Packing paths of length at least two

M. Kano\textsuperscript{a}, G.Y. Katona\textsuperscript{a,b,1}, Z. Király\textsuperscript{c,2}

\textsuperscript{a}Department of Computer and Information Sciences, Hitachi University, Hitachi 316-8511, Japan
\textsuperscript{b}Department of Computer Science and Information Theory, Budapest University of Technology and Economics, H-1364, Hungary
\textsuperscript{c}Department of Computer Science, Eötvös Loránd University, Budapest, Pázmány P. s. 1/C, H-1117 Hungary

Received 12 March 2001; received in revised form 23 December 2003; accepted 20 January 2004

Abstract

We give a simple proof for Kaneko’s theorem which gives a sufficient and necessary condition for the existence of vertex disjoint paths in a graph, each of length at least two, that altogether cover all vertices of the original graph. Moreover we generalize this theorem and give a formula for the maximum number of vertices that can be covered by such a path system.

\copyright{} 2004 Elsevier B.V. All rights reserved.

\textbf{Keywords:} Path factor; Graph factor; Packing

1. Introduction

We consider finite graphs without multiple edges and loops. Let $P_n$ denote the path which contains $n$ vertices and $n-1$ edges. For a subset $X$ of vertices of graph $G$, $G[X]$ denotes the subgraph of $G$ induced by $X$ and $G - X$ denotes the resulting graph after deleting the vertices of $X$ from $G$.

For a set $\{A,B,C,\ldots\}$ of connected graphs, a subgraph $F$ of a graph $G$ is called an $\{A,B,C,\ldots\}$-packing of $G$ if each component of $F$ is isomorphic to one of $\{A,B,C,\ldots\}$. An $\{A,B,C,\ldots\}$-packing is said to be maximum if it covers a maximum number of vertices of $G$. If $F$ is a spanning subgraph, then it is called a perfect $\{A,B,C,\ldots\}$-packing or an $\{A,B,C,\ldots\}$-factor. With this notation the well-known 1-factor (perfect matching) is a $\{P_2\}$-factor. Observe that a graph has a $\{P_3,P_4,P_5\}$-factor if and only if it has a $\{P_n\mid n \geq 3\}$-factor, which we abbreviate as $\{P_{\geq 3}\}$. We will use this fact throughout the paper.

A graph $H$ is said to be \textit{factor-critical} if $H - \{v\}$ has a 1-factor for all $v \in V(H)$. Note that factor critical graphs are connected. For a factor-critical graph $H$ with $V(H) = \{v_1,v_2,\ldots,v_n\}$, add new vertices $\{u_1,u_2,\ldots,u_n\}$ together with new edges $\{v_iu_i\mid 1 \leq i \leq n\}$ to $H$. Then the resulting graph is called a \textit{sun}. Note that $K_2$ is a sun and by definition, we regard $K_1$ also as a sun (see Fig. 1). We call a sun with one vertex a \textit{small sun}, otherwise a \textit{big sun}. We denote by $\text{Sun}(G)$ the set of sun components of $G$ and let $\text{sun}(G) = |\text{Sun}(G)|$ the number of sun components.

A vertex of degree one is called a \textit{pendant vertex}, and an edge incident with a pendant vertex is called a \textit{pendant edge}. For a vertex $v$ of a graph $G$, we denote by $\deg_G(v)$ the degree of $v$ in $G$, and by $N_G(v)$ the neighborhood of $v$ in $G$. For a subset $S \subseteq V(G)$, we define $N_G(S) := \bigcup_{x \in S} N_G(x)$.

Wang [12] characterized the \textit{bipartite} graphs having a $\{P_1\}$-factor. Kaneko recently generalized this theorem to general graphs. There are many results on component factors (for example, see [1,7]), but besides the well known theorem of Tutte [11] about $\{g\}$-factors and the more general theorem of Lovász [8] about $(g,f)$-factors all previous positive results (i.e. that gives a good characterization) allow $P_2$ as a component. Hell and Kirkpatrick [5] proved that if $H$ is a connected

\textsuperscript{1}Research supported by OTKA grant OTKA T 043520 and JSPS.
\textsuperscript{2}Research supported by OTKA grant OTKA T 029772, T 037547 and T 030059.

\textit{E-mail addresses:} kiskat@cs.bme.hu (G.Y. Katona), kiraly@cs.elte.hu (Z. Király).

0012-365X/$ - see front matter \copyright{} 2004 Elsevier B.V. All rights reserved.
Thus sun($G$) contains an $\{H\}$-factor is NP-complete. Thus, for example, we do not have a good characterization of graphs having a $\{P_3\}$-factor.

On the other hand we should mention the corresponding theorems of Hartvigsen (see [2,3]) about cycle-factors without short cycles.

**Theorem 1** (Kaneko [6]). A graph $G$ has a $\{P_{\geq 3}\}$-factor if and only if

$$\text{sun}(G - S) \leq 2|S| \text{ for all } S \subset V(G).$$

(1)

2. A simple proof of Theorem 1

The following lemma is an easy consequence of Hall’s theorem [9, Theorem 1.1.3].

**Lemma 2.** Let $B$ be a bipartite graph with bipartition $X \cup Y$ such that $|Y| = 2|X|$. $B$ has a $\{P_3\}$-factor, i.e. a factor $H$ such that $\deg_H(x) = 2$ for all $x \in X$ and $\deg_H(y) = 1$ for all $y \in Y$ if and only if

$$|N_B(S)| \geq 2|S| \text{ for all } S \subseteq X.$$

Important properties of suns are described in the following two lemmas.

**Lemma 3.** Let $D$ be a big sun, and let $vv'$ be a pendant edge of $D$. Then $D - \{vv'\}$ has a $\{P_4\}$-factor.

**Proof.** Let $V(D) = \{v, x, y, z, \ldots\} \cup \{v', x', y', z', \ldots\}$, where $v'$ is the pendant vertex connected to $v$ etc. Let $F$ be the factor-critical graph $D[\{v, x, y, z, \ldots\}]$. If $M$ denotes the perfect matching in $F - v$, then it is clear that by extending the edges of $M$ by the adjacent pendant edges we obtain a $\{P_4\}$-factor of $D - \{v, v'\}$. □

**Lemma 4.** Let $D$ be a big sun, and $v'$ a pendant vertex of $D$. Then $D - v'$ has a $\{P_4, P_5\}$-factor.

**Proof.** Using the notations of the previous proof, choose a neighbor $x$ of $v$ in $F$. Now $F - x$ has a perfect matching $M$, with some $vy \in M$. Take the path $\{y', y, v, x, x'\}$ and the $\{P_4\}$-s extending the other edges of $M$ by pendant ones. □

**Proof of Theorem 1.** Since no sun component can have a $\{P_3\}$-factor, it is easy to show that if $G$ has a $\{P_{\geq 3}\}$-factor, then (1) holds.

We now prove the sufficiency by induction on $|G| = |E(G)|$. Our method is based on the ideas of Gallai’s proof for Tutte’s theorem. Suppose that $G$ satisfies (1). By setting $S = \emptyset$, condition (1) implies that no component of $G$ is a sun. We may assume that $G$ is connected and $|G| \geq 3$. We consider some cases.

**Case 1:** There exists $\emptyset \neq S \subset V(G)$ such that $\text{sun}(G - S) = 2|S|$.

Choose a nonempty subset $S$ of $V(G)$ satisfying $\text{sun}(G - S) = 2|S|$.

Let $C$ be any non-sun component of $G - S$. Then for a subset $X \subset V(C)$, we have

$$2|S \cup X| \geq \text{sun}(G - (S \cup X)) = \text{sun}(G - S) + \text{sun}(C - X) = 2|S| + \text{sun}(C - X).$$

Thus $\text{sun}(C - X) \leq 2|X|$. Hence $C$ satisfies (1), and so $C$ has a $\{P_{\geq 3}\}$-factor by induction.

We define the bipartite graph $B$ with vertex set $S \cup \text{Sun}(G - S)$ by contracting every sun-component into a single vertex and removing multiple edges and edges inside $S$. Now $|\text{Sun}(G - S)| = 2|S|$, and we show that

$$|N_B(X)| \geq 2|X| \text{ for all } X \subseteq S.$$ (2)

Suppose that $|N_B(Y)| < 2|Y|$ holds for some $Y \subseteq S$. 

![Fig. 1. Suns.](image-url)
Then \( \text{Sun}(G - (S \setminus Y)) \subseteq \text{Sun}(G - S) \setminus N_0(Y) \) holds, and thus
\[
\text{sun}(G - (S \setminus Y)) \geq \text{sun}(G - S) - |N_0(Y)| = 2|S| - 2|Y| = 2|S \setminus Y|
\]
is implied, which contradicts the assumption (1). Thus (2) holds.

Therefore, by Lemma 2, graph \( B \) has a factor \( H \) such that \( \deg_B(x) = 2 \) for all \( x \in S \) and \( \deg_B(C) = 1 \) for all \( C \subseteq \text{Sun}(G - S) \), note that it consists of \( |S| \) copies of \( P_3 \). By making use of this factor, we can obtain a \( \{P_{p3}\} \)-factor of \( G \) in the following way. First, for each edge \( xc \) of \( H \) where \( x \in S \) and \( C \) is a sun, replace this edge with \( xc \), where \( c \) is an arbitrary vertex of \( C \) connected to \( s \). Now every \( P_1 \) of \( H \) has endvertices in two distinct suns. For every endvertex \( c \), if it is not a small sun itself, lengthen the path with the pendant edge incident to \( c \). Now we covered all the small suns and exactly one pendant edge in every big sun. The remaining parts of big suns have a \( \{P_{p3}\} \)-factor by Lemma 3 and the non-sun components have a \( \{P_{p3}\} \)-factor by induction (see Fig. 2).

**Case 2:** \( \text{sun}(G - S) < 2|S| \) for all \( \emptyset \neq S \subseteq V(G) \) and there exists \( \emptyset \neq S' \subseteq V(G) \) for which \( \text{sun}(G - S') = 2|S'| - 1 \).

Choose a subset \( S \) so that \( S \) is maximal among all subsets \( S' \) satisfying \( \text{sun}(G - S') = 2|S'| - 1 \).

Let \( C \) be any non-sun component of \( G - S \) and let \( \emptyset \neq X \subset V(C) \). Using the maximality of \( S \) we obtain
\[
2|S \cup X| - 2 \geq \text{sun}(G - (S \cup X)) = \text{sun}(G - S) + \text{sun}(C - X) = 2|S| - 1 + \text{sun}(C - X).
\]
Thus \( \text{sun}(C - X) \leq 2|X| - 1 \).

Hence \( C \) has a \( \{P_{p3}\} \)-factor by induction.

**Claim 1.** If \( G - S \) has a non-sun component then the desired \( \{P_{p3}\} \)-factor exists.

**Proof.** Let \( C \) be such a component, \( v \in S \) and \( w \in C \) such that \( vw \) is an edge. Let \( w^* \) be a new vertex and consider the graph \( H := G[C] + vw^* \). Using (3) it is easy to see that \( H \) satisfies (1) for nonempty sets: \( \text{sun}(H - X) \leq \text{sun}(C - X) + 1 \leq 2|X| - 1 + 1 \). Clearly \( ||H|| < ||G|| \), so by induction \( H \) has a \( \{P_{p3}\} \)-factor containing a path \( P \) ending with \( \{w, w^*\} \) or \( H \) itself is a sun. In the latter case, by Lemma 3, \( H \) has a \( \{P_1, P_4\} \)-factor so that the only possible \( P_2 \) is \( P = \{w, w^*\} \).

Let \( G' \) be the graph created from \( G \) by adding a new pendant edge \( vw' \). Then \( w' \) is a new sun component of \( G - S \), that is \( \text{sun}(G' - S) = 2|S| \) holds. As before, \( G' \) satisfies (1) for nonempty sets. Construct bipartite graph \( B \) from \( G' \) as in Case 1. The method of Case 1 is used to prove the fact that the empty set satisfies (1). (For \( Y = S \) we know that \( |N_0(Y)| = 2|Y| \) by the property of \( S \).) Thus the same argument shows that \( B \) satisfies (2) hence we obtain a \( \{P_3\} \)-factor of \( G' \) containing a path \( Q \) ending with \( \{v, w'\} \). Now take \( P - w^* \) in \( C \) and \( Q - w' \) in \( G \) and join them by the edge \( vw \). Using this construction we can obtain a \( \{P_{p3}\} \)-factor in the same way as in the previous case, except that for the remaining part of \( C \) we use the \( \{P_{p3}\} \)-factor found in the first paragraph, but without path \( P \).

**Claim 2.** If there exists \( v \in S \) connected to no small sun, or connected to at least two small suns in \( \text{Sun}(G - S) \), then the desired \( \{P_{p3}\} \)-factor exists.

**Proof.** Construct \( B \) as before with the additional pendant edge \( vw' \). Extend the \( \{P_1\} \)-factor of \( B \) as before to obtain a \( \{P_{p3}\} \)-factor of \( G' = G + w' \) and then delete \( w' \). The path containing \( v \) becomes shorter, if it still has at least two edges then we are done, so suppose it contains only one edge \( vw \). By the above construction, \( w \) cannot be in a big sun, because the pendant edge incident to \( w \) would also be part of this path. Therefore \( w \) is a small sun. By our assumptions another small sun \( w^* \) is also connected to \( v \), and \( w^* \) is an endvertex of another path that can be joined to \( vw \) by adding edge \( w^*v \).
Claim 3. If \( \{w\} \in \text{Sun}(G - S) \) is a small sun and \( w \) is not pendant in \( G \) then the desired \( \{P_{\geq 3}\} \)-factor exists.

Proof. As \( w \) is not pendant, it is connected to some \( v \in S \) and \( v' \in S, v' \neq v \). Construct \( B \) as before with the additional pendant edge \( vw' \). We claim that \( B' = B - vw \) satisfies (2). If \( X \subset S \) then either \( v \notin X \) and \( N_{B'}(X) = N_B(X) \) or \( v \in X \) and \( |N_{B'}(X)| \geq 2|X| \) otherwise \( S \setminus X \) would be a set with \( \text{sun}(G - (S \setminus X)) \geq 2|S \setminus X| \). For \( X = S \) we need to prove \( w \in N_{B'}(S) \) which is true because \( vw' \) is an edge. Now take the path ending in \( w' \) in the \( \{P_{\geq 3}\} \)-factor obtained using the \( \{P_{\geq 3}\} \)-factor of \( B' \), delete \( w' \) and connect the remains of this path to the path ending in \( w \) by edge \( vw \).

Summing up, we may assume from now on that there are \( |S| \) small suns in \( \text{Sun}(G - S) \) and \( |S| - 1 \) big suns. Each small sun is a pendant vertex of \( G \) and they are connected to different vertices in \( S \). Moreover every component of \( G - S \) is a sun.

Claim 4. If every vertex of \( G \) with degree \( \geq 2 \) has a pendant neighbor then the desired \( \{P_{\geq 3}\} \)-factor exists.

Proof. Let \( U \) be the set of vertices with degree \( \geq 2 \). If \( G[U] \) has a perfect matching, we are done. Otherwise there exists \( X \subset U \) such that there are more than \( |X| \) factor-critical components in \( G[U \setminus X] \), consequently \( \text{sun}(G - X) > 2|X| \), which is a contradiction.

So, we may assume there is a vertex with degree \( \geq 2 \) which has no pendant neighbor in \( G \). Clearly it is in a big sun \( D \in \text{Sun}(G - S) \). This means that \( G[D] \) has a pendant vertex \( v' \) with neighbor \( v \in D \) so that \( v' \) is connected (in \( G \)) to some \( u \in S \) and if \( |D| = 2 \) then \( v \) is also connected to \( S \). (See Fig. 3.)

**Subcase 2.1:** \( |D| = 2 \)

Let \( G' \) be the graph obtained from \( G \) by deleting the edge \( vu' \). Now \( \text{sun}(G' - S) = 2|S| \). Construct \( B \) as in Case 1. It is easy to see that (2) is satisfied, so we obtain a \( \{P_{\geq 3}\} \)-factor as in Case 1.

**Subcase 2.2:** \( |D| > 2 \)

Let the small sun neighbor of \( u \) be \( \{w\} \). Construct \( B \) as before by adding pendant edge \( uw' \). Take the \( \{P_{\geq 3}\} \)-factor of \( B \), construct a \( \{P_{\geq 3}\} \)-factor of \( G + w' \) and delete the path \( \{w, u, w'\} \). Now we have a \( \{P_{\geq 3}\} \)-factor of \( G - u - w \). If \( v' \) is an endvertex of a path of this factor then we can extend this path by adding edges \( v'u \) and \( uw \). Otherwise the path \( P \) containing \( v' \) ends with \( \{s, v', v\} \) (by our construction this is the only possibility) where \( s \in S, s \neq u \). Observe further that, because \( \text{sun}(G' - S) = 2|S| \), no other path leaves \( D \). Now delete edge \( v'v \) from \( P \) as well as all the paths inside \( D \), extend the shortened \( P \) by the edges \( v'u \) and \( uw \), and use Lemma 4 to obtain a \( \{P_{\geq 3}\} \)-factor of \( D - v' \).

Case 3: \( \text{sun}(G - S) \leq 2|S| - 2 \) for all \( \emptyset \neq S \subset V(G) \).

If \( G \) has a pendant vertex \( u \) connected to \( v \), then \( \text{sun}(G - \{v\}) \geq 1 \), which contradicts the assumption of this case. Thus \( G \) is not a tree, and so we can find an edge \( e \) for which \( G - e \) is connected. For every subset \( \emptyset \neq S \subset V(G - e) \), we have

\[
\text{sun}((G - e) - S) \leq \text{sun}(G - S) + 2 \leq 2|S| - 2 + 2 = 2|S|.
\]

Moreover, \( G - e \) is not a sun because having at least three vertices it would be a sun with at least three pendant vertices, but in this case \( G \) would have at least one pendant vertex as well. Therefore, \( G - e \) has a \( \{P_{\geq 3}\} \)-factor by the inductive hypothesis, and so does \( G \).

Consequently, the proof is complete. \( \square \)
3. The order of a maximum \( \{P_{\geq 3}\}\)-packing

Lemma 5. Let \( B \) be a bipartite graph with bipartition \( X \cup Y \) and \( Y^* \subseteq Y \). Define

\[
def(B) := \max_{Y^* \subseteq Y} (|Y^*| - 2|N_B(Y^*)|)
\]

and

\[
def^*(B) := \max_{Y^* \subseteq Y} (|Y^*| - 2|N_B(Y^*)|).
\]

If \( \def(B) = |Y| - 2|X| \) then \( B \) has a \( \{P_3\}\)-packing which covers \( |Y| - \def(B) = 2|X| \) vertices of \( Y \) including \( |Y^*| - \def^*(B) \) vertices of \( Y^* \), and every vertex of \( X \) is the middle vertex of some \( P_3 \).

Proof. Define bipartite graph \( B' = (X \cup X') \cup Y, E' \) by adding for every vertex \( x \in X \) a new vertex \( x' \in X' \) and connecting \( x' \) to all neighbors of \( x \). By Ore’s theorem [9, Theorem 1.3.1] there exists a matching \( M' \) in \( B' \) that covers

\[|Y| - \max_{Y^* \subseteq Y} (|Y^*| - 2|N_B(Y^*)|)\]

vertices of \( Y \). Since \( \def(B) = |Y| - 2|X| \) and \( |N_B(Y^*)| = 2|N_B(Y^*)| \), we have

\[|Y| - \max_{Y^* \subseteq Y} (|Y^*| - 2|N_B(Y^*)|) = 2|X| = |X \cup X'| \]

so \( M' \) covers all vertices of \( X \cup X' \). Moreover, there exists another matching \( M^* \) that covers

\[|Y^*| - \max_{Y^* \subseteq Y} (|Y^*| - 2|N_B(Y^*)|) = |Y^*| - \def^*(B)\]

vertices of \( Y^* \). It is well known that this implies the existence of a matching which covers \( X \cup X' \) and \( Y^* \) - \( \def^*(B') \) vertices of \( Y^* \) (see [10]). This gives the desired \( \{P_3\}\)-packing in \( B \) if we contract all pairs \( x, x' \). \( \square \)

Let \( k_2(H) \) denote the number of components of \( H \) which consist of an edge, in other terms the number of sun components isomorphic to a \( K_2 \).

Theorem 6. The order of a maximum \( \{P_{\geq 3}\}\)-packing in a graph \( G \) is

\[\text{pp}(G) := |V(G)| - \max_{T \subseteq S \subseteq V(G)} (\text{sun}(G - S) - 2|S| + k_2(G - T) - 2|T|).\]

Proof. It is proved first that the above expression is an upper bound on the order of a maximum \( \{P_{\geq 3}\}\)-factor. Let \( T \subseteq S \subseteq V(G) \) such that \( |V(G)| - \text{pp}(G) = \text{sun}(G - S) - 2|S| + k_2(G - T) - 2|T| \). Clearly if \( F \) is a \( \{P_{\geq 3}\}\)-factor of \( G \), then there is a vertex in at least \( \text{sun}(G - S) - 2|S| \) components in \( \text{Sun}(G - S) \) which cannot be covered by \( F \). Moreover, there are at least \( k_2(G - T) - 2|T| \) \( K_2 \)-components of \( G - T \) where none of the two vertices can be covered.

To prove the other direction choose \( T \subseteq S \subseteq V(G) \) such that \( \text{sun}(G - S) - 2|S| + k_2(G - T) - 2|T| \) is maximum.

Let \( C \) be a non-sun component of \( G - S \). Then for a subset \( X \subseteq V(C) \) we have

\[\text{sun}(G - (S \cup X)) - 2|S \cup X| + k_2(G - T) - 2|T| \leq \text{sun}(G - S) - 2|S| + k_2(G - T) - 2|T|\]

by the choice of \( S \) and \( T \). Since \( \text{sun}(G - (S \cup X)) = \text{sun}(G - S) + \text{sun}(C - X) \) and \( X \cap S = \emptyset \), it follows that \( \text{sun}(C - X) \leq 2|X| \) which implies that \( C \) has a \( \{P_{\geq 3}\}\)-factor. Hence, we may assume from now on that \( G - S \) has only sun components.

Construct a bipartite graph \( B \) from \( G \) by contracting each sun component into a single vertex and removing multiple edges and edges inside \( S \). The set of vertices which arose from the sun components in \( G - S \) is denoted by \( Y \), the set of vertices which arose from the contraction of \( K_2 \) components of \( G - T \) is denoted by \( Q \) and the set of vertices which arose from the contraction of \( K_2 \) components of \( G - S \) is denoted by \( Y^* \). The \( K_2 \) components of \( G - T \) are elements of \( \text{Sun}(G - S) \), because if there exists a \( K_2 \) component of \( G - T \) such that \( V(D) \cap (S - T) \neq \emptyset \) then we get a contradiction by considering \( S \setminus V(D) \) and \( T \). Thus \( Q \subseteq Y^* \subseteq Y \) holds.

First we show that \( \def^*(B) = k_2(G - T) - 2|T| \). Suppose that

\[R - 2|N_B(R)| > k_2(G - T) - 2|T|\]

holds for some \( R \subseteq Y^* \). Then choosing \( N_B(R) \) instead of \( T \) (and keeping \( S \)), obviously violates the choice of \( T \) and \( S \).

Next we prove that \( \def(B) = \text{sun}(G - S) - 2|S| \). Suppose that

\[R - 2|N_B(R)| > \text{sun}(G - S) - 2|S|\]
and optimal simple combinatorial proof—are not algorithmic, their results are, resulting in the first polynomial algorithm to find an
results extending our theorems for that more general problem. Moreover, in contrast with our results which—having a
holds if and only if (4) holds, hence Theorem 6 implies Theorem 1.

References

[6] A. Kaneko, A necessary and sufficient condition for the existence of a path factor every component of which is a path of length at