EMBEDDING INTO TREES: GUPTA'S RESTRICTION THEOREM

ELIOT MARTIN

Abstract. Given a weighted tree and a subset of its vertices, we wonder whether it is possible to define a weighted subtree on the chosen vertices with a metric close to that of the original tree. Gupta's Theorem answers this question in the affirmative and gives an explicit linear time algorithm. We give applications to the embedding of cycles and large girth graphs into trees.

1. Introduction

Given a weighted tree $T = (V, E, w)$ and a subset of vertices $V' \subseteq V$, one natural question is the following: does there exist a tree T' on V' and constants $c > 0$, $D > 1$ such that for any two vertices $x, y \in V'$, one has

$$c\, d_T(x, y) \leqslant d_{T'(x,y)} \leqslant cD\, d_T(x, y)$$

with $d_T, d_{T'}$ the metrics induced by the graph structure? The set $V \setminus V'$ is sometimes called the set of *Steiner nodes* after the Steiner tree problems, see [AA92] for an introduction. The restriction theorem of Gupta says that removing Steiner nodes preserves the metric up to a small universal distortion.

Theorem 1.1 ([Gup01, Theorem 1.1]). *Let $T = (V, E, w)$ be a weighted tree and $V' \subseteq V$. Then there exists $E' \subseteq \binom{V'}{2}$ and $w' : E' \to [0, +\infty)$ such that if $T' = (V', E', w')$, we have for all $x, y \in V$,*

$$\frac{1}{4} \leqslant \frac{d_{T'}(x, y)}{d_T(x, y)} \leqslant 2.$$

Let us mention a consequence of Theorem 1.1. A graph might not embed with low distortion in a single given tree, but it may embed well on average in different trees, this is contained in the following definition.

Definition 1.2. A metric space (X, d) embeds with distortion $D > 1$ in a *distribution of dominating trees* if there exist a finite family of trees $(T_i)_{i\in I}$, embeddings $f_i : X \to T_i$ and positive real numbers $(p_i)_{i\in I}$ verifying $\sum_{i\in I} p_i = 1$, such that for all $x, y \in X$,

- $d(x, y) \leqslant d_{T_i}(f_i(x), f_i(y))$,
- $d(x, y) \geqslant \frac{1}{D} \sum_{i\in I} p_i d_{T_i}(f_i(x), f_i(y))$.

The embedding is said to be bijective if all the f_i are bijections.

Combining Theorem 1.1 with a result of Fakcharoenphol, Rao and Talwar [FRT03], we get the following.

Corollary 1.3. *There exists a universal constant $c > 0$ such that any n-point metric space embeds bijectively into a distribution of dominating trees with distortion $c \log n$.*

In what follows, all considered graphs are non-oriented, finite, connected, without loops or multiple edges. An edge-weighted graph $X = (V, E, w)$ is naturally endowed with a path-length metric defined as follows: for every $x, y \in V$, set

$$d_X(x, y) = \min \left\{ \sum_{i=1}^{n-1} w(\{x_i, x_{i+1}\}) \,|\, (x_i)_{1 \leqslant i \leqslant n} \text{ is a path from } x \text{ to } y. \right\}.$$

A map $f : (X, d_X) \to (Y, d_Y)$ between metric spaces is a *bi-Lipschitz embedding* if there exist constants $c > 0$ and $D > 1$ such that for all $x, y \in X$,

$$c\, d_X(x, y) \leqslant d_Y(f(x), f(y)) \leqslant cD\, d_X(x, y).$$

The least such $D > 1$ is called the *distortion* of f, denoted $\mathrm{dist}(f)$ and is equal to $\|f\|_{\mathrm{Lip}}\|f^{-1}\|_{\mathrm{Lip}}$.

Given a weighted tree $T = (V, E, w)$ rooted at $r \in V$ and a vertex $u \in V$, $v \in V$ is a *descendant* of u if the unique geodesic from r to v passes through u. Note that in particular, u is a descendant of itself. Conversely, a vertex $u \in V$ is the *ancestor* of v if v is a descendant of u. The *subtree of $u \in V$* is the subtree induced by all descendants of u.

Let $L(T)$ denote the set of leaves of T. Up to modifying the weight function to take into account a lexicographic order induced from the root, we can suppose that all distances in the graph are distinct. Then, given a vertex $v \in V$, we define $C(v)$ to be the closest vertex in $L(T)$, and call *fringe distance of v* the quantity $h(v) = d_T(v, C(v))$. In particular, $h(v) = 0$ if and only if v is a leaf.

The rest of this document is organized as follows. Next section deals with the case where the imposed vertices are leaves of the tree, Section 3 gives a proof of Theorem 1.1, finally, the last section outlines a few applications to the embedding of graphs into trees and to a communication network problem.

2. Subtree on the set of leaves

We begin by dealing with the case where the imposed subset of vertices is a subset of the leaves, from which the general case will be deduced.

Lemma 2.1. *Given a weighted tree $T = (V, E, w)$, we can define a tree T' on the set of leaves $L(T)$ of T such that for any vertices $x, y \in L(T)$,*

$$\frac{1}{4} \leqslant \frac{d_{T'}(x, y)}{d_T(x, y)} \leqslant 2.$$

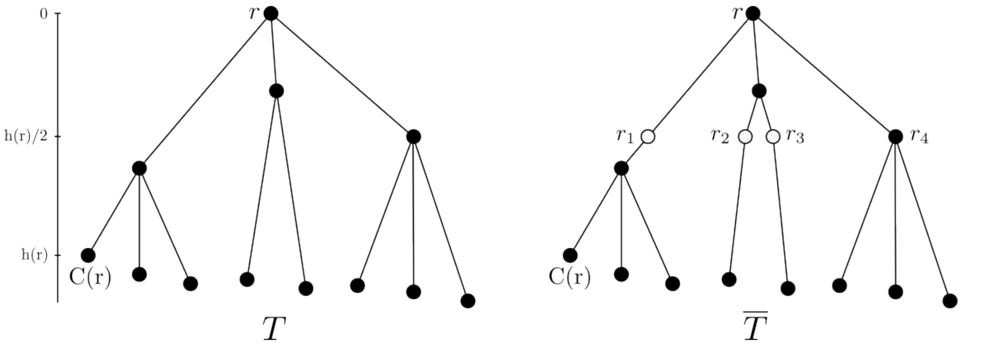

FIGURE 1. Tree splitting step

Proof of Lemma 2.1. If T is not reduced to a single edge in which case the result is straightforward, let us choose an arbitrary root r of T in the vertices of $T \setminus L(T)$. We give an algorithm to remove Steiner nodes $T \setminus L(T)$ recursively an produce a "cleaned" tree T^* whose distances are close to those in T.

If T has only one required vertex, *i.e.*, it is a path with a required vertex at its end, then T^* is set to be the isolated vertex $C(r)$. Otherwise, let the root be a vertex r of degree at least two and consider the set

$$R = \{x \in V \mid d_T(x, r) = h(r)/2\}$$

where T is seen here as its geometric realization with edge lengths given by the weight function w. This means that some of these points can lie on edges of T, in which case we split the corresponding edges to add a vertex at distance exactly $h(r)/2$ from r to obtain a tree $\overline{T}$ in which T embeds isometrically, see Figure 1. Denote $R = \{r_1, \ldots, r_k\}$ and for each $i \in \{1, \ldots, k\}$, let T_i be the subtree of r_i with the induced weight function w_i. Thus, $\overline{T}$ also contains isometrically each T_i, $i \in \{1, \ldots, k\}$. Without loss of generality, assume $C(r)$ lies in T_1, then $C(r) = C(r_1)$ and for all $j \in \{2, \ldots, k\}$, we have $h(r_i) \geqslant h(r_1)$.

We continue this decomposition procedure until the trees T_i consist of a single vertex. We claim the following.

Claim 1. Suppose that for all $i \in \{1, \ldots, k\}$, there is a tree $T'_i = (V'_i, E'_i, w'_i)$ on $V'_i = L(T_i)$ the set of leaves of T_i such that for all $x, y \in V'_i$,

$$d_{T'_i}(x, C(r_i)) \leqslant 2d_{T_i}(x, r_i) - h(r_i) \text{ and } \frac{1}{4} \leqslant \frac{d_{T'_j}(x, y)}{d_{T_j}(x, y)} \leqslant 2.$$

Let $T' = (V', E', w')$ be the tree obtained from the disjoint union $\bigsqcup_{i=1}^{k} T'_i$ by adding an edge between r_1 and r_i for all $i \in \{2, \dots, k\}$, with the weight function

$$w'(e) = \begin{cases} w_i(e) & \text{if } e \in V'_i \\ h(r_i) & \text{if } e = \{r_1, r_i\}. \end{cases}$$

Then, for all $x, y \in V'$,

$$d_{T'}(x, C(r)) \leqslant 2d_T(x, r) - h(r) \tag{1}$$

$$\frac{1}{4} \leqslant \frac{d_{T'}(x, y)}{d_T(x, y)} \leqslant 2. \tag{2}$$

Since this claim clearly holds when $L(T)$ is a singleton, the lemma follows by applying the claim inductively, starting from the end of the previous procedure where subtrees are reduced to single vertex graphs and going back up to the original tree, applying Claim 1 at each step.

Let us now prove the claim. We begin by showing that (1) holds under the provided assumptions. Let $x \in V' = \bigsqcup_{i=1}^{k} L(T_i)$ and $i \in \{1, \dots, n\}$ such that $x \in L(T_i)$.

Case 1: $i = 1$

Then

$$\begin{aligned} d_{T'}(x, C(r_1)) = d_{T'_1}(x, C(r_1)) &\leqslant 2d_{T_1}(x, r_1) - h(r_1) \text{ by assumption} \\ &= 2\left(d_T(x, r) - h(r)/2\right) - h(r)/2 \\ &\leqslant 2d_T(x, r) - h(r). \end{aligned}$$

Case 2: $i > 1$

$$\begin{aligned} d_{T'}(x, C(r_1)) &= d_{T'_i}(x, C(r_i)) + d_{T'}(C(r_i), C(r_1)) \\ &\leqslant 2d_{T_i}(x, r_i) - h(r_i) + h(r_i) \\ &= 2d_{T_i}(x, r_i) \\ &= 2\left(d_T(x, r_i) - h(r)/2\right) = 2d_T(x, r_i) - h(r). \end{aligned}$$

We now verify (2). For $x, y \in V'$, if there is some index $i \in \{1, \dots, k\}$ such that $x, y \in V'_i$, then (2) follows from the assumptions of Claim 1. Let us assume that $x \in V_i$

and $y \in V_j$ for $i \neq j$ and without loss of generality $j \neq 1$. We have

$$\begin{aligned} d_T(x,y) &\leqslant d_T(x, C(r_i)) + d_T(C(r_i), r) + d_T(r, C(r_j)) + d_T(C(r_j), y) \\ &\leqslant 4\left(d_{T'}(x, C(r_i)) + d_{T'}(y, C(r_j))\right) + d_{T_i}(r_i, C(r_i)) + d_{T_j}(r_j, C(r_j)) + d_{\overline{T}}(r_i, r) + d_{\overline{T}}(r_j, r) \\ &= 4\left(d_{T'}(x, C(r_i)) + d_{T'}(y, C(r_j))\right) + h(r_i) + h(r_j) + h(r) \\ &\leqslant 2\left(d_{T'}(r,x) + d_{T'}(r,y)\right) + h(r_i) + h(r_j) + h(r) \\ &= 2d_T(x,y) + h(r_i) + h(r_j) + h(r). \end{aligned}$$

Now, if $i = 1$, we have

$$h(r_i) + h(r_j) + h(r) = 3h(r_1) + h(r_j) \leqslant d_{T'}(r,x) + d_{T'}(r,y) = d_{T'}(x,y),$$

otherwise,

$$h(r_i) + h(r_j) + h(r) \leqslant 2h(r_i) + 2h(r_j) \leqslant d_{T'}(x,y).$$

In both cases,

$$d_T(x,y) \leqslant 4d_{T'}(x,y).$$

This gives the first inequality of (2). For the second one, if $x, y \in V_i$ for some $i \in \{1, \ldots, k\}$, the result is again straightforward from the hypotheses, so suppose that $x \in V_i$, $y \in V_j$ for $i \neq j$. We have

$$\begin{aligned} d_{T'}(x,y) &= d_{T'_i}(x, C(r_i)) + d_{T'}(C(r_i), C(r_j)) + d_{T'_j}(C(r_j), y) \\ &\leqslant d_{T'_i}(x, C(r_i)) + d_{T'_j}(y, C(r_j)) + h(r_i) + h(r_j) \\ &\leqslant 2d_{T_i}(x, r_i) - h(r_i) + 2d_{T_j}(y, r_j) - h(r_j) + h(r_i) + h(r_j) \\ &\leqslant 2d_T(x,y). \end{aligned}$$

Which concludes the proof of Claim 1 and of the Lemma. □

3. Proof of Gupta's restriction Theorem

Proof of Theorem 1.1. We first show that the set of required vertices R can be supposed to contain all leaves of $T = (V, E, w)$, then the proof is done by induction on the cardinality of $R \setminus L(T)$, the base case $R = L(T)$ being given by Lemma 2.1.

Indeed, if the set of leaves $L(T)$ is not included in the set of required vertices R, the edge incident to any leaf not in R cannot be used in any shortest path between required vertices, thus these edges can be deleted until $L(T) \subseteq R$.

Now, choose $r \in R \setminus L(T)$ as the root of T. Let $n \geqslant 2$ be the degree of r and $r_1, \ldots, r_n$ be its neighbors. For $i \in \{1, \ldots, n\}$, let $T_i = (V_i, E_i, w_i)$ be the subtree of r_i with an additional vertex being a copy of r, an additional edge $\{r, r_i\}$ and with weight function induced from w. In particular, r is a leaf of T_i. Now set $R_i = R \cap V_i$, then $|R_i \setminus L(T_i)| < |R \setminus L(T)|$

and we can apply the induction hypothesis. This yields for all $i \in \{1, \dots, k\}$ a tree $T'_i = (V'_i, E'_i, w'_i)$ such that for all $x, y \in V'_i$,

$$\frac{1}{4} \leqslant \frac{d_{T'_i}(x, y)}{d_{T_i}(x, y)} \leqslant 2.$$

Gluing together the images of r in each T'_i produces a tree $T' = (V', E', w')$. Let us show that the global distance $d_{T'}$ respects the required distortion bound. If $x, y \in V'$, identified to their preimage before the gluing, belong to the same V'_i, then the result is given by the induction hypothesis. If x and y belong respectively to V_i and V_j, with $i \neq j$, we have

$$\begin{aligned} d_{T'}(x, y) &= d_{T'}(x, r) + d_{T'}(r, y) \\ &\leqslant 2d_{T_i}(x, r) + 2d_{T_j}(y, r) = 2d_T(x, y). \end{aligned}$$

That is, the expansion of the embedding $T' \to T$ is bounded above by the maximal expansion over embeddings $T'_i \to T_i$, $i \in \{1, \dots, n\}$. Similarly, we get a lower bound on the contraction and the distortion is bounded by the maximal distortion over subtrees, which yields the result as each subtree satisfies the distortion bound. □

4. Applications

4.1. Embeddings of graphs with large girth into trees. For $n \geqslant 3$, let $C_n = (V_n, E_n)$ be the cycle on n vertices with unit edge weight and d_n be the associated metric. We begin by studying the distortion of d when embedding the cycle into a tree on the same vertex set. As an application, we show that the distortion of the embedding of a graph X into a tree grows linearly with the girth of X.

Lemma 4.1. *Let T be a tree on the vertex set V_n and d_T be the associated metric. If $f : (C_n, d_n) \to (T, d_T)$ is the identity map on vertices, then* $\operatorname{dist}(f) \geqslant n - 1$.

Proof of Lemma 4.1. Let D be the smallest distortion attained for a tree T which has minimal total edge length, let us show that this tree is in fact a path. Let $\{1, \dots, n\} = V_n$ and for $i \in V_n$, define

$$\begin{aligned} S(i) &= \left\{ i + 1, i + 2, \dots, i + \left\lfloor \frac{n}{2} \right\rfloor \right\} \\ S'(i) &= V_n \setminus \{S(i) \cup \{i\}\} \end{aligned}$$

where additions are modulo n. Assume that T has a vertex v of degree at least 3. Then there exist vertices i and j that are adjacent to v and belong to one of the semicircles $S(v)$ or $S'(v)$. Without loss of generality, suppose that $d_T(v, i) < d_T(v, j)$, the strict inequality coming from the fact that i and j belong to the same semicircle. Deleting the edge $\{v, j\}$ and adding the edge $\{i, j\}$ to T yields a graph T' which is again a tree. Indeed, if the added edge $\{i, j\}$ were contained in a cycle in T', then $\{v, j\}$ would be a chord of that cycle and that would imply the existence of a cycle in T which is impossible.

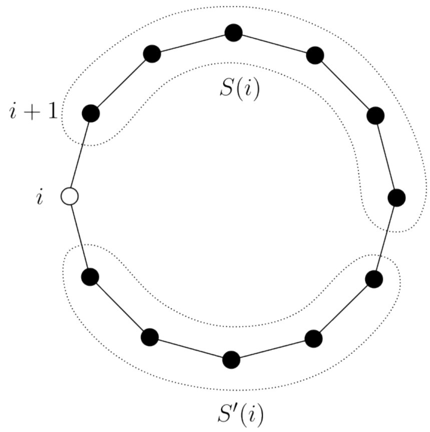

FIGURE 2. C_n split into two semicircles

Now, the distortion of the embedding into T' is not higher than that of the embedding into T, indeed a path using the edge $\{v, j\}$ in T can be rerouted along $\{v, i\}$ and $\{i, j\}$, now

$$d_{T'}(v, j) = d_{T'}(v, i) + 1 \leqslant d_T(v, j)$$

that is, the length of the path in the new tree T' has not increased. However the total edge length has gone down which contradicts the minimality of the total edge length of T, thus T has maximal degree 2. In the same way, if a vertex v of T has two neighbors in the same semicircle of v, deleting one of the edges containing v and adding one between the neighbors of v once again leads to a contradiction on the minimality of the edge length of T. Thus neighbors of v are in distinct semicircles and if $\{x, y\}$ is an edge of C_n that is not in T, the unique path from x to y in T has length $n - 1$. □

Lower bounds on the distortion of graph embeddings with given girth into trees had been discovered prior to Gupta's result, see for instance [RR98, Lemma 2.1], but Theorem 1.1 has the advantage of giving an elementary proof of the following bound.

Theorem 4.2. *Let $X = (V, E)$ be a graph with girth g and unit edge weight and let $T = (V', E')$ be a tree on $V' \supseteq V$. If $f : G \to T$ is the inclusion map on vertices, then* $\mathrm{dist}(f) \geqslant \frac{g-1}{8}$.

Proof. Assume not, then there exists a tree $T = (V', E')$ with distortion of the metric of X strictly less than $\frac{g-1}{8}$. In particular, X has a cycle C of length g such that the distortion between the metric d_C induced by C and the tree T is also less than $\frac{g-1}{8}$. Say the vertices of C understood as a subset of V' are the required vertices. Then, by Theorem 1.1, there exists a tree T' on the vertices of C such that the embedding of C into T' distorts d_C of a factor at most $8 \cdot \frac{g-1}{8} = g - 1$, which contradicts Lemma 4.1. □

4.2. **Rerouting in communication networks.** Consider the following problem in a communication network consisting of a finite set of nodes V and edge set E where the transmission time across edges is given by a weight function $w : E \to (0 + \infty)$. Some of the vertices are hosts between which we want to have communication paths. The others are routers that physically serve as relays and routing nodes in the network. Up to taking a spanning subgraph, which is well implemented in practice, we can suppose that the network is a tree T.

Multicasting a signal from a host $v \in V$ to the rest of the network works as follows: v is taken as the root of T and sends a packet to each of its children, which all do the same until the information has met every vertex. While being efficient by taking advantage of the tree structure of the network, multicasting is not always feasible as routers are often physically designed to handle only unicast, *i.e.* sending packets to only one neighbor simultaneously. On the other hand, the casting system of host being often implemented in their software and not depending on physical constraints, they are able to multicast information.

The idea to simulate multicast on a network with routers that only support unicast is to define a virtual tree on the hosts and to multicast packets according to this new tree, while using the underlying physical network. For this design to work efficiently, the transmission delays of the virtual tree should be close to those of the physical tree. Setting the routers to be the Steiner vertices, this is exactly the problem we have been discussing in this document.

References

[AA92] Noga Alon and Yossi Azar, *On-line steiner trees in the euclidean plane*, Proceedings of the eighth annual symposium on computational geometry, 1992, pp. 337–343.

[FRT03] Jittat Fakcharoenphol, Satish Rao, and Kunal Talwar, *A tight bound on approximating arbitrary metrics by tree metrics*, Proceedings of the thirty-fifth annual acm symposium on theory of computing, 2003, pp. 448–455.

[Gup01] Anupam Gupta, *Steiner points in tree metrics don't (really) help.*, Soda, 2001, pp. 220–227.

[RR98] Yuri Rabinovich and Ran Raz, *Lower bounds on the distortion of embedding finite metric spaces in graphs*, Discrete & Computational Geometry **19** (1998), no. 1, 79–94.