A note on distance labeling in planar graphs

Paweł Gawrychowski[1] and Przemysław Uznański[2]

[1]University of Haifa, Israel and University of Wrocław, Poland
[2]ETH Zürich, Switzerland

Abstract

A distance labeling scheme is an assignments of labels, that is binary strings, to all nodes of a graph, so that the distance between any two nodes can be computed from their labels and the labels are as short as possible. A major open problem is to determine the complexity of distance labeling in unweighted and undirected planar graphs. It is known that, in such a graph on n nodes, some labels must consist of $\Omega(n^{1/3})$ bits, but the best known labeling scheme uses labels of length $\mathcal{O}(\sqrt{n}\log n)$ [Gavoille, Peleg, Pérennes, and Raz, J. Algorithms, 2004]. We show that, in fact, labels of length $\mathcal{O}(\sqrt{n})$ are enough.

1 Introduction

A distance labeling scheme is an assignments of labels, that is binary strings, to all nodes of a graph G, so that the distance $\delta_G(u,v)$ between any two nodes u, v can be computed from their labels. The main goal is to make the labels as short as possible, that is to minimize the maximum length of a label. The secondary goal is to optimize the query time, that is the time necessary to compute $\delta_G(u,v)$ given the labels of u and v. We consider distance labeling for unweighted undirected graphs on n nodes. This was first considered by Graham and Pollak [1], who obtained labels of length $\mathcal{O}(n)$. The decoding time was subsequently improved to $\mathcal{O}(\log\log n)$ by Gavoille et al. [2], then to $\mathcal{O}(\log^* n)$ by Weimann and Peleg [3], and finally Alstrup et al. [4] obtained $\mathcal{O}(1)$ decoding time with labels of length $\frac{\log 3}{2}n + o(n)$. It is known that, in the general case, some labels must consist of at least $\frac{n}{2}$ bits [5, 6], so achieving sublinear bounds is not possible.

Better schemes for distance labeling are known for restricted classes of graphs. As a prime example, trees admit distance labeling scheme with labels of length $\frac{1}{4}\log^2 n + o(\log^2 n)$ bits [7], and this is known to be tight up to lower order terms [8]. In fact, any sparse graph admits a sublinear distance labeling scheme [9] (see also [10] for a somewhat simpler construction). However, the best known upper bound is still rather far away from the best known lower bound of $\Omega(\sqrt{n})$. An important subclass of sparse graphs are planar graphs, for which Gavoille et al. [2] constructed labeling schemes of length $\mathcal{O}(\sqrt{n}\log n)$. They also proved that in any such scheme some label must consist of $\Omega(n^{1/3})$ bits. Their upper bound of $\mathcal{O}(\sqrt{n}\log n)$ bits is also valid for weighted undirected planar graphs if the weights are polynomially bounded. Very recently, Abboud and Dahlgaard [11] showed a matching lower bound of $\Omega(\sqrt{n}\log n)$ in such a setting. However, determining the complexity of distance labeling in unweighted planar graphs remains to be a major open problem in this area.

We design a better distance labeling scheme for unweighted undirected planar graphs, where the labels consist of only $\mathcal{O}(\sqrt{n})$ bits. While this is only a logarithmic improvement, we believe lack of any progress in the last 12 years makes any asymptotic decrease desirable. Our improvement is based on a tailored version of the well-known planar separator lemma. We explain the idea in more detail after presenting the previous labeling scheme.

2 Previous scheme

We briefly recap the scheme of Gavoille et al. [2]. Their construction is based on the notion of separators, that is, sets of vertices which can be removed from the graph so that every remaining connected component consists of at most $\frac{2}{3}n$ nodes. By the classical result of Lipton and Tarjan [12] any planar graph on n nodes has such a separator consisting of $\mathcal{O}(\sqrt{n})$ nodes. Now the whole construction proceeds as follows: find a separator S of the graph, and let $G_1, G_2, \ldots$ be the connected components of $G \setminus S$. The label of $v \in G_i$ in G, denoted $\ell_G(v)$, is composed of i, $\ell_{G_i}(v)$ and the distances $\delta_G(v, u)$ for all $u \in S$ written down in the same order for every $v \in G$. A label of $v \in S$ consist of only the distances $\delta_G(v, u)$ for all $u \in S$.

The space complexity of the whole scheme is dominated by the space required to store $|S|$ distances, each consisting of $\log n$ bits, so $\mathcal{O}(\sqrt{n} \log n)$ bits in total. (All logarithms are in base 2.) The bound of $\mathcal{O}(\sqrt{n})$ on the size of a separator is asymptotically tight. However, the total length of the label of $v \in G$ (in bits) depends not on the size of the separator, but on the number of bits necessary to encode the distances from v to the nodes of the separator. If the separator is a simple cycle $(u_1, u_2, \ldots, u_{|S|})$ then $|\delta(v, u_i) - \delta(v, u_{i+1})| \leq 1$, for every $i = 1, 2, \ldots, |S| - 1$, and consequently writing down $\delta(v, u_1)$ explicitly and then storing all the differences $\delta(v, u_i) - \delta(v, u_{i+1})$ takes only $\mathcal{O}(\sqrt{n})$ bits in total. It is known that if the graph is triangulated, there always exists a simple cycle separator [13], so for such graphs labels of length $\mathcal{O}(\sqrt{n})$ are enough. We will show that, in fact, for any planar graph it is possible to select a separator so that the obtained sequence of differences is compressible. This will be done by triangulating the faces using carefully designed gadgets.

3 Improved scheme

We introduce the notion of weighted separators. Consider a planar graph, where every node has a non-negative weight and all these weights sum up to 1. Then a set of nodes is a weighted separator if after removing these nodes the total weight of every remaining connected component is at most $\frac{2}{3}$. We have the following well-known theorem (the result is in fact more general and allows assigning weights also to edges and faces, but this is not needed in our application):

Theorem 3.1 (see [13]). *For every 2-connected planar graph on n nodes having assigned nonnegative weights summing up to 1, there exists a simple cycle of length at most $2\sqrt{2dn}$ which is a weighted separator, where d is the maximum face size.*

Now we are ready to use this tool to show the main technical lemma of this section:

Lemma 3.2. *Any planar graph G has a separator S, such that*

$$\sum_i \log \delta_G(u_i, u_{i+1}) = \mathcal{O}(\sqrt{n})$$

for some ordering $u_1, u_2 \ldots, u_{|S|}$ of all vertices of S.

Before proving the lemma, we first describe a family of *subdivided cycles*. A subdivided cycle on $s \geq 3$ nodes, denoted D_s, consists of a cycle $C_s = (v_1, \ldots, v_s)$ and possibly some auxiliary nodes. D_3 and D_4 is simply C_3 and C_4, respectively. For $s > 4$, we add $\lceil \frac{s}{2} \rceil$ auxiliary vertices $u_1, \ldots, u_{\lceil \frac{s}{2} \rceil}$, and connect every v_i with $u_{\lceil \frac{i}{2} \rceil}$. To complete the construction, we recursively build $D_{\lceil \frac{s}{2} \rceil}$ and identify its cycle with $(u_1, \ldots, u_{\lceil \frac{s}{2} \rceil})$. (An example of such a subdivided cycle on 10 nodes is shown in Figure 1.) We have the following property.

Lemma 3.3. *For any $u, v \in C_s$, $\delta_{D_s}(u, v) \geq \log(1 + \delta_{C_s}(u, v))$.*

Proof. We apply induction on s. It is easy to check that the lemma holds when $s \leq 4$, so we assume $s \geq 5$. Let us denote $\delta_{D_s}(u, v) = d'$ and $\delta_{C_s}(u, v) = d$. We proceed with another induction on d. It is easy to check that the lemma holds when $d \leq 2$, so we assume $d \geq 3$. Consider a shortest path connecting u and v in D_s. If it consists of only auxiliary nodes except for the endpoints u and v, then we consider the immediate
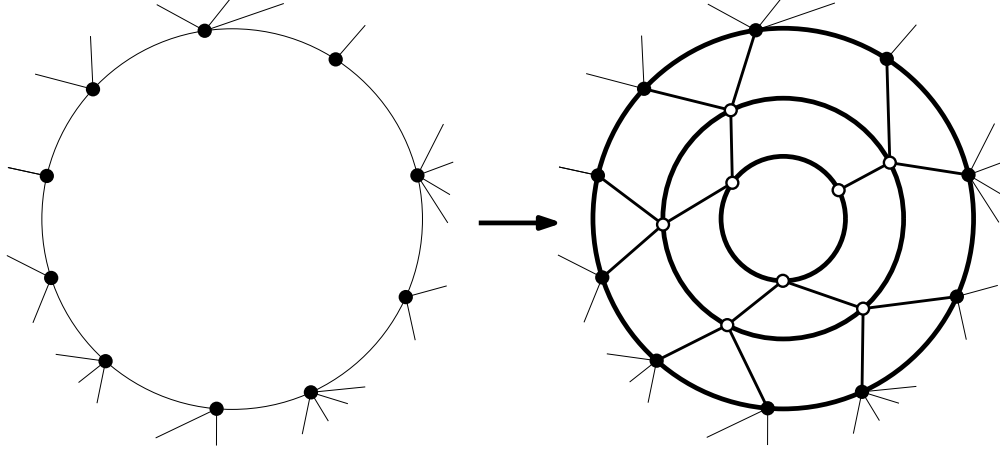
Figure 1: A face of size 10 is transformed by replacing C_{10} with D_{10} containing 8 new auxiliary nodes.

neighbors of u and v on the path, denoted u' and v', respectively. Since u' and v' must belong to the cycle of $D_{\lceil \frac{s}{2} \rceil}$ and the distance between them in the corresponding $C_{\lceil \frac{s}{2} \rceil}$ is at least $\lfloor \frac{d}{2} \rfloor$, by the inductive assumption:

$$d' \geq 2 + \log(1 + \lfloor d/2 \rfloor) = \log(4 + 4\lfloor d/2 \rfloor) \geq \log(1 + d).$$

Otherwise, let $u_0 = u, u_1, u_2, \ldots, u_s = v$ be all nodes of the cycle C_s appearing on the path, where $s \geq 2$. We have that:

$$d' = \sum_i \delta_{D_s}(u_i, u_{i+1})$$

and by the inductive assumption:

$$d' = \sum_i \delta_{D_s}(u_i, u_{i+1}) \geq \sum_i \log(1 + \delta_{C_s}(u_i, u_{i+1})).$$

Let us denote $d_i = \delta_{C_s}(u_i, u_{i+1}) \geq 1$ (as otherwise the path could have been shortened). By the triangle inequality, $\sum_i 1 + d_i \geq 1 + d$, so also $\prod_i 1 + d_i \geq 1 + d$. Therefore, $d' \geq \sum_i \log(1 + d_i) = \log \prod_i 1 + d_i \geq \log(1 + d)$. $\square$

Proof of Lemma 3.2. Let G' be the graph constructed from G by replacing every face (including the external face) with a subdivided cycle of appropriate size. More precisely, let $(v_1, v_2, \ldots, v_s)$ be a cyclic walk of a face of G. Note that nodes v_i are not necessarily distinct. We create a subdivided cycle D_s and identify its cycle C_s with $(v_1, v_2, \ldots, v_s)$. Clearly G' is also planar and each of its faces is either a triangle or a square. Additionally, every subdivided cycle is 2-connected, so also the whole G' is 2-connected. Since any subdivided cycle has at most as many auxiliary vertices as cycle vertices and lengths of all cyclic walks sum up to twice the number of edges, which is at most $3n - 6$, G' contains at most $n' = n + 2 \cdot (3n - 6) < 7n$ vertices.

We assign weights to vertices of G' in such a way that that every vertex that also appears in G has weight 1 and every new vertex has weight 0. By Theorem 3.1 there exists a weighted simple cycle separator S' in G' of size at most $2\sqrt{28n}$. Then $S = S' \cap G$ is a separator in G. Because S' is a simple cycle separator, $S = (u_1, u_2 \ldots, u_c)$, and u_i and u_{i+1} are incident to the same face of G that has been replaced with a subdivided cycle D_{s_i} such that S' connects u_i and u_{i+1} either directly or by visiting some auxiliary nodes of D_{s_i}, for every $i = 1, 2, \ldots, c$ (we assume $u_{c+1} = u_1$). Let v_i and v_i' denote nodes of D_{s_i} that have been identified with u_i and u_{i+1}, respectively. Then:

$$\sum_i \delta_{D_{s_i}}(v_i, v_i') \leq |S| = \mathcal{O}(\sqrt{n}).$$

3

By Lemma 3.3, $\delta_{D_{s_i}}(v_i, v_i') \geq \log \delta_{C_{s_i}}(v_i, v_i')$, so:

$$\sum_i \log \delta_G(v_i, v_{i+1}) \leq \sum_i \log \delta_{C_{s_i}}(v_i, v_i') \leq \mathcal{O}(\sqrt{n})$$

as required. □

Now we proceed to the main result of this section:

Theorem 3.4. *Any planar graph on n nodes admits a distance labeling scheme of size $\mathcal{O}(\sqrt{n})$.*

Proof. We proceed as in the previously known scheme of size $\mathcal{O}(\sqrt{n}\log n)$, except that in every step we apply our Lemma 3.2. In more detail, to construct the label of every $v \in G$ we proceed as follows. First, we find a separator $S = \{u_1, u_2 \ldots, u_c\}$ using Lemma 3.2. For every $v \in G$ we encode its distances to all nodes of the separator. This is done by first writing down $\delta_G(v, u_1)$ explicitly, and then $\delta_G(v, u_i) - \delta_G(v, u_{i-1})$ for $i = 2, \ldots, c$. All these numbers are encoded in binary, so by $|\delta_G(v, u_i) - \delta_G(v, u_{i-1})| \leq \delta_G(u_{i-1}, u_i)$ and the properties of our separator this encoding takes $\mathcal{O}(\sqrt{n})$ bits in total. Second, for every node we store the name of its connected component of $G \setminus S$. Third, we recurse on every connected component of $G \setminus S$ and append the obtained labels to the current labels. To calculate $\delta_G(v, v')$, we proceed recursively: we compute $\min_{u \in S}(\delta_G(v, u) + \delta_G(v', u))$ and then, if v and v' belong to the same connected component of $G \setminus S$, proceed recursively. The correctness is clear: either a shortest path between u and v is fully within one of the connected components, or it visits some $u \in S$, and in such case we can recover $\delta_G(v, u) + \delta_G(v', u)$ from the stored distances. The final size of every label is $\mathcal{O}(\sqrt{n} + \sqrt{\frac{2}{3}n} + \ldots) = \mathcal{O}(\sqrt{n})$ bits. □

References

[1] R. Graham, H. Pollak, On embedding graphs in squashed cubes, in: Graph Theory and Applications, Vol. 303 of Lecture Notes in Mathematics, Springer Berlin Heidelberg, 1972, pp. 99–110.

[2] C. Gavoille, D. Peleg, S. Pérennes, R. Raz, Distance labeling in graphs, J. Algorithms 53 (1) (2004) 85–112.

[3] O. Weimann, D. Peleg, A note on exact distance labeling, Inf. Process. Lett. 111 (14) (2011) 671–673.

[4] S. Alstrup, C. Gavoille, E. B. Halvorsen, H. Petersen, Simpler, faster and shorter labels for distances in graphs, in: 27th SODA, 2016, pp. 338–350.

[5] J. Moon, On minimal n-universal graphs, in: Proceedings of the Glasgow Mathematical Association, Vol. 7, Cambridge Univ Press, 1965, pp. 32–33.

[6] S. Kannan, M. Naor, S. Rudich, Implicit representation of graphs, SIAM J. Discrete Math. 5 (4) (1992) 596–603.

[7] O. Freedman, P. Gawrychowski, P. K. Nicholson, O. Weimann, Optimal distance labeling schemes for trees, CoRR abs/1608.00212.

[8] S. Alstrup, I. L. Gørtz, E. B. Halvorsen, E. Porat, Distance labeling schemes for trees, in: 43rd ICALP, 2016, pp. 132:1–132:16.

[9] S. Alstrup, S. Dahlgaard, M. B. T. Knudsen, E. Porat, Sublinear distance labeling, in: 24th ESA, 2016, pp. 5:1–5:15.

[10] P. Gawrychowski, A. Kosowski, P. Uznański, Sublinear-space distance labeling using hubs, in: 30th DISC, 2016, pp. 230–242.

[11] A. Abboud, S. Dahlgaard, Popular conjectures as a barrier for dynamic planar graph algorithms, CoRR abs/1605.03797.

[12] R. J. Lipton, R. E. Tarjan, Applications of a planar separator theorem, SIAM J. Comput. 9 (3) (1980) 615–627.

[13] G. L. Miller, Finding small simple cycle separators for 2-connected planar graphs, J. Comput. Syst. Sci. 32 (3) (1986) 265 – 279.