A Partical Joining Operation on Graphs and A Graphical Distance

その他（別言語等）のタイトル	グラフの部分結合と距離
著者	YAMAGUCHI Tadashi, FOULDS Les
journal or publication title	Memoirs of the Muroran Institute of Technology
volume	47
page range	103-107
year	1997-11-28
URL	http://hdl.handle.net/10258/193

Mem. Muroran Inst. Tech. 47(1997)103 ~ 107

A Partial Joining Operation on Graphs and A Graphical Distance

Tadashi YAMAGUCHI* and Les FOULDS**

(Received 9 May 1997, Accepted 20 August 1997)

The first aim of this paper is to introduce an operation between some graphs, called base graphs and a skeleton graph for the creation of a new combined graph. We call this operation a partial joining of graphs. The second aim is to analyze some distance properties related to the combined graph, base graphs, and the skeleton graph. Finally, some minimization problems concerning the distance sum of the combined graph are also considered for the special case in which the skeleton graph is a tree.

Keywords : Graph, Network, Graph Join, Distance

1 Introduction

In a natural way, graphs or networks can be used for expressing some kinds of binary relations, where vertices or nodes represent processors, system components, or individual people. On the other hand, edges represent some relationship among vertices or nodes. In these circumstances, it is often beneficial to define a new relationship among established networks, in order to combine them into a new network. In this paper we model such cases and analyze such combined graphs or networks. To achieve this we introduce an operation between certain graphs, called base graphs, and another graph, called a skeleton graph, which creates a new, combined graph. We call this operation a partial joining of graphs. We also analyze the properties of graphical distance related to the combined graph, base graphs and the skeleton graph.

Let $G = (V, E)$ denote a simple connected undirected graph with a vertex set V and an edge set E. For the graph theoretic notation and terminology used in this paper, see Foulds(1994). An edge between u and v is denoted as uv. The distance $d(x, y)$ between vertices x and y is the length of a shortest path in G between vertices x and y expressed as the number of edges. The eccentricity $e(x)$ of a vertex x is defined as $e(x) = max\{d(x, y) : y \in V\}$. The radius, denoted by $r(G)$, and the diameter, denoted by $diam(G)$, of G are defined as follows: $r(G) = min\{e(x) : x \in V\}, diam(G) = max\{e(x) : x \in V\}$.

* Common Subject Division(Mathematical Science)

** Department of Management Systems, University of Waikato

A vertex x of G is called a *center vertex* if $e(x) = r(G)$ and a peripheral vertex if $e(x) = diam(G)$ (See Buckley and Harary(1990).) The distance sum(distsum for short) of a vertex x, denoted by $d(x)$, is defined by $d(x) = \sum\{d(x, y) : y \in V\}$, and the distsum of G, denoted by $d(G)$, is defined by $d(G) = \sum\{d(x, y) : x, y \in V\}$. For a subset $U \subseteq V$, let $d(x, U) = min\{d(x, u) : u \in U\}$, and $d(U) = \sum\{d(x, U) : x \in V\}$. The eccentricity $e(U)$ of a vertex subset U is defined as $e(U) = max\{d(x, U) : x \in V\}$. The path between vertices x and y is called the $x - y$ path and the $x - U$ path is meant to represent an $x - u$ path such that $d(x, u) = d(x, U)$ and $u \in U$.

This paper is organized as follows: In Section 2, we present the definition of the partial joining of graphs. Section 3 contains some distance properties related to the combined graph, base graphs, and the skeleton graph. Section 4 contains the minimization problems of distance sum of the combined graph, for the special case in which the skeleton graph is a tree.

2 The Partial Joining of Graphs

2.1 The Definition of The Partial Joining of Graphs

Let $G_i = (V_i, E_i)(i = 1, 2, \ldots, k)$ be graphs with disjoint vertex sets, and U_i be a subset of $V_i(i = 1, 2, \ldots, k)$. A graph $S = (K_S, E_S)$; $K_S = \{1, 2, \ldots, k\}$ is also given. When $i, j \in K_S$ and $ij \in E_S$, E_{ij} is assumed given, where $E_{ij} \subseteq U_i \times U_j = \{uv : u \in U_i, v \in U_j\}$. Now we define a new graph G, called a *combined graph* as follows:

$G = (V_G, E_G)$, where $V_G = \cup_{i=1}^{k} V_i$ and $E_G = \cup_{i=1}^{k} E_i \cup E'$, where $E' = \cup_{ij \in E_S} E_{ij}$.
This operation for creating of the combined graph is called the *partial joining of graphs*, and the G_i's, S, and the elements of E', are called base graphs, a skeleton graph, and newly added edges(added edges for short), respectively.

Example : Let $G_1 = G_2 = G_3 = K_4$(the complete graph with four vertices) be base graphs. Let S be the path P_3 (the path with three vertices) as a skeleton graph. The subsets are given by: $U_1 = \{a\}, U_2 = \{b\}$,and $U_3 = \{c, d\}$. Refer to Figure 1. Let $E_{12} = \{ab\}$ and $E_{23} = \{bc, bd\}$. The combined graph is shown in Figure 1.

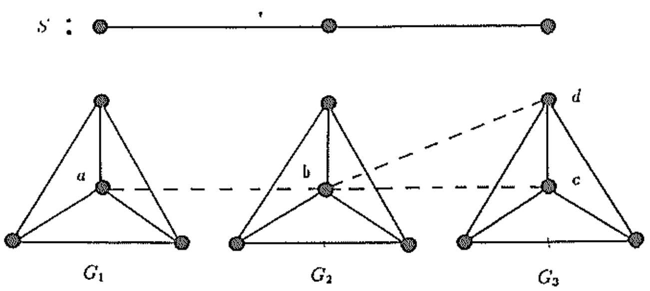

Fig. 1.
Examples of : Base graphs, a skeleton graph, and a combined graph.

2.2 Distances of Newly Added Edges

The length of the newly added edge $uv \in E_{ij}$ is considered as follows:
Case 1 : $d(u, v) = 0$,
Case 2 : $d(u, v) = 1$, and
Case 3 : $d(u, v)$ depends on u and v.
One interpretation of Case 1 concerns the identification of two end vertices of added edge. Case 2 treats edges of the base graph and the added edges at the same level. Case 3 is the most natural in practical networks. The authors plan to present results on Case 3 elsewhere.

3 The distance properties of combined graphs

In this section, some distance properties are derived. Case 1, mentioned in subsection 2.2, is considered, and some properties of Case 2 are noted.

Let $G_i = (V_i, E_i), U_i \subseteq V_i (i = 1, \ldots, k)$, and S be a tree with k vertices. Let the combined graph $G = (V_G, E_G)$ be as defined in subsection 2.1. Further, assume that each graph $(V_i \cup V_j, E_{ij})$ is connected for all $ij \in E_S$, and assume that the length of each edge in each E_i is 1 and the length of added edge of E_{ij} is 0. Moreover, $d_i(x, y)$, $e_i(x)$ and $d_i(x)$ mean the distance between vertices, the eccentricity and the distance in G_i , respectively. In Case 1, we note $d(x, y) = 0$, if $x, y \in \cup_{i=1}^{k} U_i$.

Property 1.
(1.1)If $x, y \in V_i$, then
$d(x, y) = min\{d_i(x, y), d_i(x, U_i) + d_i(U_i, y)\} (i = 1, 2, \ldots, k)$.
(1.2)If $x \in V_i$, $y \in V_j$ for $i \neq j$, then

$$d(x, y) = d_i(x, U_i) + d_j(U_j, y).$$

Proof:(1.1) By definition, the length of the $x - y$ shortest path in G_i is $d_i(x, y)$. On the other hand, the length of the $x - y$ path through vertices of U_i in the combined graph is $d_i(x, U_i) + d_i(U_i, y)$. So $d(x, y)$ is the smaller of $d_i(x, y)$ and $(d_i(x, U_i) + d_i(U_i, y))$.

(1.2) It follows from the fact that any $x - y$ shortest path between $x \in G_i$ and $y \in G_j$ is constructed by concatenating the $x - U_i$ path, the $x' - y'$ path (consisting of only newly added edges for some $x' \in U_i$ and $y' \in U_j$) and the $U_j - y$ path. However the length of this second path is zero. Thus Property (1.2) follows. □

The distance $d(x, y)$ does not satisfy one of axioms of a metric. That is, the first axiom should be changed, such that $d(x, y) = 0$ iff $x = y$ or $x, y \in \cup U_i$. But the symmetry and triangle inequality are satisfied. Because of that, the length of the added edge is zero, which means that the two end vertices of the edge are one and the same. And $d(x, y)$ is considered as if it is the distance of so-called condensed graph, which is derived by identifying all end vertices of the newly added edges.

3.1 The Diameter of Combined Graphs

The following property about the diameter of combined graph is established:

Property 2.
$diam(G) \leq 2max\{e_i(U_i) : i = 1, 2, \ldots, k\}$.

Proof: By the definition of diameter, $diam(G) = max\{d(x, y) : x, y \in V_G\}$.
Two cases are considered separately:
(2.1) $x, y \in V_i$, and (2.2) $x \in V_i, y \in V_j (i \neq j)$.
Case(2.1) : When $x, y \in V_i$, by Property(1.1),

$$\begin{aligned} d(x, y) &= min\{d_i(x, y), d_i(x, U_i) + d_i(U_i, y)\} \\ &\leq d_i(x, U_i) + d_i(U_i, y) \\ &\leq 2e_i(U_i). \end{aligned}$$

Case (2.2) : When $x \in V_i$, and $y \in V_j (i \neq j)$, by Property(1.2),

$$\begin{aligned} d(x, y) &= d_i(x, U_i) + d_j(U_j, y) \\ &\leq e_i(U_i) + e_j(U_j) \\ &\leq 2max\{e_i(U_i), e_j(U_j)\}. \end{aligned}$$

By combining the above, Property 2 follows. □

The bound in this property is sharp as can be seen in the example in Figure 2, where $e(U_1) = e(U_2) = 3$ and $diam(G) = 6$. From Property 2 it can be seen that, in order to minimize $diam(G)$, it is desirable to keep $e_i(U_i), e_j(U_j)$ as small as possible. Thus the problem is reduced to the multi-center problem of each G_i.

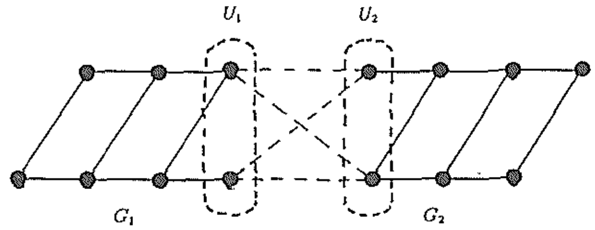

Fig. 2.
An example which shows that the inequality of Property 2 is sharp.

In particular, when $|U_i| = 1$, let $U_i = \{u_i\}(i = 1, 2, \ldots, k)$.

Property 1′.
(1.1′) If $x, y \in V_i$, then $d(x, y) = d_i(x, y)(i = 1, 2, \ldots, k)$.
(1.2′) If $x \in V_i, y \in V_j$ for $i \neq j$, then
$d(x, y) = d_i(x, u_i) + d_j(u_j, y)$.

Property 2′.
$diam(G) = max_{i,j}\{diam(G_i), diam(G_j), e_i(u_i)+e_j(u_j)\}$.

So far we have treated only the case in which the length of the added edge is zero. When this length is 1 and the number of base graphs is two, we obtain :

Property 1″.
(1) If $x, y \in V_i$, then
$d(x, y) = min\{d_i(x, y), d_i(x, U_i) + 2 + d_i(y, U_i)\}(i = 1, 2)$.
(2) If $x \in V_1, y \in V_2$, then $d(x, y) = d_1(x, U_1) + 1 + d_2(U_2, y)$.
(3) $diam(G) \leq 2max\{e_1(U_1) + 1, e_2(U_2) + 1\}$.
(4) When $|U_1| = |U_2| = 1$,

$$diam(G) = max\{diam(G_1), diam(G_2), e_1(U_1)+e_2(U_2)+1\}.$$

3.2 The Distance Sum(Transmission Number) of Combined Graphs

In this subsection, a relation between the distance sum of the combined graph and base graphs is derived. Let a graph G be defined in this section, and the assumptions about the lengths of edges be as given in the beginning of Section 3.

Then let :

$$d_i(G_i) = \sum_{x,y \in V_i} d_i(x, y)(i = 1, 2, \ldots, k),$$

$$d(G_i) = \sum_{x,y \in V_i} d(x, y)(i = 1, 2, \ldots, k)$$

and

$$d(G_i, G_j) = \sum_{x \in V_i, y \in V_j} d(x, y) \qquad (i \neq j).$$

The distance sum(transmission number) of G can be expressed as the following:

$$d(G) = \sum_{1 \leq i \leq k} d(G_i) + \sum_{1 \leq i \neq j \leq k} d(G_i, G_j).$$

By Property (1.1), for $x, y \in V_i$,

$$\begin{aligned} \Delta_i(x, y) &\equiv d_i(x, y) - d(x, y) \\ &= d_i(x, y) \dot{-} \{d_i(x, U_i) + d_i(U_i, y)\}(i = 1, 2, \ldots, k), \end{aligned}$$

$$\text{where } a \dot{-} b = \begin{cases} a - b & \text{if } a \geq b, \\ 0 & \text{if } a < b. \end{cases}$$

Then we have:

Property 3.

$$d(G_i) = d_i(G_i) - \sum_{x,y \in V_i} \Delta_i(x, y)(i = 1, 2, \ldots, k).$$

Next, for $x \in V_i, y \in V_j$, using the relation $d(x, y) = d_i(x, U_i) + d_j(U_j, y)$, we have:

$$\begin{aligned} d(G_i, G_j) &= \sum_{x \in V_i, y \in V_j} d(x, y) \\ &= \sum_{x \in V_i, y \in V_j} d_i(x, U_i) + \sum_{x \in V_i, y \in V_j} d_j(U_j, y) \\ &= |V_j| \sum_{x \in V_i} d_i(x, U_i) + |V_i| \sum_{y \in V_j} d_j(U_j, y) \\ &= |V_j| d_i(U_i) + |V_i| d_j(U_j). \end{aligned}$$

By summing up these, we obtain :

Property 4.

$$\begin{aligned} d(G) = \sum_{i=1}^{k}\{d_i(G_i) &- \sum_{x,y \in V_i} \Delta_i(x, y)\} \\ &+ \sum_{i \neq j}\{|V_j| d_i(U_i) + |V_i| d_j(U_j)\}. \end{aligned}$$

This provides a guide as to how to choose the U_i's in order to make the distance sum as small as possible:
(4.1) Choose U_i so that $d_i(U_i)$ are as small as possible, and
(4.2) take each $\Delta_i(x, y)$ as large as possible. That is, if the cardinalities of U_i are given, a realization of (4.1) is reduced to so-called multi-median problem. Moreover, in order to realize (4.2), it is necessary to make the radius of the ball B_j^i, the ball of u_j^i in G_i, as small as possible, where $B_j^i \equiv \{x \in V_i : d(x, u_j^i) < d(x, u_h^i) \text{ for all } h \neq j\}$, and u_j^i, u_h^i are the vertices of U_j.
We conjecture that the realization of (4.1) contributes to the realization of (4.2).

4 Some Minimization Problems

We now assume that for any base graphs $\{G_i\}$ and each singleton $U_i = \{u_i\}(i = 1, 2, ..., k)$, the length of any newly added edge is one, and the skeleton graph is a tree with k vertices. Each distance sum $d(G_i, G_j)$ between G_i and G_j for $i \neq j$ is described as follows:

$$d(G_i, G_j) = |V_j| d_i(u_i) + |V_i||V_j| d(u_i, u_j) + |V_i| d_j(u_j).$$

Let DS be the summation of these. That is,

$$DS = \sum_{1 \leq i \neq j \leq k} d(G_i, G_j).$$

Moreover, let $|V_i| = w_i(i = 1, 2, ..., k)$, and rewrite the above as:

$$d(G_i, G_j) = w_j d_i(u_i) + w_i w_j d(u_i, u_j) + w_i d_j(u_j),$$

and

$$DS = \sum_{1\le i\ne j\le k} w_j d_i(u_i) + \sum_{1\le i\ne j\le k} w_i w_j d(u_i, u_j) + \sum_{1\le i\ne j\le k} w_i d_j(u_j).$$

These the first and third terms are equivalent, and do not depend on the form of the skeleton graph or the correspondence of vertices between the skeleton graph and the base graphs. In order to make DS small, we have to consider the second term.

So we now analyze the following two problems:

Problem 1: Minimize the value of DS by changing the correspondence between the vertices of the skeleton graph and the base graphs.,

Problem 2: Find the form(type) of the skeleton graph among trees with k vertices, such that DS is minimized.

4.1 Minimizing of the value of DS

For the first and third terms in DS, it is necessary to minimize $d_i(u_i)$. That is, it is enough to choose u_i as a median of G_i. Let W be the summation of the second terms. That is,

$$W = \sum_{1\le i\ne j\le k} w_i w_j d(u_i, u_j).$$

We now consider the minimization of W. In other words, we wish to assign weights $\{w_i\}$ to the vertices of the skeleton graph S so that W is minimized. This problem can be modelled as a quadratic assignment problem(QAP) as follows.

Definition: Let $c_{ijst} = w_s w_t d(u_i, u_j)$ and

$$x_{si} = \begin{cases} 1 & \text{if the weight } w_s \text{ is assigned to the vertex } i, \\ 0 & \text{otherwise,} \end{cases}$$

where $i, j, s, t = 1, 2, ..., k$.

Constraints:

$$\sum_{s=1}^{k} x_{si} = 1 \quad ; \quad i = 1, 2, ..., k.$$

[Each vertex i is assigned exactly one weight.]

$$\sum_{i=1}^{k} x_{si} = 1 \quad ; \quad s = 1, 2, ..., k.$$

[Each weight w_s is assigned exactly one vertex.]

$$x_{si} = 0 \text{ or } 1 \quad ; \quad s, i = 1, 2, ..., k.$$

[Each weight w_s is either assigned to the vertex i, or it is not.]

Objective:

$$\text{Minimize} \quad W = \sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{s=1}^{k}\sum_{t=1}^{k} c_{ijst} x_{si} x_{tj}.$$

Algorithms that guarantee optimality for the QAP have been reported by Gilmore(1962) and Lawler(1963). Recent refinements have been discussed by Li*et al.* (1994). However the problem is NP-hard, which reinforces the quest for efficient heuristics for it. Kelly*et al.* (1994) have recently investigated the feasibility of employing tabu search, genetic algorithms and simulated annealing to provide a basic effective QAP heuristics. However a simple heuristic approach can be developed as follows.

Let an assignment, $X = \{x_{ij}\}_{k\times k}$, be given. Note that the vertices of S are numbered : $1, 2, ..., k$. In this assignment, let the weight of vertex i be w_i. If necessary we change numbering of vertices. Let $d_{ij} = d(u_i, u_j)$ for the sake of simplicity and let the summation of products of weights and distances be :

$$W_i = \sum_{1\le j\le k} w_j d_{ij} (i = 1, 2, ..., k).$$

Now consider a new assignment X', which is the same as X except for the fact that the weights of vertices p and q are interchanged. Let the values of W by the assignments X and X' be $W(X)$ and $W(X')$, respectively.

Then we have :

Property 5.

$$W(X) - W(X') = 2(w_p - w_q) \times \{(W_p - w_q d_{pq}) - (W_q - w_p d_{qp})\}.$$

Proof: Let $\boldsymbol{x}$ be a column vector of weights $\{w_i\}$ as its components, and $\boldsymbol{y}$ be a vector obtained by exchanging the pth component and the qth component of $\boldsymbol{x}$. Moreover, let D be the distance matrix of the skeleton graph S. Then, we get the relation:

$$W(X) = \boldsymbol{x}^t D \boldsymbol{x},$$

$$W(X') = \boldsymbol{y}^t D \boldsymbol{y}.$$

Note

$$\boldsymbol{y} = \boldsymbol{x} + (w_p - w_q)\boldsymbol{z},$$

where $\boldsymbol{z}$ is a vector in which the pth component is -1, qth component is 1, and others are 0. By combining these three equations we obtain the result. □

By this result, when $w_p > w_q$ and $W_p - W_q > (w_q - w_p)d_{pq}$, we can improve W if the weights of the pth vertex and the qth vertex are interchanged. Roughly speaking, if we assign relatively large weights to the vertices of the central part of S and relatively small weights to the vertices at the peripheral part of S, we obtain a relatively small W.

4.2 The optimal form of the skeleton

In this subsection, we wish to identify the form of the skeleton graph for minimizing W. We consider any tree with k vertices, and any assignment of weights $\{w_i\}$ to vertices of S. Now, for $uv \in E$ of tree S, let $T_{u\backslash v}$ be the maximal subtree containing u but not containing v, and $T_{v\backslash u}$ be the maximal subtree containing v but not containing u. Moreover, let the sum of weights of $T_{u\backslash v}$ and $T_{v\backslash u}$ be $bw(u\backslash v)$ and $bw(v\backslash u)$, respectively. Then the following result is derived.

Property 6.

Assume that the diameter of the skeleton graph S is greater than two. Then there exists an edge uv and a vertex v with an adjacent leaf vertex x, such that $bw(u\backslash v) > bw(v\backslash u) - w_x$, where w_x is the weight of the vertex x.

Proof: We consider an edge ab where the vertices a and b are not leaves of S and $bw(a\backslash b) \geq bw(b\backslash a)$. There exists such an edge because the diameter is greater than two. Let x be the one of leaves of $T_{b\backslash a}$ and the adjacent vertex of x be v. (It might be b.) Moreover, let the adjacent vertex of v on the $a-v$ path be u. (It might be a.) Then:

$$bw(u\backslash v) \geq bw(a\backslash b) \geq bw(b\backslash a) \geq bw(v\backslash u) > bw(v\backslash u) - w_x.$$

Hence the result follows. □

By this result, the value W can be improved if an edge vx is removed and an edge ux is added : Let S' be the tree obtained by removing vx from S and adding ux to S, and let W_S be the distance sum of S , and $W_{S'}$ be the distance sum of S'. Then we obtain the relation:

$$W_S - W_{S'} = bw(u\backslash v) - (bw(v\backslash u) - w_x) > 0.$$

By repeating this deleting and adding procedure, we obtain a tree with diameter two (the so-called *star graph.*) Moreover, if we assign the largest weight to the center of the star, we obtain the required tree and the assignment.

5 Summary

An operation, the partial joining, is proposed, and some properties concerning the graphical distance of the combined graph and base graphs are established. The instance of zero distance of edge being added is a means of identifying the end vertices of the edge.

Guide lines for decreasing the diameter and the distance sum of the combined graph are derived. By restricting the skeleton graph to a tree, a strategy for decreasing the distance sum between base graphs is established.

References

[1] Buckley, F., Harary, F. : **Distance in Graphs**, Addison-Wesley, 1990.

[2] Foulds, L. R. : **Graph Theory Applications**, 2nd edition, Springer, 1994.

[3] Gilmore, P. C. : Optimal and Suboptimal Algorithms for the Quadratic Assignment Problem, **SIAM Journal on Applied Mathematics**, Vol.10, No.2, 1962.

[4] Kelly, J. P., Laguna, M., Glover, F. : A study of diversification strategies for the quadratic assignment problem, **Computers & Ops. Res.**, Vol.21, No. 8, 1994.

[5] Lawler, E. L. : The Quadratic Assignment Problem, **Management Science**, Vol.9, No.1, 1963.

[6] Li, Y., Pardalos, P. M., Ramakrishnan, K. G. and Resende, M. G. C. : Lower bounds for the quadratic assignment problem, **Annals of OR**, Vol. 50, No.1, 1994.

グラフの部分結合と距離

山口 忠* 、L. フォルド**

概 要

本論文では、いくつかのベースグラフとスケルトングラフ間に作用する操作を提案する。これをグラフの部分結合という。この操作によって得られるグラフ（結合グラフ）のグラフ的距離に関する性質を調べる. 更に、スケルトングラフが木グラフの場合についての結合グラフの最適対応問題即ち、ベースグラフとスケルトングラフの対応問題を扱う。

キーワード： グラフ、ネットワーク、グラフ結合、距離

*共通講座（数理科学）、** ワイカト大学　経営システム学部