Discrete Optimization 15 (2015) 26–36

Contents lists available at ScienceDirect

Discrete Optimization

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

Clique partitioning with value-monotone submodular cost

José R. Correa [a], Nicole Megow [b,*]

[a] Departamento Ingeniería Industrial, Universidad de Chile, Santiago, Chile

[b] Department of Mathematics, Technische Universität Berlin, Germany

ARTICLE INFO

Article history:
Received 22 July 2011
Received in revised form 10 September 2014
Accepted 3 November 2014
Available online 8 December 2014

Keywords:
Partition into cliques
Cost coloring
Submodular functions
Cardinality constraint
Max-coloring
Batch-scheduling with compatibilities

ABSTRACT

We consider the problem of partitioning a graph into cliques of bounded cardinality. The goal is to find a partition that minimizes the sum of clique costs where the cost of a clique is given by a set function on the nodes. We present a general algorithmic solution based on solving the problem variant without the cardinality constraint. We obtain constant factor approximations depending on the solvability of this relaxation for a large class of submodular cost functions which we call value-monotone submodular functions. For special graph classes we give optimal algorithms.

1. Introduction

We consider the problem of partitioning a graph into cliques of bounded cardinality. Given is a simple graph $G = (V, E)$, a set function $f : 2^V \to \mathbb{R}^+$, and a bound $B \in \mathbb{Z}^+$. The task is to find a partition of G into cliques $K_1, \ldots, K_\ell$ of cardinality at most B. The cost of a clique K_i is given by the set function $f(K_i)$ via an oracle. Our objective is to find a partition of minimum total cost, $\sum_{i=1}^{\ell} f(K_i)$. We denote our problem as *partition into cliques of bounded cardinality* PCliq(G, f, B).

The problem PCliq(G, f, B) is a generalization of the classical *graph partitioning problem* and thus $\mathcal{NP}$-hard [1]. For arbitrary graphs the problem is even $\mathcal{NP}$-hard to approximate within a factor $|V|^{1-\varepsilon}$ for all $\varepsilon > 0$ [2] since it contains *vertex coloring* as a special case. Indeed, our problem can be formulated equivalently as finding a partition into independent sets (coloring) of bounded cardinality and minimum total cost in $\overline{G}$, the complement of G, where the cost of each independent set (color class) is defined by some cost function f. In particular, PCliq(G, f, B) with $B = \infty$ and $f(S) \equiv 1$ is equivalent to the classical vertex coloring problem in $\overline{G}$. However, in the special case that the bound on the clique size B equals 2 then PCliq(G, f, B) can be transformed into a matching problem in G and can be solved optimally in polynomial time. Since cliques in a bipartite graph have size at most 2, the graph partitioning problem corresponds in this case to a matching problem.

Graph partitioning and coloring problems are among the fundamental problems in combinatorial optimization. Imposing bounds on the size of the cliques or color classes is very natural in many applications. Such problems are often formulated also in a scheduling context where compatibilities or conflicts between jobs are expressed in a graph, and the task is to find a partition into cliques or independent sets of jobs that minimize the total cost with respect to some scheduling cost function. The problems of partitioning or coloring with a cardinality bound have been mainly studied for two cost functions: $f(S) \equiv 1$ and the max-function. The coloring problem with the objective to minimize the number of colors, i.e., $f(S) \equiv 1$, is

* Corresponding author.
E-mail addresses: correa@uchile.cl (J.R. Correa), nmegow@math.tu-berlin.de (N. Megow).

http://dx.doi.org/10.1016/j.disopt.2014.11.001
1572-5286/

known as *mutual exclusion scheduling* [3–6] or *bounded vertex coloring* [7], and when the cost function is the max-function $f(S) = \max_{v \in S} w(v)$ for some weight function $w: V \to \mathbb{R}^+$ then the coloring problem is also called *bounded max-coloring problem* [8]. The complementary problem of partitioning a graph into cliques with a max-cost function has been studied also as *max-batch scheduling with compatible jobs* [9–12].

This body of previous work implies for our problem PCliq(G, f, B) that it is already for the simple cost function $f(S) \equiv 1$ $\mathcal{NP}$-hard on cographs, co-bipartite and co-interval graphs [3], on comparability and co-comparability graphs even when B is fixed ($B \geq 6$) [6], and for the complements of line graphs also for any fixed $B \geq 3$ [13]. However, there are polynomial-time optimal algorithms for split graphs [3,14], interval graphs [3,9], as well as for the complements of forests [5] and of line graphs [15].

The problem PCliq(G, f, B) with f being the max-function is strongly $\mathcal{NP}$-hard even for $B = \infty$ for co-bipartite graphs [16], co-interval graphs [17], split graphs [10,16], and interval graphs [12]. Interval graphs found particular interest in the literature because of their wide applicability in practical applications, e.g., describing compatible processing intervals in scheduling. For this graph class there is a polynomial-time approximation scheme (PTAS) for any fixed capacity bound $B \geq 3$ [12]. A PTAS is also known for graphs that are complements of trees [8]; here the complexity status remains unsolved. And for co-bipartite graphs there is a 17/11-approximation [8].

The relaxed variant of our problem PCliq(G, f, B) without bounds on the clique size, i.e., $B = \infty$, has been widely studied in different settings. Various complexity and approximability results are known for particular cost functions and graph classes. Interestingly, many of these particular cost functions that appear in specific applications share the property of being submodular. Examples for such special problems and submodular cost functions are the above-mentioned *maximum function*, the *probabilistic cost function* [18,19], *partial q-coloring* [20,21], (modified) *chromatic entropy* [22,23], and more generally *non-decreasing weight-defined concave cost functions* [24]. Definitions and details will be given in Section 2.

While most research focused on particular cost functions, recently more general set functions have been considered in this context. Gijswijt, Jost, and Queyranne [25] introduce so-called *value-polymatroidal* set functions. They consider the graph partitioning problem without capacity restrictions, PCliq(G, f, ∞), and derive a polynomial time algorithm for interval graphs and circular arc graphs. Furthermore, Fukunaga, Halldórsson, and Nagamochi [24] assume a weight function $w: V \to \mathbb{R}^+$ and consider weight-defined *monotone concave* cost functions. They provide a general algorithmic scheme that yields a factor 4 approximation for clique partitioning without cardinality constraint, PCliq(G, f, ∞), in co-interval graphs, comparability and co-comparability graphs, and a 6-approximation for perfect graphs. In fact, they obtain a robust result in the sense that the solution they compute approximates all cost functions under consideration simultaneously.

Our results. We investigate the problem of partitioning a graph into cliques of bounded cardinality for a general class of cost functions. We consider *value-monotone submodular functions*. This subclass of submodular functions is only slightly more restricted than the class of value-polymatroidal functions introduced by Gijswijt et al. [25] in a similar context, and contains well-known cost functions for coloring or partitioning problems such as the max-function, probabilistic function, partial q-coloring, (modified) chromatic entropy, and more generally non-decreasing weight-defined concave cost functions.

We provide a very simple but general approximation algorithm for solving the graph partitioning problem for general graphs based on a solution to the relaxed problem without cardinality constraints on the cliques. We show that adding the bound on the cardinality does not diminish the solution quality too much. More precisely, if there is an α-approximate solution for the unbounded problem, then a Greedy algorithm turns this solution into an $(\alpha + 1)$-approximate solution that obeys the cardinality constraint. This general result implies several new or improved results for special graph classes. In particular, we obtain a 2-approximation for interval graphs and circular arc graphs for arbitrary value-monotone submodular cost functions. For special cost functions we obtain improved results such as, e.g., an $(e + 1)$-approximation for the general class of perfect graphs when f is the max-function. More generally, we obtain a 5-approximation for co-interval, comparability, and co-comparability graphs, and a 7-approximation for perfect graphs when f is a non-decreasing weight-defined concave function.

As a subproblem, we investigate the problem variant of partitioning a complete graph into cliques of bounded cardinality. We show that this seemingly simple problem is $\mathcal{NP}$-hard for polymatroid rank functions, a subclass of submodular functions. However, we also show that restricting to value-monotone cost functions allows a Greedy algorithm to solve the problem optimally in polynomial time.

Finally, we consider the partitioning problem on proper interval graphs. This graph class plays an important role in many applications, e.g., as compatibility graphs in the above mentioned max-batch scheduling problem [9]. We derive a dynamic programming algorithm that solves the problem optimally in running time that is exponential in the number w of different node types with respect to the cost function f. Thus, this algorithm computes an optimal solution in polynomial time if w is bounded by a constant. Moreover, using a recently shown technique to reduce the number w at the cost of a small increase in the objective function when f is the max-function [12], the algorithm leads in this case to a PTAS. Thus, we give a PTAS for max-batch scheduling with compatibilities that form a proper interval graph. This contrasts the previously known PTAS [12] which has polynomial running time only when the capacity bound B is constant. Note, however, that this result was actually derived for general interval graphs.

Outline. In Section 2 we define value-monotone submodular set functions, give examples, and discuss important properties. In Section 3 we investigate the clique partitioning problem on complete graphs. We prove optimality of a Greedy algorithm and show that the complexity status changes when allowing submodular cost functions that do not satisfy the