

Paths and Cycles in Colored Graphs*

Hajo Broersma¹, Xueliang Li², Gerhard Woeginger¹, Shenggui Zhang³

¹Department of Applied Mathematics, University of Twente,
P. O. Box 217, 7500 AE Enschede, The Netherlands

²Center for Combinatorics, Nankai University,
Tianjin 300071, P.R. China

³Department of Applied Mathematics, Northwestern Polytechnical University,
Xi'an, Shaanxi 710072, P.R. China

Abstract

Let G be an (edge-)colored graph. A path (cycle) is called monochromatic if all of its edges have the same color, and is called heterochromatic if all of its edges have different colors. In this paper, some sufficient conditions for the existence of (long) monochromatic paths and cycles, and those for the existences of long heterochromatic paths and cycles are obtained. It is proved that the problem of finding a path (cycle) with as few different colors as possible in a colored graph is **NP**-hard. Several exact and approximation algorithms for finding a path with the fewest colors are provided. The complexity of the exact algorithms and the performance ratio of the approximation algorithms are analyzed. We also pose a problem on the existence of paths and cycles with many different colors.

Keywords: (edge-) colored graph, monochromatic (heterochromatic) path (cycle), (exact) approximation algorithm

AMS Subject Classification (2000): 05C85 05C38 05C35 05C15 68Q25 68W25

1 Introduction

We use Bondy & Murty [2] for terminology and notations not defined here and consider simple graphs only.

Let $G = (V, E)$ be a graph. By an *edge-coloring* of G we will mean a function $C : E \rightarrow \mathbb{N}$, the set of nonnegative integers. If G is assigned such a coloring, then we say that G is a *colored graph*, denote the colored graph by (G, C) , and call $C(e)$ the *color* of the edge $e \in E$. All edges with the same color form a *color class* of the graph. We note that C is not necessarily a proper edge-coloring, i.e., two adjacent edges may have the same color. For a subgraph H of G , we let $C(H) = \cup_{e \in E(H)} \{C(e)\}$ and $c(H) = |C(H)|$. For a vertex v of G , the *color neighborhood* $CN(v)$ of v is defined as the set $\{C(e) : e \text{ is incident with } v\}$ and the *color degree* $d^c(v) = |CN(v)|$. A path (cycle) is called *monochromatic* if all of its edges have the same color; and it is called *heterochromatic* if all of its edges have different colors.

*Supported by NSFC.

If we regard an uncolored graph as a colored graph in which all edges have different colors, then the number of colors of a subgraph is simply the number of its edges, and the color degree of a vertex is the degree of it.

We are not aware of any papers dealing with the existence of paths and cycles with special properties in general colored graphs. All existing results we could find deal with colored complete graphs. In [9] Giraud studied the existence of monochromatic triangles and heterochromatic triangles in colored complete graph. A problem on the conditions for a colored complete graph to contain heterochromatic Hamilton cycles was mentioned in [6] by Erdős, Nešetřil & Rödl. This problem was studied by Hahn & Thomassen [10], Rödl & Winkler (see [7]), Frieze & Reed [7], and Albert, Frieze & Reed [1]. Most of the results in these papers are proved by using probabilistic methods.

This paper contains some basic results on paths and cycles in general colored graphs. In Sections 2 and 3, we give some sufficient conditions for the existence of (long) monochromatic paths and cycles, and those for the existence of long heterochromatic paths and cycles. In Section 4, we prove that the problem of finding a path (cycle) with as few different colors as possible between two given vertices in a colored graph is **NP**-hard and propose two exact algorithms and two approximation algorithms for finding a path with the fewest colors. The complexity of the two exact algorithms and the performance ratio of the approximation algorithms are analyzed. We also pose a problem on the existence of paths and cycles with many different colors in Section 5.

2 Monochromatic paths and cycles

First, let us consider the problem under what conditions a colored graph contains a monochromatic path or a monochromatic cycle. It is clear that every colored graph contains at least one monochromatic path. Moreover, it is obvious that not every colored graph contains monochromatic cycles.

The *arboricity* $a(G)$ of a graph G is defined as the minimum number of edge-disjoint forests into which G can be decomposed. Clearly, it is also the minimum number of colors necessary to color the edges of G so that no cycle is monochromatic. So we have

Proposition 1 *Let G be a colored graph. If $c(G) < a(G)$, then G contains at least one monochromatic cycle.* ■

The arboricity $a(G)$ can be determined by applying the matroid partitioning algorithm of Edmonds [11]. In [12] Picard & Queyranne showed that this parameter can be determined in at most $O(n^4)$ operations, by using network flow methods. It is (almost) trivial to check whether a colored graph contains a monochromatic cycle: for each color class E_i check whether the induced subgraph $G[E_i]$ contains a cycle.

The following result on the existence of monochromatic paths and cycles with a prescribed length is obvious.

Proposition 2 *Let G be a colored graph with color classes E_1, E_2, \dots, E_c . Then G has a monochromatic path (cycle) of length at least l if and only if for some i with $1 \leq i \leq c$, the induced subgraph $G[E_i]$ has a path (cycle) of length at least l .* ■

If we regard an uncolored graph G as a colored graph (G, C) for which all edges have the same color, then (G, C) contains a monochromatic path (cycle) of length at least l if and only if G contains a path (cycle) of length at least l . Since the problem of

deciding whether there is a path (cycle) of length at least l in an (uncolored) graph is **NP**-complete, the problem of deciding whether there is a monochromatic path (cycle) of length at least l in a colored graph is also **NP**-complete.

There are many results on the existence of long paths and cycles in (uncolored) graphs. Here we list two of them.

Theorem A (Erdős & Gallai [5]) *Let G be a graph of order n and size m . Then G contains a path of length at least $\frac{2m}{n}$.*

Theorem B (Erdős & Gallai [5]) *Let G be a graph of order n and size m such that $m \geq n$. Then G contains a cycle of length at least $\frac{2m}{n-1}$.*

Using Theorems 2, A and B, it is not difficult to prove the following results:

Proposition 3 *Suppose that G is a colored graph of order n and size m . Then G contains a monochromatic path of length at least $\frac{2m}{c(G)n}$. ■*

Proposition 4 *Suppose that G is a colored graph of order n and size m such that $m \geq c(G)n$. Then G contains a monochromatic cycle of length at least $\frac{2m}{c(G)(n-1)}$. ■*

As it was shown in [5], Theorem A is best possible. Let pK_r denote the disjoint union of p copies of K_r . This graph has $n = pr$ vertices and $m = \frac{pr(r-1)}{2}$ edges. It is easy to check that $\frac{2m}{n} = r-1$. On the other hand, pK_r contains no path of length greater than $r-1$. Of course, this graph also shows that the result in Proposition 3 in the case $c(G) = 1$ is best possible. This example can be extended to general cases to show the sharpness of the result in Proposition 3.

Let G and H be two colored graphs. The *colored Cartesian product* of G and H is the graph $G \times H$ with a coloring defined as follows: From the definition of the Cartesian product of graphs, to every vertex u of G , there corresponds a subgraph H_u of $G \times H$ such that H_u is isomorphic to H . To each edge e of H_u , assign the color of the edge corresponding to e in H . Similarly, to every vertex v of H , there corresponds a subgraph G_v of $G \times H$ such that G_v is isomorphic to G . To each edge e of G_v , assign the color of the edge corresponding to e in G . The colored Cartesian product $G_1 \times G_2 \times \cdots \times G_k$ of $k \geq 2$ colored graphs G_1, G_2, \dots, G_k can be defined inductively.

Let G_i ($1 \leq i \leq c$) be the colored graph K_r such that all the edges of G_i receive the same color i . By K_r^c we denote the colored Cartesian product $G_1 \times G_2 \times \cdots \times G_c$. It is not difficult to see that the colored graph K_r^c has $n = r^c$ vertices, $m = \frac{cr^c(r-1)}{2}$ edges and c colors. This implies that $\frac{2m}{cn} = r-1$. On the other hand, the colored graph K_r^c has no monochromatic path of length greater than $r-1$. This shows that the result in Proposition 3 is best possible. Clearly the disjoint union of some copies of the colored graph K_r^c defined above can also be used to show the sharpness of the result of Proposition 3.

Theorem B is also best possible. This can be shown by the graph $\Gamma_{p,r}$ defined as follows: The graph $\Gamma_{p,r}$ is a connected graph which has exactly $n = p(r-1) + 1$ vertices

and each of the p blocks of it is a clique on r vertices. This graph has $m = \frac{pr(r-1)}{2}$ edges and clearly $\frac{2m}{n-1} = r$. On the other hand, it has no cycle of length greater than r . Of course this example also shows that the result of Proposition 4 is best possible in the case $c(G) = 1$.

Let G_{p_i} ($1 \leq i \leq c$) be the colored graph $\Gamma_{p_i, r}$ such that all the edges of it receive the same color i . Denote by G the colored Cartesian product $G_{p_1} \times G_{p_2} \times \cdots \times G_{p_c}$. Then G has

$$n = \sum_{i=1}^c \sum_{1 \leq j_1 < j_2 < \cdots < j_i \leq c} p_{j_1} p_{j_2} \cdots p_{j_i} (r-1)^i + 1$$

vertices,

$$m = \frac{r(r-1)}{2} \sum_{i=1}^c i \sum_{1 \leq j_1 < j_2 < \cdots < j_i \leq c} p_{j_1} p_{j_2} \cdots p_{j_i} (r-1)^{i-1}$$

edges, and c colors. Therefore,

$$\begin{aligned} \frac{2m}{c(n-1)} &= \frac{2 \times \frac{r(r-1)}{2} \sum_{i=1}^c i \sum_{1 \leq j_1 < j_2 < \cdots < j_i \leq c} p_{j_1} p_{j_2} \cdots p_{j_i} (r-1)^{i-1}}{c \left(\sum_{i=1}^c \sum_{1 \leq j_1 < j_2 < \cdots < j_i \leq c} p_{j_1} p_{j_2} \cdots p_{j_i} (r-1)^i + 1 - 1 \right)} \\ &= r - \frac{\sum_{i=1}^{c-1} (c-i) \sum_{1 \leq j_1 < j_2 < \cdots < j_i \leq c} p_{j_1} p_{j_2} \cdots p_{j_i} (r-1)^{i-1}}{c \sum_{i=1}^c \sum_{1 \leq j_1 < j_2 < \cdots < j_i \leq c} p_{j_1} p_{j_2} \cdots p_{j_i} (r-1)^{i-1}}. \end{aligned}$$

It is clear that $\lceil \frac{2m}{c(n-1)} \rceil = r$. On the other hand, the colored graph G contains no monochromatic cycle of length greater than r . This shows that the result in Proposition 4 is best possible.

3 Heterochromatic paths and cycles

If we regard an uncolored graph G as a colored graph (G, C) in which all edges have different colors, then G contains a path (cycle) of length at least l if and only if (G, C) contains a heterochromatic path (cycle) of length at least l . As we mentioned earlier, the problem of deciding whether there is a path (cycle) of length at least l in an (uncolored) graph is **NP**-complete. Therefore the problem of deciding whether there is a heterochromatic path (cycle) of length at least l in a colored graph is **NP**-complete, too. In this section we will consider under what conditions there is a heterochromatic path (cycle) with a prescribed length in a colored graph.

Let G be a colored graph. By selecting precisely one edge from each color class of G , we obtain a new colored graph G' , such that all the edges of G' have different colors, and $c(G') = c(G)$. Using Theorems A and B, it is easy to prove the following results.

Proposition 5 *Let G be a colored graph of order n . Then G contains a heterochromatic path of length at least $\frac{2c(G)}{n}$. ■*

Proposition 6 *Let G be a colored graph of order n such that $c(G) \geq n$. Then G contains a heterochromatic cycle of length at least $\frac{2c(G)}{n-1}$. ■*

Clearly Propositions 5 and 6 generalize Theorems A and B, respectively.

Furthermore, we have the following two results on the existence of long heterochromatic paths.

Proposition 7 *Let G be a colored graph and k an integer. Suppose that $d^c(v) \geq k$ for every vertex v of G . Then for every vertex z of G there exists a heterochromatic z -path of length at least $\lceil \frac{k+1}{2} \rceil$.*

Proof Choose a longest heterochromatic z -path P with length l . Denote the other end-vertex of P as v . Then from the assumption of the proposition, we know that all incident edges of v with the other end not on P have colors also appearing in $E(P)$. Therefore, $d^c(v) \leq l + (l-1) = 2l-1$. On the other hand, $d^c(v) \geq k$, so we have $l \geq \lceil \frac{k+1}{2} \rceil$. ■

Proposition 8 *Let G be a colored graph and s an integer. Suppose that $|CN(u) \cup CN(v)| \geq s > 1$ for every pair of vertices u and v of G . Then G contains a heterochromatic path of length at least $\lceil \frac{s}{3} \rceil + 1$.*

Proof Choose a longest heterochromatic path P with length l . Denote the end-vertices of P as u and v . Then from the assumption of the proposition, we know that all incident edges of u and v with the other end not on P have colors also appearing in $E(P)$. Therefore, $|CN(u) \cup CN(v)| \leq l + (l-1) + (l-2) = 3l-3$. On the other hand, $|CN(u) \cup CN(v)| \geq s$, so we have $l \geq \lceil \frac{s}{3} \rceil + 1$. ■

In the following, we give a sufficient condition for the existence of heterochromatic triangles or quadrilaterals.

Proposition 9 *Let G be a colored graph of order $n \geq 4$, such that $|CN(u) \cup CN(v)| \geq n-1$ for every pair of vertices u and v of G . Then G contains at least one heterochromatic triangle or one heterochromatic quadrilateral.*

Proof If $|CN(u)| = n-1$ for every vertex u of G , then $d(u) = n-1$, and G is a complete graph. It is clear that every triangle of G is heterochromatic. So, we need only consider the case that there is some vertex $u \in V(G)$ with $|CN(u)| < n-1$.

Suppose that G contains neither heterochromatic triangles nor heterochromatic quadrilaterals. Without loss of generality, we can assume that $V(G) = \{x_1, x_2, \dots, x_k, u, v, y_{k+2}, y_{k+3}, \dots, y_{n-2}, y_{n-1}\}$, $d^c(u) = k+1 < n-1$, $C(ux_i) = i$ for $i = 1, 2, \dots, k$, $C(uv) = k+1$ and $C(vy_j) = j$ for $j = k+2, \dots, n-1$.

First, consider the vertex u and a vertex $x_i \in \{x_1, x_2, \dots, x_k\}$. Since $CN(v) \cup (CN(x_i) \cap \{C(x_ix_1), C(x_ix_2), \dots, C(x_ix_{i-1}), C(x_ix_{i+1}), \dots, C(x_ix_k), C(x_iu)\}) \subseteq \{1, 2, \dots, k+1\}$ and $|CN(u) \cup CN(x_i)| \geq n-1$, x_i must be adjacent to each vertex $y_j \in \{y_{k+2}, y_{k+3}, \dots, y_{n-1}\}$, and $C(x_iy_j) = j$ by our assumption that G contains neither heterochromatic triangles nor heterochromatic quadrilaterals.

Now consider the two vertices u and y_{n-1} . Since $CN(u) \cup (CN(y_{n-1}) \cap \{C(y_{n-1}x_1), C(y_{n-1}x_2), \dots, C(y_{n-1}x_k)\}) \subseteq \{1, 2, \dots, k+1, n-1\}$ and $|CN(u) \cup CN(y_{n-1})| \geq n-1$, we have that $y_{n-1}y_j \in E(G)$ and $C(y_{n-1}y_j) = j$ for $j = k+2, \dots, n-2$ by our assumption that G contains no heterochromatic triangles.

So, we have $CN(u) \cup (CN(y_{n-2}) \cap \{C(y_{n-2}x_1), C(y_{n-2}x_2), \dots, C(y_{n-2}x_k), C(y_{n-2}u), C(y_{n-2}v), C(y_{n-2}y_{n-1})\}) \subseteq \{1, 2, \dots, k, k+1, n-2\}$. Therefore, $|CN(u) \cup CN(y_{n-2})| \leq |\{1, 2, \dots, k, k+1, n-2\}| + |\{y_{k+2}, y_{n+3}, \dots, y_{n-3}\}| = (k+2) + (n-k-4) = n-2 < n-1$, a contradiction.

The proof of the result is complete. ■

Although the proofs of the results in Propositions 7 to 9 are easy, it can be shown that these results are best possible in the sense that there exist some graphs showing that they cannot be improved. However, we think that perhaps much stronger results are possible to obtain if one excludes some small counter-examples or simple classes of counter-examples. The proof techniques we applied here do not seem to be strong enough for obtaining such improvements. Maybe an approach using probabilistic proof techniques could yield such improvements.

4 Paths and cycles with few colors

If we regard an (uncolored) graph G as a colored graph (G, C) for which all edges have different colors, then a shortest path between two given vertices in G is a path between the two vertices with the fewest colors in the colored graph (G, C) . It is well-known that the problem of finding a shortest path between two given vertices in a (weighted) graph can be solved efficiently. There are many polynomial-time algorithms to solve this problem. In this section, we will consider the complexity aspects of finding a path between two given vertices with the fewest colors in a colored graph.

Problem 10

INSTANCE: Graph $G = (V, E)$ with a coloring $C : E \rightarrow \mathbb{N}$ and two given vertices s_0 and t_0 , positive integer $K \leq c(G)$.

QUESTION: Is there a path P from s_0 to t_0 such that $c(P) \leq K$?

According to [8], the following problem is **NP**-complete.

3-SATISFIABILITY (3-SAT)

INSTANCE: Collection $C = \{C_1, C_2, \dots, C_m\}$ of clauses on a finite set U of variables such that $|C_i| = 3$ for $1 \leq i \leq m$.

QUESTION: Is there a truth assignment for U that satisfies all the clauses in C ?

In this section, we use this result to show that Problem 10 is **NP**-complete, too.

Theorem 11 *Problem 10 is NP-complete.*

Proof It is easy to see that Problem 10 is in **NP**. One way to see this is to observe that a nondeterministic algorithm need only guess an (s_0, t_0) -path P in G , and check in linear time whether $c(P) \leq K$.

We shall now show that **3-SAT** can be polynomially transformed to Problem 10. Given a Boolean formula F consisting of m clauses C_1, C_2, \dots, C_m (with three literals per clause) and involving n variables x_1, x_2, \dots, x_n , we shall construct a graph $G =$

(V, E) with a coloring $C : E \rightarrow \mathbb{N}$ and two vertices s_0 and t_0 , such that G has an (s_0, t_0) -path P with $c(P) \leq n + 1$ if and only if F is satisfiable.

First, for the variable x_i ($1 \leq i \leq n$), we construct a subgraph A_i of G , where $V(A_i) = \{s_{i-1}, u_{i1}, u_{i2}, s_i\}$ and $E(A_i) = \{s_{i-1}u_{i1}, u_{i1}s_i, s_{i-1}u_{i2}, u_{i2}s_i\}$. Assign a special color 0 to the edges $s_{i-1}u_{i1}$ and $s_{i-1}u_{i2}$, the color i to the edge $u_{i1}s_i$ and the color i' to the edge $u_{i2}s_i$ for $i = 1, 2, \dots, n$. Then we get a colored graph $A = \cup_{i=1}^n A_i$. For the clause C_j ($1 \leq j \leq m$), we construct a subgraph B_j of G , where $V(B_j) = \{t_{j-1}, v_{j1}, v_{j2}, v_{j3}, t_j\}$ and $E(B_j) = \{t_{j-1}v_{j1}, t_{j-1}v_{j2}, t_{j-1}v_{j3}, v_{j1}t_j, v_{j2}t_j, v_{j3}t_j\}$. For $j = 1, 2, \dots, m$ and $k = 1, 2, 3$, assign the color 0 to the edge $t_{j-1}v_{jk}$, the color h to the edge $v_{jk}t_j$ if the k th literal of C_j is x_h and the color h' to the edge $v_{jk}t_j$ if the k th literal of C_j is \bar{x}_h . Then we get a colored graph $B = \cup_{j=1}^m B_j$. The colored graph G is obtained by connecting the two graphs A and B with an edge $s_n t_m$ and coloring this edge with the color 0, see Figure 1. Clearly the construction of G can be accomplished in polynomial time.

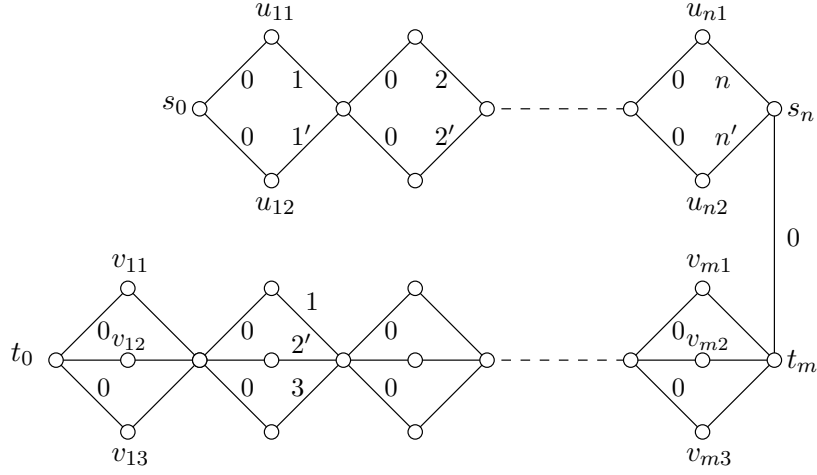


Figure 1: The graph in the proof of Theorem 11 in the case $C_2 = x_1 \bar{x}_2 x_3$.

It is not difficult to verify that there is an (s_0, t_0) -path P with $c(P) \leq n + 1$ if and only if F is satisfiable. We leave the details to the reader. ■

The following consequence of Theorem 11 is immediate.

Corollary 12 *Finding a path with as few different colors as possible in a colored graph is NP-hard.* ■

Remark 13 Broersma & Li [3] proved that the problem of finding a spanning tree with as few colors as possible in a colored graph is NP-hard by using the minimum dominating set problem. It is not difficult to see that the graph G we constructed in the proof of Theorem 11 has a spanning tree with at most $n + 1$ colors if and only if F is satisfiable. So, our technique also provides a new proof to Broersma & Li's result.

The problem of finding a cycle with as few colors as possible in a colored graph is also NP-hard. We consider the following decision problem.

Problem 14

INSTANCE: 2-connected graph $G = (V, E)$ with a coloring $C : E \rightarrow \mathbb{N}$ and a given vertex u , positive integer $K \leq c(G)$.

QUESTION: Is there a cycle C passing through u such that $c(C) \leq K$?

Theorem 15 Problem 14 is **NP**-complete.

Proof Let G a connected colored graph. Construct a 2-connected colored graph G' by adding a new vertex u , connecting u to every vertex v of G with an edge and assigning an extra color 0 to the new edges. Then G contains a path P such that $c(P) \leq K$ if and only if G' contains a cycle C passing through u such that $c(C) \leq K + 1$. It follows from Corollary 12 that Problem 14 is **NP**-complete. ■

Corollary 16 Finding a cycle with as few colors as possible in a 2-connected colored graph is **NP**-hard. ■

As we proved in Theorem 11, finding a path with as few colors as possible (minimum path) between two given vertices in a colored graph is **NP**-hard. However, if $c(G)$ is much smaller than $|V(G)|$, say, $c(G) = O(\log_2 |V(G)|)$, there will be some efficient algorithm for solving this problem.

One approach to finding a minimum path between two given vertices s_0 and t_0 is to check whether there is an (s_0, t_0) -path in the graphs $G[E_{i_1}]$ with $1 \leq i_1 \leq k$, $G[E_{i_1} \cup E_{i_2}]$ with $1 \leq i_1, i_2 \leq k$ and $i_1 \neq i_2, \dots$, and $G[E_{i_1} \cup E_{i_2} \cup \dots \cup E_{i_d}]$ with $1 \leq i_1, i_2, \dots, i_d \leq k$ and $i_p \neq i_q$ for $1 \leq p \neq q \leq k$, where E_1, E_2, \dots, E_k are the color classes, $k = c(G)$, and d is the distance between s_0 and t_0 . The complexity of such an algorithm is

$$\left[\binom{k}{1} + \binom{k}{2} + \dots + \binom{k}{\min\{d, k\}} \right] O(|V(G)|) \\ = \min\{O(k^d |V(G)|), O(2^k |V(G)|)\}.$$

Another approach is to transform the colored graph to a number of weighted graphs and use Dijkstra's Algorithm. Assign, e.g., the weights $1, |V(G)|, |V(G)|^2, \dots, |V(G)|^{k-1}$ to the edges of the colored graph such that all the edges with the same color get equal weights. There are $k!$ possibilities, so we get at most $k!$ weighted graphs. It is not difficult to see that a minimum path between two given vertices s_0 and t_0 can be found by determining the shortest paths in each of these weighted graphs. The complexity of such an algorithm is $O(k!|V(G)|^2)$.

It is also of interest to consider approximation algorithms for the minimum path problem. If we use a shortest path between two vertices as an approximate solution for a minimum path, the approximation ratio is $c(G)$. We can also design an approximation algorithm which is similar to Dijkstra's Algorithm for finding a shortest path.

Algorithm 17

Step 1. Set $C(s_0) = \emptyset$, $C(S_0) = 0$, $c(v) = \infty$ for $v \neq s_0$, $S_0 = \{s_0\}$ and $i = 0$.

Step 2. For each $v \in V \setminus S_i$, replace $C(v)$ by $C(u_i) \cup \{C(u_i v)\}$ if $c(v) > |C(u_i) \cup \{C(u_i v)\}|$ and set $c(v) = |C(v)|$. Compute $\min_{v \in V \setminus S_i} \{c(v)\}$ and let u_{i+1} denote a vertex for which this minimum is attained. Set $S_{i+1} = S_i \cup \{u_{i+1}\}$.

Step 3. If $u_{i+1} = t_0$, stop. Otherwise, replace i by $i + 1$ and go to Step 2.

The approximation factor of this algorithm can get arbitrarily large. This can be shown by the graph in Figure 2.

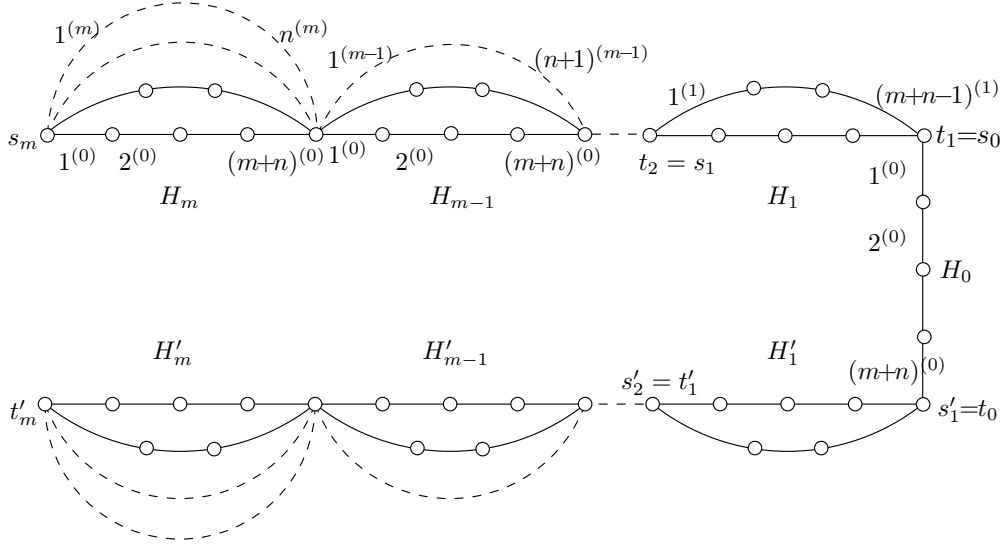


Figure 2

The graph in Figure 2 is constructed as follows. Let H_0 be an (s_0, t_0) -path P_0 with $m+n$ edges and assign the colors $1^{(0)}, 2^{(0)}, \dots, (m+n)^{(0)}$ to the edges of P_0 , respectively. For $1 \leq i \leq m$, the graph H_i is obtained from H_{i-1} by adding a new (s_{i-1}, t_{i-1}) -path P_i with $m+n-i$ edges and assigning the colors $1^{(i)}, 2^{(i)}, \dots, (m+n-i)^{(i)}$ to the edges of P_i , respectively. We denote the two vertices s_{i-1} and t_{i-1} of H_i by s_i and t_i , respectively. By denoting the vertices s_i by s'_i and t_i by t'_i , we get a new graph H'_i for each i with $1 \leq i \leq m$. The graph in Figure 2 is constructed from the graphs H_0, H_i and H'_i ($1 \leq i \leq m$) by identifying the vertices t_0 with s'_1 , the vertices s_{i-1} with t_i , and the vertices s'_i with t'_{i-1} for $1 \leq i \leq m$.

It is easy to see that the minimum path between s_m and t'_m in the graph H is of $m+n$ colors. Whereas we will get an approximate result $n(m+1) + \frac{m(m+1)}{2}$ if we apply the above algorithm to the graph H . So the approximation factor is

$$\left| 1 - \frac{n(m+1) + \frac{m(m+1)}{2}}{m+n} \right| = \frac{mn + \frac{m(m+1)}{2}}{m+n} \rightarrow \infty,$$

when $m \rightarrow \infty$ in the case $n = 1$ or $n = m^{1+\epsilon}$ ($\epsilon > 0$).

From Theorem 11, it is easy to see that the problem of finding a path with fewest colors between two given vertices in a colored directed graph is also **NP**-complete. In the following, we will give a further analysis on the approximation of this problem.

Let $C = \{C_1, C_2, \dots, C_m\}$ be a set of boolean constraints over variables $\{x_1, x_2, \dots, x_n\}$ such that each of the constraints is a boolean function over k variables. For each constrain C_i ($1 \leq i \leq m$), there corresponds a weight w_i . In the MAX k -CSP problem, we seek for an assignment of truth value to the variables $\{x_1, x_2, \dots, x_n\}$ that maximize the sum of the weights of the satisfied constraints.

First of all, for each arc e_{ij} of a colored directed graph D , we assign a variable x_{ij} . Then, for each vertex v of the digraph D there is a boolean constraint function C_v as follows:

1. for the start vertex s of D ,

$$C_s = \left\{ \bigvee_{i \in N^+(s)} [x_{si} \wedge (\bigwedge_{j \in N^+(s), j \neq i} \neg x_{sj})] \right\} \wedge \left\{ \bigwedge_{i \in N^-(s)} \neg x_{is} \right\}.$$

2. for the terminal vertex t of D ,

$$C_t = \left\{ \bigvee_{i \in N^-(t)} [x_{it} \wedge (\bigwedge_{j \in N^-(t), j \neq i} \neg x_{jt})] \right\} \wedge \left\{ \bigwedge_{i \in N^+(t)} \neg x_{ti} \right\}.$$

3. for any other vertices v of D ,

$$C_v = \left\{ \left\{ \bigvee_{i \in N^+(v)} [x_{vi} \wedge (\bigwedge_{j \in N^+(v), j \neq i} \neg x_{vj})] \right\} \wedge \left\{ \bigvee_{i \in N^-(v)} [x_{iv} \wedge (\bigwedge_{j \in N^-(v), j \neq i} \neg x_{jv})] \right\} \right\} \\ \bigvee \left\{ \left[\bigwedge_{i \in N^+(v)} \neg x_{vi} \right] \wedge \left[\bigwedge_{i \in N^-(v)} \neg x_{iv} \right] \right\}.$$

The above are the conditions for the subgraph induced by the set of arcs with truth value "1" to form a directed path from s to t in D . Now we introduce a clause Y_i for each color $i = 1, 2, \dots, c$ as follows:

$$Y_i = \bigwedge \{ \neg x_{uv} \mid e_{uv} \text{ is of color } i \}.$$

We assign a weight for each clauses C_s, C_t, C_v and Y_i such that $w(C_s) = w_1, w(C_t) = w_1, w(C_v) = w_2$ for every $v \in V(D) - \{s, t\}$ and $w(Y_i) = w_3$ for every $i = 1, 2, \dots, c$. We assume that $w_1 \gg w_2 \gg w_3$, say $w_3 = 1, w_2 = c + 1$ and $w_1 = (n - 2)(c + 1) + 1$, where n is the number of vertices of D and c is the number of colors. Then, we conclude that the weighted maximum k CSP problem for the above weighted boolean constraint functions C_s, C_t, C_v and Y_i produces a directed path of D from s to t with as many colors as possible that do not appear on the directed path. Then, from the reference [13, 14] we can get a 2^{1-k} -approximate solution in polynomial time, here k should be $Max_{v,i} \{d_v^+ + d_v^-\}$, the number of arcs of color i , i.e., a solution S such that $|S| \geq 2^{1-k} |S_{Opt}|$.

Be careful, the term "with as many colors as possible that do not appear on the directed path" is different from the term "with as few colors as possible that appear on the directed path". The latter term is exactly what we needed. So, we can formulate another problem as follows:

$X_s = \neg C_s, X_t = \neg C_t, X_v = \neg C_v$, and $Z_i = \neg Y_i = \bigvee \{x_{ij} \mid e_{ij} \text{ is of color } i\}$, where X_s and X_t have a weight w_1, X_v has a weight w_2 and Z_i has a weight w_3 , in which w_i is the same as defined above.

Then, we can see that the weighted minimum k CSP problem for these boolean constraint functions produces a directed path of D from s to t with as few colors as possible. However, unfortunately from reference [8] the weighted minimum k CSP problem does not have any approximate solution of constant factor. So, it seems that our transformation does not have any help for achieving an approximate solution of

some factor to the original problem. The interesting thing here is that if we think about the problem in the other way round, that is, to find a directed path such that the number of not used colors achieves the maximum, then there is indeed a factor 2^{1-k} approximate solution to this problem.

5 Paths and cycles with many colors

If we regard an uncolored graph as a colored graph in which all edges have different colors, then the number of colors of a subgraph is just the number of edges of it. It is well-known that the problem of finding a longest path or a longest cycle in a graph is **NP**-hard. Therefore, the problem of finding a path or a cycle with as many colors as possible in a colored graph are also **NP**-hard.

In the past decades, many sufficient conditions for the existence of long paths and cycles have been derived. The oldest result of this type is due to Dirac.

Theorem C (Dirac [4]) *Let G be graph and d an integer. If $d(v) \geq d$ for every vertex v of G , then G contains (1) a path of length at least d , and (2) a cycle of length at least $d + 1$ if $d > 1$.*

It is an interesting problem to establish whether Theorem C admits a generalization to colored graphs. This leads to the following problem.

Problem 18 *Let G be a colored graph such that $d^c(v) \geq d$ for every vertex v of G , where d is an nonnegative integer. For what values p and c such that G contains a path with at least p colors, and a cycle with at least c colors if $d > 1$?*

There are some examples which show that in Problem 18, both p and c cannot be greater than $d - 1$.

By imposing a higher connectivity, the bound on the cycle length in Theorem C can be increased.

Theorem D (Dirac [4]) *Let G be a 2-connected graph and d an integer. If $d(v) \geq d$ for every vertex v of G , then G contains either a Hamilton cycle or a cycle of length at least $2d$.*

Let $K_{n,n+1}$ be the complete bipartite graph with bipartition (X, Y) such that $|X| = n$ and $|Y| = n+1$. Assign a coloring to $K_{n,n+1}$ as follows: first color the graph $K_{n,n+1} - y$ for some vertex $y \in Y$ by a proper n -edge-coloring, then assign the same n colors of $K_{n,n+1}$ to the n edges incident to y , respectively. It is easy to show that $d^c(v) \geq n$ for each vertex v of $K_{n,n+1}$, but $K_{n,n+1}$ contains neither a Hamilton cycle nor a cycle with more than n colors. This shows that, different from Theorem D, imposing a higher connectivity on the graphs in Conjecture 18 cannot guarantee the existence of cycles with more colors.

References

- [1] M. Albert, A. Frieze and B. Reed, Multicolored Hamilton cycles, *Electronic Journal of Combinatorics* **2** (1995), #R10.

- [2] J.A. Bondy and U.S.R. Murty, *Graph Theory with Applications*, Macmillan London and Elsevier, New York (1976).
- [3] H.J. Broersma and X.Li, Spanning trees with many or few colors in edge-colored graphs, *Discuss. Math. Graph Theory* **17** (1997), 259-269.
- [4] G.A. Dirac, Some theorems on abstract graphs, *Proc. London Math. Soc.* **2** (1952), no. 3, 69-81.
- [5] P. Erdős and T. Gallai, On maximal paths and circuits in graphs, *Acta Math. Sci. Hung.* **10** (1959), 337-356.
- [6] P. Erdős, J. Nešetřil and V. Rödl, *Some problems related to partitions of edges of a graph*, in *Graphs and Other Combinatorial Topics*, Teubner, Leipzig (1983), 54-63.
- [7] A.M. Frieze and B.A. Reed, Polychromatic Hamilton cycles, *Discrete Math.* **118** (1993), 69-74.
- [8] M.R. Garey and D.S. Johnson, *Computers and Intractability*, Freeman, New York (1976).
- [9] G. Giraud, Sur les proportions respectives de triangles uni, bi ou tricolores dans un tricoloiage des arêtes du n -emble, *Discrete Math.* **16** (1976), no. 1, 13-28.
- [10] G. Hahn and C. Thomassen, Path and cycle sub-Ramsey numbers and an edge-coloring conjecture, *Discrete Math.* **62** (1986), no. 1, 29-33.
- [11] E.L. Lawler, *Combinatorial Optimization: Networks and Matroids*, Holt, Rinehart and Winston, New York (1976).
- [12] J.C. Picard and M. Queyranne, A network flow solution to some non-linear 0-1 programming problems, with applications to graph theory, *Networks* **12** (1982), 141-159.
- [13] L. Trevisan, Positive linear programming, parallel approximation and PCP's, In *Proc. 4-th European Symp. Algorithms*, **1136** (1996), 62-75.
- [14] L. Trevisan, Approximating satisfiability problems, In *Proc. 5-th European Symp. Algorithms*, **1284** (1997), 472-485.