Split Decomposition and Distance Labelling: An Optimal Scheme For Distance Hereditary Graphs

Cyril Gavoille Christophe Paul

March 27, 2001

1 Introduction

A distance labelling scheme is a distributed data-structure designed to answer queries about distance between any two vertices of a graph G. The data-structure consists in a label $L(x,G)$ assigned to each vertex x of G such that the distance $d_G(x,y)$ between any two vertices x and y can be estimated as a function $f(L(x,G),L(y,G))$. Two problems can be considered: exact distance labelling [4] and approximate distance labelling [3]. In the further one, we look for an exact value of the distance while in the later one, we estimate the distance within a multiplicative factor $s \geqslant 1$ and/or with an additive constant $r \geqslant 0$ (i.e., $d_G(x,y) \leqslant f(L(x,G),L(y,G)) \leqslant s \cdot d_G(x,y) + r$). We are interested in finding labelling scheme that assigns shortest labels (expressed in bits) for every graphs of a given family. This paper show how the use of the well known split decomposition [2] of graphs can help to design distance labelling schemes. Indeed, applying this technique enable us to apply any distance labelling scheme designed for trees to the more general family of graphs called distance hereditary graphs.

2 Split decomposition

In this paper, we consider undirected graphs and use classical notations: $N(x)$ denotes the set of neighbors of x and $N(S)$, where S is a set of vertices, denotes the set of vertices that have a neighbor in S.

A *split* [2] of a graph $G=(V,E)$ is a partition of V into two sets V_1 and V_2 such that for any $x \in \tilde{V_1} = V_1 \cap N(V_2)$ and $y \in \tilde{V_2} = V_2 \cap N(V_1)$, $(x,y) \in E$ and $|V_1| > 1$, $|V_2| > 1$. If a graph has no split, then it is called *prime*.

When a graph has a split (V_1,V_2), it can be decomposed into two graphs G_1 and G_2. The graph G_1 (resp. G_2) is induced by $V_1 \cup \{v_1\}$ (resp. $V_2 \cup \{v_2\}$) where v_1 (resp. v_2) is a virtual vertex adjacent to all the nodes of $\tilde{V_1}$ (resp. $\tilde{V_2}$). The final graphs created by the split decomposition are the *split components*. Cunninghan [2] showed that any connected graphs has a unique split decomposition into prime graphs, stars and cliques with a minimum number of split components. It is also easy to prove that the number of splits of a graph is bounded by its number of vertices [2].

To get this canonical decomposition, at each step of the decomposition process we have to choose a maximal split (i.e., maximal w.r.t. the inclusion of $\tilde{V}_1 \cup \tilde{V}_2$). Let us now present a transformation of an arbitrary graph $G = (V, E)$ into a tree-like weighted graph $T(G)$ such that for any pair of vertices x and y of V, $d_G(x, y) = d_{T(G)}(x, y)$. The transformation is a representation of the split decomposition process. It first appears in these terms in [6]. Let (V_1, V_2) be a split of G. Partition G into $G_1 = G[V_1]$ and $G_2 = G[V_2]$ and add the virtual vertices v_1 and v_2 adjacent respectively to $\tilde{V}_1$ and $\tilde{V}_2$. Then also add a virtual edge between v_1 and v_2. Repeat recursively the process on G_1 and G_2 until getting prime graphs. Notice that the split components are exactly the connected components obtained by removing the virtual edges of $T(G)$.

In order to leave to distance unchanged, we have to put weights on the edges of $T(G)$: $\omega(e) = -1$ if e is a virtual edge; $\omega(e) = 1$ otherwise.

Lemma 1 *Let x and y be two vertices of a graph G. Then $d_G(x, y) = d_{T(G)}(x, y)$, where $T(G)$ is the tree-like decomposition associated with G.*

Since the number of splits is bounded by the number of vertices, the size of $T(G)$ is linear in the size of G. Therefore if we are able to design a distance labelling scheme of size S for $T(G)$, it can be directly applied to G.

3 A distance labelling scheme for distance hereditary graphs

As an example, let us consider the family of completely decomposable graphs. These graphs are called *distance hereditary graphs* [5]: in any connected induced subgraph H of a distance hereditary graph G, $d_H(x, y) = d_G(x, y)$ for any two vertices $x, y \in V(H)$. Since they are completely decomposable graphs, the split components are either cliques or stars.

For the case of distance hereditary graph we can continue the transformation of G to get a weighted tree denoted by $\tilde{T}(G)$. It suffices to replace any k-clique with $k \geqslant 3$ by a star with k leaves (the vertices of the clique) centered on a *virtual clique vertex*. To complete the transformation, we have to weight all the edges of the new stars by $1/2$. Notice that no edge of the original k-clique is a virtual edge and thus is valued by 1. Since any edge of this k-clique is replaced by the edges of weight $1/2$, the distances leave unchanged.

Given an integer $W > 0$, let us denote by W-tree any weighted tree whose edge cost is a non-null integer, and such that the weighted diameter is at most W. We need the following result to complete our result.

Lemma 2 ([3]) *There exists a $(1 + 1/\log W)$-multiplicative (resp. exact) distance labelling scheme using labels of size $O(\log n \cdot \log\log W)$ bits (resp. $O(\log n \cdot \log W)$ bits) for the family of W-trees with at most n vertices. Moreover the scheme is polynomial time constructible, and under a word-RAM model of computation with standard arithmetic operations, the distance can be computed in constant time if $W = n^{O(1)}$.*

Now, consider the weighted tree $\tilde{T}'(G)$ composed of $\tilde{T}(G)$ in which all the weights have been incremented by two and then doubled. Formally, for any edge e of $\tilde{T}(G)$, the weight $\tilde{\omega}(e)$ is transformed in the new weight of $\tilde{T}'(G)$ setting

$\tilde{\omega}'(e) = 2(\tilde{\omega}(e)+2)$. By Lemma 1, for all $x, y \in V(G)$, $d_G(x,y) = d_{\tilde{T}(G)}(x,y) = d_{\tilde{T}'(G)}(x,y)/2 - 2$. Now observe that, since the number of splits is bounded by n, the size of $\tilde{T}'(G)$ cannot exceed $2n$, moreover $\tilde{\omega}'(e) \in \{1, \ldots, 6\}$. Thus $\tilde{T}'(G)$ is a $12n$-tree since its weighted diameter is bounded by $12n$. Applying the scheme of Lemma 2, we have:

Theorem 1 *The family of distance hereditary graphs with at most n vertices enjoys a $(1 + 1/\log W)$-multiplicative (resp. exact) distance labelling scheme using labels of size $O(\log n \log\log n)$ bits (resp. $O(\log^2 n)$ bits). Moreover the scheme is polynomial time constructible, and under a word-RAM model of computation with standard arithmetic operations, the distance can be computed in constant time.*

Theorem 1 provides labels of optimal length because the family of distance hereditary graphs contains trees, and it is known that any exact distance labelling scheme (resp. $(1 + 1/\log n)$-multiplicative distance labelling scheme) on the family of unweighted trees with at most n vertices requires some labels of size $\Omega(\log^2 n)$ bits [4] (resp. $\Omega(\log n \log\log n)$ bits [3]).

4 Conclusion

In this paper, we showed that the distances in a graphs have a nice behavior with respect to the split decomposition. The result and the technique presented here may be generalized to other families of graphs: getting a tree representation of the prime components is enough. For example, if any prime component has a k-tree spanner [1], a similar result can be obtained but for an approximating distance labelling scheme.

References

[1] L. Cai and D.Corneil. Tree spanners. *SIAM Journal on Discrete Mathematics*, 8(3):359–387, 1995.

[2] W. Cunninghan. Decomposition of directed graphs. *SIAM Journal on Algebraic and Discrete Methods*, 3:214–228, 1982.

[3] Cyril Gavoille, Michal Katz, Nir A. Katz, Christophe Paul, and David Peleg. Approximate distance labeling schemes. Research Report RR-1250-00, LaBRI, University of Bordeaux, 351, cours de la Libération, 33405 Talence Cedex, France, December 2000.

[4] Cyril Gavoille, David Peleg, Stéphane Pérennes, and Ran Raz. Distance labeling in graphs. In 12^{th} *Symposium on Discrete Algorithms (SODA)*, pages 210–219. ACM-SIAM, January 2001.

[5] P. Hammer and F. Maffray. Completely separable graphs. *Discrete Applied Mathematics*, 27:85–99, 1990.

[6] JM. Lanlignel. Autour de la decomposition en coupes. PhD Thesis Preliminary Draft, 2001.