Available online at www.sciencedirect.com
ScienceDirect
Theoretical Computer Science 363 (2006) 99–105
Theoretical Computer Science
www.elsevier.com/locate/tcs

Approximating the minimum weight weak vertex cover☆

Yong Zhang[a], Qi Ge[a], Rudolf Fleischer[a], Tao Jiang[b], Hong Zhu[a,*]

[a]Department of Computer Science and Engineering, Shanghai Key Laboratory of Intelligent Information Processing, Fudan University, China
[b]Shanghai Globalization Laboratory, IBM, China

Abstract

Accurate network flow measurement is important for a variety of network applications, where the "flow" over an edge in the network is intuitively the rate of data traffic. The problem of efficiently monitoring the network flow can be regarded as finding the minimum weight weak vertex cover for a given graph. In this paper, we present a $(2 - \frac{2}{\nu(G)})$-approximation algorithm solving for this problem, which improves previous results, where $\nu(G)$ is the cyclomatic number of G.

Keywords: Network flow measurement; Cycle-rank-proportional graph; Local-ratio theorem

1. Introduction

As the internet is growing rapidly, more and more network applications need to know the network running parameters to monitor network utilization and performance. Network flow is the most important parameter to be measured. The flow is the rate of data traffic on edges of the network. Unfortunately, accurate network flow measurement is difficult. Most monitoring technologies typically assume that the measurement instrumentation can be either intelligently distributed at different points in the underlying network [7] or placed at the endpoints of the end-to-end path whose characteristics are of interests [8]. Currently, more and more information must be collected at much higher frequencies by network monitoring processes. Then the overhead the monitoring method imposes on the underlying network can be significant and adversely impact the network's throughput as a result.

As usual, a flow f in a directed graph $G = (V, E)$ is a function on the edges with the properties $f(u, v) = -f(v, u)$ for any edge $(u, v) \in E$ (*backflow property*) and $\sum_{u \in V} f(u, v) = 0$ for any vertex $v \in V$ except sinks and sources (*flow conservation property*). The flow conservation property implies that we do not need to directly monitor all edges of the network (by a vertex cover of G) to monitor all flows in the network. If a node has degree k and the flow on $k - 1$ of its incident edges is already known, then the flow on the remaining edge can easily be calculated. We call a subset S of V a *flow monitoring set* if knowledge of the flow on all edges incident to vertices in S is sufficient to compute the flow on every edge in E.

☆ This work was supported by grants from National Natural Science Fund (Grant nos. 60273045, 60496321, 60373021, 60573025), the Ministry of Science and Technology (Grant no. 2001CCA03000), and Shanghai Science and Technology Development Fund (Grant no. 03JC14014).
* Corresponding author.
E-mail addresses: 012021133@fudan.edu.cn (Y. Zhang), qge@fudan.edu.cn (Q. Ge), rudolf@fudan.edu.cn (R. Fleischer), jiangtao@cn.ibm.com (T. Jiang), hzhu@fudan.edu.cn (H. Zhu).

0304-3975/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2006.06.009

We are interested in finding a minimum size flow monitoring set of a directed graph G. If the vertices of the graph also have some non-negative weights, we want to find a minimum weight flow monitoring set. This problem is actually equivalent to finding a minimum size (or weight) weak vertex cover of the underlying undirected graph. A similar definition exists for covering the vertices of a graph, the so-called *power dominating set* [6].

Definition 1.1 (*Breibart et al. [2]*). Let $G = (V, E)$ be an undirected graph. A subset S of V is a *weak vertex cover* of G if eventually every edge in E can be marked by performing the following algorithm.
(1) Mark all edges incident to vertices in S.
(2) As long as there is a vertex with only one unmarked incident edge, mark this edge.

Note that we can w.l.o.g. assume that G has no vertices of degree one. For the flow monitoring problem, the only edge incident to a vertex of degree one does not need to be monitored because its flow is already determined; either it is zero or the vertex is a source or sink with known flow.

If each vertex $v \in V$ has a weight w_v, the *minimum weight weak vertex cover problem* is to identify a weak vertex cover of minimum total weight. If a weak vertex cover S contains no redundant vertices, i.e., $S - v$ is not a weak vertex cover for any $v \in S$, we say S is a *minimal weak vertex cover*. Solving the minimum flow monitoring set is now equivalent to finding a minimum size (or weight) weak vertex cover.

Liu et al. showed that S is a weak vertex cover if and only if the subgraph induced by $V - S$ is a forest [9], i.e., we cannot infer the flow values on a cycle without monitoring at least one edge of the cycle. They also gave a greedy algorithm for the minimum size weak vertex cover problem with approximation ratio $2(\ln d + 1)$, where d is the maximum degree of nodes in G. Zhang and Zhu showed that the minimum size weak vertex cover problem is $\mathcal{NP}$-*hard* and gave a $(\ln d + 1)$-approximation algorithm [11]. Their algorithm is based on the concept of the cycle space of a graph (see [5], for example) which is also important in this paper.

This paper is structured as follows. In Section 2 we introduce cycle-rank-proportional graphs which are the basis of our approximation algorithm. In Section 3 we first show a local-ratio theorem and then give a $(2 - \frac{2}{\nu(G)})$-approximation algorithm for the minimum weight weak vertex cover, where $\nu(G) = |E| - |V| + 1$ is the cyclomatic number of a connected graph G. In the last section we give a lower bound for the inapproximability of Minimum Weak Vertex Cover problem.

2. Cycle-rank-proportional graphs

Let $G = \{V, E\}$ be a graph with n vertices V and m edges E. For vertices $v_0, \ldots, v_{k-1} \in V$, $C = v_0 v_1 \ldots v_{k-1} v_0$ is a *cycle* of *length* k if $(v_i, v_{i+1 \bmod k})$ are edges in E, for $i = 0, \ldots, k-1$. If the vertices of C are pairwise different, the cycle is *simple*. We assume from now on that all cycles are simple.

The *cycle space* $\mathcal{C} = \mathcal{C}(G)$ is the space spanned by all the edges of all the cycles in G. The dimension of $\mathcal{C}(G)$ is the *cyclomatic number* $\nu(G)$ of G. This number is easy to compute, it is $\nu(G) = |E| - |V| + p_G$, where p_G is the number of connected components of G [5]. For any vertex $v \in V$, let R_v be the decrease of the cyclomatic number of G after deleting v and all its incident edges and k_v the increase in the number of connected components. Then,

$$\begin{aligned} R_v &= \nu(G) - \nu(G - \{v\}) \\ &= (|E| - |V| + p_G) - [(|E| - deg(v)) - (|V| - 1) + p_{G-\{v\}}] \\ &= deg(v) - k_v - 1\ , \end{aligned}$$

where $deg(v)$ denotes the degree of v.

Definition 2.1. A vertex weight function w of a graph $G = (V, E)$ is *c-cycle-rank-proportional*, where $c > 0$ is a constant, if $w(v) = c \cdot R_v$ for all $v \in V$.

By the computation above, we could also say $w(v) = c \cdot (deg(v) - k_v - 1)$, for all $v \in V$.

Definition 2.2. A cycle C is *nearly isolated* if it contains at most one vertex incident to at least one vertex not in C.

We now show a lower and an upper bound for the weight of minimal vertex covers in cycle-rank-proportional graphs without nearly isolated cycles. This shows that minimal weak vertex covers (which we can easily compute) can serve as good approximations to the minimum weight weak vertex cover.

Lemma 2.1. *Let $G = (V, E)$ be a c-cycle-rank-proportional connected graph, for some $c > 0$. Then any weak vertex cover S of G has weight at least $c \cdot v(G)$.*

Proof. Since G is connected, $v(G) = |E| - |V| + 1$. Let S be a weak vertex cover of weight $w(S)$. Note that deleting a vertex and its incident edges cannot increase R_v of any remaining vertex v. After deleting all the vertices in S and their incident edges no cycle remains [9], i.e., the cyclomatic number of the remaining graph is zero. Thus,

$$w(S) = \sum_{v \in S} w(v) \geqslant c \cdot (|E| - |V| + 1) = c \cdot v(G). \qquad \square$$

Lemma 2.2. *Let $G = (V, E)$ be a c-cycle-rank-proportional connected graph without nearly isolated cycles, for some $c > 0$. Then any minimal weak vertex cover S of G has weight at most $2c \cdot (v(G) - 1)$.*

Proof. Let S be a minimal weak vertex cover of G of weight $w(S)$. The subgraph induced by $V - S$ is a forest F. For a tree $T \in F$ let $\delta(T)$ denote the number of edges in E connecting S to T. Since we assumed that G does not have vertices of degree one, any leaf of T must be connected to some node in S. Thus, $\delta(T) \geqslant 2$.

Let S_T be the set of vertices in S connected to T by at least two edges. Since S is minimal, each $v \in S$ belongs to some S_T (otherwise $S - \{v\}$ would also be a weak vertex cover). If T has t vertices, then

$$\sum_{v \in T} (deg(v) - 2) = \sum_{v \in T} deg(u) - 2t = \delta(T) + 2(t-1) - 2t = \delta(T) - 2.$$

Note that $S_T = \emptyset$ if $\delta(T) = 2$, otherwise G would contain a nearly isolated cycle. Since always $|S_T| \leqslant \lfloor \frac{1}{2}\delta(T) \rfloor$, we also have $\delta(T) - 2 \geqslant |S_T|$. Thus,

$$\sum_{T \in F} (\delta(T) - 2) \geqslant \sum_{T \in F} |S_T| \geqslant |S|$$

and

$$\begin{aligned}
w(S) &= c \cdot \sum_{v \in S} (deg(v) - k_v - 1) \\
&\leqslant c \cdot \sum_{v \in S} (deg(v) - 1) \\
&= c \cdot \sum_{v \in S} (deg(v) - 2) + |S| \\
&\leqslant c \cdot \sum_{v \in S} (deg(v - 2) + c \cdot \sum_{T \in F} (\delta(T) - 2) \\
&= c \cdot \sum_{v \in S} (deg(v) - 2) + c \cdot \sum_{T \in F} \sum_{v \in T} (deg(v) - 2) \\
&= c \cdot \sum_{v \in S} (deg(v) - 2) + c \cdot \sum_{v \in F} (deg(v) - 2) \\
&= c \cdot \sum_{v \in V} (deg(v) - 2) \\
&= 2c(|E| - |V|) \\
&= 2c \cdot (v(G) - 1). \qquad \square
\end{aligned}$$

The previous two lemmas immediately imply the following theorem.

Theorem 2.1. *Let $G = (V, E)$ be a c-cycle-rank-proportional connected graph without nearly isolated cycles, for some $c > 0$. Then any minimal weak vertex cover S of G is a $(2 - \frac{2}{v(G)})$-approximation to the minimum weight weak vertex cover of G.*

In the next section we will show that we can achieve the same approximation ratio for arbitrary graphs without any restrictions on the weight function.

3. The approximation algorithm

We will first introduce the local-ratio theorem [12,10], which was used by Bar Yehuda and Even to design a 2-approximation algorithm for the vertex cover problem. Then we give our algorithm to find a minimum weight weak vertex cover approximation in a graph with arbitrary positive weight function.

3.1. The local-ratio theorem

Theorem 3.1 (*[12]*). *If a feasible solution is r-approximation w.r.t a pair of weight functions* $W1$ *and* $W2$ *then it is also an r-approximation w.r.t* $W1 + W2$.

To use this theorem, we decompose our given weight function w into several weight functions.

Definition 3.1. Let $G = (V, E)$ be an undirected graph with vertex weight function w. A set of weight functions $\{w_i | i = 1, \ldots, k\}$ is a *decomposition* of w if, for $i = 1, \ldots, k$, w_i is also defined on the vertices in V and $\sum_i w_i(v) \leqslant w(v)$ for all $v \in V$.

Similar to the local-ratio theorem in [1], we propose the following theorem:

Theorem 3.2. *Let* $G = (V, E)$ *be an undirected graph with vertex weight function* w. *Let* $\{w_i | i = 1, \ldots, k\}$ *be a decomposition of* w, *and let S be a weak vertex cover of G such that* $w(S) = \sum_i w_i(S)$. *Then,*

$$\frac{w(S)}{w(Opt(G, w))} \leqslant \max_i \left\{ \frac{w_i(S)}{w_i(Opt(G, w_i))} \right\},$$

where $Opt(G, w)$ *denotes a minimum weight weak vertex cover of G with respect to weight function* w.

Proof. Since $\{w_i\}$ is a decomposition of w, for any set $U \subseteq V$, $w(U) \geqslant \sum_i w_i(U)$. Let $G(w_i)$ be the subgraph of G induced by all vertices satisfying $w_i(v) > 0$. If S is a weak vertex cover of G, then its restriction to any subgraph of G is a weak vertex cover as well. Therefore, $w_i(opt(G, w)) \geqslant w_i(Opt(G, w_i))$ for all i and

$$\frac{w(S)}{w(Opt(G, w))} \leqslant \frac{\sum_i w_i(S)}{\sum_i w_i(Opt(G, w))} \leqslant \frac{\sum_i w_i(S)}{\sum_i w_i(Opt(G, w_i))} \leqslant \max_i \left\{ \frac{w_i(S)}{w_i(Opt(G, w_i))} \right\}. \qquad \square$$

3.2. Description and analysis of the algorithm

In the following algorithm, we decompose the weight function w of G into cycle-rank-proportional weight functions. Since the weight of any minimal weak vertex cover is within $2 - \frac{2}{v(G)}$ of the weight of a minimum weight weak vertex cover in cycle-rank-proportional graphs, we obtain an approximation algorithm for the minimum weight weak vertex cover problem with an approximation ratio of $2 - \frac{2}{v(G)}$. We now state our new approximation algorithm *WVC*.

$WVC(G, w)$

find k, the number of connected components of G;
$v(G) = |E| - |V| + k$; $\quad \mathcal{S} = \emptyset$;
$j = 0$; $\quad w_j = w$; $\quad s = v(G)$;
while $V(w_j) \neq \emptyset$ and $s > 0$ **do**
    // $V(w_j)$ is the largest subset of V such that $w_j(v) > 0$ for any $v \in V(w_j)$
    if a nearly isolated cycle $v_1 \ldots v_k v_1$ exists in $G(V(w_j))$

6. **then**
7. select the vertex v_i with the minimum weight in $\{v_1, \ldots, v_k\}$;
8. $w^j(v) = \begin{cases} w_j(v_i) & \textit{if } v \in \{v_1, \ldots, v_k\}, \\ 0 & \textit{otherwise}. \end{cases}$
9. **else**
10. $\Delta = \min\left\{ \frac{w_j(v)}{deg(v, G(V(w_j))) - k_v - 1} \mid v \in V(w_j) \right\}$;
// $deg(v, G(V(w_j)))$ is the degree of v in $G(V(w_j))$
11. $w^j(v) = \Delta \cdot [deg(v, G(V(w_j))) - k_v - 1]$ for all $v \in V(w_j)$; **fi**
12. $w_{j+1} = w_j - w^j$;
13. $S_j = V(w_j) - V(w_{j+1})$;
14. compute k_{j+1}, the number of connected components of $G(V(w_{j+1}))$;
15. $s = |E(w_{j+1})| - |V(w_{j+1})| + k_{j+1}$;
16. $j = j + 1$; **od**
17. **for** $t = j - 1$ **downto** 0 **do**
18. $\mathcal{S} = \mathcal{S} \cup S_t$;
19. **for** every $v \in S_t$ **do**
20. **if** $V(w_t) - (S - v)$ induces no cycle in $G(V(w_t))$
21. **then** remove v from S;
22. **od**
23. **od**
24. **return** S

The following two observations are obvious.

Claim 3.1. *For all j,* $w(v) \geqslant \sum_j w^j(v)$ *for each* $v \in V$.

Claim 3.2. *The set S computed by algorithm WVC satisfies* $w(S) = \sum_i w^i(S)$.

We can now prove our main result.

Theorem 3.3. *For an undirected weighted graph G, algorithm WVC computes a minimum weight weak vertex cover approximation with approximation ratio* $2 - \frac{2}{\nu(G)}$.

Proof. Algorithm WVC contains two main loops. The first loop decomposes the weight function w into functions $\{w^0, \ldots, w^{j-1}\}$, where each w^i is cycle-rank-proportional for G. In line 13 we set S_i to be the set of vertices of weight zero after the ith iteration.

In the second loop, we collect the S_i, for $0 \leqslant i \leqslant j - 1$, and remove redundant vertices. If the set S computed by the algorithm is a weak vertex cover of G and its restriction to any w^i is a minimal weak vertex cover of $G(w^i)$, then the algorithm is correct.

After the first phase, $S \subseteq S_{j-1}$ because some redundant vertices will be removed. It is easy to see that S is a weak vertex cover of $G(w_{j-1})$ and its restriction to w^{j-1} is a minimal weak vertex cover. Suppose in the ith phase S is a weak vertex cover of $G(w_{j-i})$ and its restriction to w^{j-i} is a minimal weak vertex cover. Now we consider the $(i+1)$th phase. $G(w_{j-i})$ is a subgraph of $G(w_{j-i-1})$ and $V(w_{j-i-1}) = V(w_{j-i}) \cup S_{j-i-1}$. Hence, $S = S \cup S_{j-i-1}$ is a weak vertex cover of $G(w_{j-i-1})$. After the removal of redundant vertices, S is a minimal weak vertex cover of $G(w^{j-i-1})$ because $V(w_{j-i-1}) = V(w^{j-i-1})$. When the last phase ends, S is a weak vertex cover of $G(w_0) = G$ and a minimal weak vertex cover of $G(w^0)$.

Each $G(w^i)$ must be either a nearly isolated cycle or a cycle-rank-proportional graph. The analysis in Section 2 shows that the total weight of a minimal weak vertex cover of a cycle-rank-proportional graph is within $2 - \frac{2}{\nu(G)}$ of the minimum weight weak vertex cover. For a nearly isolated cycle, we choose the vertex with minimum weight among all its vertices, i.e., the vertex chosen by the algorithm has the same weight as the optimum value in the weighted subgraph.

According to Theorem 3.2, the approximation ratio of our algorithm is bounded by the maximal approximation ratio of the minimal weak vertex covers of the decompositions. Therefore, the total weight of the set $\mathcal{S}$ computed by the algorithm is within a factor of $2 - \frac{2}{v(G)}$ of the optimum. □

Theorem 3.4. *The running time of algorithm WVC is*

$$\mathcal{O}(|V| \cdot (|E| + |V|) \cdot \min\{v(G), |V|\}).$$

Proof. Line 1 can be finished in time $\mathcal{O}(|E| + |V|)$ by using $\mathcal{BFS}$ [3]. Lines 2 and 3 run in constant time. Line 5 takes time $\mathcal{O}(|V|)$ because it may check all the vertices. Since w is a vector of rank $|V|$, lines 6–8 take time $\mathcal{O}(|V|)$. Computing k_v takes $\mathcal{O}(|E| + |V|)$ by using $\mathcal{BFS}$, hence line 11 can be finished in time $\mathcal{O}(|V| \cdot (|E| + |V|))$. Lines 12–16 take time $\mathcal{O}(|V| + |E|)$ time. Notice that the while loop is executed for at most $\min\{v(G), |V|\}$ iterations. Hence the total running time of the while loop is $\mathcal{O}(|V| \cdot (|E| + |V|) \cdot \min\{v(G), |V|\})$.

Since $j \leqslant \min\{v(G), |V|\}$ and lines 18–21 take time $\mathcal{O}(|V|)$, the total running time of the for loop is $\mathcal{O}(|V| \cdot \min\{v(G), |V|\})$. Therefore, the total time of algorithm *WVC* is $\mathcal{O}(|V| \cdot (|E| + |V|) \cdot \min\{v(G), |V|\})$. □

4. Inapproximability result

In this section, we present the inapproximability result for the Minimum Weak Vertex Cover problem.

Obviously, the reduction in [11], which reduces Vertex Cover to Weak Vertex Cover, "preserves" the size of optimal solutions. That is, any optimal solution for an instance $\mathcal{I}$ of Vertex Cover is still optimal for the corresponding instance $\mathcal{I}'$ of Weak Vertex Cover made by the reduction in [11]. On the other hand, we can alter the optimal solution for the instance $\mathcal{I}'$ of Weak Vertex Cover, which is reduced from an instance $\mathcal{I}$ of Vertex Cover, to the optimal solution for $\mathcal{I}$ without increasing the number of the vertices. Formally,

Lemma 4.1. *Let τ be the reduction in* [11] *from the Vertex Cover problem to the Weak Vertex Cover problem, I be any instance of the Vertex Cover problem and $opt(I)$ be the optimal solution of I. For any constant m, n such that $0 \leqslant m, n \leqslant |I|$, we have*

$$|opt(I)| \leqslant m \Rightarrow |opt(\tau(I))| \leqslant m,$$

$$|opt(I)| \geqslant n \Rightarrow |opt(\tau(I))| \geqslant n.$$

Together with the inapproximability result for the Minimum Vertex Cover problem by Dinur and Safra [4], i.e., no polynomial-time algorithm can solve the Minimum Vertex Cover problem within approximation ratio $10\sqrt{5} - 21 \approx 1.36$, we have the following theorem, which gives a lower bound for the inapproximability of Minimum Weak Vertex Cover problem.

Theorem 4.1. *There is no polynomial time algorithm that can approximate the Minimum Weak Vertex Cover problem within $\rho = 1.36$, unless $P = NP$.*

References

[1] V. Bafna, P. Berman, T. Fujito, Constant ratio approximations of the weighted feedback vertex set problem for undirected graphs, in: Proc. of the Sixth Internat. Symp. on Algorithms and Computation (ISAAC'95), Cairns, Australia, December 4–16, 1995, pp. 142–151.

[2] Y. Breibart, C. Chan, M. Carofalakis, R. Rastogi, A. Silberschatz, Efficiently monitoring bandwidth and latency in IP network, in: Proc. of the IEEE INFOCOM 2001, Alaska, USA, 2001, pp. 933–942.

[3] T. Corman, C. Leiserson, R. Rivest, C. Stein, Introduction to Algorithms, second ed., The MIT Press, Cambridge, MA, 2001.

[4] I. Dinur, S. Safra, The importance of being biased, in: Proc. of the 34th Annu. ACM Symp. on the Theory of Computing (STOC'03), Montral, Qubec, Canada, May 19–21, 2002, pp. 33–42.

[5] M. Gondran, M. Minoux, Graph and Algorithms, Wiley, New York, 1984.

[6] T. Haynes, S. Hedetniemi, S. Hedetniemi, M. Henning, Domination in graphs applied to electric power networks, SIAM J. Discrete Math. 15 (4) (2002) 519–529.

[7] S. Jamin, C. Jin, Y. Jin, D. Raz, Y. Shavitt, L. Zhang, On the placement of internet instrumentation, in: Proc. of the IEEE INFOCOM 2000, Tel-Aviv, Israel, 2000, pp. 295–304.

[8] K. Lai, M. Baker, Measuring bandwidth, in: Proc. of the IEEE INFOCOM'99. New York, 1999, pp. 235–245.
[9] X. Liu, J. Yin, L. Tang, J. Zhao, Analysis of efficient monitoring method for the network flow, J. Software 14 (2) (2003) 300–304.
[10] R. Yehuda, S. Even, A local-ratio theorem for approximating the weighted vertex cover problem, Ann. Discrete Math. 25 (1985) 27–45.
[11] Y. Zhang, H. Zhu, An approximation algorithm for weighted weak vertex cover, J. Comput. Sci. Technol. 19 (6) (2004) 782–786.
[12] R. Yehuda, Homepage ⟨http://www.cs.technion.ac.il/∼reuven⟩.