Information Sciences 352–353 (2016) 15–34

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Hybrid evolutionary search for the minimum sum coloring problem of graphs

Yan Jin[a,b], Jin-Kao Hao[b,c,*]

[a] School of Computer Science and Technology, Huazhong University of Science and Technology, No. 1037, Luoyu Road, Wuhan, China
[b] LERIA, Université d'Angers, 2 bd Lavoisier, Angers 49045, France
[c] Institut Universitaire de France, Paris, France

ARTICLE INFO

Article history:
Received 6 March 2015
Revised 11 February 2016
Accepted 25 February 2016
Available online 4 March 2016

Keywords:
Sum coloring
Hybrid search
Evolutionary computing
Local optimization
Tabu search

ABSTRACT

Given a graph G, a proper k-coloring of G is an assignment of k colors $\{1, \ldots, k\}$ to the vertices of G such that two adjacent vertices receive two different colors. The minimum sum coloring problem (MSCP) is to find a proper k-coloring while minimizing the sum of the colors assigned to the vertices. This paper presents a stochastic hybrid evolutionary search algorithm for computing upper and lower bounds of this NP-hard problem. The proposed algorithm relies on a joint use of two dedicated crossover operators to generate offspring solutions and an iterated double-phase tabu search procedure to improve offspring solutions. A distance-and-quality updating rule is used to maintain a healthy diversity of the population. We show extensive experimental results to demonstrate the effectiveness of the proposed algorithm and provide the first landscape analysis of MSCP to shed light on the behavior of the algorithm.

1. Introduction

Given a simple undirected graph $G = (V, E)$ with vertex set $V = \{v_1, \ldots, v_n\}$ and edge set $E \subset V \times V$, let $\{1, \ldots, k\}$ be the set of k different colors. A proper k-coloring c of G is a mapping $c : V \rightarrow \{1, \ldots, k\}$ such that $c(v_i) \neq c(v_j)$, for all $\{v_i, v_j\} \in E$. Equivalently, a proper k-coloring can be defined as a partition of V into k mutually disjoint independent sets (or color classes) $V_1, \ldots, V_k$ such that no two vertices of a color class are linked by an edge, i.e., $\forall u, v \in V_i$ $(i = 1, \ldots, k), \{u, v\} \notin E$. The cardinality of a color class V_i is given by the number of its vertices, and is usually denoted by $|V_i|$. The conventional graph coloring problem (GCP) is to color a graph G with a minimum number $\chi(G)$ of colors, $\chi(G)$ is the so-called chromatic number of G. The minimum sum coloring problem (MSCP) studied in this paper is to find a proper k-coloring c of a graph G which minimizes the sum of the colors assigned to the vertices of G [22,36].

$$f(c) = \sum_{i=1}^{n} c(v_i) \text{ or } f(c) = \sum_{l=1}^{k} l|V_l| \tag{1}$$

where $|V_l|$ is the cardinality of V_l, $|V_1| \geq \cdots \geq |V_k|$ and k is larger than or equal to the chromatic number $\chi(G)$ of G in the GCP. Throughout the paper, we assume that the color classes of a k-coloring are sorted in non-increasing order of their

* Corresponding author at: LERIA, Université d'Angers, 2 bd Lavoisier, 49045 Angers, France. Tel.: +33 241735076; fax: +33 241735073.
E-mail addresses: yanjin.china@hotmail.com (Y. Jin), jin-kao.hao@univ-angers.fr, hao@info.univ-angers.fr (J.-K. Hao).

http://dx.doi.org/10.1016/j.ins.2016.02.051
0020-0255/

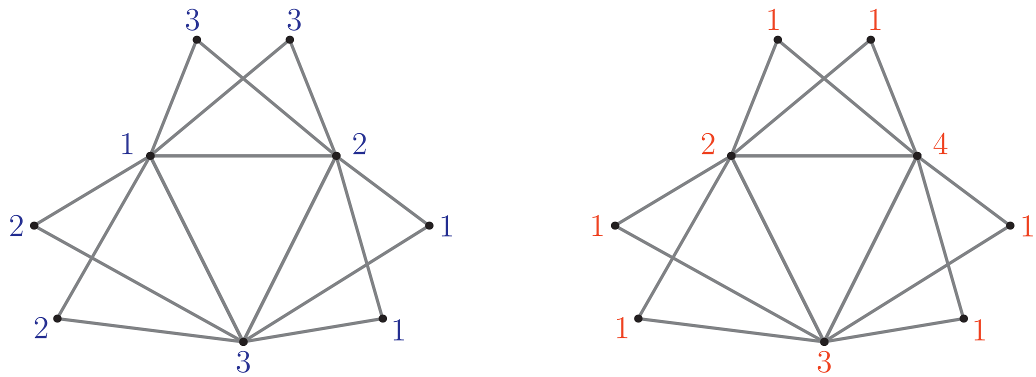

Fig. 1. An illustrative example for MSCP.

cardinality. The minimum sum of colors for MSCP is called the chromatic sum of G, and is denoted by $\Sigma(G)$. MSCP has practical applications in areas like VLSI design, scheduling and resource allocation [27].

Notice that although MSCP is tightly related to the conventional GCP, MSCP and GCP have different optimization objectives (minimization of the sum of colors vs. minimization of the number of colors). Fig. 1 provides an example for MSCP which also illustrates the relation with GCP. The graph has a chromatic number $\chi(G)$ of 3 (left figure), but requires 4 colors to achieve the chromatic sum (right figure). Indeed, with the given 4-coloring, we achieve the chromatic sum of 15 while the 3-coloring leads to a suboptimal sum of 18.

From a theoretical point of view, MSCP is proved to be NP-hard [22]. There exist a number of studies on specific graphs. For instance, a polynomial time algorithm is available for finding the chromatic sum of a tree [22]. Several approximation algorithms are also proposed for bipartite graphs, chain bipartite graphs and interval graphs, k-split graphs, and line graphs of trees [2,4,16,21,23,34]. Unfortunately, MSCP remains computationally difficult and challenging in the general case.

Given the hardness of the general MSCP, a number of heuristic algorithms have been proposed to obtain suboptimal solutions (upper bounds)or to compute lower bounds in acceptable computing time.

In 2007, Kokosiński and Kwarciany [20] presented the first parallel genetic algorithm which employs assignment and partition crossovers, first-fit mutation, and proportional selection with an island migration model. They showed the first computational results in reference to theoretical upper bounds on 16 small DIMACS graphs. In 2009, Li et al. [25] experimented two greedy algorithms (MDSAT & MRLF) which are adaptations of two well-known GCP heuristics called DSATUR [6] and RLF [24] to MSCP. Their experimental results showed that MDSAT & MRLF perform better than DSATUR & RLF. Later in 2010, Moukrim et al. [29] proposed a clique decomposition technique for computing a lower bound of MSCP. In 2010, Bouziri and Jouini [5] adapted the well-known Tabucol coloring algorithm of [15] to MSCP. The experimental results applied on seven small DIMACS instances are better than those of the greedy algorithms MDSAT & MRLF. In 2011, Helmar and Chiarandini [14] elaborated a local search heuristic (MDS(5)+LS) to find upper and lower bounds of MSCP, which combines variable neighborhood search and iterated local search to oscillate between feasible and infeasible regions. Comparative results showed that MDS(5)+LS attains new bounds for 27 out of 38 tested instances and outperforms the above three algorithms. In 2012, Benlic and Hao [3] developed a breakout local search algorithm (BLS) which jointly uses a local search and adaptive perturbation strategies. They reported improved upper bounds for 4 instances out of 27 tested graphs. In 2012, Wu and Hao [37] devised an effective heuristic using independent set extraction (EXSCOL) which performs especially well on large graphs with more than 500 vertices. Later in 2013, the same authors [38] applied this approach for lower bound computation. In 2014, Moukrim et al. [30] introduced a memetic algorithm using a two-parent crossover combined with a hill-climber and "destroy and repair" procedure (MA), and reported high quality upper and lower bounds on 81 tested instances. The same year, Jin et al. [17] presented another memetic algorithm based on tabu search and a multi-parent crossover (MASC). MASC discovered 15 new upper bounds out of 77 tested graphs.

In this work, we are interested in the computation of both upper and lower bounds of MSCP. For this, we introduce an effective stochastic hybrid evolutionary search algorithm (HESA) whose main contributions can be summarized as follows.

- From the algorithm perspective, the HESA approach integrates several special features to ensure a high search efficiency. These include an original recombination mechanism to generate offspring solutions and an iterated double-phase tabu search procedure to ensure local optimization. The solution recombination mechanism combines a diversification-guided crossover operator and a grouping-guided crossover operator to create diversified and promising offspring solutions. The double-phase tabu search procedure is designed to handle both feasible and unfeasible solutions. A dedicated perturbation mechanism is also introduced to escape local optima. Finally, a population updating procedure is employed to maintain a high-quality population with a healthy diversity.
- From the computational perspective, we evaluate the HESA approach on 94 well-known DIMACS and COLOR 2002–2004 benchmark instances. The computational results show that HESA can achieve the best-known result for most of these