

Department of Industrial Engineering and Management

Technical Report

No. 2013-6

***A note on the complexity
of the maximum edge clique partitioning
problem with respect to clique number***

Noriyoshi Sukegawa and Atsushi Miyauchi



May, 2013

Tokyo Institute of Technology

2-12-1 Ookayama, Meguro-ku, Tokyo 152-8552, JAPAN
<http://www.me.titech.ac.jp/index-e.html>

A note on the complexity of the maximum edge clique partitioning problem with respect to clique number

Noriyoshi Sukegawa^{1,1,*}, Atsushi Miyauchi²

*Graduate School of Decision Science and Technology, Tokyo Institute of Technology,
Ookayama 2-12-1, Meguro-ku, Tokyo 152-8552, Japan.*

Abstract

The maximum edge clique partitioning problem (Max-ECP) was considered by Dessmark et al. (2007) as a tool for DNA clone classification, and is known to be NP-hard on graphs with clique number 3. This complexity result was demonstrated by Punnen and Zhang (2012), however, their proof is incorrect. In this note, we give several complexity results of Max-ECP with respect to clique number, which include a new proof of the above claim.

Keywords: Clique partition, Computational complexity, Clique number

1. Introduction

For a simple and undirected graph $G = (V, E)$, a partition $\{V_1, V_2, \dots, V_m\}$ of V is called a clique partitioning if V_i is a clique of G for each $i \in \{1, 2, \dots, m\}$. The maximum edge clique partitioning problem (Max-ECP) is to find a clique partitioning that maximizes the number of edges within the cliques,

$$\sum_{i=1}^m |\{\{u, v\} \mid u, v \in V_i\}|.$$

This problem was considered by Dessmark et al. [1] as a tool for DNA clone classification. The main focus of the existing studies are on approximation algorithms [1, 4].

In this note, we study the computational complexity of Max-ECP with respect to clique number. The clique number of G is the number of vertices in a maximum clique of G and is denoted by $\omega(G)$. When $\omega(G) = 1$, since G is empty, the problem is trivial. When $\omega(G) = 2$, we see that Max-ECP can be

*Corresponding author

Email addresses: `sukegawa.n.aa@m.titech.ac.jp` (Noriyoshi Sukegawa),
`miyauchi.a.aa@m.titech.ac.jp` (Atsushi Miyauchi)

¹Postal Address: Ookayama 2-12-1-W9-520, Meguro-ku, Tokyo 152-8552, Japan.

²Postal Address: Ookayama 2-12-1-W9-521, Meguro-ku, Tokyo 152-8552, Japan.

reduced to the maximum cardinality matching problem. Hence, in this case, Max-ECP can be solved in polynomial time. When $\omega(G) = 3$, unfortunately, Max-ECP is NP-hard. This complexity result was demonstrated by Punnen and Zhang [4] by reducing the 3-dimensional matching problem (3DM). However, their proof is incorrect. In this note, we give several complexity results of Max-ECP with respect to clique number, which include a new proof of the above claim.

The remainder of this note is organized as follows. In Section 2, we explain the reduction proposed in [4] and give a counterexample. In Section 3, we modify their reduction slightly and demonstrate that this modification enables us to show that Max-ECP is NP-hard on graphs with clique number k , for $k \geq 4$. In Section 4, we show that Max-ECP is NP-hard on graphs with clique number 3 by reducing the triangle cover problem instead of 3DM. Finally, in Section 5, some remarks are given.

2. Counterexample

Punnen and Zhang [4] reduces 3DM to Max-ECP. In 3DM, we are given three disjoint sets X , Y , and Z ($|X| = |Y| = |Z| = n$), and $S \subseteq X \times Y \times Z$. The problem is to find $M \subseteq S$ such that for any two distinct elements $(x, y, z), (x', y', z') \in M$ we have $x \neq x', y \neq y', z \neq z'$, and $|M| = n$.

For a given instance (X, Y, Z, S) of 3DM, let $V := X \cup Y \cup Z$ and

$$E := \bigcup_{(x,y,z) \in S} \{\{x, y\}, \{y, z\}, \{z, x\}\}$$

and construct an undirected graph $G = (V, E)$. Since G is a tripartite graph with partite sets X , Y , and Z , $\omega(G) = 3$. Thus, we see that $3n$ is an upper bound on the optimal value of Max-ECP on G .

Punnen and Zhang [4] mentioned that an instance (X, Y, Z, S) of 3DM has a solution if and only if the optimal value of Max-ECP on G attains $3n$. If this is true, since 3DM is known to be NP-complete [2] on graphs with clique number 3, we can conclude that Max-ECP is NP-hard on graphs with clique number 3. However, the “if” part of their proof is incorrect. In fact, we see that the following 3DM instance $X = \{A, B, C\}$, $Y = \{\alpha, \beta, \gamma\}$, $Z = \{a, b, c\}$, and

$$S = \{(A, \alpha, a), (B, \beta, c), (B, \gamma, b), (C, \beta, b), (C, \gamma, c)\}$$

will be a counterexample.

First, let us verify that the optimal value of Max-ECP on G associated with (X, Y, Z, S) attains $3n$. It is clear that $\{A, \alpha, a\}$ and $\{C, \gamma, c\}$ are cliques of G . Focusing on that (B, β, c) , (B, γ, b) , and (C, β, b) are elements of S , we see that $\{B, \beta\}$, $\{B, b\}$, and $\{\beta, b\}$ are elements of E , respectively. Thus, $\{B, \beta, b\}$ is a clique of G . It follows that the partition $\{\{A, \alpha, a\}, \{B, \beta, b\}, \{C, \gamma, c\}\}$ of V is a clique partitioning. Since the objective value of this partition is 9, we see that the optimal value of Max-ECP on G attains $3n$.

However, as we see below, the instance (X, Y, Z, S) of 3DM has no solution. Suppose that there exists a solution $M^* \subseteq S$. In order to cover α , M^* must include (A, α, a) . In addition, in order to cover γ , M^* must include either (B, γ, b) or (C, γ, c) . Suppose that $(B, \gamma, b) \in M^*$ and $(C, \gamma, c) \notin M^*$. Since (B, γ, b) covers B , we see that $(B, \beta, c) \notin M^*$. In a similar manner, since (B, γ, b) covers b , $(C, \beta, b) \notin M^*$. Hence, M^* cannot cover β , which is a contradiction. Thus $(B, \gamma, b) \notin M^*$ and $(C, \gamma, c) \in M^*$. However, in an analogous way, we can prove that this again yields a contradiction. Therefore, the instance (X, Y, Z, S) has no solution.

3. Modification

A main drawback of the reduction discussed in the previous section is that the information of the given set S is lost in the process of generating G . In this section, we slightly modify their reduction and give several complexity results.

We start with the following theorem.

Theorem 1. *Max-ECP is NP-hard on graphs with clique number 4.*

Proof. For a given instance (X, Y, Z, S) of 3DM, we construct a set of vertices $V^S := \{\phi(x, y, z) \mid (x, y, z) \in S\}$ which is associated with S where $\phi : S \rightarrow V^S$ is an one-to-one mapping. Next, we construct a set of edges

$$E^S := \bigcup_{(x,y,z) \in S} \{\{x, \phi(x, y, z)\}, \{y, \phi(x, y, z)\}, \{z, \phi(x, y, z)\}\}.$$

By using these two sets, we construct an undirected graph $G = (V, E)$ where $V := (X \cup Y \cup Z) \cup V^S$ and

$$E := \left(\bigcup_{(x,y,z) \in S} \{\{x, y\}, \{y, z\}, \{z, x\}\} \right) \cup E^S.$$

The difference from the graph discussed in the previous section is the newly added sets V^S and E^S . We see that there is an one-to-one correspondence between elements of S and cliques of size 4 in G . In Figure 1, we show the clique associated with $(x, y, z) \in S$. We see that G is a 4-partite graph with partite sets $\{X, Y, Z, V^S\}$ and $\omega(G) = 4$. Thus, in this case, $6n$ is an upper bound on the optimal value of Max-ECP on G . (Here, the number 6 of $6n$ is for the number of edges included in a clique of size 4.) We will show that the instance (X, Y, Z, S) has a solution if and only if the optimal value of Max-ECP on G attains $6n$.

We start with a proof of the “only if” part. Suppose that the instance (X, Y, Z, S) of 3DM has a solution, say $\{(x_1, y_1, z_1), (x_2, y_2, z_2), \dots, (x_n, y_n, z_n)\}$. Then, we see that the partition

$$\{\{x_i, y_i, z_i, \phi(x_i, y_i, z_i)\} \mid i = 1, 2, \dots, n\}$$

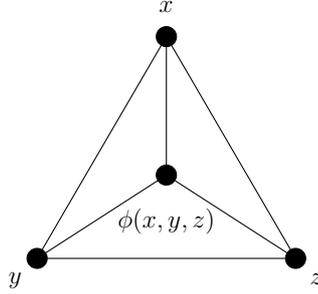


Figure 1: The clique of size 4 associated with $(x, y, z) \in S$ of G

is a clique partitioning of G and its objective value is $6n$. Thus, the optimal value of Max-ECP on G attains $6n$.

Next, we give a proof of the “if” part. Let $\Gamma = \{V_1, V_2, \dots, V_t\}$ be an optimal partition to Max-ECP on G associated with (X, Y, Z, S) . By the assumption, its objective value is $6n$. Since $\omega(G) = 4$, for each $i \in \{1, 2, \dots, t\}$ we have $|V_i| \leq 4$. For convenience, let $V^{[4]} := \{V_i \mid |V_i| = 4, i \in \{1, 2, \dots, t\}\}$ and $n^{[4]} := |V^{[4]}|$.

Suppose that $n^{[4]} = n$. Then, we see that there exists a solution to the instance (X, Y, Z, S) of 3DM. To see this, let

$$V^{[4]} = \{\{x_i, y_i, z_i, \phi(x_i, y_i, z_i)\} \mid i = 1, 2, \dots, n\}.$$

Since for any distinct $i, j \in \{1, 2, \dots, n\}$ we have $x_i \neq x_j$, $y_i \neq y_j$, and $z_i \neq z_j$ (otherwise Γ cannot be a solution of Max-ECP),

$$\{(x_1, y_1, z_1), (x_2, y_2, z_2), \dots, (x_n, y_n, z_n)\} \subseteq S$$

will be a solution to the instance (X, Y, Z, S) of 3DM. Therefore, in what follows, we show that $n^{[4]} = n$.

For a given $v \in V$, let $c_\Gamma(v)$ denote the contribution of v to the objective value of Γ . Namely, if let $W_v \in \Gamma$ be a clique such that $v \in W$,

$$c_\Gamma(v) = \binom{|W|}{2} / |W| = \frac{|W|(|W| - 1)}{2} / |W| = \frac{|W| - 1}{2}.$$

Note that $c_\Gamma(v) = 1.5$ if and only if $W_v \in V^{[4]}$. Note also that, by using this notation, the objective value of Γ can be written as $\sum_{v \in V} c_\Gamma(v)$, which equals $6n$ by the assumption. Hence, the average of the contribution is given as

$$\left(\sum_{v \in V} c_\Gamma(v) \right) / |V| = 6n/4n = 1.5.$$

Then, since $c_\Gamma(v) \leq 1.5$ for each $v \in V$, we see that $c_\Gamma(v)$ must be 1.5 for each $v \in V$. This shows that every vertex of V belongs to a clique of size 4 in Γ . Hence, $n^{[4]} = n$. \square

The idea of the graph defined in the above proof can be utilized to show the NP-hardness of Max-ECP in more general cases with respect to clique number. More specifically, we have the following result.

Theorem 2. *Max-ECP is NP-hard on graphs with clique number k , for $k \geq 4$.*

Proof. When $k = 4$, the claim is equivalent to Theorem 1. Let us consider the case when $k = 5$. We construct two copies of V^S , denoted by V_1^S and V_2^S . These two sets of vertices play the same role as V^S in the previous graph G . For each $i \in \{1, 2\}$ let

$$V_i^S := \{\phi_i(x, y, z) \mid (x, y, z) \in S\},$$

where $\phi_i : S \rightarrow V_i^S$ is an one-to-one mapping. In a similar manner, we construct two copies of E^S , denoted by E_1^S and E_2^S . We associate E_1^S and E_2^S with V_1^S and V_2^S , respectively. Namely, for each $i \in \{1, 2\}$

$$E_i^S := \bigcup_{(x,y,z) \in S} \{\{x, \phi_i(x, y, z)\}, \{y, \phi_i(x, y, z)\}, \{z, \phi_i(x, y, z)\}\}$$

Finally, we construct a new set of edges

$$E_{12}^S := \{\{\phi_1(x, y, z), \phi_2(x, y, z)\} \mid (x, y, z) \in S\}$$

and construct an undirected graph $G' = (V, E)$ where $V = (X \cup Y \cup Z) \cup (V_1^S \cup V_2^S)$ and

$$E = \left(\bigcup_{(x,y,z) \in S} \{\{x, y\}, \{y, z\}, \{z, x\}\} \right) \cup E_1^S \cup E_2^S \cup E_{12}^S.$$

We see that there is an one-to-one correspondence between elements of S and cliques of size 5 in G' . In Figure 2, we show the clique associated with $(x, y, z) \in S$. We see that G' is a 5-partite graph with partite sets $\{X, Y, Z, V_1^S, V_2^S\}$ and $\omega(G') = 5$. Hence, $10n$ is an upper bound on the optimal value of Max-ECP on G' . In an analogous way as in the above proof, we could demonstrate that the instance (X, Y, Z, S) of 3DM has a solution if and only if the optimal value of Max-ECP on G' attains $10n$.

For a general integer $k \geq 4$, preparing $k - 3$ copies $V_1^S, V_2^S, \dots, V_{k-3}^S$ of V^S and $k - 3$ copies $E_1^S, E_2^S, \dots, E_{k-3}^S$ of E^S , and

$$E_{ij}^S = \{\{\phi_i(x, y, z), \phi_j(x, y, z)\} \mid (x, y, z) \in S\}$$

for distinct $i, j \in \{1, 2, \dots, k - 3\}$ should be sufficient. \square

4. New reduction

So far, we have observed that Max-ECP is polynomial time solvable on graphs with clique number less than 3, and is NP-hard on graphs with clique number more than 3. Hence, it would be natural to ask the computational complexity of Max-ECP on graphs with clique number exactly 3. The following theorem answers the question.

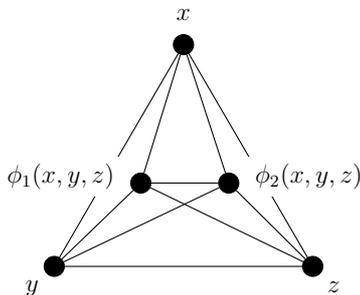


Figure 2: The clique of size 5 associated with $(x, y, z) \in S$ of G'

Theorem 3. *Max-ECP is NP-hard on graphs with clique number 3.*

Proof. We show this by reducing the triangle cover problem (TCP). In TCP, we are given a simple and undirected graph $G = (V, E)$ with $|V| = 3q$ for some $q \in \mathbb{N}$. The problem is to find a clique partitioning $\{V_1, V_2, \dots, V_q\}$ of V such that $|V_i| = 3$ (triangle) for each $i \in \{1, 2, \dots, q\}$. It is known that this problem is NP-complete even when the clique number of a given graph is restricted to 3 [2]. It is not difficult to verify that, when $\omega(G) = 3$, TCP on G has a solution if and only if the optimal value of Max-ECP on G attains $3q$. Therefore, we have the desired result. \square

Finally, we discuss a simple condition for Max-ECP on graphs with clique number 3 to be solved in polynomial time. Let $\Delta(G)$ denote the set of all the cliques of size 3 in G . Then, we have the following lemma.

Lemma 4. *If $\Delta(G)$ is pairwise vertex-disjoint, then there exists an optimal partition of Max-ECP on G which includes $\Delta(G)$.*

Proof. Let Γ be an optimal partition to Max-ECP on G . If $\Delta(G) \subseteq \Gamma$, then we have done. Hence, let us consider the case when there exists a triangle $T \in \Delta(G)$ such that $T \notin \Gamma$. Since $\Delta(G)$ is pairwise vertex-disjoint, for each $v \in T$ there exists an element $W_v \in \Gamma$ with $v \in W_v$ and $|W_v| \leq 2$. Let Γ' be a clique partitioning obtained from Γ by deleting W_v for each $v \in W$ and adding T . Note that the objective value does not decrease more than 3 by deleting E_v for each $v \in W$, while increases exactly 3 by adding W . Hence, by this operation, the objective value does not decrease and the number of triangles included in the partition increases. Therefore, by applying this operation to all the unchosen triangles repeatedly, we have the desired partition. \square

By using the above lemma, we readily see the following result.

Theorem 5. *If $\omega(G) = 3$ and $\Delta(G)$ is pairwise vertex-disjoint, then Max-ECP on G can be solved in polynomial time.*

Proof. Let M^* denote a maximum cardinality matching on $G \setminus \Delta(G)$, where $G \setminus \Delta(G)$ is a graph obtained from G by deleting all the vertices in $\Delta(G)$ and edges which are incident to these vertices. Then, by Lemma 4, we see that the partition $M^* \cup \Delta(G)$ is an optimal solution of Max-ECP on G . Since $M^* \cup \Delta(G)$ can be computed in polynomial time, we have the desired result. \square

5. Final remarks

The vertex-disjoint triangles problem is to find a set of maximum number of pairwise vertex-disjoint triangles, which is a natural extension of TCP to optimization problem. This problem is known to be NP-complete even when the input graphs are chordal, planar, line or total [3]. Since this problem can be seen as a special variant of Max-ECP, we think that it would be an interesting open problem to determine the computational complexity of Max-ECP on these graph classes.

Acknowledgement

We are grateful to Yoshio Okamoto and Yusuke Kobayashi.

References

- [1] A. Dessmark, J. Jansson, A. Lingas, E. M. Lundell, and M. Persson, On the approximability of maximum and minimum edge clique partition problems, *International Journal of Foundations of Computer Science* 18 (2007) 217–226.
- [2] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W.H. Freeman & Co., 1979.
- [3] V. Guruswami, C. P. Rangan, M. S. Chang, G. J. Chang, and C. K. Wong, The Vertex-Disjoint Triangles Problem, *Graph-Theoretic Concepts in Computer Science*, Lecture Notes in Computer Science 1517 (1998) 26–37.
- [4] A.P. Punnen and R. Zhang, Analysis of an approximate greedy algorithm for the maximum edge clique partitioning problem, *Discrete Optimization* 9 (2012) 205–208.