph can be written in the form:

$$A = \begin{bmatrix} 0 & B^T \\ B & 0 \end{bmatrix}$$

where submatrix $B$ is a rectangular matrix $|V_1| \times |V_2|$ and asymmetric.
When we deal with huge number of data, analysis of data is difficult and sometimes impossible. In these cases, clustering of data is the best preprocessing method to uniform data. Grouping data with same properties into some subsets called clustering and each group is called a cluster. In each cluster similar data are together while different data are in other clusters. For example in graph clustering, vertices with same connectivity go together[9].
Formally, clustering a graph $G(V, E)$ is represented as a set of subsets $C = \{C_1, C_2, \cdots, C_k\}$ such that:

$$C_i \subset V, \forall i = 1, \cdots, k$$

$$V = \bigcup_{i=1}^{k} C_i$$

$$C_i \bigcap C_j = \phi, \forall i \neq j$$

For bipartite graph, it is often necessary to cluster vertices $V_1$ or $V_2$. For example, in related graphs between countries and mines, often countries should be clustered. So in matrix $B$, for clustering $V_1$, rows of $B$ should be clustered and for clustering $V_2$, columns of $B$ should be clustered. Because SVD can be used for all types of matrices, B can be clustered by SVD method.

2.2. **Clustering Methods.** The most well-known clustering methods are hierarchical methods, partitioning methods, density based Methods, model-based clustering methods and fuzzy clustering. In following some methods are described:
**1) Hierarchical Methods:** These methods, cluster recursively data in either top-down or bottom-up fashion. Subdivides of these methods are as follow:
*Merging hierarchical clustering:* Each object initially is a cluster with one member, then clusters are merged together until the desired clustering is obtained.
*Divisive hierarchical clustering:* All objects initially belong to one cluster. Then

each cluster is divided into sub clusters until the desired clustering is obtained.
**2) Partitioning Methods:** In partitioning methods, algorithms start with an initial clustering and by using an optimization scale tries to replace members between clusters. In this method, the numbers of partitions are fixed and subdivides of this method are different by optimization scale.
**3) Model-based Clustering Methods:** These methods attempt to model data in some mathematical models like decision trees or neural networks and next try to cluster these new models. For more details see[8].

2.3. **Clustering measures.** Which clustering is acceptable? Unfortunately, there is not unique definition for clustering measurement and also in some applications clusters should have some properties, for example some applications need fix number of clusters or the number of members in each cluster should be almost equal. Each clustering method can satisfy some properties, for example in partitioning methods, number of clusters are fixed. In graph clustering, one can also define some clustering measurements, for example each cluster should intuitively be connected, in the other word, if $u$ and $v$ are in cluster $C$, at least a path should be between $u$ and $v$ with edges in $C$. Another measurement for graph clustering is based on density as follow:
Let graph $G(V, E)$ be clustered as $C = \{C_1, C_2, \cdots, C_k\}$:

$$InternalEdges(i) = |\{(u,v)|u,v \in C_i, (u,v) \in E\}|, 1 \leq i \leq k$$
$$ExternalEdges(i) = |\{(u,v)|u \in C_i \, and \, v \in C_j, (u,v) \in E, 1 \leq i,j \leq k, i \neq j\}|$$
$$Degree(u) = |\{(u,v)|u \neq v), (u,v) \in E\}|$$
$$Internaldegree(i) = |\{(u,v)|u,v \in C_i, u \neq v\}| = \frac{|C_i|(|C_i|-1)}{2}, 1 \leq i \leq k$$
$$InternalEdges = \sum_{i=1}^{k} InternalEdges(i)$$
$$ExternalEdges = \sum_{i=1}^{k} ExternalEdges(i)$$
$$Internaldegree = \sum_{i=1}^{k} Internaldegree(i)$$

Clearly clustering $C$ is a good clustering if $InternalEdges \simeq Internaldegree$ and $ExternalEdges \simeq 0$. In the other word, each cluster should be dense as possible and the number of edges between clusters should be small[9].

2.4. **SVD and its application in clustering.** Singular value decomposition (SVD) is recalled as a reminder.
Singular value decomposition (SVD) is decomposing matrix $A$ to three matrices $U, S$ and $V$such that $U$ and $V$ are orthogonal and $S$ is a diagonal matrix with singular values on its diagonal. SVD has so many applications in computer science such as signal processing, image compressing, clustering data and reduction noises from signal and images. By having SVD of a matrix, one can compute the pseudoinverse, least squares fitting of data, matrix approximation, determine the rank, range and null space of a matrix. For more details see [3]. SVD can be computed for all matrices not just for square or symmetric or real

matrices, so all matrices can be decomposed with SVD.

**Definition 2.1.** Let $A$ be a real $m \times n$ matrix. SVD of $A$ is:

$$A = USV^T, \tag{2.3}$$

where $U$ and $V$ are orthogonal and $S = diag(\sigma_1, \sigma_2, \cdots, \sigma_r)$ where $r = min(m; n)$ and $\sigma_1, \sigma_2, \cdots, \sigma_r > 0$. The $\sigma_i' s (1 < i < r)$ are called the singular values of $A$ and the first $r$ columns of $V$ are the right singular vectors and the first $r$ columns of $U$ are the left singular vectors of $A$. If matrix $A$ is not a real matrix, $U$ and $V$ are unitary and $S$ is diagonal[5].

For computing SVD, the best method is Golub Kahan Reinsch algorithm. This algorithm has two phases, in first phase, matrix $A$ is reduced to a bidiagonal matrix $B$ and in next phase matrix $B$ is reduced to a diagonal matrix and its time complexity is $2m^2n + 4mn^2 + \frac{9}{2}n^3$.

Another method for computing SVD is by using eigen decomposition[2]. Computing singular value decomposition by using eigen decomposition is as follow:

The left-singular vectors of matrix $A$, are a set of orthogonal eigenvectors of $AA^T$.

The right-singular vectors of matrix $A$, are a set of orthogonal eigenvectors of $A^TA$.

The non-zero singular values of $A$, are the square roots of the non-zero eigenvalues of both $A^TA$ and $AA^T$.

The eigen decomposition can give information about the connectivity of the graph, so this information also exists in SVD.

**Clustering with SVD:**

Clustering with SVD is derived from Fiedler method. In Fiedler method first the Laplacian matrix is obtained as $(A - D)$ where D is a diagonal matrix and its elements is the degree of each vertex. Next, eigen decomposition of Laplacian matrix is obtained and second eigenvector (is called Fiedler vector) is selected. By using the signs of this eigenvector, rows of matrix clustered in two clusters such that the rows with the same sign are placed in the same cluster. The Fiedler method, divides graph in two sub graph and repeat eigen decomposition for each sub graph until desired clustering is obtained. The Fiedler method has some limitations. First, Fiedler method only works for square symmetric matrices. Second, this method is iterative and any mistake in any iterative can be extend in further iterations and new eigen decompositions must be computed at every iteration, so it is expensive for large datasets. For more details see[6].

As mentioned above, SVD has relation with eigen decomposition so left singular vectors and right singular vectors also have information about matrix connectivity.

| | | | |
|---|---|---|---|
| -0.8839 | 0.053 | – | + |
| 0.2207 | 0.1961 | + | + |
| 0.089 | -0.7467 | + | – |
| 0.3701 | -0.0798 | + | – |
| 0.1585 | 0.6283 | + | + |

FIGURE 1. SVD clustering sign method

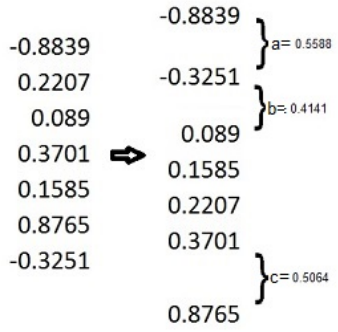


FIGURE 2. SVD clustering gap method

In SVD method of clustering for graph $G(V, E)$ first, adjacency matrix of graph should be derived and by using this matrix, graph can be clustered. SVD method clustering is based on signs of singular vector elements(entries) or gaps between singular vector elements.

In SVD clustering sign method, sign patterns of elements of the singular vectors are considered. Sign pattern of columns of $U$ can cluster rows of matrix and signs pattern of columns of $V^T$, can be used to cluster the columns of matrix. For example, in Figure 1 we have two columns of $U$ and by its signs pattern, rows 2 and 5 are in a cluster and rows 3 and 4 and in another cluster and row 1 in another cluster.

In SVD clustering gap method, elements of singular vectors are sorted and rows and columns are clustered where gaps accrued in this sorted vectors. In Figure 2, gaps in a vector of U are in a , b and c and with these gaps rows of matrix can be clustered in 4 clusters. For more details see[6]

Clustering methods with SVD limitations are the number of clusters and the number of members in each cluster. By using $k$ singular vectors, the number of clusters can be between 1 and $2^k$. In Figure 3(a), the number of clusters are big and in Figure 3(b), the number of clusters are small. Also, the number of

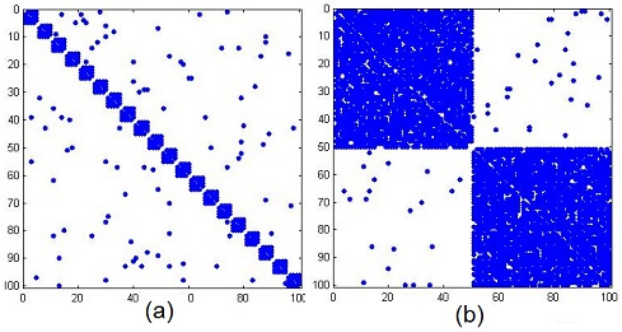


FIGURE 3. Number of clusters

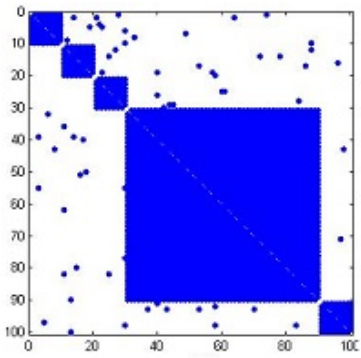

FIGURE 4. Number of members in each cluster

members in each cluster can be between 1 and the size of matrix. In Figure 4, some of clusters have so many members and some of the clusters have few members[9].

## 3. Hierarchical SVD Method

In clustering with SVD methods, the number of clusters and the number of members in each cluster are not selectable and sometimes there are clusters with just a few members or clusters with many members. These clustering are not useful. In some applications, the number of clusters are important and often the number of cluster members should be in a fix range. For example, in telecommunications network, the number of users in each station should be almost equal, otherwise, network is not optimal. In hierarchical SVD method, range of each cluster $[R_1, R_2]$ is a parameter for algorithm, so the number of clusters and the number of cluster members are selectable for users.
In proposed hierarchical SVD algorithm, SVD sign clustering method is used. In this paper, first right and left singular vectors are used. With this singular

vectors, rows and columns of matrix will be clustered into two clusters. For large datasets, one can use more singular vectors in each iterations, so often in each iteration the number of clusters are more than two. If matrix is symmetric, the SVD sign method clusters rows and columns in the same order but for asymmetric or rectangular matrices the orders are not the same, so for these matrices the range should be used for rows or columns. Here, ranges are used for row clustering. For column clustering, the input matrix should be transposed. In hierarchical SVD algorithm, after the first iteration, the SVD sign method clusters matrix $A$ into two submatrices $A_1$ and $A_2$. Submatrix $A_1$ can be in 3 states:

- The number of rows in $A_1$ is between $R_1$ and $R_2$, so $A_1$ is a good cluster and goes to ClusteredMatrix as a submatrix.
- The number of rows in $A_1$ is less than $R_1$, so $A_1$ is a small cluster and goes to RestMatrix as a submatrix.
- The number of rows in $A_1$ is more than $R_2$, so $A_1$ must be clustered recursively.

For $A_2$ also these 3 states must be checked(see algorithm 1). At last step of algorithm 1, the rest matrices must be attached to ClusteredMatrix. At first for each submatrix in RestMatrix, the number of common edges with each good submatrix must be computed. Secondly, each rest submatrix attache to a ClusteredMatrix submatrix with maximum common edges if the number of rows in ClusteredMatrix submatrix is less than $R_2$. For more details see algorithm 2.

For graph clustering with hierarchical SVD, first adjacency matrix should be derived. For simple graph, this adjacency matrix is symmetric and row clustering is same as column clustering. For bipartite graph $(V_1, V_2, E)$ after deriving adjacency matrix $A$ for clustering, it is enough to cluster submatrix $B$. For clustering $V_1$, rows of $B$ should be clustered and for clustering $V_2$, rows of $B^T$ should be clustered.

3.1. **Complexity and analysis of Algorithm HSVD.** Complexity of algorithm SVD for matrix $A_{m\times n}$, $(m > n)$ is $O(m^2n)$. When $A$ is a square matrix $(n = m)$, so complexity of algorithm of graph clustering by using SVD is $O(n^3)$. Hierarchical SVD clustering is a recursive algorithm and the best case is when in each iteration, size of submatrices $A1$ and $A2$ are $|A/2|$ and so complexity of algorithm is:

$$T(n) = O(n^3) + 2T(n/2) = O(n^3 logn) \tag{3.1}$$

When matrix is square and symmetric, first singular value and vectors can be computed in $O(n^2)$[4]. Therefore, if input matrix in hierarchical SVD algorithm is square and symmetric, and only first singular vectors is used in each iteration, complexity of algorithm is $O(n^2 logn)$.

**Algorithm 1** Hierarchical SVD

Input: $A, R$, where $A$ is a matrix $m \times n$ and $R$ is a rage $[R_1, R_2]$
Output: FinallClustring, where FinallClustring is reordered matrix $A$ that is clustered
Find $[U_1, S_1, V_1] = SVD(A, 1)$
$CR1 =$ Rows of $A$ where signs in $U_1$ is negative
$CC1 =$ Columns of $A$ where signs in $V_1$ is negative
$CR2 =$ Rows of $A$ where signs in $U_1$ is positive
$CC2 =$ Columns of $A$ where signs in $V_1$ is positive
If $|CR1| < R$ add $A(CR1, CC1)$ to RestMatrix
Else If $|CR1|$ in $R$ add $A(CR1, CC1)$ to ClusteredMatrix
Else If $|CR1| > R$ repeat HierarchicalSVD$(A(CR1, CC1), R)$
End if
If $|CR2| < R$ add $A(CR2, CC2)$ to RestMatrix
Else If $|CR2|$ in $R$ add $A(CR2, CC2)$ to ClusteredMatrix
Else If $|CR2| > R$ repeat HierarchicalSVD$(A(CR2, CC2), R)$
End if
MinClusterNum=$\lceil \frac{A}{R_2} \rceil$
If cluster number of $ClusteredMatrix < MinClusterNum$ add empty cluster to ClusteredMatrix
End if
Call FinallClustring(ClusteredMatrix, RestMatrix)

**Algorithm 2** FinallClustring

Input: ClusteredMatrix, RestMatrix,
Output: ClusteredMatrix
For each cluster $P1$ in RestMatrix
For each cluster $P2$ in ClusteredMatrix
If maximum common edges are between $P1$ and $P2$ and $|P1 + P2| \leq R_2$ add $P1$ to $P2$
End for
End for
For each cluster $P1$ in RestMatrix that is not attached to ClusteredMatrix
For each row $R$ in $P1$
For each cluster $P2$ in ClusteredMatrix
If maximum common edges are between $R$ and $P2$ add $R$ to $P2$
End for
End for
End for

This algorithm can be used to cluster data sets to desirable number of clusters and the number of members in each clusters can be selected. For square and symmetric matrices time complexity of algorithm is $O(n^2 logn)$ and for another matrices is $O(n^3 logn)$. In this algorithm the number of singular vector that be used is one and so in each iterations, each cluster divides into two clusters. Because of high complexity of algorithm for huge data, in each iterations more than one singular vectors can be used to cluster data in more cluster numbers.

## 4. Results

In experiments, the results of hierarchical SVD method are compared with SVD sign method. The algorithms are implemented in Matlab and used SVD of Matlab. The version of Matlab is R2012a and results are for simple and bipartite graph.

4.1. **Results for simple graph.** In first experiment, a random graph with 24 vertices was tried. In Figure 5, matrix (a) is the first matrix that is derived from graph (b). For this graph, we run SVD sign algorithm for 2 first singular vectors(k=2) and 3 first singular vectors(k=3) and hierarchical SVD algorithm for range [4,5]. The results are shown in Figure 5 and Table 1. As mentioned in section 2.3, a clustering is good where *InternalEdges* is large and *ExternalEdges* is low and minimum and maximum number of members in clusters be in a acceptable range. The results show, when in SVD sign method k=2, range of the number of members is between 2 and 7 and when in sign SVD method k=3, range of the number of members is between 2 and 6 but in hierarchical SVD algorithm, the range is between 4 and 5 and *InternalEdges* is more than SVD sign algorithm results and *ExternalEdges* is less than SVD Sign algorithm results. So in hierarchical SVD method, the measurement parameters for clustering is better than SVD sign method.

Table 1. Graph clustering results

| Method | ClusterNum | InternalEdges | ExternalEdges | MinMember | MaxMember |
|---|---|---|---|---|---|
| SVD k=2 | 4 | 78 | 22 | 2 | 7 |
| SVD k=3 | 7 | 24 | 76 | 2 | 6 |
| HSVD | 5 | 68 | 32 | 4 | 5 |

In the next experiment, a random matrix with 2000 vertices is used. In Figure 6 the first matrix is random matrix. For this matrix, we run SVD sign algorithm for 3 first singular vectors(k=3) and 4 first singular vectors(k=4) and hierarchical SVD algorithm for the range between 200 and 300. The results are shown in Figure 6 and Table 2. The results show, when in SVD sign method k=3, range of the number of members is between 125 and 342 and when in

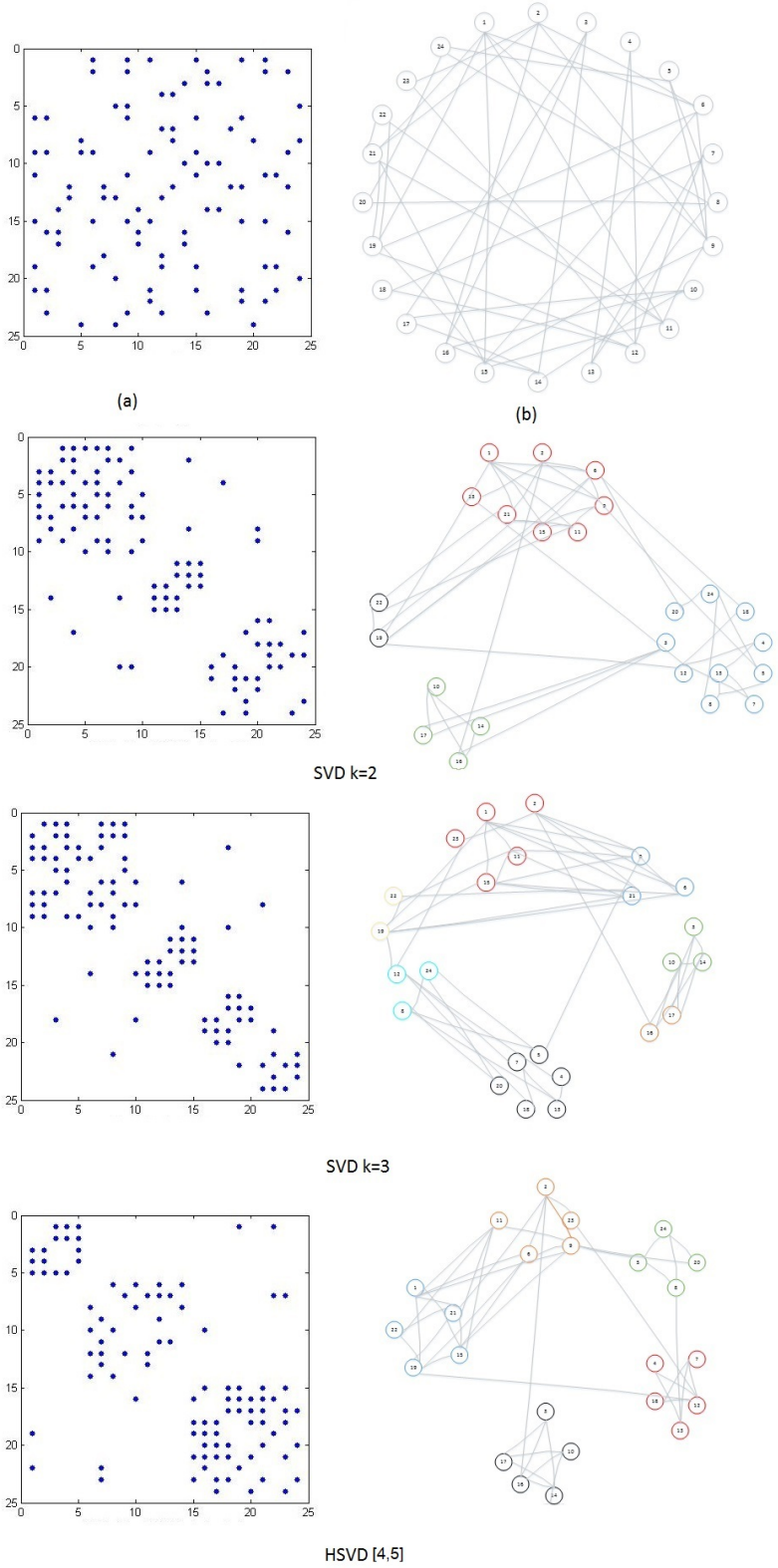


FIGURE 5. Graph clustering with SVD and HSVD

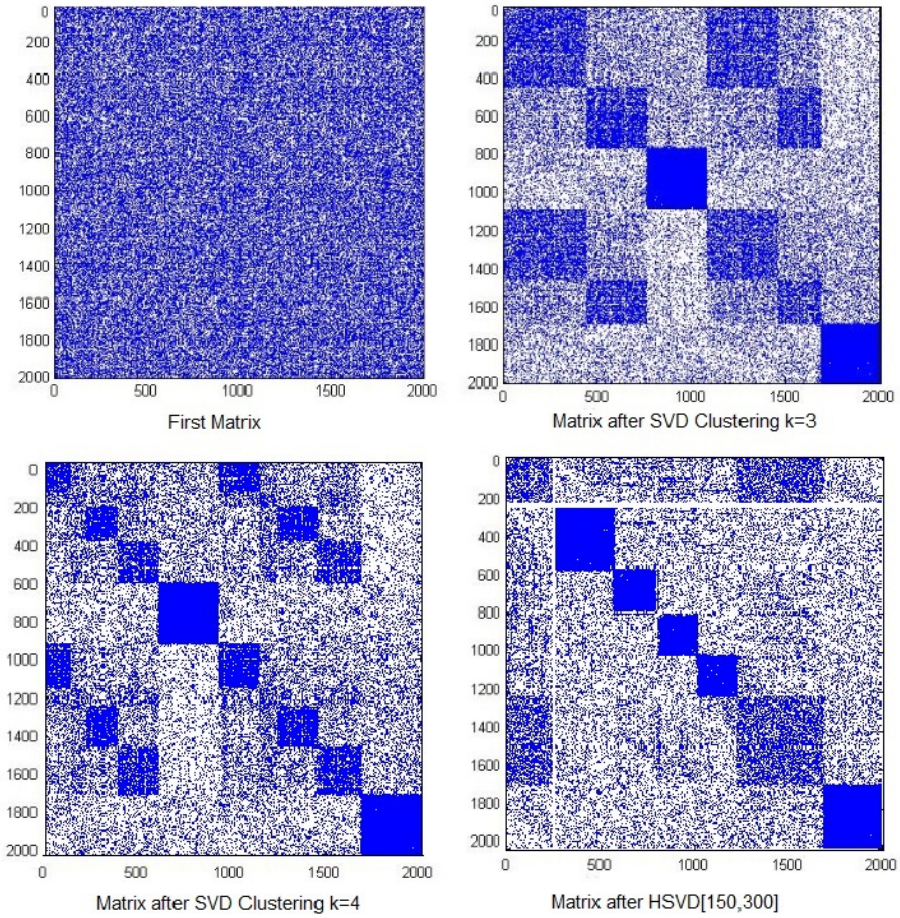


FIGURE 6. Random matrix clustering with SVD and HSVD

sign SVD method k=4, range of the number of members is between 52 and 240 but in hierarchical SVD algorithm, the range is between 207 and 291 and *InternalEdges* is more than SVD sign algorithm results and *ExternalEdges* is less than SVD sign algorithm results.

TABLE 2. Random matrix clustering results

| Method | ClusterNum | InternalEdges | ExternalEdges | MinMember | MaxMember |
|---|---|---|---|---|---|
| SVD k=3 | 8 | 40682 | 74097 | 125 | 342 |
| SVD k=4 | 16 | 27027 | 87752 | 52 | 240 |
| HSVD | 7 | 66232 | 48547 | 207 | 291 |

In next experiment, we used Facebook data for 4020 members and relations between them. In Figure 7, the first matrix is Facebook data that is changed to matrix. For this matrix, we run SVD sign algorithm for 4 first singular vectors(k=4) and 5 first singular vectors(k=5) and hierarchical SVD algorithm for range between 150 and 350. The results are shown in Figure 7 and Table 3. The results show, when in SVD sign method k=4, range of the number of members is between 2 and 2577 and when in SVD sign method k=5, range of the number of members is between 1 and 2094 but in hierarchical SVD algorithm, the range is between 160 and 325 and *InternalEdges* is more than SVD sign algorithm results and *ExternalEdges* is less than SVD sign algorithm results. So in hierarchical SVD method, the measurement parameters for clustering is better that SVD Sign method.

TABLE 3. Facebook data results

| Method | ClusterNum | InternalEdges | ExternalEdges | MinMember | MaxMember |
|---|---|---|---|---|---|
| SVD k=4 | 14 | 135254 | 33054 | 2 | 2577 |
| SVD k=5 | 23 | 132628 | 35680 | 1 | 2094 |
| HSVD | 16 | 142694 | 25614 | 160 | 325 |

4.2. **Results for bipartite graph.** In the next experiment, a random graph with $V_1 = 2000$ and $V_2 = 100$ vertices is used. In Figure 8, the first matrix is the graph adjacency matrix. For this matrix, we run SVD sign algorithm for 3 first singular vectors(k=3) and 4 first singular vectors(k=4) and hierarchical SVD algorithm for the range between 200 and 300. The results are shown in Figure 8 and Table 4. The results show, when in SVD sign method k=3, range of the number of members is between 114 and 426 and when in sign SVD method k=4, range of the number of members is between 52 and 341 but in hierarchical SVD algorithm, the range is between 220 and 299 and *InternalEdges* is more than SVD sign algorithm results and *ExternalEdges* is less than SVD sign algorithm results.

TABLE 4. Random bipartite graph clustering results

| Method | ClusterNum | InternalEdges | ExternalEdges | MinMember | MaxMember |
|---|---|---|---|---|---|
| SVD k=3 | 8 | 3118 | 4801 | 114 | 426 |
| SVD k=4 | 16 | 2300 | 5619 | 52 | 341 |
| HSVD | 8 | 3303 | 4616 | 220 | 299 |

In the next experiment, graph for relation between Iran's provinces and mines are collected. The number of provinces is 29( $V_1 = 29$) and the number

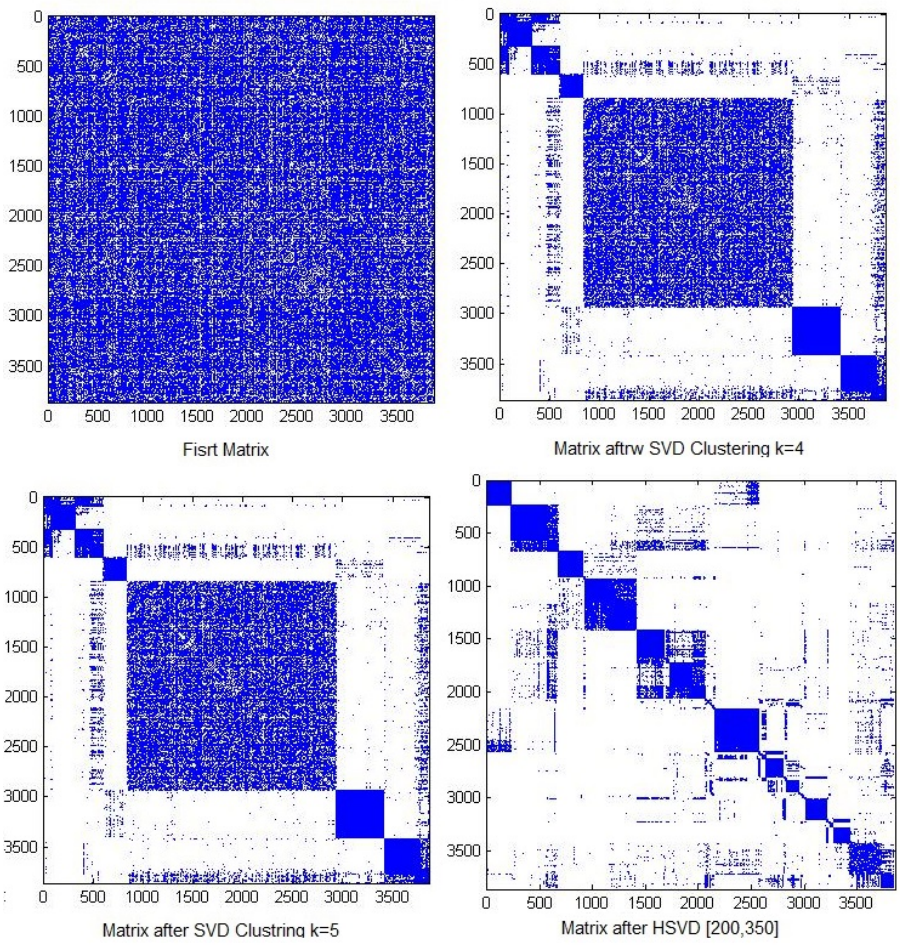


FIGURE 7. Facebook data clustering with SVD and HSVD

of mines is 76($V_2 = 76$). In Figure 9, the first matrix is the graph adjacency matrix. For this matrix, we run SVD sign algorithm for 3 first singular vectors(k=3) and 4 first singular vectors(k=4) and hierarchical SVD algorithm for the range between 4 and 6. The results are shown in Figure 9 and Table 5. The results show, when in SVD sign method k=3, range of the number of members is between 1 and 22 and when in sign SVD method k=4, range of the number of members is between 1 and 17 but in hierarchical SVD algorithm, the range is between 4 and 6 and *InternalEdges* is more than SVD sign algorithm results and *ExternalEdges* is less than SVD sign algorithm results.

In the next experiment, Movielens latest dataset is used. This dataset consists of approximately 100k ratings by 610( $V_1 = 610$) user for 9724($V_2 = 9724$)

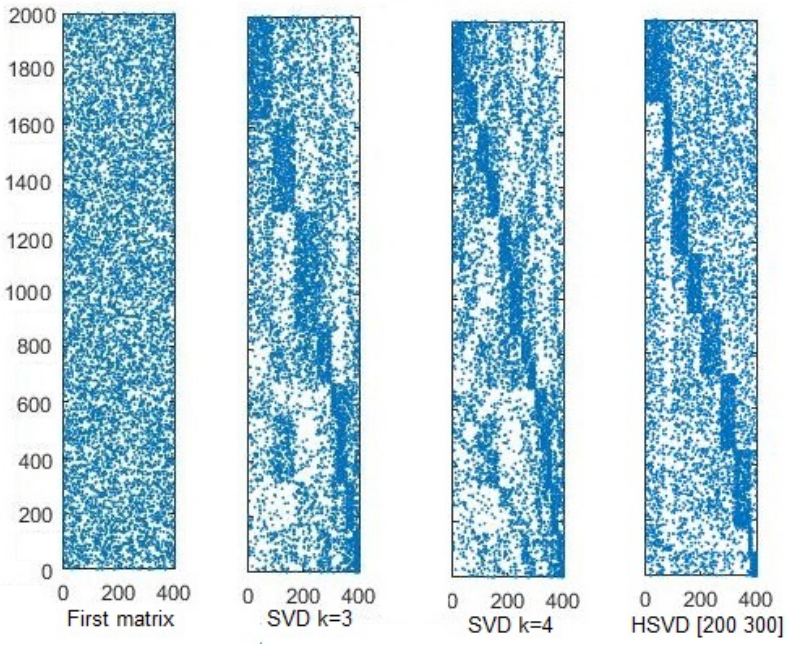


FIGURE 8. Random bipartite graph clustering with SVD and HSVD

TABLE 5. Iran's province clustering based on their mines results

| Method | ClusterNum | InternalEdges | ExternalEdges | MinMember | MaxMember |
|---|---|---|---|---|---|
| SVD k=3 | 5 | 61 | 289 | 1 | 20 |
| SVD k=4 | 7 | 40 | 310 | 1 | 17 |
| HSVD | 5 | 93 | 257 | 4 | 6 |

movies. In Figure 10, the first matrix is the graph adjacency matrix. For this matrix, we run SVD sign algorithm for 4 first singular vectors(k=4) and 5 first singular vectors(k=5) and hierarchical SVD algorithm for the range between 50 and 100. The results are shown in Figure 10 and Table 6.

TABLE 6. Movielens latest dataset clustering results

| Method | ClusterNum | InternalEdges | ExternalEdges | MinMember | MaxMember |
|---|---|---|---|---|---|
| SVD k=4 | 9 | 22384 | 78452 | 1 | 159 |
| SVD k=5 | 17 | 11332 | 89504 | 1 | 94 |
| HSVD | 7 | 13380 | 87456 | 67 | 97 |

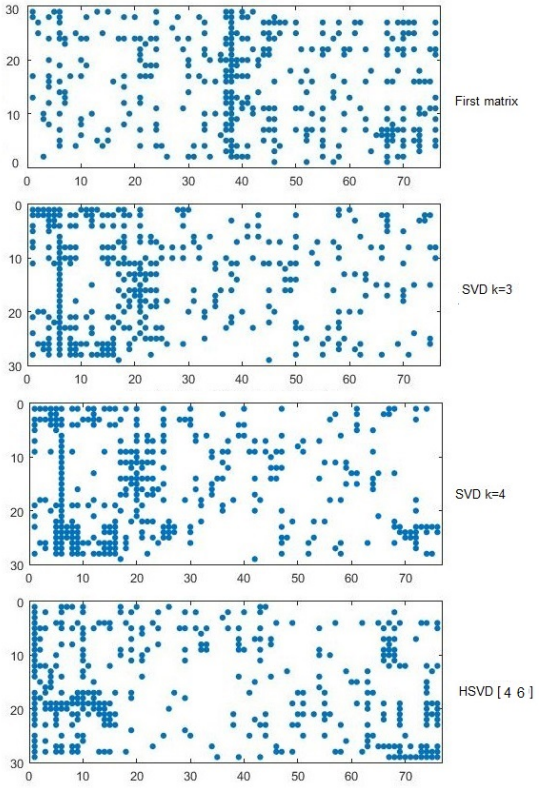


FIGURE 9. Iran's provinces clustering based on their mines with SVD and HSVD

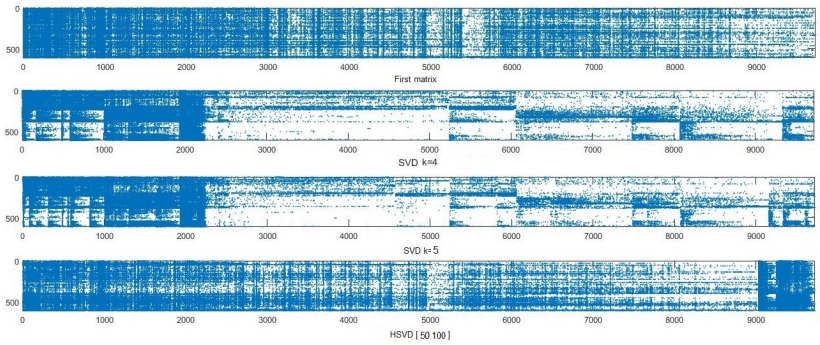


FIGURE 10. Movielens latest dataset clustering with SVD and HSVD

## 5. Conclusion

In graph clustering for some applications, it is important that each cluster be dense as possible and the number of members in each cluster be almost equal.

The SVD method clustering is one of the best tools for clustering but users do not have any chance to select the number of clusters and a range for the number of members in each cluster. In hierarchical SVD method, we changed SVD sign method with using hierarchical methods for clustering to improve SVD method clustering. In hierarchical SVD method, users can select a range for the number of members in each cluster, so the number of clusters are also desirable. The results show the measurement parameter for clustering by hierarchical SVD method is better than SVD method. This algorithm is recursive and in each iterations, with using first singular vectors, each cluster divides into tow clusters. For simple graph adjacency matrix is square and symmetric and so the time complexity of algorithm is $O(n^2 logn)$, but for another graphs, the complexity of algorithm is $O(n^3 logn)$. So for achieve desirable clustering faster, in each iterations more than one singular vectors can be used for dividing each cluster into more than two clusters.

## Acknowledgments

The authors would like to thank the referee for her/his comments.

## References

1. T. P. Cason, P. A. Absil, P. Van Dooren, Iterative methods for low rank approximation of graph similarity matrices *Linear Algebra and its Applications*,**1**(438),(2013), 1863-1882.
2. A. K. Cline, S. Dhillon, Computation of the Singular Value Decomposition, 14, The University of Texas at Austin, 2007.
3. B.Datta, *Numerical linear algebr and Applications*, Second edition, SIAM, 2010.
4. J. Demmel, Accurate singular value decomposition of structured matrices *SIAM*,**21**(2),(1997), 562-580.
5. J. Dongarra, Accuracy of computed singular values, *SIAM*,**1**(4), (1983), 712-719 .
6. E. P. Douglas, *Clustering datasets with singular value decomposition*, College of Charleston, 2008.
7. G. H. Golub, C. Reinsch, Singular value decomposition and least squares solutions, *Number Math*, **1**(14), (1973), 403-420.
8. L. Rokach, O. Maimon, Data Mining and Knowledge Discovery Handbook chapter: Clustering Methods, *Springer US*,(2005), 321-352.
9. S. E. Schaeffer, Survey: Graph clustering, *Computer Science Review*,**1**(1), (2007), 27-64.
10. G. W. Stewart, Error and perturbation bounds for subspaces associated with certain eigenvalue problems *SIAM*, **1**(15) (1973), 727-764.
11. X. Zhou, SVD-based incremental approaches for recommender systems, *Computer and System Sciences*,**1**(81), (2015), 717-733.