Journal of the Korean Institute of Industrial Engineers
Vol. 29, No. 4, pp. 253-258, December 2003.

최대 로트 그룹핑 문제의 복잡성

황 학 진†

조선대학교 산업공학과

On the Hardness of the Maximum Lot Grouping Problem

Hark-Chin Hwang

Department of Industrial Engineering, Chosun University, Gwangju, 501-709

We consider the problem of grouping orders into lots. The problem is modelled by a graph $G=(V,E)$, where each node $v\in V$ denotes order specification and its weight $w(v)$ the orders on hand for the specification. We can construct a lot simply from orders of single specification. For a set of nodes (specifications) $\theta\subseteq V$, if the distance of any two nodes in θ is at most d, it is also possible to make a lot using orders on the nodes. The objective is to maximize the number of lots with size exactly λ. In this paper, we prove that our problem is NP-Complete when $d=2, \lambda=3$ and each weight is 0 or 1. Moreover, it is also shown to be NP-Complete when $d=1, \lambda=3$ and each weight is 1, 2 or 3.

Keywords**:** lot grouping, NP-Complete, 3-dimensional matching

1. Introduction

In modern manufacturing, it is the usual case that orders in large variations of specifications come with small quantities. In order to handle this situation, we often make production facilities flexible to process orders with different specifications at the same time. Some orders with similar specifications can be grouped to a lot for production.

When making from orders of charges to slabs in steel industry, it is required not to treat charges of quite different specifications at the same time. The characteristics of charges are often described in terms of their width and thickness. In order to minimize the different kinds of specifications, the two dimensional space of width and thickness are divided into set of disjoint classes to one of which a charge may belong. Each class is defined by two interval regions of width and thickness. For instance, a class ([910, 940), [600, 650)) is for the charges with widths in the range of [910, 940) and thickness in the range of [600, 650). Let $W_1, W_2, \cdots, W_k$, ($T_1, T_2, \cdots, T_l$) be the partition of $(0,+\infty)$ in the dimension of width (thickness), respectively. Then, we can designate each class or each specification of charges by (W_i, T_j). In the following Figure 1(a), the order quantities for charges are represented in tabular form. As we can see in the figure, similar classes position in their neighborhood to each other whereas somewhat different classes in dimensions are dispersed at a distance. The similarity between a class (W_{i_1}, T_{j_1}) and another class (W_{i_2}, T_{j_2}) is measured by rectlinear distance, i.e., $|i_2-i_1|+|j_2-j_1|$. In the construction of lot with size of five charges, due to the limitation of facilities, charges from quite different classes cannot be grouped into a lot but charges of classes with some similarities, for instance, with the rectlinear distance between them at most two, can be grouped.

We generalize the situation of slab making in steel industry and formally model our problem by a graph. A graph is a pair $G=(V,E)$, where V is a finite set of nodes and E has as elements sets of two *nodes* in V

† Corresponding author : Professor Hark-Chin Hwang, Department of Industrial Engineering, Chosun University, Gwangju, 501-759, Korea, Fax : +82-62-230-7128, E-mail : hchwang@chosun.ac.kr

Received March 2003; revision received October 2003; accepted October 2003.

called *edges*. Each node v corresponds to an order specification and its weight $w(v)$ denotes the order quantity on hand for the specification. Then, orders of charges in Figure 1 can be described by the graph (grid) in Figure 1(b), where each node v_{ij} represents class or specification (W_i, T_j) and its weight is the number in the circle. We suppose that nodes are arranged in conservation of similarities as in Figure 1(b) so that the similarity between two nodes are measured by the distance between them. The *distance* of two nodes u and v is one if $\{u, v\} \in E$. In general, the distance between u and v is the distance of the shortest path from u to v.

Then, orders with specifications u and v can be grouped if the distance is no greater than the allowable *distance limit* of d. A *lot type* is any subset θ of V. If the cardinality of θ is one, it is called *homogeneous* otherwise called heterogeneous. A *heterogeneous* lot type θ is *feasible* if the distance of any two nodes in θ is at most d. Then, given a *lot size* λ (a positive integer), a lot with respect to the type θ is the order quantities corresponding to the nodes (specifications) in θ. A lot is *feasible* if its type θ is feasible and the total sum of the order quantities from the nodes in θ is exactly λ. The objective function is to maximize the number of lots, whether homogeneous or heterogeneous, with size λ. We call our problem *maximum lot grouping problem* or maximum grouping problem in short. Notice that if $\lambda=2$, $d=1$, and every $w(v)=1$, $v\in V$, our problem is the same as maximum matching problem, for which optimal algorithms are provided (Micali & Vazirani, 1980; Papadimitriou & Steiglitz, 1988).

In this paper, we consider the hardness of the maximum grouping problem. The most key factor determining the problem's intractability would be distance limit more than anything else of other parameters of lot size and weight. Firstly, we will show that the problem with $d=2$ is NP-Complete even when the graph G is bipartite, $\lambda=3$ and every $w(v)=0$ or $w(v)=1$ for $v\in V$. Next, we deal with the case of $d=1$. The problem with $d=1$ is proved to be NP-Complete in general. And it is still hard even though the graph G is bipartite, $\lambda=3$, and every $w(v)=1, 2$ or 3 for $v\in V$. In the next section, a mathematical formulation is given to describe the problem explicitly and in section 3, the hardness of the maximum grouping problem is proved. Finally, conclusion follows in section 4.

2. Problem Definition

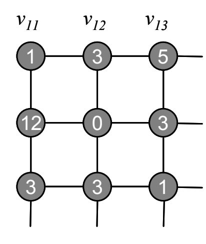
In order to provide clear definitions of parameters and variables for our problem, we present a list of notations in the following:

- V : the set of nodes which represent order specifications.
- E : the set of edges which represent similarity structures between specifications.
- $w(v)$: the weight or the order quantity on the specification $v\in V$. $w(v)$ can take any positive integer.
- λ : the size of lot.
- d : the maximum allowable distance limit.
- θ : a feasible lot type, i.e., a set of nodes whose distance between them is at most d, $\theta\subseteq V$.
- Θ : the set of all the feasible lot types. Note that the distance limit constraint of d is represented by Θ.
- $x_\theta(v)$: for a feasible lot type $\theta\subseteq V$, its lot is denoted by x_θ where $x_\theta(v)$ is the quantity from the node $v\in\theta$. Note that $\sum_{v\in\theta} x_\theta(v)=\lambda$.
- y_θ : the binary variable to identify whether a lot has been constructed or not for a feasible lot type $\theta\subseteq V$.
- y_v : the number of homogeneous lots for v.

We further understand the meaning of each parameter by the slab making illustration in the previous section for the case that $V=\{v_{ij} | i=1,2,3, j=1,2,3\}$ and edges are defined as shown in Figure 1(b). For $d=2$ and $\lambda=5$, the set $\theta=\{v_{11}, v_{12}, v_{21}, v_{22}\}$ is a feasible lot

T \ W	W1	W2	W3	...
T1	1	3	2	
T2	12	0	3	
T3	3	3	1	
...				

(a)



(b)

Figure. 1. (a) Orders in tabula form and (b) the corresponding graph.

type since all the distances between them are at most 2. Suppose that we make a lot of type θ. Then, a possible contents of it could be $x_\theta(v_{11})=1, x_\theta(v_{12})=2, x_\theta(v_{21})=2$ and $x_\theta(v_{22})=0$.

The following lemma presents a useful result which allows us to make as many homogeneous lots as possible from heterogeneous lots.

Lemma 1

Given two heterogeneous feasible lots, $\overline{x_\theta}$ and $\widehat{x_\theta}$, with respect to the same type θ, $|\theta|\geq 2$, we can make two feasible contracted lots $x_{\theta'}$ and $x_{\theta''}$, $\theta', \theta''\subseteq\theta$, where $\theta'\subset\theta$ or $\theta''\subset\theta$.

Proof.

Let $\theta=v_1,\cdots,v_k$, $k\geq 2$. Let v_i be the first node such that $\sum_{j=1}^{i-1}(\overline{x_\theta}(v_j)+\widehat{x_\theta}(v_j))<\lambda$ and $\sum_{j=1}^{i}(\overline{x_\theta}(v_j)+\widehat{x_\theta}(v_j))\geq\lambda$.

Then, we make a lot $x_{\theta'}$ where $\theta'=\{v_1,\cdots,v_i\}$ and

$$x_\theta(v_j)=\overline{x_\theta}(v_j)+\widehat{x_\theta}(v_j) \text{ for } j=1,\cdots,i-1,$$

$$x_{\theta'}(v_i)=\lambda-\sum_{j=1}^{i-1}(\overline{x_\theta}(v_j)+\widehat{x_\theta}(v_j)).$$

Next, we will construct another lot $x_{\theta''}$. First, consider the case of $\overline{x_\theta}(v_i)+\widehat{x_\theta}(v_i)-x_{\theta'}(v_i)>0$. In this case, we let $\theta''=\{v_i,\cdots,v_k\}$ and construct $x_{\theta''}$ with $x_{\theta''}(v_i)=\overline{x_\theta}(v_i)+\widehat{x_\theta}(v_i)-x_{\theta'}(v_i)$ and $x_{\theta''}(v_j)=\overline{x_\theta}(v_j)+\widehat{x_\theta}(v_j)$ for $j=i+,\cdots,k$. Next, consider the case of $\overline{x_\theta}(v_i)+\widehat{x_\theta}(v_i)-x_{\theta'}(v_i)=0$. Similarly, we let $\theta''=\{v_{i+1},\cdots,v_k\}$ and $x_{\theta''}(v_j)=\overline{x_\theta}(v_j)+\widehat{x_\theta}(v_j)$ for $j=i+1,\cdots,k$. Then, we can see that $\theta'\subset\theta$ or $\theta''\subset\theta$ with $\sum_{v\in\theta'}x_{\theta'}(v)=\sum_{v\in\theta''}x_{\theta''}(v)=\lambda$. Hence, $x_{\theta'}$ and $x_{\theta''}$, are feasible contracted lots with respect to type θ.

In most case of the real manufacturing, it is suggested that we make homogeneous lots as many as possible rather than heterogeneous ones. Applying Lemma 1 continuously to all heterogeneous lots, we can finally get the desired solution that no more than one lot is constructed for each heterogeneous type (though several lots are possibly constructed for each homogeneous type). In addition, from the lemma, we can assure that there always exists an optimal solution such that no more than one heterogeneous lot can be constructed for each heterogeneous type. Hence, it is enough to use a binary integer variable for each heterogeneous type to describe the amount of lot constructions. Let θ be a set of feasible heterogeneous types. Then, for a type $\theta\in\Theta$, we use the binary variable y_θ to identify whether a lot has been constructed or not. Note that for a heterogeneous type θ, we have $y_\theta=1$ iff $\sum_{v\in\theta}x_\theta(v)=\lambda$. The number of homogeneous lots for $\theta=\{v\}$ is just denoted by y_v. Then, keeping in mind that the distance limit is assured by the set Θ, we can formulate our problem as the following integer programming problem :

Maximize $\sum_{v\in V}y_v+\sum_{\theta\in\Theta}y_\theta$

Subject to

$\lambda y_\theta-\sum_{v\in\theta}x_\theta(v)=0$	$\theta\in\Theta$
$\lambda y_v-\sum_{\theta:v\in\theta}x_\theta(v)\leq w(v)$	$v\in V$
$x_\theta(v)\geq 0,\ x_\theta(v)$: integer for	$v\in\theta,\ \theta\in\Theta$
$y_v\geq 0,\ y_v$: integer for	$v\in V$
$y_\theta=0$ or 1 for	$\theta\in\Theta$

3. The Hardness of Maximum Lot Grouping Problem

We will show that the maximum grouping problem is NP-Complete, which further means that polynomial optimal algorithms cannot exist unless P=NP. Even in the restricted cases that a lot size, a distance limit and weights have small values or the graph is bipartite, the problem will be proven to be NP-Complete. To this purpose, we investigate the computational complexity of decision versions of the maximum grouping problem.

Theorem 2

For the maximum lot grouping problem $G=(V,E)$ with $d=2, \lambda=3$ and each weight 0 or 1, the question of deciding if there exists number of $\left\lfloor\sum_{v\in V}w(v)/\lambda\right\rfloor$ feasible lots is NP-Complete.

We prove this result by showing that the known NP-Complete problem 3-dimensional matching can be transformed to the maximum grouping problem.

3-Dimensional Matching (3DM)

Instance : Disjoint sets $A=\{a_1,\cdots,a_n\}$, $B=\{b_1,\cdots,b_n\}$, $C=\{c_1,\cdots,c_n\}$ and a family

$F=\{T_1,\cdots,T_m\}$ of triples with $|T_i\cap A|=|T_i\cap B|=|T_i\cap C|=1$ for $i=1,\cdots,m$.

Question : Does F contain a matching, that is, a subfamily F' for which $|F'|=n$ and $\cup_{T_i\in F'}T_i=A\cup B\cup C$?

For the sets A, B, C and F, we define corresponding node sets V_A, V_B, V_C and V_F as follows :

$$V_A=\{v_{a_1},\cdots,v_{a_n}\},\quad V_B=\{v_{b_1},\cdots,v_{b_n}\},$$
$$V_C=\{v_{c_1},\cdots,v_{c_n}\},\quad V_F=\{v_{T_1},\cdots,v_{T_m}\}.$$

Proof of Theorem 2

Given an instance of the above 3DM problem, we construct an instance of the maximum lot grouping problem $G=(V,E)$ with lot size $\lambda=3$, distance limit $d=2$. The node set V is given as follows:

$$V=V_A\cup V_B\cup V_C\cup V_F$$

Edges exist only when there is a corresponding triple in F: edges are constructed between the nodes v_{T_i} and v_{a_j}, v_{b_k} or v_{c_l}, that is, three edges $\{v_{T_i}, v_{a_j}\}$, $\{v_{T_i}, v_{b_k}\}$ and $\{v_{T_i}, v_{c_l}\}$ are constructed if $T_i=\{a_j, b_k, c_l\}$ is in the family F. Now, we consider the weight of each node. We let $w(v_{a_j})=1$ ($w(v_{b_k})=1$, $w(v_{c_l})=1$) for each $a_j\in A$ ($b_k\in B$, $c_l\in C$), respectively. And for each $T_i\in F$, let $w(v_{T_i})=0$. In Figure 2, the graph corresponding to an instance of 3DM is illustrated (the weight of each node is the number in the circle). Note that $n=\left\lfloor \sum_{v\in V} w(v)/\lambda \right\rfloor$.

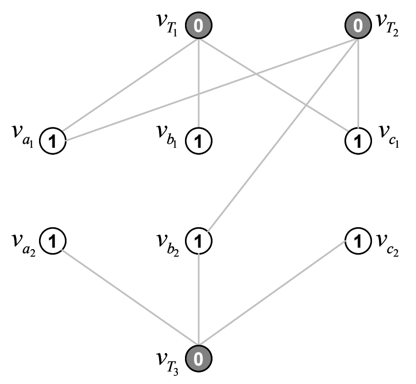
It is quite simple to show that there is number of n feasible lots if and only if there is a 3-dimensional matching. Suppose there is a matching F'. For each $T_i=\{a_j, b_k, c_l\}\in F'$, make a lot from the weights on the nodes v_{a_j}, v_{b_k} and v_{c_l}, i.e., a lot x_θ (with $y_\theta=1$) for the type $\theta=\{v_{T_i}, v_{a_j}, v_{b_k}, v_{c_l}\}$ where $x_\theta(v_{T_i})=0$, $x_\theta(v_{a_j})=x_\theta(v_{b_k})=x_\theta(v_{c_l})=1$. Note that the distance between any two nodes of θ is at most two and the size of x_θ is three. Thus, x_θ is a feasible lot. Since there are n triples in the matching, we can make the n corresponding lots.

Conversely, suppose that there are n lots in G. As the lot size λ is three and each node has weight at most one, in the n lots there are no homogeneous ones. Note that, in V, for any two nodes u, v with distance at most two, there must exist a triple $T_i=\{a_j, b_k, c_l\}\in F$ such that $u, v\in\{v_{T_i}, v_{a_j}, v_{b_k}, v_{c_l}\}$. Then, recalling the distance limit constraint $d=2$, we see that any feasible lot type θ must be a subset of nodes corresponding to a triple, say, T_i, that is, $\theta\subseteq\{v_{T_i}, v_{a_j}, v_{b_k}, v_{c_l}\}$. Let x_θ (with $y_\theta=1$) be a feasible lot where $\theta\subseteq\{v_{T_i}, v_{a_j}, v_{b_k}, v_{c_l}\}$. Since the lot size is three and the weight of v_{T_i} is zero, θ is $\{v_{a_j}, v_{b_k}, v_{c_l}\}$ or $\{v_{T_i}, v_{a_j}, v_{b_k}, v_{c_l}\}$. Thus, for the lot x_θ, we have $x_\theta(v_{a_j})=x_\theta(v_{b_k})=x_\theta(v_{c_l})=1$ or $x_\theta(v_{a_j})=x_\theta(v_{b_k})=x_\theta(v_{c_l})=1$, $x_\theta(v_{T_i})=0$. In either case, we see that for each lot x_θ, there exists exactly one corresponding triple T_i. Now, we choose n triples corresponding to the n lots. Note that each node $v_{a_j}(v_{b_k}, v_{c_l})$ has weight one. Thus, its weight or order quantity cannot be used in more than one lot, which means that the corresponding element $a_j(b_k, c_l)$ does not belong to more than one triple of the chosen n triples. Therefore, we conclude the set of n triples is a matching.

Consider again the graph $G=(V,E)$ corresponding to a 3DM in Theorem 2. Let $X=V_A\cup V_B\cup V_C$ and $Y=V_F$. Then, note that the node set V is partitioned into two disjoint sets X, Y and no edge exists between any two nodes in X and between any two nodes in Y, but edges exist between nodes in X and nodes in Y. Hence, G is a bipartite and thus the following result directly follows.

$$A=\{a_1,a_2\}\quad T_1=\{a_1,b_1,c_1\}$$
$$B=\{b_1,b_2\}\quad T_2=\{a_1,b_2,c_1\}$$
$$C=\{c_1,c_2\}\quad T_3=\{a_2,b_2,c_2\}$$
$$F=\{T_1,T_2,T_3\}$$

(a)



(b)

Figure 2. (a) A 3DM instance and (b) the corresponding graph.

Corollary 3
For the maximum lot grouping problem $G=(V,E)$ with $d=2, \lambda=3$ and each weight 0 or 1, the question of deciding if there exists number of $\left\lfloor \sum_{v\in V} w(v)/\lambda \right\rfloor$ feasible lots is NP-Complete even when G is bipartite.

Notice that when the distance limit d is one, one can find maximum number of lots using the maximum cardinality matching algorithm if $\lambda=2$, and every $w(v)=1$, $v\in V$(Micali & Vazirani, 1980; Papadimitriou & Steiglitz, 1988). However, as we shall see in the following theorem, the maximum grouping problem is still hard in general even though the distance limit is one. The proof of the following theorem is almost similar to that of Theorem 3.7 in (Garey & Johnson, 1979). We will transform our problem to 3DM as has been done in Theorem 2.

Theorem 4
For the maximum lot grouping problem $G=(V,E)$ with $d=1, \lambda=3$ and each weight 1, 2 or 3, the question of deciding if there exists number of $\left\lfloor \sum_{v\in V} w(v)/\lambda \right\rfloor$ feasible lots is NP-Complete.

Proof
For the 3-dimensional matching problem, we construct an instance of the maximum grouping problem $G=(V,E)$ with lot size $\lambda=3$. The nodes and edges in the graph G will be specified from the triples. For each triple, $T_i=\{a_j, b_k, c_l\}$, we construct a graph with edge set $E_i(|E_i|=6)$ as shown in Figure 3, where the weight of each node is the number in the circle.

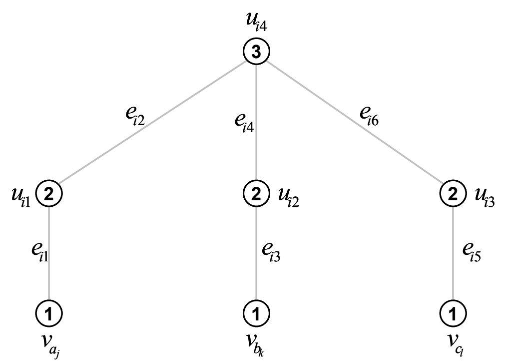


Figure 3. The graph for the triple $T_i=\{a_j, b_k, c_l\}$.

Then, the node set V and edge set E of are given as follows:

$$V=(V_A\cup V_B\cup V_C)\cup_{i=1}^{m}\{u_{ij}|1\le j\le 4\},$$

$$E=\cup_{i=1}^{m}E_i.$$

Note that the total sum of weights of V is

$$\sum_{v\in V} w(v)=|V_A\cup V_B\cup V_C|+9|F|=3(n+3m)$$

and $\left\lfloor \sum_{v\in V} w(v)/\lambda \right\rfloor = n+3m$.

We want to show that there is $n+3m$ feasible lots if and only if there is a 3-dimensional matching. Suppose there is a matching F' from F for A, B and C. From this matching we can find number $n+3m$ lots from G, as described in the following :
if $T_i=\{a_j, b_k, c_l\}$ is in the subfamily F', then the corresponding lots (three heterogeneous lots and one homogeneous lot) are given by

$$y_{e_{i1}}=1\ \ (x_{e_{i1}}(v_{a_j})=1,\ x_{e_{i1}}(u_{i1})=2),$$
$$y_{e_{i3}}=1\ \ (x_{e_{i3}}(v_{b_k})=1,\ x_{e_{i3}}(u_{i2})=2),$$
$$y_{e_{i5}}=1\ \ (x_{e_{i5}}(v_{c_l})=1,\ x_{e_{i5}}(u_{i3})=2),$$
$$y_{u_{i4}}=1$$

otherwise if T_i is not in the subfamily F', then the corresponding lots (three heterogeneous lots) are given by

$$y_{e_{i2}}=1\ \ (x_{e_{i2}}(u_{i1})=2,\ x_{e_{i2}}(u_{i4})=1),$$
$$y_{e_{i4}}=1\ \ (x_{e_{i4}}(u_{i2})=2,\ x_{e_{i4}}(u_{i4})=1),$$
$$y_{e_{i6}}=1\ \ (x_{e_{i6}}(u_{i3})=2,\ x_{e_{i6}}(u_{i4})=1).$$

Conversely, suppose that there is a solution with $n+3m$ lots, that is, $\sum_{v\in V} y_v+\sum_{\theta\in\Theta} y_\theta=n+3m$. Then, the corresponding matching is given by choosing those $T_i\in F$ such that $y_{u_{i4}}=1$.

We consider again the maximum grouping problem $G=(V,E)$ corresponding to a 3DM in Theorem 4. Let X and Y be defined as follows:

$$X=V_A\cup V_B\cup V_C\cup_{T_i\in F}\{u_{i4}\},$$
$$Y=\cup_{T_i\in F}\{u_{i1}, u_{i2}, u_{i3}\}.$$

Then, note that the node set V is partitioned into two disjoint sets X, Y and no edge exists between any two nodes in X and between any two nodes in Y, but edges exist between nodes in X and nodes in Y and thus the following result directly follows.

Corollary 5
For the maximum lot grouping problem $G=(V,E)$ with $d=1, \lambda=3$ and each weight 1, 2 or 3, the

question of deciding if there exists number of $\left\lfloor \sum_{v \in V} w(v)/\lambda \right\rfloor$ feasible lots is NP-Complete even when G is bipartite.

4. Conclusion

In this paper, we introduced the maximum lot grouping problem and modeled it by graph. The similarities between nodes have been represented by the distances between them and the constraint of lot grouping between nodes of specifications is imposed by distance limit. For the distance limit of two, the problem was shown to be NP-Complete even when the graph is bipartite, the lot size is three and each weight is 0 or 1. Next, we considered the case that the distance limit is one. Also in this case, the problem was proved to be still NP-Complete even when the graph is bipartite, the lot size is three and each weight is 1, 2 or 3.

It is open question whether an optimal algorithm exists for the case that $d=1$ and the graph is grid. In general, we need to develop efficient approximation algorithms for maximum grouping problem.

Modeling the constraint on lot grouping by clique size (not by distance limit), we have another interesting research topic. It might be natural to represent the constraint by just establishing edges between any two nodes of similar specifications and then by setting maximum limit on the number of different nodes when constructing a lot, i.e., the maximum clique size limit for constructing heterogeneous lots. The further research would be to show the hardness and devise approximation algorithms for the revised model.

Acknowledgements

This study was supported by research funds from Chosun University 2003. The authors are grateful to the anonymous referees for the constructive comments to improve the paper.

References

Garey M.R., Johnson D.S. (1979), Computers and Intractability: A Guide to the theory of NP-Completeness, San Francisco, CA: Freeman .

Micali S., Vazirani V.V. (1980), An $O(\sqrt{|V|}|E|)$ Algorithm for Finding Maximum Matching in General Graphs, Proc. Twenty-first Annual Symposium on the Foundations of Computer Science 17-27.

Papadimitriou, C.H., Steiglitz K. (1988), Combinatorial Optimization, Mineola, NY: DOVER PUBLICATIONS, INC.