Applied Mathematics and Computation 316 (2018) 167–173

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

Short Communication

Primal dual based algorithm for degree-balanced spanning tree problem

Yingli Ran[a], Zhihao Chen[b], Shaojie Tang[c], Zhao Zhang[b,*]

[a] College of Mathematics and System Sciences, Xinjiang University, Urumqi, Xinjiang 830046, China
[b] College of Mathematics Physics and Information Engineering, Zhejiang Normal University, Jinhua, Zhejiang 321004, China
[c] Naveen Jindal School of Management, University of Texas at Dallas, Richardson, TX 75080, USA

ARTICLE INFO

Keywords:
Degree-balanced spanning tree
Nonlinear objective function
Primal dual algorithm

ABSTRACT

This paper studies approximation algorithm for the degree-balanced spanning tree (DBST) problem. Given a graph $G = (V, E)$, the goal is to find a spanning tree T such that $\sum_{v\in V} deg_T(v)^2$ is minimized, where $deg_T(v)$ denotes the degree of node v in tree T. The idea of taking squares on node degrees is to manifest the role of nodes with large degree, and thus minimizing the sum will result in a comparatively balanced degree distribution. This is a non-linear objective function. We prove that DBST is NP-hard, and then develop a primal–dual based algorithm with a guaranteed performance ratio.

1. Introduction

The degree-balanced spanning tree problem can find its applications in many domains such as wireless networks and distributed streaming system. In the study of wireless sensor networks [4], tree structure is often used to serve as a backbone guiding message broadcasting or information aggregation etc. One major concern in constructing routing trees in wireless networks is to maximize the network lifetime, it is highly desirable to balance the transmission load among all wireless nodes, because otherwise there will be large overhead at some nodes. In addition, low transmission delay is also crucial for network operations, however, the node with highest degree can easily become the bottleneck that dominates the entire transmission delay. An effective approach to cope with these problems is to construct a degree-balanced routing tree.

Another motivation is from live distributed streaming systems [2] in which a source diffuses a content to peers via a tree. Since such system often experiences high churn, peers frequently joining and leaving the system, it is desirable to repair the diffusion tree to allow an efficient data distribution, where an efficient diffusion tree must ensure that node degrees are bounded due to limited bandwidth.

In this paper, we propose using the sum of degree squares to measure the degree of balance of a tree. The idea is to emphasize the contribution of nodes with large degrees, and thus one needs to minimize the objective value in order to obtain a tree with a balanced degree distribution.

Definition 1.1 (Degree-balanced spanning tree (DBST))**.** Given a graph $G = (V, E)$, our goal is to find a spanning tree T such that $SDS(T) = \sum_{v\in V} deg_T(v)^2$ is minimized, where $deg_T(v)$ denotes the degree of v in tree T.

* Corresponding author.
E-mail addresses: hxhzz@sina.com, hxhzz@163.com, zhaozhang@zjnu.cn (Z. Zhang).

http://dx.doi.org/10.1016/j.amc.2017.08.016
0096-3003/

A related problem is the *degree-concentrated spanning tree* problem (DCST), the goal of which is to maximize the sum of degree squares of a spanning tree. DCST was proposed in [3] as an application of Discrete DC programming. A local optimal solution to DCST can be found in polynomial time.

As to the degree-balanced spanning tree problem, we shall prove that it is NP-hard, and then present a primal–dual algorithm with a guaranteed performance ratio.

2. Preliminaries

Lemma 2.1. *The DBST problem is NP-hard.*

Proof. We claim that a graph G on n vertices has a Hamiltonian path if and only if the optimal solution to DBST on G has value $4n-6$.

For this purpose, it suffices to show that among all trees on $n \geq 3$ vertices, a path has the minimum SDS value. This claim can be proved by showing that any tree can be transformed step by step into a path on the same number of vertices such that in each step the SDS value does not increase. Suppose T is a tree which is not a path. Let u be a leaf of T and p_u be the unique neighbor of u. Let v be another leaf of T. Trim u from p_u and append u to v. Denote the resulting tree as T'. Then $SDS(T)-SDS(T')=deg_T(p_u)^2-(deg_T(p_u)-1)^2+1^2-2^2=2deg_T(p_u)-4\geq 0$ since p_u is not a leaf. Hence $SDS(T')\leq SDS(T)$. Recursively using such an operation will result in a path. □

3. Primal–dual programme for DBST

For a vertex $v\in V$ and an edge $e\in E$, $e\sim v$ indicates that e is incident to vertex v. For a vertex set $\emptyset\neq S\subsetneq V$, we use $\delta(S)$ to denote the set of edges with exactly one end in S.

The following programme is a relaxed program of DBST. In fact, in the integer programme of DBST, $x_e=1$ if e is taken into tree T, and $x_e=0$ otherwise; y_v is the degree of vertex v in tree T. When we relax the programme allowing the variables to take real values, it suffices to require $x_e\geq 0$ ($x_e\leq 1$ is not needed since an optimal solution to (1) must have $x_e\leq 1$).

$$\min \sum_{v\in V} y_v^2$$
$$s.t. \begin{cases} \sum_{e\in\delta(S)} x_e \geq 1, & \emptyset\neq S\subsetneq V \\ y_v=\sum_{e\sim v} x_e, & v\in V \\ x_e\geq 0, & e\in E \end{cases} \tag{1}$$

We shall prove that the following programme can serve as a dual of (1).

$$\max \sum_{\emptyset\neq S\subsetneq V} p_S - \sum_{v\in V} z_v^2$$
$$s.t. \begin{cases} 2\sum_{v:e\sim v} z_v \geq \sum_{S:e\in\delta(S)} p_S, & e\in E, \\ z_v\geq 0, & v\in V \\ p_S\geq 0, & \emptyset\neq S\subsetneq V \end{cases} \tag{2}$$

Lemma 3.1. *Let opt_1 and opt_2 be the optimal values of (1) and (2) respectively. Then $opt_2\leq opt_1$.*

Proof. It is easy to see that

$$y_v^2\geq 2y_vz_v-z_v^2,\ \forall v\in V \tag{3}$$

For a fixed vector $\mathbf{z}=(z_v)_{v\in V}$, we have the following program.

$$\min\ 2\sum_{v\in V} y_vz_v$$
$$s.t. \begin{cases} \sum_{e\in\delta(S)} x_e \geq 1, & \emptyset\neq S\subsetneq V \\ y_v=\sum_{e\sim v} x_e, & v\in V \\ x_e\geq 0, & e\in E \end{cases} \tag{4}$$

This is a linear program in variables $\{x_e\}_{e\in E}$ and $\{y_v\}_{v\in V}$. Its dual is as follows.

$$\max \sum_{\emptyset\neq S\subsetneq V} p_S$$
$$s.t. \begin{cases} 2\sum_{v:e\sim v} z_v \geq \sum_{S:e\in\delta(S)} p_S, & e\in E, \\ p_S\geq 0, & \emptyset\neq S\subsetneq V \end{cases} \tag{5}$$