ARTICLE IN PRESS
JID: APM [m3Gsc;December 15, 2017;17:5]

Applied Mathematical Modelling 000 (2017) 1–10

Contents lists available at ScienceDirect

Applied Mathematical Modelling

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

LP-rounding approximation algorithms for two-stage stochastic fault-tolerant facility location problem☆

Sai Ji[a], Dachuan Xu[a,*], Donglei Du[b], Yijing Wang[a]

[a] Department of Information and Operations Research, College of Applied Sciences, Beijing University of Technology, 100 Pingleyuan, Chaoyang District, Beijing 100124, PR China
[b] Faculty of Business Administration, University of New Brunswick, Fredericton NB E3B 5A3, Canada

ARTICLE INFO

Article history:
Received 17 September 2016
Revised 15 July 2017
Accepted 5 December 2017
Available online xxx

Keywords:
Stochastic
Fault-tolerant
LP-rounding
Facility location problem

ABSTRACT

In this paper, we study the weighted two-stage stochastic fault-tolerant facility location problem. We present a deterministic LP-rounding 5-approximation algorithm by exploiting both of its stochastic and fault-tolerant structures. We further offer an improved randomized LP-rounding 3.8617-approximation algorithm along with the corresponding derandomized version with the same approximation ratio.

1. Introduction

The facility location problem (FLP) has been widely studied in the operations research literature. In the classical FLP, we are given a facility set F and a client set C. There are an open cost $f_i \in R_+$ for each $i \in F$ and a service cost $c_{ij} \in R_+$ for each $i \in F$ and $j \in C$. We assume that c is a metric. The objective is to find an open facility subset S and an assignment $\sigma : C \to S$ of the clients to open facilities such that the total cost including the open cost and service cost is minimized.

Since the FLP is NP-hard, one cannot obtain the optimal solution in polynomial time unless $P = NP$. There are many existing work on approximation algorithms for this problem. Shmoys et al. [1] give the first constant approximation ratio of 3.16 by using the LP-rounding technique, followed by several improved approximation algorithms [2–8]. The currently best approximation ratio is 1.488 [9] based on LP-rounding combined with dual-fitting techniques. Guha and Khuller [4] and Sviridenko [10] show that the lower bound is 1.463 under the assumption $P \neq NP$. For more discussion of the variants of the FLP, we refer to [11–14] and the references therein.

Among these variants, we are particularly interested in the fault-tolerant facility location problem (FTFLP) and the two-stage stochastic facility location problem (2-SFLP). The FTFLP and 2-SFLP are introduced by Jain and Vazirani [15] and Ravi and Sinha [16], respectively.

In the FTFLP, each client j has an integer connection requirement r_j. The objective is to open some facilities and connect each client j to r_j different open facilities such that the total open cost and service cost is minimized. The FTFLP is reduced to the FLP if $r_j = 1$ for all j. Byrka et al. [17] obtain an improved dependent LP-rounding 1.7245-approximation algorithm.

☆ This paper has been presented on the 10th International Conference on Optimization: Techniques and Applications (ICOTA10).

* Corresponding author.
E-mail addresses: jisai@emails.bjut.edu.cn (S. Ji), xudc@bjut.edu.cn (D. Xu), ddu@unb.ca (D. Du), yjwang@emails.bjut.edu.cn (Y. Wang).

https://doi.org/10.1016/j.apm.2017.12.009
0307-904X/© 2017 Published by Elsevier Inc.

Guha et al. [18] introduce a general weight version for the FTFLP in which the service cost of each client j is a weighted sum of its distance to the r_j different facilities. For the restricted weight version with nonincreasing weight vectors, they present an LP-rounding 4-approximation algorithm and further improve to 3.16 and 2.408 using randomization and greedy argumentation techniques. Hajiaghayi et al. [19] give two LP-rounding 4-, 3.16-approximation algorithms for the general weight FTFLP without the nonincreasing restriction on weight vectors.

In the 2-SFLP, we are given two stages, possible scenarios together with the corresponding clients and probability distributions. Each facility can be opened in the first stage to serve any client in any scenario. The facilities opened in a given scenario can only be used to serve clients in that particular scenario. Usually, the open cost of each facility in the first stage is more expensive comparing with that in the second stage. The objective is to assign each client in each scenario to an open facility which is opened either in the first stage or in the second stage such that the expected total open cost and service cost is minimized. Ravi and Sinha [16] present an LP-rounding 8-approximation algorithm, which is further investigated by Shmoys and Swamy [20,21]. The currently best approximation ratio for the 2-SFLP is 1.8526 by Ye and Zhang [22].

Wu et al. [23] study the two-stage stochastic fault-tolerant facility location problem (2-SFTFLP) and give an LP-rounding 5-approximation algorithm for the restricted weight version with nonincreasing weight vectors. Wu et al. [24] proposing an LP (location problem)-rounding approximation algorithm with 2.3613 per-scenario bound for this problem. In this paper, we consider the general weight 2-SFTFLP without any restriction on the weight vectors. Following the LP-rounding technique by Hajiaghayi et al. [19], we present a deterministic LP-rounding 5-approximation algorithm. Using a standard randomization rounding technique (cf. [1]), we further offer an improved randomized LP-rounding 3.8617-approximation algorithm along with the corresponding derandomized version with the same approximation ratio. During the design and analysis for our algorithms, we carefully exploit the combined stochastic and fault-tolerant structure to obtain the desired results.

The rest of this paper is organized as followed. In Section 2, we describe the formulation for the 2-SFTFLP and the corresponding LP relaxation. In Section 3, we present a deterministic LP-rounding algorithm followed by 5 approximation ratio analysis. In Section 4, we offer a randomized LP-rounding 3.8617-approximation algorithm along with the corresponding derandomized version. Some discussions are given in Section 5.

Throughout the remainder of this paper, we use $[l]$ to denote the set $\{1, \ldots, l\}$ for any given positive integer l.

2. Formulation

In the 2-SFTFLP, we are given facility set F, client set C, and K scenarios $1, 2, \ldots, K,$ together with the corresponding clients set $C_1, C_2, \ldots, C_K \subseteq C$ and probability distributions $p_k (k \in [K])$ with $\sum_{k\in[K]} p_k = 1$. The open cost of facility $i \in F$ in the first stage is f_i^0 . The open cost of facility $i \in F$ in the second stage of scenario k is f_i^k $(k \in [K])$. Each client $j \in C_k$ has an integer connection requirement $r_j^k \le n := |F|$. Each client $j \in C_k$ has a nonnegative weight vector $w_j^k = \{w_j^{k(1)}, \ldots, w_j^{k(r_j^k)}\}$. Suppose that we are given a feasible solution consisting of a facility subset F_0 which is opened in the first stage, a facility subset F_k which is disjoint with F_0 and opened in the second stage of scenario k for each $k \in [K]$. The service cost for each client $j \in C_k$ is calculated as follows. Sort all the facilities in $F_0 \cup F_k$ according to the nondecreasing order of their distance to j and denote $F_0 \cup F_k$ as $\{i_1, i_2, \ldots, i_h\}$ with $h = |F_0 \cup F_k|$. The service cost of $j \in C_k$ is $\sum_{t\in[r_j^k]} w_j^{k(t)} d(i_t, j)$. We define the open cost and service cost for each scenario $k \in [K]$ as follows,

$$\text{cost}_k := \sum_{i\in F_k} f_i^k + \sum_{j\in C_k} \sum_{t\in[r_j^k]} w_j^{k(t)} d(i_t, j).$$

The objective is to find $K+1$ facility subsets $F_0, F_1, \ldots, F_K \subseteq F$ such that the total cost including the open cost of F_0 and the expectation of summation of the open cost and service cost over all scenarios, i.e.,

$$\sum_{i\in F_0} f_i^0 + \sum_{k\in[K]} p_k \cdot \text{cost}_k$$

is minimized.

Without loss of generality, we assume that only one entry of the vector $w_{(j,k)}^k$ is nonzero for each $j \in C_k$ and $k \in [K]$. In fact, we create r_j^k copies for each $j \in C_k$ and $k \in [K]$ with the first copy associated with the weight vector $\{w_j^{k(1)}, 0, \ldots, 0\}$, ..., and the r_j^k th copy associated with the weight vector $\{0, \ldots, 0, w_j^{k(r_j^k)}\}$. One can prove that this instance is equivalent to the original one. We refer to [19] for the similar discussion on this equivalence.

We use $\pi(j, k, t)$ to denote the facility in F which is the one that is tth closest to $j \in C_k$ for each $k \in [K]$. Let $N(j, k, t) := \{\pi(j, k, 1), \ldots, \pi(j, k, t)\}$ and $c_{jt}^k := d(j, \pi(j, k, t))$. Let $c_{j0}^k := 0$ and $c_{j(n+1)}^k := c_{jn}^k$ for all $k \in [K]$ and $j \in C_k$. We use indicator variable z_{jt}^k to denote the event whether client $j \in C_k$ is satisfied by $N(j, k, t)$, i.e., at least r_j^k facilities among $N(j, k, t)$ are opened. We further introduce binary decision variables x and y as follows. Let x_{ij}^{0k} (or x_{ij}^{kk}) denote whether client $j \in C_k$ is served by facility i in the first stage (or in the second stage of scenario k). Let y_i^0 (or y_i^k) denote whether a facility i is opened in the first stage (or in the second stage of scenario k).