Balanced $k$-decompositions of graphs

Hsiang-Chun Hsu $^{a, *}$, Gerard Jennhwa Chang $^{a,b,c}$

$^a$ Department of Mathematics, National Taiwan University, Taipei 10617, Taiwan
$^b$ Taida Institute for Mathematical Sciences, National Taiwan University, Taipei 10617, Taiwan
$^c$ National Center for Theoretical Sciences, Taipei Office, Taipei, Taiwan

Abstract

For a given integer $k \geq 2$, a balanced $k$-coloring of a graph $G$ is a mapping $c: V(G) \to \{0, 1, 2, \ldots, k\}$ such that $|A_j| = |A'_j|$ for $1 \leq j < j' \leq k$, where $A_j = \{v \in V(G): c(v) = j\}$ for $0 \leq j \leq k$. The balanced $k$-decomposition number $f_k(G)$ of $G$ is the minimum integer $s$ with the property that for any balanced $k$-coloring $c$ there is a partition $V(G) = V_1 \cup V_2 \cup \cdots \cup V_r$ such that $V_i$ induces a connected subgraph with $|V_i| \leq s$ and $|V_i \cap A_j| = |V_i \cap A'_j|$ for $1 \leq i \leq r$ and $1 \leq j < j' \leq k$. In this paper, we determine $f_k(G)$ for some graphs of high connectivity, trees and complete multipartite graphs.

© 2012 Elsevier B.V. All rights reserved.
where $c$ runs over all balanced $k$-colorings of $G$ and $\{V_1, V_2, \ldots, V_t\}$ runs over all balanced decompositions of $(G; c)$. A balanced $k$-coloring $c$ of $G$ is optimal when every balanced decomposition of $(G; c)$ has size at least $f_k(G)$.

Notice that the balanced 2-decomposition number $f_2(G)$ of $G$ is the same as the balanced decomposition number $f(G)$ introduced by Fujita and Nakamigawa [5]. They established interesting results including (i) $f(G) = 2$ if and only if $G$ is a complete graph of at least two vertices, (ii) $f(T) = n$ for any tree $T$ of $n$ vertices, (iii) $f(K_{m,n}) = \lceil \frac{n-2}{m} \rceil + 3$ for any complete bipartite graph $K_{m,n}$ with $2 \leq m \leq n$, (iv) $f(C_n) = \lceil \frac{n}{2} \rceil$ for any cycle $C_n$ with $n \geq 3$. Then they proposed a conjecture that $f(G) \leq \lceil \frac{k}{2} \rceil + 1$ for any 2-connected graph $G$ of $n$ vertices. They in fact confirmed the conjecture for the case of $|A_1| = |A_2| = 2$. The conjecture was then verified for generalized $\theta$-graphs [2], for 3-connected planar graphs and some special graphs [3], for $TK_4$ and series–parallel graphs [4], and finally for all 2-connected graphs [1]. It was also proved in [2] that for a graph $G$ of $n \geq 3$ vertices, $f(G) = 3$ if and only if $G$ is $\lceil \frac{k}{2} \rceil$-colorable but is not a complete graph.

The purpose of this paper is to study the balanced $k$-decomposition number $f_k(G)$ for general $k$. In particular, we determine $f_k(G)$ for some graphs of high connectivity, trees and complete $r$-partite graphs.

2. Graphs of high connectivity

It is evident that $f_k(G)$ is small when $G$ has high connectivity, as shown in the characterizations of graphs $G$ with $f(G) = 2$ (see [5]) and $f(G) = 3$ (see [2]).

Suppose $G$ is a connected graph of $n$ vertices. For the case when $n < k$, there is only one balanced $k$-coloring, i.e., $A_k = V(G)$. In this case, $f_k(G) = 1$. So, we may only consider the case when $n \geq k$. In this case, $k \leq f_k(G) \leq n$. We now characterize graphs $G$ for which $f_k(G) = k$.

Proposition 1. For any connected graph $G$ of $n \geq k \geq 2$ vertices, the following statements are equivalent:

1. $f_k(G) = k$.
2. $G[S]$ is connected for any $k$-vertex subset $S \subseteq V(G)$, or equivalently, $G$ is $(n-k+1)$-connected.
3. The complement of $G$ does not contain $K_{p,q}$ for any positive integers $p$ and $q$ with $p + q = k$.

Proof. (1) $\Rightarrow$ (2). For any $S = \{v_1, v_2, \ldots, v_k\} \subseteq V(G)$, consider the balanced $k$-coloring $c$ with $A_j = \{v_j\}$ for $1 \leq j \leq k$. Then since $f_k(G) = k$, there exists a balanced decomposition $\{V_1, V_2, \ldots, V_t\}$ of size at most $k$ for $(G; c)$. For some $i$, we have $V_i \supseteq S$, so $k \geq |V_i| \geq |S| = k$. Hence, $S = V_i$ and $G[S]$ is connected.

(2) $\Rightarrow$ (1). Suppose $(A_1, A_2, \ldots, A_k)$ is a balanced $k$-coloring of $G$ with $|A_j| = t$ for $1 \leq j \leq k$. We can choose mutually disjoint $k$-vertex sets $S_i, 1 \leq i \leq t$, such that $|S_i \cap A_j| = 1$ for all $i$ and $j$. By the assumption, each $G[S_i]$ is connected and so $(G; A_1, A_2, \ldots, A_k)$ has a balanced decomposition which consists of $S_1, S_2, \ldots, S_t$ and one-vertex sets. Therefore, $f_k(G) \leq k$ and so $f_k(G) = k$.

(2) $\Rightarrow$ (3). If $G \supseteq K_{p,q}$ for some $p$ and $q$ with $p + q = k$, then taking $S$ to be the vertices of the $K_{p,q}$, we have $G[S]$ is disconnected, a contradiction.

(3) $\Rightarrow$ (2). If $G[S]$ is disconnected for some $S \subseteq V(G)$ with $|S| = k$, then $G[S] \supseteq K_{p,q}$ for some $p$ and $q$ with $p + q = k$, a contradiction. □

Another relation between small balanced $k$-decomposition number and high connectivity of a graph $G$ is the following.

Proposition 2. If $G$ is a connected graph of $n \geq k \geq 2$ vertices and $f_k(G) \leq 2k - 1$, then $G$ is $\lceil \frac{n}{k} \rceil$-connected.

Proof. Suppose to the contrary that $G$ is not $\lceil \frac{n}{k} \rceil$-connected. Then $G$ has a cut set $C$ of $\lceil \frac{n}{k} \rceil - 1$ vertices. Since $|V(G) - C| = n - |C| \geq (k-1)|C| + k \geq |C| + 2$, there exist two vertex sets $A, B \subseteq V(G) - C$ with $|A| + |B| = |C| + 2$ such that there is no edge between $A$ and $B$. We color $A$ and $|B| - 1$ vertices of $C$ by 1, color $B$ and $|A| - 1$ vertices of $C$ by 2, and choose arbitrary $k - 2$ mutually disjoint subsets of size $|C| + 1$ from $V(G) - (A \cup B \cup C)$ as the other $k - 2$ color classes. This is permissible since $|V(G) - (A \cup B \cup C)| \geq (k - 2)(|C| + 1)$. As $C$ is a cut set, any balanced decomposition for this coloring has a balanced set using at least two vertices colored by 1. Hence, this balanced set has size at least $2k$ and so $f_k(G) \geq 2k$, a contradiction. □

Notice that a connected graph $G$ with $f_k(G) = 2k$ may have small connectivity. For instance, consider the graph $G$ obtained from $K_{n-1}$ by adding a new vertex adjacent to a vertex in $K_{n-1}$. If $n \geq 2k$, then $f_k(G) = 2k$ but the connectivity of $G$ is only 1.

3. Trees

In this section, we determine the balanced $k$-decomposition number of a tree $T$ with $n \geq k \geq 2$ vertices. Recall that Fujita and Nakamigawa [5] proved that $f_2(T) = n$. 
Theorem 3. Suppose $T$ is a tree with $n \geq k \geq 2$ vertices, $\ell$ leaves and $m = n - \ell$ non-leaves. If $\ell = kq + r$, where $q$ is a nonnegative integer and $0 \leq r \leq k - 1$, then

$$f_k(T) = \begin{cases} n - r, & \text{if } m + r \leq k - 1; \\ n, & \text{if } m + r \geq k. \end{cases}$$

Proof. For $k = 2$, the formula in the theorem is $f_2(T) = n$, which was proved by Fujita et al. [5]. Now we may assume that $k \geq 3$.

For the case of $m + r \leq k - 1$, any balanced $k$-coloring of $T$ has at least $r$ uncolored leaves. So, there is a balanced decomposition whose size is at most $n - r$. Hence, $f_k(T) \leq n - r$. This establishes the upper bound for $f_k(T)$.

For the lower bound, we need to construct a balanced $k$-coloring of $T$ such that any balanced decomposition of $(T; c)$ has size at least $n$ if $m + r \geq k$, and at least $n - r$ if $m + r \leq k - 1$. We consider two cases.

Case 1. $r = 0$ or $m + r \geq k$.

We draw $T$ as a plane graph inside a circle $\Omega$ such that the leaves of $T$ are on $\Omega$. We order the leaves as $v_1, v_2, \ldots, v_\ell$ along $\Omega$. Let $t = [\ell/k]$. For $r = 0$ or $r \geq 3$, let $A_1 = \{v_1, v_2, \ldots, v_t\}$, $A_2 = \{v_{t+1}, v_{t+2}, \ldots, v_{2t}\}$ and $A_3 = \{v_{2t+1}, v_{2t+2}, \ldots, v_{3t}\}$. For $r = 2$, since $m \geq k - r = k - 2 \geq 1$, we can choose $A_1 = \{v_1, v_2, \ldots, v_t\}$, $A_2 = \{v_{t+1}, v_{t+2}, \ldots, v_2t\}$ and $A_3 = \{v_{2t+1}, v_{2t+2}, \ldots, v_{3t-1}, u\}$ where $u \in N(v_{3t-1})$. For $r = 1$, since $m \geq k - r = k - 1 \geq 2$, we have at least two non-leaf vertices $x$ and $y$ adjacent to some leaves. Hence we can choose an ordering of leaves $v_1, v_2, \ldots, v_\ell$ such that $x \in N(v_{3t-1})$ and $y \in N(v_{2t})$, we have to redraw $T$ if necessary. Let $A_1 = \{v_1, v_2, \ldots, v_t\}$, $A_2 = \{v_{t+1}, v_{t+2}, \ldots, v_{2t-1}, x\}$ and $A_3 = \{v_{2t}, v_{2t+1}, \ldots, v_{3t-2}, y\}$. In all the above cases, we can choose the other $k - 3$ color classes $A_4, A_5, \ldots, A_k$ such that all leaves are colored. We claim that this balanced $k$-coloring is optimal. For a balanced decomposition of $(T; A_1, A_2, \ldots, A_k)$, there is a balanced set $S$ containing a path $P$ from $v_{t+1}$ to $A_2$. Since all paths from $A_2$ to $A_1$ must intersect $P$, all colors 1 and 2 vertices, and hence all colored vertices, must be in $S$. Because all leaves are colored, $S$ contains all leaves and $T[S]$ is connected. Therefore, $S = V(T)$ and so $f_k(T) = n$.

Case 2. $r \neq 0$ and $m + r \leq k - 1$.

We first claim that we can delete $r$ leaves from $T$ such that all non-leaves remain non-leaves after the deletion. Let $B_i$ be the set of non-leaves with exactly $i$ leaves as its neighbors and $b_i = |B_i|$ for $i \geq 0$. When we delete $i - 1$ leaves from neighbors of a vertex in $B_i$, all non-leaves remain as non-leaves. Hence the previous claim is equivalent to that $\sum_{i=1}^r (i - 1) b_i \geq r$. Suppose to the contrary that $\sum_{i=1}^r (i - 1) b_i < r$. Then $kq + r = \ell = \sum_{i=1}^r i b_i = \sum_{i=1}^r b_i + \sum_{i=2}^r (i - 1) b_i < m + r \leq k - 1$. This happens only when $q = 0$ and $r = \ell$. Hence $n = m + \ell = m + r \leq k - 1$, a contradiction to the assumption that $n \geq k$.

Therefore we can delete $r$ leaves from $T$ to get a tree $T'$ such that $T'$ has the same $m$ non-leaves as $T$ and $\ell' = \ell - r$ leaves, where $\ell' = kq + \ell'$ with $\ell' = 0$. Hence $f_k(T') = n - r$ since $T'$ satisfies $r' = 0$ in Case 1. So we have an optimal balanced $k$-coloring $(A_1, A_2, \ldots, A_k)$ for $T'$ such that the only balanced decomposition of $(T'; A_1, A_2, \ldots, A_k) = \{V(T')\}$. Therefore, any balanced decomposition of $(T; A_1, A_2, \ldots, A_k)$ has a component containing $V(T')$. This implies that $f_k(T) \geq n - r$. □

4. Complete multipartite graphs

For complete multipartite graphs, Fujita and Liu [2] proved that $f_2(K_{n_1, n_2, \ldots, n_r}) = \left\lfloor \frac{n_1-2}{2} \right\rfloor + 3 = \frac{n - 2}{\sum_{i=1}^r n_i}$, where $r \geq 2, n_1 \geq n_2 \geq \ldots \geq n_r \geq 1$ and $n = \sum_{i=1}^r n_i$. The following theorem considers $f_k(K_{n_1, n_2, \ldots, n_r})$ for $k \geq 2$.

Theorem 4. For $r \geq 2$, if complete $r$-partite graph $G = K_{n_1, n_2, \ldots, n_r}$ has $n \geq k \geq 2$ vertices, where $n_1 \geq n_2 \geq \cdots \geq n_r \geq 1$ and $m = n - n_1$, then

$$f_k(G) = \begin{cases} k + kt, & \text{if } t \leq \frac{n - k}{km} < t + \frac{1}{k} < \frac{1}{k} \leq \frac{n - k}{km} < t + 1 \text{ where } t \in \mathbb{Z}^+ \cup \{0\}. \\ k + kt + 1, & \text{if } \frac{1}{k} \leq \frac{n - k}{km} < t + 1 \text{ where } t \in \mathbb{Z}^+ \cup \{0\}. \end{cases}$$

Proof. Let $S_i$ be the partite set of $G$ with $|S_i| = n_i$ and $S_0 = V(G) - S_i$ for $1 \leq i \leq r$. First, we give a balanced $k$-coloring $c$ of $G$ and prove that any balanced decomposition of $(G; c)$ has sufficiently large size.

For the case when $t \leq \frac{n - k}{km} < t + \frac{1}{k}$ with $t \in \mathbb{Z}^+ \cup \{0\}$, we have $|S_i| = n - m \geq (kt - 1)m + k = (tm - m + 1)k + (k-1)(tm + 1) .

Consider a balanced $k$-coloring of $G$ that colors $m$ vertices of $S_1$ and $tm - m + 1$ vertices of $S_1$ by $1$ for each color $i$ from 2 to $k$, and leaves the other vertices uncolored. Since some vertex in $S_1$ must be adjacent to $\lceil \frac{tm+1}{m+1} \rceil = t + 1$ vertices colored by $k$ in any balanced decomposition of $(G; c)$, we have $f_k(G) \geq k + kt$.

For the case when $t + \frac{1}{k} \leq \frac{n - k}{km} < t + 1$ with $t \in \mathbb{Z}^+ \cup \{0\}$, we have $|S_i| = n - m \geq k(tm + 1)$. Consider a balanced $k$-coloring $c$ of $G$ that colors $tm + 1$ vertices of $S_1$ by $i$ for each color $i$ from 1 to $k$, and leaves the other vertices uncolored. Since some vertex in $S_1$ must be adjacent to $\lceil \frac{tm+1}{m+1} \rceil = t + 1$ vertices colored by $k$ in any balanced decomposition of $(G; c)$, we have $f_k(G) \geq k + kt + 1$. 

On the other hand, given a balanced \( k \)-coloring \( c \) of \( G \), we have to find a balanced decomposition of \((G; c)\) with small size. Choose, as many as possible, disjoint balanced sets of \( k \) colored vertices. Let \( \beta \) be the maximum. Let \( U \) be the union of these \( \beta \) balanced sets. If \( S_i - U \) and \( S_i^* - U \) both have some colored vertices for some \( i \), then we can choose another balanced set of \( k \) vertices outside \( U \), a contradiction to the maximality of \( \beta \). Hence for \( i \neq j \), both \( S_i - U \) and \( S_j - U \) cannot contain colored vertices. If \( S_j - U \) contains colored vertices of colors from 1 to \( k \), then any one of the previous \( \beta \) balanced sets of \( k \) vertices has exactly one vertex in \( S_j \), otherwise we can change the balanced \( k \)-sets such that \( \beta \) is not the maximum.

By the above observations, we may consider three cases: (i) colored vertices are all in \( S_j \cup U \) but not all in \( U \) for some \( i \neq 1 \), and (iii) colored vertices are all in \( S_j \cup U \) but not all in \( U \) for some \( i \neq 1 \), and (iii) colored vertices are all in \( S_j \cup U \) but not all in \( U \) for some \( i \neq 1 \). For case (i) when \( U \) contains all colored vertices, we have a balanced decomposition of size at most \( k \). For case (ii) when \( S_i - U \) contains some colored vertices for some \( i \neq 1 \), any of the \( \beta \) balanced \( k \)-sets has one vertex in \( S_i^* \). Since \( S_i^* \) is a subset of \( S_j^* \), has at least \( n_1 - \beta \) uncolored vertices and \( S_j^* - U \) has at most \( n_1 - (k - 1) \beta \leq n_1 - \beta \) colored vertices, we have a balanced decomposition of size \( k + 1 \). For case (iii), \( S_j - U \) contains some colored vertices, say \( S_j - U \) contains \( \ell k \) colored vertices. Since any of the \( \beta \) balanced \( k \)-sets has only one vertex in \( S_j^* \), there are \( \alpha = m - \beta \) uncolored vertices in \( S_j^* \).

Now, for the case when \( t \leq \frac{n - k}{km} < t + 1 \) with \( t \in \mathbb{Z}^+ \cup \{0\} \), we have \( \ell k \leq n_1 - \beta (k - 1) \leq n_1 = n - m \leq km + k - 1 \) and so \( \ell \leq \left\lfloor \frac{km + k - 1}{k} \right\rfloor = tm = t(\alpha + \beta) \). Now, consider the balanced decomposition formed by grouping at most \( \alpha t k \) of the \( \ell k \) colored vertices with the \( \alpha \) uncolored vertices of \( S_j^* \), and at most \( \beta t k \) of the \( \ell k \) colored vertices with the \( \beta \) balanced \( k \)-sets. This has size at most \( k + \alpha t k \), which gives \( f_k(G) \leq k + kt \).

For the case when \( t + 1 \leq \frac{n - k}{km} < t + 2 \) with \( t \in \mathbb{Z}^+ \cup \{0\} \), we have \( \ell k \leq n_1 - \beta (k - 1) = n - m - \beta (k - 1) \leq \left( (k + 1) m + k - \beta (k - 1) \right) - \beta (k - 1) = (t + 1) m - \beta = (t + 1) \alpha + t \beta \).

By a similar argument, there is a balanced decomposition of size at most \( k + kt + 1 \) which gives \( f_k(G) \leq k + kt + 1 \). \( \square \)

We remark that in the formula for \( f_k(K_{n_1,n_2,\ldots,n_r}) \), we have \( t = \left\lfloor \frac{n - k}{km} \right\rfloor \) and so \( f_k(K_{n_1,n_2,\ldots,n_r}) = k + \left\lfloor \frac{n - k}{km} \right\rfloor \left( \left\lfloor \frac{n - k}{km} \right\rfloor + \left\lfloor \frac{n - k}{km} \right\rfloor + 1 \right) \). \( \square \)

When \( H \) is a connected subgraph of \( G \), we may not have \( f_k(H) \geq f_k(G) \). For instance, \( f_2(P_3) = 3 \) and \( f_2(P_2) = 2 \). But when \( H \) is a connected spanning subgraph of \( G \), we have the following.

**Proposition 5.** If \( H \) is a connected spanning subgraph of \( G \), then \( f_k(H) \geq f_k(G) \).

**Proof.** This is obvious, since all balanced \( k \)-colorings and corresponding balanced decompositions of \( H \) are also those of \( G \), the assertion holds. \( \square \)

**Corollary 6.** If \( n = n_1 + m \geq 2m \) and \( G \) is a connected graph of \( n \geq k \geq 2 \) vertices such that \( K_{n_1, m} \subseteq G \subseteq K_{n_1, 1, 1, \ldots, 1} = K_n - E(K_{n_1, m}) \), then \( f_k(G) = k + k \left\lfloor \frac{n - k}{km} \right\rfloor \left( \left\lfloor \frac{n - k}{km} \right\rfloor + \left\lfloor \frac{n - k}{km} \right\rfloor + 1 \right) \).

**Proof.** The corollary follows from \( V(K_{n_1, m}) = V(G) = V(K_{n_1, 1, 1, \ldots, 1}) \) and the fact that \( f_k(K_{n_1, m}) = f_k(K_{n_1, 1, 1, \ldots, 1}) = k + k \left\lfloor \frac{n - k}{km} \right\rfloor \left( \left\lfloor \frac{n - k}{km} \right\rfloor + \left\lfloor \frac{n - k}{km} \right\rfloor + 1 \right) \) by the previous theorem. \( \square \)

**Corollary 7.** If \( G \) is a connected graph of \( n \geq k \geq 2 \) vertices such that \( \alpha(G) \geq \alpha = n - m \), then \( f_k(G) \geq k + k \left\lfloor \frac{n - k}{km} \right\rfloor \left( \left\lfloor \frac{n - k}{km} \right\rfloor + \left\lfloor \frac{n - k}{km} \right\rfloor + 1 \right) \).

**Proof.** The corollary follows from that \( G \) is a spanning subgraph of \( K_{n_1, 1, 1, \ldots, 1} = K_n - E(K_{n_1}) \). \( \square \)

**Acknowledgments**

The authors thank the referees for many constructive suggestions.

**References**


