Neighborhood unions and extremal spanning trees

Evelyne Flandrin*
Tomáš Kaiser†‡
Roman Kužel†‡
Hao Li§¶
Zdeněk Ryjáček†‡

Abstract

We generalize a known sufficient condition for the traceability of a graph to a condition for the existence of a spanning tree with a bounded number of leaves. Both of the conditions involve neighborhood unions. Further, we present two results on spanning spiders (trees with a single branching vertex). We pose a number of open questions concerning extremal spanning trees.

1 Introduction

There are several well-known conditions ensuring that any sufficiently 'dense' graph is *traceable* (admits a Hamilton path). Viewing a Hamilton path as an 'extremal' spanning tree (one with only two leaves), one may ask for similar conditions ensuring the existence of a spanning tree with at most m leaves.

*L.R.I., UMR8623 CNRS–Université Paris-Sud, Bât. 490, Université Paris-Sud, 91405 Orsay cedex, France. E-mail: evelyne.flandrin@lri.fr.

†Department of Mathematics, University of West Bohemia, and Institute for Theoretical Computer Science (ITI), Charles University, Univerzitní 8, 306 14 Plzeň, Czech Republic. E-mail: {kaisert,rkuzel,ryjacek}@kma.zcu.cz.

‡Research supported by project 1M0545 and Research Plan MSM 4977751301 of the Czech Ministry of Education.

§L.R.I., UMR8623 CNRS–Université Paris-Sud, Bât. 490, Université Paris-Sud, 91405 Orsay cedex, France, and School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, China. E-mail: li@lri.fr.

¶Partially supported by NSFC 60373012.

An early result of this type by Las Vergnas [9] gives a degree condition that guarantees that any forest in G of limited size and with a limited number of leaves can be extended to a spanning tree of G whose number of leaves is also limited in an appropriate sense. Specifically, this result implies as a corollary that G has a spanning tree with at most m leaves provided that

$$\sigma_2(G) \geq n - m + 1$$

(we refer to Section 2 for the definition of the parameter $\sigma_k(G)$ and other notation).

An alternative way of generalizing traceability is to bound the number of *branching vertices* (vertices of degree at least 3) in a spanning tree, for a Hamilton path is just a tree with no branchings. Following [7], we call a spanning tree with at most one branching vertex a *spanning spider* and remark that the investigation of this sort of spanning trees was catalyzed by problems in the construction of optical networks. Yet another related constraint on spanning trees is an upper bound on the maximum degree. Sufficient conditions for the existence of extremal spanning trees of the above types have been studied e.g. in [1], [7], [8] or [12].

Gargano et al. [7] prove a sufficient condition for a graph G without an induced $K_{1,3}$ to admit a spanning tree with a bounded number of branching vertices. The result subsumes known conditions for the traceability of such a graph G from [10] and [11]:

Theorem 1 *[7] If a graph G with no induced $K_{1,3}$ satisfies $\sigma_{k+3}(G) \geq n - k - 2$, then G admits a spanning tree with at most k branching vertices.*

The following 'neighborhood union' condition for traceability is an easy consequence of a similar condition for hamiltonicity of 2-connected graphs from [2] (see also [5]):

Theorem 2 *Any connected graph G with n vertices and $N_2(G) > \frac{2}{3}(n-2)$ is traceable.*

The first result of the present paper is a generalization of this statement that applies to spanning trees with at most m leaves:

Theorem 3 *Let G be a connected graph with n vertices and let $m \geq 2$ be an integer. If*

$$N_m(G) > \frac{m}{m+1} \cdot (n-m),$$

then G has a spanning tree with at most m leaves.

A proof is given in Section 3. It is easy to see that the condition is sharp: just consider the graph $(m+1)K_k + K_1$, consisting of $m+1$ cliques of size $k+1$, all sharing a vertex and otherwise disjoint.

In Section 4, we turn to spanning spiders and give two sufficient conditions for the existence of a spanning spider centered at a prescribed vertex. The conditions are sharp (the first one up to an additive constant) and involve 'localized' versions of the parameter σ_k.

2 Notation

We shall deal with simple undirected graphs. We write $V(G)$ for the vertex set and $E(G)$ for the edge set of a graph G.

As usual, the neighborhood $N(X)$ of a set $X \subset V(G)$ is defined to be the set of all vertices with at least one neighbor in X. We write $N(v)$ for $N(\{v\})$. The degree of a vertex v is denoted by $d(v)$. Let $k \geq 1$ an integer. We define

$$\begin{aligned} \sigma_k(G) &= \min_I \sum_{v \in I} d(v), \\ N_k(G) &= \min_I |N(I)|, \end{aligned}$$

where in both cases, I ranges over sets of k independent vertices in G. Thus, $\sigma_1(G) = N_1(G)$ is the minimum degree of G. In general, we have $N_k(G) \leq \sigma_k(G)$.

Let T be a tree and $v, w \in V(T)$. The unique path between v and w in T will be denoted by $[v, w]$. We shall often make use of the following notion, illustrated in Fig. 1. The *predecessor* u^- of a vertex $u \in V(T) - \{v\}$ *relative to* v is the neighbor of u in $[v, u]$. (For brevity, the vertex v is not indicated by the notation, but it will be always clear from the context.) Intuitively, u^- is the vertex that is 'one step closer' to v than u is. If $U \subset V(T) - \{v\}$, we set $U^- = \{u^- : u \in U\}$.

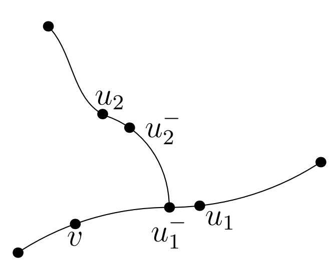


Figure 1: Predecessors of u_1 and u_2 relative to v.

We use the standard notation for paths. A path on vertices $x_1, \ldots, x_k$ is written as $x_1 \ldots x_k$. If x, y are vertices of a path P, then xPy denotes the subpath of P with endvertices x and y. The concatenation of two paths is represented by the concatenation of the corresponding sequences. For instance, the sequence $xPyzQw$ (where x, y, z, w are vertices and P, Q are paths) denotes the path that starts at x, follows P as far as y, uses the edge yz, and finally follows Q as far as w. (Of course, we are assuming here that the result of the concatenation is indeed a path, rather than a walk with self-intersections.)

3 Spanning trees with few leaves

We shall make use of the following well-known lemma. We include a proof for convenience.

Lemma 4 *For any graph G and $k \geq 1$,*

$$\frac{\sigma_{k+1}(G)}{k+1} \geq \frac{\sigma_k(G)}{k}.$$

Proof. Let $I \subset V(G)$ be an independent set of $k+1$ vertices whose degrees sum up to $\sigma_{k+1}(G)$, and let b be a vertex whose degree is maximal in I. For a set $X \subset V(G)$, let $a(X)$ denote the average degree of vertices in X. Clearly, $a(I-b) \leq a(I) = \sigma_{k+1}(G)/(k+1)$, while on the other hand, $a(I-b)$ is at least $\sigma_k(G)/k$ since $I-b$ is independent. The lemma follows. □

We can now proceed to the proof of the main result of this section.

Proof of Theorem 3. The result is well known for $m = 2$ (the case of a Hamilton path), so we may assume $m \geq 3$.

Let T be a tree in G with at most m leaves such that it spans as many vertices of G as possible, and (subject to this condition) it has the least possible number of leaves. We assume that T is not spanning, and choose a vertex $x_0 \notin V(T)$.

If T had fewer than m leaves, then we could extend it to some vertex in its neighborhood without making the number of leaves exceed m. We may thus assume that T has exactly m leaves $x_1, \ldots, x_m$.

We begin by noting that the set $X = \{x_0, \ldots, x_m\}$ is independent. Indeed, an edge between two vertices in X would allow us to either extend T to x_0, or to decrease the number of leaves of T, contradicting in both cases the extremal property of T.

We shall now prove, in several steps, the following estimate on the neighborhood sizes for the sets $X - x_k$:

$$|N(X - x_k)| \leq n - d(x_k) - m \tag{1}$$

for all $k = 0, \ldots, m$. The proof of (1) is given for $k \in \{0, 1\}$; observe that the remaining cases are analogous to the case $k = 1$ since the leaves $x_2, \ldots, x_m$ play a role symmetric to that of x_1.

The predecessor of a vertex $v \in V(T)$ was defined in Section 2. In this proof, all predecessors will be relative to the vertex x_1. Thus, v^- denotes the predecessor of a vertex $v \in V(T) - \{x_1\}$ relative to x_1. (It should be noted that when proving (1) for $k > 1$, one has to work with predecessors relative to x_k.)

Claim 1 *We have*

$$|N(x_k)^-| = |N(x_k)|$$

for $k = 0, 1$.

First, let $k = 0$. Since x_1 is not contained in $N(x_0)$, all we need to show is that the mapping $v \mapsto v^-$ is injective. Assuming $v^- = w^-$, extend T to cover x_0 by replacing the edge vv^- by vx_0 and wx_0. The resulting tree has m leaves and spans more vertices.

It remains to prove the claim for $k = 1$. For every neighbor v of x_1, the predecessor v^- must have degree at most 2 in T. Otherwise, the tree obtained by replacing vv^- with vx_1 has fewer leaves than T. The injective property of the mapping $v \mapsto v^-$ follows.

Before proceeding to the next claim, define a vertex $v \in N(x_k)$ to be *minimal* if v is contained in the path $[x_1, w]$ for all $w \in N(x_k)$. (See Fig. 2 for an illustration.)

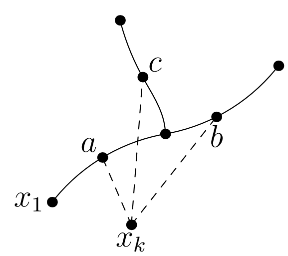
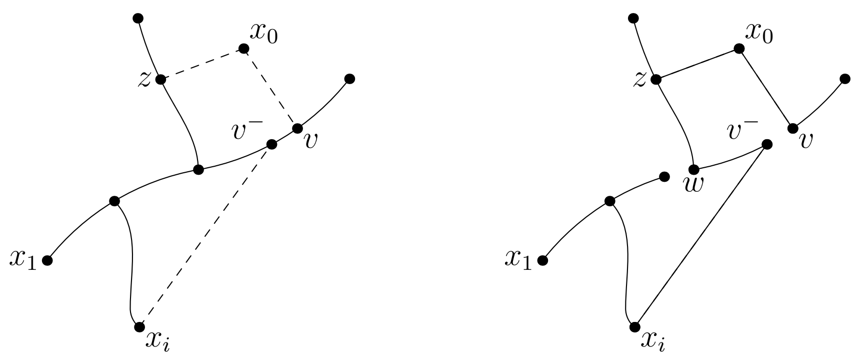


Figure 2: Among the vertices $a, b, c \in N(x_k)$, the vertex a is minimal, while b and c are not.

Claim 2 *For* $k = 0, 1$ *and any vertex* $v \in N(x_k)$ *which is* not *minimal, we have* $v^- \notin N(X - x_k) \cup X$.

Assume $v \in N(x_k)$. It is easy to see that $v^- \notin X$. Indeed, the only vertex v with $v^- \in X$ is the unique neighbor x_1^+ of x_1 in T, and this vertex is necessarily minimal.

Thus, we aim to prove that $v^- \notin N(X - x_k)$. Assume, to the contrary, that $v^- \in N(x_i)$, where $i \neq k$. (The argument is illustrated in Fig. 3.) We distinguish two cases: $k = 0$ and $k = 1$. First, suppose $k = 0$. By the non-minimality of v, we may choose some $z \in N(x_0)$ such that v is not contained in $[x_1, z]$. Form a new tree T' from T by adding the vertex x_0 and replacing the edge vv^- with edges vx_0 and x_0z. Since v and z are clearly in different components of $T - vv^-$, T' is indeed a tree. Note that it may have one leaf more than T since the degree of v^- decreased. The addition of x_iv^- to T' creates a unique cycle C. By our assumption that m, the number of leaves of T, is at least 3, it follows that T' is not a path, and so C contains a vertex w with $d_{T'}(w) \geq 3$. Remove one of the edges of C incident with w from $T' + x_iv^-$ to obtain a tree T''. It is easy to see that T'' has at most m leaves while it covers more vertices than T, a contradiction.



(a) The tree T. (b) The tree T''.

Figure 3: An illustration to the proof of Claim 2 ($k = 0$).

If $k = 1$, we form T' by replacing vv^- with vx_1 in T. By the same argument as above, T' is a tree with at most m leaves. The addition of the edge x_iv^- to T' creates a unique cycle C unless $i = 0$, in which case we have extended T to a tree with at most m leaves spanning more vertices. For $i > 1$, remove an edge $e \in E(C)$ incident with a vertex of degree 3 to get a tree T'' with at most $m - 1$ leaves, spanning all of $V(T)$. This contradiction concludes the proof of Claim 2.

Claim 3 *The intersection of $N(x_k)^-$ and $N(X - x_k) \cup X$ contains at most one vertex for $k = 0, 1$.*

By Claim 2, if $v^- \in N(x_k)^- \cap (N(X - x_k) \cup X)$, then v must be minimal. It is easy to see that there is at most one minimal vertex in $N(x_k)$ $(k = 0, 1)$: if u and u' are both minimal, then $u \in [x_1, u']$ and $u' \in [x_1, u]$, and so $u = u'$. This proves Claim 3.

We now show that Claim 3 implies (1). Clearly,

$$|N(x_k)^-| + |N(X - x_k) \cup X| \leq n + 1.$$

Furthermore, the size of $N(X - x_k) \cup X$ equals $|N(X - x_k)| + m + 1$. By Claim 1, $|N(x_1)^-| = d(x_1)$. Combining these facts together, the case $k = 1$ of inequality (1) follows. As regards $k = 0$, if x_0 has d' neighbors outside T, then $|N(x_0)^-|$ is only $d(x_0) - d'$. On the other hand, we can include the d' neighbors in the total sum as none of them is in $N(X - x_0)$, so the result is the same. Thus, (1) is established.

It is now easy to finish the argument. By the independence of X, we have $|N(X - x_k)| \geq N_m(G)$ for all k. Furthermore, the sum of the degrees of vertices in X is at least $\sigma_{m+1}(G)$. It follows that summing (1) over $k = 0, \ldots, m$, we get

$$(m + 1)N_m(G) \leq (m + 1)n - \sigma_{m+1}(G) - m(m + 1),$$

and so

$$N_m(G) \leq n - \frac{\sigma_{m+1}(G)}{m + 1} - m \leq n - \frac{\sigma_m(G)}{m} - m,$$

by Lemma 4. However, it is clear that $N_m(G) \leq \sigma_m(G)$, and so the above yields

$$N_m(G) \leq \frac{m}{m + 1}(n - m),$$

which contradicts the hypothesis of the theorem. It follows that the tree T spans all of $V(G)$ and the proof is finished. □

4 Spanning spiders

Recall from Section 1 that a tree T is a *spider* if it has at most one *branching vertex* (vertex whose degree in T exceeds 2). The spider T is *centered at v* (where $v \in V(T)$) if none of its vertices, except possibly for v, are branching. It follows that T is centered at a unique vertex, unless T is a path, in which case it is considered as centered at each vertex. If T is a spider centered at v, than a *branch* (or *leg*) of T is any path from v to a leaf of T. (If T is a path, this notion depends on the choice of the 'central' vertex, which will always be clear from the context.)

Since a spanning tree with at most 3 leaves is necessarily a spanning spider, we have already proved one result on spiders: the case $m = 3$ of Theorem 3. In this section, we prove two results concerning the existence of a spanning spider with a prescribed center u. Each of them gives a sufficient condition based on a 'localized' version of the σ_k parameter.

For a vertex $u \in V(G)$ and a positive integer k, define

$$\sigma_k^u(G) = \min_I \sum_{v \in I} d(v)$$

with I ranging over vertex sets of size k such that $I \cup \{u\}$ is independent.

In the following result, the parameter $\sigma_1^u(G)$ is simply the minimum degree of a vertex non-adjacent to u.

Theorem 5 *Let G be a graph of order n. Then for any vertex $u \in V(G)$, there exists a spider in G centered at u and spanning all vertices w of G with $d(w) > n - d(u)$.*

In particular, if $\sigma_1^u(G) > n - d(u)$, then G has a spanning spider centered at u.

Proof. Let W be the set of all vertices w satisfying $d(w) \geq n - d(u)$. Fix a spider S centered at u that covers the maximum number of vertices from W and, subject to this condition, S has as few branches as possible. If S spans W then we are done, so assume this is not the case and take any vertex $w \in W - V(S)$. Let m be the number of branches of S.

All predecessors (see Section 2) considered in this proof will be in the tree S and relative to u. We retain the notation x^- for the predecessor of $x \in V(S) - \{u\}$.

Assume that u has a neighbor v such that v^-w is an edge of G, and that v is contained in a branch P of S with endvertex z. Replacing P by two branches $P_1 = uvPz$ and $P_2 = uPv^-w$, we obtain a spider spanning more vertices, which is a contradiction. Thus, if we let A be the set of the $d(u) - m$ neighbors of u (in G) which are non-adjacent to u in S, then their predecessors are non-adjacent to w in G and they are pairwise distinct.

There are some further vertices which are non-adjacent to w, but are not found in A^-: namely, all the leaves of S and the vertex u. Taking into account A^- and the possibility that u itself is a leaf, we have found $d(u)$ vertices non-adjacent to w, so that $d(w) \leq n - d(u)$, contradicting the assumption. Hence S spans W.

As for the second half of the theorem, the σ_1^u condition implies the existence of a spider S that is centered at u and spans all non-neighbors of u. It is easy to extend S to a spanning spider. $\square$

As a corollary, we obtain the following sufficient condition for the existence of a spanning spider in terms of the minimum degree $\delta(G)$ and the maximum degree $\Delta(G)$ of the graph G.

Corollary 6 *If a connected graph G of order n satisfies $\delta(G) + \Delta(G) \geq n$, then it admits a spanning spider.*

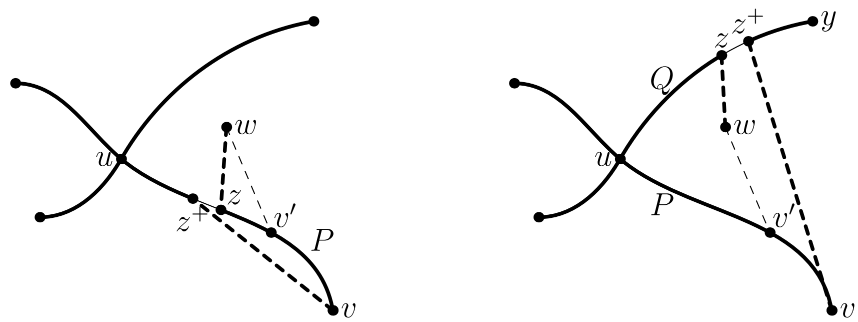
The following example shows that the first half of Theorem 5 is sharp. Consider the complete bipartite graph $K_{m,m+2}$ with the larger partite class denoted by B_1 and the smaller one by B_2. Choose a vertex u in B_1. No spanning spider S has u as the center, since each branch of S would contain at least as many vertices from B_2 as from $B_1 - u$, and on the other hand, $|B_1 - u| > |B_2|$. Now this also implies that no spider with center u covers all vertices of degree at least $n - d(u) = m+2$, since such a spider would necessarily be spanning. The same example shows that the second half of Theorem 5 is sharp up to a small additive constant.

Theorem 7 *Let u be a vertex of a connected graph G on n vertices. If $\sigma_2^u(G) \geq n-1$, then G has a spanning spider centered at u.*

Proof. Take a spider S centered at u with the maximum number of vertices and, subject to this condition, the maximum number of branches. Assuming S does not span G, we may choose $w \notin V(S)$ with a neighbor v' in S since G is connected. Clearly $uw \notin E(G)$, so we may let v be the endvertex of the branch P of S containing v'. Note that v is adjacent neither to w (since otherwise we could extend S to w) nor to u (since we could replace the edge incident with v in S by uv, increasing the number of branches). It follows that $X = \{u, v, w\}$ is an independent set.

In this argument, we shall consider predecessors in S relative to v. For $x \in N_S(w)$, denote by x^+ the unique vertex whose predecessor relative to v is x. This is well-defined since w is adjacent neither to u nor to any leaf of S. For $x \in N(w) - V(S)$, we put $x^+ = x$. Setting $N(w)^+ = \{x^+ : x \in N(w)\}$, we aim to show that $N(v) \cap N(w)^+ = \emptyset$.

Thus let $z \in N(w)$ with $z^+ \in N(v)$ as illustrated in Fig. 4. Clearly, $z \in V(S)$, for otherwise we could use the edge vz to extend S to a spider spanning more vertices. If $z \in V(P)$, then the replacement of P with uPz^+vPzw extends S to w without changing the number of branches. If z is on some branch $Q \neq P$ (ending with, say, y), then we may replace Q with $uQzw$ and P with $uPvz^+Qy$, increasing the number of vertices in the spider. This shows that $N(v)$ and $N(w)^+$ are disjoint as desired.



(a) The case $z \in V(P)$.

(b) The case $z \notin V(P)$.

Figure 4: An illustration to the proof of Theorem 7. The modified spiders are shown in bold.

Now since $|N(w)^{+}| = d(w)$, $|N(v)| = d(v)$, and the vertices u and v are in neither of the sets, we obtain $d(v)+d(w)+2 \leq n$, or equivalently, $d(v)+d(w) \leq n-2$. This contradicts our hypothesis. $\square$

The bound in Theorem 7 is sharp. For the graph $K_{m,m+2}$ and the vertex u from the example given for Theorem 5, one has $\sigma_2^u(K_{m,m+2}) = 2m = n-2$.

5 Problems

We conclude with several open questions. The first of them is a variant of Theorem 3 for graphs of larger connectivity in the spirit of the well-known hamiltonicity condition of Fraisse [6].

Problem 8 Is it true that if G is a κ-connected graph and

$$N_{m+\kappa-1}(G) \geq \frac{m+\kappa-1}{m+\kappa} \cdot (n-m),$$

then G has a spanning tree with at most m leaves?

It is not hard to see that an affirmative answer to this question would generalize the following theorem of Win [13] (conjectured by M. Las Vergnas), which in turn extends the well-known result of Chvátal and Erdős [4] that every κ-connected graph G with independence number $\alpha(G) \leq \kappa+1$ has a Hamilton path:

Theorem 9 *Every κ-connected graph G has a spanning tree with at most $\alpha(G)-\kappa+1$ leaves.*

The following conjecture has been stated in [7]:

Proposition 10 *Any connected graph G with $\sigma_{k+2}(G) \geq n-1$ has a spanning tree with at most k branch vertices.*

Theorem 1 shows that σ_{k+2} may be replaced with σ_{k+3} for $K_{1,3}$-free graphs, and an example in [7] proves that the bound in the theorem does not hold for graphs that may contain induced $K_{1,3}$. On the other hand, the following possibility does not seem to be ruled out.

Problem 11 Is there a constant $C = C(k)$ such that every connected graph G with $\sigma_{k+3}(G) \geq n + C$ has a spanning tree with at most k branch points?

Another question inspired by Theorem 1 is the following.

Problem 12 Does every connected $K_{1,4}$-free graph G with $\sigma_4(G) \geq n$ contain a spanning spider with at most three branches?

It seems plausible that one could find density conditions for the existence of a spanning spider with at most one 'long' leg:

Problem 13 Find a degree condition for the existence of a spanning spider all of whose legs, except possibly one, consist of a single edge.

Proposition 10 shows that any graph with minimum degree no less than $n/3$ admits a spanning spider. If the graph is assumed 2-connected, it is likely that smaller degrees will ensure the existence of a spanning spider. We propose the following question.

Problem 14 Does every 2-connected graph with n vertices and minimum degree at least $n/6$ possess a spanning spider?

Finally, it is natural to ask if there is an analogue of the well-known Bondy–Chvátal closure [3] for spanning trees with few leaves.

Problem 15 Does there exist a function $c(m)$ of m such that $c(m) < 1$ and the following holds: Given any pair of non-adjacent vertices x, y of a graph G with $d(x) + d(y) > c(m) \cdot n$, the graph G has a spanning tree with at most m leaves if and only if the graph $G + xy$ (obtained by adding the edge xy to G) has one?

Acknowledgment

We are indebted to the anonymous referees for their helpful comments. We also thank Aung Kyaw who brought Theorem 9 to our attention.

References

[1] M. Aung and A. Kyaw. Maximal trees with bounded maximum degree in a graph. *Graphs Combin.* **14** (1998), 209–221.

[2] D. Bauer, G. Fan and H. J. Veldman. Hamiltonian properties of graphs with large neighborhood unions. *Discrete Math.* **96** (1991), 33–49.

[3] J. A. Bondy and V. Chvátal. A method in graph theory. *Discrete Math.* **15** (1976), 111–135.

[4] V. Chvátal and P. Erdős. A note on Hamiltonian circuits. *Discrete Math.* **2** (1972), 111–113.

[5] R. J. Faudree, R. J. Gould, M. S. Jacobson and R. H. Schelp. Neighborhood unions and hamiltonian properties in graphs. *J. Combin. Theory Ser. B* **47** (1989), 1–9.

[6] P. Fraisse. A new sufficient condition for Hamiltonian graphs. *J. Graph Theory* **10** (1986), 405–409.

[7] L. Gargano, M. Hammar, P. Hell, L. Stacho and U. Vaccaro. Spanning spiders and light-splitting switches. *Discrete Math.* **285** (2004), 83–95.

[8] A. Kyaw. A sufficient condition for a graph to have a k-tree. *Graphs Combin.* **17** (2001), 113–121.

[9] M. Las Vergnas. Sur une propriété des arbres maximaux dans un graphe, C. R. Acad. Sci. Paris Série A **272** (1971), 1297–1300.

[10] Y. Liu, F. Tian and Z. Wu. Some results on longest paths and cycles in $K_{1,3}$-free graphs. *J. Changsha Railway Inst.*, **4** (1986), 105–106.

[11] M. M. Matthews and D. P. Sumner. Longest paths and cycles in $K_{1,3}$-free graphs. *J. Graph Theory* **9** (1985), 269–277.

[12] V. Neumann Lara and E. Rivera-Campo. Spanning trees with bounded degrees. *Combinatorica* **11** (1991), 55–61.

[13] S. Win. On a conjecture of Las Vergnas concerning certain spanning trees in graphs. *Result. Math.* **2** (1979), 215–224.