Journal of Combinatorics
Volume 7, Number 2–3, 271–283, 2016

The Erdős-Sós conjecture for spiders of four legs

Genghua Fan* and Zhenxiang Huo

The Erdős-Sós Conjecture states that if G is a graph with average degree more than $k-1$, then G contains every tree of k edges. A special case of the conjecture is the well-known Erdős-Gallai theorem: if G is a graph with average degree more than $k-1$, then G contains a path of k edges. A spider is a tree with at most one vertex of degree more than 2, called the center of the spider (if no vertex of degree more than two, then any vertex can be the center). A leg of a spider is a path from the center to a vertex of degree 1. Thus, a path can be regarded as a spider of 1 or 2 legs. In this paper, we prove that if G is a graph with average degree more than $k-1$, then G contains every spider of 4 legs.

Keywords and phrases: Erdős-Sós conjecture, trees, spiders.

1. Introduction

The graphs considered in this paper are finite, undirected, and simple (no loops or multiple edges). The sets of vertices and edges of a graph G are denoted by $V(G)$ and $E(G)$, respectively, and $e(G) = |E(G)|$. The following conjecture is one of the most challenging problems in extremal graph theory.

Erdős-Sós Conjecture. *Let T be a tree of k edges. If G is a graph on n vertices with $e(G) > \frac{k-1}{2}n$, then G contains a copy of T.*

A special case of the conjecture is the well-known Erdős-Gallai theorem [1]: if G is a graph on n vertices with $e(G) > \frac{k-1}{2}n$, then G contains a path of k edges. The conjecture has been investigated on two directions. One is to verify the conjecture for certain families of graphs. For instance, Brandt and Dobson [2] proved that the conjecture is true for graphs without cycles of length less than 5, which was extended by Saclé and Woźniak [7] to graphs without cycles of length 4. Another direction is to verify the conjecture for certain families of trees. The above-mentioned Erdős-Gallai theorem is a classical result on this direction. A *spider* is a tree with at most one vertex of degree more than 2, called the *center* of the spider (if no vertex of degree

*Research supported by NSFC Grant 11331003.

more than two, then any vertex can be the center). A *leg* of a spider is a path from the center to a vertex of degree 1. Thus, a path can be regarded as a spider of 1 or 2 legs. Woźniak [8] proved that the conjecture is true if T is a spider in which each leg has at most 2 edges, which was extended by Fan and Sun [5] to spiders in which each leg has at most 4 edges. In the same paper, Fan and Sun [5] proved that the conjecture is true if T is a spider of 3 legs. In this paper, we prove that

Theorem. *If G is a graph on n vertices with $e(G) > \frac{k-1}{2}n$, then G contains every spider of 4 legs.*

If $xy \in E(G)$, we say that x is *adjacent* to y and that y is a *neighbor* of x. For a subgraph H of G, $N_H(x)$ is the set of the neighbors of x which are in H, and $d_H(x) = |N_H(x)|$ is the *degree* of x in H. When no confusion can occur, we shall write $N(x)$ and $d(x)$, instead of $N_G(x)$ and $d_G(x)$. The *maximum degree* of G is defined by $\Delta(G) = \max\{d(v) : \ v \in V(G)\}$. We use $G-H$ to denotes the graph obtained from G by deleting all the vertices of H together with all the edges with at least one end in H. For two subgraphs A and B in G, $A+B$ denotes the subgraph induced by $V(A) \cup V(B)$. $E(A,B)$ is the set, and $e(A,B)$ is the number, of edges with one end in A and the other end in B. The *length* of a path/cycle is the number of the edges in it. We also use $|C|$, instead of $e(C)$, to denote the length of a cycle C.

Let $C = v_0v_1 \cdots v_c$ be a cycle. For $x \in V(C)$, say $x = v_i$, we use x^+ for v_{i+1} and x^- for v_{i-1}. For $x, y \in V(C)$, $C[x,y]$ denotes the segment $xx^+ \cdots y^-y$. These notations are also applied to vertices of a path P. Thus, if $x, y \in V(P)$, then $P[x,y]$ is the segment of P from x to y.

2. Lemmas

Several lemmas are given in this section, which will be needed in the proof of the main theorem. The first one appeared in [6]. Since the proof is simple, we present it here.

Lemma 1. *Let G be a 2-connected graph on n vertices. Then every vertex of G is contained in a cycle of length at least $\frac{2e(G)}{n-1}$.*

Proof. Let $x \in V(G)$. Construct a new graph G' by adding a new vertex x' joined to, and only to, all neighbors of x. Clearly, $G^{'}$ is also a 2-connected. The average degree of the vertices of $V(G') \setminus \{x, x'\}$ is

$$\frac{\sum_{v \in V(G') \setminus \{x,x'\}} d_{G'}(v)}{|V(G') \setminus \{x,x'\}|} = \frac{2e(G)}{n-1}.$$

By Theorem 1 in [3], x and x' are connected by a path of length at least $\frac{2e(G)}{n-1}$ in G'. Since $N_{G'}(x') = N_{G'}(x)$, this path yields a cycle in G containing x, of the same length $\frac{2e(G)}{n-1}$. $\square$

Lemma 2. *Let G be a graph on n vertices and let T be a k-edge spider of s legs. For any vertex $x \in V(G)$ with $d(x) \geq s$, if $d(v) \geq k$ for all $v \in V(G) \setminus \{x\}$, then G has a copy of T centered at x.*

Proof. Use induction on k. If $k = 1$, it is trivially true. If T is a star, it is also trivially true. Suppose that $k > 1$ and T is not a star. Let w be a leaf ($d(w) = 1$ in T) with $uw \in E(T)$, where u is not the center of T. Set $T' = T - w$, which is a $(k-1)$-edge spider. By the induction, G has a copy of T' centered at x. Since $d(u) \geq k$ in G, and thus u is joined to some vertex $z \in V(G) \setminus V(T')$. Then, $T' + uz$ is the required copy of T in G. $\square$

Lemma 3. *Let T be a k-edge spider. Let C be a cycle in a graph G and $v_0 \in V(C)$. If $|N(v_0) \cap V(C)| \geq k$, then G has a copy of T centered at v_0.*

Proof. Let $C = v_0v_1 \cdots v_cv_0$, where $|C| \geq |N(v_0) \cap V(C)| + 1 \geq k+1$. Suppose that T has s legs. We use induction on s. For $s = 1$, since $|C| \geq k+1$, the lemma holds. Assume thus that $s \geq 2$. Let L be a shortest leg of T with $e(L) = \ell$. Let $Q = v_1v_2 \cdots v_\ell$. If $k \leq 2$, the result trivially holds. Assume that $k \geq 3$, and so $k - \ell \geq 2$. Thus, there is an integer q, $\ell < q < c$, such that $v_0v_q \in E(G)$. Choose such a q as small as possible, and let $C' = v_0v_qv_{q+1} \cdots v_cv_0$ and $G' = G - Q$. Then

$$|N_{G'}(v_0) \cap V(C')| \geq k - e(v_0, Q) \geq k - \ell.$$

Let $k' = k - \ell$ and T' be the k'-edge spider obtained from T by deleting the leg L, except for the center. By the induction, G' has a copy of T' centered at v_0, which together with v_0v_1Q gives a copy of T centered at v_0. $\square$

Lemma 4. *Let C be a maximal cycle in G, containing a given vertex v_0. Let $P = v_0u_1u_2 \cdots u_\ell$ be a path in which $u_i \in V(G - C)$, $1 \leq i \leq \ell$. If $|C| \geq k$ and $e(u_\ell, C - v_0) \geq \frac{k+1}{2} - \ell$, then for any given k-edge spider T of 3 legs, G has a copy of T centered at v_0.*

Proof. Let L_1, L_2, L_3 be the three legs of T, where $e(L_1) \leq e(L_2) \leq e(L_3)$. Let $C = v_0v_1v_2 \cdots v_c$. Let $y_1, y_2, \cdots, y_t$ be the neighbors of u_ℓ on $C - v_0$, with that order around C, where

$$t = e(u_\ell, C - v_0) \geq \frac{k+1}{2} - \ell. \tag{1}$$

If $\ell \geq e(L_1)$, then P gives a copy of L_1 and C gives L_2, L_3, which together yield a copy of T centered at v_0. Assume thus that $\ell \leq e(L_1) - 1 \leq \frac{k}{3} - 1$, and thus $t \geq 2$. For each y_b, $1 \leq b \leq t$, the cycle $C_b = v_0 P u_\ell y_b y_b^- \cdots v_1 v_0$ has length

$$|C_b| = e(P) + 1 + e(C[v_0, y_1]) + e(C[y_1, y_b]).$$

By the maximality of C, $e(C[v_0, y_1]) \geq \ell + 1$, and y_i, y_{i+1} are not adjacent on C, which implies that $e(C[y_1, y_b]) \geq 2(b-1)$. Consequently,

$$|C_b| \geq \ell + 1 + (\ell + 1) + 2(b-1) = 2(\ell + b). \tag{2}$$

In particular, $|C_t| \geq 2(\ell + t)$, and by (1),

$$|C_t| \geq 2(\ell + (\frac{k+1}{2} - \ell)) = k + 1 \geq e(L_1) + e(L_2) + 1.$$

So we may choose b such that $|C_b| \geq e(L_1) + e(L_2) + 1$, and subject to this, b is as small as possible. If $b > 1$, by the minimality of b, $|C_{b-1}| \leq e(L_1) + e(L_2)$. But, by (2) with b replaced by $b-1$, we have that $|C_{b-1}| \geq 2(\ell + b - 1)$, and thus

$$b \leq \frac{e(L_1) + e(L_2)}{2} - \ell + 1. \tag{3}$$

If $b = 1$, since $\ell \leq e(L_1) - 1$, clearly we have (3). Using $e(L_1) + e(L_2) = k - e(L_3)$ in (3), we obtain that

$$b \leq \frac{k - e(L_3)}{2} - \ell + 1,$$

which together with (1) yields that

$$t - b \geq \frac{k+1}{2} - \ell - (\frac{k - e(L_3)}{2} - \ell + 1) = \frac{e(L_3) - 1}{2}.$$

By the maximality of C, $e(C[y_b^+, y_t]) \geq (2(t-b)-1)$ and $e(C[y_t, v_0]) \geq \ell + 1$. Hence, the segment of C from y_b^+ to v_0 of length

$$e(C[y_b^+, y_t]) + e(C[y_t, v_0]) \geq (2(t-b) - 1) + (\ell + 1) \geq e(L_3) - 1 + \ell \geq e(L_3),$$

which gives a path of length $e(L_3)$, starting at v_0. Since $|C_b| \geq e(L_1) + e(L_2) + 1$, we see that C_b gives two paths of lengths $e(L_1)$ and $e(L_2)$, respectively, starting at v_0. Thus G has a copy of T centered at v_0. $\square$

A cycle C in a graph G is 2-*dominating* if $G-C$ consists of isolated vertices and isolated edges. C is *non-extendable* if for each edge $xy \in E(C)$, there is no path $xv_1v_2\cdots v_sy$ with all $v_i \in V(G-C)$, $1 \le i \le s$. The following technical lemma was proved in [4].

Lemma 5. *Let T be a k-edge spider. Let C be a 2-dominating non-extendable cycle in a graph G and $v_0 \in V(C)$ with $d_G(v_0) \ge k$. Suppose that for each component H in $G-C$ with $V(H) \cap N(v_0) \neq \emptyset$, we have that $e(H,C) \ge \frac{k}{2}$ if H is a single vertex, and $e(H,C) \ge k-1$ if H is a single edge. Then G has a copy of T.*

3. Proof of the theorem

Proof. If this is not true, let G be a counterexample with a minimum number of vertices. Let T be a k-edge spider of four legs. For any subgraph G' of G, since a copy of T in G' is also a copy in G, by the minimality of G, we have that $e(G') \le \frac{k-1}{2}|V(G')|$. For any vertex $v \in V(G)$, taking $G' = G - v$, we obtain that

$$d_G(v) \ge \frac{k}{2}, \tag{4}$$

and for any edge $xy \in E(G)$, taking $G' = G - x - y$, we have that

$$d_G(x) + d_G(y) \ge k+1. \tag{5}$$

Suppose that L_1, L_2, L_3 and L_4 are the four legs of T. Let $e(L_i) = \ell_i$, $1 \le i \le 4$. We may assume that $1 \le \ell_1 \le \ell_2 \le \ell_3 \le \ell_4$. If $\ell_4 \le 4$, we are done by [5]. Suppose therefore that $\ell_4 \ge 5$, and thus $k \ge 8$. By the minimality, G is connected. Since $e(G) > \frac{(k-1)n}{2}$, we have that $\Delta(G) \ge k$. We shall prove that

Claim 1. *There is a vertex $x \in V(G)$ with $d_G(x) \ge k$ and x is contained in a cycle of length at least k.*

Proof. If G is 2-connected, let $x \in V(G)$ with $d(x) = \Delta(G) \ge k$. By Lemma 1, x is contained in a cycle C of length at least $|C| \ge \frac{2e(G)}{n-1} > k-1$. By integrality, $|C| \ge k$.

If G is not 2-connected, let H be an endblock of G with b as the unique cut vertex and set $h = |V(H)|$. Let $R = G - (H - b)$. We have that $V(R) \cap V(H) = b$ and $E(R) \cap E(H) = \emptyset$. Then,

$$e(R) = e(G) - e(H) > \frac{(k-1)n}{2} - e(H).$$

If $e(H) \leq \frac{k-1}{2}(h-1)$, then $e(R) > \frac{k-1}{2}|V(R)|$, contradicting the minimality of G. Suppose thus that

$$e(H) > \frac{k-1}{2}(h-1).$$

Note that H is 2-connected. By Lemma 1, every vertex in H is contained in a cycle C (in H) of length at least $|C| \geq \frac{2e(H)}{h-1} > k-1$, and by integrality, $|C| \geq k$. Thus, if there is a vertex $x \in V(H)$ with $d_G(x) \geq k$, then we have the claim. Suppose therefore that $d_H(v) = d_G(v) \leq k-1$ for each $v \in V(H-b)$ and $d_H(b) \leq d_G(b) - 1 \leq k-2$, which means that

$$e(H) \leq \frac{(k-1)(h-1)+k-2}{2} = \frac{(k-1)h-1}{2}, \tag{6}$$

with equality only if $d_H(v) = k-1$ for each $v \in V(H-b)$ and $d_H(b) = k-2$.

Since b is contained by a cycle C of length at least k in H, we see that H has a copy of $L_3 \cup L_4$ centered at b. Noting that $\ell_1 + \ell_2 \leq k/2$, if $d_R(b) \geq 2$, then by (4) and Lemma 2, R has a copy of $L_1 \cup L_2$ centered at b. Consequently, G has a copy of T centered at b. Therefore we assume that $d_R(b) = 1$. Let $B = R - b$. By the minimality of G, $e(B) \leq \frac{(k-1)}{2}|V(B)|$, and so

$$e(H) + 1 = e(G) - e(B) > \frac{k-1}{2}n - \frac{k-1}{2}|V(B)| = \frac{k-1}{2}h,$$

which gives that

$$e(H) \geq \frac{(k-1)h-1}{2}.$$

Thus we have equality in (6). So, $d_H(v) = k-1$ for each $v \in V(H-b)$ and $d_H(b) = k-2 \geq 3$. By Lemma 2 with $s = 3$, H has a copy of $L_2 \cup L_3 \cup L_4$ centered at b, and by (4), R has a path of length ℓ_1, starting at b. Together G has a copy of T centered at b. This contradiction proves Claim 1. □

By Claim 1, we may let

$$C = v_0 v_1 v_2 \cdots v_{c-1} v_c v_0$$

be a cycle in G, where $|C| \geq k$ and $d_G(v_0) \geq k$. Suppose that C has been chosen such that $|C|$ is as large as possible. If $N(v_0) \subseteq V(C)$, then by Lemma 3, we are done. If for each component H of $G - C$ with $V(H) \cap N(v_0) \neq \emptyset$, we have that $|V(H)| \leq 2$, then let G' be the graph obtained from G by removing all components F of $G - C$ with $V(F) \cap N(v_0) = \emptyset$. Then, by the maximality of C and by (4) and (5), C is a cycle in G' satisfying the conditions of Lemma 5. By Lemma 5, G' has a copy of T, and so does G. Assume therefore that there is a component H of $G-C$ with $V(H)\cap N(v_0)\neq\emptyset$ and $|V(H)|\geq 3$. Let

$$P = v_0u_1u_2\cdots u_{\ell-1}u_\ell$$

be a path with $u_i\in V(H)$, $1\leq i\leq \ell$. Suppose that P has been chosen such that ℓ is as large as possible, and subject to this, $d_G(u_\ell)$ is maximum. Since $|V(H)|\geq 3$, we see that $\ell\geq 2$.

Let s be the smallest integer such that $u_su_\ell\in E(G)$, where $u_0=v_0$. Let $x_1,x_2,\cdots,x_p$ be the neighbors of u_ℓ in the order on P, where $x_1=u_{\ell-1}$, $x_p=u_s$, and $p=e(u_\ell,P)$. Let $y_1,y_2,\cdots,y_t$ be the neighbors of u_ℓ on $C-v_0$, with that order around C, where $t=e(u_\ell,C-v_0)$. By the maximality of P, $N_G(u_\ell)\subseteq V(C)\cup V(P)$, and hence,

$$t = d_G(u_\ell)-p. \tag{7}$$

By (4),

$$t\geq \frac{k}{2}-p. \tag{8}$$

By definitions, $p\leq\ell$. If $p\leq \ell-1$, then from (8) above,

$$t\geq \frac{k}{2}-\ell+1. \tag{9}$$

If $p=\ell$, then $u_iu_\ell\in E(G)$ for all i, $0\leq i\leq \ell-1$, which implies that $v_0u_1\cdots u_{\ell-2}u_\ell u_{\ell-1}$ is a path of the same length as P, and then by the choice of P, $d_G(u_{\ell-1})\leq d_G(u_\ell)$, and by (5), we have that $d_G(u_\ell)\geq\frac{k+1}{2}$. Applying this to (7) yields that

$$t\geq \frac{k+1}{2}-\ell. \tag{10}$$

In either case, we have (10). Combining (10) and (8) yields that $2t\geq k-(\ell+p)+\frac{1}{2}$, and by integrality,

$$2t\geq k-(\ell+p)+1. \tag{11}$$

The rest of the proof is divided into two parts, according to the values of t.

Part I. $t\geq 2$.

Case 1. $p\leq \ell_1$. As seen in the proof of (2), for each y_b, $1\leq b\leq t$, the cycle $C_b=v_0Pu_\ell y_by_b^-\cdots v_1v_0$ has length $|C_b|\geq 2(\ell+b)$. Let b be the smallest integer such that $|C_b| \geq \ell_1 + \ell_2 + 1$. If $b > 1$, as seen in (3),

$$b \leq \frac{\ell_1 + \ell_2}{2} - \ell + 1. \tag{12}$$

Combining this with (10), we have that

$$t - b \geq \frac{k+1}{2} - \frac{\ell_1 + \ell_2}{2} - 1 \geq \frac{\ell_3 + \ell_4 - 1}{2}.$$

If $b = 1$, by (8) and using $p \leq \ell_1$, we have that

$$t - b = t - 1 \geq \frac{k}{2} - \ell_1 - 1 \geq \frac{\ell_3 + \ell_4}{2} - 1.$$

In either case, using $\ell_4 \geq 5$, $t - b \geq \frac{\ell_3+3}{2}$. Hence, the segment $u_\ell y_{b+1} \cdots y_{b+2} \cdots y_t y_t^+$ has length at least $2(t-b) > \ell_3$. Let

$$Q_1 = u_\ell y_{b+1} \cdots y_{b+2} \cdots y_m y_m^+,$$

where m is the smallest integer such that $e(Q_1) \geq \ell_3$. By this choice of m, we have that $m - b \leq \frac{\ell_3+1}{2}$, and so

$$m \leq \frac{\ell_3 + 1}{2} + b. \tag{13}$$

If $b > 1$, by (12),

$$m \leq \frac{\ell_1 + \ell_2 + \ell_3 + 3}{2} - \ell = \frac{k - \ell_4 + 3}{2} - \ell.$$

That is, $2m \leq k - \ell_4 - 2\ell + 3$, which together with (11) gives that

$$2(t - m) \geq \ell_4 + \ell - p - 2. \tag{14}$$

Since $p \leq \ell$ and $\ell_4 \geq 5$, we have that $t - m > 0$. Thus, we may let

$$Q_2 = u_\ell y_{m+1} \cdots y_t \cdots v_c.$$

We have that

$$e(Q_2) = 1 + e(C[y_{m+1}, y_t]) + e(C[y_t, v_c]).$$

Noting that $e(C[y_{m+1}, y_t]) \geq 2(t - m - 1)$ and $e(C[y_t, v_c]) \geq \ell$,

$$e(Q_2) \geq 2(t - m) + \ell - 1. \tag{15}$$

By (14),

$$e(Q_2) \geq \ell_4 + (\ell - p) + (\ell - 3).$$

If $p \leq \ell - 1$ or $\ell \geq 3$, then $e(Q_2) \geq \ell_4$. Since $|C_b| \geq \ell_1 + \ell_2 + 1$, there is a copy of $L_1 \cup L_2$ centered at u_ℓ in C_b, which together with Q_1 and Q_2 yields a copy of T centered at u_ℓ. Thus we have that $p = \ell = 2$, which means that $v_0 u_2 \in E(G)$, and by the maximality of P and since $|V(H)| \geq 3$, H is a star centered at u_1. Let $u' \in V(H) \setminus \{u_1, u_2\}$. Then $P' = v_0 u_2 u_1 u'$ is a path of length $e(P) + 1$, contradicting the choice of P. This shows that $b = 1$, and then by (13),

$$m \leq \frac{\ell_3 + 3}{2}. \tag{16}$$

Combining this with (8), we have that $t - m \geq \frac{k - 2p - \ell_3 - 3}{2}$. Since $2p \leq 2\ell_1 \leq \ell_1 + \ell_2$, we obtain that $t - m \geq \frac{\ell_4 - 3}{2} > 0$. Thus, we have the path Q_2 in (15). Combining (16) with (11) and using $p \leq \ell_1$,

$$2(t - m) \geq k - (\ell + \ell_1 + \ell_3) - 2.$$

Applying this to (15) gives that

$$e(Q_2) \geq k - (\ell_1 + \ell_3) - 3 = \ell_4 + \ell_2 - 3. \tag{17}$$

If $\ell_2 \geq 3$, then $e(Q_2) \geq \ell_4$, and as above we are done. We shall show that this is indeed the case. If not, then $\ell_1 \leq \ell_2 \leq 2$. Let $Q_2' = Q_2 \cup (v_c v_0 v_1)$. Using (17),

$$e(Q_2') \geq e(Q_2) + 2 \geq \ell_4 + \ell_2 - 1 \geq \ell_4.$$

Since $|V(H)| \geq 3$ and u_ℓ is the end of a maximum path, there is a path P_1 in H, starting at u_ℓ and $e(P_1) \geq 2$. Let $P_2 = u_\ell y_1 y_1^-$. Then, P_1, P_2, Q_1, Q_2' together give a copy of T centered at u_ℓ. This proves Case 1.

Case 2. $\ell_1 + 1 \leq p \leq \ell_1 + \ell_2$. Let $P' = u_\ell u_{\ell-1} \cdots u_r$ with $e(P') = \ell_1$. Then $e(u_\ell, P') \leq \ell_1$, and so there is $i < r$ such that $u_i u_\ell \in E(G)$. Choose such an i as large as possible and consider the cycle

$$C_b' = u_\ell u_i u_{i-1} \cdots u_1 v_0 v_1 \cdots y_1 \cdots y_b u_\ell.$$

Then

$$|C_b'| = e(P[u_i, v_0]) + e(C[v_0, y_1]) + e(C[y_1, y_b]) + 2.$$

Using that $e(P[u_i, v_0]) \geq p - \ell_1 - 1$, $e(C[v_0, y_1]) \geq \ell + 1$, and $e(C[y_1, y_b]) \geq 2(b-1)$, we obtain that

$$|C'_b| \geq 2b + \ell + p - \ell_1. \tag{18}$$

In particular, $|C'_t| \geq 2t + \ell + p - \ell_1$, and by (11), $|C'_t| \geq k - \ell_1 + 1 > \ell_2 + \ell_3 + 1$. Hence we may choose the smallest b such that $|C'_b| \geq \ell_2 + \ell_3 + 1$.

If $b > 1$, then $|C'_{b-1}| \leq \ell_2 + \ell_3$. By (18) with $b-1$ in the place, we have that

$$2(b-1) + \ell + p - \ell_1 \leq \ell_2 + \ell_3.$$

That is,

$$2(b-1) \leq \ell_1 + \ell_2 + \ell_3 - (\ell + p) = k - \ell_4 - (\ell + p).$$

By (11), $2(b-1) \leq 2t - 1 - \ell_4$, which gives that

$$2(t-b) \geq \ell_4 - 1. \tag{19}$$

So $t - b > 0$ and we may let

$$Q = u_\ell y_{b+1} \cdots y_t \cdots v_{c-1} v_c.$$

We have that

$$e(Q) = 1 + e(C[y_{b+1}], y_t]) + e(C[y_t, v_c]) \geq 1 + 2(t-b-1) + \ell. \tag{20}$$

By (19), $e(Q) \geq \ell_4 + \ell - 2 \geq \ell_4$.

If $b = 1$, then since $t \geq 2$, we directly have that $t - b > 0$, and as above, we also have the path Q. Substituting $b = 1$ in (20) yields that $e(Q) \geq 2t + \ell - 3$, and by (11), $e(Q) \geq k - p - 2$. Since $p \leq \ell_1 + \ell_2$, we have that $e(Q) \geq \ell_4 + \ell_3 - 2$. If $\ell_1 = \ell_2 = \ell_3 = 1$, then it is easy to see that G has a copy of T centered at v_0. We assume thus that $\ell_3 \geq 2$, and hence, $e(Q) \geq \ell_4$.

In either case ($b > 1$ or $b = 1$), we have the path Q with $e(Q) \geq \ell_4$. Then, Q gives a path of length ℓ_4 starting at u_ℓ, P' is a path of length ℓ_1 starting at u_ℓ, and C'_b has a copy of $L_2 \cup L_3$ centered at u_ℓ. Together, we have a copy of T centered at u_ℓ.

Case 3. $p \geq \ell_1 + \ell_2 + 1$. Let $q = \ell_1 + \ell_2$. The cycle $C_1 = P[u_\ell, x_q] \cup x_q u_\ell$ has length at least $\ell_1 + \ell_2 + 1$, which gives a copy of $L_1 \cup L_2$ centered at u_ℓ. The cycle $C_2 = u_\ell x_{q+1} \cdots x_p \cdots u_1 v_0 v_1 \cdots y_1 \cdots y_t u_\ell$ has length

$$|C_2| \geq p - (\ell_1 + \ell_2) + (\ell + 1) + 2(t-1) + 1 = p + \ell - (\ell_1 + \ell_2) + 2t.$$

By (11), $|C_2| \geq k - (\ell_1 + \ell_2) + 1 = \ell_3 + \ell_4 + 1$. So C_2 has a copy of $L_3 \cup L_4$ centered at u_ℓ. Together we have a copy of T centered at u_ℓ. This completes the proof of Part I.

Part II. $t \leq 1$.

By (10), $\ell \geq \frac{k-1}{2}$. If $v_0u_\ell \in E(G)$, then $C' = P \cup \{v_0u_\ell\}$ is a cycle of length at least $|C'| = \ell + 1 \geq \frac{k+1}{2} > \ell_1 + \ell_2$. So C' gives a copy of $L_1 \cup L_2$ centered at v_0. Together with a copy of $L_3 \cup L_4$ centered at v_0 in C, we have a copy of T centered at v_0. Thus $v_0u_\ell \notin E(G)$. That is, $x_p \neq v_0$.

Let $q = \ell_1+\ell_2+\ell_3$. If $p \geq q$, consider the cycle $C_q = x_q \cdots u_{\ell-1}u_\ell x_q$. Then C_q contains q neighbors of u_ℓ. It follows from Lemma 3 that there is a copy of $L_1 \cup L_2 \cup L_3$ centered at u_ℓ. If $p \geq q+1$, then $u_\ell x_{q+1} \cdots x_p \cdots u_1v_0v_1v_2 \cdots v_c$ give a path of length ℓ_4 starting at u_ℓ; if $p = q$ and $t = 1$, then $u_\ell y_1 y_1^+ \cdots v_c v_0 v_1 \cdots y_1^-$ gives a path of length ℓ_4 starting at u_ℓ. In either case, we have a copy of T centered at u_ℓ. In what follows, we assume that,

$$p \leq \ell_1 + \ell_2 + \ell_3,$$

with equality only if $t = 0$. By (8), $k \leq 2p + 2t$, and so

$$\ell_1 + \ell_2 + \ell_3 = k - \ell_4 \leq 2p + 2t - \ell_4.$$

Combining the above two inequalities, we obtain that $\ell_4 \leq p + 2t$, with equality only if $t = 0$. Noting that $t \leq 1$, we have that

$$\ell_4 \leq p + 1. \tag{21}$$

Let P_1 be the segment of P from x_p to u_ℓ. Let P_0 be a shortest path from v_0 to x_p with all internal vertices in $V(H - P_1)$. (Such a path exists since $v_0u_1u_2 \cdots x_p$ is a candidate.) Let $P' = P_0 \cup P_1$. Then $e(P') \geq p + 1$, and by (21), $e(P') \geq \ell_4$, which means that P' has a copy of L_4 centered at v_0.

If $v_0u_r \in E(G)$ for some $u_r \in V(P_1 - x_p)$, then let j be the largest integer with $j < r$ and $u_ju_\ell \in E(G)$ and consider the cycle

$$C' = v_0u_ru_{r+1} \cdots u_\ell u_j u_{j-1} \cdots x_p P_0 v_0.$$

Then $N_H(u_\ell) \subseteq V(C')$ and $\{u_\ell, v_0\} \subseteq V(C')$, and hence $|C'| \geq p + 2$. But, by (8) and since $t \leq 1$, we have that $p \geq \frac{k}{2} - 1$, and so $|C'| \geq \frac{k}{2} + 1$, which means that C' has a copy of $L_1 \cup L_2$ centered at v_0. As seen before, G has a copy of T centered at v_0. This shows that $N(v_0) \cap V(P_1 - x_p) = \emptyset$. Moreover, by the minimality of P_0, $|N(v_0) \cap V(P_0)| = 1$, and hence

$$|N(v_0) \cap V(P')| = 1. \tag{22}$$

Let $u \in V(G) \setminus (V(P') \cup V(C))$. If $e(u, P_0) \geq 4$ (so it must be that $u \in V(H - P')$), then $P_0 + u$ contains a shorter path from v_0 to x_p through u, contradicting the choice of P_0. Thus,

$$e(u, P_0) \leq 3. \tag{23}$$

Let $T' = L_1 \cup L_2 \cup L_3$, $k' = |T'| = k - \ell_4$, and $G' = G - (P' - v_0)$. By (22), $d_{G'}(v_0) \geq d_G(v_0) - 1$. If $N_{G'}(v_0) \subseteq V(C)$, then by Lemma 3, G' has a copy of T' centered at v_0, which together with P' gives a copy of T centered at v_0. Thus, there is $w_1 \in V(G' - C)$ and $v_0 w_1 \in E(G')$. Let $W = v_0 w_1 w_2 \cdots w_m$ be a longest path in G' in which $w_i \in V(G' - C)$, $1 \leq i \leq m$. If, in G, $w_m u_r \in E(G)$ for some $u_r \in V(P_1 - x_p)$, then as seen in C' above (with $v_0 u_r$ replaced by $v_0 W w_m u_r$), we have a copy of T centered at v_0. Therefore, in G,

$$e(w_m, P_1 - x_p) = 0.$$

By (23), $e(w_m, P_0) \leq 3$, and consequently

$$e(w_m, P') \leq 3. \tag{24}$$

By the maximality of W, $N_{G'}(w_m) \subseteq V(W) \cup V(C)$, and so

$$e(w_m, C - v_0) \geq d_G(w_m) - e(w_m, P' \cup W).$$

By (24) and noting that $V(P') \cap V(W) = \{v_0\}$, we have that $e(w_m, P' \cup W) \leq m + 2$, and so $e(w_m, C - v_0) \geq \frac{k}{2} - m - 2$. Using $\ell_4 \geq 5$, we obtain that

$$e(w_m, C - v_0) \geq \frac{k' + 1}{2} - m.$$

It follows from Lemma 4 that G' has a copy of T' centered at v_0, which together with P' gives a copy of T centered at v_0. This completes the proof of the theorem. □

References

[1] P. Erdős and T. Gallai, On maximal paths and circuits of graphs, *Acta. Math. Acad. Sci. Hung* 10 (1959) 337–356. MR0114772

[2] S. Brandt and E. Dobson, The Erdős-Sós conjecture for graphs of girth 5, *Discrete Math.* 150 (1996) 411–414. MR1392751

[3] G. Fan, Long cycles and the codiameter of a graph I, *J. Combin. Theory Ser. B* 49 (1990) 151–180. MR1064673

[4] G. Fan, The Erdős-Sós conjecture for spiders of large size, *Discrete Math.* 313 (2013) 2513–2517. MR3095425

[5] G. Fan and L. Sun, The Erdős-Sós conjecture for spiders, *Discrete Math.* 307 (2007) 3055–3062. MR2371077

[6] B. Ning, Long cycles passing through specified vertices in graphs, M.S. Thesis, Fuzhou University, 2011.

[7] J. F. Saclé and M. Woźniak, The Erdős-Sós conjecture for graphs without C_4, *J. Combin. Theory Ser. B* 70 (1997) 367–372. MR1459879

[8] M. Woźniak, On the Erdős-Sós conjecture, *J. Graph Theory* 21 (1996) 229–234. MR1368749

Genghua Fan
Center for Discrete Mathematics
Fuzhou University
Fuzhou, Fujian 350108
China
E-mail address: fan@fzu.edu.cn

Zhenxiang Huo
Institute of Disaster Prevention
Sanhe, Hebei 065201
China
E-mail address: ganwuqingkong@163.com

Received 31 December 2014