Europ. J. Combinatorics (2000) **21**, 761–766
Article No. 10.1006/eujc.1999.0381
Available online at http://www.idealibrary.com on IDEAL

A Note on Distance Approximating Trees in Graphs

VICTOR CHEPOI AND FEODOR DRAGAN

Let $G = (V, E)$ be a connected graph endowed with the standard graph-metric d_G and in which longest induced simple cycle has length $\lambda(G)$. We prove that there exists a tree $T = (V, F)$ such that

$$|d_G(u, v) - d_T(u, v)| \leq \left\lfloor \frac{\lambda(G)}{2} \right\rfloor + \alpha$$

for all vertices $u, v \in V$, where $\alpha = 1$ if $\lambda(G) \neq 4, 5$ and $\alpha = 2$ otherwise. The case $\lambda(G) = 3$ (i.e., G is a chordal graph) has been considered in Brandstädt, Chepoi, and Dragan, (1999) *J.Algorithms* **30**. The proof contains an efficient algorithm for determining such a tree T.

All graphs $G = (V, E)$ occurring in this note are connected, undirected, loopless, and without multiple edges (but not necessarily finite). 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 is the length of a shortest (u, v)-path, and the *interval* between these vertices is the set

$$I(u, v) = \{w \in V : d_G(u, v) = d_G(u, w) + d_G(w, v)\}.$$

For each integer $k \geq 0$, let $B_k(u)$ denote the *ball* of radius k centered at u:

$$B_k(u) = \{v \in V : d_G(u, v) \leq k\}.$$

Let $S_k(u)$ denote the *sphere* of radius k centered at u:

$$S_k(u) = \{v \in V : d_G(u, v) = k\}.$$

A *leveling* of G with respect to some basepoint u is a partition of V into the spheres $S_k(u)$, $k = 0, 1, 2, \ldots$. We will say that a tree $T = (V, F)$ is a *distance δ-approximating tree* of a graph $G = (V, E)$ if $|d_G(x, y) - d_T(x, y)| \leq \delta$ for each pair of vertices $x, y \in V$. Finally, by $\lambda(G)$ we denote the length of a longest induced simple cycle of G.

THEOREM. *Given a graph $G = (V, E)$ with $\lambda(G) > 0$ and an arbitrary basepoint $u \in V$, there is a distance $\left(\left\lfloor \frac{\lambda(G)}{2} \right\rfloor + \alpha\right)$-approximating tree $T = (V, F)$ of G preserving the distances to u, where $\alpha = 1$ if $\lambda(G) \neq 4, 5$ and $\alpha = 2$ otherwise.*

PROOF. The case $\lambda(G) = 3$ has been considered in [1], whose idea is generalized here. Thus assume $\lambda(G) \geq 4$. Consider the leveling of G from u. For each $k \geq 0$ define a graph $\mathcal{S}_k$ with the kth sphere $S_k(u)$ as a vertex set. Two vertices $x, y \in S_k(u)$ $(k \geq 1)$ are adjacent in $\mathcal{S}_k$ if and only if they can be connected by a path outside the ball $B_{k-1}(u)$. Define a graph Γ whose vertex-set is the collection $\mathcal{S}$ of all connected components of the graphs $\mathcal{S}_k$, $k = 0, 1, 2, \ldots$, and two vertices are adjacent in Γ if and only if there is an edge of G between the corresponding components (see Figure 1 for an example). Clearly, two adjacent in Γ connected components lie in consecutive levels in the leveling of G.

0195–6698/00/060761 + 06 $35.00/0

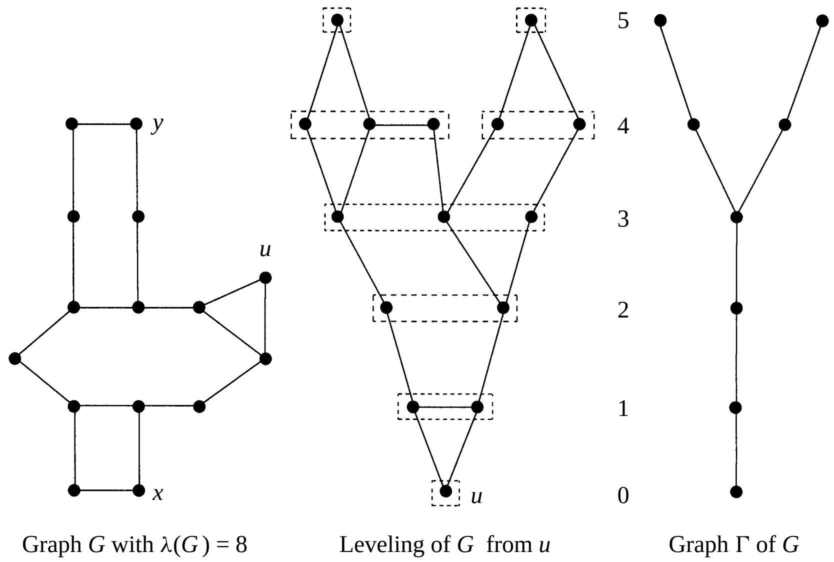

FIGURE 1.

Claim 1. Γ is a tree.

PROOF. It suffices to show that any connected component Q of $\mathcal{S}_k$ $(k > 0)$ is adjacent in Γ with exactly one connected component of $\mathcal{S}_{k-1}$. Suppose not, and let Q be adjacent to the connected components Q' and Q'' of $\mathcal{S}_{k-1}$. Then we will find the vertices $x' \in Q'$ and $x'' \in Q''$ which are adjacent to some vertices y' and y'' of Q. Take a path connecting the vertices y', y'' and lying outside the ball $B_{k-1}(u)$. Adding the edges $x'y'$ and $x''y''$, we will get a (x', x'')-path outside the ball $B_{k-2}(u)$, contrary to the assumption that x', x'' are in different connected components of $\mathcal{S}_{k-1}$. □

We will assume that Γ is rooted with $Q^* := \mathcal{S}_0 = \{u\}$ as a root.

Claim 2. If the vertices x, y (not necessarily distinct) belong to a common connected component of $\mathcal{S}_k$ and x', y' are some of their neighbors in the sphere S_{k-1}, then $d_G(x', y') \leq \left\lfloor \frac{\lambda(G)}{2} \right\rfloor$.

PROOF. First, we may assume that x' and y' are distinct non-adjacent vertices, for otherwise $d_G(x', y') \leq 1 \leq \lfloor \lambda(G)/2 \rfloor$. By the definition of $\mathcal{S}_k$, there is a path connecting x' and y' whose interior vertices are outside the ball $B_{k-1}(u)$. Among all such paths, choose a chordless one P_1. Since x' and y' connect to u by paths inside $B_{k-1}(u)$, there is a path connecting x' and y' whose interior vertices are inside $B_{k-2}(u)$. Among all such paths, choose a chordless one P_2. Then P_1 and P_2 together form a chordless cycle C passing via x' and y'. Thus $d_G(x', y') \leq d_C(x', y') \leq \lfloor \lambda(G)/2 \rfloor$. □

To construct a tree $T = (V, F)$, for a connected component Q of a graph $\mathcal{S}_k$ $(k \geq 1)$ we select a vertex v_Q of $S_{k-1}(u)$ which is adjacent in G to at least one vertex of Q, and make v_Q adjacent in T to all vertices of Q (see Figure 2). From Claim 1 we conclude that T is indeed a tree. Assume T is rooted at u. We denote the distance function in T by d_T. The *discrepancy function* $c(x, y)$ is now defined by

$$c(x, y) := |d_G(x, y) - d_T(x, y)|.$$

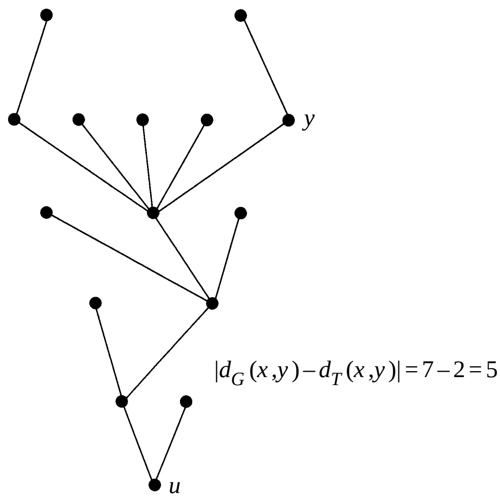

FIGURE 2. Distance 5-approximating tree for G from Figure 1.

By induction on $d_G(u,x)$ one can easily show that $d_G(u,x) = d_T(u,x)$ for every vertex x. From Claim 2 we deduce that if xy is an edge of T and not an edge of G, then $d_G(x,y) \leq \lfloor \frac{\lambda(G)}{2} \rfloor + 1$, i.e., $c(x,y) \leq \lfloor \frac{\lambda(G)}{2} \rfloor$. Conversely, let x and y be adjacent in G but not adjacent in T. If x, y lie in the same level, then they are in a common connected component Q, thus in T both x, y are adjacent to v_Q, showing that $d_T(x,y) = 2$. Now suppose that x and y lie in consecutive levels, say $x \in Q', y \in Q''$, where Q', Q'' are connected components of respective levels. Then necessarily $v_{Q'} \in Q''$, thus both $v_{Q'}$ and y are adjacent in T to $v_{Q''}$. This shows that $d_T(x,y) = 3$ in this case. Therefore, if xy is an edge of G or of T, then $c(x,y) \leq \lfloor \frac{\lambda(G)}{2} \rfloor$.

Finally pick the vertices x, y such that xy is an edge neither in G nor in T. Let $d_G(u,x) = n$, $d_G(u,y) = m$. Suppose that x belongs to the connected component Q' of $\mathcal{S}_n$ and y belongs to the connected component Q'' of $\mathcal{S}_m$. If $Q' = Q''$, then $d_G(x,y) \leq \lfloor \frac{\lambda(G)}{2} \rfloor + 2$ by Claim 2 and $d_T(x,y) = 2$ by the construction of T. Therefore $c(x,y) \leq \lfloor \frac{\lambda(G)}{2} \rfloor$ in this case. Thus Q' and Q'' are distinct. Let Q be the nearest common ancestor of Q' and Q'' in the tree Γ (as usual, the nearest common ancestor of two vertices in a rooted tree is the root of the smallest subtree that contains both vertices).

First assume that $Q \neq Q^*$ and $Q' \neq Q \neq Q''$. Denote by Q_0, Q_1, and Q_2 the neighbors of Q in the tree Γ on the paths connecting Q with Q^*, Q', and Q'', respectively. One can easily show that every (x,y)-path of G will share vertices with each connected component in $\mathcal{S}$ which lies on the unique path connecting Q' and Q'' in Γ. In particular, every shortest (x,y)-path will intersect the sets Q_1, Q, and Q_2. Since $d_G(x,z) \geq n-k$, $d_G(y,z) \geq m-k$ for every vertex $z \in Q$ (here $k := d(u,z)$), from this and Claim 2 we conclude that

$$n+m-2k \leq d_G(x,y) \leq n+m-2k+\left\lfloor \frac{\lambda(G)}{2} \right\rfloor + 2.$$

On the other hand,

$$d_T(x,y) = \begin{cases} n+m-2k & \text{if } v_{Q_1} = v_{Q_2}, \\ n+m-2k+2 & \text{otherwise.} \end{cases}$$

Comparing the expressions for $d_G(x,y)$ and $d_T(x,y)$, we obtain the desired estimation, except the case when $\lambda(G) > 5$, $d_T(x,y) = n+m-2k$, and $d_G(x,y) = n+m-2k+\lfloor \frac{\lambda(G)}{2} \rfloor + 2$.

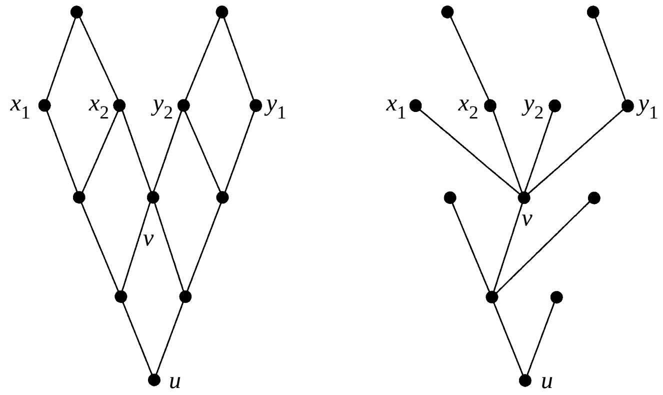

FIGURE 3. A graph G with $\lambda(G) = 4$ and its distance 4-approximating tree.

We assert that this cannot happen. Let x' and y' be closest to x and y vertices of Q in G, i.e., $d_G(x, x') = n-k$ and $d_G(y, y') = m-k$. By Claim 2 necessarily $d_G(x', y') = \lfloor \frac{\lambda(G)}{2} \rfloor + 2 \geq 5$, otherwise we are done. Pick the vertices $x'', y'' \in Q_0$, $w_1 \in Q_1, w_2 \in Q_2$ (they necessarily exist) such that $x'x'', x'w_1, y'y'', y'w_2 \in E$. Denote by z_1 and z_2 some neighbors of the vertex $v := v_{Q_1} = v_{Q_2}$ in the connected components Q_1 and Q_2. Finally, let z be a vertex of $I(x'', u) \cap I(y'', u)$ located as far as possible from u. Pick two shortest (x'', z)- and (y'', z)-paths $P(x'', z)$ and $P(y'', z)$. Among the paths connecting the vertices w_1, z_1 and w_2, z_2 outside the ball $B_k(u)$ let $P(w_1, z_1)$ and $P(w_2, z_2)$ have minimal length l_1 and l_2. Additionally assume that the pairs z_1, w_1 and z_2, w_2 have been selected so that l_1 and l_2 are as small as possible. Denote by C the simple cycle of G formed by these four paths and the edges $w_1x', x'x'', w_2y', y'y'', vz_1, vz_2$. Arguing as in the proof of Claim 2 and taking into account that $d_G(x', y') \geq 5$, we deduce that every possible chord of C has either the form vs with $s \in \{x', x'', y', y''\}$ or the form ab with $a \in P(x'', z)$ and $b \in P(y'', z)$. Since $d_G(x', y') = \lfloor \frac{\lambda(G)}{2} \rfloor + 2$, no pair of vertices $\{x', y'\}, \{x', y''\}, \{x'', y'\}$ lies on a common induced cycle. This implies that C is not induced, and, moreover, that v is adjacent to at least one of the vertices x', x'' and to at least one of the vertices y', y''. Then we get a path of length at most 4 between x' and y', contrary to the assumption that $d_G(x', y') \geq 5$. This establishes the case $Q \neq Q^*$ and $Q' \neq Q \neq Q''$. In the remaining cases the proof is similar, even simpler. □

We continue with two examples. First, we show that in the case $\lambda(G) = 4$ our method can construct distance 4-approximating trees of G. Note also that for the graph G with $\lambda(G) = 8$ from Figure 1 our method may construct a distance 5-approximating tree. Second, we present a chordal graph without distance 1-approximating trees, thus answering the question posed in [1]. Recall that G is a chordal graph iff $\lambda(G) = 3$.

EXAMPLE 1. Let G be a graph presented in Figure 3 and leveled with respect to the bottom vertex u. The graphs $\mathcal{S}_1$ and $\mathcal{S}_2$ are connected, while the graph $\mathcal{S}_3$ has two connected components $Q' = \{x_1, x_2\}$, $Q'' = \{y_1, y_2\}$. Since v is adjacent to x_2 and y_2, it may happen that $v_{Q'} = v = v_{Q''}$. But in this case $c(x_1, y_1) = 4$.

EXAMPLE 2. Consider a chordal graph G whose maximal cliques all have the same size $s \geq 4$. Additionally assume that every two maximal cliques can be connected by a chain of

TABLE 1.
Results.

$\lambda(G)$	δ
< 3	$=0$
$=3$	$=2$
$\in \{4, 5, 6, 7\}$	≤ 4
$\in \{8, 9\}$	≤ 5
...	...
$\in \{2k, 2k+1\}, k \ge 4$	$\le k+1$

maximal cliques such that every two consecutive cliques share an $(s-1)$-clique, and that G has diameter at least 4 (one can easily draw such examples; every 3-tree of diameter 4 has those properties). We claim that G does not contain distance 1-approximating trees. Suppose not, and let T be such a tree. Take a maximal clique R of G. In T either all vertices of R are adjacent to a vertex outside R, or there is a vertex of R which is adjacent to the remaining vertices of R. In both cases, R is embedded in T as a star. Now, if two maximal cliques share a triangle, then in T their stars must have a common center. From this we immediately conclude that T is a star. If we will take two vertices x, y with $d_G(x, y) = 4$, then obviously $d_T(x, y) \le 2$, contrary to the choice of T.

In Table 1 we summarize our results on distance δ-approximating trees for graphs with longest induced cycle of length $\lambda(G)$. Note that for chordal graphs our method is optimal in the sense that a chordal graph may not have a distance 1-approximating tree. It remains an interesting open question to characterize/recognize the graphs admitting distance 1-approximating trees.

REMARK 1. In the case of finite graphs, the proof of the theorem provides a linear algorithm for determining a tree T. The most expensive step is the construction of the connected components of the graphs $\mathcal{S}_k$ $(k = 0, 1, \ldots)$. We start from the sphere $S_n(u)$ of largest radius, find its connected components and contract each of them into a vertex. Then find the connected components in the graph induced by $S_{n-1}(u)$ and the set of contracted vertices, contract each of them and descend to the lower level, until we will come to the vertex u.

REMARK 2. Our result is in the vein of the following general result of Gromov (for a proof and definitions see Chapitre 2 in [2]): *Let (X, d) be a δ-hyperbolic metric space with at most $2^k + 2$ points for some positive integer k. Then there exist a tree $T = (X, F)$ rooted at u such that $d(x, u) = d_T(x, u)$ and*

$$d(x, y) - 2k\delta \le d_T(x, y) \le d(x, y)$$

for all $x, y \in X$.

REMARK 3. The result for the case $\lambda(G) = 4$ can be refined. Let G be a graph with $\lambda(G) = 4$. If G contains neither a house (i.e., the complement of an induced path on five vertices) nor a domino (the graph obtained from two induced 4-cycles by identifying an edge in one cycle with an edge in the other cycle) as induced subgraphs, then G admits a distance 2-approximating tree. If G does not contain only a domino as an induced subgraph, then it admits a distance 3-approximating tree. This follows from the proof of the theorem and from the fact (which is easy to prove) that in the case $v_{Q_1} = v_{Q_2}$, we have $d_G(x', y') \le \lfloor \lambda(G)/2 \rfloor = 2$, if G is house- and domino-free, and $d_G(x', y') \le \lfloor \lambda(G)/2 \rfloor + 1 = 3$, if G is domino-free.

REMARK 4. The result can be applied to provide efficient approximate solutions of several problems. In [1] we outlined how to compute the entries of the distance matrix of a chordal graph with an error of at most 2 in total optimal time $O(|V|^2)$ (it is unknown whether the exact calculation can be done within the same time bounds). More generally, the distance matrix and the diameter of a graph whose largest induced cycle has length $\lambda(G)$ can be computed in optimal time with an error given in the theorem. 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 the tree T constructed in the theorem, we will find an optimal covering radius r of T and a set of centers X. Then $|r - r^*| \leq \lfloor \lambda/2 \rfloor + \alpha$ and X can be taken as an approximate solution.

ACKNOWLEDGEMENTS

The authors are indebted to the referees for constructive suggestions and comments improving the presentation. The research of the second author was supported by the DFG.

REFERENCES

1. A. Brandstädt, V. Chepoi and F. Dragan, Distance approximating trees for chordal and dually chordal graphs, *J. Algorithms*, **30** (1999), 166–184.
2. E. Ghys and P. de la Harpe, Les Groupes Hyperboliques d'après Mikhael Gromov, *Prog. Math.*, **83** (1990).

Received 12 November 1998 and accepted 29 November 1999

VICTOR CHEPOI
Laboratoire d"Informatique de Marseille,
Université d'Aix Marseille II,
Faculté des Sciences de Luminy,
163, Avenue de Luminy, F–13288 Marseille Cedex 8,
France
E-mail: chepoi@lim.univ-mrs.fr
AND

FEODOR DRAGAN
Universität Rostock, Fachbereich Informatik,
Lehrstuhl für Theoretische Informatik,
Albert-Einstein Str. 21,
D–18051 Rostock,
Germany
E-mail: dragan@informatik.uni-rostock.de