Distance Approximating Trees: Complexity and Algorithms

Feodor F. Dragan and Chenyu Yan

Department of Computer Science, Kent State University, Kent, OH 44242
dragan@cs.kent.edu, cyan@cs.kent.edu

Abstract. Let $\Delta \geq 1$ and $\delta \geq 0$ be real numbers. A tree $T = (V, E')$ is a *distance* (Δ, δ)*–approximating tree* of a graph $G = (V, E)$ if $d_H(u, v) \leq \Delta\ d_G(u, v) + \delta$ and $d_G(u, v) \leq \Delta\ d_H(u, v) + \delta$ hold for every $u, v \in V$. The *distance* (Δ, δ)*-approximating tree problem* asks for a given graph G to decide whether G has a distance (Δ, δ)-approximating tree. In this paper, we consider unweighted graphs and show that the distance $(\Delta, 0)$-approximating tree problem is NP-complete for any $\Delta \geq 5$ and the distance $(1, 1)$-approximating tree problem is polynomial time solvable.

1 Introduction

Many combinatorial and algorithmic problems are concerned with distances in a finite metric space induced by an undirected graph (possible weighted). An arbitrary metric space (in particular a finite metric defined by a general graph) might not have enough structure to exploit algorithmically. A powerful technique that has been successfully used recently in this context is to embed the given metric space in a simpler metric space such that the distances are approximately preserved in the embedding. New and improved algorithms have resulted from this idea for several important problems [1, 2, 7, 11, 12, 20]. Tree metrics are a very natural class of simple metric spaces since many algorithmic problems become tractable on them. If we approximate the graph by a tree such that the distance between a pair of vertices in the tree is at most some small factor of their distance in the graph, we can solve the problem on the tree and the solution interpret on the original graph.

Approximating general graph–distance d_G by a simpler distance (in particular, by tree–distance d_T) is useful also in such areas as communication networks, data analysis, motion planning, image processing, network design, and phylogenetic analysis. The goal is, for a given graph $G = (V, E)$, to find a sparse graph $H = (V, E')$ with the same vertex set, such that the distance $d_H(u, v)$ in H between two vertices $u, v \in V$ is reasonably close to the corresponding distance $d_G(u, v)$ in the original graph G. There are several ways to measure the quality of this approximation, two of them leading to the notion of a spanner. For $t \geq 1$ a spanning subgraph H of G is called a *multiplicative* t*–spanner* of G [9, 23, 24] if $d_H(u, v) \leq t\ d_G(u, v)$ for all $u, v \in V$. If $r \geq 0$ and $d_H(u, v) \leq d_G(u, v) + r$ for all $u, v \in V$, then H is called an *additive* r*–spanner* [19].

When H is a tree, one gets the notions of *multiplicative tree* t*–spanner* and *additive tree* r*–spanner*, respectively. Tree spanners of graphs were considered

T. Calamoneri, I. Finocchi, G.F. Italiano (Eds.): CIAC 2006, LNCS 3998, pp. 260–271, 2006.

in [6, 10, 25]. It was shown in [6] that for a given graph G and integer t, the problem to decide whether G has a multiplicative tree t–spanner is NP–complete for $t \geq 4$ and is linearly solvable for $t = 1, 2$. The status of the case $t = 3$ is open.

For many applications (e.g. in numerical taxonomy or in phylogeny reconstruction) the condition that H must be a spanning subgraph of G can be dropped (see [3, 26, 27]). In this case there is a striking way to measure how sharp d_H approximates d_G, based on the notion of a pseudoisometry between two metric spaces [20, 4]. Let $\Delta \geq 1$ and $\delta \geq 0$ be real numbers. Two graphs $G = (V, E)$ and $H = (V, E')$ are said to be *(Δ, δ)–pseudoisometric* [4] if for all $u, v \in V$, $d_H(u, v) \leq \Delta\ d_G(u, v) + \delta$ and $d_G(u, v) \leq \Delta\ d_H(u, v) + \delta$ hold. H is then said to be a *distance (Δ, δ)–approximating graph* for G (and vice-versa, G is a *distance (Δ, δ)–approximating graph* for H).

In this paper, continuing the line of research started in [4, 8], we will be interested in two special cases, when H is a tree and either $\Delta = 1$ or $\delta = 0$. A tree $T = (V, E')$ is a *distance $(\Delta, 0)$–approximating tree* of $G = (V, E)$ if $\frac{1}{\Delta} d_G(u, v) \leq d_T(u, v) \leq \Delta\ d_G(u, v)$ for all $u, v \in V$. A tree $T = (V, E')$ is a *distance $(1, \delta)$–approximating tree* of $G = (V, E)$ (or, simply, a distance δ-approximating tree of G) if $|d_G(u, v) - d_T(u, v)| \leq \delta$ for all $u, v \in V$. The *distance (Δ, δ)-approximating tree problem* asks for a given graph G to decide whether G has a distance (Δ, δ)-approximating tree.

In this paper, we consider unweighted graphs and show that the distance $(\Delta, 0)$-approximating tree problem is NP-complete for any $\Delta \geq 5$ and the distance $(1, 1)$-approximating tree problem is polynomial time solvable. The latter solves (algorithmically) the problem posed in [8] which asked to characterize/recognize the graphs admitting distance $(1, 1)$-approximating trees.

1.1 Previous Results and Their Implications

Let $G = (V, E)$ be a connected, undirected, loopless, and without multiple edges graph. The *length* of a path from a vertex u to a vertex v is the number of edges in this path. The *distance* $d_G(u, v)$ between the vertices u and v in G is the length of a shortest (u, v)-path.

A graph G is called *chordal* if no induced cycle of G has four or more edges. It is known that the class of chordal graphs does not admit any good tree spanners. Independently McKee [21] and Kratsch et al. [16] showed that, for every fixed integer t, there is a chordal graph without tree t–spanners (additive as well as multiplicative). Furthermore, recently Brandstädt et al. [5] have shown that, for any $t \geq 4$, the problem to decide whether a given chordal graph G admits a multiplicative tree t-spanner is NP-complete.

In contrast, in [4], Brandstädt et al. proved that every chordal graph G admits a tree $T(G)$ (constructable in linear time) which is both a $(3, 0)-$ and a $(1, 2)-$approximating tree of G. So, from the metric point of view chordal graphs do look like trees, but the notion of tree spanners failed to capture this. Note that the result is optimal in the sense that there are chordal graphs which do not admit any distance $(1, 1)$–approximating trees [8].

The result was used in [4, 8, 13] to provide efficient approximate solutions for several problems on chordal graphs. It is known that the (exact) distance matrix $D(G)$ of a chordal graph $G = (V, E)$ cannot be computed in less than "matrix-multiplication" time. Using a distance $(1, 2)$-approximating tree $T(G)$ of G, after a linear time preprocessing of G (and then of $T(G)$), in only $O(1)$ time, one can compute $d_G(x, y)$ with an error of at most 2 for any $x, y \in V$ (see [4] for further details). As another application, consider the p-center problem: given a graph G (or, more generally, a metric space) and an integer $p > 0$, we are searching for smallest radius r^* and a subset of vertices X of G with $|X| \leq p$ such that $d_G(v, X) \leq r^*$ for every vertex v of G. The problem is NP-hard even for chordal graphs. Solving the p-center problem on a distance $(1, 2)$-approximating tree $T(G)$ of G (on trees this problem is polynomial time solvable [15]), we will find an optimal covering radius r of $T(G)$ and a set of centers Y with $|Y| \leq p$. Then, Y can be taken as an approximate solution for G since $d_G(v, Y) \leq r + 2 \leq r^* + 4$ for all $v \in V$ (see [8] for further details). Clearly, similar results can be obtained for any graph admitting a good distance approximating tree.

The result was also used by Gupta in [13] for bandwidth approximation in chordal graphs. If a graph G has a distance (Δ, δ)–approximating tree $T(G)$ for some constants Δ and δ, then the bandwidth of a linear arrangements of G will be within some constant of the bandwidth of the same arrangement for $T(G)$. Gupta developed in [13] a simple randomized $O(log^{2.5} n)$-approximation algorithm for bandwidth minimization on trees and used it to get an approximation algorithm with a similar performance guarantee for chordal graphs (see [13] for further details). In [18], Krauthgamer et al. used the existence of good distance approximating trees for chordal graphs to obtain an embedding of any chordal graph into l_2 with a small r-dimensional volume distortion.

Later, in [8], Chepoi and Dragan extended the method of [4] from chordal graphs to all k-chordal graphs. A graph G is said to be *k-chordal* if no induced cycle of G has more than k edges. It was proven that, for every k-chordal graph $G = (V, E)$, there exists a tree $T = (V, F)$ (constructable in linear time) such that $|d_G(u, v) - d_T(u, v)| \leq \lfloor \frac{k}{2} \rfloor + \alpha$ for all vertices $u, v \in V$, where $\alpha = 1$ if $k \neq 4, 5$ and $\alpha = 2$ otherwise. Clearly, this result can be used to provide efficient approximate solutions for several problems on k-chordal graphs. Here, we will mention only one implication provided in [17]. Krauthgamer and Lee, in [17], proved first that the *Levin's conjecture on intrinsic dimensionality of graphs* holds for trees. Then, relying on low-distortion embeddings of k-chordal graphs into trees, due to [8], they extended that result to all k-chordal graphs: the *Levin's conjecture on intrinsic dimensionality of graphs* holds for all k-chordal graphs with bounded k (see [17] for further details).

Motivated by those applications of distance approximating trees, in this paper, we investigate the question how hard for a given graph G to find a good distance (Δ, δ)-approximating tree (for small Δ and δ). We prove that the distance $(\Delta, 0)$-approximating tree problem is NP-complete for any $\Delta \geq 5$ and the distance $(1, 1)$-approximating tree problem is polynomial time solvable. Due to space limitation, in this conference version, we present only the second result. The

NP-completeness proof will be given in the journal version. We reduce 3SAT to our problem. The reduction is too technical, involves complicated gadgets for the Boolean variables and hence omitted in this version.

1.2 Basic Notions, Notation and Facts

Let $G = (V, E)$ be a graph endowed with the shortest path metric $d_G(u, v)$. The *eccentricity* $ecc_G(v)$ of a vertex v is the maximum distance from v to any vertex in G. The *radius* $rad(G)$ of a graph G is the minimum eccentricity of a vertex in G and the *diameter* $diam(G)$ of G is the maximum eccentricity of a vertex.

For a subset $S \subseteq V$ of vertices of a graph G, by $G(S)$ we denote the subgraph of G induced by S. Let, for simplicity, $G - v := G(V \setminus \{v\})$ and $G - v - u := G(V \setminus \{v, u\})$, where v and u are vertices of G. Let also $G - uv$ denote the graph obtained from G by removing edge uv of G, i.e., $G - uv := (V, E \setminus \{uv\})$. A graph G is said to be *3-connected* if $G - u - v$ is connected for any pair of vertices $u, v \in V$. A graph G is said to be *2-connected* if $G - u$ is connected for any vertex $u \in V$. In a 2-connected graph G, if for some pair of vertices $x, y \in V$ the graph $G - x - y$ is disconnected, then we say that $\{x, y\}$ is a *2-cut* of G. In a connected graph G, if for some vertex $x \in V$ the graph $G - x$ is disconnected, then we say that x is a *1-cut vertex* (or, simply, *1-cut*) of G.

It is easy to see from the definitions of distance approximating trees that the following holds.

- A tree $T = (V, F)$ is a distance $(\Delta, 0)$-approximating tree of a graph $G = (V, E)$ if and only if $d_T(x, y) \leq \Delta$ holds for each edge $xy \in E$ and $d_G(u, v) \leq \Delta$ holds for each edge $uv \in F$.
- If T is a distance $(1, \delta)$-approximating tree for G, then T is a distance $(\delta + 1, 0)$-approximating tree for G.

2 Distance $(1, 1)$-Approximating Trees

In this section, we show that the distance $(1, 1)$-approximating tree problem is polynomial time solvable. For simplicity, in what follows, we will use the notion "distance 1-approximating tree" as a synonym to "distance $(1, 1)$-approximating tree".

2.1 3-Connected Graphs

A *star* is a tree with a vertex adjacent to all other vertices. We call that vertex *the center of the star*. Equivalently, a *star* is a tree of diameter at most 2.

Lemma 1. *For a 3-connected graph G, the following statements are equivalent.*

1. *G has a distance 1-approximating tree.*
2. *G has a distance 1-approximating tree which is a star.*
3. *$diam(G) \leq 3$ and $rad(G) \leq 2$.*

Proof. (1$\Longleftrightarrow$2) Let T be a distance 1-approximating tree of G. If T is not a star, then there exists a path in T with length 3. Let (x', x, y, y') be such a path. Consider subtrees T_x and T_y obtained from T by removing edge xy, and assume that x belongs to T_x and y belongs to T_y. Since for any $u \in V(T_x) \setminus \{x\}$ and $v \in V(T_y) \setminus \{y\}$, $d_T(u,v) \geq 3$, we have $uv \notin E(G)$. This implies that $\{x, y\}$ is a 2-cut of G, contradicting with the 3-connectedness of G. Hence, T must be a star.

(2$\Rightarrow$3) Let T be a distance 1-approximating tree of G which is a star. Then, for any $x, y \in V$, we have $d_T(x,y) \leq 2$ and, therefore, $d_G(x,y) \leq 3$. Hence, $diam(G) \leq 3$. Let now u be the center of T. Then, for each $x \in V$, $d_T(x,u) \leq 1$, and therefore $d_G(x,u) \leq 2$. The latter implies $rad(G) \leq 2$.

(3$\Rightarrow$2) If $rad(G) \leq 2$, then, by definition, there exists a vertex $u \in V$ such that $d_G(x,u) \leq 2$, for any $x \in V$. Pick such a vertex u and construct a tree $T = (V, E')$ where each vertex $v \in V \setminus \{u\}$ is adjacent to u, i.e., construct a star on vertices V with the center u. Obviously, $0 \leq d_G(x,y) - d_T(x,y) \leq 1$, for any $x \in V \setminus \{u\}$. Moreover, since $diam(G) \leq 3$, we have $d_G(x,y) \leq 3$ for any $x, y \in V \setminus \{u\}$. As, for those vertices x and y, $d_T(x,y) = 2$, we conclude $d_G(x,y) - d_T(x,y) \leq 3 - 2 = 1$ and $d_G(x,y) - d_T(x,y) \geq 1 - 2 = -1$. Hence, T is a distance 1-approximating tree of G. □

Corollary 1. *Let G be an arbitrary (not necessarily 3-connected) graph. Then, G has a distance 1-approximating tree which is a star if and only if $diam(G) \leq 3$ and $rad(G) \leq 2$.*

2.2 2-Connected Graphs

A vertex of a tree is *inner* if it is not a leaf. An edge of a tree is an *inner edge* if it is not incident to a leaf.

Lemma 2. *If T is a distance 1-approximating tree of a connected graph G, then any inner edge of T is a 2-cut of G.*

Proof. For any inner edge xy of T, let T_x and T_y be the two subtrees of T obtained from T by removing edge xy. Let also x belong to T_x and y belong to T_y. Then, since T is a distance 1-approximating tree of G, for all $u \in V(T_x) \setminus \{x\}$ and $v \in V(T_y) \setminus \{y\}$, $uv \notin E(G)$. This implies that $\{x, y\}$ is a 2-cut of G separating $V(T_x) \setminus \{x\}$ from $V(T_y) \setminus \{y\}$. □

A *bistar* is a tree with only one inner edge. Equivalently, a *bistar* is a tree of diameter 3. The proof of the following lemma is omitted.

Lemma 3. *If T is a distance 1-approximating tree of a 2-connected graph G, then $diam(T) \leq 3$, i.e., T is a star or a bistar.*

To characterize 2-connected graphs admitting distance 1-approximating trees, we will need also the following easy observations (proofs are omitted).

Lemma 4. *Assume a graph G has a distance 1-approximating bistar T with the inner edge c_1c_2. Then, the following properties hold:*

1. *$diam(G) \leq 4$ and $rad(G) \leq 3$;*
2. *for any $j = 1, 2$ and $x, y \in V(T_{c_j}) \cup \{c_1, c_2\}$, $d_G(x,y) \leq 3$ and $d_G(x, c_j) \leq 2$;*

3. *if* $A_1, \ldots, A_k$ *are the connected components of the graph* $G - c_1 - c_2$ *and* T_{c_1}, T_{c_2} *are the connected components of* $T - c_1c_2$, *then, for any* $i = 1, \ldots, k$, $V(A_i)$ *is entirely contained either in* $V(T_{c_1})$ *or in* $V(T_{c_2})$.

Let now G be a graph with a 2-cut $\{a, b\}$ and $A_1, \ldots, A_k$ be the connected components of the graph $G - a - b$. For given 2-cut $\{a, b\}$ of G we can construct a new graph $H_{a,b}$ as follows. The vertex set of $H_{a,b}$ is $\{a, b, a_1, \ldots, a_k\}$. Edge aa_i $(i = 1, \ldots, k)$ exists in $H_{a,b}$ if and only if for each $x, y \in V(A_i) \cup \{b\}$, $d_G(x, y) \leq 3$ and $d_G(x, a) \leq 2$ hold. Edge ba_i $(i = 1, \ldots, k)$ exists in $H_{a,b}$ if and only if for each $x, y \in V(A_i) \cup \{a\}$, $d_G(x, y) \leq 3$ and $d_G(x, b) \leq 2$ hold. Edge a_ia_j $(i, j = 1, \ldots, k, i \neq j)$ exists in $H_{a,b}$ if and only if for each vertex $x \in V(A_i)$ and each vertex $y \in V(A_j)$, $d_G(x, y) \leq 3$ holds. No other edges exist in $H_{a,b}$.

The following lemma gives a characterization of those 2-connected graphs that admit distance 1-approximating trees. Denote the complement of a graph H by $\overline{H}$.

Lemma 5. *For a 2-connected graph* G, *the following statements are equivalent.*

1. G *has a distance 1-approximating tree.*
2. G *has a distance 1-approximating tree which is a star or a bistar.*
3. $diam(G) \leq 3$ *and* $rad(G) \leq 2$ *or* $diam(G) \leq 4$ *and there exists a 2-cut* $\{a, b\}$ *in* G *such that the graph* $\overline{H_{a,b}}$ *is bipartite.*

Proof. $(1 \Longleftrightarrow 2)$ is given by Lemma 3.

$(2 \Rightarrow 3)$ If G has a distance 1-approximating tree which is a star, then, by Corollary 1, $diam(G) \leq 3$ and $rad(G) \leq 2$. Assume now that a distance 1-approximating tree T of G is a bistar. Then, by Lemma 4, $diam(G) \leq 4$. Lemma 4 (together with Lemma 2) implies also that G has a 2-cut $\{a, b\}$ (which is the inner edge of T) such that for any connected component A_i $(i \in \{1, \ldots, k\})$ of $G - a - b$, either $V(A_i) \subset V(T_a)$ or $V(A_i) \subset V(T_b)$ holds. Since vertices $V(T_a) \cup \{b\}$ form a star in T with the center a, we have $d_G(x, y) \leq 3$ and $d_G(x, a) \leq 2$ for any $x, y \in V(T_a) \cup \{b\}$. By construction of $H_{a,b}$, vertices $\{a\} \cup \{a_i : V(A_i) \subset V(T_a)\}$ of $H_{a,b}$ will form a clique. Analogously, vertices $\{b\} \cup \{a_i : V(A_i) \subset V(T_b)\}$ form a clique in $H_{a,b}$. Since these two cliques cover all vertices of $H_{a,b}$, the complement $\overline{H_{a,b}}$ of $H_{a,b}$ is bipartite.

$(3 \Rightarrow 2)$ Clearly, if $diam(G) \leq 3$ and $rad(G) \leq 2$ then, by Corollary 1, G has a distance 1-approximating star. Assume now that $diam(G) \leq 4$ and there exists a 2-cut $\{a, b\}$ in G such that the graph $\overline{H_{a,b}}$ is bipartite. Let $A_1, \ldots, A_k$ be the connected components of the graph $G - a - b$. Vertices of $H_{a,b}$ can be partitioned into two cliques C_1 and C_2. Since a and b are not adjacent in $H_{a,b}$, they must be in different cliques. Assume, $a \in C_1$ and $b \in C_2$. By construction of $H_{a,b}$, for all $x, y \in \cup\{V(A_i) : a_i \in C_1\} \cup \{b\}$, $d_G(x, y) \leq 3$ and $d_G(x, a) \leq 2$ holds. Similarly, for all $x, y \in \cup\{V(A_i) : a_i \in C_2\} \cup \{a\}$, $d_G(x, y) \leq 3$ and $d_G(x, b) \leq 2$ holds. Hence, we can construct a bistar T of G as follows. Vertices a and b will form the inner edge of T. Vertices of A_i with $a_i \in C_1$ will be attached (i.e., made adjacent in T) to a. Vertices of A_i with $a_i \in C_2$ will be attached to b. It is easy to see that T is a distance 1-approximating tree of G. The only interesting case to mention here is when $x \in V(A_i)$, where $a_i \in C_1$, and $y \in V(A_j)$, where $a_j \in C_2$. For those x and y, we have $d_T(x, y) = 3$ and $2 \leq d_G(x, y) \leq 4$ (since $diam(G) \leq 4$ and x and y are separated by $\{a, b\}$ in G). Thus, $-1 \leq c_T(x, y) \leq 1$ holds. □

Corollary 2. *Let G be an arbitrary (not necessarily 2-connected) graph. Then, G has a distance 1-approximating tree which is a star or a bistar if and only if $diam(G) \leq 3$ and $rad(G) \leq 2$ or $diam(G) \leq 4$ and there exists a 2-cut $\{a, b\}$ in G such that the graph $\overline{H_{a,b}}$ is bipartite.*

Lemma 5 implies also that the problem of checking whether a given 2-connected graph G has a distance 1-approximating tree is polynomial time solvable. More specifically, we have

Corollary 3. *It is possible, for a given 2-connected graph $G = (V, E)$, to check in $O(|V|^4)$ time whether G has a distance 1-approximating tree and, if such a tree exists, construct one within the same time bound.*

Proof. We can find in $O(|V||E|)$ time the distance matrix of G and all 2-cuts [14, 22] of G. Then, to check whether $diam(G) \leq 3$ and $rad(G) \leq 2$ and, if so, to construct a distance 1-approximating star of G as described in the proof of Lemma 1, one needs at most $O(|V|^2)$ time in total. To check if $diam(G) \leq 4$ and whether there exists a 2-cut $\{a, b\}$ of G with $\overline{H_{a,b}}$ bipartite, one needs $O(|V|^4)$ total time. We just need, for each 2-cut $\{a, b\}$, to construct the graph $\overline{H_{a,b}}$ and check if it is bipartite. Construction of $\overline{H_{a,b}}$ for a given 2-cut $\{a, b\}$ and checking whether it is bipartite will take no more than $O(|V|^2)$ time (given the distance matrix of G). Since any graph G has at most $O(|V|^2)$ 2-cuts, to check if G has a distance 1-approximating bistar, one needs at most $O(|V|^4)$ time. If G admits such a bistar, then we can find one in linear time as described in the proof of Lemma 5. □

2.3 Connected Graphs

In this subsection, we assume that G is a connected graph but not 2-connected. Therefore, there exists a vertex $v \in V(G)$, such that $G - v$ contains at least two connected components.

From Lemma 3 and its proof, the following lemma is obvious.

Lemma 6. *Let T be a distance 1-approximating tree of a connected graph G and (a, b, c) be a path in T. If both a and c are inner vertices of T, then at least one of these vertices is a 1-cut of G. Moreover, assuming c is a 1-cut, c separates vertices $V(T_c) \setminus \{c\}$ from other vertices of G, where T_c is the subtree of $T - bc$ containing c.*

A *2-connected component* of a graph G is a maximal by inclusion 2-connected subgraph of G or an edge uv of G such that both u and v are 1-cuts of G (such an edge is called a *bridge* of G). Two 2-connected components of G are neighbors if they share a common vertex (a 1-cut) of G.

Lemma 7. *Let G be a connected graph admitting a distance 1-approximating tree T and A be a 2-connected component of G. Then, for any two vertices $x, y \in V(A)$, $d_T(x, y) \leq 3$. Moreover, if there exist vertices $x, y \in V(A)$, such that $d_T(x, y) = 3$, then $T(V(A))$ is a bistar.*

Proof. Assume that, for some vertices $x, y \in V(A)$, $d_T(x, y) \geq 4$ holds. Then, one can connect x and y in T with a path $P_T(x, y)$ of length at least 4. Pick three consecutive inner vertices a, b, c of path $P_T(x, y)$, they necessarily exist. According to Lemma 6, a or c is a 1-cut of G separating x from y in G. The latter is in contradiction with the assumption that $x, y \in V(A)$ and A is a 2-connected component of G. Hence, $d_T(x, y) \leq 3$, for any $x, y \in V(A)$, is proven.

Assume now that there exist vertices $x, y \in V(A)$, such that $d_T(x, y) = 3$. Then, one can find two vertices $\{c_1, c_2\}$ in G such that $T(V(A) \cup \{c_1, c_2\})$ is a bistar with the inner edge c_1c_2. Let $xc_1, yc_2 \in E(T)$. We will show that both c_1 and c_2 are in A.

Suppose, neither c_1 nor c_2 is in A. Assume $c_1 \in V(B)$, $c_2 \in V(C)$, where B and C are 2-connected components of G. Let $V(B) \cap V(A) = \{v\}$ and $V(C) \cap V(A) = \{u\}$. We claim that $B = C$ or at least $v = u$. Suppose $B \neq C$ and $v \neq u$. Then, since $V(B) \cap V(C) = \emptyset$ (otherwise, A, B and C will be parts of one 2-connected component of G), $d_G(c_1, c_2) \geq 3$. As $d_T(c_1, c_2) = 1$, a contradiction with T being a distance 1-approximating tree of G arises. So, c_1, c_2 must be either in one 2-connected component of G or in two 2-connected components B and C such that $V(B) \cap V(A) = V(C) \cap V(A)$.

Without loss of generality, assume v is attached (i.e., adjacent in T) to c_1. Since $d_T(y, c_2) = 1$, we have $d_G(y, c_2) \leq 2$ and, hence, $yv \in E(G)$. On the other hand, $d_T(y, v) = 3$, contradicting the assumption that T is a distance 1-approximating tree of G.

Assume now that $c_1 \in V(A)$ and $c_2 \in V(B) \setminus \{v\}$. For any vertex $x' \in V(A)$ which is attached to c_1 and any vertex $y' \in V(A) \setminus \{c_1\}$ which is attached to c_2, $x'y' \notin E(G)$ must hold. Moreover, since $V(A) \cap V(B) = \{v\}$, one concludes that for all $x' \in V(A) \setminus \{v\}$, $x'c_2 \notin E(G)$. Hence, any path of A connecting a vertex attached to c_1 with a vertex attached to c_2 must use vertex c_1. Since there exist vertices $x, y \in V(A)$ such that $xc_1, yc_2 \in E(T)$, this is in contradiction with the assumption that A is 2-connected.

Thus, we conclude that $T(V(A))$ is a bistar. □

Corollary 4. *Let G be a connected graph admitting a distance 1-approximating tree T and A be a 2-connected component of G. Then, either $T(V(A))$ is a bistar or $T(V(A) \cup \{c\})$ is a star centered at some vertex c of G.*

In what follows, we will show that among all possible distance 1-approximating trees of G there is a tree T such that, for any 2-connected component A of G, $T(V(A))$ is connected, i.e., if $T(V(A) \cup \{c\})$ is a star for some vertex c of G, then c must be in A. To show that, we will need two lemmata (proofs can be found in the journal version).

A sequence $(B_0 := B, B_1, \ldots, B_{k-1}, B_k := A)$ is called *the chain of 2-connected components of G between A and B* if each B_i is a 2-connected component of G, B_i and B_j are different for $j \neq i$, B_{i-1}, B_i are neighbors sharing a 1-cut $v_i := V(B_{i-1}) \cap V(B_i)$ of G for any $i \in \{1, \ldots, k\}$, and $v_i \neq v_j$ for any $i \neq j$. Clearly, this chain is unique for any A and B.

Lemma 8. *Let G be a connected graph admitting a distance 1-approximating tree T, A and B be 2-connected components of G and $(B_0 := B, B_1, \ldots, B_{k-1}, B_k := A, Z)$ be the chain of 2-connected components of G between Z and B. If $T(V(A) \cup \{c\})$ is a star with the center c belonging to $V(Z) \setminus V(A)$, then for any $i \in \{0, \ldots, k-1\}$, $T(V(B_i))$ is a star centered at a 1-cut $v_{i+1} := V(B_{i+1}) \cap V(B_i)$ of G. Moreover, for any $i \in \{0, \ldots, k-1\}$ and any $x \in V(B_i)$, $xv_{i+1} \in E(G)$ must hold.*

Lemma 9. *Let G be a connected graph admitting a distance 1-approximating tree T and let A, Z be 2-connected components of G such that $V(A) \cap V(Z) = \{v\}$. Let also A' be that connected component of the graph $G - v$ which contains $A - v$. If $T(V(A) \cup \{c\})$ is a star centered at $c \in V(Z) \setminus \{v\}$, then for any vertices $x \in V(A')$, $y \in (V(G) \setminus V(A')) \setminus \{c, v\}$, $xy \notin E(T)$ holds. In particular, for any two vertices $y, z \in V(G) \setminus V(A')$, the path $P_T(x, y)$ between x and y in T does not contain any vertices of A'.*

In what follows, let G be a connected graph admitting a distance 1-approximating tree and let T denote a distance 1-approximating tree of G with minimum $|E(T) \setminus E(G)|$, i.e., with minimum number of non-graph edges. We will show that this tree T has a number of nice properties.

Theorem 1. *If T is a distance 1-approximating tree of G with minimum $|E(T) \setminus E(G)|$, then for any 2-connected component A of G, $T(V(A))$ is a star or a bistar.*

Proof. Since A is a 2-connected component of G, by Corollary 4, either $T(V(A))$ is a bistar or $T(V(A) \cup \{c\})$ is a star centered at some vertex c of G. By way of contradiction, assume that for A, $T(V(A) \cup \{c\})$ is a star centered at a vertex c of G not belonging to A. Let c belong to some 2-connected component Z of G. Necessarily, A and Z are neighbor (2-connected) components. Let $v := V(A) \cap V(Z)$ and A' be a connected component of $G - v$ containing $V(A) \setminus \{v\}$. By Lemma 8, for any 2-connected component B of G, which is different from A and belongs to A', $T(V(B))$ is a star centered at a 1-cut of G lying in B and closest to A. Moreover, if v' is that 1-cut, then for any $x \in V(B)$, $xv' \in E(G)$ holds (see Fig. 1). We have also that v is adjacent in G to c and to any vertex a $(a \neq v)$ of A (see Lemma 8).

We can transform tree T into a new tree T' as follows. Set $E(T') := E(T)$ and $V(T') := V(T)$. For each vertex $a \in V(A) \setminus \{v\}$, let $E(T') := (E(T') \setminus \{ac\}) \cup \{av\}$ (i.e., replace edge ac with edge av). We claim that T' is a distance 1-approximating tree of G, too. We need to show that $|d_{T'}(x, y) - d_G(x, y)| \leq 1$ holds for any two vertices $x, y \in V(G)$.

If $x, y \in V(A')$ then, by Lemma 8 and the way we transformed T into T', $d_{T'}(x, y) = d_T(x, y)$. If $x, y \in V(G) \setminus V(A')$ then, by Lemma 9 and the way T was transformed into T', $d_{T'}(x, y) = d_T(x, y)$. Hence, in these cases, $|d_{T'}(x, y) - d_G(x, y)| = |d_T(x, y) - d_G(x, y)| \leq 1$.

Consider now the case when $x \in V(A')$ and $y \in V(G) \setminus V(A')$. By Lemma 8, $d_{T'}(x, v) = d_G(x, v)$. Since v is a 1-cut of G, $d_G(x, y) = d_G(x, v) + d_G(v, y)$. By Lemma 9 and the way we transformed T into T', one concludes that $d_{T'}(x, y) =$

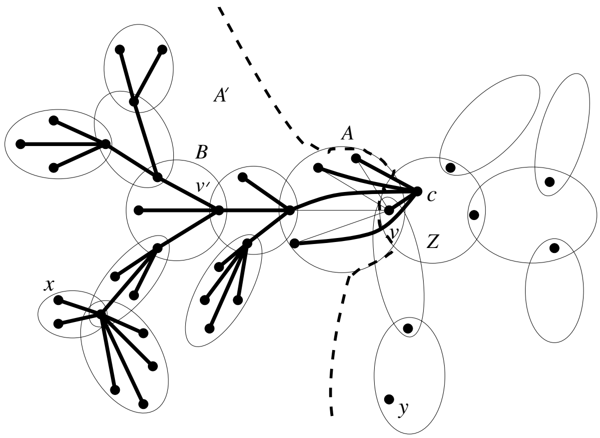

Fig. 1. Illustration to the proof of Theorem 1. A part of the tree T is shown using thick edges. Thin edges show some graph edges.

$d_{T'}(x, v) + d_{T'}(v, y)$. Combining these equalities, we get $|d_{T'}(x, y) - d_G(x, y)| = |d_{T'}(x, v) + d_{T'}(v, y) - (d_G(x, v) + d_G(v, y))| = |d_{T'}(v, y) - d_G(v, y)|$. But, by Lemma 9, $d_{T'}(v, y) = d_T(v, y)$. Hence, we get $|d_{T'}(x, y) - d_G(x, y)| = |d_T(v, y) - d_G(v, y)| \leq 1$.

Thus, T' is a distance 1-approximating tree of G. Since T' has original graph edges more than T has ($|E(T') \setminus E(G)| < |E(T) \setminus E(G)|$), a contradiction with the choice of T arises. Hence, the center c of star $T(V(A) \cup \{c\})$ must be in A. □

Lemma 10. *Let T be a distance 1-approximating tree of G with minimum $|E(T) \setminus E(G)|$ and A be a 2-connected component of G such that $T(V(A))$ is a bistar. Then, for any other 2-connected component B of G, $T(V(B))$ is a star centered at a 1-cut of G which is closest to A (among all 1-cuts of G located in B).*

Corollary 5. *If T is a distance 1-approximating tree of G with minimum $|E(T) \setminus E(G)|$, then there is at most one 2-connected component A in G such that $T(V(A))$ is a bistar.*

The following lemma and its corollaries show that a distance 1-approximating tree T of G with $T(V(A))$ being a star for any 2-connected component A of G has also a very deterministic structure.

Lemma 11. *Let T be a distance 1-approximating tree of G with minimum $|E(T) \setminus E(G)|$ and A and B be two neighbor 2-connected components of G with $v := V(A) \cap V(B)$. If $T(V(A))$ is a star centered not at v, then $T(V(B))$ is a star centered at v.*

Proof. Since $T(V(A))$ is a star centered at some vertex $c \in V(A) \setminus \{v\}$, there must exist a vertex a in A such that $av \in E(G) \setminus E(T)$. By Lemma 10, $T(V(B))$ cannot be a bistar. If $T(V(B))$ is a star centered at some vertex $c' \in V(B) \setminus \{v\}$, then there must exist a vertex b in B such that $bv \in E(G) \setminus E(T)$. For these vertices a and b, $d_G(a,b) = 2$ and $d_T(a,b) = d_T(a,v) + d_T(v,b) = 2 + 2 = 4$ hold, contradicting with T being a distance 1-approximating tree of G. Hence, the center of $T(V(B))$ must be v. □

Corollary 6. *Let T be a distance 1-approximating tree of G with minimum $|E(T) \setminus E(G)|$ and A be a 2-connected component of G such that $T(V(A))$ is a star. If the center of this star $T(V(A))$ is not a 1-cut of G, then for any other 2-connected component B of G, $T(V(B))$ is a star centered at a 1-cut of G which is closest to A (among all 1-cuts of G located in B).*

Corollary 7. *Let T be a distance 1-approximating tree of G with minimum $|E(T) \setminus E(G)|$. If for every 2-connected component A of G, $T(V(A))$ is a star centered at a 1-cut of G, then there exists a 1-cut v in G such that*

a) for any 2-connected component A of G containing v, $T(V(A))$ is a star centered at v,
b) for any 2-connected component B of G not containing v, $T(V(B))$ is a star centered at a 1-cut of G which is closest to v (among all 1-cuts of G located in B).

Clearly, if $T(V(A))$ is a star for a 2-connected component A of G, then $diam(A) \leq 3$ and $rad(A) \leq 2$. And, if $T(V(B))$ is a bistar for a 2-connected component B of G, then $diam(B) \leq 4$ and $rad(B) \leq 3$.

Using all these auxiliary results, one can prove the following theorem (its proof is omitted in this conference version).

Theorem 2. *It is possible, for a given connected graph $G = (V, E)$, to check in $O(|V|^4)$ time whether G has a distance 1-approximating tree and, if such a tree exists, construct one within the same time bound.*

3 Conclusion

In this paper, we proved that the distance $(\Delta, 0)$-approximating tree problem is NP-complete for any $\Delta \geq 5$ and the distance $(1, 1)$-approximating tree problem is polynomial time solvable.

It remains an interesting open question to characterize/recognize the graphs admitting distance (Δ, δ)–approximating trees for $\Delta = 2, 3, 4$ and $\delta = 2, 3, 4$, or to prove that the problem remains NP-hard even for some of these small Δs and δs.

References

1. Y. Bartal, Probabalistic approximation of metric spaces and its algorithmic applications, *FOCS* 1996, pp. 184-193.
2. Y. Bartal, A. Blum, C. Burch, and A. Tomkins, A polylog(n)competitive algorithm for metrical task systems, *STOC* 1997, pp 711–719.
3. J.-P. Barthélemy and A. Guénoche, Trees and Proximity Representations, *Wiley,* New York, 1991.

4. A. Brandstädt, V. Chepoi, and F.F. Dragan, Distance Approximating Trees for Chordal and Dually Chordal Graphs, *Journal of Algorithms* 30 (1999), 166–184.
5. A. Brandstädt, F. Dragan, H.-O. Le, and V.B. Le, Tree Spanners on Chordal Graphs: Complexity and Algorithms, *Theor. Comput. Science* 310 (2004), 329-354.
6. L. Cai and D.G. Corneil, Tree spanners, *SIAM J. Disc. Math.* 8 (1995), 359–387.
7. M. Charikar, C. Chekuri, A. Goel, S. Guha, and S. Plotkin, Approximating a Finite Metric by a Small Number of Tree Metrics, *FOCS* 1998, pp. 379–388.
8. V. Chepoi and F.F. Dragan, A note on distance approximating trees in graphs, *European Journal of Combinatorics* 21 (2000), 761–766.
9. L.P. Chew, There are planar graphs almost as good as the complete graph, *J. of Computer and System Sciences*, 39 (1989), 205–219.
10. Y. Emek and D. Peleg, Approximating Minimum Max-Stretch Spanning Trees on Unweighted Graphs, SODA 2004, pp. 261-270.
11. J. Fakcharoenphol, S. Rao, and K. Talwar, A tight bound on approximating arbitrary metrics by tree metrics, *STOC* 2003, pp. 448-455.
12. U. Feige, Approximating the Bandwidth via Volume Respecting Embeddings, *J. Comput. System Sci.* 60 (2000), 510–539.
13. A. Gupta, Improved bandwidth approximation for trees and chordal graphs, *Journal of Algorithms* 40 (2001), 24–36.
14. J.E. Hopcroft and R.E. Tarjan, Dividing a graph into triconnected components, *SIAM J. Comput.* 2 (1973), 135–158.
15. O. Kariv and S.L. Hakimi, An algorithmic approach to network location problems, I: the p-centers, *SIAM J. Appl. Math.* 37 (1979), 513–538.
16. D. Kratsch, H.-O. Le, H. Müller, E. Prisner, and D. Wagner, Additive tree spanners, *SIAM J. Discrete Math.* 17 (2003), 332-340.
17. R. Krauthgamer and J.R. Lee, The intrinsic dimensionality of graphs, *STOC* 2003, pp. 438–447.
18. R. Krauthgamer, N. Linial, and A. Magen, Metric Embedding – Beyond one-dimensional distortion, *Discrete and Computational Geometry* 31 (2004), 339–356.
19. A.L. Liestman and T. Shermer, Additive graph spanners, *Networks,* 23 (1993), 343-364.
20. N. Linial, E. London, and Y. Rabinovich, The geometry of graphs and some its algorithmic applications, *Combinatorica* 15 (1995), 215–245.
21. T.A. McKee, personal communication to E. Prisner, 1995.
22. G.L. Miller and V. Ramachandran, A new graph triconnectivity algorithm and its parallelization, *Combinatorica* 12 (1992), 53–76.
23. D. Peleg and A.A. Schäffer, Graph Spanners, *J. Graph Theory*, 13(1989), 99-116.
24. D. Peleg and J.D. Ullman, An optimal synchronizer for the hypercube, *PODC* 1987, 77–85.
25. E. Prisner, Distance approximating spanning trees, *STACS'97*, LNCS 1200, 1997, pp. 499–510.
26. P.H.A. Sneath and R.R. Sokal, Numerical Taxonomy, *W.H. Freeman,* San Francisco, California, 1973.
27. D.L. Swofford and G.J. Olsen, Phylogeny reconstruction, In *Molecular Systematics (D.M. Hillis and C. Moritz, editors),* Sinauer Associates Inc., Sunderland, MA., 1990, 411–501.