Neural Networks 22 (2009) 58–66

Contents lists available at ScienceDirect

Neural Networks

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

A deterministic annealing algorithm for approximating a solution of the min-bisection problem

Chuangyin Dang [a,*], Wei Ma [a], Jiye Liang [b]

[a] Department of Manufacturing Engineering and Engineering Management, City University of Hong Kong, Kowloon, Hong Kong

[b] School of Computer and Information Technology, Shanxi University, Taiyuan, China

A R T I C L E I N F O

Article history:
Received 25 April 2008
Received in revised form
12 September 2008
Accepted 17 September 2008

Keywords:
Log-cos barrier function
Descent direction
Deterministic annealing
Iterative algorithm

A B S T R A C T

The min-bisection problem is an NP-hard combinatorial optimization problem. In this paper an equivalent linearly constrained continuous optimization problem is formulated and an algorithm is proposed for approximating its solution. The algorithm is derived from the introduction of a logarithmic-cosine barrier function, where the barrier parameter behaves as temperature in an annealing procedure and decreases from a sufficiently large positive number to zero. The algorithm searches for a better solution in a feasible descent direction, which has a desired property that lower and upper bounds are always satisfied automatically if the step length is a number between zero and one. We prove that the algorithm converges to at least a local minimum point of the problem if a local minimum point of the barrier problem is generated for a sequence of descending values of the barrier parameter with a limit of zero. Numerical results show that the algorithm is much more efficient than two of the best existing heuristic methods for the min-bisection problem, Kernighan–Lin method with multiple starting points (MSKL) and multilevel graph partitioning scheme (MLGP).

1. Introduction

Consider an undirected graph $G = (V, E)$, where $V = \{1, 2, \ldots, n\}$ is the node set of G and E the edge set of G. We denote by (i, j) an edge between nodes i and j. Let

$$W = \begin{pmatrix} 0 & w_{12} & \cdots & w_{1n} \\ w_{21} & 0 & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & 0 \end{pmatrix}$$

be a given symmetric weight matrix such that $w_{ij} > 0$ if $(i, j) \in E$ and $w_{ij} = 0$ if $(i, j) \notin E$. Assume that G has an even number of nodes. The min-bisection problem is to partition V into two sets, S and $V \setminus S$, of equal cardinality such that

$$w(S) = \sum_{i \in S, j \in V \setminus S} w_{ij}$$

is minimized. This problem is an NP-hard problem (Murty & Kabadi, 1987) and has many applications. A variant of it called the bisection problem with k-resource sets can be found in Ishii, Iwata, and Nagamochi (2007). Due to its computational complexity, the min-bisection problem is very difficult to solve to optimality with an exact algorithm.

The min-bisection problem is to partition a graph into $p = 2$ parts of equal cardinality, a natural generalization of which is the case when $p \geq 2$, yielding the graph partitioning problem. This problem has attracted much attention in recent years, because of its extensive applications in such areas as scientific computing, VLSI design, etc. According to Karypis and Kumar (1998), algorithms for graph partitioning are of three major types, which are spectral partitioning methods (Amin, 2005), geometric partitioning algorithms, and multilevel graph partitioning schemes with three phases: coarsening, partitioning of the coarsest graph, and refining (Hendrickson & Leland, 1993). During the past few decades, meta-heuristics have become popular, among which are simulated annealing (SA) and evolutionary algorithms (EAs), two most remarkable algorithms. An application of meta-heuristic algorithms to the graph partitioning problem can be found in many literatures. For details, please refer to Bui and Moon (1996), Gil, Ortega, Diaz, and Montoya (1998), Jerrum and Sorkin (1998) and Soper, Walshaw, and Cross (2004). In addition, some software packages such as JOSTLE, METIS, CHACO are available. In Loiola, Abreu, Boaventura-Netto, Hahn, and Querido (2007), the graph partitioning problem was formulated as a special case of the quadratic assignment problem (QAP). Furthermore, with its speciality in mind, the min-bisection problem has enjoyed some privileges, theoretical and algorithmic. For details, they are referred to Saab and Rao (1992) and Feig and Kuauthgamer (2006).

* Corresponding author. Tel.: +852 27888429; fax: +852 27888423.
E-mail address: mecdang@cityu.edu.hk (C. Dang).

0893-6080/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2008.09.008

Neural Networks, since their emergence, have experienced significant advances in both theory and applications, especially in optimization, among which combinatorial optimization, due to Hopfield and Tank (1985), has become a popular topic in the literature of neural computation. Although initial results were disappointing, modified network dynamics and better problem mapping contribute significantly to solution quality (Gee, Aiyer, & Prager, 1993). Peterson and Soderberg (1989) mapped the graph partition problem onto a neural network with the graded neurons encoding, which can reduce the solution space by one dimension. Gee et al. (1993) presented a problem mapping evaluation method without recourse to purely experimental means. Gee and Prager (1994) proposed a rigorous mapping for quadratic 0–1 programming problems with linear equality and inequality constraints. After transforming variables with exponential functions, Urahama (1996) presented an analog solver for nonlinear programming problems with linear constraints. A feasible solution construction mechanism was introduced in Horio, Ikeguchi, and Aihara (2005) to improve the performance of the Hopfield-type chaotic neuro-computer system for quadratic assignment problems (QAPs). Bout and Miller (1990) and Wu (2004) applied the mean field annealing (MFA) algorithm for a solution of the graph bisection problem. Waugh and Westervelt (1993) introduced a neural network architecture that is applicable in optimization. By combining deterministic annealing, self-amplification, algebraic transformations, clocked objectives, and softassign, an optimizing network architecture was constructed in Rangarajan, Gold, and Mjolsness (1996). A Hybrid of Lagrange and transformation approaches (Hybrid LT) was proposed in Xu (1994) for solving combinatorial optimization problems whose constraints were separated into linear-constant-sum constraints and binary constraints and they were, respectively, treated by Lagrange approach and penalty or barrier functions. Furthermore, special network models were constructed for the traveling salesman problem (Aiyer, Niranjan, & Fallside, 1990; Dang & Xu, 2001; Durbin & Willshaw, 1987; Wacholder, Han, & Mann, 1989; Wolfe, Parry, & MacMillan, 1994). Statistical mechanics as the underlying theory of optimization neural networks was studied in Simic (1990). A systematic investigation of such neural computational models for combinatorial optimization can be found in Berg (1996) and Cichocki and Unbehaunen (1993). Most of these algorithms are of deterministic annealing type, which is a heuristic continuation method that attempts to find the global minimum of the effective energy at a high temperature and track it as the temperature decreases. There is no guarantee that the minimum at a high temperature can always be tracked to the minimum at a low temperature, but the experimental results are encouraging (Yuille & Kosowsky, 1994).

In this paper we adapt the idea of deterministic annealing for approximating a solution of the min-bisection problem. An equivalent linearly constrained continuous optimization problem is formulated and an algorithm is proposed for approximating its solution. The algorithm is derived from the introduction of a logarithmic-cosine barrier function, where the barrier parameter behaves as temperature in an annealing procedure and decreases to zero from a sufficiently large positive number satisfying that the barrier function is convex. The algorithm searches for a better solution in a feasible descent direction, which has a desired property that lower and upper bounds are always satisfied automatically if the step length is a number between zero and one. We prove that the algorithm converges to at least a local minimum point of the problem if a local minimum point of the barrier problem is generated for a sequence of descending values of the barrier parameter with a limit of zero. The main differences between the proposed algorithm and existing neural computational models are the barrier function and the feasible descent direction. Numerical results show that the algorithm is much more efficient than MSKL and MLGP, two best existing heuristic methods for the min-bisection or graph partitioning problem.

The rest of this paper is organized as follows. In Section 2 we formulate an equivalent linearly constrained continuous optimization problem to the min-bisection problem, introduce a logarithmic-cosine barrier function, and derive several important theoretical results. In Section 3 we describe a deterministic annealing algorithm for approximating a solution of the min-bisection problem. In Section 4 we present some numerical results to show that the algorithm is effective and efficient. Finally, we conclude the paper with some remarks in Section 5.

2. A logarithmic-cosine barrier function

It is clear that the min-bisection problem is equivalent to

$$\begin{aligned} &\min \frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n}(1-x_i x_j)w_{ij} \\ &\text{subject to } \sum_{i=1}^{n} x_i = 0, \\ &\qquad x_i \in \{-1, 1\}, \quad i = 1, 2, \ldots, n. \end{aligned} \tag{1}$$

Let

$$\xi_i = \max_{1\le j\le n} w_{ji}$$

for $i = 1, 2, \ldots, n$, and $\xi = (\xi_1, \xi_2, \ldots, \xi_n)^\top$. Then the min-bisection problem is equivalent to

$$\begin{aligned} &\min f(x) = -\frac{1}{2}x^\top (W + \Xi + \alpha I)x \\ &\text{subject to } \sum_{i=1}^{n} x_i = 0, \\ &\qquad x_i \in \{-1, 1\}, \quad i = 1, 2, \ldots, n, \end{aligned} \tag{2}$$

where Ξ is the diagonal matrix formed by the components of ξ, α any given positive number, and I an $n \times n$ identity matrix. A continuous relaxation of (2) yields

$$\begin{aligned} &\min f(x) = -\frac{1}{2}x^\top (W + \Xi + \alpha I)x \\ &\text{subject to } \sum_{i=1}^{n} x_i = 0, \\ &\qquad -1 \le x_i \le 1, \quad i = 1, 2, \ldots, n. \end{aligned} \tag{3}$$

Let

$$b(x) = -\sum_{i=1}^{n} \ln\cos\left(\frac{\pi}{2}x_i\right),$$

which will be used as a barrier term to incorporate $-1 \le x_i \le 1$, $i = 1, 2, \ldots, n$, into the objective function. For any given positive number β, consider

$$\begin{aligned} &\min h(x;\beta) = f(x) + \beta b(x) \\ &\text{subject to } \sum_{i=1}^{n} x_i = 0. \end{aligned} \tag{4}$$

Let $F = \{x \mid \sum_{i=1}^{n} x_i = 0\}$,

$B = \{x \mid -1 \le x_i \le 1, i = 1, 2, \ldots, n\}$,

and

$\text{int}(B) = \{x \mid -1 < x_i < 1, i = 1, 2, \ldots, n\}$.

Note that

$$\frac{\partial b(x)}{\partial x_i} = \frac{\pi}{2}\tan\left(\frac{\pi}{2}x_i\right).$$