Discrete Mathematics and Theoretical Computer Science DMTCS vol. **17:3**, 2016, 227–244

On the complexity of edge-colored subgraph partitioning problems in network optimization

Xiaoyan Zhang[1,3*] Zan-Bo Zhang[2,3†] Hajo Broersma[3] Xuelian Wen[4‡]

[1] *Institute of Mathematics & School of Mathematical Science,Nanjing Normal University, China*
[2] *Department of Computer Engineering, Guangdong Industry Polytechnic, China*
[3] *Faculty of Electrical Engineering, Mathematics and Computer Science,University of Twente,The Netherlands*
[4] *School of Economics and Management, South China Normal University, China*

received 14th Jan. 2014, revised 5th July 2015, accepted 24th May 2016.

Network models allow one to deal with massive data sets using some standard concepts from graph theory. Understanding and investigating the structural properties of a certain data set is a crucial task in many practical applications of network optimization. Recently, labeled network optimization over colored graphs has been extensively studied. Given a (not necessarily properly) edge-colored graph $G = (V, E)$, a subgraph H is said to be *monochromatic* if all its edges have the same color, and called *multicolored* if all its edges have distinct colors. The monochromatic clique and multicolored cycle partition problems have important applications in the problems of network optimization arising in information science and operations research. We investigate the computational complexity of the problems of determining the minimum number of monochromatic cliques or multicolored cycles that, respectively, partition $V(G)$. We show that the minimum monochromatic clique partition problem is APX-hard on monochromatic-diamond-free graphs, and APX-complete on monochromatic-diamond-free graphs in which the size of a maximum monochromatic clique is bounded by a constant. We also show that the minimum multicolored cycle partition problem is NP-complete, even if the input graph G is triangle-free. Moreover, for the weighted version of the minimum monochromatic clique partition problem on monochromatic-diamond-free graphs, we derive an approximation algorithm with (tight) approximation guarantee $\ln |V(G)| + 1$.

Keywords: network optimization, monochromatic clique, multicolored cycle, partition, forbidden induced subgraphs

*Partially supported by NSFC (11471003), the Priority Academic Program Development of Jiangsu Higher Education Institutions and Qinglan Project.

†Partially supported by NSFC (11471342), Chinese Government Scholarship (201508440189), Natural Science Foundation of Guangdong Province (2016A030313829) and Science and Technology Project of Guangdong Province (2015A020214025). Email:eltonzhang2001@gmail.com.

‡Partially supported by NSFC (11201158).

1365–8050 © 2016 Discrete Mathematics and Theoretical Computer Science (DMTCS), Nancy, France

1 Introduction

Graph based data mining is defined as the science and the art of extracting useful knowledge like patterns and outliers provided by an underlying complex system in order to draw meaningful conclusions regarding the system's properties [1, 11]. The vertices of a complex network denote the entities in a system, and the edges between the vertices represent some kind of relationship between the entities. Network clustering is an important task frequently arising with the aim of partitioning a network into clusters of elements with some similar relationship. In many cases, one can investigate specific properties of a data set by detecting special information in the corresponding clusters, for instance, cliques, cycles, spanning trees and connected components. In particular, edge-colored connected components are often used for solving the important clustering problem arising in data mining, which essentially represents partitioning the set of elements of a certain data set into a number of clusters of objects according to some kind of relationship. For example, a major application of edge-colored graph descriptions of structure arose in sociometry. Social network analysis has grown to be a field in its own right with widely accepted methods used in an increasing variety of applications [3, 13, 25]. Wasserman and Faust [34] describe the main methods and underlying philosophies as well as giving a range of illustrative problems.

1.1 Motivation

In social networks, vertices represent people and edges the relations between them. Different kinds of relationship are distinguished with different colors. The social network analysts need to survey each person about their friends, ask for their approval to publish the data and keep a trace of that population for years. Also the applications, implemented on internet, that use the concept of links between friends and friends of friends, like Google+ which is built on this foundation of "Circles", provide such large structured data sets. One person on Google+ may be connected to many people which can be divided into different circles such as a circle of "Family", a circle of "Friends", a circle of "Employees", and a circle of "Customers". Colored edges can be used to describe the different relations of circles between members. Since a cluster is typically understood as a "tightly knit" group of elements, the graph theoretic concept of a *monochromatic clique*, which is a subset of vertices inducing a monochromatic complete subgraph, is a natural formalization of a cluster that has been used within the context. These monochromatic cliques define cohesive subgroups of some kind of relationship and provide a useful start to the analysis of the structure of social networks, which gives some basis for the study of information exchange and patterns of influence in social networks. The proposed Socratic query is the following: How many "circles" (or monochromatic cliques) are needed to cover the whole graph? Moreover, colors may have weights which can be assigned by the strength or the levels of irrelevant relations between members. This might lead to the problem of partitioning the graph into minimum weight-sum clusters with the highest possible level of cohesiveness which is a natural generalization of the classical clique partition problem.

A *cycle partition* or *cycle cover* of a graph is a spanning subgraph such that each vertex is part of exactly one simple cycle. A special case of the cycle cover problem is the *traveling salesman problem* (TSP), where the goal is to compute a Hamiltonian tour of maximum or minimum weight. The problem of cycle partition is an important tool for the design of approximation algorithms for different variants of the traveling salesman problem [6, 8, 9, 19]. Computing cycle partitions is an important task in the fields of information science, graph theory and combinatorial optimization [23, 26].

Although many studies have been carried out to analyze the complexity of cycle partition problems and the design of approximation algorithms for it [28, 29], few of them considered a more general case of

the problem for edge-colored graphs or edge-labeled graphs. For the edge-colored case, *the multicolored cycle partition* problem is to seek a minimum number of vertex-disjoint multicolored cycles (all the edges of any cycle have distinct colors) in G such that every vertex is in at least one multicolored cycle [20].

1.2 Related results

A variant of the traveling salesman problem called MaxLTSP, where the goal is to compute a Hamiltonian tour of maximum number of colors in an edge-colored complete graph, has been considered in [4, 10]. It is easy to see that MaxLTSP and the multicolored cycle partition problem have a similar relationship to that of TSP and the cycle partition problem. The multicolored cycle partitioning models naturally the need of a maximum covering problem with a certain network structure (in our case such a structure is a multicolored cycle). For example, consider the situation of designing some metropolitan peripheral ring roads, where every color represents a different sub-urban area that a certain link would traverse. In order to minimize the number of peripheral rings such that each of them can cover different sub-urban areas, we wish to use as few multicolored cycles as possible to partition a given edge-colored graph (network).

Labeled network optimization over colored graphs has been extensively studied [16, 17, 30, 21]. Several variations of such problems, and in particular their computational complexity, have been well studied. MacGillivray and Yu [27] studied a general graph partitioning problem including graph coloring, homomorphism to H, conditional coloring, contractibility to H, and partition into cliques as special cases, and investigated their complexity. Yegnanarayanan [35] considered three coloring parameters of a graph G in connection with the computational complexity, partitions, algebra, projective plane geometry and analysis. Jin et al. [18] investigated the computational complexity of the problem of partitioning complete multipartite 2-edge-colored graphs into the minimum number of vertex-disjoint monochromatic cycles, paths and trees, respectively. For more general coloring and partitioning problems, the readers could refer to Garey and Johnson [14], and Kano and Li [20].

Monochromatic clique and multicolored cycle partition problems have important applications in the problems of network optimization arising in information science and operations research mentioned above. We abbreviate the problems of partitioning the vertex set of a (not necessarily properly) edge-colored graph into a minimum number of monochromatic cliques and multicolored cycles to MCLP and MCYP, respectively.

1.3 Diamond-free graphs

In graph theory, many important families of graphs can be described by a finite set of individual graphs that do not belong to the family and further exclude all graphs from the family which contain any of these forbidden graphs as (induced) subgraph or minor. Diamond-free graphs belong to such kinds of important families of graphs. The *diamond graph* is obtained from a complete graph K_4 by removing one edge. A graph is *diamond-free* if it has no diamond as an induced subgraph. The triangle-free graphs are diamond-free graphs, since every diamond contains a triangle. Much research about diamond-free graphs has focused on graph coloring. Characterizations of (subclasses of) diamond-free graphs and their structural properties have been considered both from a theoretical and applications angle. For example, cactus graphs as well as the family of pseudoforests are diamond-free graphs. The former graph family is downwardly closed under graph minor operations and may be characterized by a single forbidden minor which is the diamond graph [12]. They represent electrical circuits that have useful properties [31] and have also recently been used in comparative genomics as a way of representing the relationship between different genomes or parts of genomes [33]. The latter graph family with the characters of both the

butterfly graph and the diamond graph being forbidden minors. Pseudoforests also form graph-theoretic models of functions and occur in several algorithmic problems. Pseudoforests are sparse graphs which have very few edges relative to their number of vertices and their matroid structure allows several other families of sparse graphs to be decomposed as unions of forests and pseudoforests. Pseudoforests also play a key role in parallel algorithms for graph coloring and related problems [15, 22]. Li and Zhang [24] showed that both the problems of determining the minimum number of monochromatic cliques and the minimum number of multicolored cycles that partition $V(G)$ for edge-colored diamond-free graphs are NP-complete.

1.4 Our contribution

In this paper, we prove, by reduction from vertex cover in 3-regular connected graphs, that MCLP is APX-hard for graphs that are monochromatic-diamond-free. Previously, it had been shown in [21] that MCLP is NP-complete for diamond-free (and implicitly also for monochromatic-diamond-free) graphs.

We observe that the algorithms from [24] (polynomial algorithm for finding a largest monochromatic clique, and $O(\log m)$-approximation for MCLP, where m is the size of a largest monochromatic clique) do not work for diamond-free graphs, as claimed in [24], but only for monochromatic-diamond-free graphs (See the Appendix). Furthermore, the algorithmic ideas of [24] are extended to show that for monochromatic-diamond-free graphs, one can enumerate all maximal monochromatic cliques in polynomial time and find a $(\log(|V|)+1)$-approximation algorithm even for a weighted version of MCLP.

We prove, by reduction from set cover, that MCYP is NP-hard even for triangle-free graphs. (The graphs constructed in the previous NP-hardness proof from [21] were diamond-free but not triangle-free.)

2 Preliminaries

Let $G=(V,E)$ be a connected undirected simple graph. If G is assigned a mapping $\ell : E \to \mathbb{N}$, we say that G is an *edge-colored graph*. We call $\ell(e)$ the color of the edge $e \in E$, and we use $\ell(H)$ to denote the number of different colors in the set $\{\ell(e)|e \in E(H)\}$ for a subgraph H of G. A *complete graph* is a graph in which every two distinct vertices are adjacent. We denote by K_m a complete graph on m vertices, and by C_m a cycle on m vertices. A *clique* of G is a nonempty subset of $V(G)$ that induces a complete subgraph of G. A clique CL of G is called a *monochromatic clique* if all the edges of the corresponding subgraph of G have the same color. A cycle CY of G is called a *multicolored cycle* if $\ell(CY) = |E(CY)|$, i.e., if no two edges of CY have the same color. Note that a single vertex can be viewed as a degenerate monochromatic clique or multicolored cycle. We simply call it a *vertex-clique* or *vertex-cycle*.

A subgraph H of an edge-colored graph G is called *monochromatic-induced* if H is monochromatic with edge color c and for any pair of vertices $u, v \in V(H)$, uv is an edge of H if and only if uv is an edge of G with color c. And a graph G is called *monochromatic-diamond-free* if it does not contain a monochromatic diamond as a monochromatic-induced subgraph. Note that the properties of being diamond-free and monochromatic-diamond-free do not imply each other. For example, a K_4 with one edge colored ℓ_1 and the others colored ℓ_2 is diamond-free, but not monochromatic-diamond-free. However, a monochromatic cycle on 4 vertices with a chord of a different color is monochromatic-diamond-free, but not diamond-free. A vertex u is *color-adjacent* to a vertex v of a monochromatic clique CL if the edge uv has the same color as the edges of CL. A clique CL of G is called a *maximal monochromatic clique* if there is no vertex u of G color-adjacent to each vertex of CL.

The paper is organized as follows. In Section 3, we show that MCLP is APX-complete on monochromatic-diamond-free graphs with maximum degree 6. In Section 4, we generalize MCLP to its weighted version WMCLP, and present a greedy scheme that yields an $\ln |V(G)| + 1$-approximation algorithm for WMCLP on monochromatic-diamond-free graphs. We also provide an example to show that the approximation guarantee is tight. In Section 5, we investigate the complexity of MCYP and show that the problem is NP-complete, even if the input is a triangle-free graph. In the final section, we present some concluding remarks and propose some open problems for further research.

3 Inapproximability of MCLP on monochromatic-diamond-free graphs

Given a graph $G = (V, E)$ and a positive integer k, the *Partition into cliques* (PIC) decision problem consists of deciding whether there exists a partition of V into k disjoint subsets such that the subgraph induced by each part of the subsets is a clique of G. Garey and Johnson [14] proved that PIC is NP-complete for K_4-free graphs. Cerioli et al. [5] establish both the NP-completeness of PIC for planar cubic graphs and the Max SNP-hardness of PIC for cubic graphs. They also presented a deterministic polynomial time 5/4-approximation algorithm for finding clique partitions in maximum degree three graphs. It is easily seen that PIC is a special case of the decision version of MCLP if the graph G is colored by a unique color. Li and Zhang [24] have proved that MCLP is NP-complete, even when the input is restricted to diamond-free graphs. They showed a polynomial algorithm to find an approximation solution for MCLP in diamond-free graphs with performance ratio $\ln m + 1$, where m is the size of a maximum monochromatic clique in the input graph. However, the algorithm actually works for monochromatic-diamond-free graphs instead of diamond-free graphs. We added Appendix A to explain the details.

Hence, if the input graph for MCLP is monochromatic-diamond-free with the size of a maximum monochromatic clique bounded by a constant, we have an approximation algorithm with constant performance ratio.

We further investigate the inapproximability of MCLP. Alimonti and Kann [2] have shown that the Vertex Cover problem restricted to 3-regular connected graphs is APX-complete. This implies that there is some small $\epsilon > 0$ such that the existence of a polynomial time approximation algorithm for finding a minimum cardinality vertex cover in a connected 3-regular graph with performance guarantee $1 + \epsilon$ would imply $\mathrm{P} = \mathrm{NP}$.

L-reduction is a transformation of optimization problems which linearly preserves approximability features and it is one type of approximation-preserving reduction. L-reductions in studies of approximability of optimization problems play a similar role to that of polynomial reductions in the studies of computational complexity of decision problems. L-reductions preserve membership in APX for the minimizing case only, as a result of implying AP-reductions. Its definition is given as follows [32].

Let f be a polynomial-time transformation from a minimization optimization problem Π to a minimization optimization problem Π' , We say that f is an L-reduction if there are constants $\alpha, \beta > 0$ such that for each instance I of Π:

a) The optima of I and $f(I)$, $OPT(I)$ and $OPT(f(I))$ respectively, satisfy $OPT(f(I)) \leq \alpha OPT(I)$.

b) For any solution of $f(I)$ with cost c, we can find in polynomial time a solution of I with cost at most $OPT(I) + \beta[c - OPT(f(I))]$.

The constant β will be usually 1. We now give an approximation preserving L-reduction from the Vertex Cover problem in 3-regular connected graphs to MCLP and draw the following conclusions.

Theorem 3.1 *MCLP is*

(1) APX-hard on monochromatic-diamond-free graphs, and

(2) APX-complete on monochromatic-diamond-free graphs with the size of a maximum monochromatic clique bounded by a constant.

3.1 Proof of Theorem 3.1

Consider an arbitrary instance of the Vertex Cover problem in 3-regular connected graphs. So let $G = (V, E)$ be a 3-regular connected graph, with $|V| = 2n$ and $|E| = 3n$ for some $n \in \mathbb{Z}^+$. A corresponding MCLP instance on an edge-colored graph $H = (V_H, E_H)$ is constructed from G in the following way. H is obtained from G by replacing every edge $(u, v) \in E$ by a gadget $g(u, v)$ consisting of the vertices u and v as well as two new vertices $e^1_{u,v}$ and $e^2_{u,v}$, and the edges $(u, e^1_{u,v})$, $(u, e^2_{u,v})$, $(v, e^1_{u,v})$ and $(v, e^2_{u,v})$. Furthermore, for a vertex u with neighbors v, w and x, the vertices $e^1_{u,v}$, $e^1_{u,w}$ and $e^1_{u,x}$ are made mutually adjacent in H. For every vertex $u \in V$, we define a color $\ell(u)$, and for every edge $(u, v) \in E$, we define two colors $\ell(u, e_{u,v})$ and $\ell(v, e_{u,v})$, where all the colors we define are different. For an edge $(u, v) \in E$, the corresponding edges in H are colored as follows. The edge $(u, e^1_{u,v})$ is assigned color $\ell(u)$ and the edge $(v, e^1_{u,v})$ is assigned color $\ell(v)$. The edge $(u, e^2_{u,v})$ is assigned color $\ell(u, e_{u,v})$ and the edge $(v, e^2_{u,v})$ is assigned color $\ell(v, e_{u,v})$. For a vertex u with neighbors v, w and x in G, the edges $(e^1_{u,v}, e^1_{u,w})$, $(e^1_{u,w}, e^1_{u,x})$ and $(e^1_{u,x}, e^1_{u,v})$ in H are all assigned color $\ell(u)$. This completes the construction and edge-coloring of the graph H (See Figure 1). It is easy to observe that a largest monochromatic clique in H corresponds to a K_4, and that H is monochromatic-diamond-free, with maximum degree 6. Note that the degree of the vertices u, v, and $e^1_{v,u}$ is exactly 6 for every gadget $g(u, v)$.

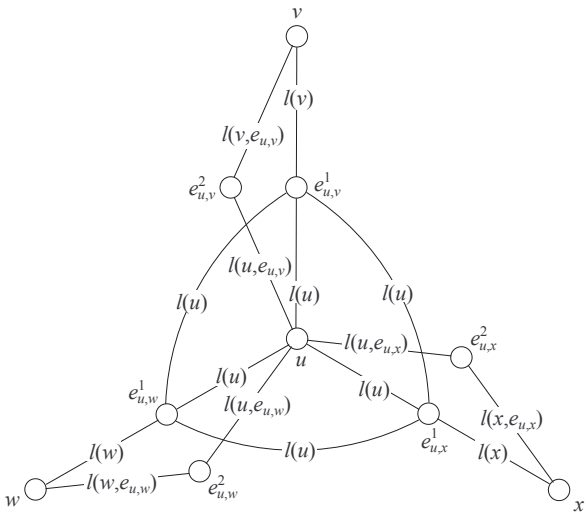

Fig. 1: Gadgets near the vertex u

Let V_c^* be a minimum vertex cover of G, and let P^* be a minimum monochromatic clique partition of H. Then, we have the following inequalities.

Lemma 3.2 $|P^*| \leq 8n \leq 8|V_c^*|$

Proof: Since every vertex in G is incident with exactly three edges, V_c^* has at least $|E|/3 = n$ vertices. There are $|V| + 2|E| = 8n$ vertices in H, so H can be partitioned into $8n$ vertex-cliques. Hence, $|P^*| \leq 8n \leq 8|V_c^*|$. □

Suppose P is an arbitrary monochromatic clique partition of H. We further have the following conclusion.

Lemma 3.3 *P can always be turned into a new monochromatic clique partition P' such that $|P'| \leq |P|$ and for every edge (u, v) of G, $e^1_{u,v} \in K(u)$ or $e^1_{u,v} \in K(v)$ holds and there is no vertex-clique u or v in P'. Here $K(v)$ denotes a vertex-clique v or a (nontrivial) monochromatic clique containing v.*

Proof: First suppose that $K(e^1_{u,v})$ is a vertex-clique in P, or is a monochromatic clique with color $\ell(u)$ that does not contain u. We can execute one of the following operations on P to merge $K(e^1_{u,v})$ into $K(u)$ or u into $K(e^1_{u,v})$, without increasing the cardinality of P. If u forms a vertex-clique or is contained in a monochromatic clique with color $\ell(u)$, then $K(e^1_{u,v})$ can be combined with $K(u)$ to obtain a larger monochromatic clique with color $\ell(u)$. If u is contained in a monochromatic clique with a color different from $\ell(u)$, then u can be taken away from $K(u)$ and combined with $K(e^1_{u,v})$ to form a new clique with color $\ell(u)$.

Therefore, we may assume that $e^1_{u,v} \in K(u)$ or $K(v)$ for all edges (u, v) in G. If there exists a vertex-clique v in a gadget $g(u, v)$ after executing the above operations, then $e^1_{u,v} \in K(u)$, and $e^2_{u,v}$ forms a vertex-clique. Hence, v can be combined with $e^2_{u,v}$ to form a new monochromatic clique with color $\ell(v, e_{u,v})$, and the cardinality of the partition is decreased.

After applying the above operations we have obtained a new partition P' with $|P'| \leq |P|$, satisfying the conditions in the lemma. □

Let $g(u, v)$ be a gadget in H. Without loss of generality we may assume that $e^1_{u,v} \in K(u)$ with color $\ell(u)$ in P'. Then $e^2_{u,v}$ forms either a vertex-clique, or a clique with v of color $\ell(v, e^2_{u,v})$.

Let V_c be composed of all the vertices $u \in V$ such that for some edge $(u, v) \in E$, $e^1_{u,v} \in K(u)$ in P'. Since for every edge $(u, v) \in E$, $e^1_{u,v} \in K(u)$ or $e^1_{u,v} \in K(v)$ in P', at least one of u and v is in V_c. Hence V_c is a vertex cover of G.

For every edge $(u, v) \in E$, $e^2_{u,v}$ forms either a vertex-clique or a monochromatic clique together with v or u in P'. There are totally $|E|$ such cliques. Each of the other cliques in P' contains exactly one vertex $u \in V$ and at least one vertex $e^1_{u,v}$ for some neighbor v of u in G, and hence corresponds to a vertex $u \in V_c$. Consequently,

$$|V_c| = |P'| - |E| = |P'| - 3n \leq |P| - 3n. \tag{1}$$

On the other hand, we can obtain a monochromatic clique partition $\widetilde{P}$ of H from a minimum vertex cover V_c^* of G, as follows. For a gadget $g(u, v)$ in H, if (u, v) is covered by exactly one end vertex in V_c^*, say u, then let $e^1_{u,v}$ be in the same clique with u in $\widetilde{P}$, and hence $K(u)$ is of color $\ell(u)$. If (u, v) is covered by both u and v, then let $e^1_{u,v}$ be in the same clique with either u or v in $\widetilde{P}$ arbitrarily.

Since every edge is covered by at least one vertex, every vertex of type $e^1_{u,v}$ is contained in either $K(u)$ or $K(v)$. We claim that for every vertex $u \in V_c^*$, $K(u)$ contains at least one vertex of type $e^1_{u,v}$. For if there exists a $u_0 \in V_c^*$ such that $K(u_0)$ contains no vertex $e^1_{u_0,v}$ for every neighbor v of u_0 in G, then $V_c^* \backslash \{u_0\}$ is a vertex cover of G with cardinality less than V_c^*, contradicting the minimality of V_c^*. Since no two vertices in G can be in the same clique in $\widetilde{P}$, there are exactly $|V_c^*|$ cliques in $\widetilde{P}$ containing vertices of type $e^1_{u,v}$.

For any vertex v of G that is not contained in V_c^*, let v form a clique in $\widetilde{P}$ with a vertex $e^2_{u,v}$ for a neighbor u of v in G. Note that such a vertex $e^2_{u,v}$ is always available for v, since any neighbor of v in G must be in V_c^*.

Finally, we let the remaining vertices of type $e^2_{u,v}$ be vertex-cliques in $\widetilde{P}$. $\widetilde{P}$ consists of $|E|$ cliques containing vertices of type $e^2_{u,v}$, and $|V_c^*|$ cliques containing vertices of type $e^1_{u,v}$, therefore

$$|V_c^*| = |\widetilde{P}| - |E| = |\widetilde{P}| - 3n. \tag{2}$$

Thus,

Lemma 3.4 *$\widetilde{P}$ is a minimum monochromatic clique partition.*

Proof: For if there exists a monochromatic clique partition P with $|P| < |\widetilde{P}|$, then by the above discussion we can always obtain a vertex cover V_c of G with $|V_c| \leq |P| - 3n < |\widetilde{P}| - 3n = |V_c^*|$, contradicting the minimality of V_c^*. □

Lemma 3.5 *The existence of a polynomial time approximation scheme for MCLP restricted to monochromatic-diamond-free graphs with maximum monochromatic clique K_4 would imply the existence of a polynomial time approximation scheme for the Vertex Cover problem restricted to 3-regular connected graphs.*

Proof: Given an instance of the Vertex Cover problem restricted to 3-regular connected graphs, we have known that it can be turned into an instance of MCLP restricted to monochromatic-diamond-free graphs with maximum monochromatic clique K_4. We can assume that the monochromatic clique partition P we find satisfies the condition that every vertex of type $e^1_{u,v}$ is contained in $K(u)$ or in $K(v)$, and that there is no vertex-clique u or v in P. Then from P we can obtain a solution V_c for the instance of the Vertex Cover problem, in the way we discussed above. We keep using the notations P^* and V_c^* to denote the optimal solutions for both problems. We have $|P| = |V_c| + 3n$ and $|P^*| = |V_c^*| + 3n$. Further, using $|P^*| \leq 8|V_c^*|$ we have $3n \leq 7|V_c^*|$.

Suppose there exists a small positive ϵ such that $|P| \leq (1+\epsilon)|P^*|$. Substituting V_c and V_c^* into the inequality, we get $|V_c|+3n \leq (1+\epsilon)(|V_c^*|+3n)$, that is, $|V_c| \leq (1+\epsilon)|V_c^*|+3n\epsilon \leq (1+\epsilon)|V_c^*|+7\epsilon|V_c^*| = (1+8\epsilon)|V_c^*|$.

Therefore, we complete the proof of this lemma. □

Since the Vertex Cover problem restricted to 3-regular connected graphs is APX-complete, we have that MCLP is APX-hard on monochromatic-diamond-free graphs. Finally, we have the algorithm from [24] that works out a solution with a constant approximation ratio for MCLP in monochromatic-diamond-free graphs in which the size of a maximum monochromatic clique is bounded by a constant. Therefore, statement (2) holds.

4 An approximation algorithm for WMCLP

We generalize MCLP to its weighted version WMCLP. Let G be an edge-colored graph with colors $\ell(G)$. Each color $c \in \ell(G)$ is associated with a non-negative cost $w(c)$. Every monochromatic clique CL of G with at least two vertices has the same non-negative cost as its color, denoted by $w(CL)$. As any vertex v of G is viewed as a degenerate monochromatic clique, we also assign it a non-negative cost $w(v)$, with $w(v) \leq \min\{w(c)|c \in \ell(G)\}$. WMCLP asks for a monochromatic clique partition such that the sum of the costs of all cliques in the partition is minimal among all the possible partitions. Obviously, MCLP is the special case of WMCLP in which all the costs are equal to 1.

Li and Zhang [24] presented a polynomial algorithm, denoted by Alg(clique), which was claimed to calculate all maximal monochromatic cliques in a diamond-free graph and return a maximum one. However, their claim holds only when diamond-free is replaced by monochromatic-diamond-free. We refer the reader to Appendix A for details.

In this paper, we use Alg(clique) as a subroutine to find all maximal monochromatic cliques in our $(\ln |V(G)|+1)$-approximation algorithm (Algorithm 1) for solving WMCLP restricted to monochromatic-diamond-free graphs. In Algorithm 1, Alg(clique) is implemented from Step 2 to Step 9.

Let G be a monochromatic-diamond-free graph, and let CL_1 and CL_2 be two distinct maximal monochromatic cliques in G. Suppose that there is at least one common edge (u, v) with color c of CL_1 and CL_2. Since CL_1 and CL_2 are maximal, there must be at least one vertex $w \in V(CL_1)\backslash V(CL_2)$ and one vertex $x \in V(CL_2)\backslash V(CL_1)$ such that w and x are not adjacent by an edge of color c. But then u, v, w and x span a monochromatic-diamond in G, a contradiction. Therefore, any two distinct maximal monochromatic cliques in a monochromatic-diamond-free graph do not share a common edge.

We note that MCLP can be considered as a variant of the Set Cover problem, in which the (possibly exponentially many) subsets are the vertex sets of all the monochromatic cliques and vertex-cliques of the input graph G, and the objective is to find a minimum collection of pairwise disjoint subsets covering the vertex set of G. Hence, it is natural that our design of a greedy approximation algorithm for WMCLP is inspired by the greedy algorithm for the weighted Set Cover problem in [7].

Theorem 4.1 *Algorithm 1 runs in polynomial time and achieves the performance ratio* $\ln |V(G)| + 1$ *for WMCLP on a monochromatic-diamond-free graph* G.

Proof: In this proof, we do not distinguish between a clique and its vertex set.

First we claim that the set C contains all maximal monochromatic cliques after the execution of the loop from Step 2 to Step 9 in Algorithm 1. Since any two maximal monochromatic cliques do not share an edge in G, every edge belongs to one maximal monochromatic clique. Hence, we can start from the end vertices of any edge, and find out the maximal monochromatic clique containing the edge through the loop from Step 5 to Step 7. Then, all edges of this clique are removed from $E(G)$. Repeating this process until $E(G)$ becomes empty, all maximal monochromatic cliques of G are found. The running time of the loop from Step 2 to Step 9 is at most $O(|E||V|^2) = O(|V|^4)$.

Assume that the loop from Step 11 to Step 17 is iterated r times. Let the vertex-clique or maximal monochromatic clique picked in Step 12 at the i-th iteration of the loop be q_i, for $1 \leq i \leq r$. Let $G = G_1$ and $G_{i+1} = G_1 \setminus \{q_1 \cup q_2 \cup \ldots \cup q_i\} = G_i \setminus q_i$, for $1 \leq i \leq r-1$. The algorithm outputs $D = \{q_i, 1 \leq i \leq r\}$ as a solution.

It is easy to prove by induction that Q contains all the vertex-cliques and maximal monochromatic cliques of G_{i+1} at the i-th iteration after the execution of Step 16, for $1 \leq i \leq r-1$. We denote by

Algorithm 1 An approximation algorithm for WMCLP on monochromatic-diamond-free graphs

Input: A monochromatic-diamond-free graph G;
Output: A monochromatic vertex-disjoint clique partition D of G;

1: Let $C := \emptyset, D := \emptyset$;
2: **repeat**
3: Select an edge $(v_i, v_j) \in E(G)$;
4: Let $S := \{v_i,\ v_j\}$;
5: **while** there is a vertex v_k which is color-adjacent to each vertex of the monochromatic clique $G[S]$ **do**
6: $S = S \cup \{v_k\}$;
7: **end while**
8: $C = C \cup S$, $E(G) = E(G) \setminus E(S)$, where $E(S)$ denotes the edges of G with both end vertices in S;
9: **until** no edge in $E(G)$.
10: Let $Q := V(G) \cup C$;
11: **repeat**
12: Pick $q \in Q$ such that the ratio $w(q)/|q|$ is minimum, where $w(q)$ denotes the weight of the monochromatic clique $G[q]$;
13: Let $C' := \emptyset$;
14: **for all** $c \in C$ **do**
$c = c \setminus q$, $C' = C' \cup c$;
15: **end for**
16: $D = D \cup \{q\}, V(G) = V(G) \setminus q, Q = V(G) \cup C', C = C'$;
17: **until** $V(G) = \emptyset$.
18: **return** D.

P_i an optimal monochromatic clique partition of G_i and $w(P_i)$ the cost of P_i, for $1 \leq i \leq r$. Note that P_1 is an optimal solution of the problem. Let j be an integer such that $1 \leq j \leq r-1$. Let the number of cliques in P_j be t, and $P_j = \{p_{j1},\ p_{j2},\ \ldots,\ p_{jt}\}$. Then $P'_j = \{p_{ji} \setminus q_j : 1 \leq i \leq t\}$ is a monochromatic clique partition of G_{j+1}. Hence, $w(P_{j+1}) \leq w(P'_j) = \sum_{i=1}^{t} w(p_{ji} \setminus q_j)$. Furthermore, for each $p_{ji} \in P_j$, if $|p_{ji} \setminus q_j| \geq 2$, $w(p_{ji}) = w(p_{ji} \setminus q_j)$; otherwise $w(p_{ji}) \geq w(p_{ji} \setminus q_j)$. Therefore, $w(P_{j+1}) \leq w(P'_j) \leq \sum_{i=1}^{t} w(p_{ji}) = w(P_j)$.

Note that for any monochromatic clique q and $q' \subseteq q$ with $|q'| \geq 2$, the relation $\frac{w(q)}{|q|} \leq \frac{w(q')}{|q'|}$ holds. Therefore, the clique q_i picked in Step 12 at the i-th iteration has the minimum ratio $\frac{w(q_i)}{|q_i|}$ over all vertex-cliques and monochromatic cliques of G_i. So we have $\frac{w(q_i)}{|q_i|} \leq \frac{w(p_{ij})}{|p_{ij}|}$, for $1 \leq j \leq t$. Therefore,

$$w(P_i) = \sum_{j=1}^{t} \frac{|w(p_{ij})|}{|p_{ij}|}|p_{ij}| \geq \frac{w(q_i)}{|q_i|} \sum_{j=1}^{t} |p_{ij}| = \frac{w(q_i)}{|q_i|}|V(G_i)|,$$

or, $w(q_i) \leq \frac{w(P_i)|q_i|}{|V(G_i)|}$. Hence,

$$\begin{aligned} w(D) &= \sum_{i=1}^{r} w(q_i) \leq \sum_{i=1}^{r} \frac{w(P_i)|q_i|}{|V(G_i)|} \leq w(P_1) \sum_{i=1}^{r} \frac{|q_i|}{|V(G_i)|} \\ &\leq w(P_1) \sum_{i=1}^{r} \sum_{k=0}^{|q_i|-1} \frac{1}{|V(G_i)|-k} = w(P_1) H(|V(G)|) \\ &\leq w(P_1)(\ln |V(G)| + 1), \end{aligned}$$

where $H(|V(G)|)$ is the $|V(G)|$-th harmonic number. □

To show that the above approximation ratio is tight, we present an example to demonstrate that the approximation algorithm may find a solution with cost $H(n)$ times the optimum, where n is the number of vertices of the graph.

Let G be an edge-colored complete graph with vertices v_1, v_2, …,v_n. And let the cost of every vertex-clique v_i be $\frac{1}{i+\epsilon}$ for $i = 1, 2, \ldots, n$, where ϵ is a very small positive number. All edges $e \in E(G)$ have the same color of cost 1. It is not difficult to verify that Algorithm 1 finds a solution consisting of all vertex-cliques in the order v_n, v_{n-1}, …, v_1, with total cost $\sum_{i=1}^{n} \frac{1}{i+\epsilon}$, and hence arbitrarily close to $H(n)$, whereas the optimal solution picks G directly, with cost 1.

5 MCYP is NP-complete for triangle-free graphs

We first consider several trivial cases of MCYP. Let G be an edge-colored graph on n vertices. If G is colored with a small number of colors, say $\ell(G) = 1$ or $\ell(G) = 2$, then we can only partition G into n vertex-cycles. When G is colored by the largest possible number of colors, that is, $\ell(G) = |E(G)|$, then finding the minimum multicolored cycle partition of G is at least as hard as finding a Hamiltonian cycle of G, which is well-known to be NP-hard. For $\ell(G) \geq 3$, Li and Zhang [24] showed that MCYP is NP-complete, even if the input graph G is diamond-free. Their proof is based on a reduction from the Exact Cover By 3-Sets problem. We achieve a further strengthening by showing that MCYP is NP-complete, even if the input graphs are restricted to triangle-free graphs, a proper subclass of diamond-free graphs.

Theorem 5.1 *MCYP is NP-complete when restricted to triangle-free graphs.*

5.1 Proof of Theorem 5.1

MCYP on triangle-free graphs is clearly in NP: a nondeterministic algorithm needs only guess a set of cycles of the input graph, and check in polynomial time whether the cycles in the set are vertex-disjoint multicolored cycles that cover all the vertices of the graph, and whether the number of cycles in the set is no larger than a given positive number.

Our proof of the NP-completeness of MCYP is based on a reduction from the Minimum Set Cover problem. In an instance of the Minimum Set Cover problem, a universe U of n elements, a collection of subsets of U, $S = \{s_1, ..., s_m\}$ where $\cup_{i=1}^{m} s_i = U$, and a positive integer $k \leq \min\{m, n\}$ are given. The question is whether there exists a subcollection C of S with $|C| \leq k$ that covers all the elements of U.

Suppose now that we are given an instance of the Minimum Set Cover problem, with the universe $U = \{u_i | 1 \leq i \leq n\}$ and the subset collection $S = \{s_j | 1 \leq j \leq m\}$, where $\cup_{i=1}^{m} s_i = U$. We construct an edge-colored triangle-free graph G as follows.

The vertex set of G is the union of the sets $U_i = \{u_i^1, u_i^2\} \cup \{u_{ij}^1, u_{ij}^2 | 1 \leq j \leq m\}, 1 \leq i \leq n$, and $S_i = \{s_{ij} | 0 \leq j \leq n+1\}, 1 \leq i \leq m$.

We define the following colors for the edges of G.

1. For $1 \leq i \leq n$, $1 \leq j \leq 2m+1$, define the colors c_{ij}.
2. For $0 \leq i \leq n$, define the colors ℓ_i.
3. For $1 \leq i \leq m$, define the colors ℓ_{i1} and $\ell_{i(n+1)}$.
4. For $1 \leq i \leq m$, supposing the set s_i contains t_i elements, define the colors d_{i1}, d_{i2}, ..., $d_{i(t_i-1)}$.

The edges of G and their colors are given below.

1. For $1 \leq i \leq m$, form the cycles $CY(s_i) = s_{i0}s_{i1} \ldots s_{i(n+1)}s_{i0}$. Assign color ℓ_{i1} to edge (s_{i0}, s_{i1}), and color $\ell_{i(n+1)}$ to edge $(s_{i0}, s_{i(n+1)})$. Each of the other n edges on the cycle is assigned a color from the set $\{c_{j1} : u_j \notin s_i, 1 \leq j \leq n\} \cup \{\ell_0\} \cup \{d_{ij} : 1 \leq j \leq t_i - 1\}$, so that no two edges have the same color. Note that we have exactly n colors in the color set, because t_i is the number of u_j's contained in s_i. We denote by $P(s_{i1}, s_{i(n+1)})$ the path $s_{i1}s_{i2} \ldots s_{i(n+1)}$.
2. For $1 \leq i \leq n$, form the cycles $CY(u_i) = u_i^1 u_{i1}^1 u_{i2}^1 \ldots u_{im}^1 u_{i1}^2 u_{i2}^2 \ldots u_{im}^2 u_i^2 u_i^1$. Assign color ℓ_i to the edge (u_i^1, u_i^2). Denote the paths $u_i^1 u_{i1}^1 u_{i2}^1 \ldots u_{im}^1 u_{i1}^2 u_{i2}^2 \ldots u_{im}^2 u_i^2$ by $P(u_i^1, u_i^2)$. Assign colors $c_{ij}, 1 \leq j \leq 2m+1$ to the $2m+1$ edges of $P(u_i^1, u_i^2)$ successively.
3. For $1 \leq j \leq n$, $1 \leq i \leq m$, if s_i contains u_j, join u_j^1 and s_{i1} by an edge and assign color ℓ_{i1} to the edge, and join u_j^2 and $s_{i(n+1)}$ and assign color $\ell_{i(n+1)}$ to the edge.
4. For $1 \leq j < r \leq n$, if there exists a set s_i, $1 \leq i \leq m$, containing both u_j and u_r, then join u_j^1 and u_r^2, and u_j^2 and u_r^1, and assign color ℓ_j to the edges.

Figure 2 is an illustration of G. It is easy to verify that G is triangle-free and that the construction can be accomplished in polynomial time.

Now suppose that G has a multicolored cycle partition P with $k' + m \leq k + m$ multicolored cycles. We list the following properties of P, which are crucial for our proof.

1. For $1 \leq i \leq m$, the vertex s_{i0} either forms a vertex-cycle, or is contained in the cycle $CY(s_i)$ of P, and hence we have m cycles in P, each containing one s_{i0}. To see this, suppose that s_{i0} is contained in a multicolored cycle CY. CY must contain the edges (s_{i0}, s_{i1}) and $(s_{i(n+1)}, s_{i0})$. Since every edge associated with s_{i1} has color ℓ_{i1} except for the edge (s_{i1}, s_{i2}), CY must contain (s_{i1}, s_{i2}). Similarly CY must contain the edge $(s_{in}, s_{i(n+1)})$ and hence $CY = CY(s_i)$.
2. For $1 \leq i \leq m$, the path $P(s_{i1}, s_{i(n+1)})$ is contained in a multicolored cycle; otherwise, the vertices s_{i0}, s_{i1}, ..., $s_{i(n+1)}$ must form $n+2$ vertex-cycles in P. However, we have m multicolored cycles to cover all s_{i0} by Property 1. Together we need at least $m + n + 2 - 1 = m + n + 1 > k' + m$ cycles in P, a contradiction.
3. For any $1 \leq i \neq j \leq m$, the two paths $P(s_{i1}, s_{i(n+1)})$ and $P(s_{j1}, s_{j(n+1)})$ cannot be contained in the same multicolored cycle, since both of them have an edge of color ℓ_0.

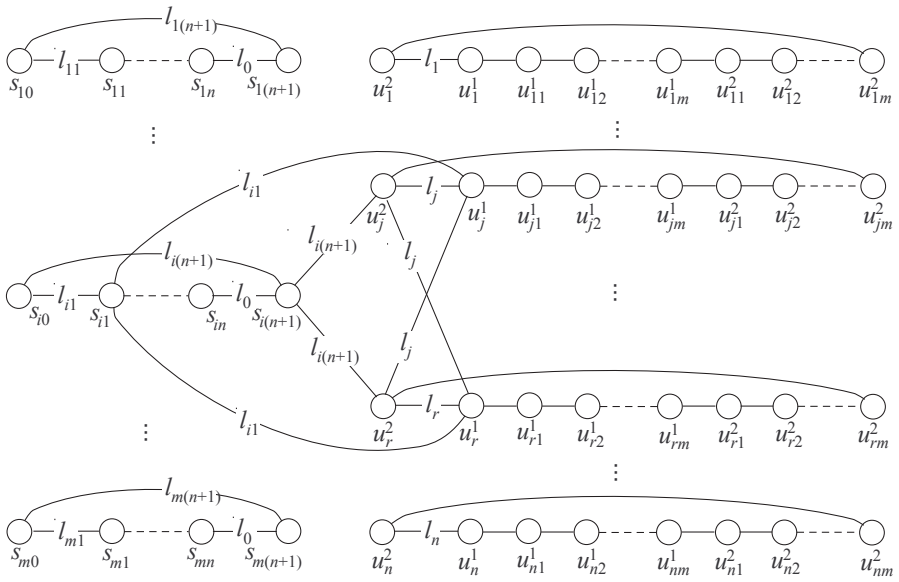

Fig. 2: The graph G constructed from an instance of the Minimum Set Cover problem, in which $u_j,\ u_r \in s_i$ and $j < r$. Labels on some edges denote the colors of the edges. We assume that the edges $(s_{tn}, s_{t(n+1)})$ are colored l_0, $1 \leq t \leq m$.

4. For $1 \leq i \leq n$, the path $P(u_i^1, u_i^2)$ is contained in a multicolored cycle; otherwise, the $2m$ vertices u_{ij}^1 and u_{ij}^2, $j = 1, \ldots, m$ would form $2m$ vertex-cycles in P. Then, there would be at least $2m + 1 > k' + m$ cycles in P, a contradiction.

5. If there is a multicolored cycle CY in P containing only vertices in $\cup_{j=1}^{n} U_j$, then $CY = CY(u_i)$, for some $1 \leq i \leq n$. To see this, let j_0 be the smallest index such that CY contains some vertices in U_{j_0}. By Property 4, CY contains the paths $P(u_{j_0}^1, u_{j_0}^2)$. If CY contains some more vertices in $\cup_{j=1}^{n} U_j$, then there must be two edges joining $P(u_{j_0}^1, u_{j_0}^2)$ to the other part of CY, which must be associated with $u_{j_0}^1$ and $u_{j_0}^2$. However, both edges have the same color ℓ_{j_0}, contradicting the multicolored property of CY. Therefore $CY = CY(u_{j_0})$.

6. If s_j does not contain u_i, then the path $P(u_i^1, u_i^2)$ cannot be in a multicolored cycle which contains the path $P(s_{j1}, s_{j(n+1)})$. The reason is that the first edge of $P(u_i^1, u_i^2)$ is colored c_{i1}, while there is also an edge of color c_{i1} on the path $P(s_{j1}, s_{j(n+1)})$.

Lemma 5.2 *There are only four possible kinds of multicolored cycles in the partition P, as follows.*

(1) The cycles $CY(s_i)$, for some $1 \leq i \leq m$.

(2) The cycles which contain a path $P(s_{i1}, s_{i(n+1)})$ and several paths $P(u_j^1, u_j^2)$, where $u_j \in s_i$ for some $1 \leq i \leq m$ and $1 \leq j \leq n$.

(3) The vertex-cycles formed by s_{i0} for some $1 \leq i \leq m$.

(4) The cycles $CY(u_i)$, for some $1 \leq i \leq n$.

Proof: By Property 1, every vertex s_{i0} must be contained in cycle kind of (1) or (3). If s_{i0} forms a vertex-cycle, then by Property 2 and Property 3, the path $P(s_{i1}, s_{i(n+1)})$ must form a multicolored cycle with some vertices from $\cup_{i=1}^{n} U_i$ in P. Furthermore, by Property 4 and Property 6, $P(s_{i1}, s_{i(n+1)})$ must form a multicolored cycle with some paths $P(u_j^1, u_j^2)$, where $u_j \in s_i$ for some $1 \leq j \leq n$. All such cycles belong to cycle kind of (2). Finally, the cycles that contains only vertices from $\cup_{i=1}^{n} U_i$ is of kind (4), by Property 5.

□

Lemma 5.3 *Given a positive integer $k \leq min\{m, n\}$, there is a covering $C \subseteq S$ of U with no more than k subsets, if and only if G has a multicolored cycle partition P with $k' + m \leq k + m$ multicolored cycles.*

Proof: Suppose G has a multicolored cycle partition P with $k' + m \leq k + m$ vertex-disjoint multicolored cycles. By Property 1, there must be m cycles of the first or third kind in P, each covering one s_{i0}, for $1 \leq i \leq m$. Let t and t' denote the number of multicolored cycles in P of the second kind and the fourth kind, respectively. Then $t + t' = k' \leq k$. Every multicolored cycle of the second kind contains a path $P(s_{i1}, s_{i(n+1)})$ for some $s_i \in S$; there are t such s_i's. For every multicolored cycle $CY(u_i)$ of the fourth kind, we can always find a subset s_j in S containing u_i, since $\cup_{j=1}^{m} s_j = U$; there are at most t' such s_j's. The subcollection of S composed of the s_i's and s_j's covers U, and has at most $t + t' \leq k' \leq k$ elements.

Conversely, let there be a covering C of U with $k' \leq k$ subsets. Without loss of generality, let $C = \{s_1, \ldots, s_k\}$. Let $s'_1 = s_1$, and $s'_i = s_i \backslash \cup_{j=1}^{i-1} s_j, i \geq 2$. Then $s'_1, \ldots, s'_k$ are k' disjoint subsets (some of which may be empty) whose union covers U. For every $1 \leq i \leq k'$, if $s'_i \neq \emptyset$, take a multicolored cycle of the second kind constituted by the path $P(s_{i1}, s_{i(n+1)})$ and the paths $P(u_j^1, u_j^2)$ for all $u_j \in s'_i$. Since $\cup_{i=1}^{k} s'_i$ covers U, all vertices of G in U_i, $1 \leq i \leq n$, are covered by these multicolored cycles. Finally take m multicolored cycles of the first kind $CY(s_i)$ or the third kind s_{i0}, covering all vertices of G that are left. Then a partition of at most $k' + m \leq k + m$ multicolored cycles for G is obtained. □

Since the Minimum Set Cover problem is NP-complete ([14]), by Lemma 5.3, we have that MCYP with restriction to triangle-free graphs is NP-complete.

6 Concluding Remarks

The monochromatic clique and multicolored cycle partition problems have many practical applications in information science and operations research. It is an important task of network optimization to understand the structural properties of a certain data set. In this paper we obtained results on the inapproximability and complexity of MCLP and MCYP restricted to graphs avoiding some induced subgraphs, and we presented a $(1 + \ln |V(G)|)$-approximation algorithm for WMCLP restricted to monochromatic-diamond-free graphs. A natural suggestion for further research is to consider possible approximation algorithms for MCYP or its weighted version on triangle-free graphs, and alternative algorithms that might improve the approximation ratio $(1 + \ln |V(G)|)$ for WMCLP on monochromatic-diamond-free graphs. Another interesting direction is to study the computational complexity of similar problems, e.g., on minimum monochromatic or multicolored path and tree partition problems, restricted to graphs avoiding some induced subgraphs.

Acknowledgements

The authors would like to thank the anonymous referees for their valuable suggestions and comments that have helped a lot to improve the quality and presentation of the paper.

References

[1] C. C. Aggarwal, H. Wang, Graph data management and mining: A survey of algorithms and applications, Managing and Mining Graph Data Advances in Database Systems **40**, 13-68, (2010).

[2] P. Alimonti and V. Kann, Some APX-completeness results for cubic graphs, Theor. Comput. Sci. **237**, 123–134 (2000).

[3] H. K. Anheier, J. Gerhards and F. P. Romo, Forms of capital and social structure in cultural fields: examining Bourdieu's social topography, American Journal of Sociology, **100**, 859–903, (1995).

[4] H. Broersma, X. Li, G. J. Woeginger and S. Zhang, Paths and Cycles in Colored Graphs, Australasian Journal on Combinatorics **31**, 299–311, (2005).

[5] M. R. Cerioli, L. Faria, T. O. Ferreira, C. A. J. Martinhon, F. Protti, and B. A. Reed, Partition into cliques for cubic graphs: Planar case, complexity and approximation, Discrete Applied Mathematics **156**, 2270–2278, (2008)

[6] L. S. Chandran and L. S. Ram, On the relationship between ATSP and the cycle cover problem, Theoretical Computer Science **370**, 218–228, (2007).

[7] V. Chvátal, A greedy heuristic for the set-covering problem, Mathematics of Operations Research, **4**, 233–235 (1979).

[8] Z. Chen, Y. Okamoto, and L. Wang, Improved deterministic approximation algorithms for Max TSP, Information Processing Letters, **95**, 333–342, (2005).

[9] Z. Chen and T. Nagoya, Improved approximation algorithms for metric MaxTSP, Journal of Combinatorial Optimization **13**, 321–336, (2007).

[10] B. Couëtoux, L. Gourvès, J. Monnot, and O. A. Telelis, On labeled traveling salesman problems, Chapter Algorithms and Computation, Lecture Notes in Computer Science **5369**, 776–787, (2008).

[11] M. Deng, Q. Liu, J. Wang, and Y. Shi, A general method of spatio-temporal clustering analysis, Science China Information Sciences **56**, 1-14, (2013).

[12] E. S. El-Mallah and C. J. Colbourn, The complexity of some edge deletion problems, IEEE Transactions on Circuits and Systems **35**, 354-362, (1988).

[13] L. Freeman, The Development of Social Network Analysis, Vancouver: Empirical Press, 2006.

[14] M. R. Garey and D. S. Johnson, Computers and intractability: A guide to the theory of NP-completeness, W. H. Freeman and co., New York, 1979.

[15] A. V. Goldberg, S. A. Plotkin, G. E. Shannon, Parallel symmetry-breaking in sparse graphs, SIAM Journal on Discrete Mathematics **1**: 434-446, (1988).

[16] R. Hassin, J. Monnot and D. Segev, Approximation algorithms and hardness results for labeled connectivity problems, J. Comb. Opt. **14**, 437–453, (2007).

[17] R. Hassin, J. Monnot and D. Segev, The complexity of bottleneck labeled graph problems, In Proc. of WG'07, Lecture Notes in Computer Science **4769**, Springer, (2007).

[18] Z. Jin, M. Kano, X. Li, and B. Wei, Partitioning 2-edge-colored complete multipartite graphs into monochromatic cycles, paths and trees, J. Comb. Optim. **11**, 445–454 (2006).

[19] H. Kaplan, M. Lewenstein, N. Shafrir and M. I. Sviridenko, Approximation algorithms for asymmetric TSP by decomposing directed regular multigraphs, Journal of the ACM, **52**, 602–626, (2005).

[20] M. Kano and X. Li, Monochromatic and heterochromatic subgraphs in edge-colored graphs - a survey, Graphs Combin. **24**, 237–263 (2008).

[21] S. O. Krumke and H. C. Wirth, On the Minimum Label Spanning Tree Problem, Inf. Process. Lett. **66**, 81–85, (1998).

[22] C. P. Kruskal, L. Rudolph and M. Snir, Efficient parallel algorithms for graph problems, Algorithmica **5**, 43–64, (1990).

[23] P. L. Lai and M. Y. Chiu, Uniform disjoint cycle covers on a hierarchical multicomputer system, Security-Enriched Urban Computing and Smart Grid, CCIS **223**, 141–148, (2011).

[24] X. Li and X. Y. Zhang, On the minimum monochromatic or multicolored subgraph partition problems, Theor. Comput. Sci. **385**, 1–10 (2007).

[25] B. Liu, Web data mining: exploring hyperlinks, Contents, and Usage Data, Springer, 2011.

[26] L. Lovász and M. D. Plummer, Matching theory, Elsevier, 1986.

[27] G. MacGillivray and M. Yu, Generalized partitions of graphs, Discrete Appl. Math. **91**, 143–153 (1999).

[28] B. Manthey, On approximating restricted cycle covers, SIAM J. Comput. **38**, 181–206, (2008).

[29] B. Manthey, Minimum-weight cycle covers and their approximability, Discrete Applied Mathematics **157**, 1470–1480, (2009).

[30] J. Monnot, The labeled perfect matching in bipartite graphs, Inf. Process. Lett. **96**, 81–88, (2005).

[31] T. Nishi, L. O. Chua, Topological proof of the Nielsen-Willson theorem, IEEE Transactions on Circuits and Systems **33**, 398–405, (1986).

[32] C. H. Papadimitriou, M. Yannakakis, Optimization, approximation, and complexity classes, J. Comput. System Sci. **43**, 425–440, (1991).

[33] B. Paten, M. Diekhans, D. Earl, J. St. John, J. Ma, B. Suh, D. Haussler, Cactus graphs for genome comparisons, Lecture Notes in Bioinformatics, **6044**, 410–425, (2010).

[34] S. Wasserman, K. Faust, Social network analysis: Methods and applications (structural analysis in the social sciences), Cambridge: Cambridge University Press, 1994.

[35] V. Yegnanarayanan, Graph colorings and partitions, Theor. Comput. Sci. **263**, 59–74 (2001).

Appendix A

This section contains an explanation to a problem in reference [24].

In Section 2 of the paper, the authors proved that the MCLP problem is NP-complete for diamond-free edge-colored graphs where each color is assigned to at most 3 edges. They designed a polynomial algorithm to approximate MCLP problem in a diamond-free graph G, within a performance ratio $\ln m+1$, where m is the size of the largest monochromatic clique in G.

However, the algorithm given actually works for monochromatic-diamond-free graphs, instead of diamond-free graphs.

The key problem lies in the proof of Lemma 2.2 in the paper. In the lemma, it is firstly proved that any two maximal cliques do not share an edge in a diamond-free graph. And it is assumed that when the diamond-free graph is edge-colored, two maximal monochromatic cliques do not share an edge. However, this does not always hold. For example, let H be an edge-colored K_4 with all edges colored red but one colored blue. Then H is diamond-free. But the two maximal monochromatic cliques in H share a common edge.

However, if the input graph G is monochromatic-diamond-free, we can use similar arguments as in the proof of Lemma 2.2 to prove that any two maximal monochromatic cliques in G do not share an edge. Furthermore, the two algorithms given in Section 2 and Theorem 2.3 are valid for monochromatic-diamond-free edge-colored graphs, instead of diamond-free edge-colored graphs.

It is not hard to see that in the proof of Theorem 2.1, when reducing an instance of the Exact Cover By 3-Sets problem to an instance of the MCLP problem, the graphs constructed are not only diamond-free but also monochromatic-diamond-free. Therefore, we can strengthen Theorem 2.1 as below.

Theorem A.1 *MCLP is NP-complete, even if the input graph G is diamond-free, monochromatic-diamond-free and each color is assigned to at most 3 edges.*

In the other parts of Section 2, we can replace all "diamond-free "with "monochromatic-diamond-free ", and change the proof of Lemma 2.2 accordingly. In particular, Lemma 2.2 becomes:

Theorem A.2 *The above algorithm can output a largest monochromatic clique for any monochromatic-diamond-free graph G in polynomial time.*

And the two algorithms given should work on monochromatic-diamond-free graphs to find an approximation solution of MCLP problem in polynomial time, with a performance ratio $\ln m + 1$, where m is the size of the largest monochromatic clique in the input graph G.