The Hub Number of a Graph

Tracy Grauman[*], Stephen G. Hartke[†], Adam Jobson[‡], Bill Kinnersley[§], Douglas B. West[¶], Lesley Wiglesworth[‖], Pratik Worah[**], Hehui Wu[††]

February 20, 2008; revised June 2, 2008

Abstract

A *hub set* in a graph G is a set $U \subseteq V(G)$ such that any two vertices outside U are connected by a path whose internal vertices lie in U. We prove that $h(G) \le h_c(G) \le \gamma_c(G) \le h(G)+1$, where $h(G), h_c(G), \gamma_c(G)$ respectively are the minimum sizes of a hub set in G, a hub set inducing a connected subgraph, and a connected dominating set. Furthermore, all graphs with $\gamma_c(G) > h_c(G) \ge 4$ are obtained by substituting graphs into three consecutive vertices of a cycle; this yields a polynomial-time algorithm to check whether $h_c(G) = \gamma_c(G)$.

Keywords: graph algorithms, complexity, hub number, connected domination

1 Introduction

Introduced by Walsh [3], a *hub set* in a graph G is a set U of vertices in G such that any two vertices outside U are connected by a path whose internal vertices lie in U. Adjacent vertices

[*]Department of Computer Science, University of Illinois, grauman2@uiuc.edu.

[†]Department of Mathematics, University of Nebraska, shartke2@math.unl.edu. Research partially supported by a Maude Hammond Fling Faculty Research Fellowship from the University of Nebraska Research Council.

[‡]Department of Mathematics, University of Louisville, asjobs01@louisville.edu.

[§]Department of Mathematics, University of Illinois, wkinner2@uiuc.edu.

[¶]Department of Mathematics, University of Illinois, west@math.uiuc.edu. Research partially supported by the National Security Agency under Award No. H98230-06-1-0065.

[‖]Department of Mathematics, University of Louisville, l0well01@louisville.edu.

[**]Department of Computer Science, University of Illinois, pworah2@uiuc.edu. Research partially supported by NSF grant DMS-0528086.

[††]Department of Mathematics, University of Illinois, hehuiwu2@uiuc.edu.

are joined by a path with *no* internal vertices, so the condition holds vacuously for them.

The *hub number* of G, denoted $h(G)$, is the minimum size of a hub set in G. A *connected set* in G is a vertex set S such that the subgraph of G induced by S (denoted $G[S]$) is connected. The *connected hub number* of G, denoted $h_c(G)$, is the minimum size of a connected hub set in G. Various related notions of connection, including these, were studied for integer lattices by Hamburger, Vandell, and Walsh [2]. Walsh [3] studied the hub number for several classes of graphs and showed that the hub number is at least the girth minus 3 (the *girth* is the length of the shortest cycle).

Placing transmitters at the vertices of a hub set would enable communication among all vertices outside the set; this motivates seeking a small hub set. The same idea motivates studying the connected domination number of a graph. We show in this note that these problems are almost the same.

We review the well-known definitions for domination. A *dominating set* in a graph G is a set S of vertices such that every vertex in G outside S has a neighbor in S. The *domination number* of G, denoted $\gamma(G)$, is the minimum size of a dominating set in G. The *connected domination number*, denoted $\gamma_c(G)$, is the minimum size of a connected dominating set. Connected dominating sets and connected hub sets exist only in connected graphs.

A hub set S need not be a dominating set; in particular, there may be an undominated vertex v whose neighborhood is $V(G) - S - \{v\}$. On the other hand, in a connected graph, every connected hub set is a hub set, and every connected dominating set is a connected hub set. Thus $h(G) \le h_c(G) \le \gamma_c(G)$. We prove that also $\gamma_c(G) \le h(G) + 1$, obtained independently and concurrently by multiple subsets of the listed authors. Although no two of the parameters are the same, the connected hub number and connected domination number are almost always equal. We describe the structure of graphs G such that $\gamma_c(G) > h_c(G) \ge 4$ and use this to give a polynomial-time algorithm for determining whether $h_c(G) = \gamma_c(G)$.

2 Near-Equality of the Parameters

Theorem 2.1. *For any connected graph G,*

$$h(G) \le h_c(G) \le \gamma_c(G) \le h(G) + 1$$

Proof. It remains to prove that $\gamma_c(G) \le h(G) + 1$. Let U be a smallest hub set in G. We will construct a connected dominating set U' of size at most $|U| + 1$.

Since U is a smallest hub set, $U \ne V(G)$. Hence we may choose a vertex v outside U with a neighbor in U. Initialize $U' = \{v\}$. Let $C_1, \ldots, C_k$ be the components of $G[U]$ that contain no neighbor of v, and let D be the set of all vertices in other components of $G[U]$.

Add all of D to U'. From each C_i, choose a vertex u_i having a neighbor v_i outside C_i; note that $v_i \neq v$. Let w_i be a vertex of C_i farthest from u_i in C_i; note that $C_i - w_i$ is connected (or empty). The set $V(C_i) - \{w_i\} \cup \{v_i\}$ is connected; add it to U'.

After starting with v, we added to U' vertex sets having the same size as each component of $G[U]$. Thus $|U'| \leq |U| + 1$ (the vertices of the form v_i need not be distinct).

It remains to show that U' is connected and dominating. Since $v \in U'$, both claims are proved by finding, for each vertex z, a U'-internal z, v-path (that is, a path from z to v whose internal vertices all lie in U'). By construction, such a path exists with $z \in D$. For $z \in V(G) - U$, there must be a U-internal z, v-path P, since U is a hub set. If z is not adjacent to v, then the internal vertices of P must lie in D, and $D \subseteq U'$.

Finally, consider $z \in V(C_i)$. The previous case found a U'-internal v_i, v-path. Since $v_i \in U'$, it suffices to find a U'-internal z, v_i-path. Indeed, for each such z there is a z, v_i-path whose internal vertices all lie in $V(C_i) - \{w_i\}$, and these vertices lie in U'. □

For the path P_n with n vertices, the three parameters have the same value: $h(P_n) = h_c(P_n) = \gamma_c(P_n) = n - 2$. On the other hand, the 3-dimensional cube Q_3 has a hub set of size 3, but its connected hub and connected domination numbers equal 4. Finally, note that $h(C_n) = h_c(C_n) = n - 3 = \gamma_c(C_n) - 1$.

For another example with γ_c exceeding the others, let T be a tree of diameter 3, with central edge uv. (The *diameter* of a graph is the maximum distance between vertices; a tree of odd diameter has a *central edge*, which is the shared central edge on all longest paths.) Form G from the disjoint union of K_r and T by making each vertex of K_r adjacent to each leaf in T. Now $\{u, v\}$ is a connected hub set of size 2 (all nonadjacent pairs outside $\{u, v\}$ are pairs of leaves of T, joined by paths through $\{u, v\}$). However, G has no connected dominating set of size 2. We generalize this construction in the next section.

3 Distinguishing h_c and γ_c

We introduce several additional concepts that will aid in characterizing the graphs where $\gamma_c > h_c \geq 4$. We use $N(x)$ for the set of vertices adjacent to x in G. Also $G - v$ denotes the graph obtained from a graph G by deleting a vertex v.

Definition 3.1. Let G and H be graphs with disjoint vertex sets. A graph G' is obtained from G by *substituting* H for v in G if G' is obtained from the disjoint union $(G - v) + H$ by making every neighbor of v in G adjacent to every vertex of $V(H)$. A *swollen k-cycle* is a graph obtained from the cycle C_k with vertices $x_1, \ldots, x_k$ in order by substituting a complete graph for x_2 and substituting any graphs for x_1 and x_3.

The graphs constructed in the last example of Section 2 are swollen 5-cycles. A *thread* in a graph G is a path whose internal vertices have degree 2 in G. In a swollen k-cycle with vertices indexed as in Definition 3.1, the path with vertices $x_4, \dots, x_k$ is a thread.

Lemma 3.2. *If G is a swollen k-cycle, for $k \ge 4$, then $h_c(G) = k - 3 = \gamma_c(G) - 1$, and the thread formed by the $k - 3$ unsubstituted vertices is a connected hub set.*

Proof. Let P be the thread of unsubstituted vertices in G, indexed as $x_4, \dots, x_k$ in order. Let F and F' be the graphs substituted for x_1 and x_3, respectively, and let Q be the complete graph substituted for x_2. Since the only nonadjacent vertices outside P lie in $F \cup F'$, the path P is a connected hub set of size $k-3$. To complete the proof, it suffices by Theorem 2.1 to show that $\gamma_c(G) > k - 3$.

Let S be a connected dominating set in G. Let V_i be the set of vertices in G arising from vertex x_i of the original cycle in the construction of G. Since S is connected, the sets in $\{V_1, \dots, V_k\}$ that S intersects must be consecutive (cyclically). If three sets are missed, then the vertices in the middle set are not dominated. Hence $|S| \ge k - 2$. □

Other examples arise for graphs with very small hub sets. Note that $h_c(K_n) = 0$. When $h_c(G) = 1$ and $\gamma_c(G) = 2$, let x be a connected hub set. The vertices outside $N(x)$ must form a nonempty clique, and each must be adjacent to all of $N(x)$. Any edges can be added within $N(x)$ as long as no dominating vertex is created. Since $N(x)$ may be connected, G need not be a swollen 4-cycle. Similar examples occur when $h_c(G)$ is 2 or 3, with G differing from a swollen $(h_c(G) + 3)$-cycle by having additional edges in the neighborhood of the endpoints of the thread. For $h_c(G) > 3$, there is no such flexibility.

Theorem 3.3. *If G is a connected graph with $\gamma_c(G) > h_c(G) = r \ge 4$, then G is a swollen $(r + 3)$-cycle.*

Proof. Let U be a smallest connected hub set, and let $H = G[U]$. We show first that H is a thread. Let $W = N(U) - U$ and $Q = G - (U \cup W)$. Since $h_c(G) < \gamma_c(G)$, the set U does not dominate G, and hence $V(Q) \ne \varnothing$. Since U is a hub set, Q is a complete subgraph, and its vertices are adjacent to all of W. Since G is connected, $W \ne \varnothing$. Choose $v \in V(H)$ having a neighbor $w \in W$. Choose any spanning tree T of H rooted at v, let S be the set of (non-root) leaves of T, and let y be a vertex of Q (see Figure 1). Now $(U - S) \cup \{w, y\}$ is a connected dominating set of G. Since $\gamma_c(G) > |U|$, we have $|S| \le 1$. Since this holds for every choice of T, H must be a path with v at one end. Since this also holds for every choice of v having a neighbor outside H, in fact H is a thread. Let u be the other endpoint of H.

If u has no neighbor in W, then $U - \{u\}$ plus any vertex of W forms a connected dominating set of size r. Hence $N(u) \cap W \ne \varnothing$. At this point, G is a swollen $(r + 3)$-cycle unless u and v have a common neighbor or some edge has endpoints in $N(u) \cap W$ and $N(v) \cap W$. Let Z consist of those endpoints or of a vertex in $N(u) \cap N(v)$. In either case, let Y be a set of two adjacent internal vertices of H; this exists since $r \geq 4$. Now $(U \cup Z) - Y$ is a connected dominating set with size at most r, which is a contradiction. We conclude that G is a swollen $(r+3)$-cycle. □

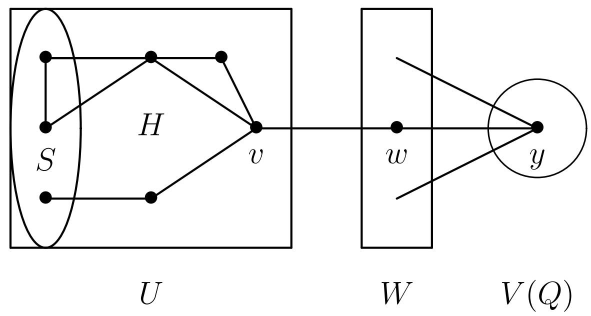

Figure 1: A dominating set that is too small

Theorem 3.4. *Given a graph G, there exists an algorithm to decide, in polynomial time, whether or not $h_c(G) = \gamma_c(G)$.*

Proof. By checking all sets of size at most 3, we may compute $h_c(G)$ if $h_c(G) < 4$. In this case, we can also check sets of size at most 3 to test whether $\gamma_c(G) = h_c(G)$. If $h_c(G) \geq 4$, then it suffices to determine whether G is a swollen cycle. For each edge e of G, we can find the longest thread containing e, find the neighborhoods of the endpoints u and v, check whether those neighborhoods (outside the thread) are disjoint and not joined by any edges, and check whether the remaining set of vertices is a nonempty clique whose vertices are all adjacent to the neighbors of u and v that are not in the thread. If these properties do not all hold for some edge, then $\gamma_c(G) = h_c(G)$, by Theorem 3.3. □

This theorem yields, as a corollary, an immediate proof of a complexity result stronger than that found in [3],

Corollary 3.5. *Approximating the hub number or the connected hub number within a factor of $\ln D$ is NP-hard even on the family of bipartite graphs with maximum degree at most D.*

Proof. Chlebík and Chlebíková [1] showed that approximating the connected domination number within a factor $\ln D - c \ln \ln D$ is NP-hard on this class. Computing the hub number or connected hub number approximates it within 1. □

References

[1] M. Chlebík and J. Chlebíková, Approximation hardness of dominating set problems. In *Algorithms—ESA 2004, Lecture Notes in Comput. Sci.* 3221 (Springer, Berlin, 2004), 192–203.

[2] P. Hamburger, R. Vandell, and M. Walsh, Routing sets in the integer lattice, *Discrete Appl. Math.* 155 (2007), 1384–1394.

[3] M. Walsh, The hub number of a graph, *Intl. J. Mathematics and Computer Science* 1 (2006), 117–124.