Discrete Applied Mathematics 239 (2018) 193–199

Contents lists available at ScienceDirect

Discrete Applied Mathematics

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

Note

Improved complexity results for the robust mean absolute deviation problem on networks with linear vertex weights

Igor Averbakh [a,b,*], Oded Berman [b], Marina Leal [b,c]

[a] Department of Management, University of Toronto Scarborough, 1265 Military Trail, Toronto, Ontario M1C 1A4, Canada
[b] Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, Ontario M5S3E6, Canada
[c] Department of Statistics and Operations Research, Faculty of Mathematics, Campus Reina Mercedes, 41011, University of Seville, Seville, Spain

ARTICLE INFO

Article history:
Received 28 November 2016
Received in revised form 3 August 2017
Accepted 28 December 2017
Available online 1 February 2018

Keywords:
Minmax regret
Robust location
Facility location under uncertainty
Network algorithms

ABSTRACT

In a recent paper Lopez-de-los-Mozos et al. (2013), an algorithmic approach was presented for the robust (minmax regret) absolute deviation single-facility location problem on networks with node weights which are linear functions of an uncertain or dynamically changing parameter. The problem combines the mean absolute deviation criterion, which is the weighted average of the absolute deviations of individual customer–facility distances from the mean customer–facility distance and is one of the standard measures of "inequality" between the customers, with the minmax regret approach to optimization under uncertainty. The uncertain data are node weights (demands) which are assumed to change in a correlated manner being linear functions of a single uncertain parameter. The analysis in Lopez-de-los-Mozos et al. (2013) presented complexity bounds that are polynomial but too high to be of practical value. In this note, we present improvements of the analysis that significantly reduce the computational complexity bounds for the algorithm.

1. Introduction

In the last two decades, much attention has been focused on the minmax regret approach to optimization problems with data uncertainty (e.g., [2,3]), where it is required to find a solution with an objective function value reasonably close to the optimal one for all possible realizations of uncertain parameters. The literature on minmax regret facility location largely focuses on minmax regret versions of classical location objectives such as the median or center objectives. In [4], the minmax regret version of a location problem with an equity-based objective was considered. The mean absolute deviation is the weighted average of the absolute deviations of customer–facility distances from the mean distance between the facility and the customers, and is a measure of "inequality" between the customers. Lopez-de-los-Mozos et al. [4] consider this objective in the context of dynamic evolution (or linearly-correlated uncertainty) of node weights which are assumed to be linear functions of a changing parameter (time, for instance), and study the minmax and minmax regret versions of the problem. For the minmax regret version, they develop a polynomial time algorithm, which, in spite of being polynomial, cannot be considered practically useful due to very high order of complexity obtained by the analysis in [4] (slightly higher than $O(n^8 \log n)$ for trees and slightly higher than $O(m^3 n^8 \log n)$ for general networks, where n and m are the numbers of vertices and edges, respectively). In this paper, we present improvements of the analysis that reduce the complexity bounds to $O(m^2 n^3 (\log^* n)^2 \log n)$ for general networks, $O(n^4 (\log^* n)^2 \log n)$ for trees, and $O(n^3 (\log^* n)^2 \log n)$ for paths, where $\log^* n$ is

* Corresponding author at: Department of Management, University of Toronto Scarborough, 1265 Military Trail, Toronto, Ontario M1C 1A4, Canada.
E-mail addresses: averbakh@utsc.utoronto.ca (I. Averbakh), berman@rotman.utoronto.ca (O. Berman), mleal3@us.es (M. Leal).

https://doi.org/10.1016/j.dam.2017.12.036

the iterated logarithm which is a function that grows extremely slowly and can be considered as a constant for practical purposes.

To comply with space requirements of a note, we keep the presentation compact and technical. An expanded version of this paper, with additional intuitive explanations, is available in [1].

2. Notation and the problem

Consider a network $N(V, E)$ with the set of nodes $V = \{v_1, \ldots, v_n\}$ and the set E of m rectifiable edges. We denote by l_e the length of the edge $e \in E$, by N the set of all points of the network, or the network itself, and by $d(x, y)$ the shortest-path distance in the network between points x and y. We also use the following notation (consistent with [4]):

- $w_i(t) = \alpha_i t + \beta_i \geq 0$ is the weight of node v_i which is a linear function of a real-valued parameter t (time, for instance) that can take any value from an interval $[t^-, t^+]$, for some constants $t^-, t^+,\ t^- \leq t^+$;
- $W(t) = \sum_{i=1}^{n} w_i(t) = At + B$, where $A = \sum_{i=1}^{n} \alpha_i$, $B = \sum_{i=1}^{n} \beta_i$, is the total weight, which is assumed to be strictly positive for any $t \in [t^-, t^+]$;
- $M(x, t) = \frac{1}{W(t)} \sum_{i=1}^{n} w_i(t) d(v_i, x)$ is the dynamic median function, for $t \in [t^-, t^+]$ and $x \in N$;
- $F(x, t) = \frac{1}{W(t)} \sum_{i=1}^{n} w_i(t) |d(v_i, x) - M(x, t)|$ is the dynamic mean absolute deviation (MAD) function;
- $F^*(t) = \min_{x \in N} F(x, t)$.

Following [4], the minmax regret mean absolute deviation location problem (MMR-MAD) is

$$\min_{x \in N} \max_{t \in [t^-, t^+]} \left(F(x, t) - F^*(t) \right).$$

MMR-MAD is the minmax regret version of the problem of finding a location x that minimizes the mean absolute deviation $F(x, t)$ considering t as uncertain parameter that can take values in $[t^-, t^+]$.

For any integer i, j, $i \leq j$, let $[i : j]$ denote the set $\{i, i+1, \ldots, j\}$. For the interpretation of the standard notation $\log^* n$, $\lambda_s(n)$ and $\alpha(n)$ from computational geometry and some related facts that we use for complexity estimation, see Appendix.

Lopez-de-los-Mozos et al. [4] present an algorithm for MMR-MAD and analyze its complexity. Their analysis obtains the complexity bound $O(mn^3 \lambda_6(m^2 n^5) \log^* n \log n)$ (even though stated as $O(mn^3 \lambda_6(n^5) \log^* n \log n)$ in [4], as we will discuss in Section 5) for general networks, which is slightly higher than $O(m^3 n^8 \log n)$, and the bound $O(n^3 \lambda_6(n^5) \log^* n \log n)$ for trees which is slightly higher than $O(n^8 \log n)$. These orders of complexity, although polynomial, are clearly of little practical value, and the purpose of this paper is to improve them.

To facilitate reading the paper, in the next section we give an informal sketch of the approach of [4] with minimum notation and details. The purpose of this is to outline a "big picture" and some ideas without delving into technicalities. For further details and explanations we refer the reader to [4] or to our more detailed description of the algorithm and analysis of [4] in [1]. In Sections 4 and 5 we present our improvements. In Section 6, we make some concluding remarks.

3. A sketch of the ideas of the approach of [4]

The logic of the algorithm from [4] for MMR-MAD is essentially the same for a tree and for a general network. On a tree, the problem restricted to a single edge is considered; on a general network, each edge is partitioned into $O(n)$ primary regions where all functions $d(v_i, x)$, $i \in [1 : n]$ are linear, and the problem restricted to a primary region of an edge is considered. Then the best of the optimal solutions for different edges/primary regions is chosen.

Suppose that N is a tree. Consider the function $F(x, t)$ in the domain D_e which is the Cartesian product of an edge e of the tree N and the interval $[t^-, t^+]$. In the interior of D_e, function $F(x, t)$ is differentiable everywhere except the points of at most n breakpoint curves $\mathcal{B}_{e,i}$, $i \in [1 : n]$ that satisfy $d(v_i, x) - M(x, t) = 0$. The breakpoint curves are hyperbolas or straight lines and partition the domain D_e into full-dimensional cells inside which $F(x, t)$ is differentiable. This partition has $O(n^2)$ cells and $O(n^2)$ vertices (intersection points of the breakpoint curves with other breakpoint curves or the boundary of D_e) and is denoted $P(D_e)$. In the paper, we use the term "vertices" for vertices of geometric arrangements, while the term "nodes" is used for the nodes of N.

In Phase 1 of the algorithm from [4], function $F^*(t)$ is obtained. For any fixed $t \in [t^-, t^+]$ the minimum of $F(x, t)$ over all $x \in N$ is attained at some breakpoint curve or at a node of N, since $F(x, t)$ as a function of x is piece-wise linear convex on e with breakpoints at the breakpoint curves. Hence, for obtaining the function $F^*(t)$, function $F(x, t)$ needs to be considered only for the points (x, t) of the breakpoint curves or points that correspond to $x \in V$.

The t-coordinates of all vertices of all partitions $P(D_e)$, $e \in E$ partition the interval $[t^-, t^+]$ into $O(n^3)$ *basic subintervals*. When (x, t) moves along a breakpoint curve $\mathcal{B}_{e,i}$, $F(x, t)$ can be viewed as a function of only t which has an analytic expression (quotient of two polynomials) that does not change inside any basic subinterval but may be different for different basic subintervals. Considering this function over each basic subinterval as a separate function, Lopez-de-los-Mozos et al. [4] obtain overall $O(n^5)$ simple functions of t (partially defined over $[t^-, t^+]$), as there are $n-1$ domains D_e, $e \in E$, each domain has $O(n)$ breakpoint curves, and $F(x, t)$ over each breakpoint curve is split into $O(n^3)$ pieces that correspond to the basic subintervals. (By "simple functions" we mean differentiable functions, rational or algebraic, each having a single analytic expression over its domain, such that the number of intersection points between any two of them is bounded by a constant.) Then, they