IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 12, Issue 6 Ver. II (Nov. - Dec.2016), PP 55-62
www.iosrjournals.org

Cluster of tangle graph

H. El-Zohny*, S.Radwan*, Z.M.Hagrass*

Abstract: *In this paper we will introduce how to cluster of tangle graph. We will study clustering by normalized tangle graph cut and markov cluster. We will explain the cluster by examples. We will illustrate algorism to understand the cluster.*
Keywords: *Tangle graph, cluster, normalized, markov model.*

I. Introduction

Clustering can be considered the most important *unsupervised learning* problem; so, as every other problem of this kind, it deals with finding a *structure* in a collection of unlabeled data. A loose definition of clustering could be "the process of organizing objects into groups whose members are similar in some way".A *cluster* is therefore a collection of objects which are "similar" between them and are "dissimilar" to the objects belonging to other clusters. the goal of clustering is to determine the intrinsic grouping in a set of unlabeled data. But how to decide what constitutes a good clustering? It can be shown that there is no absolute "best" criterion which would be independent of the final aim of the clustering. Consequently, it is the user which must supply this criterion, in such a way that the result of the clustering will suit their needs.For instance, we could be interested in finding representatives for homogeneous groups (*data reduction*), in finding "natural clusters" and describe their unknown properties (*"natural" data types*), in finding useful and suitable groupings (*"useful" data classes*) or in finding unusual data objects (*outlier detection*).

PossibleApplications:

- *Marketing*: finding groups of customers with similar behavior given a large database of customer data containing their properties and past buying records;
- *Biology*: classification of plants and animals given their features;
- *Libraries*: book ordering;
- *City-planning*: identifying groups of houses according to their house type, value and geographical location;
- *Earthquake studies*: clustering observed earthquake epicenters to identify dangerous zones.

Definition1:

Tangle graph: Let D be a unit cube, so D={(x,y,z):0<x,y,z<1} on the top face of cube place n points $a_1, a_2, \ldots, a_n$ similarly place on bottom face $b_1, b_2, \ldots, b_n$, now join the points $a_1, a_2, \ldots, a_n$ with $b_1, b_2, \ldots, b_n$ by arcs $d_1, d_2, \ldots, d_n$ these arcs are disjoint and each d_i connects some a_j to b_k not connect a_j to a_k or b_j to b_k this called tangle [1].

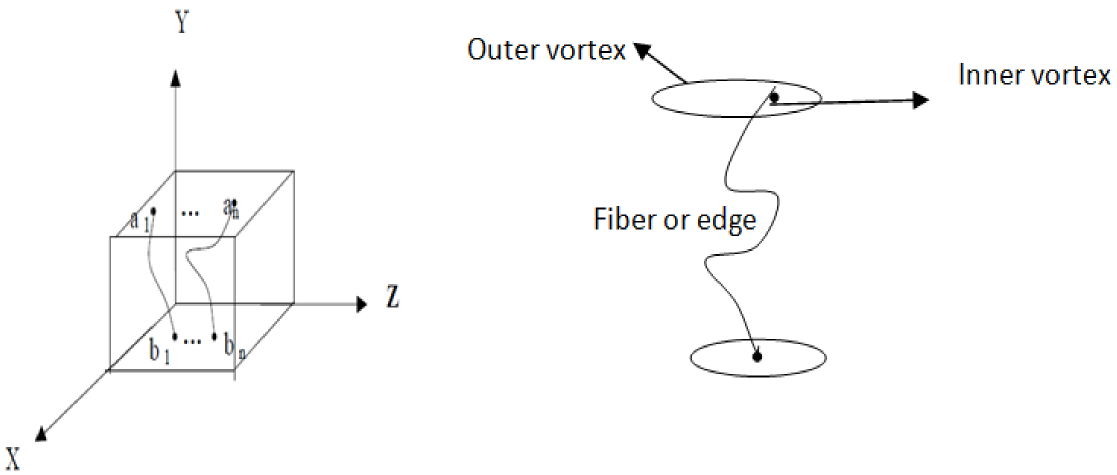


Fig.(1)

In tangle graph T(V,E) , whose vertices consists of inner and outer vertices, and the edges(E) subsets e of V.

Definition 2:
Markov Chain:

A sequence of variables X1, X2, X3, etc (in our case, the probability matrices) where, given the present state, the past and future states are independent. random walk is an example of a Markov Chain, using the transition probability matrices [2].

Definition 3:
Inflation operator [3]**:**

DEFINITION 3. *Given a matrix* $M \in \mathbb{R}^{k\times l}$, $M \geq 0$, *and a real nonnegative number* r, *the matrix resulting from rescaling each of the columns of* M *with power coefficient* r *is called* $\Gamma_r M$, *and* Γ_r *is called the* **inflation** *operator with power coefficient* r. *Formally, the action of* $\Gamma_r : \mathbb{R}^{k\times l} \to \mathbb{R}^{k\times l}$ *is defined by*

$$(\Gamma_r M)_{pq} = (M_{pq})^r / \sum_{i=1}^{k} (M_{iq})^r$$

If the subscript is omitted, it is understood that the power coefficient equals 2. □

Algorism of MCL:

1. Input is an un-directed graph, power parameter *e*, and inflation parameter *r*.
2. Create the associated matrix
3. Add self loops to each node (optional)
4. Normalize the matrix
5. Expand by taking the e^{th} power of the matrix
6. Inflate by taking inflation of the resulting matrix with parameter *r*
7. Repeat steps 5 and 6 until a steady state is reached (convergence).
8. Interpret resulting matrix to discover clusters.

Main results:
Weighted tangle graph:
A weighted tangle graph is a graph that has appositive number w (e) associated with each edge (fiber) called the weight of edge. Denote a weighted tangle graph by T (V, E, w).
Degree of a vortex:
Degree of a vortex is defined by the relation $d(v) = \sum_{e\in E | v\in e} w(e)\, h(v,e)$
Where h (v, e) =1 express if we find relation between vortex and edge i.e. (v∊e) and 0 otherwise**.**
Note:
The tangle graph is special case of hyper graph.

Normalized hyper graph cut:

For a vertex subset $S \subset V$, let S^c denote the compliment of S. A cut of a hypergraph $G = (V, E, w)$ is a partition of V into two parts S and S^c. We say that a hyperedge e is cut if it is incident with the vertices in S and S^c simultaneously.

Given a vertex subset $S \subset V$, define the *hyperedge boundary* ∂S of S to be a hyperedge set which consists of hyperedges which are cut, i.e. $\partial S := \{e \in E | e \cap S \neq \emptyset, e \cap S^c \neq \emptyset\}$, and define the *volume* $\mathrm{vol}\, S$ of S to be the sum of the degrees of the vertices in S, that is, $\mathrm{vol}\, S := \sum_{v\in S} d(v)$. Moreover, define the volume of ∂S by

$$\mathrm{vol}\, \partial S := \sum_{e\in\partial S} w(e) \frac{|e \cap S||e \cap S^c|}{\delta(e)}. \quad (1)$$

Adjacency matrix A:
A= H W HT-Dv Where H is the incident matrix (v X E) , W denote the diagonal matrix containing of the weight of edge of tangle graph , HT is the transpose of H and Dv denote the diagonal matrix of vortex degree.

Example1:

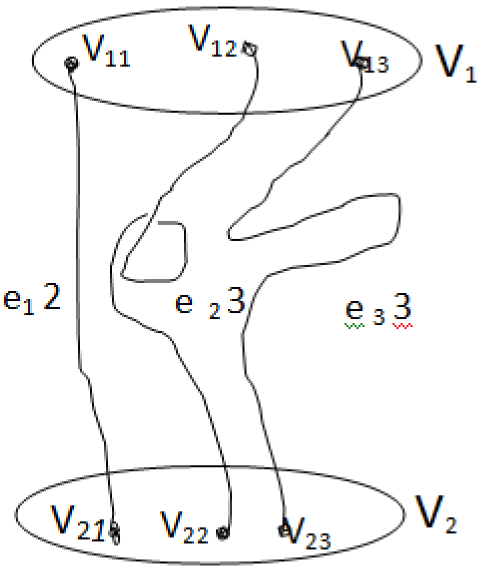


Fig.(2)

Note:

1ijk express the relation between vortex and edge and 0 other wise, where I express the number of fiber, j express number of rolls and k express the number of curves.

The incident matrix:

$$H=\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1_{11}^{1} & 0 \\ 0 & 0 & 1_{03}^{1} \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1_{11}^{1} & 0 \\ 0 & 0 & 1_{03}^{1} \end{bmatrix} \quad \Big| \quad H^{T}=\begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1_{11}^{1} & 0 & 1 & 0 & 1_{11}^{1} & 0 \\ 1 & 0 & 0 & 1_{03}^{1} & 1 & 0 & 0 & 1_{03}^{1} \end{bmatrix}$$

$$W=\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

$D_v=$

$$\begin{matrix} dv_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & dv_{11} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & dv_{12} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & dv_{13} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & dv_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & dv_{21} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & dv_{22} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & dv_1 \end{matrix}$$

Where **d(v)** is degree of vortex such that **d(v)** = $\sum_{e\in E|v\in e} w(e)\ h(v,e)$.

$d(v_1)=W(e_1)\,h(v_1,e_1)+W(e_2)\,h(v_1,e_2)+W(e_3)\,h(v_1,e_3)$.

$d(v_1) = 2+3+3 = 8$

$d(v_2) = 8$

$d(v_{11}) = 2$

$d(v_{12}) = 3^* 1^1_{11}$

$d(v_{13}) = 3^* 1^1_{03}$

$d(v_{21}) = 2$

$d(v_{22}) = 3^* 1^1_{11}$

$d(v_{23}) = 3^* 1^1_{03}$

$$\begin{bmatrix} 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 * 1^1_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 * 1^1_{03} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 * 1^1_{11} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 * 1^1_{03} \end{bmatrix}$$

Adjacent matrix:

$A = H\ W\ H^T - D_v$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1^1_{11} & 0 \\ 0 & 0 & 1^1_{03} \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1^1_{11} & 0 \\ 0 & 0 & 1^1_{03} \end{bmatrix} * \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} * \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1^1_{11} & 0 & 1 & 0 & 1^1_{11} & 0 \\ 1 & 0 & 0 & 1^1_{03} & 1 & 0 & 0 & 1^1_{03} \end{bmatrix}$$

$$- \begin{bmatrix} 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 * 1^1_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 * 1^1_{03} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 * 1^1_{11} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 * 1^1_{03} \end{bmatrix}$$

$$=$$

$$\begin{bmatrix} 8 & 2 & 3 * 1^1_{11} & 3 * 1^1_{03} & 6 & 2 & 2 + 3 * 1^1_{11} & 3 * 1^1_{03} \\ 2 & 2 & 0 & 0 & 2 & 2 & 2 & 0 \\ 3 * 1^1_{11} & 0 & 3 * (1^1_{11})^2 & 0 & 3 * 1^1_{11} & 0 & 3 * (1^1_{11})^2 & 0 \\ 3 * 1^1_{03} & 0 & 0 & 3 * (1^1_{03})^2 & 3 * 1^1_{03} & 0 & 0 & 3 * (1^1_{03})^2 \\ 8 & 2 & 3 * 1^1_{11} & 3 * 1^1_{03} & 6 & 2 & 2 + 3 * 1^1_{11} & 3 * 1^1_{03} \\ 2 & 2 & 0 & 0 & 2 & 2 & 2 & 0 \\ 3 * 1^1_{11} & 0 & 3 * (1^1_{11})^2 & 0 & 3 * 1^1_{11} & 0 & 3 * (1^1_{11})^2 & 0 \\ 3 * 1^1_{03} & 0 & 0 & 3 * (1^1_{03})^2 & 3 * 1^1_{03} & 0 & 0 & 3 * (1^1_{03})^2 \end{bmatrix}$$

$$\begin{bmatrix} 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 * 1^1_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 * 1^1_{03} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 * 1^1_{11} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 * 1^1_{03} \end{bmatrix}$$

Where $(1^i_{jk})^2 = 1^i_{jk}$

The adjacent matrix (A) is equal to:

$$\begin{bmatrix} 0 & 2 & 3*1_{11}^{1} & 3*1_{03}^{1} & 6 & 2 & 2+3*1_{11}^{1} & 3*1_{03}^{1} \\ 2 & 0 & 0 & 0 & 2 & 2 & 2 & 0 \\ 3*1_{11}^{1} & 0 & 0 & 0 & 3*1_{11}^{1} & 0 & 3*1_{11}^{1} & 0 \\ 3*1_{03}^{1} & 0 & 0 & 0 & 3*1_{03}^{1} & 0 & 0 & 3*1_{03}^{1} \\ 0 & 2 & 3*1_{11}^{1} & 3*1_{03}^{1} & 6 & 2 & 2+3*1_{11}^{1} & 3*1_{03}^{1} \\ 2 & 0 & 0 & 0 & 2 & 2 & 2 & 0 \\ 3*1_{11}^{1} & 0 & 0 & 0 & 3*1_{11}^{1} & 0 & 3*1_{11}^{1} & 0 \\ 3*1_{03}^{1} & 0 & 0 & 0 & 3*1_{03}^{1} & 0 & 0 & 3*1_{03}^{1} \end{bmatrix}$$

Normalized fiber (edge) of tangle graph:

In tangle graph the vertices consisting of inner and outer vertices.I f we partition the vertices into two sets, we consider the outer vortex and all vertices contained in it one of two sets (S) and others are consider the complement (S^{c})). We say that the edge (e) is cut if it is incident with the vertices in(S) and (S^{c}) .

The edge boundary:

Is the edge set which consists of edges which are cut.

$$\partial\ (s) = \{e \in E/\ e \cap S \neq \emptyset\ ,\ e \cap S^{c} \neq \emptyset\}$$

The volume of S (vol(s) $=\sum_{v \in S} d(v)$.

The volume of the boundary vol($\partial\ (S)$ $=\sum_{e \in \partial\ (s)} W(e)\ \frac{|e \cap S||e \cap Sc|}{\delta(e)}$

Where $\delta\ (e) = \sum h\ (v, e)$

In example 1:

$S = \{V_1,\ v_{11}, v_{12}\ , v_{13}\}$,$S^{c} = \{\ V_2\ , v_{21}\ , v_{22}\ ,\ v_{13}\}$

$E = \{e_1\ , e_2\ , e_3\}$.

The edges $\{e_1\ ,\ e_2\ , e_3\}$ are incident with S and S^{c} .

$\partial(S) = \{e1\ , e2\ , e3\}$

$Vol(S) = d\ V_1 + d\ v_{11} + d\ v_{12} + d\ v_{13}$

$= 8+2+3 * 1_{11}^{1} + 3 * 1_{03}^{1} = 10 + 3\ (1_{11}^{1} + 1_{03}^{1})$.

$Vol\ (S^{c}) = d\ v_2 + dv_{21} + dv_{22} + dv_{23}$.

$$=8+2+3*1^{1}_{11}+3*1^{1}_{03}=10+3\,(1^{1}_{11}+1^{1}_{03}).$$

$Vol\,\partial(S) = w(e_1)\,(|e1 \cap S||e1 \cap S^c|)/(\delta(e1)) + W(e_2)\,(|e2 \cap S||e2 \cap S^c|)/(\delta(e2)) + W(e_3)\,(|e3 \cap S||e3 \cap S^c|)/(\delta(e3))$

$$=\frac{2*2*2}{4}+\frac{3*2*2}{2*1^{1}_{11}+2}+\frac{3*2*2}{2+2*1^{1}_{02}}=2+\frac{6}{1+1^{1}_{11}}+\frac{6}{1+1^{1}_{02}}$$

$C(S)= Vol\,\partial(S)\,(1/(Vol(S)) + 1/Vol\,(S)^c)$

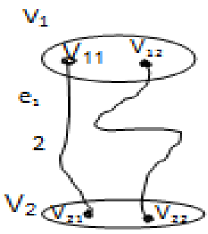
$$=\frac{7+41^{1}_{02}+4\,1^{1}_{11}+1^{1}_{02}\,1^{1}_{11}}{(1+1^{1}_{02})(1+1^{1}_{11})}\left(\frac{1}{10+3\,(1^{1}_{11}+1^{1}_{02}).}+\frac{1}{10+3\,(1^{1}_{11}+1^{1}_{02}).}\right)$$

Note:

To calculate these problems ,we must substitute 1^{i}_{jk} by (j+k).

C(S) = 0.22

Example 2:



d $(V_1) = 5$, d $(V_2) = 5$,d $(V_{11}) = 2$,d $(V_{12}) = 3*1^{1}_{02}$,d $(V_{21}) = 2$

,d $(V_{22}) = 3*1^{1}_{02}$.

$S = \{V_1, v_{11}, v_{12}\}$,$S^c = \{V_2, v_{21}, v_{22}\}$, $\partial(s) = \{e_1, e_2\}$,|

$$vol\ \partial(s) = 2\frac{2*2}{4} + 3\,2*\frac{2*2}{2*1^{1}_{02}+2} = 2+\frac{12}{1^{1}_{02}+1}.$$

$Vol\,S = 5+2+3*1^{1}_{02} = 7+3*1^{1}_{02}$, $Vol\,(S^c) = 5+2+3*1^{1}_{02} = 7+3*1^{1}_{02}$

$$C(S) = 2+\frac{12}{1^{1}_{02}+1}\left(\frac{2}{7+3*1^{1}_{02}}\right)=\frac{2*1^{1}_{02}+2+12}{1^{1}_{02}+1}\left(\frac{2}{7+3*1^{1}_{02}}\right)=0.9$$

Algorism 1:

H = (V,E,W) , a partition $\alpha > 1$.

Output: a cut set C =(S ,S^c).

Step 1:

While v ϵ S, u ∈ S^c e_1 ϵ E and $e_1 \cap S \neq \emptyset$ and $e_1 \cap S^c \neq \emptyset$ do.

Output C= { e_1}

Otherwise return to step 1.

Continue until

C = {e_i ,i=1,2,3,…,n}.

End.

Cluster by markov chain:

We will use cluster by markov chain on braid which be consider special case of tangle graph (fiber not contain any curves or rolls).

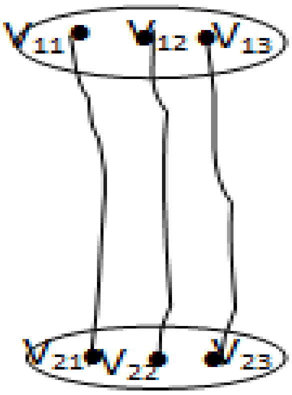
We will apply markov clustering algorism on braid graph.

Algorism2:

1. Input is an un-directed graph, power parameter *e*, and inflation parameter *r*.
2. Create the associated matrix
3. Add self loops to each node (optional)
4. Normalize the matrix
5. Expand by taking the e^{th} power of the matrix
6. Inflate by taking inflation of the resulting matrix with parameter *r*
7. Repeat steps 5 and 6 until a steady state is reached (convergence).
8. Interpret resulting matrix to discover clusters.

Example3:

1- Creat associated matrix:



$$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

II. Adding self loop

$$\begin{matrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{matrix}$$

III. Normalized the matrix

$$\begin{bmatrix} 0.5 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0.5 \\ 0.5 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0.5 \end{bmatrix}$$

IV. Expand the matrix

$$\begin{bmatrix} 0.5 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0.5 \\ 0.5 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0.5 \end{bmatrix} \begin{bmatrix} 0.5 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0.5 \\ 0.5 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0.5 \end{bmatrix}$$

$$\begin{bmatrix} 0.5 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0.5 \\ 0.5 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0.5 \end{bmatrix}$$

V. Inflate by taking inflation of the resulting matrix with parameter r =2

$$\begin{bmatrix} 0.25 & 0 & 0 & 0.25 & 0 & 0 \\ 0 & 0.25 & 0 & 0 & 0.25 & 0 \\ 0 & 0 & 0.25 & 0 & 0 & 0.25 \\ 0.25 & 0 & 0 & 0.25 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0.25 \end{bmatrix}$$

$$\begin{bmatrix} 0.5 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0.5 \\ 0.5 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0.5 \end{bmatrix}$$

Sum of the column = 1 and every raw contain at least one value. In this case we reach convergence of the matrix.

VI. Interpret resulting matrix to discover clusters

In this example we note that we have three clusters, these clusters are:
{ v_{11} ,v_{21}} ,{v_{12} ,v_{22}} ,{v_{13} ,v_{23}}.

Co lorry:

In braid graph the cluster can be defined as if we have vertex on top face say v_{ij} and another vertx say v_{lm} on bottom face then { v_{ij} ,v_{lm} }called cluster.

References

[1]. M.El-Ghoul &M.M.Al-Shamiri, A study on graphs and Knots.

[2]. Van Dongen, S. (2000) *Graph Clustering by Flow* Simulation. PhD Thesis, University of Utrecht, The Netherlands.