Journal of Global Optimization **6:** 193–205, 1995.

On Complexity of Subset Interconnection Designs*

DING-ZHU DU† and DEAN F. KELLEY
Computer Science Department, University of Minnesota, Minneapolis, MN 55455, U.S.A.

(Received: 15 June 1992; accepted: 18 November 1994)

Abstract. Given a set X and subsets $X_1, \dots, X_m$, we consider the problem of finding a graph G with vertex set X and the minimum number of edges such that for $i = 1, \dots, m$, the subgraph G_i induced by X_i is connected. Suppose that for any α points $x_1, \dots, x_\alpha \in X$, there are at most βX_i's containing the set $\{x_1, \dots, x_\alpha\}$. In the paper, we show that the problem is polynomial-time solvable for $(\alpha \leqslant 2, \beta \leqslant 2)$ and is NP-hard for $(\alpha \geqslant 3, \beta = 1)$, $(\alpha = 1, \beta \geqslant 6)$, and $(\alpha \geqslant 2, \beta \geqslant 3)$.

Key words: Subset interconnection design, globally optimal, polynomial time.

1. Introduction

Given a set X and subsets $X_1, \dots, X_m$, we consider the problem of finding a graph G with vertex set X and the minimum number of edges such that for $i = 1, \dots, m$, the subgraph G_i induced by X_i is connected. We will refer this problem as the SID (subset interconnection designs). This combinatorial optimization problem has many applications in real world [1,2,5]. Let us mention one of them as follows.

A vaccum system contains many valves. The function of valves is to give different connections for different work at the different stage. Now, we use a vertex to represent a part separated by valves and an edge to represent a valve. Then the SID corresponds to the following: Given m connection requirements, design a vaccum system with minimum number of valves. The importance of decreasing the number of valves in the vaccum system is not only on saving money but also on increasing the degree of vaccum.

The SID is NP-hard. Du [3] gave a sufficient optimality condition and indicated that the condition is necessary for $m = 2$. Tang [11] showed that the condition is also necessary for $m = 3$, but is not necessary for $m \geqslant 4$. In this paper, we consider the restriction denoted by (α, β) as follows: For any α points $x_1, \dots, x_\alpha \in X$, there are at most βX_i's containing the set $\{x_1, \dots, x_\alpha\}$ where α and β are two given natural numbers.

---

* Support in part by the NSF under grant CCR-9208913 and CCR-8920505.

† Part work was done while this author was visiting at DIMACS and on leave from Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing.

Suppose $\alpha' \geqslant \alpha$ and $\beta' \geqslant \beta$. Clearly, (α, β) implies (α', β'), then it is polynomial-time solvable for (α, β) and if the problem is NP-hard for (α, β), then it is NP-hard for (α', β'). We will show that the problem is polynomial-time solvable for ($\alpha \leqslant 2$ and $\beta \leqslant 2$) and is NP-hard for ($\alpha \geqslant 3$ and $\beta = 1$), ($\alpha = 1$ and $\beta \geqslant 6$), and ($\alpha \geqslant 2$ and $\beta \geqslant 3$).

A graph G with vertex X is called a *feasible graph* for $(X_1, X_2, \ldots, X_m)$ if for any $i = 1, 2, \ldots, m$, the subgraph G_i induced by X_i is connected. A feasible graph is *minimum* if G is an optimal solution for the SID. For a graph G, we denote by $V(G)$ the vertex set of G, by $E(G)$ the edge set of G, and by $||G||$ the number of edges in G. For a set Y, we denote by $|Y|$ the number of elements in Y. For example, $|E(G)| = ||G||$. Let $G = (V, E)$ and $G' = (V', E')$. Then the union, the intersection, and the difference of graphs G and G' are defined by $G \cup G' = (V \cup V', E \cup E")$, $G \cap G' = (V \cap V', E \cap E')$, and $G \backslash G' = (V, E \backslash E')$, respectively. The symmetric difference of G and G' is defined by $G \oplus G'\text{“}(G \backslash G') \cup (G' \backslash G)$. Before presenting our results, we make four conventions:

(1) We assume $X = \cup_{i=1}^{m} X_i$, without loss of generality, since every minimum feasible graph has no edge incident to a point in $X \backslash \cup_{i=1}^{m} X_i$.
(2) We assume $|X_i| \geqslant 2$ for all i since X_i with $|X_i| = 1$ can be deleted.
(3) We assume that every feasible graph G satisfies $G = \cup_{i=1}^{m} G_i$ since an edge not in $\cup_{i=1}^{m} G_i$ can be deleted without changing the feasibility.
(4) All cycles and paths we talk about in this paper are simple.

2. Preliminary

It was proved in [3] that every graph satisfying the following condition is a minimum feasible graph.

($*$) *For any* $i, j = 1, 2, \ldots, m$, *the subgraph* G_{ij} *induced by* $X_i \cap X_j$ *is a tree.*

For $(\alpha = 1, \beta = 2)$, a graph satisfying the condition ($*$) can be constructed in the following way: For every pair $i, j \in \{1, 2, \ldots, m\}$ with $i \neq j$ and $X_i \cap X_j \neq \emptyset$, connect all points in $X_i \cap X_j$ into a tree. Let G' denote the resulting graph. Since $\alpha = 1$ and $\beta = 2$, for different pairs $\{i, j\} \neq \{i', j'\}$, $(X_i \cap X_j) \cap (X_{i'} \cap X_{j'}) = \emptyset$. Thus, for all i, the graph G_i' induced by X_i from G' is a forest. Connecting each forest G_i' into a tree, we obtain a graph G satisfying the condition ($*$).

For $(\alpha = 2, \beta = 2)$, it is not so easy to find a minimum feasible graph. In the next section, we will present a polynomial-time algorithm to compute it. Before doing so, we prove some general properties in the rest of this section.

Let $K(i)$ be the number of (connected) components of a graph with vertex set X_i obtained by joining all vertices in $X_i \cap X_j$ into a tree for all $j = 1, \ldots, m$ and $j \neq i$. Let $X_{i1}, \ldots, X_{iK(i)}$ denote the vertex sets of these $K(i)$ components. Clearly, $X_{i1}, \ldots, X_{iK(i)}$ form a disjoint partition of X_i. Let T be a set of $K(i) - 1$ edges which interconnect the $K(i)$ components into a connected graph.

LEMMA 2.1. *If $K(i) \geqslant 2$, then any disjoint union of T and a minimum feasible graph for $(X_1, \ldots, X_{i-1}, X_{i1}, \ldots, X_{iK(i)}, X_{i+1}, \ldots, X_m)$ must be a minimum feasible graph for $(X_1, \ldots, X_m)$.*

Proof. Without loss of generality, we assume $i = m$ since, otherwise, we can rearrange the indices. Let G be a minimum feasible graph for $(X_1, \ldots, X_m)$. Note that

$$\begin{aligned}||G|| &= || \cup_{j=1}^{m-1} G_j|| + ||G_m \setminus \cup_{j=1}^{m-1} G_j|| \\ &= || \cup_{j=1}^{m-1} G_j|| + ||G_m \setminus \cup_{j=1}^{m-1} (G_j \cap G_m)||.\end{aligned}$$

Suppose that $\cup_{j=1}^{m-1}(G_j \cap G_m)$ has k^* components. By the minimality of G, $G_m \setminus \cup_{j=1}^{m-1} (G_j \cap G_m)$ contains exactly $k^* - 1$ edges which interconnect those k^* components. Note that the vertex set of each component of $\cup_{j=1}^{m-1}(G_j \cap G_m)$ is a subset of X_{mk} for some $k = 1, \ldots, K(m)$. Thus, we can use $k^* - 1$ edges interconnecting the k^* components in the following way: First, for $k = 1, \ldots, K(m)$, we join the components with vertex set in X_{mk} together. Then we use T to interconnect the resulting $K(m)$ components. Let G' be the obtained graph. Clearly, G' is also a minimum feasible graph. In addition, G' is a disjoint union of T and a minimum feasible graph for $(X_1, \ldots, X_{m-1}, X_{m1}, \ldots, X_{mK(m)})$. Since every disjoint union of T and a minimum feasible graph for $(X_1, \ldots, X_{m-1}, X_{m1}, \ldots, X_{mK(m)})$ is a feasible graph for $(X_1, \ldots, X_m)$ with the number of edges as G' has, it must also be a minimum feasible graph for $(X_1, \ldots, X_m)$. □

Denote $I(x) = \{i | x \in X_i\}$ for every vertex $x \in X$ and $I(u) = I(x) \cap I(y)$ for every edge u with endpoints x and y.

By Lemma 2.1, we may assume that $K(i) = 1$ for all i. With this assumption, the SID has the following property.

LEMMA 2.2. *If $K(i) = 1$ for all $i = 1, \ldots, m$, then there exists a minimum feasible graph G such that $|I(u)| \geqslant 2$ for every edge u of G.*

Proof. Suppose G is a minimum feasible graph and u is an edge of G such that $|I(u)| = 1$. Without loss of generality, assume $I(u) = \{m\}$. Clearly, u belongs to $G_m \setminus \cup_{j=1}^{m-1} (G_j \cap G_m)$. Deleting u breaks G_m into two parts. Since $K(m) = 1$, we can connect the two parts into one by an edge v with $|I(v)| \geqslant 2$. □

3. $(\alpha = 2, \beta = 2)$

We first consider the restriction $(\alpha = 2, \beta = 2)$, i.e., for any two distinct points x and y in X, $|I(x) \cap I(y)| \leqslant 2$. We also assume $|K(i)| = 1$ for all i. By Lemma 2.2, there exists a minimum feasible graph for $X_1, \ldots, X_m)$ such that every edge u contains two indices, i.e., $|I(u)| = 2$.

Let K^* be a graph with vertex set X and all edges which contains two indices. Then K^* has a subgraph which is a minimum feasible graph for $(X_1, \ldots, X_m)$. To find it, let us first study a graph $H(G)$ constructed for a given feasible graph G in the following way:

(1) For each edge $u \in E(K_i^*\backslash G)$, choose a cycle in $G_i \cup u$ containing u. Then choose a set of cycles in G_i together with the already chosen cycles to form a maximal independent set C_i of cycles in K_i^*, where K_i^* is the subgraph of K^*, induced by X_i. Here, a set of cycles is said to be *independent* if they are linear independent in a linear vector space generated by them. (See [9] for detail.) Each chosen cycle Q with the index i together forms a vertex in $H(G)$, denoted by $\langle Q, i\rangle$. Note that for $i \neq j$, $\langle Q, i\rangle$ and $\langle Q, j\rangle$ denote different vertices. The vertex subset of $H(G)$ is exactly the collection of such pairs, i.e., $H(G) = \{\langle Q, i\rangle | Q \in C_i,\}$.

(2) $H(G)$ has an edge between $\langle Q, i\rangle$ and $\langle Q', i'\rangle$ if and only if $i \neq i'$ and Q and Q' have at least one edge in common.

Assume $v \in E(K^*\backslash G)$ and $I(v) = \{i, i'\}$. Then $H(G)$ has two vertices $\langle Q, i\rangle$ and $\langle Q', i'\rangle$ such that both Q and Q' contain edge v. Hence, $H(G)$ has an edge between $\langle Q, i\rangle$ and $\langle Q', i'\rangle$. We denote this edge by $m(v)$. Note that edge $m(v)$ may not be uniquely determined. In the case of existence of many choices, we choose one arbitrarily to be $m(v)$. Now, let $M(G) = \{m(v) | v \in E(K^*\backslash G)\}$.

LEMMA 3.1. *$M(G)$ is a matching in graph $H(G)$.*

Proof. Let u and $\bar{u}$ be two distinct edges in $K^*\backslash G$. Let $\langle Q, i\rangle$ and $\langle \bar{Q}, bari\rangle$ be endpoints of $m(u)$ and $m(\bar{u})$, respectively. Since $\bar{u}$ is not in G and $\bar{Q}\backslash u$ is in G, Q does not contain $\bar{u}$. However, $\bar{Q}$ contains $\bar{u}$. Therefore, Q and $\bar{Q}$ are different. Hence, $\langle Q, i\rangle \neq \langle \bar{Q}, \bar{i}\rangle$. This means that any two edges in M cannot have an endpoint in common. Thus, M is a matching in H. □

LEMMA 3.2. *If a feasible graph G is not minimum, then $M(G)$ is not the maximum matching of $H(G)$.*

Proof. Since G is not minimum, there exists a feasible graph G^* such that $||G^*|| < ||G||$. We will find a matching M^* in $H(G)$ such that $|M^*| = ||K^*\backslash G^*|| > ||K^*\backslash G|| = |M(G)|$. To do so, consider an edge $u \in E(K_i^*\backslash G_i)$. Let x and y be two endpoints of u. Then both x and y belong to X_i. Since G is a feasible graph, G_i has a path connecting x and y. This path together with u forms a cycle Q_u in $G_i \cup u$ containing u.

Clearly, the set of cycles Q_u for $u \in K_i^*\backslash G_i$ is independent. In fact, if we write each Q_u into a 0-1 raw vector such that each component of the vector corresponding to an edge of K_i^* and an edge is in Q_u if and only if the corresponding component equals 1, then all columns corresponding to edges in $K_i^*\backslash G^*$ form the identity matrix of order $|K_i^*\backslash G^*|$. From this fact, we can also see that in C_i, the number of cycles which contain edges in $K_i^*\backslash G^*$ is at least $||K_i^*\backslash G^*||$. In fact, all Q_u's can be written as linear combinations of cycles in C_i, so, if we put all cycles in C_i as raw vectors into a matrix, then all columns corresponding to edges in $K_i^*\backslash G^*$ form a submatrix of rank $||K_i^*\backslash G^*||$ and hence this submatrix has at least $||K_i^*\backslash G^*||$ nonzero rows.

Similarly, for each subset S of edges in $K_i^*\backslash G^*$, the number of cycles in C_i, containing some edges in S, is at least $|S|$. By König-Hall's theorem, we can find $\|K_i^*\backslash G^*\|$ distinct cycles $Q_{i,u}$ from C_i such that $Q_{i,u}$ contains edge u in $K_i^*\backslash G^*$.

Now, for each edge u in $K^*\backslash G^*$, let $I(u) = \{i, i'\}$. Then, we have two cycles $Q_{i,u}$ and $Q_{i',u}$ in C_i, both containing u. By the definition of $H(G), H(G)$ has an edge between $\langle Q_{i,u}, i\rangle$ and $\langle Q_{i',u}, i'\rangle$. This edge is denoted by $m^*(u)$. Define $m^* = \{m^*(u)|u \in E(K^*\backslash G)\}$. We claim that M^* is a matching in $H(G)$.

To prove the claim, let us consider two distinct edges u and $\bar{u}$ in $K^*\backslash G^*$. Suppose that $\langle Q_{i,u}, i\rangle$ and $\langle Q_{\bar{i},\bar{u}}, \bar{i}\rangle$ are endpoints of u and $\bar{u}$, respectively. If $i \neq \bar{i}$, then it is clear that $\langle Q_{i,u}, i\rangle \neq \langle Q_{\bar{i},\bar{u}}, \bar{i}\rangle$. If $i = \bar{i}$, then we can see from the above choice that $Q_{i,u} \neq Q_{i,\bar{u}}$ since $u \neq \bar{u}$. Therefore, $\langle Q_{i,u}, i\rangle \neq \langle Q_{\bar{i},\bar{u}}, \bar{i}\rangle$. This means that any two edges in M^* have no endpoint in common, that is, M^* is a matching. Clearly, $|M^*| = \|K^*\backslash G^*\|$. Thus, $M(G)$ is not maximum. □

By Lemma 3.2, if a feasible graph G is not minimum, then there exists a matching M^* in $H(G)$ such that $|M^*| > |M(G)|$. The symmetric difference $M(G) \oplus M^*$ is a disjoint union of paths and cycles. Since $|M(G)| < |M^*|, M(G) \oplus M^*$ has a path P with more edges in M^*. Clearly, this path P must satisfy the following conditions:

(A1) The path is alternating for $M(G)$, i.e., the edges in the path are alternatively in $H(G)\backslash M(G)$ and $M(G)$. Thus, it has even number of vertices. We may write it as $\{\langle Q_1, i_1\rangle, \ldots, \langle Q_{2k}, i_{2k}\rangle\}$.

(A2) $\langle Q_1, i_1\rangle$ and $\langle Q_{2k}, i_{2k}\rangle$ are not covered by $M(G)$. (Therefore, Q_1 and Q_{2k} are in G. In fact, a vertex $\langle Q, i\rangle$ of $H(G)$ is not covered by $M(G)$ if and only if Q is in G_i).

(A3) G has k distinct edges $v_1, \ldots, v_k$ such that $v_j \in E(Q_{2j-1} \cap Q_{2j})$ for $j = 1, \ldots, k$.

Such a path $\{\langle Q_1, i_1\rangle, \ldots, \langle Q_{2k}, i_{2k}\rangle\}$ in $H(G)$ will be called an *augmenting path* in $H(G)$.

LEMMA 3.3. *A feasible graph G is a minimum feasible graph if and only if $H(G)$ has no augmenting path.*

Proof. Suppose that G is not a minimum feasible graph. By Lemma 3.2 and the discussion after Lemma 3.2, $H(G)$ has an augmenting path.

Conversely, suppose that $H(G)$ has an augmenting path $\{\langle Q_1, i_1\rangle, \ldots, \langle Q_{2k}, i_{2k}\rangle\}$ we prove by induction on k that G is not a minimum feasible graph. For $k = 1$, since Q_1 and Q_2 are in $G, G\backslash v_1$ is a feasible graph where $v_1 \in E(Q_1 \cap Q_2)$. Thus, G is not minimum. For $k > 1$, assume that for $j = 1, \ldots, k-1$, $m(u_j)$ is the edge between vertices $\langle Q_{2j}, i_{2j}\rangle$ and $\langle Q_{2j+1}, i_{2j+1}\rangle$. Denote $G' = (G\backslash v_1) \cup u_1$ Since $G \cup u_1$ contains cycles Q_1 and Q_2 and $I(v_1) = \{i_1, i_1\}$, G' is feasible. We will find a sequence of cycles $Q'_1, \ldots, Q'_{2k'} (k' < k)$ such that there exists an even integer $J^*, 2 \leqslant j^* \leqslant 2k-2$, satisfying the following conditions.

(a) For any $l = 1, \ldots, 2k'$, Q'_l contains $v_{\lfloor (j^*+l)/2 \rfloor}$. For any $l = 2, \ldots, 2k'-1$, Q'_l contains $u_{\lfloor (j^*+l)/2 \rfloor}$ and $Q'_l \backslash u_{\lfloor (j^*+l)/2 \rfloor}$ is in G'. (Hence, $2k' \leqslant 2k - j^* \leqslant 2k-2$.)

(b) Q'_1 and $Q'_{2k'}$ are in G' and $Q'_l \cup u_{\lfloor (j^*+l)/2 \rfloor}$ is in G' for $l = 2, \ldots, 2k'-1$.

Once the sequence of cycles is found, we construct the graph $H(G')$ containing $\langle Q'_1, i_{j*+1} \rangle, \ldots, \langle Q'_{2k'}, i_{j^*+2k'} \rangle$ as vertices. Then those vertices form an augmenting path in $H(G)$. By the induction hypothesis, the graph G' is not a minimum feasible graph and neither is G since $\|G\| = \|G'\|$. Next, we describe how to compute the sequence of cycles $Q'_1, \ldots, Q'_{2k'}$.

begin
  $j^* := 2; j := 3; l := 1;$
  found:=false;
  while (*found* = *false*) **do begin**
    Case 1: Q_j does not contain v_1 {Removal v_1 does not destroy Q_j. We keep Q_j.}
    begin
      $Q'_l := Q_j; l := l + 1; j := j + 1;$
    end;
    Case 2: Q_j contains v_1 {Q_j is destroyed by removing v_1. We need to find a new cycle replacing Q_j. Note that $i_j = i_1$ or i_2. Thus, either $Q_j \oplus Q_1$ or $Q_j \oplus Q_2$ contains a cycle which contains $v_{\lceil j/2 \rceil}$. This cycle is put into the sequence and denoted by Q^*_j.}
    Subcase 2.1: $j < 2k$ and Q^*_j contains $u_{\lfloor j/2 \rfloor}$ {Q^*_j is not contained in G' So, Q^*_j cannot end the sequence.}
    begin
      $Q'_l := Q^*_j; l : l + 1; j := j + 1;$
    end
    Subcase 2.2: $j = 2k$ or Q^*_j does not contain $u_{\lfloor j/2 \rfloor}$ when $j < 2k$ {In either situation, Q^*_j is contained in G'. When j is even, Q^*_j will end the required sequence. When j is odd, it means taht the previous search fails. However, in this situation, we must have $j < 2k$ so that Q^*_j contains $v_{(j+1)/2}$. Thus, we can start over with Q^*_j again.}
    begin
      if j is even **then** set $Q_l := Q^*_j$ and found := truth;
      if j is odd **then** set $j^*; +j - 1, l := 1, Q'_1 := Q^*_j$ and $j := j + 1;$
    end
  end {while loop};
end

Clearly, the above algorithm terminates when the required sequence is found. This completes the proof of the lemma. □

From the proof of lemma 3.3, we can also conclude that when an augmenting path $\{\langle Q_1, i_1 \rangle, \ldots, \langle Q_{2k}, i_{2k} \rangle\}$ exists, we can improve the current feasible graph G by deleting all edges v_j's and adding all edges u_j's where v_j's are in the definition of the augmenting path and $m(u_j)$'s are edges in the augmenting path. To see this, let us consider three types of paths in $H(G)$:

1. the augmenting path $\langle Q_1, i_1\rangle, \ldots, \langle Q_{2k}, i_{2k}\rangle\}$,
2. the alternating path $\{\langle Q_1, i_1\rangle, \ldots, \langle Q_{2k-1}, i_{2k-1}\rangle\}$ satisfying condition (A3) and that Q_1 is in G,
3. the alternating path $\{\langle Q_2, i_2\rangle, \ldots, \langle Q_{2k-1}, i_{2k-1}\rangle\}$.

For any type of path, we have u_j's and v_j's related to the path like before.

LEMMA 3.4. *For any type of path, we can delete all v_j's and add all u_j's in the path without changing feasibility.*

Proof. This lemma is proved by induction on the length of the path. For the path length equal to one or two, we can verify it easily. For the path length larger than two, let us first assume that the path is an augmenting path. We look at the proof of Lemma 3.3 again. Note that $G' = (G \backslash v_1) \cup u_1$ and $H(G')$ will be constructed by using all cycles obtained in the search. Now, in Subcase 2.2, if j is even, than an augmenting path is found; if j is odd, then an alternating path of even length is found and this path is of type 2. Moreover, if j is odd, then an alternating path of even length is found and this path is of type 2. Moreover, if j is even and $j < 2k$, then either $Q_j \oplus Q_1$ or $Q_j \oplus Q_2$ contains a cycle Q_j^{**} which contains $u_{j/2}$, then we can start a new search with Q_j^{**}. This search will end up on a path of type 2 or type 3. In this way, we can obtain a collection of disjoint paths of the three types such that all v_j for $j = 2, \ldots, k$ and all u_j for $j = 2, \ldots, k-1$ will appear in paths of the collection and play a similar role as v_j's and u_j's in the original augmenting path. By the induction hypothesis, we can replace all v_j's by all u_j's preserving the feasibility. Similarly, we can deal with the path of type 2 or type 3. □

Now, we face the problem of how to find an augmenting path in $H(G)$. Since an alternating path respect to $M(G)$ in $H(G)$ may be non-augmenting, $H(G)$ can also contain alternating path for an optimal G. Should we enumerate all alternating paths satisfying condition (A2) and check if one of them satisfies condition (A3)? It may be a way. However, we would like to provide a better way.

Besides $H(G)$, we construct another graph $B(G) = (V_1 \cup V_2, E_1 \cup E_2 \cup E_3)$ where

$$\begin{aligned}
V_1 &= V(H(G)), \\
V_2 &= \{\langle v, i\rangle, \langle v, i'\rangle | I(v) = \{i, i'\} \text{ and } \exists \langle Q, i\rangle, \langle Q', i'\rangle \\
&\in V_1, v \in \exists E(Q \cap Q' \cap G)\}, \\
E_1 &= M(G), \\
E_2 &= \{(\langle v, i\rangle, \langle v, i'\rangle) | \langle v, i\rangle, \langle v, i'\rangle \in V_2 \text{ and } I(v) = \{i, i'\}\}, \\
E_3 &= \{(\langle Q, i\rangle, \langle v, i\rangle) | \langle Q, i\rangle \in V_1, \langle v, i\rangle \in V_2 \text{ and } Q \text{ contains } v\}.
\end{aligned}$$

Clearly, $E_1 \cup E_2$ is a matching in $B(G)$ and a more important fact is stated in the next lemma.

LEMMA 3.5. *$H(G)$ has an augmenting path with respect to $M(G)$ if and only if $B(G)$ has an augmenting path with respect to $E_1 \cup E_2$.*

Proof. Suppose that $\{\langle Q_1, i_1\rangle, \dots, \langle Q_{2k}, i_{2k}\rangle\}$ is an augmenting path in $H(G)$. By condition (2), Q_1 and Q_{2k} are in G. Thus, $\langle Q_1, i_1\rangle$ and $\langle Q_{2k}, i_{2k}\rangle$ are not covered by $M(G)$ and hence not covered by $E_1 \cup E_2$, i.e., they are non-saturated. Let $v_1, \dots, v_k$ be edges in condition (3) of the definition of the augmenting path. Then it is easy to verify that $\{\langle Q_1, i_1\rangle, \langle v_1, i_1\rangle, \langle v_1, i_2\rangle, \langle Q_2, i_2\rangle, \langle Q_3, i_3\rangle, \langle v_2, i_3\rangle, \dots, \langle Q_{2k}, i_{2k}\rangle\}$ is an augmenting path with respect to $E_1 \cup E_2$.

Conversely, consider an augmenting path P with respect to $B(G)$. Since every vertex in V_2 is saturated, the two endpoints of P belong to V_1 and can be denotes as $\langle Q, i\rangle$ and $\langle Q', i'\rangle$. Note that $\langle Q, i\rangle$ is saturated in $H(G)$ with respect to $M(G)$ if and only if it is saturated in $B(G)$ with respect to $E_1 \cup E_2$. Thus, $\langle Q, i\rangle$ is not saturated in $H(G)$ with respect to $M(G)$. It follows that Q is in G. Similarly, Q' is in G. Next, noting the form of edges in E_3, it is easy to see that the jth edge in path P is in E_3 if j is odd, is in E_2 if $j \equiv 2 \pmod 4$, and is in E_1 if $j \equiv 0 \pmod 4$. It follows that P has an even number of edges in E_3. Let $(\langle Q_1, i_1\rangle, \langle v_1, i_1\rangle), (\langle v_1, i_2\rangle, \langle Q_2, i_2\rangle), (\langle Q_3, i_3\rangle, \langle v_2, i_3\rangle, \dots, \langle v_k, i_{2k}\rangle, \langle Q_{2k}, i_{2k}\rangle)$ be the $2k$ edges of P in E_3 where $Q_1 = Q$ and $Q_{2k} = Q'$. Since $|I(v_j)| = 2$ for all j and an alternating path always simple, $v_1, \dots, v_k$ are distinct. Therefore, $\{\langle Q_1, i_1\rangle, \dots, \langle Q_{2k}, i_{2k}\rangle\}$ is an augmenting path in $H(G)$. □

THEOREM 3.6. *When $\alpha = \beta = 2$, the minimum feasible graph for $(X_1, \dots, X_n)$ can be computed in $O(|X|^6)$ time.*

Proof. Note that the minimum feasible graph can be computed in the following way: Initially, using Lemma 2.1 reduces the problem to one satisfying the condition that for every $i = 1, \dots, m, k(i) = 1$ and then set G : 'K^*. At each iteration, construct graph $B(G)$ and look for an alternating path with respect to $E_1 \cup E_2$ with two non-saturated endpoints. If such a path does exist, then G is minimum. If such a path exists, then we improve the feasible graph by deleting all v_j's and adding all u_j's, where $(\langle v_j, i_{2j-1}\rangle, \langle v_j, i_{2j}\rangle)$ and $m(u_j)$ are edges in the alternating path.

To estimate the computing time, we first note that for any graph R, the number of independent cycles equals $1 + |E(R)| - |V(R)|$. Thus, K_i^* has at most $\|K_i^*\|$ independent cycles. Since $\alpha = \beta = 2$, each edge of K^* appears in at most two k_i^*'s. Hence, $|V_1| \leqslant \sum_{i=1}^m \|K_i^*\| \leqslant 2\|K^*\|$. Therefore, $(|V(B(G))| = |V_1| + |V_2| = O(|X|^2) + O(|X|^2) = O(|X|^2)$. Since each (simple) cycle in K^* contains at most $|X|$ edges, $|E(B(G))| = O(|X|^3)$. By an algorithm given by Micali and Vazirani in [10], we can compute the alternating path in $B(G)$ in $O(|X|^4)$ time. Moreover, it is easy to construct $B(G)$ in $O(|X|^4)$ time. Therefore, at each iteration, the computation takes $O(|X|^4)$ time. Since there are at most $\|K^*\| = O(|X|^2)$ iterations, total running time after the initial step is $O(|X|^6)$. Now, we examine the initial step. Since $\alpha = \beta = 2, |I(x)| < 2|X|$ for any $x \in X$. Thus, $m \leqslant 2|X|^2$. This implies that the initial step can also be computed within $O(|X|^5)$ time. □

We end this section by presenting an example. Let $X = \{a,b,c,d,e\}, X_1 = \{a,b,c,d\}, X_2 = \{b,c,e\}, X_3 = \{c,d,e\}, X_4 = \{a,c\}, X_5 = \{a,d\}, X_6 = \{a,b\}, X_7 = \{a,e\}$, and $X_8 = \{b,e\}$. It is easy to verify that $K(1) = \cdots = K(8) = 1$. The graph K^* is as shown in Figure 1.

At the first iteration, we have $G = K^*$ and $B(G)$ as shown in Figure 1. There exist three augmenting paths. We consider path $\{\langle bce, 2\rangle, \langle (c,e), 2\rangle, \langle (c,e), 3\rangle, \langle cde, 3\rangle\}$ and obtain G' by deleting (c,e). The graph $B(G')$ is constructed as shown in Figure 1. There exists a unique alternating path in $B(G')$. Deleting edges (b,c) and (c,d) and adding edge (c,e), we obtain G''. Finally, constructing $B(G'')$, we find that $B(G'') = \emptyset$ and hence G'' is a minimum feasible graph.

4. ($\alpha \geqslant 3, \beta$=1), ($\alpha$=1, $\beta \geqslant 6$), and ($\alpha \geqslant 2, \beta \geqslant 3$)

We show that the SID is NP-hard for $(\alpha \geqslant 3, \beta = 1), (\alpha = 1, \beta \geqslant 6)$, and $(\alpha \geqslant 2, \beta \geqslant 3)$. To do so, we consider a special case of the set packing problem as follows:

Set Packing

Instance: Collection Γ of finite sets, a positive integer $K \leqslant |\Gamma|$.

Question: Does Γ contain at least K mutually disjoint sets?

In the special case we consider, the instance is required to satisfy the following conditions:

(C1) For any $A \in \Gamma, |A| \leqslant 3$.

(C2) For any different $A, B \in \Gamma, |A \cap B| \leqslant 1$.

(C3) For any distinct $A, B, C \in \Gamma, |A \cap B \cap C| = 0$.

This special case is denoted by 3-1-SP in [6], which was first proved to be NP-complete in [8] with name vertex packing on cubic graphs. By (C3), each element is in at most two sets in Γ. Since an element in only one set can be deleted without loss of generality, we can furthermore assume the following.

(C3') Every element is in exactly two sets in Γ.

The decision version of the SID is as follows.

The Decision Version of the SID

Instance: m sets $X_1, \ldots, X_m$ and a positive integer K'.

Question: Is there a feasible graph G for $(X_1, \ldots, X_m)$ such that $||G|| \leqslant K'$?

THEOREM 4.1. *The decision version of the SID is NP-complete for* $(\alpha \geqslant 3, \beta = 1), (\alpha = 1, \beta \geqslant 6)$, *and* $(\alpha \geqslant 2, \beta \geqslant 3)$.

Proof. Clearly, the decision version of the SID belongs to NP. We reduce 3-1-SP to the decision version of SID. Consider an instance of 3-1-SP consists of a collection Γ of sets and a positive number K. Denote $M = \cup_{A \in \Gamma} A$. Without loss

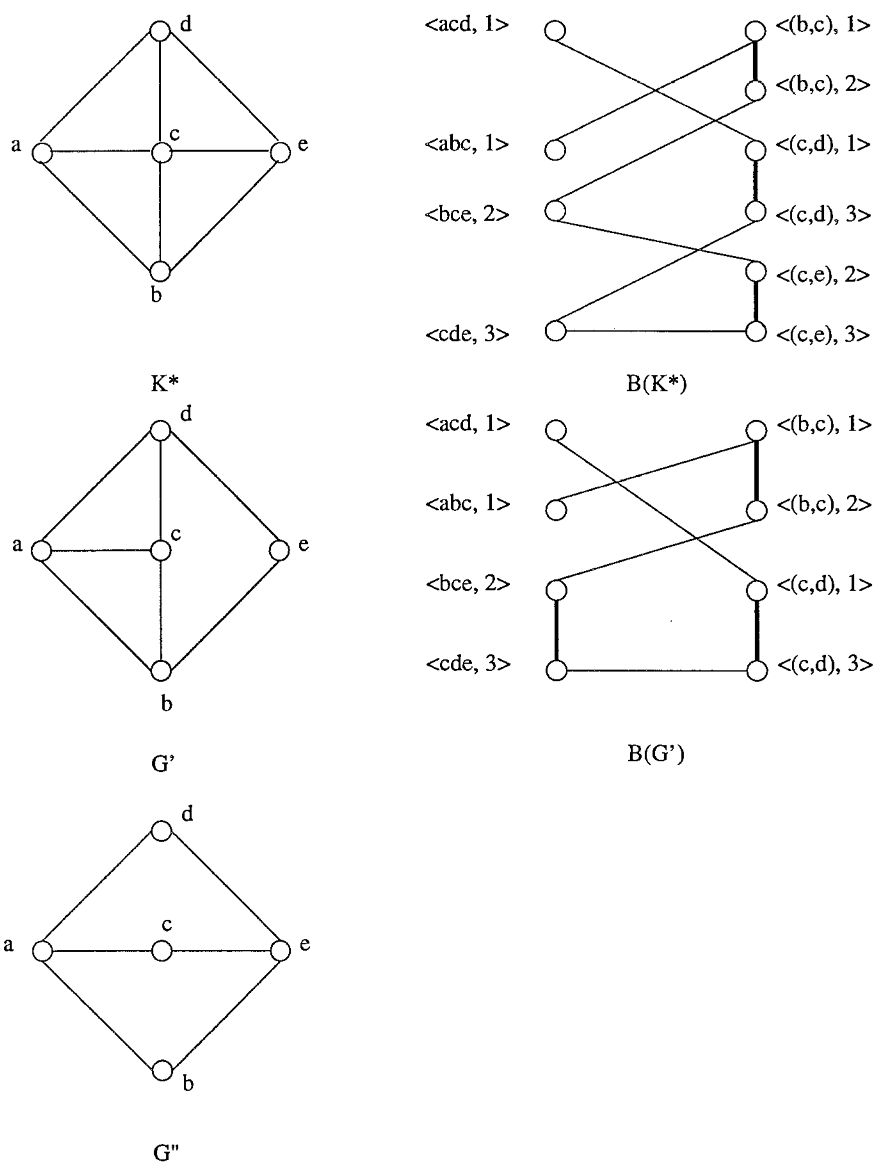

Fig. 1. An example.

of generality, assume that all sets in Γ consists of natural numbers in $\{1, \ldots, |M|\}$. We first define an instance of the decision version of SID as follows.

(1) For each $A \in \Gamma$, we introduce two points x_A and y_A and for each $i \in M$, we introduce a point z_i. Then for each $i \in M$, define $X_i = \{x_A, y_A | i \in A\} \cup \{z_i\}$. Note that each $i \in M$ is in exactly two sets $A(1)$ and $A(2)$ in Γ. Now, for each $i \in M$, we also define $X_{|M|+3i} = \{z_i, x_{A(1)}\}, X_{|M|+3i+1} = \{y_{A(1)}, y_{A(2)}\}$, and $X_{|M|+3i+2} = \{z_i, x_{A(2)}\}$.

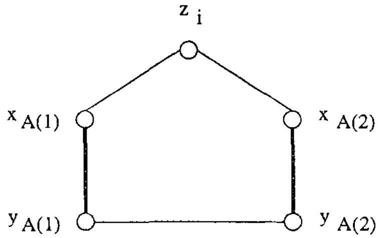

Fig. 2. H_i is a cycle.

(2) Set $m = 4|M|$ and $K' = |\Gamma| + 3|M| - K$.

Now, we claim that the collection Γ has at least K mutually disjoint sets if and only if there exists a feasible graph G for $(X_1, \ldots, X_m)$ such that $||G|| \geqslant K'$. To prove this claim, let us first construct a feasible graph H for $(X_1, \ldots, X_m)$ as follows:

$$\begin{aligned} V(H) &= \{x_A, y_A | A \in \Gamma\} \cup \{z_i | i \in M\} \text{ and} \\ E(H) &= \{(x_A, y_A) | A \in \Gamma\} \cup \{(z_i, x_{A(1)}), (y_{A(1)}, y_{A(2)}), (x_{A(2)}, z_i) | i \in M\}. \end{aligned}$$

Clearly, H has $|\Gamma| + 3|M|$ edges and for each $A \in \Gamma, I((x_A, y_A)) = A$.

First, suppose that Γ has K mutually disjoint sets $B(1), \ldots, B(K)$. This means that the index sets of $u_{B(1)}, \ldots, u_{B(K)}$ are mutually disjoint. Moreover, for each $i \in B(j)$, the subgraph H_i induced by X_i is a cycle having edge $u_{B(j)}$. (See Figure 2.) Thus, all edges $u_{B(1)}, \ldots, u_{B(K)}$ can be deleted preserving the feasibility. This gives a feasible graph G with $||G|| \leqslant K'$.

Conversely, suppose that there exists a feasible graph G for $(X_1, \ldots, X_m)$ such that $||G|| \leqslant K'$. We first want to show that G can be assumed to be a subgraph of H. To see this, we note that the following two facts hold:

(b) By (C2), H contains all edges with at least two indices.

(a) For every $i = 1, \ldots, m, K(i) = 1$, i.e., H_i is connected. In fact, H_i is a cycle for $i \in M$ and H_i is an edge for $|M| < i \leqslant m = 4|M|$.

By Lemma 2.2, H contains a minimum feasible graph which can be chosen as the feasible graph G with $||G|| \leqslant K'$. Next, note that G must contain all edges in $\{(z_i, x_{A(1)}), (y_{A(1)}, y_{A(2)}), (x_{A(2)}, z_i) | i \in M\}$. Suppose that all edges in $H \backslash G$ are $(x_{B(1)}, y_{B(1)}), \ldots, (x_{B(k'')}, y_{B(k'')})$ where $K'' = ||H|| - ||G|| \geqslant K$. We show that K'' sets $B(1), \ldots, B(K'')$ are mutually disjoint. To do this, note that each H_i for $i \in M$ is a cycle as shown in Figure 2. Thus, $H_i \backslash G$ contains at most one edge because G_i is connected. This means that for every $i \in M$, there exists at most one $B(j)$ containing i. It follows that $B(j)$ for $j = 1, \ldots, K''$ are mutually disjoint.

Finally, we remark that the instance of the decision version of SID, constructed as above, satisfies $(\alpha = 3, \beta = 1)$, $(\alpha = 1, \beta = 6)$, and $(\alpha = 2, \beta = 3)$. In fact, we have the following.

(1) Every X_i has at most three elements. Thus, $(\alpha = 3, \beta = 1)$ is satisfied.

(2) By (C1), $|A| \leqslant 3$ for every $A \in \Gamma$. Thus every x_A (or y_A) is in at most three X_i's for $i \in M$ and at most three X_i's for $i \notin M$. Moreover, every z_i appears in at most three X_i's Hence, $(\alpha = 1, \beta = 6)$ is satisfied.

(3) By (C1), every edge (x_A, y_A) has at most three indices. Moreover, every edge in $\{(z_i, x_{A(1)}).(y_{A(1)}, y_{A(2)}), (x_{A(2)}, z_i) | i \in M\}$ has exactly two indices and every edge not in H has at most one index. Thus, $(\alpha = 2, \beta = 3)$ is satisfied. □

5. Discussion

We leave an open question on the computational complexity of the SID for $(\alpha = 1, 3 \leqslant \beta \leqslant 5)$. Tang [11] showed that for $m = 3$ the condition $(*)$ is necessary and sufficient for a feasible graph to be minimum. This implies that $m = 3$ the minimum feasible graph is polynomial-time computable. From this evidence, we believe that the SID is polynomial-time solvable for $(\alpha = 1, \beta = 3)$. However, we also believe that the SID is NP-hard for $(\alpha = 1, \beta = 4)$. An application of result for $(\alpha = 2, \beta = 2)$ is to construct approximation solution for the general SID. There are several ways. The first one is to divide the collection of subsets $X_1, \ldots, X_m$ into several small collections satisfying $(\alpha = 2, \beta = 2)$, construct a minimum feasible graph for each small collection and take the union of all of them. The second one is as follows: When a pair of points appear in more than two subsets X_i's, we stick them together. In this way, we can reduce original collection of subsets to a new one satisfying condition $(\alpha = 2, \beta = 2)$. After a minimum feasible graph for the new collection of subsets is found, we break stuck pairs by adding some edges.

It is still an open problem whether the SID has a bounded polynomial-time heuristic or not.

Acknowledgements

The authors wish to thank referees for their insightful comments on an earlier draft of this paper and to thank Mr. Zhicheng Liu and Dr. Bo Chen for many helpful discussions.

References

1. S.-C. Chao and D.-Z. Du (1983), A Sufficient Optimality Condition for the Valve-Placement Problem, *Journal of North-East Heavy Machinery Institute*, 4 (in Chinese).
2. D.-Z. Du and Y.-M. Chen (1976), Placement of Valves in Vacuum Systems, *J. Electric Light Sources* 4 (in Chinese).
3. D.-Z. Du (1986), An Optimization Problem, *Discrete Appl. Math.* **14**, 101–104.
4. D.-Z. Du and Z. Miller (1988), Matroids and Subset Interconnection Design, *SIAM J. Disc. Math.* **1**, 416–424.
5. D.-Z. Du and X.-F. Du (1991), A Special Case of Valve-Placement Problem, *Acta Mathematics Applicatae Sinica* **4** (in Chinese).
6. D.-Z. Du and Ker-I Ko (1989), On the Complexity of an Optimal Routing Tree Problem, *Acta Mathematics Applicatae Sinica (English Series)* **5**, 68–80.

7. M. R. Garey and D. S. Johnson (1979), *Computers and Intractability, A Guide to the Theory of NP-Completeness,* Freeman, San Francisco.
8. M. R. Garey, D. S. Johnson and L. Stockmeyer (1976), Some Simplified NP-Complete Graph Problems, *Theoretical Computer Science* **1** 237–267.
9. M. Gondran and M. Ninoux (1984), *Graphs and Algorithms*, John Wiley, New York.
10. S. Micali and V. Vazirani (1980), An $O(\sqrt{|V|}|E|)$ Algorithm for Finding Maximum Matching in General Graphs, *21st Annual Symposium on Foundation of Computer Science,* Syracuse NY.
11. T.-Z. Tang (1989), An Optimality Condition for Minimum Feasible Graphs, *Applied Mathematics* **2**, 21–24 (in Chinese).