International Journal of Mathematics
and Computer Science, 1(2006), 117-124

The hub number of a graph

Matthew Walsh

Department of Mathematical Sciences
Indiana-Purdue University
Fort Wayne, Indiana 46805, USA
e-mail: walshm@ipfw.edu

(Received August 24, 2005, Accepted November 2, 2005)

Abstract

Let G be a graph; a *hub set* H of G is a set of vertices with the property that for any pair of vertices outside of H, there is a path between them with all intermediate vertices in H. The *hub number* $h(G)$ is then defined to be the size of a smallest hub set of G. In this paper the hub number for several classes of graphs is computed; bounds in terms of other graph parameters are also determined.

1 Hub sets in graphs

In what follows all graphs may be assumed to be simple; for notation and standard definitions see [2].

Imagine that we have a graph G which represents the buildings in a large industrial complex, with an edge between two buildings if it is an easy walk from one to the other. The corporation is considering implementing a rapid-transit system, and wants to place its stations in buildings (which will then be used only for this purpose) so that to travel between two non-adjacent buildings (which are not stations), one need only walk to an adjacent station, take the RTS, and walk to the desired building. The corporation would like to implement this plan as cheaply as possible, which involves converting as few buildings as possible into transit stations.

Suppose that $S \subseteq V(G)$ and let $x, y \in V(G)$. An *S-path* between x and y is a path where all intermediate vertices are from S. (This includes the degenerate cases where the path consists of the single edge xy or a single vertex x if $x = y$; call such an S-path *trivial*.) A set $S \subseteq V(G)$ is a *hub set* of G if it has the property that, for any $x, y \in V(G) - S$, there is an S-path in G between x and y. The problem in the previous paragraph can then be rephrased: *what is the smallest size of a hub set in G?* We shall call this the *hub number* of G, and denote it by $h(G)$. It is clear that $h(G)$ is well-defined for any G, since $V(G)$ is a hub set. In all situations of interest, we will assume G to be connected; if G is a disconnected graph then any hub set must contain all of the vertices in all but one of the components, as well as a hub set in the remaining component.

Note that if the hypothetical corporation in question wanted to find a less draconian solution that permitted buildings to remain in use for other purposes,

---

Key words and phrases: vertex contraction, connectivity, routing.
AMS (MOS) Subject Classifications: Primary 05C40, Secondary 05C69
ISSN 1814-0424

then they would probably add the proviso that the graph induced by their hub set should be connected; this way it would be possible for someone working in one of the station-buildings to get to any other building in the complex while leaving the RTS at most once. Such a set will be called, naturally enough, a *connected hub set*, and the size of a smallest connected hub set will be the *connected hub number* $h_c(G)$ of the graph G; this is well-defined if G is connected, since again $V(G)$ is a connected hub set. (We can define this equally well as follows: S is a connected hub set iff there as an S-path between any two vertices x, y if either $x, y \in S$ or $x, y \in V - S$.

It is obvious that $h_c(G) \geq h(G)$, since every connected hub set is a hub set. Equality is not always achieved; as an example, the hub number of the 3-cube $h(Q_3) = 3$ but $h_c(Q_3) = 4$.

2 Vertex contraction

To determine the hub number of a graph G, the graph operation introduced in this section will be useful. Let x be a vertex in G; define the *contraction of x in G* (denoted by G/x) as being the graph obtained by deleting x and putting a clique on the (open) neighbourhood of x. (Note that this operation does not create parallel edges; if two neighbours of x are already adjacent, then they remain simply adjacent.)

We call this operation *vertex contraction* by analogy with edge contraction. In fact, the two operations are related in a stronger sense than the metaphorical.

Theorem 2.1. *Let G be a graph with an edge x, and let $\mathcal{L}(G)$ denote the line graph of G. Then $\mathcal{L}(G/x) = \mathcal{L}(G)/x$.*

Proof. Two vertices a, b in $\mathcal{L}(G/x)$ are adjacent precisely when either ab or axb is a path in G. Those vertices of the first type will also be adjacent in $\mathcal{L}(G)$, and hence in $\mathcal{L}(G)/x$ since neither a nor b is equal to x, and vertex contraction only deletes edges adjacent to the vertex being contracted.

If ab is not an edge in $\mathcal{L}(G)$ but axb was a path in G, then axb will still be a path in the line graph. Then if we contract the vertex x, any two vertices adjacent to x will become adjacent to each other in the contraction; specifically, ab will be an edge in $\mathcal{L}(G)/x$.

This takes into account the only two ways that a pair of vertices in $\mathcal{L}(G)/x$ can be adjacent; therefore, the two graphs must be identical. $\square$

Recall that we can contract a set of edges in arbitrary order, and the result will always be the same. A similar fact holds true for vertex contraction.

Theorem 2.2. *Let G be a graph containing vertices u, v. Then $(G/u)/v = (G/v)/u$.*

The proof is straightforward and tedious, and for those reasons it is omitted.

Note that this implies that we can talk about G/S, the contraction of a set of vertices in G, without having to specify an ordering on that set.

Theorem 2.3. *Let G be a connected graph, and denote by $d(G)$ the diameter of G. If x is a vertex in G, then $d(G) - 1 \leq d(G/x) \leq d(G)$.*

Proof. Let P be a maximum shortest path in G; either P passes through x or it does not. In the former case, let u and v be the neighbours of x in P; in G/x, u will be adjacent to v, so P will have its length reduced by 1. (Note that x can only have two neighbours in P if x is on P, since otherwise we could find a shorter path that connects the endpoints of P.) Similarly, if x is in fact one of the end vertices of P, then the length of P will be reduced by 1 in G/x.

So suppose that x is not in P. It is clear that any two neighbours of x contained in P can be at most distance 2 apart, since otherwise P is not a shortest path between its endpoints.

From this observation it is easy to see that contracting x will either leave P as it was, or produce a new path between P's end vertices of length one less than the length of P.

Further, it is easy to see that any shortest path in G/x is either a shortest path in G or arises from a path in G of length one greater. (It is worth noting for future reference that in the latter case we can arrange for x to be on the aforementioned path in G, which may be of length either equal to the distance in G between the end vertices of the path, or one greater than that. This observation will be used in the proof of Theorem 4.6) The conclusion of the theorem now follows. □

Theorem 2.4. *Let $g(G)$ denote the girth of G. If $g(G) \geq 4$, then for any $x \in V(G)$, $g(G/x) \leq g(G)$.*

Proof. First note that if $g(G) = \infty$ (that is, if G is acyclic), then $g(G/x)$ is either ∞ (if $\deg(x) \leq 2$) or 3. So let us assume that G is a triangle-free graph which contains at least one cycle. Choose a vertex $x \in V(G)$; one of three things must be true.

Case 1: $\deg(x) \geq 3$. Then $N(x)$ will be a clique of size at least 3 in G/x, and so contain a triangle; since G was triangle-free, the girth is therefore lower.

Case 2: $\deg(x) = 2$ *and x is in a cycle.* Let C be the shortest cycle containing x. Then the image of C in G/x will be a cycle shorter than C by one vertex. So if $|C| = g(G)$ then $g(G/x) = g(G) - 1$; otherwise the girth will remain unchanged.

Case 3: $\deg(x) \leq 2$ *and no cycle of G contains x.* Clearly the girth of x remains unchanged in this case.

Since the girth is never increased in any of the above cases, the result follows. □

3 Computing the hub number

The relationship between vertex contraction and the hub number of a graph is as follows.

Theorem 3.1. *Let S be a subset of $V(G)$. Then G/S is complete if and only if S is a hub set of G.*

Proof. Let S be a hub set. Choose $x, y \in V(G) - S$, and suppose that they are not identical. Then if they are not adjacent, there must be a path in G between x and y with all intermediate vertices in S. This path will now be contracted to a single edge, hence x and y are adjacent in G/S. If they were already adjacent, of course, then they remain adjacent in the contraction.

Now suppose that G/S is a complete graph, and choose $x, y \in V(G/S)$. Then x and y are vertices in the original graph G. Suppose that they are neither adjacent nor identical; then since they are adjacent in G/S, there must be a path between them in G that was contracted into an edge. But then such a path can only go through vertices from S. Since our choice of x, y was arbitrary, S must be a hub set of G. □

Corollary 3.2. *G is a complete graph if and only if $h(G) = h_c(G) = 0$.*

Corollary 3.3. *Let G be a graph with hub set S, and let $x \in S$. Then $S - x$ is a hub set in G/x.*

Theorem 3.4. *$G = (V, E)$ is a graph with $h(G) = h_c(G) = 1$ if and only if G has the following structure:*

- *G is not a complete graph.*
- *$V(G) = A \dot{\cup} B \dot{\cup} \{x\}$, where $\{x\}$ is a hub set.*
- *x is adjacent to every vertex in A and no vertex in B.*
- *Every vertex in A is adjacent to every vertex in B.*
- *$G[B]$ is a complete graph.*

Proof. The reverse direction is clear. Assume that $h(G) = 1$; then the first three points are obvious: we must have some lone vertex which is a hub set, so call it x and call its sets of neighbours and non-neighbours A and B, respectively. Since the adjacencies of vertices in B are unaffected by the contraction of x, yet G/x is complete, then clearly any vertex in B must already be adjacent to everything (except x, by assumption.) □

Theorem 3.5. *Let $G = K_{k_1,k_2,\ldots,k_n}$ with $n \geq 2$ and $k_i > 1$ for some i. Then $h(G) = h_c(G) = 1$ if and only if some $k_i \leq 2$; otherwise $h(G) = h_c(G) = 2$.*

Proof. Note that if we have some part consisting of a single vertex, then that vertex is adjacent to every other vertex, and hence contracting that vertex yields a complete graph. Similarly, suppose that some part consists of the vertices u and v; then we can invoke the previous theorem with $x = u$, $B = \{v\}$, and $A = V(G) - \{u, v\}$.

If all part sizes are at least 3, however, then we cannot contract a single vertex to get a complete graph. We can, however, contract two vertices in different parts, so $h(G) = 2$ in this case. (And the hub set we found is connected, so $h_c(G) = 2$ also.) □

Theorem 3.6. $h(C_n) = h_c(C_n) = n - 3$.

Proof. Note that, for any vertex $x \in V(C_n)$, $C_n/x = C_{n-1}$ for $n \geq 4$. $\square$

Theorem 3.7. *Let T be a tree with n vertices and l leaves. Then $h(G) = h_c(G) = n - l$.*

Proof. It is clear that the set N of all non-leaf vertices in T forms a hub set, since the unique path between any two leaves never passes through another leaf. Note that any proper subset of N cannot be a hub set, since every non-leaf vertex is a cut-vertex, and therefore separates a pair of leaves. To complete the proof, we will show that we can find a minimum hub set which contains no leaf vertices.

Let S be a minimum hub set which contains a leaf vertex x. Since the set of non-leaf vertices in T form a hub set, if S is minimum it must exclude at least one vertex of degree 2 or greater; choose such a vertex y so that the path from x to y in T has all intermediate vertices in S. Then $S' = (S - x) \cup \{y\}$ is also a hub set, since any path through S between y and another vertex can now be extended to be a path through S' between x and that vertex.

Note that we removed a vertex of degree 1 without adding another; therefore, we can repeat this process to find a minimum hub set S containing no vertices of degree 1. However, we know that the only such hub set is the set N of all non-leaf vertices; therefore N must be minimum, and $h(G) = n - l$. Finally, note that $T[N]$ is connected, so $h_c(G) = h(G)$. $\square$

4 Bounding the hub number

There is a natural relationship between the two hub parameters and the more familiar domination and connected domination numbers, $\gamma(G)$ and $\gamma_c(G)$, for a graph G.

Lemma 4.1. *For any graph G, $\gamma(G) \leq h(G) + 1$.*

Proof. Let S be a hub set in G, and consider the set of vertices W which are not adjacent to anything in S (nor in S themselves). The only S-paths between vertices in W must therefore be trivial; since we know that there are S-paths between all pairs of vertices in $V - S$, this implies that $G[W]$ is complete. Therefore, $S \cup \{w\}$ must be a dominating set for any $w \in W$. $\square$

Lemma 4.2. *For any connected graph G, $h_c(G) \leq \gamma_c(G)$.*

Proof. If S is a connected dominating set, then for any $x, y \in V$ there is an S-path between them. Since this is a stronger condition that that for a connected hub set, any connected dominating set is also a connected hub set. $\square$

In addition, the earlier result about diameter and vertex contraction gives is a relationship between diameter and hub number.

Lemma 4.3. *Let $d(G)$ denote the diameter of G. Then $h(G) \geq d(G) - 1$, and the inequality is sharp.*

Proof. The diameter of a complete graph is 1. If S is a hub set, then G/S must have diameter 1; however, each single vertex contraction reduces the diameter by at most 1. Therefore, we require at least $d(G)-1$ contractions before reaching a complete graph, and thus $h(G) \geq d(G) - 1$.

To prove sharpness, note that the path on n vertices has diameter $d = n-1$ and 2 leaf vertices; by Theorem 3.7 its hub number is $n - 2 = d - 1$. □

Given the above lemma following from Theorem 2.3, it seems reasonable to assume that something similar follows from Theorem 2.4.

Conjecture 4.4. *For any connected non-tree graph G, $h(G) \geq g(G) - 3$.*

Our next major result is most easily shown using a technical lemma on hub sets and cut sets, noticed by Vandell [3] and the author.

Lemma 4.5. *Let G be a graph, C a cut-set in G which separates the vertex sets T and U, and S a hub set of G such that $S \cap C$ is empty. Then either $T \subseteq S$ or $U \subseteq S$.*

Proof. Suppose that we can find vertices $t \in T - S$ and $u \in U - S$. Then there must be an S-path between t and u. But all (t, u)-paths must intersect positively with C, which no S-path can do. □

Theorem 4.6. *Let G be a graph such that $h(G) = d(G) - 1$. Then $h_c(G) = h(G)$, and there is a minimum hub set which induces a path in G.*

Proof. We begin by noting that, from the proof of Theorem 2.3, if a and b are a pair of vertices of maximum distance in G and x is a vertex such that $d(G/x) < d(G)$, then there is a shortest (a, b)-path which goes through x. Using this fact recursively, if S is a hub set of size $d(G)-1$ then we can find a shortest path between a and b in G which contains every vertex in S. Let us call one such path P. If neither a nor b is in S, then $P - \{a, b\}$ is a connected hub set of size $h(G)$, and we have our result; so we shall assume for definiteness that $b \in S$.

Let x be the vertex in $P-S$ which is closest (in G) to b, and let $k = d_G(b, x)$. Consider the set $C = \{v \in V(G) | d(b, v) = k\}$; clearly C is disjoint from S, since $S \subset P$ and the only vertex in $C \cap P$ is x. Denote by A and B the sets of vertices of distance greater than and less than k, respectively, from b; C is a cut-set which separates A and B. From this, Lemma 4.5 tells us that either $A \subseteq S$ or $B \subseteq S$. The former is impossible, since we are given that there is a second vertex in $P - S$ which is farther from b than x (and is thus in A); therefore $B \subseteq S$, and B consists only of vertices from P.

If $|B| > 1$, then the vertex b provides no connectivity to anything outside of S, and the set $S^* = (S - b) \cup \{x\}$ is a (minimum) hub set. If $a \notin S^*$ then we are done, since S^* is connected; otherwise, we repeat the construction from the other direction.

Suppose then that $B = \{b\}$, so b and x are neighbours; let y be the vertex in $P - S$ furthest from b, and z the other neighbour of x in P. If $y \neq z$, then all of the vertices in C must be connected to y via S-paths, and hence must be adjacent to z. Hence we can construct S^* as above with no loss of connectivity.

If $y = z$, then for S to be a hub set we still require that all vertices in C be adjacent to z. Then we construct the set $S^\dagger = (S - b) \cup \{z\}$, which is a connected hub set. □

Theorem 4.7. *Let $\Delta(G)$ denote the maximum degree of G. Then if G is a connected graph $h(G) \leq |V(G)| - \Delta(G)$, and the inequality is sharp.*

Proof. We note first of all that, from Theorem 3.5, $h(K_{1,n}) = 1$ for $n \geq 2$, so equality in this case is obvious.

Now, let G be a connected graph with minimum hub set S, and choose $v \in S$ to be a vertex of highest degree in S. We shall show that if the contraction of v decreases the maximum degree of G, then either $S = \{v\}$ or we can find another minimum hub set S' which does not include v.

Suppose first that v is of maximum degree in G. Then any neighbour of v in G will have degree at least $\Delta(G) - 1$ in G/v. If any of these vertices was adjacent to something other than neighbours of v, then it will now have at least $\Delta(G)$ neighbours, and so the maximum degree has not decreased. Otherwise, by the connectedness of G the graph G/v must be complete, and thus $|S| = 1$.

Suppose now that v has less than maximum degree. Then the only way that the contraction of v could possibly lower the maximum degree of G is if, for any vertex u of maximum degree, v is adjacent to u and all of the neighbours of v are also adjacent to u. But then it is clear that the set $S' = (S \cup \{u\}) - \{v\}$ is also a hub set of G.

Therefore, we can find a minimum hub set such that, if we contract the vertices one at a time, we need only lower the maximum degree of the graph once, and that time only by one. Therefore $|V(G)| - h(G) \geq \Delta(G)$. □

We have seen easy computations for hub numbers in certain restricted classes of graphs; as with most domination-type parameters, however, the general computation is a hard problem.

Theorem 4.8. *Determining whether a graph G has a (connected) hub set of size k is $\mathcal{NP}$-complete.*

Proof. (Due to Wayne Goddard, [1], and the author) Recognition of a hub set can be done in polynomial time, since vertex contraction and clique recognition are both polynomial-time operations. To show $\mathcal{NP}$-completeness, we reduce from "domination": determining whether there is a dominating set in a graph of a given size k, a known $\mathcal{NP}$-complete problem.

Let $G = (V, E)$ be a graph, and define a graph $H = (V', E')$ where

$$\begin{aligned} V' =& V \times \{0,1\} \\ E' =& \{\{(x,0),(y,0)\} | x,y \in V, x \neq y\} \\ & \cup \{\{(x,0),(x,1)\} | x \in V\} \\ & \cup \{\{(x,0),(y,1)\} | \{x,y\} \in E\} \end{aligned}$$

For convenience, we shall define $V_0 = \{(v,0)|v \in V\}$ and $V_1 = \{(v,1)|v \in V\}$. Suppose that S is a hub set of H; we show that we can always find a hub set of size at most $|S|$ consisting entirely of vertices in V_0.

So let $(v,1) \in S$, and suppose that $(v,0) \notin S$; since $N[(v,1)] \subset N[(v,0)]$, we see that $S' = (S - (v,1)) \cup \{(v,0)\}$ is also a hub set: any path between two vertices which passes through $(v,1)$ can be rerouted to pass through $(v,0)$ instead, and the path from $(v,1)$ to other non-hub set vertices will work equally well from $(v,0)$.

Now suppose that both $(v,0)$ and $(v,1)$ are in S; again, any S-path that uses $(v,1)$ can be rerouted through $(v,0)$ if the latter is not used already, and if it is then $(v,1)$ can be omitted altogether. Therefore, in this case the set $S' = S - (v,1)$ is a hub set. We can use this pair of arguments repeatedly to find a hub set entirely contained in V_0.

Using this observation, we can see that if $k \leq |V|$, then there exists a hub set of H with size k that consists solely of vertices in V_0, since a hub set with extra vertices added is still a hub set. But the image of such a hub set in the original graph G will be a dominating set: by our observation, every vertex in G must be equal to or adjacent to something in the image of S. Therefore, if we can find a hub set of size k in a graph in polynomial time, then we can find a dominating set of size k in a graph in polynomial time. □

Acknowledgements

The author would like to thank P. D. Johnson Jr. and an anonymous reviewer for helpful comments on earlier drafts of this work.

References

[1] W. Goddard, personal correspondence to author, 2002.

[2] T. W. Haynes, S. T. Hedetniemi, P. J. Slater, Fundamentals of Domination in Graphs, Marcel Dekker, New York, 1998.

[3] R. Vandell, personal correspondence to author, 2004.

[4] M. P. Walsh, A Problem in Network Stability, doctoral dissertation, Department of Discrete and Statistical Sciences, Auburn University, 2002.