Results on the min-sum vertex cover problem

Ralucca Gera,[1] **Craig Rasmussen, Pantelimon Stănică**[1]
Naval Postgraduate School
Monterey, CA 93943, USA
{rgera, ras, pstanica}@nps.edu
and
Steve Horton
United States Military Academy
West Point, NY 10996, USA
steve.horton@usma.edu

ABSTRACT

Let G be a graph with the vertex set $V(G)$, edge set $E(G)$. A *vertex labeling* is a bijection $f : V(G) \to \{1, 2, \ldots, |V(G)|\}$. The *weight* of $e = uv \in E(G)$ is given by $g(e) = \min\{f(u), f(v)\}$. The min-sum vertex cover (msvc) is a vertex labeling that minimizes the vertex cover number $\mu_s(G) = \sum_{e \in E(G)} g(e)$. The minimum such sum is called the msvc *cost*. In this paper, we give both general bounds and exact results for the msvc cost on several classes of graphs.

Key Words: labeling, vertex cover, independence.

AMS Subject Classification: 05C15, 05C69.

1 Introduction and motivation

Let G be a graph with the vertex set $V(G)$ and edge set $E(G)$. A *vertex labeling* is a bijection $f : V(G) \to \{1, 2, \ldots, |V(G)|\}$, and the *weight* of an edge $e = uv \in E(G)$ is given by $g(e) = \min\{f(u), f(v)\}$. For simplicity, we call a vertex labeling a *labeling* of G. The *cost* of a labeling f is $\mu_f(G) = \sum_{e \in E(G)} g(e)$. A *min-sum vertex cover (msvc)* or an *msvc labeling* is a labeling that minimizes μ_f over all choices of f. Formally, $\mu_s(G) = \min_f \mu_f(G)$, where $\mu_s(G)$ is called the *msvc cost* of the graph G. Given a labeling f, we define a *cost set* $S_f(G) = \{u \in V(G) : \exists e \in E(G), f(u) = g(e)\}$. That is, $S_f(G)$ is the subset of $V(G)$ that induces the weights on the edges. A cost set associated with an msvc labeling is called an *msvc set*. Note that an msvc set does not have to be a minimum size cost set (See section 6).

Consider the graph in Figure 1 below. The function $f(v_1) = 4, f(v_2) = 5, f(v_3) = 1, f(v_4) = 2$, and $f(v_5) = 3$ is a labeling of G with cost $\mu_f(G) = 12$. Under f, $\{v_3, v_4, v_5\}$ is a cost set. On the other hand, the function f^* with $f^*(v_i) = i$ is another labeling of G with cost $\mu_{f^*}(G) = 9$. Under f^*, $\{v_1, v_2\}$ is a cost set. It is easy to see that $\mu_s(G) = \mu_{f^*}(G) = 9$, so we say $\{v_1, v_2\}$ is a minimum cost set, and that graph G has an msvc cost of 9.
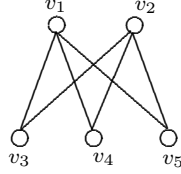

Figure 1: Example for Discussion of Labeling and Cost Sets

Given a graph G, an *independent set* $I(G)$ is a subset of $V(G)$ such that no two vertices in $I(G)$ are adjacent. The maximum independent set problem is to find an independent set with the largest number of vertices in a given graph. We use the notation $\alpha(G)$ to denote the maximum cardinality of an independent set in a graph G. A *vertex cover* $C(G)$ is a subset of $V(G)$ such that each edge of G is incident with a vertex in $C(G)$. The minimum vertex cover problem seeks a set $C(G)$ of smallest cardinality. We use the notation $\beta(G)$ to denote the minimum cardinality of a vertex cover in a graph G. The decision versions of these problems are both NP-Complete [3].

It is easy to see that any *minimal* vertex cover is a cost set. Given a graph G and a minimal vertex cover $S \subseteq V(G)$, select a function $f : V(G) \rightarrow \{1, 2, \ldots, |V(G)|\}$ so that vertices in S get labels in $\{1, 2, \ldots, |S|\}$. Then g is dictated by f. Each vertex in S is adjacent to at least one vertex in $V(G) - S$. Edges connecting a vertex $u \in S$ with a vertex $v \in V(G) - S$ will get weight $f(u)$. Since S is a minimal vertex cover, each weight in $\{1, 2, \ldots, |S|\}$ will be realized for some edge, and $g(e) \leq |S|$ holds for all $e \in E(G)$. This allows us to make the following observation.

Observation 1.1 *For any graph G, $\mu_s(G)$ is bounded from above by the cost of any labeling f that uses labels $\{1, 2, \ldots, |S|\}$ on the vertices of a minimal vertex cover S.*

Since *any* labeling of G provides an upper bound on $\mu_s(G)$, this observation is certainly not surprising. But we will see that for certain classes of graphs, vertex covers and also independent sets will play a role in establishing improved bounds.

2 Bounds on μ_s for connected graphs

Proposition 2.1 *For a connected graph G,*

$$\mu_s(G) \geq \frac{\beta^2 + 3\beta - 2}{2}, \tag{1}$$

where $\beta = \beta(G)$ is the vertex cover number of G. The bound is sharp.

Proof. Let G be a connected graph. Let $B = \{v_1, v_2, \ldots, v_\beta\}$ be a minimum vertex cover. Thus for each $v_i \in B$ $(1 \leq i \leq \beta)$ there is an edge e_i that is incident with only one vertex in B (otherwise B is not minimum vertex cover). Also, since G is connected, there are at least $\beta - 1$ edges different from e_i, say g_j $(1 \leq j \leq \beta - 1)$. Since each edge e_i will receive the weight $f(v_i)$ with $f(v_i) \neq f(v_{i'})$ $(1 \leq i, i' \leq \beta)$, and each edge g_j will receive a weight of at least 1, it follows that

$$\mu_s(G) \geq \sum_{i=1}^{\beta} i + (\beta - 1) = \frac{\beta(\beta + 1)}{2} + (\beta - 1) = \frac{\beta^2 + 3\beta - 2}{2}.$$

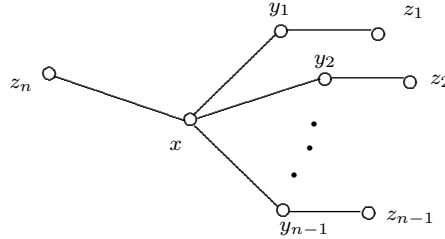To see the sharpness of the bounds, consider the graph G in Figure 2 below.

Figure 2: The graph G

Then the msvc set $\{x\} \cup \{z_i : 1 \leq i \leq n - 1\}$ gives the sharpness of the bound (1). $\qquad \square$

Proposition 2.2 *For a connected graph G with n vertices,*

$$n - 1 \leq \mu_s(G) \leq \frac{n(n^2 - 1)}{6}, \tag{2}$$

and the bounds in (2) are sharp.

Proof. Let G be a connected graph of order n. Then G has at least $n - 1$ edges of weight at least 1, so $\mu_s(G) \geq n - 1$. For the second inequality, let

G be any connected graph. The vertex labeled i has at most $n-i$ neighbors with labels exceeding i. Therefore at most $n - i$ edges have weight i, and it follows that

$$
\begin{aligned}
\mu_s(G) &\leq \sum_{i=1}^{n} i(n - i) = \sum_{i=1}^{n}(ni - i^2) \\
&= n \cdot \frac{n(n + 1)}{2} - \frac{n(n + 1)(2n + 1)}{6} = \frac{n(n^2 - 1)}{6}.
\end{aligned}
$$

Sharpness of the first inequality follows from labeling the center vertex of $K_{1,n-1}$ with 1, and sharpness of the second follows from any labeling of K_n. $\square$

Note, though, that if $G \neq K_n$, then

$$
\mu_s(G) \leq \frac{n(n^2 - 1)}{6} - (n - 1) = \frac{(n - 1)(n - 2)(n + 3)}{6}.
$$

Thus not all integer pairs (n, k), with $n - 1 \leq k \leq \frac{n(n^2-1)}{6}$, can be realized as the order and cost of some connected graph G; for an example, consider the pair $(n, n - 1)$. In particular, if k cannot be written as $\sum_{i=1}^{n-1} a_i \cdot i$ with $a_i \leq n - i$, then there is no graph G of order n for which $\mu_s(G) = k$. Although $k = \sum_{i=1}^{n-1} a_i \cdot i$, with $a_i \leq n - i$, is a necessary condition, it is not sufficient. For example, there is no connected graph G of order 4 for which $\mu_s(G) = 7$.

Note that if H is a proper induced subgraph of G, then the min-sum vertex cover of G is a set cover of H (possibly minimum), which gives the following relation for the costs of the graphs: $\mu_s(H) \leq \mu_s(G)$. Moreover, since $H \neq G$ the inequality is strict, and $\mu_s(H) < \mu_s(G)$.

3 Bounds on μ_s for regular graphs

In this section, we consider μ_s for r-regular graphs.

Theorem 3.1 *If G is an r-regular graph with n vertices, then*

$$
r\frac{\alpha(\alpha + 1)}{2} \leq \mu_s(G) \leq r\frac{\beta(\beta + 1)}{2},
$$

Where $\alpha = \alpha(G)$ and $\beta = \beta(G)$. The bounds are sharp.

Proof. For the first inequality, since G is r-regular, each label on $V(G)$ can be given to at most r edges. If we label the vertices of any independent set $I(G)$ with labels $1, 2, \ldots, |I(G)|$, then each weight in the set

$\{1, 2, \ldots, |I(G)|\}$ gets assigned to exactly r edges. Then, selecting a maximum independent set and simply adding up the contribution at each independent vertex, a lower bound on $\mu_s(G)$ is $r \sum_{1 \le i \le \alpha(G)} i = r \frac{\alpha(G)(\alpha(G)+1)}{2}$. The second inequality is just the application of Observation 1.1 to the r-regular case. The double inequality is sharp, with equality for the complete bipartite graph $K_{r,r}$. $\qquad \square$

In essence, the first inequality represents the best possible use of the labels $\{1, 2, \ldots, \alpha(G)\}$. Also note that for r-regular graphs with $\alpha(G) = \beta(G)$, the inequalities of Theorem 3.1 collapse around $\mu_s(G)$. But since the complement of any independent set is a vertex cover, we know $\alpha(G) + \beta(G) = |V(G)| = n$. This implies that for a regular graph $\alpha(G) = \beta(G) = \frac{n}{2}$ for such graphs, and a corollary follows. The double inequality is sharp, with equality for the complete bipartite graph $K_{r,r}$.

Corollary 3.2 *If G is an r-regular graph with $\alpha(G) = \beta(G)$, then*

$$\mu_s(G) = \frac{r(n^2 + 2n)}{8}.$$

Proof. Substitute $\frac{n}{2}$ for α and β in Theorem 3.1. $\qquad \square$

Since computing $\alpha(G)$ and $\beta(G)$ is not trivial, the result of Theorem 3.1 can be used to get the following.

Corollary 3.3 *If G has girth $2k + 3$, $k \ge 2$ then*

$$\mu_s(G) \ge \frac{r \left(\frac{n}{2}\right)^{\frac{k-1}{k}} r^{\frac{1}{k}} \left(\left(\frac{n}{2}\right)^{\frac{k-1}{k}} r^{\frac{1}{k}} + 1 \right)}{2}. \tag{3}$$

A simpler bound (under the previous assumptions) but slightly weaker is

$$\mu_s(G) \ge \frac{r \sqrt{nr/6} \left(\sqrt{nr/6} + 1 \right)}{2}. \tag{4}$$

Proof. Shearer in [4] states that, under the above assumptions, the independence number has the lower bound

$$\alpha(G) \ge \left(\frac{n}{2}\right)^{\frac{k-1}{k}} r^{\frac{1}{k}},$$

which gives us equation (3).

If G has no cycles of length 3 or 5, Denley [1] proved the simpler result $\alpha(G) \ge \sqrt{nr/6}$ producing the bound in equation (4). $\qquad \square$

Proposition 3.4 *If G is an r-regular graph with n vertices, then*

$$\frac{rn(n+2)}{8} \leq \mu_s(G) \leq \frac{n(n+1)}{6}$$

and the inequalities are sharp.

Proof. Feige et al. [2] showed that $\mu_s(G)$ satisfies

$$\frac{e(n+2)}{4} \leq \mu_s(G) \leq \frac{e(n+1)}{3},$$

where e is the number of edges of an r-regular graph G. Since $2e = nr$, we obtain

$$\frac{n(n+2)r}{8} \leq \mu_s(G) \leq \frac{n(n+1)r}{6}.$$

The double inequality is sharp, with lower equality for the complete bipartite graph $K_{r,r}$, and upper equality for the complete graph K_n. □

4 Elementary results on μ_s

Proposition 4.1 (a) *For the star $K_{1,n-1}$, we have $\mu_s(K_{1,n-1}) = n - 1$.*

(b) *For the path P_n on n vertices, we have $\mu_s(P_n) = \left\lfloor \frac{n^2}{4} \right\rfloor$.*

(c) *For the cycle C_n on n vertices, we have*

$$\mu_s(C_n) = \begin{cases} \frac{n(n+2)}{4} & \text{if } n \text{ is even} \\ \frac{(n+1)^2}{4} & \text{if } n \text{ is odd.} \end{cases}$$

(d) *For the wheel $W_{1,n}$ on $n+1$ vertices, we have*

$$\mu_s(W_{1,n}) = \begin{cases} \frac{n(n+6)}{4} & \text{if } n \text{ is even} \\ \frac{(n+1)^2 + 4n}{4} & \text{if } n \text{ is odd.} \end{cases}$$

Proof.

(a) The labeling that assigns 1 to the central vertex, and the rest of the labels to the other vertices is an msvc label that gives the result.

(b) Let $P_n : v_1, v_2, \ldots, v_n$ be the path on n vertices, and define a labeling $f : V(P_n) \to \{1, 2, \ldots, n\}$ such that $f(v_{2i}) = i$ for each $1 \leq i \leq \left\lfloor \frac{n}{2} \right\rfloor$,

and any unused labels to the rest of the vertices. This induces the weights on the edges of the graph, and so for n odd

$$\mu_s(P_n) \le 2\sum_{i=1}^{\frac{n-1}{2}} i = 2\frac{\frac{n-1}{2}\left(\frac{n-1}{2}+1\right)}{2} = \frac{n^2-1}{4} = \left\lfloor \frac{n^2}{4} \right\rfloor,$$

and for even n

$$\mu_s(P_n) \le 2\sum_{i=1}^{\frac{n}{2}-1} i + \frac{n}{2} = 2\frac{\left(\frac{n}{2}-1\right)\frac{n}{2}}{2} + \frac{n}{2} = \left\lfloor \frac{n^2}{4} \right\rfloor.$$

Since $\deg v_i \le 2$, each label on a vertex induces weights on at most two edges, so the above labeling is an msvc labeling. The result follows.

(c) Similar to (b).

(d) The labeling that assigns the label 1 to the central vertex, and 2 through $n+1$ to every other vertex of the outer cycle is an msvc labelinmg with $\mu_s(W_{1,n}) = n + \mu_s(C_n) = \frac{n(n+6)}{4}$ if n is even, and $\mu_s(W_{1,n}) = n + \mu_s(C_n) = \frac{(n+1)^2+4n}{4}$ if n is odd. $\qquad\square$

The star can be generalized to the multi-star $K_m(a_1, a_2, \ldots, a_m)$, which is formed by joining $a_i \ge 1$ ($1 \le i \le m$) pendant vertices to each vertex x_i of a complete graph $K_m : x_1, x_2, \ldots, x_m$. The 2-star and 3-star are shown in Figure 3.
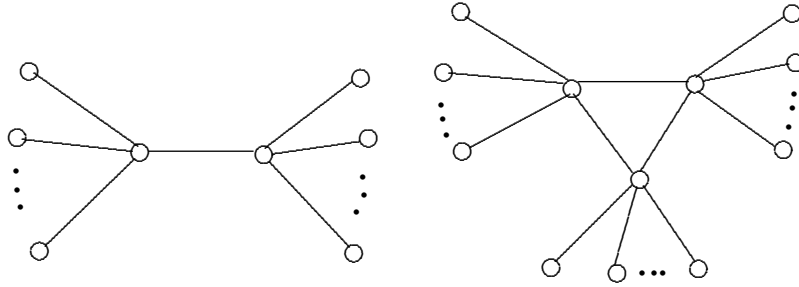

Figure 3: 2-star $K_2(a_1, a_2)$ and 3-star $K_3(a_1, a_2, a_3)$

The first claim of Proposition 4.1 can be generalized to any multi-star.

Proposition 4.2 *The multi-star $K_m(a_1, \ldots, a_m)$, with $a_1 \geq a_2 \geq \cdots \geq a_m \geq 1$, has*

$$\mu_s(K_m(a_1, \ldots, a_m)) = \sum_{i=1}^{m} i a_i + \frac{m(m^2 - 1)}{6}.$$

Proof. Since $a_1 \geq a_2 \geq \cdots \geq a_m \geq 1$, it follows that the central vertices of the multistar $K_m(a_1, \ldots, a_m)$ form a minimum msvc set

$$
\begin{aligned}
\mu_s(K_m(a_1, \ldots, a_m)) &= (a_1 + m - 1) + 2(a_2 + m - 2) + \cdots \\
&\quad + (m-1)(a_{m-1} + 1) + m a_m \\
&= \sum_{i=1}^{m} i(a_i + m - i) = \sum_{i=1}^{m} i a_i + \frac{m(m^2 - 1)}{6},
\end{aligned}
$$

which proves the result. $\qquad\square$

5 μ_s for biregular bipartite graphs

A graph G is *biregular* if $V(G) = V_1 \cup V_2$ where all vertices in V_1 have degree r, and all vertices in V_2 have degree s, for natural numbers r, s. Note that if $r = s$, then we obtain a regular graph. If $r \neq s$ then we call G *strictly biregular*. If the graph G happens to be bipartite, then we call it *biregular bipartite* or *strictly biregular bipartite*, accordingly.

Theorem 5.1 *Let G be a biregular bipartite graph on partite sets $V_1 = \{v_1, \ldots, v_p\}$ and $V_2 = \{w_1, \ldots, w_q\}$, where $p \leq q$, all vertices in V_1 have degree r, and all vertices in V_2 have degree s, for natural numbers r and s. Then*

$$\mu_s(G) = \frac{r \cdot p(p+1)}{2}.$$

Proof. We first consider G to be a strictly biregular bipartite graph. Since $p < q$ and $pr = qs$, it follows that $r > s$. Since G is connected, $s \geq 1$. Since any labeling of G with V_1 as a cost set yields the same cost, it follows that $\mu_s(G) \leq \frac{r \cdot p(p+1)}{2}$. We now show that this is optimal. Given a cost set S, let N_i denote the number of edges of weight i, where $1 \leq i \leq n$. Clearly $\sum_{i=1}^{n} N_i = |E(G)| = rp$ and, since $\Delta(G) = r$, we know that $N_i \leq r$ for all i. With any labeling f as described above, we have $N_i = \begin{cases} r, & \text{if } i \leq p; \\ 0, & \text{if } i > p. \end{cases}$

The resulting cost is $\mu_f(G) = \dfrac{rp(p+1)}{2}$. Now suppose that we have some labeling h for which $S_h \bigcap V_2 \neq \emptyset$. Then $N_i < r$ for at least one index $i \leq p$, so $\sum_{i=1}^{p} N_i < rp$ and therefore $N_k \geq 1$ for some $k > p$. But then $\mu_h(G) \geq \mu_f(G)+1$, and it follows that $V_2 \bigcap S = \emptyset$ in every optimal solution.

If G is a biregular graph that is not strictly biregular graph, then $r = s$ and so either V_1 or V_2 is an msvc set with the msvc cost of

$$\mu_s(G) = \frac{r \cdot p(p+1)}{2},$$

as desired. □

The following two corollaries are direct consequences of Theorem 5.1.

Corollary 5.2 *Let Q_n be the n-cube with 2^n vertices, where $n \geq 2$. Then $\mu_s(Q_n) = n2^{n-2}(2^{n-1}+1)$.*

Corollary 5.3 *Let $K_{a,b}$ be the complete bipartite graph with $a+b$ vertices $(a \leq b)$. Then $\mu_s(K_{a,b}) = \dfrac{ba(a+1)}{2}$.*

Proposition 5.4 *Let $G \cong C_n \times K_2$ be the prism graph with $2n$ vertices. Then*

$$\mu_s(G) \;\; = \;\; \begin{cases} \frac{3n(n+1)}{2}, & \text{if } n \text{ is even;} \\ \frac{3n^2+3n+2}{2}, & \text{if } n \text{ is odd.} \end{cases}$$

Proof. If n is even, we observe that G is 3-regular with $\alpha(G) = \beta(G) = n$, and we can apply Corollary 3.2 (although note that the term n is used in different senses).

If n is odd, then a maximum independent set of G has $n-1$ elements; this is seen easily by observing that one cannot have more than $(n-1)/2$ independent vertices either on the inside or the outside cycle, and one can take exactly $(n-1)/2$ vertices on either of the two cycles, which form an independent set. We display such a choice in Figure 4, where the solid vertices form an independent set.

We now label the independent set in Figure 4 with labels $1, 2, \ldots, n-1$ for a contribution to the cost of $3\frac{n(n-1)}{2}$. This labeling uses all but 3 edges which may form a copy of P_3, or a copy of P_2 together with a copy of K_2, or 3 copies of K_2. In the first two cases, these three edges will have the same contribution to the cost, namely $2n + (n+1)$, and in the last case they will have a contribution of $n + (n+1) + (n+2)$. Regardless whether the last case occurs, the msvc cost will be given by one of the first two choices.
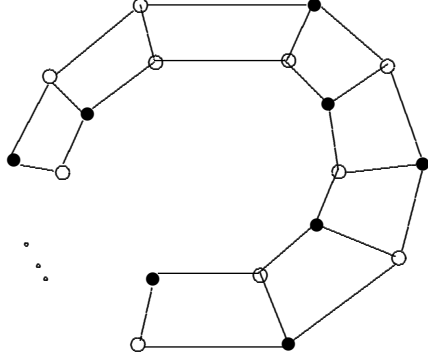
9

Figure 4: The independent set for the prism $C_n \times K_2$

If there is an msvc set of G containing at most $n-2$ independent vertices, say that $n - 1 - j$ ($j \geq 1$) vertices contribute the weight of 3 edges each to the cost, and the rest of the vertices will contribute the weight of at most 2 edges, then the msvc cost is greater than

$$3 \sum_{i=1}^{n-1-j} i + 2 \sum_{\ell=1}^{3(j+1)/2} (n - j + \ell - 1) = \frac{3}{2}\left(n^2 + n + \frac{j^2}{2} + j + \frac{1}{2}\right), \text{ or}$$

$$3 \sum_{i=1}^{n-1-j} i + 2 \sum_{\ell=1}^{(3j+2)/2} (n - j + \ell - 1) + n + 3 + \frac{j}{2} = \frac{3}{2}\left(n^2 + n + \frac{j^2}{2} + j + 2\right),$$

depending upon whether j is odd or even, respectively. In either case, it is a greater cost than the one we previously obtained. Therefore, the msvc cost of the prism $C_n \times K_2$, if n is odd, is

$$\mu_s(C_n \times K_2) = 3\frac{n(n-1)}{2} + 2n + n + 1 = \frac{3n^2 + 3n + 2}{2},$$

which proves the result. □

6   Closing remarks

Note that if a graph G has a minimal vertex cover that happens to also be an independent set, then this set is a cost set, and an upper bound for the cost of the graph can be easily computed. However, it is not always the case that even a minimum vertex cover that is independent is an msvc cost set. To see this, consider the graph G obtained from the star $K_{1,n}$ by subdividing

all but one edges as shown Figure 2. The set $\{z_n\} \cup \{y_i : 1 \leq i \leq n-1\}$ is an independent vertex cover, but not an msvc cost set. However, the set $\{x\} \cup \{z_i : 1 \leq i \leq n-1\}$ is both an independent vertex cover and an msvc cost set.

We have frequently used independent vertices as an initial subset in finding an msvc set in this paper. That raises the question whether there is a connected graph G such that

(a) no minimum cardinality cost set of G includes a maximum independent set of G, or

(b) no msvc set includes a maximum independent set of G.

To see this consider the double star $K_2(a_1, a_2)$ of Figure 3, for $a_1, a_2 \geq 2$. Observe then that the unique maximum independent set is the set of end vertices, but the unique minimum cardinality msvc set is given by the central vertices. Thus the answer is yes to both questions.

There are also regular graphs whose maximum independent set is not a subset of the minimum cardinality msvc set, as we can see in the graph of Figure 5.
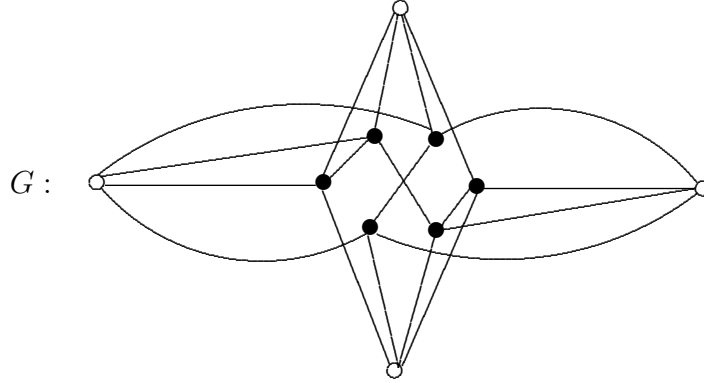

Figure 5: Maximum independent set is not a subset of the msvc set

For this graph, the clear vertices form the unique maximum independent set, which is not a subset of the set of solid vertices that form the minimum cardinality msvc set S. However, there is a larger msvc set that produces the same cost of the graph as S, and it includes the maximum independent set of the graph.

Problem 6.1 *Is there a connected regular graph G for which no msvc set contains a maximum independent set?*

Another question might be whether choosing the vertices of maximum degree as a subset of the cost set will always produce an msvc set. The answer is no, as we discuss below. Note that caterpillars are a class of graphs for which the greedy algorithm does not always produce a min-sum vertex cover. To see this, consider the two labelings of the graphs below ($cost(G)$ is the cost associated with the implied labeling).
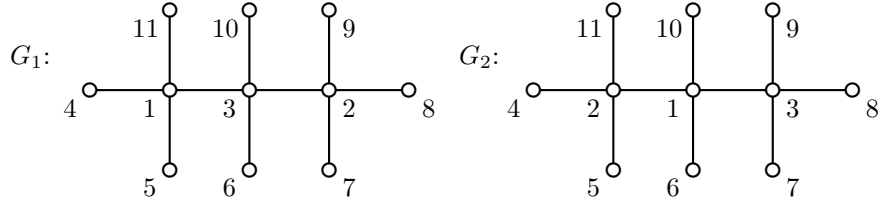

Figure 6: $cost(G_1) = 18$ and $cost(G_2) = 19$

References

[1] T. Denley, *The Independence Number of Graphs with Large Odd Girth*, Electronic J. Combin. **1** (1994) #R9 (12pp).

[2] U. Feige, L. Lovász, P. Tetali, *Approximating min-sum set cover*, Algorithmica **40** (2004), 219–234.

[3] M. R. Garey and D. S. Johnson, *Computers and Intractability*, W. H. Freeman and Co (1978).

[4] J.B. Shearer, *The Independence Number Of Dense Graphs With Large Odd Girth*, Electronic J. Combin. **2** (1995), #N2 (3pp).