Approximate Distance Labeling Schemes

Cyril Gavoille and Christophe Paul

LaBRI, Université Bordeaux I
351, cours de la Libération, 33405 Talence Cedex, France.

Abstract

In this paper we show how to label the nodes of every interval graphs of diameter D with integers of size $\lceil \log D \rceil + 1$ bits so that the distance up to an additive factor 1 between two distinct nodes can be computed in constant time from their labels only. The technique we propose may be generalized to prove additive $O(1)$-approximate distance labeling schemes for more general class of graphs including various intersecting classes of graphs, like chordal and circle graphs.

Key words: Distance, interval graphs, node-labeling

1 Introduction and Definitions

Given two nodes u, v of a graph G, we denote by $\text{dist}(u,v)$ the distance between u and v in G. A *distance labeling* for a graph G is a pair (L, f) where L is a labeling of the nodes of G by positive integers, and where f is an integer function (called the *distance function*) such that for every two labels L_u, L_v assigned to distinct nodes u, v, $f(L_u, L_v) = \text{dist}(u,v)$. The distance labeling is *additive k-approximate* if $\text{dist}(u,v) \leqslant f(L_u, L_v) \leqslant \text{dist}(u,v) + k$.

A *distance labeling scheme* for a class of graphs $\mathcal{F}$ is a pair $(\mathcal{L}, f)$ such that for every $G \in \mathcal{F}$, the pair $(\mathcal{L}(G), f)$ is a distance labeling for G. Note that the distance function of a distance labeling scheme, f, depends on $\mathcal{F}$ but is independent of any particular graph of $\mathcal{F}$. Our goal is, given a family of graphs $\mathcal{F}$, to minimize the length (in bits) of the labels used. As is it written in the first paper addressing this problem [Pel99], the labeling problem studied in this paper seems to have some relevance in the context of communication networks.

In this paper we present an additive 1-approximate distance labeling for interval graphs. The technique we use may be naturally extended to a treedecomposition providing an additive $O(1)$-approximate distance labeling schemes for more general class of graphs including various intersecting classes of graphs, like chordal and circle graphs.

Recently, [KKP00] proposed an exact distance labeling scheme for interval and permutation graphs with $O(\log^2 n)$ bit labels. Note that interval/permutation graphs do not include trees, so the recent $\Omega(\log^2 n)$-lower bound of Gavoille *et al.* [GPPR00] does not apply. Determining the complexity of label size for exact distance labeling in interval and permutation graphs is an open problem.

2 An Upper Bound for Interval Graphs

It is well-known that graphs without asteroidal triple (AT-free for short) have a *dominating pair*, that is a pair $\{x, y\}$ of nodes such that every path from x to y is a dominating set [COS97]. According to this path structure, we can label all the nodes of any AT-free graphs to give a simple additive 4-approximate distance labeling scheme. The length of these labels does not exceed $\lceil \log(D+1) \rceil$ bits, for graphs of diameter D. AT-free graphs includes interval, permutation, bounded tolerance, co-comparability graphs, ...

Let $\mathcal{U} = \{U_1, U_2, \ldots\}$ be any ordered set of cliques of a graph. $\mathcal{U}$ is a *clique chain* if $U_i \cap U_{i+1} \neq \varnothing$ for every i. $\mathcal{U}$ has the *interval property* if the set of indices of cliques of $\mathcal{U}$ containing any given node consists of an interval of consecutive integers. A clique chain having the interval property is called a *clique path*.

It is well-known that a graph G is an interval graph if and only if its set of maximal cliques can be ordered in a clique path. Let $\mathcal{C} = \{C_1, C_2, \ldots\}$ be such a clique path. For all $i < j$, we denote $u \in [C_i C_j]$ the fact that u belongs to every clique C_k with $i \leqslant k \leqslant j$.

Lemma 1 *For all $i < j < k$, there is no edge connecting any node of $C_i \setminus C_j$ to any node of $C_k \setminus C_j$.*

Let us define the subset of cliques $\mathcal{A} = \{A_1, A_2, \ldots\}$ as follows: $A_1 = C_1$, and for every $i \geqslant 1$, $A_{i+1} = C_k$ where $k = \max\{j \mid C_j \cap A_i \neq \varnothing\}$. By construction, $\mathcal{A}$ is a clique path. For convenience, we denote $A_i = C_{f(i)}$ for every i. By construction of $\mathcal{A}$, f is strictly increasing.

Lemma 2 *For all i, a, b such that $a \leqslant f(i)$, and $f(i+1) < b$, there is no edge connecting any node of C_a to any node of $C_b \setminus A_{i+1}$.*

Proof. Assume that there is an edge between $x \in C_a$, and $y \in C_b \setminus A_{i+1}$. By maximality of the cliques every edge is contained in a clique of $\mathcal{C}$. Let k be an index such that $(x, y) \in C_k$. By Lemma 1, since $a < f(i+1) < b$, there is no edge connecting $C_a \setminus A_{i+1}$ and $C_b \setminus A_{i+1}$. So, $x \in A_{i+1}$, and thus $x \in A_i$ as well since $x \in [C_a C_{f(i+1)}]$ and $a < f(i) < f(i+1)$. We have $k \leqslant f(i+1)$ because C_k intersects A_i, and $f(i+1)$ is, by construction of $\mathcal{A}$, the largest index such that $C_{f(i+1)}$ intersects A_i. By assumption, $f(i+1) < b$. Therefore $y \in [C_k C_b]$ implies $y \in C_{f(i+1)}$, a contradiction with the fact that $y \in C_b \setminus A_{i+1}$. ∎

For every node u, we set $b(u) = \min\{i \mid u \in C_i\}$, and let $r(u) = \min\{i \mid b(u) \leqslant f(i)\}$. Note that $r(u)$ exists because $C_t \in \mathcal{A}$, where $t = |\mathcal{C}|$. Moreover, since $C_1 \in \mathcal{A}$, it follows that for every $C_j \in \mathcal{C}$, there exists i such that either $j = f(i)$, or $f(i) < j < f(i+1)$. Since $A_i \cap A_{i+1} \neq \varnothing$, we have that C_j intersects A_i.

Lemma 3 *Let $u_0, \ldots, u_k$ be any path between u_0 and u_k. For every t such that $0 \leqslant t < k$, $r(u_{t+1}) \leqslant r(u_t) + 1$.*

Proof. Let $a = b(u_t)$ and $b = b(u_{t+1})$. We have $u_t \in C_a$, and $u_{t+1} \in C_b$. Let $i = r(u_t)$. $a \leqslant f(i)$ by definition of $r(u_t)$. Assume $b > f(i+1)$. It follows that $u_{t+1} \notin C_{f(i+1)}$, i.e., $u_{t+1} \in C_b \setminus A_{i+1}$. By Lemma 2, there is no edge between u_t and u_{t+1}, a contradiction. Therefore, $b \leqslant f(i+1)$, i.e., $b(u_{t+1}) \leqslant f(r(u_t)+1)$. Recall that $b(w) \leqslant f(i)$ implies $r(w) \leqslant i$. So, $r(u_{t+1}) \leqslant r(u_t) + 1$. ∎

Lemma 4 *Let x, y be two distinct nodes such that $r(x) \leqslant r(y)$. Let us set $z = 0$ if $x \in A_{r(x)}$, and $z = 1$ otherwise. Then $r(y) - r(x) + z \leqslant \operatorname{dist}(x, y) \leqslant r(y) - r(x) + z + 1$.*

Proof. Because $\mathcal{A}$ is a clique path, $A_i \cap A_{i+1} \neq \varnothing$ for every i. Hence, it suffices to take a node $w_i \in A_i \cap A_{i+1}$, for every $i \in \{r(x), \ldots, r(y) - 1\}$, and any node $w \in A_{r(x)}$ adjacent to x in order to form a path $P = (x, w, w_{r(x)}, w_{r(x)+1}, \ldots, w_{r(y)-1}, y)$ between x and y. The length of P is at most $r(y) - r(x) + 2$. We remark that if $x \in A_{r(x)}$ ($z = 0$), x is adjacent to any node of $A_{r(x)} \cap A_{r(x)+1}$. In particular x is adjacent to $w_{r(x)}$ allowing to remove w from P. So, in general, $\operatorname{dist}(x, y) \leqslant r(y) - r(x) + 1 + z$.

Let $P' = (x = u_0, \ldots, u_\delta = y)$ be a shortest path between x and y. Note that $z \in \{0, 1\}$, so $u_z \in \{u_0, u_1\}$. By application of Lemma 3, $r(u_\delta) \leqslant r(u_z) + \delta - z$, showing that the length of P', $\delta = \operatorname{dist}(x, y)$, satisfies $\delta \geqslant r(y) - r(u_z) + z$. It suffices now to remark that $r(u_z) = r(x)$. Indeed, the case $z = 0$ is done by definition. Otherwise, if $x \notin A_{r(x)}$, then $b(x) < f(r(x))$. It follows by Lemma 1 that there is no edge between $x \in C_{b(x)} \setminus A_{r(x)}$, and any node of $C_i \setminus A_{r(x)}$, for any $i > f(r(x))$. So, $b(u_1) \leqslant f(r(x))$, and thus $r(u_1) = r(x)$. ∎

Lemma 5 *If G is of diameter D then $|\mathcal{A}| \leqslant D$.*

An immediate consequence of Lemma 4 and of Lemma 5, is that:

Theorem 1 *There exists an additive* 1*-approximate distance labeling scheme for the class of interval graphs of diameter* D *using labels of size* $\lceil \log D \rceil + 1$*. Moreover the distance function is computable in constant time.*

3 A Lower Bound for General Graphs

Let us show that the previous distance labeling scheme for interval graphs is optimal in the length of the labels. Actually the lower bound we present applies for any family of graphs.

Theorem 2 *Let* k *be any integer,* $k \geqslant 0$*. Let* G *be a graph of diameter* D*,* $D \geqslant 3+k$*. Every additive* k*-approximate distance labeling on* G *uses labels of size at least* $\lceil \log \lceil (D+1)/(k+1) \rceil \rceil$*.*

Proof. Let (L, f) be any additive k-approximate distance labeling on G, a graph of diameter D, with $D \geqslant 3+k$ and $k \geqslant 0$. It suffices to show that $N = |\{L_u \mid u \in V(G)\}|$, the number of different labels used for G, is at least $(D+1)/(k+1)$. Let P be any shortest path from two nodes at distance D. P contains $D+1$ nodes. If $N < (D+1)/(k+1)$, there exists two nodes of P at distance at least $k+1$, say x and y, with $L_x = L_y$. Since $D \geqslant 3+k$, there exists $z \in P \setminus \{x, y\}$ such that $|\mathrm{dist}(z, x) - \mathrm{dist}(z, y)| > k$ (in particular a neighbor of x or of a neighbor y): a contradiction. ∎

References

[COS97] D.G. Corneil, S. Olariu, and L. Stewart, *Asteroidal triple-free graphs*, SIAM J. Discrete Math., vol. 10, 1997, pp. 399–430.

[GPPR00] C. Gavoille, D. Peleg, S. Pérennes, and R. Raz, *Distance labeling in graphs*, Research Report RR-1239-00, LaBRI, University of Bordeaux, 351, cours de la Libération, 33405 Talence Cedex, France, April 2000.

[KKP00] M. Katz, N. Katz, and D. Peleg, *Distance labeling schemes for well-separated graph classes*, in STACS '00, vol. 1770 of LNCS, Springer, February 2000, pp. 516–528.

[Pel99] D. Peleg, *Proximity-preserving labeling schemes and their applications*, in WG '99, vol. 1665 of LNCS, Springer, June 1999, pp. 30–41.