An Algorithm for Outerplanar Graphs with Parameter

Zhu Binghuan
Department of Computer Science
Zhongshan University, P.R.C.

Wayne Goddard
Department of Mathematics
Massachusetts Institute of Technology

Abstract

For n-vertex outerplanar graphs, it is proven that $O(n^{2.87})$ is an upper bound on the number of breakpoints of the function which gives the maximum weight of an independent set, where the vertex weights vary as linear functions of a parameter. An $O(n^{2.87})$ algorithm for finding the solution is proposed.

1 Introduction

We consider only undirected graphs $G = (V, E)$. A set $S \subseteq V$ is independent if whenever $u \in S$ and $v \in S$ it holds that $(u, v) \notin E$. Finding the maximum size of an independent set is one of the canonical NP-complete problems [3]. Nevertheless, if the graph has some nice structure then we may hope to make progress.

The independence problem may be extended: given $w\colon V \to \mathbf{R}^+$, we must find an independent set S of G such that $\sum_{v \in S} w(v)$ is maximized. We call this sum the weight of the set S, and abbreviate it $\mathrm{MISW}(G)$. The problem can be further extended: given, for each $v \in V$, a weight function $w_v(t)$ of the parameter t, we must determine, as a function of t, an independent set of maximum weight $\mathrm{MISW}(G, t)$.

Even if $w_v(t)$ is a linear function of t for each v, the problem of determining $\mathrm{MISW}(G, t)$ is complicated. In [4] it was observed that $\mathrm{MISW}(G, t)$ is a piecewise linear concave function for each G. It is not easy to estimate the number of breakpoints of this curve, though if there is some restriction on the structure of the graph we might obtain meaningful results.

In [2] Fernández-Baca and Slutzki discussed a number of algorithms for a tree with parameter. The main idea was to make use of the centroid decomposition technique. Here we use this method to solve the parametrized independence problem for outerplanar graphs.

A planar embedding of a graph is a drawing in the plane so that curves corresponding to edges intersect only at the point corresponding to a vertex mutually incident with them. In a natural way, this embedding breaks up the plane into connected *regions*: the unbounded one is called *exterior* and the remaining ones are *interior*. The *boundary* of a region is the set of vertices that touch it. A graph is *outerplanar* if it can be embedded in the plane such that every vertex lies on the unbounded exterior region.

We show that, for any outerplanar graph G, the number of breakpoints of $\mathrm{MISW}(G, t)$ is $O(n^{2.87})$, and provide an algorithm to calculate this curve. This also solves the dual of this viz. the parametrized vertex cover problem; for a set T is a vertex cover if and only if $V - T$ is an independent set.

2 Decomposition

We discuss here the application of the centroid decomposition technique to outerplanar graphs. A quadratic time algorithm to do this is presented as well.

Throughout this section, let $G = (V, E)$ denote an outerplanar graph on n vertices. We assume also that we have an embedding of G in the plane such that every vertex is on the exterior region. Note that in such an embedding, the boundary of every interior region is a cycle. In fact, there is a 1–1 correspondence between interior regions and induced cycles, so that knowing the latter is equivalent to knowing the embedding.

For any subgraph H, let $m(H)$ denote the number of vertices in the largest component. Then we say that a vertex v is a *central vertex* if $m(G - \{v\}) \leq n/2$. (If $S \subset V$ then $G - S$ denotes the graph induced by the vertex set $V - S$.) Similarly, we say that an interior region $\mathcal{R}$, with boundary C, is a *central region* if $m(G - C) \leq n/2$. Further, the components of $G - C$ are termed the *ears* of the region $\mathcal{R}$.

Now, any two regions in an outerplanar embedding that are not disjoint have in common, either a single vertex, or two vertices joined by an edge. Thus:

Lemma 1 *An ear can be separated from the rest of the graph by the removal of at most two vertices, and if two are required, they are adjacent.*

The decomposition is based on the following lemma.

Lemma 2 *Let G be an outerplanar graph on n vertices. Then there exists either a central region or a central vertex. Further, we can find such a region/vertex in time $O(n^2)$.*

Proof. For the purposes of this proof, X_i will denote either a vertex or a region, and $G - X_i$ will mean G with the vertex, or the boundary of the region, removed (as the case may be).

Start with any (interior) region or cut-vertex in the largest component of G and call it X_1. If $m(G - X_1) \leq n/2$ then we are done. Otherwise, $G - X_1$ has a component H_1 of size more than $n/2$.

We must find a region or cut-vertex X_2 such that $m(G - X_2) < m(G - X_1)$. In fact, as $G - V(H_1)$ contains less than $n/2$ vertices, we need only choose X_2 such that: (i) it separates $H_1 - X_2$ from the rest of G, and (ii) it contains at least one vertex of H_1.

This is easily done. Let F_1 denote the graph induced by the vertices of H_1 and X_1. We handle two cases.

1. X_1 *is a vertex:* If X_1 has degree one in F_1, then let X_2 be its neighbor. Otherwise, X_1 is adjacent to u and v (say) in F_1. But then u and v were connected in H_1, so that X_1 and u (say) lie together on the boundary of (at least) one (interior) region X_2 of F_1.

2. X_1 *is a region:* If X_1 has only one vertex x which is adjacent to a vertex of H_1, then let $X_2 = x$. Otherwise, X_1 has exactly two vertices u and v (and these are consecutive) which are adjacent to vertices of H_1 (though maybe

to the same vertex). Then u and v lie together on the boundary of another region X_2 of F_1.

It is easily checked that, under each circumstance, conditions (i) and (ii) are satisfied.

Thus we may iteratively construct $X_1, X_2, \ldots, X_r$ such that $m(G-X_1) > m(G-X_2) > \ldots > m(G-X_r)$, until $m(G-X_r) \leq n/2$ as required.

Clearly $r \leq n/2$. Thus to show that finding the central vertex or region X_r takes $O(n^2)$ time, we must demonstrate that each iteration takes only $O(n)$ time. Now, an outerplanar graph has a linear number of edges. Thus it is sufficient to know that you can in time linear in the number of edges: (a) find connected components, and (b) find a region whose boundary contains two specific consecutive vertices. If x and y are the specified vertices, then (b) can be solved by finding a shortest path from x to y in the graph with the edge $e = xy$ removed. Thus both tasks can be performed in the required $O(n)$ time (by breadth-first-searches, for example—see [1]). QED

3 The Independence Number

Here we discuss the number of breakpoints of $\mathrm{MISW}(G,t)$ of an outerplanar graph G, and outline an algorithm for finding $\mathrm{MISW}(G,t)$ and the independent set of maximum weight.

Let $G = (V, E)$ be an outerplanar graph on n vertices. For each $v \in V$, we have a weight function $w_v(t) = a_v t + b_v$, where $a_v, b_v \geq 0$. Let $b(G)$ denote the number of breakpoints of $\mathrm{MISW}(G,t)$ where $t \in \mathbf{R}^+$. Then we define $b(n)$ to be the maximum of $b(G)$ taken over all n-vertex outerplanar graphs. Thus we have $b(m) \leq b(n)$ if $m \leq n$.

The following is easily shown:

Lemma 3 [2] *Let $b(h)$ denote the number of breakpoints of a piecewise linear function h with domain $\mathbf{R}^+$, and let f and g be such functions. Then $b(f+g) \leq b(f) + b(g)$, and $b(\max\{f,g\}) \leq b(f) + b(g) + 1$.*

Then we may establish:

Lemma 4 *Let G be defined as above, and let $S \subset V$. Let the components of $G - S$ have sizes $m_1, m_2, \ldots, m_r$, and let a be the number of subsets of S that are independent in G. Then*

$$b(G) \leq a \cdot (\, b(m_1) + b(m_2) + \ldots + b(m_r) \,) + a - 1.$$

PROOF. Let G_i be the component of $G-S$ of size m_i, $(i = 1, \ldots, r)$. Let T_j be one of the independent subsets of S. Then, to find the independent set of maximum weight among those whose intersection with S is T_j, one need only find (separately) the maximum weight independent set in each of the subgraphs $G_i - N(T_j)$ (where $N(T_j)$ denotes the set of all vertices adjacent to some vertex of T_j). Thus $\mathrm{MISW}(G, t)$ can be computed as the maximum of a sums. By the above lemma, each sum has at most $\sum_{i=1}^{r} b(m_i)$ breakpoints, so that, by the above lemma again, the proof is complete. QED

We are now in a position to prove our main result.

Theorem 1 *It holds that $b(n)$ is $O(n^{\varepsilon})$ where $\varepsilon = 2.87$.*

PROOF. We prove by induction that $b(n) \leq dn^{\varepsilon}$, where d is some constant. Our main task is to show using Lemma 2 that there always a good choice of S for invoking Lemma 4.

Consider first the case where there is a *central vertex* v. By convexity the worst that can happen is that you have only two components in $G-S$. Thus by Lemma 4

$$b(G) \leq 2(b(n/2) + b(n/2)) \leq 0.9998dn^{\varepsilon},$$

with some to spare.

So we may assume that you have a *central region* $\mathcal{R}$ with boundary C. Let the maximum size of an ear be c. Depending on c/n, one of several tactics is appropriate.

1. *Isolate one ear:* Let S be the vertex or pair of adjacent vertices whose removal separates the largest ear from the rest of the graph. Then

   $$b(G) \leq 3(b(c) + b(n-c)) + 2 \leq 0.9998dn^{\varepsilon},$$

   provided $c \geq 0.3575n$.

2. *Isolate one ear and remove a vertex on C:* Let T be the vertex or pair of adjacent vertices whose removal separates the largest ear from the rest of the graph G'. Then remove another vertex v on C which splits G' as equally as possible into two components of size $x \geq y$. We can ensure that $x - y \leq c$, so that $x \leq n/2$. With two applications of Lemma 4 we get:

   $$\begin{aligned} b(G) &\leq 3(b(c) + 2(b(x) + b(y)) + 1) + 2 \\ &\leq 3dc^{\varepsilon} + 6d(n/2)^{\varepsilon} + 6d(n/2 - c)^{\varepsilon} + 5 \\ &\leq 0.9998dn^{\varepsilon}, \end{aligned}$$

   provided $0.2375n \leq c \leq 0.3575n$.

3. *Remove three vertices on C:* Consider three vertices well-placed on C whose removal yields components with sizes $x \geq y \geq z$. Combining the two smaller pieces first we get that

$$\begin{aligned} b(G) &\leq 4(b(x) + 2(b(y) + b(z)) + 1) + 3 \\ &\leq 4dx^{\varepsilon} + 8dy^{\varepsilon} + 8dz^{\varepsilon} + 7. \end{aligned}$$

We can ensure that $x - z \leq c$. Thus the worst that can happen, for a fixed x, is that $x = y$ or $x = z + c$. Some calculations show that for $c = 0.2375n$, the RHS has a maximum of at most $0.9998dn^{\varepsilon}$, obtained at $x = y = z + c$. Thus $b(G) \leq 0.9998dn^{\varepsilon}$, provided $c \leq 0.2375n$.

This completes the proof. QED

The actual curve $\mathrm{MISW}(G, t)$ can be computed in a similar amount of time. If $T(n)$ denotes the most time needed for an n-vertex outerplanar graph, then it satisfies the same recursive upper bounds used above for $b(n)$, except that there is an $O(n^2)$ overhead for the decomposition and $O(b(n))$ overhead for the bookkeeping.

4 Comments

The constant $\varepsilon = 2.87$ can easily be improved, though for each small improvement the complexity of the case study increases.

Nevertheless, there is a limit to how far we can get with an oblivious induction process. Let G be the outerplanar graph formed by taking a cycle and triangulating it such that the maximum degree is four. We claim that (one of) the best cut-up scheme(s) for G is to remove one copy of K_2 leaving two equal pieces. The best ε such that $3((1/2)^{\varepsilon} + (1/2)^{\varepsilon}) \leq 1$ is $\varepsilon = 1 + \log_2 3 \approx 2.58$. So a pure induction argument could not be used to show anything better than $b(n) \leq n^{2.58}$.

It would be interesting to determine the actual value of $b(n)$ for outerplanar graphs. This would, of course, entail finding a lower bound for this quantity.

Another question is: for what other classes of graphs can this decomposition approach be used to show polynomial upper bounds on $b(G)$? In particular, for what classes of graphs does there exist a constant k such that any graph G in this class has a set S of size at most k such that $m(G - S) \leq n/2$?

Acknowledgements

The authors would like to thank G.W. Peck for considerable assistance and guidance. The second author was supported in part by NSF grant DMS-8606225.

References

[1] T.H. Cormen, C.E. Leiserson and R.L. Rivest, "Introduction to Algorithms," MIT Press, Cambridge, 1990.

[2] D. Fernández-Baca and G. Slutzki, "Solving Parametric Problems on Trees," J. Algorithms **10** (1989), 381–402.

[3] M. Garey and D. Johnson, "Computers and Intractability: A Guide to the Theory of NP-Completeness," Freeman, San Francisco, 1979.

[4] D. Gusfield, "Sensitivity Analysis for Combinatorial Optimization," Memo UCB/ERL M80/22, Electronics Research Laboratory, University of California, Berkeley, 1980.