International Journal of Computer Mathematics
Vol. 00, No. 0, Month 2004, pp. 1–6

Taylor & Francis
Taylor & Francis Group

THE COMPLEXITY FOR PARTITIONING GRAPHS BY MONOCHROMATIC TREES, CYCLES AND PATHS

ZEMIN JIN and XUELIANG LI*

Research Center for Combinatorics, LPMC, Nankai University, Tianjin 300071, P.R. China

(Revised 31 October 2003; In final form 17 June 2004)

Let G be an edge-colored graph. We show in this paper that it is NP-hard to find the minimum number of vertex disjoint monochromatic trees which cover the vertices of the graph G. We also show that there is no constant factor approximation algorithm for the problem unless $P = \mathrm{NP}$. The same results hold for the problem of finding the minimum number of vertex disjoint monochromatic cycles (paths, respectively) which cover the vertices of the graph.

Keywords: Combinatorial problems; Computational complexity; Vertex disjoint; Tree, Cycle and path; Monochromatic

CR Categories: G.2.2; F.2.2

1 INTRODUCTION

Many combinatorial problems can be described as finding a partition of the vertices of a given graph into subsets satisfying certain properties. Many graph partition problems and their corresponding computational complexity problems have been well studied [1–5], most of which are shown to be NP-complete. MacGillivray and Yu [6] considered a general graph partition problem, the (H, C)-partition problem, which contains some well-known graph partition problems as special cases. Feder *et al.* [7] introduced a parameterized family of graph problems, which also includes several well-known graph partition, problems as special cases. A list of graph partition problems can be found in the Ref. [4].

Some researchers also focused on the graph partition problems in edge-colored graphs. The aim is to find some kind of vertex disjoint monochromatic subgraphs (*e.g.* trees, cycles or paths) to cover all the vertices of the given graph. Gyárfás [8] considered the problem of covering edge-colored graphs by vertex disjoint monochromatic paths and cycles. Erdös *et al.* [9] showed that if the edges of a finite complete graph are colored with r colors, then the vertex set of the graph can be covered by at most $cr^2 \log r$ vertex disjoint monochromatic cycles, where c is a constant. Haxell and Kohayakawa [10] proved that at most r vertex disjoint monochromatic trees are needed to cover the vertex set of the complete graph K_n for n large enough, if the edges of the finite complete graph are colored with r colors. Haxell [11]

* Corresponding author. E-mail: x.li@eyou.com

ISSN 0020-7160 print; ISSN 1029-0265 online
DOI: 10.1080/00207160412331290685

considered the analog problem in edge-colored complete bipartite graphs, whereas Kaneko *et al.* [12] considered the analog problem in two-edge-colored complete multipartite graphs.

Motivated by the results in Refs. [9–12], in this paper we consider the following optimal problems: Given an edge-colored graph G, find the minimum number of vertex disjoint monochromatic trees, cycles and paths, respectively, which cover the vertices of G. For convenience, we simply call the three problems the PGMT, PGMC and PGMP problems, respectively. By transforming the set cover problem to each of the problems in polynomial time, we show that all the three problems are NP-hard. And we show that there does not exist constant factor approximation algorithm for any of the three problems unless $P = \mathrm{NP}$.

We will see in Sections 2 and 3 that the previously defined problems contain several well-known problems as special cases. A more general graph partition problem is to find the minimum number of some kind of vertex disjoint monochromatic subgraphs which cover the vertices of G, which is a very hard problem in general. Some recent publications on the edge-colored graphs can be found in Refs. [13, 14].

2 MONOCHROMATIC TREES

In this section we focus on the PGMT problem. The corresponding decision version is defined formally as follows.

THE PGMT PROBLEM

INSTANCE: An edge-colored graph G and a positive integer k.
QUESTION: Are there k or less vertex disjoint monochromatic trees which cover the vertices of the graph G?

Note that the PGMT problem looks like the problem of partitioning a graph into induced forests [4]. But actually it is not the case. The following facts are easily seen. If G is colored properly, *i.e.* adjacent edges receive different colors, the PGMT problem is equivalent to the maximum matching problem, which can be solved in polynomial time [4, 5]. If G is colored with one color, the PGMT problem is equivalent to the spanning tree problem, and can be solved in polynomial time. However, in general, by transforming the set cover problem to the PGMT problem in polynomial time, we have the following result.

THEOREM 2.1 *The PGMT problem is* NP-*complete.*

Proof The problem is clearly in NP, since a nondeterministic algorithm needs only to guess a set of trees and to check in polynomial time that the trees in the set are vertex disjoint monochromatic ones and cover the vertices of the given graph.

Now we transform the set cover problem to the PGMT problem. Let an arbitrary instance of the set cover problem be given by the set $S = \{v_1, v_2, \dots, v_n\}$ and subsets $C_1, C_2, \dots, C_m$ of S. Here we construct an edge-colored graph G such that there are k or less subsets of S which cover all the elements of S if and only if G contains $k + 1$ or less vertex disjoint monochromatic trees which cover the vertices of G.

The graph G is constructed as follows. The vertex set of G is $V(G) = \{v_1, v_2, \dots, v_n, C_1, C_2, \dots, C_m, v\}$. The edge set of G consists of the edges vC_i, $1 \le i \le m$, and the edges v_iC_j if and only if $v_i \in C_j$, $1 \le i \le n$, $1 \le j \le m$. Color the edges of G by $m + 1$ colors $c_0, c_1, c_2, \dots, c_m$ as follows. At first, color all the edges vC_i, $1 \le i \le m$, by the same color c_0. Next, color all other edges incident to the vertex C_j with the color c_j. It is easy to show that the construction can be accomplished in polynomial time. We claim that there are k or less subsets of S which cover all the elements of S if and only if G contains $k+1$ or less vertex disjoint monochromatic trees which cover the vertices of G.

If there are k subsets $C_{i_1}, C_{i_2}, \ldots, C_{i_k}$ of S which cover all the elements of S, it is easy to find that G contains $k+1$ or less vertex disjoint monochromatic trees which cover the vertices of G.

Suppose that G contains $k+1$ vertex disjoint monochromatic trees, denoted by $\Gamma = \{T_1, T_2, \ldots, T_{k+1}\}$, which cover the vertices of G. If $k \geq m$, the claim is true, since $\bigcup_{i=1}^{i=m} C_i = S$. So, we assume $k < m$.

Let T_i, $i = 1, 2, \ldots, t$, $t \leq k < m$, be the trees, each of which contains an edge $v_p C_q$ for some $1 \leq p \leq n$ and $1 \leq q \leq m$. Without loss of generality, we can assume that $C_i \in T_i$, $i = 1, 2, \ldots, t$. If for any $1 \leq j \leq k+1$, T_j is not composed of a single vertex v_r for some $1 \leq r \leq n$, then it is easy to see that C_i, $i = 1, 2, \ldots, t$ cover the set S. Suppose that T_{t+i}, $i = 1, 2, \ldots, t'$, is composed a single vertex. Without loss of generality, we can assume that $T_{t+i} = v_i$, $i = 1, 2, \ldots, t'$. Then $t + t' \leq k < m$, since v cannot lie in any T_j for $1 \leq j \leq t + t'$. For each v_i, $1 \leq i \leq t'$, find a subset C_{p_i} such that $v_i \in C_{p_i}$. It is easy to see that the subsets $C_1, C_2, \ldots, C_t, C_{p_i}$, $i = 1, 2, \ldots, t'$ form a cover of S. This completes the proof. ■

From the transformation described in the proof of Theorem 2.1, we see that the graph G constructed in the above-mentioned proof is bipartite. Therefore, we have the following result as a corollary.

COROLLARY 2.2 *The PGMT problem remains to be NP-complete for edge-colored bipartite graphs.*

Similarly, we can show that to find a good approximate optimal solution for the PGMT problem is also an NP-hard problem.

THEOREM 2.3 *There is no constant factor approximation algorithm for the PGMT problem unless $P =$ NP.*

Proof Assume that there exists an approximation algorithm Π of constant factor for the PGMT problem. Let α be the exact constant. Denote by OPT_I the optimal solution of the instance I.

Let I be an arbitrary instance of the set cover problem given by a set S and its subsets $C_1, C_2, \ldots, C_m$. Without loss of generality, we can assume that $S = \bigcup_{i=1}^{i=m} C_i$. As in the proof of Theorem 2.1, we can construct an instance I' of the PGMT problem in polynomial time, where I' is given by the edge-colored graph G. And we have that there are k or less given subsets of S which cover all the elements of S if and only if G contains $k+1$ or less vertex disjoint monochromatic trees which cover the vertices of G. Then $\mathrm{OPT}_{I'} = \mathrm{OPT}_I + 1$. When running the algorithm Π on I', Π finds a solution $s' \leq \alpha \cdot \mathrm{OPT}_{I'} \leq 2\alpha \mathrm{OPT}_I$. Thus, we obtain a 2α-approximation algorithm for the instance I. This contradicts to the fact that there is no constant factor approximation algorithm for the set cover problem unless $P =$ NP [16, 17]. ■

3 MONOCHROMATIC CYCLES AND PATHS

In this section we focus on the PGMC and PGMP problems. The corresponding decision versions are defined formally as follows.

THE PGMT (PGMP) PROBLEM

INSTANCE: An edge-colored graph G and a positive integer k.
QUESTION: Are there k or less vertex disjoint monochromatic cycles (paths) which cover the vertices of the graph G?

If G is colored with only one color, the PGMC problem and the PGMP problem is equivalent to the problem of finding minimum number of vertex disjoint cycles and paths, respectively, to cover the vertices of G. Moreover, by setting $k = 1$, the PGMC problem and the PGMP problem is equivalent to the Hamiltonian cycle problem and the Hamiltonian path problem, respectively, which are NP-hard problems. If G is colored properly, *i.e.* adjacent edges receive different colors, the PGMP problem is equivalent to the maximum matching problem, which can be solved in polynomial time [4, 15]. However, in general, by transforming the set cover problem to each of the two problems in polynomial time, we have the following results.

THEOREM 3.1 *The PGMC problem is NP-complete.*

Proof The problem is clearly in NP, since a nondeterministic algorithm needs only to guess a set of cycles and to check in polynomial time that the cycles in the set are vertex disjoint monochromatic ones and cover the vertices of the given graph.

Now we transform the set cover problem to the PGMC problem. Let an arbitrary instance of the set cover problem be given by the set $S = \{v_1, v_2, \ldots, v_n\}$ and its subsets $C_1, C_2, \ldots, C_m$. Here we construct an edge-colored multiple graph G such that there are k or less subsets of S which cover all the elements of S if and only if G contains $k + 1$ or less vertex disjoint monochromatic cycles which cover the vertices of G.

Construct the multiple graph G and color its edges by $m + 1$ colors $c_1, c_2, \ldots, c_m, c$ as follows. The vertex set of G is $V(G) = \{v_1, v_2, \ldots, v_n, v_1^*, v_2^*, \ldots, v_n^*, v_1^{**}, v_2^{**}, \ldots, v_n^{**}, C_1, C_2, \ldots, C_m, u, v, w\}$. The edges of G and their colors are defined as follows.

(1) If $v_i \in C_j$, $1 \le i \le n$, $1 \le j \le m$, then each of the vertices v_i, v_i^* and v_i^{**} is connected to C_j, and color the three edges by the color c_j.
(2) If $v_i \in C_j$, construct a triangle on the three vertices v_i, v_i^*, and v_i^{**}, and color the three edges with the color c_j. Note that if the element v_i also belongs to another subset C_r, then we double the edges of the triangle on the three vertices v_i, v_i^*, and v_i^{**}, and color its edges by the color c_r. If v_i belongs to more subsets, we simply construct more copies of triangles with different colors according to the subsets. So, in general what we constructed is a multiple graph.
(3) If $v_i, v_r \in C_j$, connect each of the vertices v_i, v_i^* and v_i^{**} to each of the vertices v_r, v_r^* and v_r^{**}, and color the edges with the color c_j.
(4) The vertices $\{C_1, C_2, \ldots, C_m, u, v, w\}$ form a complete graph and color all its edges by a same color c.

It is easy to show that the construction can be accomplished in polynomial time. We claim that there are k or less subsets of S, which cover all the elements of S if and only if G contains $k + 1$ or less vertex disjoint monochromatic cycles which cover the vertices of G.

If there are k subsets $C_{i_1}, C_{i_2}, \ldots, C_{i_k}$ of S, which cover all the elements of S, it is easy to find that G contains $k + 1$ or less vertex disjoint monochromatic cycles which cover the vertices of G.

Suppose that G contains $k + 1$ vertex disjoint monochromatic cycles, denoted by $\Gamma = \{T_1, T_2, \ldots, T_{k+1}\}$, which cover the vertices of G. If $k \ge m$, the claim is true, since $S = \bigcup_{i=1}^{i=m} C_i$. So, we assume $k < m$.

Let T_i, $i = 1, 2, \ldots, t$, be the cycles containing some edges $v_p C_q$, $v_p^* C_q$ or $v_p^{**} C_q$. From the construction we have that none of u, v and w can lie in any T_i, $i = 1, 2, \ldots, t$ and each T_i, $i = 1, 2, \ldots, t$, contains a unique vertex of some C_q. Without loss of generality, we can assume that $C_i \in T_i$, $i = 1, 2, \ldots, t$. If T_j contains no vertex of $\{v_1, v_2, \ldots, v_n\}$ for any $t + 1 \leq j \leq k + 1$, then it is easy to see that C_is, $i = 1, 2, \ldots, t$ cover the set S. So, assume that T_{t+i}, $i = 1, 2, \ldots, t'$, are the cycles, each of which only contains vertices of $\{v_1, v_2, \ldots, v_n, v_1^*, v_2^*, \ldots, v_n^*, v_1^{**}, v_2^{**}, \ldots, v_n^{**}\}$. It is easy to see that $t + t' \leq k$, since none of u, v and w can lie in any T_{t+i}, $i = 1, 2, \ldots, t'$. Consider the cycle T_{t+i}, $i = 1, 2, \ldots, t'$. Denote the color of the cycle T_{t+i} by c_{r_i}, $i = 1, 2, \ldots, t'$. Let $I_i = \{j | T_{t+i} \cap \{v_j, v_j^*, v_j^{**}\} \neq \emptyset\}$. Then $\{v_j | j \in I_i\} \subseteq C_{r_i}$. It is easy to see that the subsets $C_1, C_2, \ldots, C_t, C_{r_i}, i = 1, 2, \ldots, t'$ form a cover of S. This completes the proof. ■

Similarly, we can show that to find a good approximate optimal solution for the PGMC problem is also an NP-hard problem.

THEOREM 3.2 *There is no constant factor approximation algorithm for the PGMC problem unless $P = NP$.*

Proof Assume that there exists an approximation algorithm Π of constant factor for the PGMC problem. Let α be the exact constant. By similar analysis as in the proof of Theorem 2.3, we obtain a 2α-approximation algorithm for the set cover problem, which also contradicts to the fact that there is no constant factor approximation algorithm for the set cover problem unless $P =$ NP [16, 17] ■

Though the PGMP problem is equivalent to the Hamiltonian path problem if G is colored with only one color and $k = 1$, using the analog technique in the proof of Theorem 3.1, we can show that it remains to be NP-complete if $k \geq 2$. We have the following result and the detailed proof is omitted.

THEOREM 3.3 *The PGMP problem is NP-complete.*

Similarly, we can show that to find a good approximate optimal solution for the PGMP problem is also an NP-hard problem. We have the following result and the detailed proof is omitted.

THEOREM 3.4 *There is no constant factor approximation algorithm for the PGMP problem unless $P = NP$.*

One can think about other kind of monochromatic subgraphs in a edge-colored graph G to cover the vertices of G. For example, the same proof as for Theorem 2.1 can be employed to show that to find the minimum number of vertex disjoint monochromatic stars to cover the vertices of G is NP-hard, and there does not exist any constant factor approximation algorithm for the problem.

4 CONCLUSIONS

In this paper, we show that to find the minimum number of vertex disjoint monochromatic trees, cycles and paths, respectively, in an edge-colored graph G which covers the vertices of G is NP-hard. We also show that there does not exist constant factor approximation algorithm for any of the problems unless $P = \mathrm{NP}$. A more general graph partition problem is to find the minimum number of some kind of vertex disjoint monochromatic subgraphs in an edge-colored graph G, which cover the vertices of G. We believe that most of them are NP-hard. Though there is no constant factor approximation algorithm for any of the PGMT, PGMC and PGMP problems, to find heuristic procedures with fixed performance guarantee is an interesting problem for further research. For some special edge-colorings or some special graphs, for example, G is colored by 2 or c colors where c is a constant, or each color class has at most 2 colors, or G is complete or multipartite complete, the computational complexity of the problems remains open.

Acknowledgements

This Research was supported by NSFC.

References

Q1 [1] Brandstadt, A. (1996). Partitions of graphs into one or two independent sets and cliques. *Discrete Math.*, **152**, 47–54.
[2] Enomoto, H. (2001). Graph partition problems into cycles and paths. *Discrete Math.*, **233**, 93–101.
[3] Feder, T. and Motwani, R. (1995). Clique partitions, graph compression and speeding-up algorithms. *J. Comput. Syst. Sci.*, **51**, 261–272.
[4] Garey, M. R. and Johnson, D. S. (1979). *Computers and Intractability*. W.H. Freeman, San Francisco.
[5] Holyer, L. (1981). The NP-compleeteness of some edge-partition problems. *SIAM J. Comput.*, **10**, 713–717.
[6] MacGillivray, G. and Yu, M. L. (1999). Generalized partitions of graphs. *Discrete Appl. Math.*, **91**, 143–153.
[7] Feder, T., Hell, P., Klein, S. and Motwani, R. Complexity of graph partition problems. *Proceedings of Thirty-First Annual ACM Symposium on Theory of Computing, STOC'99*, 464–472.
[8] Gyárfás, A. (1983). Vertex coverings by monochromatic paths and cycles. *J. Graph Theory*, **7**, 131–135.
[9] Erdös, P., Gyárfás, A. and Pyber, L. (1991). Vertex coverings by monochromatic cycles and trees. *J. Comb. Theory Ser. B*, **51**, 90–95.
[10] Haxell, P. E. and Kohayakawa, Y. (1996). Partitioning by monochromatic trees. *J. Comb. Theory Ser. B*, **68**, 218–222.
[11] Haxell, P. E. (1997). Partitioning complete bipartite graphs by monochromatic cycles. *J. Comb. Theory Ser. B*, **69**, 210–218.
[12] Kaneko, A., Kano, M. and Suzuki, K. Partitioning complete multipartite graphs by monochromatic trees, personal communication.
[13] Broersma, H. J. and Li, X. (1997). Spanning trees with many or few colors in edge-colored graph. *Discuss. Math. Graph Theory*, **17**(2), 259–269.
[14] Chang, R. and Leu, S. (1997). The minimum labeling spanning trees. *Info. Process. Lett.*, **63**, 277–282.
[15] Lawler, E. L. (1976). *Combinatorial Optimization: Networks and Matroids*, Holt, Rinehart and Winston, New York.
[16] Feige, U. (1998). A threshold of ln n for approximating set cover. *J. ACM*, **45**, 634–652.
[17] Hochbaum, D. S. (Ed.) (1997). *Approximation Algorithm for NP-Hard Problems*. PWS Publishing Company, Boston, MA.

Journal ...**International Journal of Computer Mathematics** Article ID ...**GCOM 041114**

TO: CORRESPONDING AUTHOR

AUTHOR QUERIES - TO BE ANSWERED BY THE AUTHOR

The following queries have arisen during the typesetting of your manuscript. Please answer the queries.

Q1	References have been renumbered. Please check the list and citations in the text.	

Production Editorial Department, Taylor & Francis Ltd.
4 Park Square, Milton Park, Abingdon OX14 4RN

Telephone: +44 (0) 1235 828600
Facsimile: +44 (0) 1235 829000