Information Processing Letters 108 (2008) 175–178

Contents lists available at ScienceDirect

Information Processing Letters

www.elsevier.com/locate/ipl

Distortion lower bounds for line embeddings

Claire Mathieu, Charalampos Papamanthou *

Department of Computer Science, Brown University, P.O. Box 1910, Providence, RI, USA

ARTICLE INFO

Article history:
Received 24 January 2008
Received in revised form 6 May 2008
Available online 20 May 2008
Communicated by C. Scheideler

Keywords:
Combinatorial problems
Theory of computation

ABSTRACT

In this paper, we show how we can derive lower bounds and also compute the exact distortion for the line embeddings of some special metrics, especially trees and graphs with certain structure. Using linear programming to formulate a simpler version of the problem gives an interesting intuition and direction concerning the computation of general lower bounds for distortion into the line. We also show that our lower bounds on special cases of metrics are a lot better than previous lower bounds.

1. Introduction

Embedding a metric induced by an unweighted graph into the line with minimum distortion is a well-known NP-hard problem [4]. Approximation algorithms and hardness results are discussed in [3,4,8]. Improved algorithms for optimal embeddings are presented in [6], where, however, the embedding is required to be a bijection. In [1], embeddings of series-parallel graphs into ℓ_1 and embeddings between two line metrics are discussed. In this paper, we show how we can derive lower bounds for the line distortion of some special metrics, especially trees and graphs with certain structure. To do this, we use linear programming to formulate another simpler version of the problem, where we know the permutation of the vertices of the graph on the line in advance. This approach gives an interesting intuition and direction concerning the computation of general lower bounds for distortion into the line.

Problem. Let $G = (V, E)$ be an unweighted graph with the implied shortest path metric $d(u, v)$. A map f of V into the real line is *non-contracting* if $|f(u) - f(v)| \geqslant d(u, v)$ for all u, v. The *minimum distortion problem* is to find a map f with minimum distortion, where the distortion of a non-contracting map f is defined as $\max_{u,v \in V} |f(u) - f(v)|/d(u, v)$.

Previous work. There are some results on lower bounds concerning embeddings into higher dimensions. For example you can embed the complete binary tree in the Euclidean space with distortion $\geqslant \Omega(\sqrt{\lg \lg n})$ [5], where n is the size of the tree. A short proof for that is also presented in [11]. Also upper bounds for embedding the binary tree into Euclidean space are presented in [5]. Additionally, in [2], lower bounds of embeddings of trees into $\Re^2$ are discussed. Lower bounds of low distortion embeddings to higher dimensional spaces are also presented in [5].

However, there are few results for the computation of lower bounds concerning embeddings into the line. Actually, in [4,9], it is proved that the least distortion required to embed the star into the line is $\Omega(n)$. Also, in [3] a lower bound method that applies to all unweighted graphs is presented.

In [3], an $O(\sqrt{n})$ approximation algorithm for computing embeddings of unweighted graphs (n is the size of the graph) into the line is described. This algorithm partitions the nodes of the graph into distinct clusters and explicitly computes the intervals between the nodes on the line.

Define the local density of an unweighted graph $G = (V, E)$ to be the quantity Δ such that

$$\Delta = \max_{v \in V, r \in \Re_{>0}} \left\{ \frac{|B(v, r)| - 1}{2r} \right\}$$

* Corresponding author.
E-mail addresses: claire@cs.brown.edu (C. Mathieu), cpap@cs.brown.edu (C. Papamanthou).

0020-0190/$ – see front matter
doi:10.1016/j.ipl.2008.05.003

where $B(v,r) = \{u \in V : d(u,v) \leqslant r\}$. In [3], the following general lower bound on the distortion of the embedding of an unweighted graph is presented:

Theorem 1. *(See [3].) Let $G = (V, E)$ be an unweighted graph with local density Δ. Then, any embedding of G into the line has distortion more or equal than Δ.*

The local density was also extensively used and studied in the context of bandwidth (see for example [7]).

Our results. We present some unweighted graph cases for which we are able to prove exact distortion based on lower bounds produced by using a linear programming formulation. We compute the optimal distortion by computing an embedding of a distortion c and then by proving that c is also a lower bound on the optimal distortion of these graph instances. More specifically we compute the optimal distortion for a structured tree (we call it two-path tree) and the $n \times j$ square grid. We apply our results to a star metric, giving a better lower bound and finally we set up a conjecture for the binary tree.

2. An explicit formula for the minimum distortion

Theorem 2. *Given an unweighted graph $G = (V, E)$ then the minimum distortion embedding of G into the line has distortion equal to*

$$c^* = \min_{\pi : V \to \{1,\ldots,n\}} \left\{ \max_{\{u,v\} \in E} \left\{ \sum_{k=\pi(u)}^{\pi(v)-1} r_k^\pi \right\} \right\}$$

where $r_i^\pi = d(\pi^{-1}(i), \pi^{-1}(i+1))$.

We start the proof with two simple observations.

Lemma 3. *The distortion of a map f that maps the vertices of an unweighted graph $G = (V, E)$ into the line equals $\max_{\{u,v\} \in E} |f(u) - f(v)|$.*

Proof. For $u, v \in V$, let $u = y_1, y_2, \ldots, y_k = v$ be a shortest path from u to v. Then

$$\begin{aligned}\bigl|f(u) - f(v)\bigr| &\leqslant \sum_{j=1}^{k-1} \bigl|f(y_j) - f(y_{j+1})\bigr| \\ &\leqslant c(k-1) = cd(u,v)\end{aligned}$$

where $c = \max_{\{u,v\} \in E} |f(u) - f(v)|$. Since for every edge $\{u,v\} \in E$ it is $d(u,v) = 1$, the result follows. $\square$

Lemma 4. *The map f is non-contracting if and only if $f(y_{i+1}) - f(y_i) \geqslant d(y_{i+1}, y_i)$ for all i, where $f(y_i)$ is a labelling of the vertices in order of their embedding on the line by f, i.e., $f(y_1) < f(y_2) < \cdots < f(y_n)$.*

Proof. For $u, v \in V$, let i, j be such that $u = y_j$ and $v = y_i$, and without loss of generality $i < j$. Then

$$\begin{aligned}f(u) - f(v) &= \sum_{k=i}^{j-1} \bigl(f(y_{k+1}) - f(y_k)\bigr) \\ &\geqslant \sum_{k=i}^{j-1} d(y_{k+1}, y_k) \geqslant d(y_j, y_i) = d(u,v).\end{aligned}$$

The inverse holds trivially by the definition of a non-contracting embedding and the hypothesis. $\square$

Combining the two lemmas, we can write a linear program to compute the minimum distortion c of a non-contracting map, given the permutation π such that $\pi(u)$ is the rank of u in the embedding. Here x_i is the distance on the real line between the ith point and the $(i+1)$th point of the embedding.

$$\min(c) \quad \text{s.t.} \quad \begin{cases} \sum_{k=\pi(u)}^{\pi(v)-1} x_k \leqslant c & \forall \{u,v\} \in E, \\ x_i \geqslant d(\pi^{-1}(i), \pi^{-1}(i+1)) & \forall i, \\ x_i, & c \geqslant 0. \end{cases}$$

Recall that $r_i^\pi = d(\pi^{-1}(i), \pi^{-1}(i+1))$. It is easy to see that the minimum is reached when $x_i = r_i^\pi$, and given that, that the minimum is obtained for $c = \max_{\{u,v\} \in E} \{\sum_{k=\pi(u)}^{\pi(v)-1} r_k^\pi\}$. Minimizing over all permutations π yields the proof. $\square$

Theorem 5. *Let G be an unweighted graph. Let v_1, v_2, v_3 be three vertices, and p, p', p'' be three paths from v_1 to v_3, from v_3 to v_2 and from v_2 to v_1 respectively. Then*

$$c^* \geqslant 2 \min\Bigl\{ \min_{u \in p} \bigl\{ d(v_2, u) \bigr\}, \min_{u \in p'} \bigl\{ d(v_1, u) \bigr\}, \min_{u \in p''} \bigl\{ d(v_3, u) \bigr\} \Bigr\}$$

where c^* is given in Theorem 2.

Proof. Let t denote the minimum of the three quantities. Let π be an arbitrary permutation and f be the optimal embedding respecting π. First assume that in π, v_2 is between v_1 and v_3, and so, $f(v_2)$ is between $f(v_1)$ and $f(v_3)$. Since p is a path from v_1 to v_3 in G, there must exist an edge (x, x') of p such that $f(v_2)$ is between $f(x)$ and $f(x')$. Then, since f is non-contracting:

$$\begin{aligned}d(x', v_2) + d(v_2, x) &\leqslant \bigl(f(x') - f(v_2)\bigr) + \bigl(f(v_2) - f(x)\bigr) \\ &= f(x') - f(x) = \sum_{k=\pi(x)}^{\pi(x')-1} r_k^\pi.\end{aligned}$$

Here the last equality comes from the fact that for f optimal, the second set of inequalities of the linear program are tight. Therefore for edge (x, x')

$$\sum_{k=\pi(x)}^{\pi(x')-1} r_k^\pi \geqslant d(x', v_2) + d(v_2, x) \tag{1}$$

which can be more simplified into (decreasing the lower bound a little bit) $\sum_{k=\pi(x)}^{\pi(x')-1} r_k^\pi \geqslant 2t$. The other cases are similar. Thus, for every permutation π, there exists an edge (x, x') such that $\sum_{k=\pi(x)}^{\pi(x')-1} r_k^\pi \geqslant 2t$. By Theorem 2 this implies $c^* \geqslant 2t$. $\square$

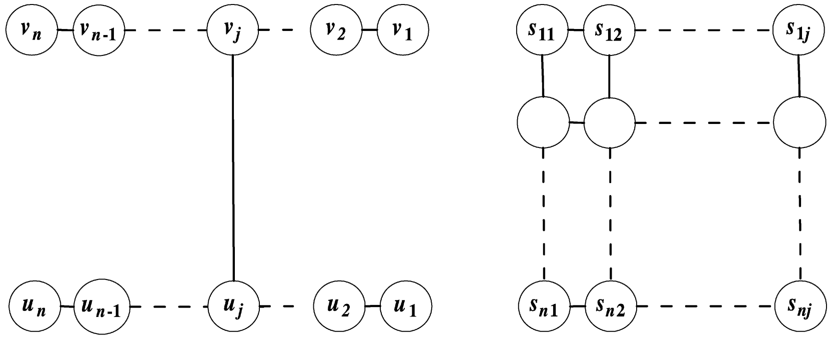

Fig. 1. A two-path tree $T(n,j)$ (left) and the grid $S(n,j)$ (right).

3. Examples

3.1. The two-path tree

Theorem 6. *A two-path tree $T(n,j)$ is an unweighted graph which is the tree consisting of two paths $V = v_1, v_2, \dots, v_n$ and $U = u_1, u_2, \dots, u_n$ and an edge $\{v_j, u_j\}$ such that $j \leqslant (n+1)/2$. Let $c^*(T(n,j))$ denote the optimal distortion into the line of a 2-path tree $T(n,j)$. Then*:

$$2j-2 \leqslant c^*\big(T(n,j)\big) \leqslant 2j-1.$$

Proof. To prove the lower bound, use Theorem 5 for the nodes v_1, u_1, v_n to yield $c^* \geqslant 2\min\{j-1, j, n-j\} = 2(j-1)$. Also we can get an even better lower bound equal to $2j-1$, by using directly Eq. (1). To prove the upper bound, embed the nodes in the following order (see Fig. 1):

$$P(n,j) = v_n(\mathbf{1})v_{n-1}(\mathbf{1})\dots(\mathbf{1})v_1(\mathbf{j})u_j(\mathbf{1}) \\ \times u_{j-1}(\mathbf{1})\dots(\mathbf{1})u_1(\mathbf{j})u_{j+1}(\mathbf{1})u_{j+2}(\mathbf{1})\dots(\mathbf{1})u_n$$

where $x(\mathbf{d})y$ means that node y is embedded after node x at distance d. Note that the embedding is non-contracting and also

$$\frac{|f(u_j)-f(v_j)|}{d(u_j,v_j)} = \frac{2j-1}{1} = 2j-1.$$

Hence $c^*(T(n,j)) \leqslant 2j-1$ and the presented embedding is optimal. $\square$

Note that Theorem 1 would give a lower bound of $\Omega(1)$ only, whereas our lower bound is essentially tight.

3.2. The rectangular grid

Theorem 7. *Let $S(n,j)$, $n \geqslant j$ denote the unweighted graph with nj nodes consisting of the $n \times j$ rectangular grid. Let $c^*(S(n,j))$ denote the optimal distortion. Then*:

$$2j-2 \leqslant c^*\big(S(n,j)\big) \leqslant 2j-1.$$

Proof. Each node of the grid is denoted s_{kl}, $i = 1,\dots,n$, $l = 1,\dots,j$. For the lower bound, use Theorem 5 with nodes s_{11}, s_{n1}, s_{nj}. The paths we are using are

$$s_{11}, s_{21}, \dots, s_{n1},$$

$$s_{n1}, s_{n2}, \dots, s_{nj},$$

$$s_{nj}, s_{(n-1)j}, \dots, s_{1j}, s_{1(j-1)}, \dots, s_{11}.$$

By using again Eq. (1) we can easily get the best lower bound equal to $2j-1$. For the upper bound, consider the embedding (see Fig. 1) represented with the following permutation (without loss of generality we assume n is odd):

$$p = s_{11}, s_{12}, \dots, s_{1j}, s_{2j}, s_{2(j-1)}, \dots, \\ s_{21}, \dots, s_{nj}, s_{n(j-1)}, \dots, s_{n1}.$$

For this permutation π, by the linear programming formulation we have that

$$c^*\big(S(n,j)\big) = \max_{\{u,v\}\in E}\left\{\sum_{k=\pi(u)}^{\pi(v)-1} r_k^\pi\right\} \\ = \sum_{k=\pi(s_{11})}^{\pi(s_{21})-1} r_k^\pi = \sum_{k=\pi(s_{(n-1)1})}^{\pi(s_{n1})-1} r_k^\pi = 2j-1.$$

Hence the optimal distortion is $2j-1$. $\square$

3.3. Star metrics

Consider a star metric that consists of a star with $O(1)$ branches of length $O(n)$ with the following properties: Then, our lower bound is $\Omega(n)$, hence tight, whereas the lower bound by [3] is $\Omega(1)$ (there is no ball of constant radius that contains $O(n)$ nodes).

3.4. The complete binary tree

Next, we give some thoughts that may help in order to prove some bounds for the optimal distortion of the complete binary tree.

Corollary 8. *The complete binary tree of n nodes and height ℓ cannot be embedded in the line with a distortion less than $\frac{n-1}{2(\lg(n+1)-1)} = \frac{2^\ell-1}{\ell} = \Omega(\frac{n}{\lg n}) = \Omega(\frac{2^\ell}{\ell})$.*

Proof. Follows from Theorem 1 if we choose the root as the center of the ball. $\square$

Also, we can easily prove the following:

Theorem 9. *Let T be a complete binary tree of height ℓ. Then, an inorder traversal of the tree gives a non-contracting embedding of distortion* $2^\ell - 1$.

Finally we believe that the following conjecture holds:

Conjecture 10. *The complete binary tree can be embedded into the line with distortion* $\Theta(2^\ell)$.

Optimal embeddings of a complete binary tree under a different setting though (where non-contractness is not an issue and also expansion is allowed) are presented in [10].

4. Concluding remarks

In this paper we propose a technique to compute good lower bounds for the optimal distortion into the line of certain metrics. We give better lower bounds than the ones derived when using an already proposed general lower bound. To do this we use a linear programming formulation that gives the optimal solution when the optimal permutation of the embedding is known. We give explicit constructions for various metrics, such as trees, grids and star metrics. Also, we propose an interesting conjecture. Finally we note that it might be possible to resolve the proposed conjecture by taking some ideas from [10] in order to compute an optimal upper bound.

Acknowledgements

The authors would like to thank Aris Anagnostopoulos and Eli Upfal for their useful feedback at the early stages of this work.

References

[1] M. Ali Safari, D-width, metric embedding, and their connections, PhD Thesis, University of British Columbia, 2007.

[2] R. Babilon, J. Matousek, J. Maxova, P. Valtr, Low-distortion embeddings of trees, Journal of Graph Algorithms and Applications 7 (4) (2003) 399–409.

[3] M. Badoiu, K. Dhamdhere, A. Gupta, Y. Rabinovich, H. Raecke, R. Ravi, A. Sidiropoulos, Approximation algorithms for low-distortion embeddings into low-dimensional spaces, in: Proc. of the ACM–SIAM Symposium on Discrete Algorithms (SODA), 2005, pp. 119–129.

[4] M. Badoiu, J. Chuzhoy, P. Indyk, A. Sidiropoulos, Low-distortion embeddings of general metrics into the line, in: Proc. of the ACM Symposium on Theory of Computing (STOC), 2005, pp. 225–233.

[5] J. Bourgain, On Lipschitz embeddings of finite metric spaces in Hilbert space, Israel Journal of Mathematics 52 (1–2) (1985) 46–52.

[6] N. Chandran, R. Moriarty, R. Ostrovsky, O. Pandey, A. Sahai, Improved algorithms for optimal embeddings, in: Electronic Colloquium on Computational Complexity, ECCC Report TR06-110, 2006.

[7] F.R.K. Chung, P.D. Seymour, Graphs with small bandwidth and cutwidth, Discrete Mathematics 75 (1–3) (1989) 113–119.

[8] W. Evans, M. Ali Safari, *k*-separable permutations and their applications to metric embedding, pattern matching, and stack sorting, Personal communication, 2005.

[9] A. Gupta, Embedding tree metrics into low dimensional Euclidean spaces, in: Proc. of the ACM Symposium on Theory of Computing (STOC), 1999, pp. 694–700.

[10] R. Heckmann, R. Klasing, B. Monien, W. Unger, Optimal embedding of complete binary trees into lines and grids, Parallel and Distributed Computing 49 (1) (1998) 40–56.

[11] N. Linial, M. Saks, The Euclidean distortion of complete binary trees, Discrete & Computational Geometry 29 (1) (2003) 19–21.