Draft

Spanning Trees with Leaf Distance at Least Four

Atsushi Kaneko, M. Kano[1] and Kazuhiro Suzuki[1]
[1]Department of Computer and Information Sciences
Ibaraki University
Hitachi, Ibaraki 316-8511 Japan
e-mail: kano@cis.ibaraki.ac.jp

Abstract

For a graph G, we denote by $i(G)$ the number of isolated vertices of G. We prove that for a connected graph G of order at least five, if $i(G-S) < |S|$ for all $\emptyset \neq S \subseteq V(G)$, then G has a spanning tree T such that the distance in T between any two leaves of T is at least four. This result was conjectured by Kaneko in *"Spanning trees with constrains on the leaf degree",* Discrete Applied Math., **115** (2001) 73–76. Moreover, the condition in the result is sharp in a sense that the condition $i(G-S) < |S|$ cannot be replaced by $i(G-S) \leq |S|$.

1 Introduction

We consider finite graphs which have neither multiple edges nor loops. Let G be a graph with vertex set $V(G)$ and edge set $E(G)$. We denote by $I(G)$ the set of isolated vertices of G, and by $i(G)$ the number of isolated vertices of G, that is, $i(G) = |I(G)|$. The degree of a vertex v in G is denoted by $\deg_G(v)$. A vertex of degree one is called a *pendant vertex*, and a pendant vertex of a tree is usually called a *leaf*. An edge incident with an pendant vertex is called a *pendant edge*.

There are many results on the existence of spanning trees with some given properties. For example, Win [5] proved that if $\omega(G-S) \leq (k-2)|S|+2$ for all $\emptyset \neq S \subseteq V(G)$, then G has a spanning tree with maximum degree at most k, where $k \geq 2$ and $\omega(G-S)$ denotes the number of components of $G-S$. Ellingham, Nam and Voss [3] showed that every m-edge connected graph has spanning tree T such that $\deg_T(v) \leq 2 + \lceil \deg_G(v)/m \rceil$ for every vertex v of G. Other similar results can be found a recent survey [4] by Kouider and Vestergaard.

In this paper, we consider a spanning tree with given leaf distance. Let T be a spanning tree of a graph. The *leaf distance* of T is defined to be the minimum of distances between any two leaves of T (see Figure 1). The *leaf degree* of a vertex v in T is the number of leaves of T incident with v, and the *maximum leaf degree* of T is the maximum leaf degree among the vertices of T.

Kaneko made the following conjecture.

Conjecture 1 ([1]) *Let $d \geq 4$ be an integer and G be a connected graph with order at least $d+1$. If*

$$i(G-S) < \frac{2|S|}{d-2} \qquad \text{for all} \quad \emptyset \neq S \subseteq V(G), \tag{1}$$

then G has a spanning tree with leaf distance at least d.

Moreover, he showed that if the above conjecture is true for an even integer d, then the condition (1) is sharp in a sense that there exist connected graphs G' that have no spanning tree with leaf distance at least d and satisfy $i(G'-S) \leq 2|S|/(d-2)$ for all $\emptyset \neq S \subseteq V(G')$. Other class of such graphs for $d = 4$ is the following: the corona $cor(K_n)$ of a complete graph K_n, which is obtained from K_n by adding a pendant edge to each vertex of K_n, satisfies $i(cor(K_n) - X) \leq |X|$ for every $\emptyset \neq X \subseteq V(cor(K_n))$, but has no spanning tree with leaf distance at least four.

Note that Conjecture 1 is true for $d = 3$ by the following theorem with $m = 1$ since the leaf distance of a spanning tree T is at least three if and only if the maximum leaf degree of T is one.

Theorem 2 ([1]) *Let G be a connected graph and $m \geq 1$ be an integer. Then G has a spanning tree with maximum leaf degree at most m if and only if*

$$i(G-S) < (m+1)|S| \qquad \text{for all} \quad \emptyset \neq S \subseteq V(G). \tag{2}$$

In this paper, we shall prove the following theorem, which implies that Conjecture 1 is true for $d = 4$.

Theorem 3 *Let G be a connected graph with order at least five. If*

$$i(G-S) < |S| \qquad \text{for all} \quad \emptyset \neq S \subseteq V(G), \tag{3}$$

then G has a spanning tree with leaf distance at least four (Figure 1).

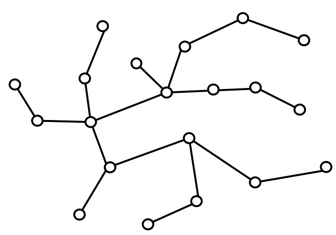

Figure 1: A tree with leaf distance four

2 Proof of Theorem 3

Let G be a graph. For a subset $S \subseteq V(G)$, we write $G-S$ for the subgraph of G induced by $V(G)-S$. For two disjoint vertex subsets S and T of G, we denote by $E_G(S,T)$ the set of edges of G joining a vertex in S to a vertex in T. We denote by $N_G(v)$ the neighborhood of v, and so $\deg_G(v) = |N_G(v)|$. The neighborhood of S is defined by $N_G(S) = \bigcup_{x \in S} N_G(x)$. We call a component with order at least two a *non-trivial component*.

For a set X and its subset Y, we write $Y \subset X$ when Y is a proper subset of X.

Lemma 4 (Kaneko and Yoshimoto [2]) *Let G be a connected bipartite graph with bipartition $A \cup B$, and f be a function defined by $f : A \to \{2, 3, 4, \ldots\}$. Then G has a spanning tree T such that $\deg_T(x) \geq f(x)$ for all $x \in A$ if and only if*

$$|N_G(S)| \geq \sum_{x \in S} f(x) - |S| + 1 \qquad \text{for all} \quad \emptyset \neq S \subseteq A. \tag{4}$$

In particular, if $|N_G(S)| \geq |S| + 1$ for all $\emptyset \neq S \subseteq A$, then G has a spanning tree all whose leaves are contained in B.

Proof of Theorem 3. We prove Theorem 3 by induction on the lexicographic order of $(|V(G)|, |E(G)|)$, and thus, when we consider a graph G, we may assume that the theorem holds for a graph H with either $|V(H)| < |V(G)|$ or $|V(H)| = |V(G)|$ and $|E(H)| < |E(G)|$. By (3), G has no pendant vertices, that is, $\deg_G(x) \geq 2$ for every $x \in V(G)$. We often use this fact without mentioning it.

Claim 1. *We may assume that G has order at least eight.*

If G has order 5 or 6, then we can easily show that G has a Hamiltonian path, which is obviously the desired spanning tree. Suppose that G has order 7. In this case, we shall show that G has a Hamiltonian path, or G is a graph with vertex set $\{u, v_1, v_2, v_3, v_4, v_5, v_6\}$ and edge set $\{uv_i \mid 1 \leq i \leq 6\} \cup \{v_1v_2, v_3v_4, v_5v_6\}$, which is called the 7-*windmill* and has a spanning tree with leaf distance four. Let k be the length of a longest cycle of G. If $k \geq 4$, then by considering the cases of $k = 7,\ 6,\ 5$ and $k = 4$ one by one, we can show that G has a Hamiltonian path. If $k = 3$, then we can show that G has a Hamiltonian path or G is the 7-windmill. Hence Claim 1 follows.

Since every vertex has degree at least two, we have $|E(G)| \geq |V(G)|$. If $|E(G)| = |V(G)|$, then G must be a cycle, and so G has a Hamiltonian path, which is obviously the desired spanning tree. Hence we may assume $|E(G)| \geq |V(G)| + 1$, in particular, G has an edge e_0 such that $G - e_0$ is connected.

Claim 2. *There exists a subset $\emptyset \neq R \subset V(G)$ such that $1 \leq |R| - i(G-R) \leq 2$ and $i(G-R) \geq 1$. Moreover, if $|R| = 3$ and $i(G-R) = 1$, then the two end-vertices of e_0 form an component of $G-R$ or two vertices of R are joined by an edge of G.*

If $G - e_0$ satisfies (3), then $G - e_0$ has the desired spanning tree by the inductive hypotheses. Hence we may assume that $G - e_0$ does not satisfy (3), which implies that there exists a subset $\emptyset \neq S \subset V(G)$ such that $i(G - e_0 - S) \geq |S|$. Since $i(G - e_0 - S) \leq i(G - S) + 2$ and (3), we have

$$1 \leq |S| - i(G - S) \leq |S| - i(G - e_0 - S) + 2 \leq 2.$$

If $i(G-S) \geq 1$, then $R = S$ satisfies the first statement of the claim. If $i(G-S) = 1$ and $|S| = 3$, then $i(G - e_0 - S) = 3$ and e_0 joins two isolated vertices of $G - e_0 - S$. Thus the two end-vertices of e_0 form an component of $G - S$, and hence $R = S$ also satisfies the second statement. Therefore we may assume $i(G - S) = 0$.

If $|S| - i(G - S) = 1$, then $|S| = 1$ and $i(G - e_0 - S) = 1$ or 2, and so e_0 joins the unique isolated vertex y_1 of $G - e_0 - S$ to a non-isolated vertex z_1 of $G - e_0 - S$, or e_0 joins the two isolated vertices y_1 and y_2 of $G - e_0 - S$ (see Figure 2 (a)). Thus we have either $i(G - S \cup \{z_1\}) = |\{y_1\}| = 1$ and $|S \cup \{z_1\}| - i(G - S \cup \{z_1\}) = 1$ or $i(G-S\cup\{y_2\}) = |\{y_1\}| = 1$ and $|S\cup\{y_2\}|-i(G-S\cup\{y_2\}) = 1$. Therefore $R = S\cup\{z_1\}$ or $S \cup \{y_2\}$ satisfies the first tatement of the claim.

If $|S| - i(G - S) = 2$, then $|S| = 2$, $i(G - e_0 - S) = 2$ and e_0 joins the two isolated vertices y_3 and y_4 of $G-e_0-S$ (see Figure 2 (b)). Thus $i(G-S\cup\{y_4\}) = |\{y_3\}| = 1$ and $|S\cup\{y_4\}|-i(G-S\cup\{y_4\}) = 3-1 = 2$. Hence $R = S\cup\{y_4\}$ satisfies the first statement of the claim. Furthermore, since $\deg_G(y_4) \geq 2$ and y_4 is an isolated vertex of $G - e_0 - S$, there exists at least one edge joining y_4 to S. Hence $R = S\cup\{y_4\}$ also satisfies the second statement.

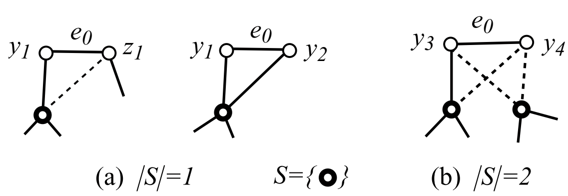

Figure 2: $i(G - S) = 0$; Broken lines are edges or not.

We define an integer m by

$$m = \min\{|X| - i(G - X) \mid R \subseteq X \subset V(G),\ i(G - X) \geq 1\},$$

where R is the vertex subset given in Claim 2. Since $i(G - X) < |X|$ by (3), we have $m \geq 1$. By Claim 2, we have $1 \leq m \leq 2$. Let S_0 be a maximal subset of $V(G)$ subject to $|S_0| - i(G - S_0) = m$, $R \subseteq S_0$ and $i(G - S_0) \geq 1$. Then we have

$$m < |X| - i(G - X) \qquad \text{for every} \quad S_0 \subset X \subset V(G) \ \text{ with } \ i(G - X) \geq 1. \qquad (5)$$

Claim 3. *Every non-trivial component D of $G - S_0$ satisfies $i(D - T) < |T|$ for every subset $\emptyset \neq T \subseteq V(D)$. In particular, $|D| \geq 3$, and D has a Hamiltonian cycle if $3 \leq |D| \leq 4$; and otherwise D has a spanning tree with leaf distance at least four.*

Let D be a non-trivial component of $G - S_0$. For every $\emptyset \neq T \subset V(D)$, by (5) we have

$$m = |S_0| - i(G - S_0) < |S_0 \cup T| - i(G - (S_0 \cup T)),$$

which implies $i(D - T) < |T|$. Hence the Claim follows by induction.

Before going to the next step, we need some definition. Let T be a rooted tree with root v. Then for every vertex $x \in V(T) - v$, we define the *parent* of x as the neighbor of x lying on the x-v path. Similarly, we define the *children* of x as the neighbors of x not lying on the x-v path. Note that there exists exactly one parent for every $x \in V(T) - v$, but there may exist none or more than one children for some $x \in V(T) - v$.

We shall consider the following two cases accroding to m.

Case 1. $m = 1$.

In this case, $|S_0| = i(G - S_0) + 1$. Let B be the bipartite graph with vertex set $I(G-S_0) \cup S_0$ and edge set $E_G(I(G-S_0), S_0)$, which is the set of edges of G joining $I(G-S_0)$ to S_0. If $|N_B(X')| \leq |X'|$ for some subset $\emptyset \neq X' \subseteq I(G - S_0)$, then $i(G - N_B(X')) \geq |X'| \geq |N_B(X')|$, which contradicts (3). Hence

$$|N_B(X)| > |X| \qquad \text{for all} \quad \emptyset \neq X \subseteq I(G - S_0). \tag{6}$$

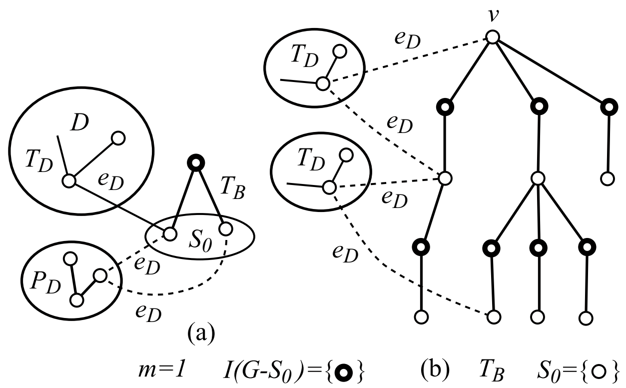

Figure 3: S_0, $I(G - S_0)$ and a spanning tree T_B.

By Lemma 4, B has a spanning tree T_B such that all the leaves of T_B are contained in S_0. If all the vertices in S_0 are leaves of T_B, then $|S_0| = 2$, and there exists at least one non-trivial component of $G - S_0$ by Claim 1. Every non-trivial component D of $G - S_0$ has a Hamiltonian cycle C_D or a spanning tree T_D with leaf distance at least four by Claim 3, and G has an edge e_D joining D to a vertex of S_0 (Figure 3 (a)). If D has a Hamiltonian cycle C_D, then D has a Hamiltonian path P_D such that e_D joins a leaf of P_D to S_0. Hence

$$T = \bigcup_D \ (P_D + e_D \text{ or } T_D + e_D) \quad + T_B,$$

where the union is taken over all the non-trivial components D of $G - S_0$, is the desired spanning tree of G with leaf distance at least four (see Figure 3 (a)).

Therefore we may assume that S_0 contains a vertex v that is not a leaf of T_B. We regard T_B as a rooted tree with root v. Since all the leaves of T_B are contained in S_0 and $m = |S_0| - i(G - S_0) = 1$, for every vertex $y \in I(G - S_0)$, there exists exactly one child of y in S_0 (Figure 2 (b)). Let D be a non-trivial component of $G - S_0$. By the same argument as above, D has a Hamiltonian path P_D or a spanning tree T_D with leaf distance at least four, and a leaf of P_D or a vertex of T_D are joined to S_0 by an edge e_D of G. Then

$$T = \bigcup_D \left(P_D + e_D \text{ or } T_D + e_D\right) \quad + T_B, \tag{7}$$

where the union is taken over all the non-trivial components D of $G - S_0$, is the desired spanning tree of G with leaf distance at least four.

Case 2. $m = 2$.

Let B, T_B, D, P_D, T_D and e_D be the same as in the proof of Case 1. Note that the existence of them are guaranteed by the same argument as in the proof of Case 1.

If all the vertices in S_0 are leaves of T_B, then $|S_0| = 3$ and $i(G - S_0) = 1$. Put $S_0 = \{u_1, u_2, u_3\}$ and $I(G - S_0) = \{w\}$ (see Figure 4 (a)). By Calim 2, $|R| \leq |S_0| = 3$ and by the definition of m, we have $2 = m \leq |R| - i(G - R) \leq 2$, $|R| = 3$ and $i(G - R) = 1$, which imply $S_0 = R$. By the second statement of Claim 2 and Claim 3, two vertices of S_0 are joined by an edge of G. Without loss generality, we may assume that u_1 and u_2 are joined by an edge of G (see Figure 4 (a)).

Assume that there exists a non-trivial component D of $G - S_0$ such that the edge e_D joins D to u_3. If there exists another non-trivial D of $G - S_0$ which is joined to $u_i \in \{u_1, u_2\}$ by an edge e_D, then for a spanning tree $T_B = \{u_3 w, w u_j, u_j u_i\}$ of B, where $\{u_i, u_j\} = \{u_1, u_2\}$,

$$T = \bigcup_D \left(P_D + e_D \text{ or } T_D + e_D\right) \quad + T_B,$$

forms the desired spanning tree of G. If for every non-trivial component D of $G - S_0$, the edge e_D joins D to $\{u_1, u_2\}$, then by choosing $u_i \in \{u_1, u_2\}$ so that at least one e_D is incident with u_i and by letting $T_B = \{u_3 w, w u_j, u_j u_i\}$, where $\{u_i, u_j\} = \{u_1, u_2\}$, $T = \bigcup_D \left(P_D + e_D \text{ or } T_D + e_D\right) \quad + T_B$, forms the desired spanning tree of G. Therefore we may assume that S_0 contains at least one vertex of T_B which is not a leaf of T_B.

Let $v \in S_0$ be a vertex that is not a leaf of T_B. We regard T_B as a rooted tree with root v. Since all the end-vertices of T_B are contained in S_0 and $m = |S_0| - |I(G - S_0)| = 2$, $I(G - S_0)$ has exactly one vertex x that has exactly two children, say y_1 and y_2, and every other vertex of $I(G - S_0) - x$ has exactly one child. (Fig. 2 (b) (c)). If at least one of $\{y_1, y_2\}$ is not an pendant vertex of T_B (Fig. 2 (c)), then we can obtain the desired spanning tree $T = \bigcup_D \left(P_D + e_D \text{ or } T_D + e_D\right) \quad + T_B$, which is given in the proof of Case 1. Hence we hereafter assume that both y_1 and y_2 are end-vertices of T_B (Figure 2 (b)).

If G has an edge joining $\{y_1, y_2\}$ to a non-trivial component D of $G - S_0$, then by choosing e_D to be such an edge, we obtain the desired spanning tree $T = \bigcup_D \left(P_D +\right.$

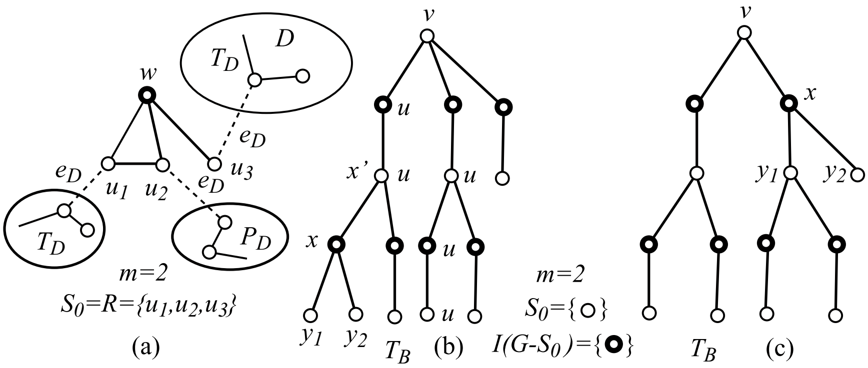

Figure 4: S_0, $I(G - S_0)$ and a spanning tree T_B.

e_D or $T_D + e_D)$ $+ T_B$ of G. So we may assume that no edge of G joins $\{y_1, y_2\}$ to $V(G) - (S_0 \cup I(G - S_0))$. If y_1 and y_2 are adjacent in G, then

$$T = \bigcup_D \ (P_D + e_D \text{ or } T_D + e_D) \ + T_B + y_1y_2 - xy_2$$

is the desired spanning tree of G. Thus we may assume that y_1 and y_2 are not adjacent in G. Since $\deg_G(y_1) \geq 2$, y_1 is adjacent to a vertex u of T_B, which is not y_2. Let x' be the parent of x. Then

$$T = \bigcup_D \ (P_D + e_D \text{ or } T_D + e_D) \ + T_B + y_1u - xx'$$

is the desired spanning tree of G.

Consequently the proof is complete.

References

[1] A. Kaneko, Spanning trees with constraints on the leaf degree, *Discrete Applied Math.*, **115** (2001), 73–76.

[2] A. Kaneko and K. Yoshimoto, On spanning trees with restricted degrees. *Inform. Process. Lett.* **73** (2000),163–165.

[3] M.N. Ellingham, Y. Nam, H.-J. Voss, Connected $[g, f]$-factors, *J. Graph Theory*, **39** (2002), 62–75.

[4] M. Kouider and P. D. Vestergaard, Connected factors in graphs — a survey, *Graphs and Combi.*, **21** (2005) 1–26.

[5] S. Win, On a connection between the existence of k-trees and the toughness of a graph, *Graphs and Combin.*, **5** (1989) 201–205.