An Improved Parameterized Algorithm for the Independent Feedback Vertex Set Problem

Yinglei Song∗

Abstract

In this paper, we develop a new parameterized algorithm for the Independent Feedback Vertex Set (IFVS) problem. Given a graph $G = (V, E)$, the goal of the problem is to determine whether there exists a vertex subset $F \subseteq V$ such that $V - F$ induces a forest in $G$ and $F$ is an independent set. We show that there exists a parameterized algorithm that can determine whether a graph contains an IFVS of size $k$ or not in time $O(4^k n^2)$. To our best knowledge, this result improves the known upper bound for this problem, which is $O^*(5^k n^{O(1)})$.

Keywords: independent feedback vertex set, parameterized algorithm, dynamic programming

1 Introduction

A feedback vertex set in a graph $G = (V, E)$ is a vertex subset $F \subseteq V$ such that vertices in $V - F$ induce a forest in $G$. Feedback vertex sets have important applications in deadlock recovery in the development of operating systems [20] and database management systems [15]. Moreover, $F$ is an independent feedback vertex set if it induces an independent set in $G$.

In the past two decades, the Feedback Vertex Set (FVS) problem has been intensively studied. The goal of the problem is to determine whether a graph contains a FVS of size $k$ or not. In [13], it is shown that the problem is NP-complete. So far, a few exact algorithms have been developed to compute a minimum FVS in a graph. For example, in [18], an algorithm that needs $O(1.9053^n)$ time was developed to find a minimum FVS in a graph, this algorithm is also the first one that breaks the trivial $O(2^n n)$ bound. In [8], an elegant algorithm that can enumerate all induced forests in a graph is used to find the minimum FVS. This algorithm needs $O(1.7548^n)$ computation time. Recently, the upper bound of this problem is improved to $O(1.7347^n)$ [9].

Parameterized computation provides another potentially practical solution for problems that are computationally intractable. Specifically, one or a few parameters in some intractable problems can be identified and parameterized computation studies whether efficient algorithms exist for these problems while some or all of the parameters are small. A parameterized problem may contain a few parameters $k_1, k_2, \ldots, k_l$ and the problem is fixed parameter tractable if it can be solved in time $O(f(k_1, k_2, \ldots, k_l)n^c)$, where $f$ is a function of $k_1, k_2, \ldots, k_l$, $n$ is the size of the problem and $c$ is a constant independent of all parameters. For example, the VERTEX COVER problem is to determine whether a graph $G = (V, E)$ contains a vertex cover of size at most $k$ or not. The problem is NP-complete. However, a simple parameterized algorithm can solve the problem in time $O(2^k |V|)$ [8]. In practice, this algorithm can be used to efficiently solve the VERTEX COVER problem when the parameter $k$ is fixed and small. On the other hand, some problems do not have known efficient parameterized solutions and are therefore parameterized intractable. Similar to the conventional complexity theory, a

∗School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu 212003, China. Email: yingleisong@gmail.com.
The graphs in this paper are undirected graphs without loops. For a given graph $G = (V, E)$ and a vertex subset $V' \subseteq V$, $G[V']$ is the subgraph induced on the vertices in $V'$. A vertex $u$ is a neighbor of $V'$ if it is joined to one of the vertices in $V'$ by an edge in $G'$. We use $N_G[V']$ to denote the set of neighbors of $V'$ in $G$. We may omit the subscript $G$ when the underlying graph is clear in the context. Given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, we use $G_1 \cup G_2$ to denote the graph $(V_1 \cup V_2, E_1 \cup E_2)$. Furthermore, to simplify the notation, we use $G - V'$ to denote subgraph $G[V - V']$. A vertex subset $V'$ is a feedback vertex cover in $G$ if $G - V'$ is a forest. The objective of our algorithm is to find a feedback vertex cover that contains minimum number of vertices.

2 Notations and Preliminaries

The graphs in this paper are undirected graphs without loops. For a given graph $G = (V, E)$ and a vertex subset $V' \subseteq V$, $G[V']$ is the subgraph induced on the vertices in $V'$. A vertex $u$ is a neighbor of $V'$ if it is joined to one of the vertices in $V'$ by an edge in $G'$. We use $N_G[V']$ to denote the set of neighbors of $V'$ in $G$. We may omit the subscript $G$ when the underlying graph is clear in the context. Given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, we use $G_1 \cup G_2$ to denote the graph $(V_1 \cup V_2, E_1 \cup E_2)$. Furthermore, to simplify the notation, we use $G - V'$ to denote subgraph $G[V - V']$. A vertex subset $V'$ is a feedback vertex cover in $G$ if $G - V'$ is a forest. The objective of our algorithm is to find a feedback vertex cover that contains minimum number of vertices.
Figure 1: The conversion the algorithm uses to convert a tree in $G[V - F]$ into a binary tree. (a) An internal node $u$ that has $d$ children $c_1, c_2, \ldots, c_d$; (b) A path of $d - 1$ white vertices $w_1, w_2, \ldots, w_{d-1}$ are added to convert the tree into a binary tree.

3 The Algorithm

Our algorithm is based on the following Lemma.

Lemma 3.1 Given a graph $G = (V, E)$ and an FVS $F$ in $G$, a minimum IFVS in $G$ can be computed in time $O(4^{|F||V|})$.

Proof. Since $F$ is an FVS, $G[V - F]$ is a forest. As the first step of the algorithm, we arbitrarily choose a vertex in each tree in $G[V - F]$ as the root of the tree. We color each vertex in $V - F$ with black. We then check the number of children of each vertex in the tree and find those that have more than two children. For each such vertex $u$, we remove the edges that join it with its children $c_1, c_2, \ldots, c_d$, where $d$ is the number of children in the tree. We then create a path of white vertices $w_1, w_2, \ldots, w_{d-1}$, $v_{d-1}$ is joined to $u$ with an edge. To construct a binary tree, $w_1$ is joined to $c_1$ and $c_2$ with two edges and each white vertex $w_t$ $(2 \leq t \leq d - 1)$ in the path is joined to $c_{t+1}$ with an edge. Each white vertex $w_t$ is said to be equal to $u$. We use $H$ to denote the resulting forest. The number of white vertices in $H$ is at most

$$|V| - |F| \leq 2(|V| - |F|)$$

where $d_i$’s are the degrees of all vertices in $G[V - F]$. The number of vertices in $H$ is at most $3(|V| - |F|)$.

Without loss of generality, we assume $G[F - S]$ contains $l$ connected components $C_1, C_2, \ldots, C_l$ and let $C = \{C_1, C_2, \ldots, C_l\}$. Since $H$ is a forest, for each tree $T_i \in H$, we use a dynamic programming algorithm to find a minimum subset $S'_i$ in $T_i$ such that $(T_i - S'_i) \cup (F - F')$ is acyclic. We then obtain $S'$ by taking the union of $S'_i$’s for all $T_i$’s in $G - F$.

We use $r_i$ to denote the root of $T_i$. The algorithm starts with the leaves of the $T_i$ and follows a
bottom-up order to scan through all vertices in \( T_i \). For a tree node \( u \) in \( T_i \) and a set \( S_u \subseteq C \), we consider the following two cases.

1. \( u \) is not in the IFVS, there exists a path that connects \( u \) to each component in \( S_u \) in the subtree rooted at \( u \) in \( T_i \). Moreover, none of the vertices in the path are in the IFVS,

2. \( u \) is in the IFVS.

The algorithm maintains dynamic programming table \( M[u, S_u] \) for case 1 and table \( P[u] \) for case 2. The two tables store the minimum number of vertices that can be included in the IFVS in the subtree rooted at \( u \).

For a leaf node \( v \) of \( T_i \) and a subset \( S_u \) of connected components in \( G[V - F] \), the values of \( M[v, S_u] \) are initialized to be 0 if \( S_u \) is the set of components that are directly connected to \( v \) with an edge. Otherwise, \( M[v, S_u] \) is set to be \( |V| + 1 \). If \( v \in N_G(S) \), \( v \) is not in the IFVS, we thus set \( P[u] \) to be \(|V| + 1 \), otherwise \( P[u] \) is set to be 1.

For an internal node \( u \) in \( T_i \), since \( T_i \) is a binary tree, we use \( c_1, c_2 \) to denote its children in \( T_i \). In the case where \( u \in N_G(S) \), \( u \) cannot be in the IFVS and we thus set \( P[u] \) to be \(|V| + 1 \). Otherwise, both \( M[u, P[u]] \) and \( P[u] \) need to be determined with the recursion relations of the dynamic programming.

Next, we determine the recursion relations of the dynamic programming. For a black vertex \( u \), we use \( W_u \) to denote the set of connected components in \( G[F - S] \) that are directly connected to \( u \) with an edge. If \( W_u \) contains a component that is not in \( S_u \), \( M[u, S_u] \) is set to be \(|V| + 1 \), since \( u \) is connected to any component in \( W_u \) regardless of its child nodes in \( T_i \). For a white vertex \( u \), \( W_u = W_b \), where \( b \) is the black vertex that is equal to \( u \) in \( T_i \). The algorithm thus only needs to compute the values of \( M[u, S_u] \) for these \( S_u \)'s that are supersets of \( W_u \).

We now assume \( W_u \subseteq S_u \) and \( u \) has two children. Based on the color of \( u \), we use the following recursion relations to compute \( M[u, W_u] \) and \( M[u, S_u] \); if \( u \) is a black vertex, the recursion relation to compute \( M[u, S_u] \) are

\[
M[u, W_u] = \min \{M[c_1, \epsilon] + M[c_2, \epsilon], M[c_1, \epsilon] + P[c_2], M[c_1,\epsilon] + M[c_2, \epsilon], P[c_1] + P[c_2]\} \tag{2}
\]

\[
M[u, S_u] = \min \{ \min_{S_{c_1}, S_{c_2}} \{ M[c_1, S_{c_1}] + M[c_2, S_{c_2}] \}, M[c_1, S_u - W_u] + P[c_2], P[c_1] + M[c_2, S_u - W_u]\} \tag{3}
\]

where \( S_{c_1} \) and \( S_{c_2} \) are disjoint subsets of \( S_u \) and \( S_u - W_u = S_{c_1} \cup S_{c_2} \). The algorithm exhaustively enumerates all such subsets and computes the minimum value of \( M[c_1, S_{c_1}] + M[c_2, S_{c_2}] \) over all possible subsets \( S_{c_1} \) and \( S_{c_2} \).

If \( u \) is a white vertex and both \( c_1 \) and \( c_2 \) are black vertices, the recursion relation to compute \( M[u, S_u] \) is the same as above. If \( u \) is white and one of \( c_1, c_2 \) is white, without loss of generality, we assume \( c_1 \) is white. Since \( u \) and \( c_1 \) represents the same vertex in \( G \), if \( u \) is not in the IFVS, \( c_1 \) cannot be in the IFVS. The recursion relation is thus

\[
M[u, W_u] = \min \{M[c_1, W_u] + M[c_2, \epsilon], M[c_1, W_u] + P[c_2]\} \tag{4}
\]

\[
M[u, S_u] = \min \{ \min_{S_{c_1}, S_{c_2}} \{ M[c_1, S_{c_1} \cup W_u] + M[c_2, S_{c_2}]\}, M[c_1, S_u] + P[c_2]\} \tag{5}
\]

where \( S_{c_1} \) and \( S_{c_2} \) are disjoint subsets of \( S_u \) and \( S_{c_1} \cup S_{c_2} = S_u - W_u \).

If \( u \) is black and \( u \notin N_G(S) \), since none of \( c_1 \) and \( c_2 \) are included in the IFVS, the recursion relation to compute \( P[u] \) is as follows.

\[
P[u] = \min_{S_u} \{ M[c_1, S_u] \} + \min_{S_u} \{ M[c_2, S_u] \} + 1 \tag{6}
\]

where the minimum is taken over all possible subsets \( S_u \) of \( C \).
If $u$ is white, $u \not\in N_G(S)$, and both $c_1$ and $c_2$ are black, the recursion for computing $P[u]$ is the same as above. If $u$ is white and one of $c_1, c_2$ is white, without loss of generality, we assume $c_1$ is white. Since $u$ and $c_1$ represents the same vertex and $c_2$ is not included in the IFVS, the recursion relation is

$$P[u] = P[c_1] + \min_{S_u} \{ M[c_2, S_u] \}$$

where the minimum is taken over all possible subsets $S_u$ of $C$.

If $u$ has only one child, $u$ must be black and we use $c$ to denote its child. If $c$ is black, we use the following recursion relations to compute $M[u, W_u]$, $M[u, S_u]$, and $P[u]$.

$$M[u, W_u] = \min \{ M[c, e], P[c] \}$$
$$M[u, S_u] = M[c, S_u - W_u]$$
$$P[u] = \min_{S_u} \{ M[c, S_u] \} + 1$$

where the minimum is taken over all possible subsets $S_u$ of $C$.

If $c$ is white, $c$ is equal to $u$, we use the following recursion relations to compute $M[u, S_u]$ and $P[u]$.

$$M[u, S_u] = M[c, S_u]$$
$$P[u] = P[c]$$

After the dynamic programming tables are filled by the algorithm, for each $T_i(1 \leq k \leq l)$, we can find the minimum of $P[r_1]$, $M[r_1, S_{r_1}]$ for all subsets $S_{r_1}$ of $C$. We then follow a top-down trace back procedure to find $S'_i$. Note that the above recursions split $S_u$ into two disjoint subsets $S_{c_1}$ and $S_{c_2}$, which can guarantee that no cycles are formed if $u$ is not included in the IFVS. The correctness of this algorithm can thus be easily proved with the principle of induction and its recursion relations.

Before the dynamic programming starts, the algorithm needs to construct $H$ from the trees in $G[V - F]$, it needs at most $O(|V|)$ time. In addition, the algorithm also needs to compute $W_u$ for each internal node in the trees in $H$ and $N_G(S)$. However, since $S$ contains at most $|F|$ vertices, this needs at most $O(|F||V|)$ time.

During the dynamic programming process, for a given node $u$ and $S_u$, the algorithm needs to consider all possible disjoint pairs of subsets $S_{c_1}$ and $S_{c_2}$, such that $S_{c_1} \cup S_{c_2} = S_u - W_u$. We denote $t = |S_u|$, the number of such pairs is at most $2^t$. The total amount of computation time needed by the dynamic programming on one single vertex in $T_i$ is thus at most

$$O(\sum_{t=0}^{l} 2^t \binom{l}{t}) = O(3^t)$$

where the right hand side of the equation can be directly obtained with the binomial theorem.

Since $H$ contains at most $3(|V| - |F|)$ vertices, the dynamic programming algorithm can find $S'$ in time $O(3^t|V|)$. For each $S \subseteq F$ such that $G[F - S]$ is acyclic, we can find a feedback vertex cover $S \cup S'$. The one that contains the minimum number of vertices is a minimum IFVS. The total amount of the computation time needed by the algorithm is thus at most

$$O(\sum_{i=1}^{|F|} 3^i \binom{|F|}{i} |V|) = O(4^{|F|}|V|)$$

The right hand side of the equation again is the direct result of applying the binomial theorem to the left hand side. The theorem has been proved. \qed
Based on the algorithm we have developed in Lemma 3.1 we can use the iterative compression technique developed in [19] to develop a parameterized algorithm that can determine whether a graph contains an IFVS of size \( k \) in time \( O(4^k n^2) \). We thus have the following theorem.

**Theorem 3.1** Given a graph \( G = (V, E) \) and a positive integer parameter \( k \), there exists a parameterized algorithm that can determine whether \( G \) contains an IFVS of size \( k \) or not in time \( O(4^k |V|^2) \).

**Proof.** We use the iterative compression technique to develop the algorithm. Let \( n = |V| \), the algorithm can be sketched as follows.

1. Starting with an empty graph \( G_0 \), add vertices one by one to reconstruct graph \( G \), when a vertex \( u_i \) is added to \( G_i \) (\( 0 \leq i < n \)), edges are added to \( G_i \) to join \( u_i \) with its neighbors in \( G \) that have been included in \( G_i \) to construct \( G_{i+1} \). We thus can obtain a series of graphs \( G_0, G_1, \ldots, G_n \), where \( G_0 \) is an empty graph and \( G_n = G \);

2. Arbitrarily choose one vertex from \( G_2 \) and include it in set \( F_2 \);

3. For each \( i = 2 \) to \( n - 1 \), given \( F_i \cup \{ u_i \} \), use the algorithm presented in the proof of Lemma 3.1 to compute a minimum IFVS \( F_{i+1} \) in \( G_{i+1} \), if \( F_{i+1} \) contains more than \( k \) vertices, return “no”;

4. return “yes”.

For computation time, step 1 can be accomplished in time \( O(n^2) \). From Lemma 3.1 the computation time needed by step 3 is at most

\[
\sum_{i=2}^{n} O(4^k i) = O(4^k n^2)
\]

The theorem has been proved.

In [14], it is shown that the FVS problem can be reduced to the IFVS problem through a simple polynomial time reduction that also preserves the parameter \( k \). Theorem 3.1 thus implies there exists a dynamic programming based algorithm that solves the FVS problem in time \( O(4^k n^2) \). In fact, it is not difficult to see that the algorithm presented in the proof of Lemma 3.1 can be slightly changed to solve the FVS problem.

### 4 Conclusions

In this paper, we develop a new parameterized algorithm that can determine whether a graph contains an IFVS of size \( k \) or not in time \( O(4^k n^2) \). To our best knowledge, this result improves the previously known upper bound for this problem, which is \( O^*(5^k n^{O(1)}) \). Our technique combines a dynamic programming approach with the recently developed iterative compression technique, which has been proved to be useful for developing parameterized algorithms for many NP-complete problems.

Our algorithm does not employ the method of branching, which is an often used approach for developing exact and parameterized algorithms. Indeed, the IFVS is a subset with certain global structural properties. In general, a branching algorithm employs the local structure of a graph and can hardly be used to further improve the upper bound of the problem. However, we are still looking for the possibility to combine our method and the branching approach to further improve the upper bound of this problem. In fact, recent work [9] has shown that this combination can lead to improved upper bound results for many NP-hard graph optimization problems.
References

[1] Bodlaender H., “On Disjoint Cycles”, *International Journal on Foundations of Computer Science*, 5(1):59-68, 1994.

[2] Y. Cao, J. Chen, and Y. Liu, “On Feedback Vertex Set New Measure and New Structures”, *SWAT 2010*, 93-104, 2010.

[3] J. Chen, F. V. Fomin, Y. Liu, S. Lu, and Y. Villanger, “Improved algorithms for feedback vertex set problems”, *Journal of Computer System Science*, 74(7): 1188-1198, 2008.

[4] F. K. H. Dehne, M. R. Fellows, M. A. Langston, F. A. Rosamond, and K. Stevens, “An O(2^{O(k)}n^3) Algorithm for The Undirected Feedback Vertex Cover Problem”, *Proceedings of the 11th Annual International Conference on Computing and Combinatorics*, pp. 859-869, 2005 (COCOON 2005).

[5] M. Dom, J. Guo, F. Huffner, R. Niedermeier, and A. Trub, “Fixed-parameter tractability results for feedback set problems in tournaments”, *Algorithms and complexity*, volume 3998 of Lecture Notes in Computer Science, pages 320-331. Springer, Berlin, 2006.

[6] Downey, R., Fellows, M., “Fixed parameter tractability and completeness, *Complexity Theory: Current Research*, pp. 191225. Cambridge University Press, 1992.

[7] Downey, R., Fellows, M., *Parameterized Complexity*, Springer-Verlag, New York, 1999.

[8] F. V. Fomin, S. Gaspers, and A. V. Pyatkin, “Finding A Minimum Feedback Vertex Cover in Time O(1.754^n)”, *Proceedings of 2006 International Workshop on Parameterized and Exact Computation*, 2006 (IWPEC 2006).

[9] F. V. Fomin, S. Gaspers, S. Saurabh, and A. A. Stepanov, “On Two Techniques of Combining Branching and Treewidth”, *Proceedings of the 17th International Symposium on Algorithm and Computation*, 2006 (ISAAC 2006).

[10] J. Guo, R. Niedermeier, and S. Wernicke, “Parameterized Complexity of Generalized Vertex Cover Problems”, *Proceedings of the 9th Workshop on Algorithms and Data Structures*, pp. 36-48, 2005 (WADS 2005).

[11] J. Guo, J. Gramm, F. H ?uffner, R. Niedermeier, and S. Wernicke, “Compression-based fixed-parameter algorithms for feedback vertex set and edge bipartization”, *Journal of Computer and System Science*, 72(8):13861396, 2006.

[12] I. Kanj, M. Pelsmajer, and M. Schaefer, “Parameterized Algorithms for Feedback Vertex Set”, *Proceedings of 2004 International Workshop on Parameterized and Exact Computation*, pp. 234-247, 2004 (IWPEC 2004).

[13] R. M. Karp, “Reducibility Among Combinatorial Problems”, *Complexity of Computer Computations*, Plenum Press, New York, pp. 85-103, 1972.

[14] N. Misra, G. Philip, V. Raman, and, S. Saurabh, “On Parameterized Independent Feedback Vertex Set”, *Proceedings of COCOON 2011*, pp.98-109, 2011.

[15] H. G. Molina, J. D. Ullman, and J. Widom, *Database Systems: The Complete Book (2nd Edition)*, Pearson Prentice Hall, 2009.

[16] V. Raman, S. Saurabh, and C. R. Subramanian, “Faster Fixed Parameter Tractable Algorithms for Undirected Feedback Vertex Set”, *ACM Transactions on Algorithms*, 2(3): 403-415, 2006.
[17] Raman, V., Saurabh, S., Subramanian, C., “Faster fixed parameter tractable algorithms for undirected feedback vertex set”, ISAAC 2002, LNCS vol. 2518, pp. 241-248. Springer, 2002.

[18] I. Razgon, “Exact Computation of Maximum Induced Forest”, Proceedings of the 10th Scandinavian Workshop on Algorithm Theory, pp. 241-248, 2002.

[19] B. Reed, K. Smith, and A. Vetta, “Finding odd cycle transversals”, Operations Research Letters, 32(4):299-301, 2004.

[20] A. Silberschatz, P. B. Galvin, and G. Gagne, Operating System Concepts, 9th Edition, John Wiley and Sons, Inc., 2012.