Minimum feedback vertex sets in circle graphs

Fanica Gavril
Computer Science Dept., Technion, Haifa 32000, Israel, gavril@cs.technion.ac.il

ABSTRACT: We describe a polynomial time algorithm to find a minimum feedback vertex set, or equivalently, a maximum induced forest, in a circle graph. The circle graphs are the overlap graphs of intervals on a line.

KEY WORDS: minimum feedback vertex set, maximum induced forest, circle graph

1. Introduction

We consider only finite graphs $G(V,E)$ with no parallel edges and no self-loops, where $V$ is the set of vertices and $E$ the set of edges. For $U \subseteq V$, $G(U)$ is the vertex subgraph induced by $U$. A tree is an acyclic connected graph. A forest is a collection of trees. A tree $t$ is rooted if it is directed and has exactly one vertex of in-degree zero called root. A cocomparability graph is the complement of a transitively orientable graph [3].

A feedback vertex set of an undirected graph $G$ is a vertex set whose deletion leaves an induced forest. The problem of finding a minimum feedback vertex set, or equivalently a maximum induced forest, of a graph $G$ is NP-complete for bipartite and planar graphs, but polynomial for interval, cocomparability, AT-free and chordal graphs [1,8,11,13].

A graph $G$ is an intersection graph of a family $S$ of subsets of a set if there is a one-to-one correspondence between the vertices of $G$ and the subsets in $S$ such that two vertices are adjacent iff their corresponding subsets in $S$ have a non-empty intersection. Interval graphs are intersection graphs of families of intervals on a line. Two intervals are said to overlap if they intersect and none is contained into the other. A graph $G$ is called a circle graph or overlap graph if there exists a family $I$ of intervals on a line and a one-to-one correspondence between the vertices of $G$ and the intervals in $I$ such that two vertices are adjacent iff their corresponding intervals in $I$ overlap. Circle graphs [2,5], have polynomial time algorithms for recognition, maximum clique and maximum independent set, while their minimum coloring and minimum covering by cliques problems are NP-complete [4,5,7,12]. These graphs are of interest in computer science, genetics and ecology [9,10].
the present paper we describe a polynomial time algorithm to find a maximum induced forest, or equivalently a minimum feedback vertex set, in a circle graph.

In a Euclidean plane $PL$, consider a line $L$ defined by $y=0$, drawn from left to right. On $L$, consider a family of closed intervals $I$. For an interval $[l,r] \in I$, we define an interval-curve $c$ in $PL$ as a continuous function $c:[l,r] \rightarrow \mathbb{R}^+$ having $c(l)=c(r)=0$: an interval-curve $c$ starts and ends at the endpoints of $[l,r]$ and is delimited by them. Consider a family $ff$ of interval-curves fulfilling that $\cup c \in ff$ is a continuous curve: $a=\cup_{c \in ff} c$ is called an interval-filament. An interval-filament is delimited in $PL$ by its two extreme endpoints in $L$, hence, if two intervals are disjoint, their interval-filaments do not intersect. Clearly, the union of two intersecting interval-filaments is an interval-filament. Let $V=\{v\mid i(v) \in I\}$ be a vertex set. For each interval $i(v) \in I$ we consider an interval-filament $a(v)$ connecting the left and right endpoints $l_{i(v)}, r_{i(v)}$ of $i(v)$; $FI=\{a(v)\mid i(v) \in I\}$ is a family of interval-filaments and its intersection graph $G(V,E)$ is an interval-filament graph, as introduced in Gavril [6].

Consider a circle graph $G(V,E)$ represented as the overlap graph of a family $I=\{i(v)\mid v \in V\}$ of closed intervals on a line $L$; w.l.o.g. we assume that no two intervals have a common endpoint. We denote $l_v=l_{i(v)}, r_v=r_{i(v)}$ and $EP=\{x\mid x=l_v \text{ or } x=r_v, v \in V\}$. In the plane $PL$ of $L$, above $L$, we connect the endpoints of each interval $i(v)$ by an interval-filament $a(v)$ consisting of a simple arc (Fig. 1a), such that if $i(u) \subset i(v)$, then the arcs $a(u), a(v)$ do not intersect. Clearly, $G$ is the intersection graph of this family of interval-filaments, since $a(u)\cap a(v)\neq\emptyset$ iff $i(u)\cap i(v)\neq\emptyset$, $i(u)\subset i(v)$ and $i(v)\subset i(u)$.

In Section 2 we analyze the structure of a maximum induced forest in a circle graph, using the above representation by interval-filaments. In Section 3 we describe a polynomial time algorithm to find a maximum induced forest in a circle graph.

### 2. Analysis of maximum induced forests in circle graphs

Let $G(V,E)$ be a circle graph represented as an intersection graph of interval-filaments, each filament $a(v)$ being an arc connecting the endpoints of an interval $i(v) \in I$. Consider a maximum induced forest $F_G$ of $G$ (Fig. 1b, without the dashed edges). Let $t_1, t_2, ..., t_r$ denote the trees of $F_G$ in order from left to right of the left endpoints of the intervals $i(t_j)$ delimiting the interval-filaments $\cup a(u) \cup_{u \in t_j}$. We take as root of $t_1$ a vertex $v_1$ having maximal interval $i(v_1)$. For every other tree $t_j$ of $F_G$ we take as root its vertex $v_j$ whose interval has leftmost left endpoint, hence $i(v_j)$ is maximal for $t_j$. The interval-
filaments corresponding to two trees $t_i, t_j$ of $FG$ do not intersect and their intervals are either disjoint or are contained one into another. Note that if $u$ is a successor of $w$ in some $t_j$, then $i(w)\subseteq i(u)$, otherwise $i(v_j)\subseteq i(u)$ contradicting the maximality of $i(v_j)$. Let $x$ be the endpoint of an interval-filament $a(u)$, $u \in t_1$, which is at the left of $i(v_2)$ and is closest to the left endpoint of $i(v_2)$. If $x=r_w$, we attach $v_2$ to $u$ by a dashed edge as a dummy son denoted $DR_u$, and if $x=l_w$, we attach $v_2$ to $u$ by a dashed edge as a dummy son denoted $DL_u$; in Fig. 1 $v_2=DR_u$, $v_3=DL_d$, $v_4=DR_w$. We assume now that $t_1$ includes $t_2$, and we continue recursively with $t_3$ and so on to $t_r$. In this way, $FG$ becomes a tree rooted at $v_1$, called oriented form of $FG$. Note that every vertex $u$ has at most two dummy sons, one $DR_u$ and one $DL_u$, and there are no endpoints of intervals $i(v), v \in FG$, between $r_u$ and $l_{DR_u}$ and between $l_u$ and $l_{DL_u}$. For the remaining of this Section we assume that $FG$ is in oriented (tree) form rooted at $v_1$; we denote by $FG(u)$ the subtree containing $u$ and its successors in $FG$.

The sons $u$ of a vertex $w$ in $FG$ can be partitioned as follows, according to whether interval $i(u)$ contains the right or left endpoint of $i(w)$:

$$SR_w(FG) = \{ u \mid u \text{ son of } w, \ r_w \in i(u) \text{ and } l_w \notin i(u) \} \cup \{DR_w\},$$

$$SL_w(FG) = \{ u \mid u \text{ son of } w, \ l_w \in i(u) \text{ and } r_w \notin i(u) \} \cup \{DL_w\}. $$

In each of the sets $SR_w(FG)$ and $SL_w(FG)$, the interval-filaments corresponding to their vertices (Fig. 1a) form a sequence $a(u_1), a(u_2), \ldots, a(u_s)$ of non-intersecting interval-filaments having $i(u_1)\subseteq i(u_2)\subseteq \ldots \subseteq i(u_s)$; if $w$ has a dummy son, then $u_1$ is the dummy son. In Fig. 1a, $SR_w(FG)=\{c,d,v_4\}$ and $i(v_4)\subseteq i(d)\subseteq i(c)$. Let $u_{Rw}, u_{Lw}$ be the vertices with maximal intervals in $SR_w(FG)$, $SL_w(FG)$, respectively; we denote the intervals delimiting $FG(u_{Rw})$, $FG(u_{Lw})$ by $IR(w)=[x,y], IL(w)=[z,q]$, respectively. If $SR_w(FG)=\{DR_w\}$ or $SL_w(FG)=\{DL_w\}$, then $x=r_w, z=l_w$, respectively. If $SR_w(FG)=\emptyset$ or $SL_w(FG)=\emptyset$, then $x=y=r_w, z=q=l_w$, respectively. In Fig. 1, $u_{Rw}=c, u_{Lw}=u, IR(w)=[lg,rc]$ and $IL(w)=[lu,re]$. Clearly $z\leq l_w \leq r_w \leq y$ and the intervals $IR(w)=[x,y], IL(w)=[z,q]$ are disjoint. For every vertex $w$ of $FG$, we denote the interval-filament formed by the arc $a(w)$ and the intervals $IR(w), IL(w)$ by $fil_w(FG)=a(w)\cup IR(w)\cup IL(w)$. For two brothers $u,v$, both in $SR_w(FG)$ or both in $SL_w(FG)$, $fil_u(FG)$ and $fil_v(FG)$ are non-intersecting and the intervals $i(fil_u(FG))$, $i(fil_v(FG))$ are contained one in the other. In Fig. 1a, $i(fil_u(FG))=[l_u,r_u] \subseteq [r_w,r_d]=[r_{IL(c)},l_{IR(c)}]$. For two brothers $u,v$, one in $SR_w(FG)$ the other in $SL_w(FG)$, $fil_u(FG)$ and $fil_v(FG)$ are non-intersecting and $i(fil_u(FG))$, $i(fil_v(FG))$ are disjoint. Each family $\{fil_v(FG) \mid v \in SR_w(FG)\}$, $\{fil_v(FG) \mid v \in SL_w(FG)\}$ and their union is a family of mutually non-intersecting interval-filaments.
Figure 1: Induced forest $F_G$ in oriented form (b) and its representation (a) by intersections of interval-filaments; $fil_w(F_G) = a(w) \cup AR(w) \cup AL(w)$. For simplicity, an interval-filament $a(u)$ is denoted by $u$.

We now characterize the vertex sets of $G$ which contain each subset of sons of $w$ in $F_G$. The non-dummy sons of $w$ are contained (since $z \leq l_w \leq q < x \leq r_w \leq y$) in

\[
VR_w[x,y] = \{ u | r_w \in i(u) \subseteq [x,y] \}, \quad VL_w[z,q] = \{ u | l_w \in i(u) \subseteq [z,q] \}.
\]

The unique dummy sons of $w$, and their successors, are contained in

\[
VDR_w(r_w,y) = \{ u | i(u) \subseteq (r_w,y) \}, \quad VDL_w(l_w,q) = \{ u | i(u) \subseteq (l_w,q) \}.
\]

We denote separately the subforests of $F_G(w)$ defined by $w$ and its sons in $SR_w(F_G)$ and in $SL_w(F_G)$, respectively, by:

\[
FR(F_G(w),[x,y]) = \bigcup \{ FG(u) | u \in SR_w(F_G) \} \cup \{w\},
\]
\[
FL(F_G(w),[z,q]) = \bigcup \{ FG(u) | u \in SL_w(F_G) \} \cup \{w\}.
\]

We also denote $\text{weight}(fil_w(F_G)) = |FR(F_G(w),[x,y]) \cup FL(F_G(w),[z,q])|$.

Let $V_w[x,y] = \{ u | i(u) \subseteq [x,y] \} \cup \{w\}, V_w[z,q] = \{ u | i(u) \subseteq [z,q] \} \cup \{w\}$.

**Lemma 1.** The subgraphs $FR(F_G(w),[x,y]), FL(F_G(w),[z,q])$ of $F_G$ are maximum induced forests in oriented form, rooted at $w$, in $G(V_w[x,y]), G(V_w[z,q])$, respectively.

**Proof:** If not, we can replace them by bigger forests, to obtain a bigger forest $F_G$.\[\]

For every vertex $v \in VR_w[x,y]$ and for every two disjoint intervals $[z',q'], [x',y']$, $z', q', x', y' \in EP$, such that $l_v \in [z',q'] \subseteq [x,r_w)$ and $r_v \in [x',y'] \subseteq (r_w,y]$, let $FR_v[x',y'], FL_v[z',q']$ be maximum induced forests in oriented form, rooted at $v$, in $G(V_v[x',y']), G(V_v[z',q'])$.

Observe that we use the notation $FR$ in two distinct forms: one as $FR(F_G(w),[x,y])$ for the subforest defined by $w$ and its $SR(F_G(w))$ sons in the specific maximum induced forest $F_G$.
of \( G \), and another \( FR_v(x',y') \) to denote any maximum induced forest of \( G \) rooted at \( v \) in \( G(V, [x', y']) \); same for \( FL \). We denote \( \text{fil}_v([z', q'], [x', y']) = a(v) \cup [z', q'] \cup [x', y'] \) and assign \( \text{weight}(\text{fil}_v([z', q'], [x', y'])) = |FR_v(x', y') \cup FL_v(z', q')| \).

For every \( y' \in (r_w, y] \), \( y' \in EP \), let \( FR_v(r_w, y'] \) be the maximum induced forest in oriented form, in \( G(VDR_w(r_w, y')] \). We assign \( \text{weight}(r_w, y'] = |FR(r_w, y']| \).

Let \( HR_w[x, y] \) be the weighted intersection graph of the family of weighted interval-filaments\( \{\text{fil}_v(FG) \mid v \in VR_w[x, y], z', q', x', y' \in EP\} \cup \{[r_w, y'] \mid r_w < y' \leq y, y' \in EP\} \) where \( l_v \in [z', q'] \subseteq [x, r_w) \) and \( r_v \in [x, y'] \subseteq (r_w, y] \).

**Lemma 2.** Every maximum weight independent set of \( HR_w[x, y] \), together with \( w \), defines an induced forest \( FR_w(x, y] \) which can replace \( FR(FG(w), [x, y]) \) in \( FG \) to obtain a maximum induced forest of \( G \). Thus, the vertex set corresponding to the family of filaments \( \{\text{fil}_v(FG) \mid v \in SR_w(FG)\} \) is a maximum weight independent set of \( HR_w[x, y] \).

**Proof:** Every maximum weight independent set of \( HR_w[x, y] \) defines a maximum induced forest \( FR_w[x, y] \) with interval-filaments delimited by \( [x, y] \) and disjoint from any other interval-filament of \( FG - FR(FG(w), [x, y]) \); adding \( w \) to \( FR_w[x, y] \) \(-\{w\} \), we obtain a maximum induced forest \( FR_w[x, y] \) which can replace \( FR(FG(w), [x, y]) \) in \( FG \) to obtain a maximum induced forest of \( G \). \( \square \)

**Lemma 3.** \( HR_w[x, y] \) is a cocomparability graph.

**Proof:** The intervals of two non-intersecting filaments corresponding to vertices of \( HR_w[x, y] \) are contained one in another, since both contain \( r_w \). Hence, for every three filaments \( \text{fil}_1, \text{fil}_2, \text{fil}_3 \), if \( i(\text{fil}_1) \subseteq i(\text{fil}_2) \subseteq i(\text{fil}_3) \), \( \text{fil}_1 \cap \text{fil}_2 = \phi \) and \( \text{fil}_2 \cap \text{fil}_3 = \phi \), then \( \text{fil}_1 \cap \text{fil}_3 = \phi \). Thus the complement of \( HR_w[x, y] \) is transitively orientable. \( \square \)

We can find a maximum weight independent set in \( HR_w[x, y] \), or equivalently a maximum clique in its transitively oriented complement, by the greedy algorithm in [3]. Similarly for \( VL_w[z, q] \), \( \{\text{fil}_v(FG) \mid v \in SL_w(FG)\}, FL(FG(w), [z, q]) \) and \( HL_w[z, q] \).

### 3. Algorithm for maximum induced forests in circle graphs

Our purpose is to describe a dynamic programming algorithm to find a maximum induced forest \( F \) of \( G(V,E) \) in oriented form. We partition the vertex set \( V \) into a family of subsets as follows: denote by \( A_0 \) the set containing the vertices with minimal intervals in \( I \), delete \( A_0 \) from \( V \), denote by \( A_1 \) the set containing the vertices with minimal intervals in the
remaining set of intervals, and so on \( A_0, A_1, ..., A_k \); for every \( i \), denote \( V_i = A_0 \cup A_1 \cup \ldots \cup A_i \).

For \( w \in V_i \), we denote by \( VR_{w,i}[x,y] \), \( VL_{w,i}[z,q] \), \( VDR_{w,i}(r_w,y) \), \( VDL_{w,i}(l_w,q) \), \( V_w,i[x,y] \), \( V_w,i[z,q] \) the sets of the previous Section, defined on \( G(V_i) \).

For \( w \in V_i \), we denote by \( VR_{w,i}[x,y] \), \( VL_{w,i}[z,q] \), \( VDR_{w,i}(r_w,y) \), \( VDL_{w,i}(l_w,q) \), \( V_w,i[x,y] \), \( V_w,i[z,q] \) the sets of the previous Section, defined on \( G(V_i) \).

The algorithm works by dynamic programming on the levels \( i \). For every \( i \), \( 0 \leq i \leq k \), for every \( w \in V_i \) and for every pair of disjoint intervals \( [z,q],[x,y] \), fulfilling \( z,q,x,y \in EP \), \( l_w \in [z,q] \), \( r_w \in [x,y] \), the algorithm constructs two maximum induced subforests \( FL_{w,i}[z,q] \), \( FR_{w,i}[x,y] \), in oriented form, rooted at \( w \), in \( G(V_i[z,q]) \), \( G(V_i[x,y]) \).

At level \( i \) we assume that for every \( v \) in \( V_{i-1} \) and every pair of disjoint intervals \( [z,q],[x,y] \), \( z,q,x,y \in EP \), \( l_v \in [z,q] \subseteq [x,r_w) \), \( r_v \in [x',y') \subseteq (r_w,y] \), we have maximum induced subforests \( FL_{v,i-1}[z,q] \), \( FR_{v,i-1}[x,y] \), in oriented form with \( v \) as root, including the cases \( z=l_v \), \( x=r_v \). Now, we evaluate them for every \( w \) in \( V_i \).

We go on \( L \) from right to left in order of right endpoints; assume that we are at some \( w \) in \( V_i \). For every \( y > r_w \), \( y \in EP \), we evaluate \( FR_{w,i}(r_w,y) \) as follows: For every vertex \( v \in VDR_{w,i}(r_w,y) \), we have, by recursion, the maximum induced forest \( FR_{v,i}[z,q'] \) in \( V_i \) with root \( v \) and \( IL(v) \subseteq [l_v,r_v) \), \( IR(v) \subseteq (l_v,y) \), \( IL(v) \cap IR(v) = \emptyset \), \( r_v \in IR(v) \); we take \( FR_{w,i}(r_w,y) \) to be the largest among these maximum induced forests and set \( weight([r_w,y]) \) to its size. Similarly, for every \( q \in EP \), \( l_w < q < r_w \), we evaluate \( FL_{w,i-1}(l_w,q) \) and \( weight([l_w,q]) \).

Now, for every interval \( [x,y] \), \( l_w < x \leq r_w \leq y \), we evaluate the maximum induced forest \( FR_{w,i}[x,y] \), rooted at \( w \), in \( G(V_w,i[x,y]) \). The vertices \( v \in VR_{w,i}[x,y] \) have right endpoints at the right of \( r_w \) and for every two disjoint intervals \( [z',q'],[x',y'] \), \( z',q',x',y' \in EP \), \( l_v \in [z',q'] \subseteq [x,r_w) \), \( r_v \in [x',y') \subseteq (r_w,y] \), we already have \( FR_{v,i}[x',y'],FL_{v,i-1}[z',q'] \); we assign
\[
weight(fil_v([z',q'],[x',y'])) = |FR_{v,i}[x',y'] \cup FL_{v,i-1}[z',q']|.
\]
By Lemma 3, the intersection graph \( HR_{w,i}([z',q'],[x',y']) \) of the family of interval-filaments \( \{ fil_v([z',q'],[x',y']) | v \in VR_{w,i}[x,y], z',q',x',y' \in EP \} \cup \{ [r_w,y'] | r_w < y' \leq y \} \) is a weighted cocomparability graph in which we can find a maximum weight independent set, which together with \( w \), gives us \( FR_{w,i}[x,y] \), by Lemma 2. When \( i=0 \), \( FR_{v,0}[x',y'] \) is a collection of induced paths starting with \( v \) having \( weight(fil_v([l_v,l_v],[x',y'])) = |FR_{v,0}[x',y']| \); \( FR_{w,0}[x,y] \) is obtained by taking \( FR_{v,0}[x',y'] \) for \( fil_v([l_v,l_v],[x',y']) \) of maximum weight and attaching \( v \) to \( w \).

When \( w \in V_i \), we also evaluate (as above) \( FL_{w,i-1}[z,q] \) using the already evaluated \( FR_{w,i}[x',y'], FL_{v,i-1}[z',q'] \) for every \( v \in VL_{w,i-1}[z,q] \) and every two disjoint intervals \( [z',q'],[x',y'], z',q',x',y' \in EP \), \( l_v \in [z',q'] \subseteq [x,l_v) \), \( r_v \in [x',y') \subseteq (l_w,y] \). Similarly we evaluate \( FR_{w,i-1}[z,q] \).
Now, we go on from left to right in order of left endpoints; assume that we are at some $w$ in $V_i$. For every interval $[z,q]$, $z \leq l_w < r_w$, we evaluate the maximum induced forest $FL_{w,[z,q]}$, rooted at $w$, in $G(V_{w,[z,q]})$. The vertices $v \in VL_{w,[z,q]}$ have left endpoints at the left of $l_w$ and for every two disjoint intervals $[z',q'],[x',y']$, $z',q',x',y' \in EP$, $l_v \in [z',q'] \subseteq [z,l_w)$, $r_v \in [x',y'] \subseteq (l_w,q]$, we already evaluated $FR_{v,i-1}[x',y']$ and $FL_{v,i}[z',q']$; we assign
\[ \text{weight}(fil_v([z',q'],[x',y'])) = |FR_{v,i-1}[x',y'] \cup FL_{v,i}[z',q']|. \]

By Lemma 3, the intersection graph $HL_{w,i}(z',q'] \cup [x',y'])$ of the family of interval-filaments is a weighted cocomparability graph in which we can find a maximum weight independent set, which together with $w$, gives us $FL_{w,[z,q]}$, by Lemma 2. When $i=0$, $FL_{0,[z',q']}$ is an induced path ending in $v$ with $\text{weight}(fil_v([z',q'],[r_v,r_v])) = |FL_{v,0}[z',q']|$. $FL_{w,0}[z,q]$ is obtained by taking $FL_{v,0}[z',q']$ for maximum weight $fil_v([z',q'],[r_v,r_v])$ and attaching $v$ to $w$.

Finally, we take $F_{w,i}(z,q],x,y] = FL_{w,i}[z,q] \cup FR_{w,i-1}[x,y]$ as maximum induced forest at stage $i$ for $w$ and the intervals $x,y]$. When $i=0$, $F_{w,0}(z,q],x,y]$ is a collection of induced paths containing $w$. We obtain:

**Theorem 4.** For every $i$, $0 \leq i \leq k$, every vertex $w$ and every pair of disjoint intervals $[z,q], [x,y]$, $z,q,x,y \in EP$, $l_w \in [z,q]$, $r_w \in [x,y]$, $F_{w,i}(z,q],x,y]$ is a maximum induced forest, in oriented form in $V_i$, with $w$ as root.

**Proof:** By induction on $i$ and Lemma 2.□

The maximum induced forest of $G$ is the maximum of the induced forests $F_{w,k}(z,q],x,y])$. Keeping track of the sinks in the transitive complements of the graphs $FR_{w,i-1}$, $FL_{w,i-1}$, in each interval $[x,y]$, $[z,q]$, when going from $i-1$ to $i$, the algorithm works in $O(|V|^4)$ time. If in the algorithm we do not include $FR_{w,i}(r_w,q]$ and $FL_{w,i}(l_w,y]$, i.e., we do not include dummy sons, then we obtain a maximum induced tree of $G$.

**REFERENCES**


