ELSEVIER

Information Processing Letters 52 (1994) 259–263

Information Processing Letters

A fuzzy clustering algorithm for graph bisection

Jin-Tai Yan *, Pei-Yung Hsiao

Department of Computer and Information Science, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu, Taiwan, ROC

Communicated by D. Gries; received 27 August 1993; revised 14 June 1994

Abstract

A fuzzy clustering algorithm based on global connection information is proposed to solve the graph bisection problem.

Keywords: Algorithms; Graph bisection; Fuzzy membership; Fuzzy clustering

1. Introduction

Let $G=(V, E)$ be an undirected connected edge-weighted graph. In general, a partition of G is a partition of its vertex set V. Hence, if the ends of an edge e in E belong to two different subsets of the partition, e will be cut by a partition (V_1, V_2) of V. The cut of a partition (V_1, V_2) for graph G is defined as the sum of the weights of all the edges cut by the partition

$$\mathrm{Cut}(V_1, V_2) = \sum_{i \in V_1} \sum_{j \in V_2} c_{ij},$$

where c_{ij} is the weight of the edge $\{i, j\}$ in G. Therefore, a min-cut partition for graph G is a partition (V_1, V_2) of V with minimum cut. However, a min-cut partition always yields an unbalanced partition, and an unbalanced partition is inefficient on many applications. Therefore, balanced-partition graph bisection is formulated as follows: A partition (V_1, V_2) of V is said to be a graph bisection (GB) if $|V_1| = |V_2|$ when $|V|$ is even or $|V_1| = |V_2| - 1$ when $|V|$ is odd.

Due to the size constraint on the partition, GB is NP-complete [4]. Many heuristic approaches have been suggested for GB. In 1970, Kernighan and Lin [5] proposed a two-way "group-migration" improvement algorithm with a constraint on the subset size. They randomly started with two subsets and iteratively applied pairwise swapping on all pairs of nodes. Subsequently, Fiduccia and Mattheyses [3] reduced the time complexity to $O(P)$ with respect to the number of pins P. Saab and Rao [7,8] also proposed heuristic algorithms to solve GB. Generally speaking, the Kernighan–Lin based algorithm [1] is quite efficient, but it does not focus on the global connection information of the given graph. Therefore, it is difficult for it to obtain an optimal or near-optimal graph bisection.

In this paper, we propose a fuzzy clustering algorithm based on global connection information to solve GB. For an undirected connected edge-

* Corresponding author.

SSDI 0020-0190(94)00148-0

weighted graph, we introduce two groups of fuzzy memberships on the vertex set and define the clustering distance between any pair of vertices in the graph according to global connection. Based on fuzzy c-means clustering [2,6], two-way fuzzy graph clustering generates two groups of converged fuzzy memberships for the vertex set. Finally, according to the grades of the memberships, the vertices in the graph can be separated into two even subsets with minimum cut.

2. Fuzzy membership on vertices

Given is an undirected edge-weighted graph $G=(V, E)$, where $V=\{x_1, x_2, \dots, x_n\}$ and $E=\{y_1, y_2, \dots, y_m\}$. Let $\mathbb{R}^+$ the set of nonnegative reals and M_{2n} the set of real $2\times n$ matrices. First, fuzzy memberships and fuzzy functions for the vertex set V are introduced. Every fuzzy function $u_i: V \to [0,1]$ assigns grades of fuzzy memberships onto the vertices in V. Function u_i is called the ith fuzzy set in V. There are infinitely many fuzzy sets associated with V. Every fuzzy set in V represents a possible fuzzy clustering. Hence, for two-way partitioning, two fuzzy sets in V will be applied to partition the vertex set V.

In order to partition V by means of fuzzy sets, we need some clustering constraints between the two fuzzy sets. For example, for each x_k in V, the sum of the fuzzy memberships in the two fuzzy sets is restricted to be 1. Formally, a two-way fuzzy clustering for two-way partitioning can be represented by a fuzzy matrix U in M_{2n} whose entries satisfy the following clustering constraints:

(1) Row i of U, say $U_i=(u_{i1}, u_{i2})$, exhibits fuzzy set i of V.
(2) Column j of U, say $U_j=(u_{1j}, u_{2j})$ exhibits the values of the 2 fuzzy sets of the jth datum in V.
(3) u_{ik} shall be interpreted as $u_i(x_k)$, the value of fuzzy set i for the kth datum.
(4) The sum of the membership values for each x_k is 1 ($u_{1k}+u_{2k}=1$, for all k).
(5) No fuzzy set is empty (row sum $\Sigma_k u_{ik}>0$, for all i).
(6) No fuzzy set is all of V (row sum $\Sigma_k u_{ik}<n$, for all i).

3. Clustering distance

Due to the primary min-cut operation in graph partitioning, it is sure that any pair of connected vertices with larger weight will be clustered into the same cluster to reduce the partitioning result. Hence, for graph partitioning, the larger the weight of the edge, the less its clustering distance. A related clustering graph can be generated by modifying the edge weights of the original edge-weighted graph. For $G=(V, E)$, the related clustering graph $G^*=(V^*, E^*)$ is an undirected edge-weighted graph, where $V^*=V$, $E^*=E$, and the edge weight c^*_{ij} of the edge $\{i, j\}$ is defined by $c^*_{ij}=1/c_{ij}$.

Since there is no geometrical distance between any pair of vertices in a graph, it is critical for fuzzy clustering on a graph structure to estimate the clustering distance of any pair of vertices. Simply speaking, for an undirected edge-weighted graph, the distance of any pair of vertices in the graph is the distance of the shortest path. Furthermore, the clustering distance of any pair of vertices in the related clustering graph can be computed by running a shortest-path algorithm. The clustering distance will indicate the clustered degree of the pair of vertices in the same cluster. For any pair of vertices i and j in G^*, the clustering distance d^*_{ij} between vertex i and j can be further obtained as

$$d^*_{ij}=\begin{cases} c^*_{ij} \\ \quad \text{if } \{i, j\} \text{ is an edge in the graph,} \\ \text{Short\_path}(i, j) \\ \quad \text{if } \{i, j\} \text{ is not an edge in the graph,} \end{cases}$$

where Short_path(s, t) is the sum of the weights on the shortest path from vertex s to vertex t. Clearly, the clustering distance of all pairs of vertices in the graph must be obtained for fuzzy clustering. Hence, all clustering distances can be computed by running an all-pairs shortest-path algorithm.

4. Optimality of fuzzy clustering

Based on fuzzy c-means clustering, two-way fuzzy graph clustering can be transformed into a mathematical optimization problem for the mapped objective function. Using the fuzzy memberships of the vertex set and the clustering distance between any pair of vertices, the objective function for two-way fuzzy graph clustering can be formulated as follows: Let U in M_{2n} be a fuzzy graph partition of V, and let $v = (v_1, v_2)$ be the cluster centers. Objective function $J_i : M_{2n} \times V \to \mathbb{R}^+$ is defined as

$$J_i(U, v_i) = \sum_{k=1}^{n} (u_{ik})^2 (d_{ik})^2.$$

Further, objective function $J : M_{2n} \times V^2 \to \mathbb{R}^+$ is defined as

$$J(U, v) = J_1(U, v_1) + J_2(U, v_2) = \sum_{k=1}^{n} \sum_{i=1}^{2} (u_{ik})^2 (d_{ik})^2,$$

where U in M_{2n} is a fuzzy graph clustering of V, $v = (v_1, v_2)$ in V^2 is the cluster centers, and $d_{ik} = \| x_k - v_i \|$ is the clustering distance between x_k and v_i. Note the several parameters in the definition of the objective function. The squared clustering distance is weighted by the second power of the membership of datum x_k in cluster i. Thus, function J is a squared error criterion, and its minimization produces a fuzzy clustering matrix U that is optimal in a generalized least-squared error sense.

Since two-way fuzzy graph clustering can be transformed into a mathematical optimization problem for the mapped objective function, two-way fuzzy graph clustering can be stated as an approach that attempts to find a solution for the following mathematical program:

Minimize

$$J(U, v) = \sum_{k=1}^{n} \sum_{i=1}^{2} (u_{ik})^2 (d_{ik})^2$$

subject to

- $u_{1k} + u_{2k} = 1$,
- $u_{ik} \geqslant 0$, $1 \leqslant i \leqslant 2$, $1 \leqslant k \leqslant n$,
- $x_i \in V$, $1 \leqslant i \leqslant n$ are vertices in the graph,
- $v_j \in V$, $1 \leqslant j \leqslant 2$ are unknown cluster centers,
- $U = \{u_{ik}\}$ is a $2 \times n$ matrix, where u_{ik} is referred to as the grade of membership of x_k in row i of matrix U.

Objective function J is a nonlinear multi-variable function, and it is difficult for two-way fuzzy graph clustering to obtain an optimal matrix U. For minimizing J, iterative optimization on U and clustering center v can be applied to approximate the minima of the function. In the following lemmas, we discuss necessary conditions on U and v for the mapped objective function.

Lemma 4.1. *Consider the following problem*:

Minimize

$$J(U, v) = \sum_{k=1}^{n} \sum_{i=1}^{2} (u_{ik})^2 (d_{ik})^2$$

subject to

$u_{1k} + u_{2k} = 1$, $1 \leqslant k \leqslant n$,
$u_{ik} \geqslant 0$, $1 \leqslant i \leqslant 2$, $1 \leqslant k \leqslant n$,

*where v is fixed. Then $U = \{u_{ik}\}$ is a global minimum of the problem if for $1 \leqslant k \leqslant n$,
if $x_k \neq v_1$ and $x_k \neq v_2$ then*

$$u_{ik} = \frac{d_{1k}^2 * d_{2k}^2}{d_{ik}^2 \left(d_{ik}^2 + d_{2k}^2 \right)} \quad (\textit{for } 1 \leqslant i \leqslant 2),$$

else

$$u_{ik} = \begin{cases} 1 & \textit{if } x_k = v_i, \\ 0 & \textit{if } x_k \neq v_i \ (\textit{for } 1 \leqslant i \leqslant 2). \end{cases}$$

Proof. By the definition of fuzzy membership, the columns in matrix U are independent. Therefore,

$$\operatorname{Min}\{J(U, v)\} = \operatorname{Min}\left\{ \sum_{k=1}^{n} \sum_{i=1}^{2} (u_{ik})^2 (d_{ik})^2 \right\} = \sum_{k=1}^{n} \operatorname{Min}\left\{ \sum_{i=1}^{2} (u_{ik})^2 (d_{ik})^2 \right\}.$$

As mentioned in the previous definitions, the restricted condition for each column in U is $\sum_{i=1}^{2} u_{ik} = 1$. Further, the minimum function can be formulated as a function F and solved by the Lagrange Multiplier method,

$$F = \sum_{i=1}^{2} (u_{ik})^2 (d_{ik})^2 + \lambda \left(\sum_{i=1}^{2} u_{ik} - 1 \right).$$

The first-order sufficient and necessary conditions for optimality are

$$\frac{\partial F}{\partial \lambda} = \left(\sum_{i=1}^{2} u_{ik} - 1 \right) = 0, \tag{1}$$

$$\frac{\partial F}{\partial u_{ik}} = \left[2(u_{ik})(d_{ik})^2 - \lambda \right] = 0. \tag{2}$$

By (2), we obtain

$$u_{ik} = \frac{\lambda}{2(d_{ik})^2}. \tag{3}$$

Substitute (3) into (1):

$$\sum_{i=1}^{2} u_{jk} = \sum_{j=1}^{2} \left(\frac{\lambda}{2} \right) \frac{1}{(d_{jk})^2} = \frac{\lambda}{2} \sum_{j=1}^{2} \frac{1}{(d_{jk})^2} = 1.$$

Therefore,

$$\frac{\lambda}{2} = \frac{d_{1k}^2 + d_{2k}^2}{d_{1k}^2 * d_{2k}^2}. \tag{4}$$

Substitute (4) into (3), we obtain

$$u_{ik} = \frac{1}{(d_{ik})^2} * \frac{d_{1k}^2 + d_{2k}^2}{d_{1k}^2 * d_{2k}^2} = \frac{d_{1k}^2 * d_{2k}^2}{d_{ik}^2 (d_{1k}^2 + d_{2k}^2)}.$$

The fuzzy membership assignment can be further classified into two different cases. If x_k corresponds to v_i, the fuzzy membership of x_k on cluster i is 1 and that on the other cluster is 0. Thus, u_{ik} is assigned as follows: for $1 \leqslant k \leqslant n$, if $x_k \neq v_1$ and $x_k \neq v_2$ then

$$u_{ik} = \frac{d_{1k}^2 * d_{2k}^2}{d_{ik}^2 (d_{1k}^2 + d_{2k}^2)} \quad (\text{for } 1 \leqslant i \leqslant 2),$$

else

$$u_{ik} = \begin{cases} 1 & \text{if } x_k = v_i, \\ 0 & \text{if } x_k \neq v_i \ (\text{for } 1 \leqslant i \leqslant 2). \end{cases}$$ □

Lemma 4.2. *Consider the following problem*:

Minimize

$$J(U, v) = \sum_{k=1}^{n} \sum_{i=1}^{2} (u_{ik})^2 (d_{ik})^2$$

subject to

$u_{1k} + u_{2k} = 1,\ 1 \leqslant k \leqslant n,$
$u_{ik} \geqslant 0,\ 1 \leqslant i \leqslant 2,\ 1 \leqslant k \leqslant n,$

where U is fixed. Then $v = (v_1, v_2)$ is a global minimum of the problem if v_i is in V such that $J_i(U, y_i)$ is the least.

Proof. Due to U being fixed, all rows in matrix U are independent. Therefore,

$$\begin{aligned} \text{Min}\{J(U, v)\} &= \text{Min}\{J_1(U, v_1) + J_2(U, v_2)\} \\ &= \text{Min}\{J_1(U, v_1)\} \\ &\quad + \text{Min}\{J_2(U, v_2)\}. \end{aligned}$$

Furthermore, the minimization of $J(U, v)$ will depend on the minimization of $\text{Min}\{J_1(U, v_1)\} + \text{Min}\{J_2(U, v_2)\}$. Thus, the center of cluster i for $1 \leqslant i \leqslant 2$ can be assigned by v_i such that $J_i(U, v_i)$ is the least. □

5. Fuzzy clustering and graph bisection

According to Lemmas 4.1 and 4.2, two-way fuzzy graph clustering, via iterative optimization of $J(U, v)$ on U and v, produces a feasible fuzzy graph partition of $V = \{x_1, x_2, \ldots, x_n\}$. The basic steps of the algorithm are as follows:

Algorithm Fuzzy_Graph_Clustering

1. Determine the clustering distance d_{ij}^* between x_i and x_j, $1 \leqslant i,j \leqslant n$.
2. Initialize an arbitrary partition and establish a fuzzy matrix U,
3. Calculate the centers $v = (v_1, v_2)$ using U as follows:
   (1) Determine v_1 such that $J_1(U, v_1)$ is the least,
   (2) Determine v_2 such that $J_2(U, v_2)$ is the least.
4. Calculate a new fuzzy matrix U' using $v = (v_1, v_2)$ as follows:

for $1 \leqslant k \leqslant n$,
if $x_k \neq v_1$ and $x_k \neq v_2$ then

$$u'_{ik} := \frac{d_{1k}^2 * d_{2k}^2}{d_{ik}^2 (d_{1k}^2 + d_{2k}^2)} \quad (\text{for } 1 \leqslant i \leqslant 2),$$

else

$$u'_{ik} := \begin{cases} 1 & \text{if } x_k = v_i, \\ 0 & \text{if } x_k \neq v_i \text{ (for } 1 \leqslant i \leqslant 2), \end{cases}$$

5. If $|u'_{ik} - u_{ik}| < \varepsilon$, for $1 \leqslant i \leqslant 2$, $1 \leqslant k \leqslant n$, then stop; otherwise, $U := U'$, and repeat at step 3.

After U converges, two groups of fuzzy memberships can be generated for all the vertices in the graph. According to the grades of any one group of fuzzy memberships, a vertex ordering will be constructed by sorting the selected group of fuzzy membership, and all the vertices will be separated into two even subsets with minimum cut for graph bisection.

Acknowledgement

We thank David Gries and the anonymous referees for many helpful suggestions.

References

[1] T. Bui, C. Heigham, C. Jones and T. Leighton, Improving the performance of the Kernighan-Lin and simulated annealing graph bisection algorithms, in *Proc. ACM / IEEE 26th Design Automation Conf.* (1989) 775–778.
[2] R.L. Cannon, J.V. Dave and J.C. Bezdek, Efficient implementation of the fuzzy c-means clustering algorithms, *IEEE Trans. Pattern Analysis Machine Intelligence* **8** (1986) 248–255.
[3] C.M. Fiduccia and R.M. Mattheyses, A linear-time heuristic for improving network partitions, in *Proc. ACM / IEEE 19th Design Automation Conf.* (1982) 175–181.
[4] M.R. Garey and D.S. Johnson, Computers and Intractability (Freeman, San Francisco, CA, 1979).
[5] B.W. Kernighan and S. Lin, An efficient heuristic procedure for partitioning graphs, *Bell Systems Tech. J.* **49** (1970) 291–307.
[6] T. Kim, J.C. Bezdek and R. J. Hathaway, Optimality tests for fixed points of the fuzzy c-means algorithm, *Pattern Recognition* **21** (1988) 651–663.
[7] Y.G. Saab and V.B. Rao, Fast effective heuristics for the graph bisectioning problem, *IEEE Trans. Computer-Aided Design* **9** (1990) 91–98.
[8] Y.G. Saab and V.B. Rao, On the graph bisection problem, *IEEE Trans. Circuits Systems I Fund. Theory Appl.* **39** (1992) 760–762.