The Greedy Algorithm for Domination in Cubic Graphs

Suzanne M. Seager
Mount Saint Vincent University, Halifax, NS, Canada

Abstract. We show that for a cubic graph on n nodes, the size of the dominating set found by the greedy algorithm is at most $\frac{4}{9}n$, and that this bound is tight.

1 Introduction

Let G be a graph, witn n nodes and e edges. Let $N(v)$ be the set of neighbours of a node v and $N[v] = N(v) \cup \{v\}$. Let $d(v) = |N(v)|$ be the *degree* of v. G is *r-regular* if $d(v) = r$ for all v; if $r = 3$ then G is *cubic*. Let $N[S] = \bigcup_{v \in S} N[v]$ and let $N(S) = \bigcup_{v \in S} N(v)$. A *dominating set* for G is a set D with every node in $N[D]$. The *domination number* $\gamma(G)$ is the minimum size of a dominating set of G. The *girth* of G is the minimum size of a cycle in G.

The problem of finding the domination number of a graph is NP-hard, even when restricted to cubic graphs. One simple heuristic is the greedy algorithm, initially analyzed by Chvatal [1], Johnson [4], and Lovász [5] in the context of set coverings. Let d_g be the size of the dominating set returned by the greedy algorithm. In 1991 Parekh [6] showed that $d_g \leq n + 1 - \sqrt{2e+1}$.

Recently, new bounds have been discovered on $\gamma(G)$ for cubic graphs. Reed [7] proved that $\gamma(G) \leq \frac{3}{8}n$. Fisher et al. [2,3] repeated this result and showed that if G has girth at least 5 then $\gamma(G) \leq \frac{5}{14}n$. In the light of these bounds on γ, we consider bounds on d_g for cubic graphs.

2 The Greedy Algorithm

Let $V(G) = \{1, 2, ..., n\}$. The greedy algorithm for finding a dominating set D as described by Parekh [6] starts with $D = \emptyset$ and all nodes uncovered. At each iteration, put into D the least indexed node that covers (i.e., equals or is adjacent to) the maximum number of uncovered nodes, until all nodes are covered; the least indexed node is chosen as a way of breaking ties. When the graph G has maximum degree 3, this can be formalized as follows:

Algorithm GREEDY to find a dominating set D for graph G
Set $D_k = \emptyset, k = 4, 3, 2, 1$. Set $U = V(G)$.
For $k = 4, 3, 2, 1$
    For $v = 1, 2, ..., n$
        If $|N[v] \cap U| = k$, add v to D_k and set $U = U - N[v]$.
Set $D = \bigcup_{k=1}^{4} D_k$.

Theorem 1 *For a cubic graph G, $d_g \leq \frac{4}{9}n$.*

Proof. Let $d_k = |D_k|, k = 4, 3, 2, 1$. Then $d_g = |D| = d_4 + d_3 + d_2 + d_1$. Let $G_4 = G$ and $G_k = G_{k+1} - N[D_{k+1}]$ for $k = 3, 2, 1$; so $U = V(G_k)$ at the start of the iteration for k. Note that G_k has maximum degree at most $k-1$, since any node of degree k in G_k would have been uncovered with k uncovered neighbours during the iteration for $k + 1$, and thus added to D_{k+1}. Moreover, no node in G has three edges to G_2 (or it would have been in D_3) or two edges to G_1.

Each node in D_4 covers exactly four uncovered nodes, so G_3 has $n - 4d_4$ nodes; similarly G_2 has $n - 4d_4 - 3d_3$ nodes and G_1 has $n - 4d_4 - 3d_3 - 2d_2$ nodes. Each of the $3d_4$ nodes in $N(D_4)$ has degree 3 in G_4, so there are at most $9d_4$ edges incident with $N(D_4)$ in G_4. These include all edges within $N[D_4]$ and all edges between $N[D_4]$ and G_3, so G_3 has at least $e - 9d_4$ edges. After the first iteration ($k = 4$), every node in $N[D_4]$ is covered and has at least one covered neighbour, leaving at most two uncovered neighbours, so no node in $N[D_4]$ can be in D_3. Thus every node in D_3 must be a node of degree 2 in G_3, which means there are $2d_3$ nodes in $N(D_3)$, each of degree at most 2 in G_3. So there are at most $4d_3$ edges within $N[D_3]$ and between $N[D_3]$ and G_2, leaving at least $e - 9d_4 - 4d_3$ edges in G_2. Let x be the number of edges in G_2; then $x \geq e - 9d_4 - 4d_3$. Since G_2 has $n - 4d_4 - 3d_3$ nodes and maximum degree 1, it consists of x copies of K_2 and $n - 4d_4 - 3d_3 - 2x$ isolated nodes, so $2x \leq n - 4d_4 - 3d_3$. Thus

$2(e - 9d_4 - 4d_3) \leq n - 4d_4 - 3d_3$ and so $5d_3 + 14d_4 \geq 2e - n$. Since G is cubic, $2e = 3n$ and so $5d_3 + 14d_4 \geq 2n$ (1).

Each node in D_2 is either in $G - G_2$ with two neighbours in G_2, or is in G_2 with one neighbour in G_2. Let y be the number of nodes in $D_2 \cap (G - G_2)$, leaving $d_2 - y$ nodes in $D_2 \cap G_2$. Then there are $2y$ neighbours of $D_2 \cap (G - G_2)$ in G_2. Let z be the number of these neighbours which are isolated nodes in G_2; the remaining $2y - z$ must each have degree 1 in G_2. Since G_1 has no edges, each of the x edges of G_2 must be incident with a node in $N[D_2]$, either with one of the $2y - z$ nodes of degree 1 in G_2 which have a neighbour in $D_2 \cap (G - G_2)$, or with one of the $d_2 - y$ nodes in $D_2 \cap G_2$. Thus $x \leq (2y - z) + (d_2 - y)$ and so $d_2 - x + y - z \geq 0$ (2).

Each node in G_1 has three neighbours in $G - G_1$, which cannot be in D_4, D_3, or D_2. Since no node in G has two edges to G_1, there must be at least $3|G_1|$ nodes in $(G - G_1) - D_4 - D_3 - D_2$. Thus $3(n - 4d_4 - 3d_3 - 2d_2) \leq (4d_4 + 3d_3 + 2d_2) - d_4 - d_3 - d_2$ and so $7d_2 + 11d_3 + 15d_4 \geq 3n$ (3). Moreover, since G_2 has maximum degree 1, each node in G_1 can have at most one neighbour in G_2 and so must have at least two neighbours in $G - G_2$, and these cannot be in D_4, D_3, or $D_2 \cap (G - G_2)$. Thus $2(n - 4d_4 - 3d_3 - 2d_2) \leq (4d_4 + 3d_3) - d_4 - d_3 - y$ and so $4d_2 + 8d_3 + 11d_4 - y \geq 2n$ (4).

Finally, consider the edges between $N(D_4)$ and $G_2 - G_1$. Since G_2 has $2x$ nodes with two edges to $N(D_4) \cup N(D_3)$, and $n - 4d_4 - 3d_3 - 2x$ nodes with three edges to $N(D_4) \cup N(D_3)$, there are $2(2x) + 3(n - 4d_4 - 3d_3 - 2x) = 3n - 12d_4 - 9d_3 - 2x$ edges between G_2 and $N(D_4) \cup N(D_3)$. Each node in $N[D_3]$ has degree at most 2 in G_3 and so must have at least one edge to $N(D_4)$. So each of the $2d_3$ nodes of $N(D_3)$ has one edge to D_3 and at least one edge to $N(D_4)$, leaving altogether at most $2d_3$ edges between $N(D_3)$ and G_2. Thus there are at least $(3n - 12d_4 - 9d_3 - 2x) - 2d_3 = 3n - 12d_4 - 11d_3 - 2x$ edges between $N(D_4)$ and G_2. Since there are $3d_4$ nodes in $N(D_4)$ and each has at most two edges to G_2, by the Pigeonhole Principle at least $(3n - 12d_4 - 11d_3 - 2x) - 3d_4 = 3n - 15d_4 - 11d_3 - 2x$ of the nodes in $N(D_4)$ must have two edges to G_2. But no node in G can have two edges to G_1, so at least $3n - 15d_4 - 11d_3 - 2x$ nodes in $N(D_4)$ must have a neighbour in $G_2 - G_1$. Now consider the nodes in $G_2 - G_1 = N[D_2] \cap G_2$. Each of the y nodes in $D_2 \cap (G - G_2)$ has two neighbours in $N(D_2)$, so these y nodes cannot be in $N(D_3)$ and thus must be in $N(D_4)$. Moreover, of the $2y$ neighbours of $D_2 \cap (G - G_2)$ in $N(D_2)$, z are isolated in G_2 so have at most two more neighbours in $N(D_4)$, and $2y - z$ have one neighbour in G_2 so at most one more neighbour in $N(D_4)$. Each of the $d_2 - y$ nodes in $D_2 \cap G_2$ has one neighbour in G_2, so these $d_2 - y$ nodes and their $d_2 - y$ neighbours in G_2 each have at most two neighbours in $N(D_4)$. Thus the number of neighbours in $N(D_4)$ of nodes in $G_2 - G_1$ is at most $y + 2z + (2y - z) + 2(d_2 - y) + 2(d_2 - y) = 4d_2 - y + z$. Thus $3n - 15d_4 - 11d_3 - 2x \leq 4d_2 - y + z$ and so $4d_2 + 11d_3 + 15d_4 + 2x - y + z \geq 3n$ (5).

Now we want the maximum value of $d_g = d_1 + d_2 + d_3 + d_4 = (n - 4d_4 - 3d_3 - 2d_2) + d_2 + d_3 + d_4 = n - d_2 - 2d_3 - 3d_4$; alternatively we want the minimum value of $n - d_g = d_2 + 2d_3 + 3d_4$. So putting together (1) to (5), we have the following linear program:

minimize $d_2 + 2d_3 + 3d_4$ such that
$5d_3 + 14d_4 \geq 2n$
$d_2 - x + y - z \geq 0$
$7d_2 + 11d_3 + 15d_4 \geq 3n$
$4d_2 + 8d_3 + 11d_4 - y \geq 2n$
$4d_2 + 11d_3 + 15d_4 + 2x - y + z \geq 3n$
with $d_2, d_3, d_4, x, y, z \geq 0$.

Solving this gives $n - d_g = \frac{5}{9}n$ with $d_2 = \frac{5}{54}n, d_3 = \frac{1}{27}n, d_4 = \frac{7}{54}n$ $x = \frac{5}{27}n, y = \frac{5}{54}n, z = 0$ (and so $d_1 = \frac{5}{27}n$). This can be verified using the fundamental duality theorem for linear programming, by checking that $(\frac{1}{27}, \frac{4}{27}, \frac{1}{27}, \frac{2}{27}, \frac{2}{27})$ is a feasible solution to the dual linear program, with objective value $\frac{5}{9}n$. Thus $d_g \leq n - \frac{5}{9}n = \frac{4}{9}n$. □

The bound in Theorem 1 is tight. Consider an 8-cycle *abcdefgh* with additional edge *ae*, and a 6-cycle *tuvwxy* with additional node z and edges tz, wz. For any $k \geq 1$ take $5k$ copies of the modified 8-cycles numbered from 0 to $5k - 1$, and

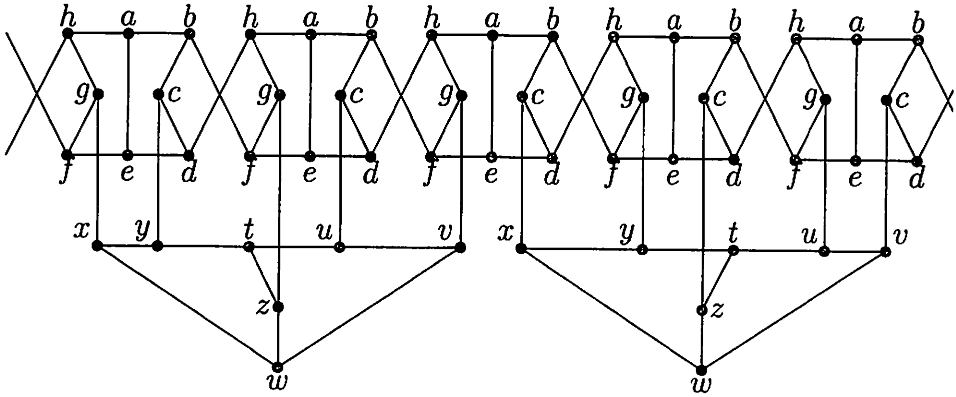

Figure 1: Constructing a class of graphs for which $d_g = \frac{4}{9}n$.

$2k$ copies of the modified 6-cycles. Add edges bf and dh between 8-cycles i and $i+1 \pmod{5k}$ for $i = 0, 1, ..., 5k-1$ and join each of the $2k$ copies of u, v, x, y, z to a different one of the $5k$ copies of c, g. The result is a cubic graph of girth 5 with $n = 5k(8) + 2k(7) = 54k$ (see Figure 1).

Number the nodes of this graph from 1 to n in the following order: all a's, all t's, all w's, all e's, all c's, all g's, and then the rest in any order. The first pass through the greedy algorithm will put the a's and t's into D_4, the second pass will put the w's into D_3, the third pass will put the e's into D_2, and the final pass will put the c's and g's into D_1, resulting in a dominating set with $(5k+2k)+(2k)+(5k)+(5k+5k) = 24k$ nodes, or $\frac{4}{9}n$. Note that the domination number is considerably lower. Let T consist of all t's and w's; all b's and f's from evenly numbered 8-cycles; and all d's and h's from oddly numbered 8-cycles. If k is even then T is a dominating set of size $14k$; if k is odd then adding d from the last 8-cycle makes T a dominating set of size $14k+1$. Since these graphs are cubic, at least $\frac{n}{4}$ nodes are needed for any dominating set, so we have $13.5k \leq \gamma \leq 14k+1$ for all k.

The result for cubic graphs can be partially extended to r-regular graphs. For a graph G of maximum degree r, replace

4, 3, 2, 1 in the greedy algorithm for maximum degree 3 with $r+1, r, ..., 2, 1$.

Theorem 2 *For an r-regular graph G with $r \geq 3$, $d_g \leq \frac{r^2+4r+1}{(2r+1)^2}n$.*

Proof. Extend the notation from Theorem 1 to $k = r+1, r, ..., 2, 1$. The same argument as in Theorem 1 gives $n - (r+1)d_{r+1}$ nodes and at most $e - r^2 d_{r+1}$ edges in G_r, and $n - (r+1)d_{r+1} - rd_r$ nodes and at most $e - r^2 d_{r+1} - (r-1)^2 d_r$ edges in G_{r-1}. For $k = r-1, r-2, ..., 3$, there are kd_k nodes in $N[D_k] \cap G_k$ each of degree at most $k-1$ in G_k, so at most $k(k-1)d_k$ edges within $N[D_k] \cap G_k$ and between $N[D_k] \cap G_k$ and G_{k-1}. Thus G_2 has $n - (r+1)d_{r+1} - rd_r - ... - 3d_3$ nodes and $x \geq e - r^2 d_{r+1} - (r-1)^2 d_r - (r-1)(r-2)d_{r-1} - ... - k(k-1)d_k - ... - 6d_3$ edges. G_2 has maximum degree 1, so $2x \leq n - (r+1)d_{r+1} - rd_r - ... - 3d_3$. Since G is r-regular, $2e = rn$, and thus $9d_3 + 20d_4 + ... + (2k-3)kd_k + ... + (2r-5)(r-1)d_{r-1} + (2r-1)(r-2)d_r + (2r+1)(r-1)d_{r+1} \geq (r-1)n$.

Each of the $n-(r+1)d_{r+1}-rd_r-...-3d_3-2d_2$ nodes in G_1 has r neighbours in $G - G_1$, which cannot be in $D_{r+1}, D_r, ..., D_3, D_2$. Since no node in G has two neighbours in G_1, there must be at least $r|G_1|$ nodes in $(G - G_1) - D_{r+1} - D_r - ... - D_3 - D_2$. Thus $r[n-(r+1)d_{r+1}-...-2d_2] \leq [(r+1)d_{r+1}+...+2d_2]-d_{r+1}-...-d_2$ and so $(2r+1)d_2 + (3r+2)d_3 + ... + (kr+k-1)d_k + ... + (r^2+2r)d_{r+1} \geq rn$. As in Theorem 1, we get the following linear program:

minimize $n - d_g = d_2 + 2d_3 + ... + rd_{r+1}$ such that
$9d_3 + 20d_4 + ... + (2k-3)kd_k + ... + (2r-5)(r-1)d_{r-1}$
$+(2r-1)(r-2)d_r + (2r+1)(r-1)d_{r+1} \geq (r-1)n$
$(2r+1)d_2+(3r+2)d_3+...+(kr+k-1)d_k+...+(r^2+2r)d_{r+1} \geq rn$
with $d_2, d_3, ..., d_{r+1} \geq 0$.

Solving this gives $n - d_g = \frac{3r^2}{(2r+1)^2}n$ with $d_2 = \frac{r(r-1)}{(2r+1)^2}n, d_3 = d_4 = ... = d_r = 0, d_{r+1} = \frac{1}{2r+1}n$. This can be verified using the fundamental duality theorem for linear programming, by checking that $(\frac{r}{(2r+1)^2}, \frac{1}{2r+1})$ is a feasible solution to the dual linear program, with objective value $\frac{3r^2}{(2r+1)^2}n$. Thus $d_g \leq n - \frac{3r^2}{(2r+1)^2}n = \frac{r^2+4r+1}{(2r+1)^2}n$. □

This bound is not tight. For $r = 3$ (G cubic), Theorem 2 gives $d_g \leq \frac{22}{49}n$, which is about $.0045n$ greater than the bound in Theorem 1.

References

1. V. Chvatal, A greedy heuristic for the set covering problem, *Math. Oper. Res* **4** (1979) 223-235.
2. D. Fisher, K. Fraughnaugh, S. Seager, Domination of graphs with maximum degree three, *Proceedings of the Eighth Quadrennial International Conference on Graph Theory, Combinatorics, Algorithms and Applications*, vol I (1998) 411-421.
3. D. Fisher, K. Fraughnaugh, S. Seager, The domination number of cubic graphs of larger girth, to appear in *Proceedings of the Ninth Quadrennial International Conference on Graph Theory, Combinatorics, Algorithms and Applications.*
4. D.S. Johnson, Approximating algorithms for combinatorial problems, *J. Comput. System Sci.* **9** (1974) 256-278.
5. L. Lovász, On the ratio of optimal integral and fractional covers, *Discrete Math.* **13** (1975) 383-390.
6. A.K. Parekh, Analysis of a greedy heuristic for finding small dominating sets in graphs, *Information Processing Letters* **39** (1991) 237-240.
7. B. Reed, Paths, Stars, and the Number Three, *Combin. Probab. Comput.*, **5** (1996) 277-295.