BALANCE IN RANDOM TREES

Azer Akhmedov, Warren Shreve

ABSTRACT: We prove that a random labeled (unlabeled) tree is balanced. We also prove that random labeled and unlabeled trees are strongly k-balanced for any $k \geq 3$.

Definition: Color the vertices of graph G with two colors. Color an edge with the color of its endpoints if they are colored with the same color. Edges with different colored endpoints are left uncolored. G is said to be *balanced* if neither the number of vertices nor and the number of edges of the two different colors differ by more than one.

1. Introduction

The notion of a balanced graph is defined [LLT] as follows:

Definition 1.1. Let $G = (V, E)$ be a finite simple graph, $k \geq 2$ be an integer, $c : V \to \{1, \ldots, k\}$ be a map. For all $i \in \{1, \ldots, k\}$, we write $V_i(c) = c^{-1}(\{i\}), E_i(c) = \{uv \in E \mid u, v \in V_i(c)\}$. We also write $v_i(c) = |V_i(c)|, e_i(c) = |E_i(c)|$. The map c is called *a coloring*.

The case of $k = 2$ is especially interesting. In this case, the sets $V_1(c), V_2(c), E_1(c), E_2(c)$ are called the sets of black vertices, white vertices, black edges, and white edges respectively. If the coloring c is fixed we may drop it in the notation.

Definition 1.2. A finite simple graph $G = (V, E)$ is called *balanced* if there exists a coloring $c : V \to \{1, 2\}$ such that $|v_1(c) - v_2(c)| \leq 1$ and $|e_1(c) - e_2(c)| \leq 1$. A map $c : V \to \{1, 2\}$ satisfying this condition is called *a balanced coloring*.

The graph in Figure 1 is balanced since we have shown a balanced coloring of it.

It is not difficult to see that:

a) The complete graph K_n is balanced iff $n \leq 3$ or n is even.

b) The star S_n is balanced iff $n \leq 5$; see Fig.2 for a balanced coloring of S_5.

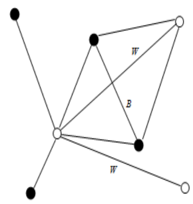

FIGURE 1. with the given coloring, the graph has 4 black and 3 white vertices; it also has 2 white edges (labeled with a “W”) and 1 black edge (labeled with a “B”)

c) The double star $S_{p,q}$ is balanced iff $|p-q| \leq 3$. (The double star $S_{p,q}$ is a connected graph where some two adjacent vertices have degree p and q, and all other vertices have degree 1.)

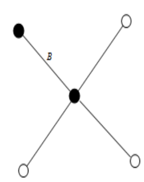

FIGURE 2. a balanced coloring of S_5

In [Cah1], the author introduces a somewhat similar notion of a cordial graph, a generalization of both graceful and harmonious graphs. It has been conjectured by A. Rosa, G. Ringel and A. Kotzig that every tree is graceful (*Graceful Tree Conjecture*, [Ga]), and it has been conjectured by R. Graham and N. Sloane that every tree is harmonious (see [GrSl]). While these conjectures are still open, in [Cah2] it is proved that every tree is cordial.

Not every tree is balanced; in this paper, we will be interested in the property of being balanced for a random labeled and unlabeled tree, as well as for random labeled graphs.

The main results of the paper are Theorem A and Theorem B, stated below.

Theorem A. A random labeled (unlabeled) tree is balanced; more precisely, if $t_n(\tau_n)$ denotes the number of all labeled (unlabeled) trees on n vertices, and $b'_n(b''_n)$ denotes the number of all balanced labeled (unlabeled) trees on n vertices, then $\lim\limits_{n\to\infty} \dfrac{b'_n}{t_n} = 1$ and $\lim\limits_{n\to\infty} \dfrac{b''_n}{\tau_n} = 1$.

Remark 1.3. In this paper, for simplicity, we consider only uniform models of random graphs and random trees. The results can be extended to a large class of non-uniform models as well. Note that $t_n = n^{n-2}$(see [Cay] or [W]) and $\tau_n \sim C\alpha^n n^{-5/2}$ for some positive constants C and α (see [O]).

We also would like to introduce the notion of k-balanced graphs.

Definition 1.4. Let $k \geq 2$. A finite simple graph $G = (V, E)$ is called *k-balanced* if there exists a coloring $c : V \to \{1, 2, \dots, k\}$ such that $|v_i(c) - v_j(c)| \leq 1$ and $|e_i(c) - e_j(c)| \leq 1$ for all distinct $i, j \in \{1, 2, \dots, k\}$. The map c will be called a *k-balanced coloring*.

Definition 1.5. Let $k \geq 2$. A finite simple graph $G = (V, E)$ is called *strongly k-balanced* if there exists a coloring $c : V \to \{1, 2, \dots, k\}$ such that $|e_i(c)| = 0, 1 \leq i \leq k$, and $|v_i(c) - v_j(c)| \leq 1$ for all distinct $i, j \in \{1, 2, \dots, k\}$. The map c will be called a *strongly k-balanced coloring*.

In more popular terms, a finite simple graph is strongly k-balanced iff it is *k-equitably colorable*. In Section 5 we study some basic properties of k-balanced graphs. We prove the following theorem.

Theorem B. For all $k \geq 3$, a random (labeled) tree is strongly k-balanced.

Remark 1.6. Let us emphasize that Theorem B is originally due to B. Bollobás and R. Guy (see [BG]). Our proof in this paper is very different with some ingredients which might be interesting independently.

Remark 1.7. It has been proved by I. Ben-Eliezer and M. Krivelevich (see [BK]) that a random graph is balanced. For $k \geq 3$, it seems quite plausible that a random graph is indeed k-balanced. However, notice that the clique number of a random graph on n vertices is at least $\log_2(n)$ (see [B]) thus a random graph is not strongly k-balanced.

Acknowledgment: We would like to thank to B. Gittenberger for the discussion and for bringing the reference [GoSh] to our attention; to I. Pak for his comments; to M. Krivelevich for bringing [BG] to our attention. We also would like to extend our gratitude to the anonymous referee for many helpful remarks, and for pointing out a flaw in the proof of Lemma 5.3.

Notes: **1.** For any finite simple graph G, we will denote the maximum degree of G by $\Delta(G)$.

2. A vertex of degree one will be called a *leaf vertex* or simply a *leaf*. A non-leaf vertex v is called a *pre-leaf vertex* if it is adjacent exactly to $m-1$ leaves where $m = deg(v)$. A pre-leaf vertex of degree two is called *special*.

3. For $n \geq 2$, there exists a unique tree up to isomorphism with n vertices and maximum degree at most two; we will call this tree a *path on n vertices*, and denote it with P_n.

4. For a finite simple graph $G = (V, E)$ and for a subset $A \subseteq V$, the induced subgraph $G[A]$ will be called the *full subgraph* of G on A.

5. For a tree $G = (V, E)$ and a non-leaf vertex $v \in V$, a subset $A \subseteq V$ will be called a *branch* of G with respect to v if A is a maximal subset such that the full subgraph $G[A]$ is connected and $d(v, A) = 1$ where $d(.,.)$ denotes the distance in the tree G.

2. Characterization of Balanced Graphs

In this section we observe some basic facts on balanced and k-balanced graphs. Let us first prove a very simple lemma which provides a necessary and sufficient condition for a graph to be balanced.

Lemma 2.1. *Let G be a finite simple graph with n vertices, and degrees $d_1, \ldots, d_n$. G is balanced if and only if there exists a partition $\{1, \ldots, n\} = I \sqcup J$ such that*

(i) $\big||I| - |J|\big| \leq 1$

(ii) $|\sum_{k \in I} d_k - \sum_{k \in J} d_k| \leq 2$

Proof. Let $G=(V,E), V=\{v_1,\dots,v_n\}, \deg(v_i)=d_i, 1\le i\le n$.

Assume G is balanced with a balanced coloring $c: V\to\{1,2\}$.

Let $I=\{i \mid 1\le i\le n, c(v_i)=0\}, J=\{i \mid 1\le j\le n, c(v_i)=1\}$.

Since c is a balanced coloring, we get $\big||I|-|J|\big|\le 1$ so condition (i) is satisfied.

For every $i\in I$, we denote

$$p_i=|\{k\in I: v_iv_k\in E\}|, q_i=|\{k\in J: v_iv_k\in E\}|,$$

and for every $j\in J$, we denote

$$m_j=|\{k\in I: v_jv_k\in E\}|, n_j=|\{k\in J: v_jv_k\in E\}|$$

Then $\sum_{i\in I} q_i=\sum_{j\in J} m_j=|E\backslash(E_1\cup E_2)|$. On the other hand, since G is balanced, we have $\sum_{i\in I} p_i=2|E_1|, \sum_{j\in j} n_j=2|E_2|$.

Then $|\sum_{k\in I} d_k-\sum_{k\in J} d_k|=|\sum_{k\in I}(p_k+q_k)-\sum_{k\in J}(m_k+n_k)|=2\big||E_1|-|E_2|\big|\le 2$. Thus condition (ii) is also satisfied.

To prove the converse, assume conditions (i) and (ii) are satisfied. We define the coloring $c: V\to\{1,2\}$ as follows: for every $i\in I$ we set $c(v_i)=1$ and for every $j\in J$ we set $c(v_j)=2$.

Then we have $|E_1|=\frac{1}{2}\sum_{i\in I} p_i,\ |E_2|=\dfrac{1}{2}\sum_{j\in J} n_j,$ and $\sum_{i\in I} q_i=\sum_{j\in J} m_j=|E\backslash(E_1\cup E_2)|$.

On the other hand,

$$\sum_{k\in I} d_k=\sum_{k\in I}(p_i+q_i) \text{ and } \sum_{k\in J} d_k=\sum_{k\in J}(m_j+n_j)$$

Then by condition (ii), we get $\big||E_1|-|E_2|\big|=\frac{1}{2}|\sum_{k\in I} d_k-\sum_{k\in J} d_k|\le 1$.

■

Corollary 2.2. It is proved in [LLT] that an r-regular finite simple graph with n vertices is balanced iff n is even or $r=2$. This fact also follows immediately from Lemma 2.1. In [KLST], the authors deduce the same fact from their characterization of balanced graphs.

Lemma 2.1 shows that the balancedness of a graph completely depends on the degree sequence of it. This is no longer the case for k-balanced graphs for $k \geq 3$. In fact, the trees G_1 and G_2 in Figure 3 have the same degree sequence $(1,1,1,1,1,1,1,1,1,1,1,2,2,2,11)$, and it is not difficult to see that G_1 is 3-balanced while G_2 is not.

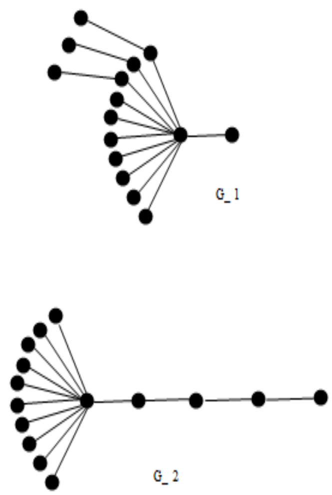

FIGURE 3. The trees G_1 and G_2 have the same degree sequence; G_1 is 3-balanced while G_2 is not.

The fact that, for $k \geq 3$, the k-balancedness is not determined by the degree sequence causes difficulties in proving that random graphs are k-balanced. It also seems plausible that, generically, k-balancedness is a weaker condition than balancedness, although it does not seem easy to describe (with a good sufficient condition) when exactly is this true. It is useful to point out the following simple fact.

Proposition 2.3. *For all distinct $m, n \geq 2$ there exists a finite simple graph which is m-balanced but not n-balanced.*

Proof. Let p be a prime number such that $p > \max\{m, n\}$.

Let us first assume that $m > n$. If n divides m, then the graph K_{m+1} is m-balanced but not n-balanced. If n does not divide m then the graph K_{mp} is m-balanced but not n-balanced.

Now assume that $m < n$. Then the graph K_{mp} is m-balanced but not n-balanced. ■

3. Combinatorial Lemmas

Let $M_n = \{\bar{d} = (d_1, \ldots, d_n) : d_i \in \mathbb{N}, 1 \le d_i \le n, 1 \le i \le n, \}$. The elements of M_n consist of sequences of positive integers of length n such that no term is bigger than n. We denote $\max(\bar{d}) = \max\limits_{1 \le i \le n} d_i$.

Now we introduce the notion of balanced sequences

Definition 3. A sequence (an element) $\bar{d} \in M_n$ is called *balanced* if and only if there exists a partition $\{1, \ldots, n\} = I \sqcup J$ such that
(a) $\big||I| - |J|\big| \le 1$
(b) $|\sum\limits_{k \in I} d_k - \sum\limits_{k \in J} d_k| \le 2$

The partition $\{1, \ldots, n\} = I \sqcup J$ will be called *a balanced partition.*

In these new terms, Lemma 2.1 states that a graph is balanced if and only if its degree sequence is balanced.

When the sequence is not balanced, we would like to measure how far it is from being balanced.

Definition 3.1. Let $\bar{d} = (d_1, \ldots, d_n)$ be any finite sequence of non-negative integers. The quantity

$$F(\bar{d}) = \min_{\{1,\ldots,n\} = I \sqcup J, \big||I|-|J|\big| \le 1} \left| \sum_{k \in I} d_k - \sum_{k \in J} d_k \right|$$

will be called *the balance of* $\bar{d}$.

Remark 3.2. By Lemma 2.1, a sequence $\bar{d} \in M_n$ is balanced if and only if $0 \le F(\bar{d}) \le 2$. The quantity $F(\bar{d})$, somewhat roughly, measures how far the sequence is from being balanced. For an example, let $n = 8$ and $\bar{d} = (1, 3, 12, 2, 1, 1, 4, 3)$ be a sequence of length 8. It is easy to see $F(\bar{d}) = |(12 + 1 + 1 + 1) - (2 + 3 + 3 + 4)| = 3$.

The following easy lemma will be useful

Lemma 3.3. *Let $\bar{d} = (d_1, \ldots, d_n)$ be any finite sequence of non-negative integers. Then $F(\bar{d}) \le \max(\bar{d})$.*

Proof. We will present a constructive proof.

Without loss of generality, we may assume that $d_1 \le d_2 \le \ldots \le d_n$. First, let us assume that n is even, so let $n = 2m$. We will build two subsets I, J of $\{1, \ldots, n\}$ inductively such that $\{1, \ldots, n\} = I \sqcup J, |I| = |J|$, and $|\sum_{k \in I} d_k - \sum_{k \in J} d_k| \le \max(\bar{d})$.

We divide the sequence into pairs $(d_1, d_2), \dots, (d_{2m-1}, d_{2m})$, and we will abide by the rule that exactly one element of each pair belongs to I and the other element belongs to J. We start by letting $I_1 = \{d_{2m}\}, J_1 = \{d_{2m-1}\}$. Assume now we have built the subsets $I_k, J_k, 1 \le k \le m-1$ such that $\{d_{2m}, d_{2m-1}, \dots, d_{2m-2k+2}, d_{2m-2k+1}\} = I_k \sqcup J_k$ and $|\{d_{2m-2i-2}, d_{2m-2i-1}\} \cap I_k| = 1$ for all $1 \le i \le k$.

Let $S(I_k) = \sum_{i \in I_m} d_i, S(J_k) = \sum_{I \in J_m} d_i$. If $S(I_k) > S(J_k)$ then we let $I_{k+1} = I_k \sqcup \{d_{2m-2k-1}\}, J_{k+1} = J_k \sqcup \{d_{2m-2k}\}$ but if $S(I_k) \le S(J_k)$ then we let $I_{k+1} = I_k \sqcup \{d_{2m-2k}\}, J_{k+1} = J_k \sqcup \{d_{2m-2k-1}\}$, and we proceed by induction. Then we let $I = I_m, J = J_m$. Clearly, we have $F(\bar{d}) \le |\sum_{k \in I} d_k - \sum_{k \in J} d_k| \le \max(\bar{d})$.

If n is odd, then we may replace $\bar{d}$ by $\bar{d}' = (0, d_1, \dots, d_n)$ and apply the previous argument. ■

We will need the following notations

Definition 3.4. Let $\bar{d} = (d_1, \dots, d_n) \in M_n$. We will denote

$$u(\bar{d}) = \{1 \le i \le n \mid d_i = 1\}, \; v(\bar{d}) = \{1 \le i \le n \mid d_i = 2\}$$

Lemma 3.5. *Let $\bar{d} = (d_1, \dots, d_n) \in M_n$ such that $|u(\bar{d})| \ge \max(\bar{d})$ and $|v(\bar{d})| \ge \max(\bar{d})$. Then $\bar{d}$ is balanced.*

Proof. Let $\max(\bar{d}) = m$. Without loss of generality we may assume that $d_1 = \dots = d_m = 1, d_{m+1} = \dots = d_{2m} = 2$. If $n = 2m$ then $\bar{d}$ is clearly balanced so let $n > 2m$ and let $\bar{d}' = (d_{2m+1}, \dots, d_n)$.

By Lemma 3.3, $F(\bar{d}') \le m$ hence there exists a partition $\{d_{2m+1}, \dots, d_n\} = I' \sqcup J'$ such that $\big||I'| - |J'|\big| \le 1$ and $|\sum_{k \in I'} d_k - \sum_{k \in J'} d_k| \le m$. Then there exists a partition $\{d_1, \dots, d_{2m}\} = I'' \sqcup J''$ such that $|I''| = |J''|$ and $|(\sum_{k \in I''} d_k - \sum_{k \in J''} d_k) - (\sum_{k \in I'} d_k - \sum_{k \in J'} d_k)| \le 2$. By letting $I = I' \sqcup I'', J = J' \sqcup J''$ we obtain that $\{1, \dots, n\} = I \sqcup J$, $\big||I| - |J|\big| \le 1$, and $|\sum_{k \in I} d_k - \sum_{k \in J} d_k| \le 2$. ■

4. Proof of Theorem A

First, we will discuss the case of labeled trees. The following theorem of J.W.Moon will play a crucial role

Theorem 4.1 (See (M))**.** *If $\epsilon > 0$ is a fixed positive constant, then in a random labeled tree G with n vertices, the maximum degree $\Delta(G)$ satisfies the following inequality*

$$(1-\epsilon)\frac{\log n}{\log\log n} < \Delta(G) < (1+\epsilon)\frac{\log n}{\log\log n}$$

Remark 4.2. By choosing $\epsilon = 0.1$ we obtain that

$$0.9\frac{\log n}{\log\log n} < \Delta < 1.1\frac{\log n}{\log\log n}$$

in a random tree with n vertices.

We will use only the upper bound in the inequality of Remark 4.2. Besides the upper bound on the maximum degree in random trees, we also need a lower bound on the number of vertices with degree 1, and with degree 2. Notice that, since the sum of degrees of a tree with n vertices is exactly $2n-2$, at least half of the vertices have degree either 1 or 2. However, we need a linear lower bound for the number of vertices of degree 1 and for the number of vertices of degree 2 separately.

Let $X_i(T), 1 \leq i \leq 2$ be the random variable which denotes the number of vertices of degree i in a labeled tree T with n vertices. Also let $\mu = \frac{n}{e}, \sigma_1^2 = \frac{n}{e}(1-\frac{2}{e}), \sigma_2^2 = \frac{n}{e}(1-\frac{1}{e})$. It has been proved by A. Rényi (see [R]) that the asymptotic distribution of random variable $\frac{X_1-\mu}{\sigma_1}$ is normal with mean μ and variance σ_1^2. A similar result has been proved for the random variable $\frac{X_2-\mu}{\sigma_2}$, by A. Meir and J.W. Moon (see [MM]), namely, that the asymptotic distribution of the random variable $\frac{X_2-\mu}{\sigma_2}$ is normal with mean μ and variance σ_2^2. Combining these two results we can state the following theorem (due to A. Rényi and A. Meir-J.W. Moon)

Theorem 4.3. *Let α, β be fixed real numbers, $\alpha < \beta$; and for $i \in \{1, 2\}$, let $P_i(\alpha, \beta)$ denotes the probability that $\alpha < \frac{X_i-\mu}{\sigma_1} < \beta$. Then*

$$\lim_{n\to\infty} P_i(\alpha, \beta) = \frac{1}{\sqrt{2\pi}}\int_\alpha^\beta e^{-\frac{1}{2}t^2}dt$$

We need the following immediate corollary of this theorem

Corollary 4.4. In a random labeled tree with n vertices, for all $i \in \{1, 2\}$, $X_i \geq 2\frac{\log n}{\log\log n}$.

Now, in the case of random labeled trees, Theorem A immediately follows from Theorem 4.1, Corollary 4.4, and Lemma 3.5.

The case of unlabeled trees: We will use the results analogous to Theorem 4.1 and Theorem 4.3. The analogue of Theorem 4.1 is proved by W. Goh and E. Schmutz:

Theorem 4.5 (See [GoSh)**.** *] There exists positive constants c_1, c_2 such that in a random unlabeled tree T with n vertices, the maximum degree $\Delta(T)$ satisfies the inequality $c_1\log(n) < \Delta(T) < c_2\log(n)$.*

Now, for any $k \in \mathbb{N}$ let the random variable Y_k denotes the number of vertices of degree k in a random unlabeled tree with n vertices. The following theorem is due to M. Drmota and B. Gittenberger; in the case of $k \in \{1, 2\}$, as a special case, it provides an analogue of Theorem 4.3.

Theorem 4.6 (See [DG)**.** *] For arbitrary fixed natural k, there exist positive constants μ_k and σ_k such that the limiting distribution of Y_k is normal with mean $\mu(n) \sim \mu_k n$ and variance $\sigma(n) \sim \sigma_k^2 n$.*

Corollary 4.7. For all $c > 0$ and $i \in 1, 2$, in a random unlabeled tree with n vertices $Y_i > c\log(n)$.

Now, in the case of unlabeled trees, the claim of Theorem A follows from Theorem 4.5, Lemma 3.5, and Corollary 4.7.

5. k-balanced trees: proof of Theorem B

In this section we will assume that $k \geq 3$. The fact that the k-balancedness is not determined by the degree sequence causes significant difficulties in proving that random graphs are balanced. We nevertheless prove that random trees are strongly k-balanced by more careful study of k-balancedness.

First, we need to prove the following technical lemma.

Lemma 5.1. *Let $G = (V, E)$ be a tree and u, v be distinct vertices of G with degree at least $\frac{|G|}{3}$. If p, q are distinct pre-leaf vertices of G then there exists a strongly 3-balanced coloring $c : V \to \{1, 2, 3\}$ of G such that $c(u) \neq c(v)$ and $c(p) \neq c(q)$.*

Proof. The proof is by induction on $n = |G|$. For $n \leq 5$ the claim is obvious (since, in this case, G will be isomorphic either to a path or to the double star $S_{3,2}$) so we will assume that $n \geq 6$ and the claim holds for all trees of order less than n.

Assume that at least one of the following two conditions hold:
(c1) there exists $z \in \{p, q\} \backslash \{u, v\}$ such that $deg(z) \geq 3$;
(c2) there exists a leaf vertex not adjacent to any of the vertices u, v, p, q.
Then there exists a leaf w such that if G' is a full subgraph on $V \backslash \{w\}$, then, in the tree G', we have $\min\{deg(u), deg(v)\} \geq \frac{|G'|}{3}$, and p, q are still pre-leaf vertices.

By inductive hypothesis, there exists a strongly 3-balanced coloring $c_0 : V \backslash \{w\} \to \{1, 2, 3\}$ of G' such that $c_0(u) \neq c_0(v)$ and $c_0(p) \neq c_0(q)$. Let w_0 be the unique vertex of G adjacent to w. Without loss of generality, we may assume that $c_0(w_0) = 1$ and $|c_0^{-1}(2)| \leq |c_0^{-1}(3)|$.

If $|c_0^{-1}(1)| \geq |c_0^{-1}(2)|$ then we let $c(w) = 2$ thus extending c_0 to a strongly 3-balanced coloring $c : V \to \{1, 2, 3\}$ of G' such that $c(u) \neq c(v)$ and $c(p) \neq c(q)$.

If, however, $|c_0^{-1}(1)| < |c_0^{-1}(2)|$ then there exists $r \in \{u, v\}$ such that $c_0(r) \neq 1$; also, since $deg(r) \geq \frac{|G|}{3}$, there exists a branch B of G' with respect to r which is disjoint from $c_0^{-1}(1)$. Let x be a leaf vertex in B. Then $x \notin \{u, v, p, q\}$ and $c_0(x) \neq 1$. We define $c : V \to \{1, 2, 3\}$ as follows:

$$c(\omega) = \begin{cases} c_0(\omega) & \text{if } \omega \in V \backslash \{w, x\} \\ 1 & \text{if } \omega = x \\ c_0(x) & \text{if } \omega = w \end{cases}$$

Notice that because of the inequality $|c_0^{-1}(1)| < |c_0^{-1}(2)| \leq |c_0^{-1}(3)|$, we have $|c_0^{-1}(2)| = |c_0^{-1}(3)|$ and $|c_0^{-1}(1)| = |c_0^{-1}(2)| - 1$. Then the map $c : V \to \{1, 2, 3\}$ is a strongly 3-balanced coloring.

Now, suppose that neither of the conditions (c1) and (c2) hold. Let P be the path in G starting at u and ending at v (it may possibly consist of just the vertices u and v). Then the tree G satisfies the following conditions: there exists two vertices z_1, z_2 in P and paths R_1, R_2 starting at z_1, z_2 respectively such that any vertex of G either belongs to one of the paths P, R_1, R_2 or it is a leaf vertex adjacent to one of the vertices u, v. Then it is straightforward to build a strongly 3-balanced coloring $c : V \to \{1, 2, 3\}$ satisfying the conditions $c(u) \neq c(v)$ and $c(p) \neq c(q)$. ■

The following proposition is interesting in itself; it will also play a key role in proving Theorem B.

Proposition 5.2. *If $G = (V, E)$ is a tree with $\Delta(G) \le \frac{n}{3}$ then G is strongly 3-balanced. Moreover, for any two distinct pre-leaf vertices p and q of G there exists a strongly 3-balanced coloring $c : V \to \{1, 2, 3\}$ such that $c(p) \ne c(q)$.*

Proof. The proof will be by induction on $n = |G|$. For $n \le 8$ we have $\Delta(G) \le 2$ hence G is isomorphic to a path. Thus, the claim is obvious. Let us now assume that $n \ge 9$, and the claim holds for all trees G' of order less than n with $\Delta(G') \le \frac{|G'|}{3}$.

Let $G = (V, E)$ and $n = 3k + r, r \in \{0, 1, 2\}$. We will consider the following three cases separately:

Case 1. $r = 1$.

Let v be a leaf of G, $V' = V\backslash\{v\}$, and let $G' = (V', E')$ be the full subgraph of G on V'. Then we have

$$\Delta(G') \le \Delta(G) \le k \le \frac{|G'|}{3}.$$

By inductive hypothesis, there exists a strongly 3-balanced coloring $c' : V' \to \{1, 2, 3\}$ of G'.

On the other hand, v is adjacent to exactly one vertex in G; let u be this vertex. Let j be any element of $\{1, 2, 3\}\backslash\{c'(u)\}$. We extend the coloring c' of G' to a strongly 3-balanced coloring $c : V \to \{1, 2, 3\}$ by defining $c(v) = j$.

Case 2. $r = 2$.

Let v_1, v_2 be distinct leaves and u_1, u_2 be the only vertices of G adjacent to v_1, v_2 respectively (u_1 and u_2 are not necessarily distinct). Let also G' be the full subgraph of G on the set $V\backslash\{v_1, v_2\}$. Then we still have the inequality $\Delta(G') \le \Delta(G) \le k \le \frac{|G'|}{3}$. Hence, by inductive assumption, there exists a strongly 3-balanced coloring $c' : V' \to \{1, 2, 3\}$ of G'.

Then there exist distinct $j_1, j_2 \in \{1, 2, 3\}$ such that $j_1 \ne c'(u_1)$ and $j_2 \ne c'(u_2)$. Thus we can extend c' to a strongly 3-balanced coloring of G by defining $c(v_1) = j_1$ and $c(v_2) = j_2$.

Case 3. $r = 0$.

The major difference in this case compared with the previous two cases is that when we obtain G' by deleting some arbitrary three leaves v_1, v_2, v_3 from G, we may lose the inequality $\Delta(G') \le \frac{|G'|}{3}$. (Notice that G possesses three leaf vertices unless it is isomorphic to a path). Suppose u_1, u_2, u_3 are the vertices adjacent to v_1, v_2, v_3 respectively. Note that u_1, u_2, u_3 are not necessarily distinct. If we have the inequality $\Delta(G') \leq \frac{|G'|}{3}$ then by inductive assumption we would have a strongly 3-balanced coloring $c' : V \backslash \{v_1, v_2, v_3\} \to \{1, 2, 3\}$. However, if $c'(u_1) = c'(u_2) = c'(u_3)$ then it becomes problematic to extend c' to a strongly 3-balanced coloring $c : V \to \{1, 2, 3\}$. Thus we need to employ different and more careful tactics.

We will prove the following lemma which suffices for the proof of Proposition 5.2 in the case $r = 0$.

Lemma 5.3. *Let $G = (V, E)$ be a tree with $n = 3k$ vertices where $\Delta(G) \leq k$. If p, q are distinct pre-leaf vertices of G then there exists a strongly 3-balanced coloring $c : V \to \{1, 2, 3\}$ such that $c(p) \neq c(q)$.*

Proof. The proof of the lemma will be again by induction on k. The "$c(p) \neq c(q)$ part" of the claim will be needed to make the step of the induction. For $k \leq 2$, the graph G is isomorphic to a path thus the claim is obvious. For $k = 3$ it can be seen by a direct checking. (We leave this to a reader as a simple exercise.) Thus let us assume that $k \geq 4$.

Let $W = \{v \in V \mid d(v) = k\}$. Let also $deg(p) \leq deg(q)$. We will consider the following cases (the notations in each case will be independent of the notations of other cases):

Case A: The vertices p and q are the only pre-leaf vertices of G.

The claim is obvious when $deg(q) \leq 3$, so we may assume that $deg(q) \geq 4$. We will consider two sub-cases:

Sub-case 1: $deg(p) \geq 3$.
Then there exist leaves w_1, w_2, w_3 such that w_1 is adjacent to p and w_2, w_3 are adjacent to q. Let G' be the full subgraph of G on $V' = V \backslash \{w_1, w_2, w_3\}$. Notice that p and q are still pre-leaf vertices of G'. By the inductive hypothesis, G' has a strongly 3-balanced coloring $c' : V' \to \{1, 2, 3\}$ such that $c'(p) \neq c'(q)$. We may assume that $c'(p) = 1, c'(q) = 2$. Then we extend c' to a strongly 3-balanced coloring $c : V \to \{1, 2, 3\}$ by letting $c(w_1) = 2, c(w_2) = 1, c(w_3) = 3$.

Sub-case 2: $deg(p) = 2$.
In this case there exists a path $(w, p, v_1, \ldots, v_r, q)$ in G where w is a leaf, $p, v_1, \ldots, v_r$ are vertices of degree two, and

$$V = \{w, p, v_1, \ldots, v_r, q, u_1, \ldots, u_s\}$$

where $u_1, \dots, u_s$ are the leaves adjacent to q. Then we have the inequalities $s = deg(q) - 1 \leq k - 1$ and $r \geq 2k - 2$. We will construct the required strongly 3-balanced coloring explicitly as follows.

First, for all $1 \leq i \leq s$ we let $c(u_i) = 1$ when i is odd, and $c(u_i) = 3$ when i is even. We also let $c(q) = 2$ and $c(v_{r-2i+1}) = 2, 1 \leq i \leq k - 1$.

Now we need to define the coloring on the remaining set

$$D = \{w, p\} \cup (\{v_1, \dots, v_r\} \backslash \{v_{r-2i+1} \mid 1 \leq i \leq k - 1\}).$$

Let $D = \{x_1, \dots, x_{n-s-k}\}$ where $d(x_i, w) < d(x_j, w)$ for all $1 \leq i < j \leq n - s - k$. (Thus we are reorder the elements of the set D from closest to the farthest from the leaf w.) Then, for all $1 \leq i \leq n - s - k$, we let $c(x_i) = 3$ when i is odd, and $c(x_i) = 1$ when i is even.

For the rest of the proof we will assume that G has more than two pre-leaf vertices.

Case B: $W = \emptyset$ and p is not special.

Let v_1, v_2, v_3 be distinct leaves such that v_1 is adjacent to p, v_2 is adjacent to q, and v_3 is adjacent to a vertex w distinct from p and q. We let G' be the full subgraph on $V \backslash \{v_1, v_2, v_3\}$. Then $|G'| = 3(k - 1)$ and we have $\Delta(G') \leq k - 1$. By inductive hypothesis, there exists a strongly 3-balanced coloring $c_0 : V \to \{1, 2, 3\}$ such that $c_0(p) \neq c_0(q)$. Without loss of generality we may assume that $c_0(p) = 1, c_0(q) = 2$. Then we extend c_0 to a strongly 3-balanced coloring $c : V \to \{1, 2, 3\}$ as follows: if $c_0(w) \in \{1, 2\}$ then we let $c(v_1) = 2, c(v_2) = 1, c(v_3) = 3$; and if $c_0(w) = 3$ then we let $c(v_1) = 3, c(v_2) = 1, c(v_3) = 2$.

Case C: $W = \emptyset$ and p is special.

Let v_1 be the only leaf adjacent to p, u be the unique non-leaf vertex adjacent to p, v_2 be a leaf vertex not adjacent to u, and w be the unique vertex adjacent to v_2. We let G' be the full subgraph on $V \backslash \{v_1, v_2, p\}$. Then $|G'| = 3(k - 1)$ and $\Delta(G') \leq k - 1$. By inductive hypothesis, there exists a strongly 3-balanced coloring $c_0 : V \to \{1, 2, 3\}$. Then we extend c_0 to a strongly 3-balanced coloring $c : V \to \{1, 2, 3\}$ as follows: we let $c(p) \in \{1, 2, 3\}$ such that $c(p)$ is distinct from $c_0(u)$ and $c_0(q)$. Then we define $c(v_2) \in \{1, 2, 3\}$ such that $c(v_2)$ is distinct from $c_0(w)$ and $c(p)$. Finally we let $c(v_1) \in \{1, 2, 3\}$ such that $c(v_1)$ is distinct from $c(p)$ and $c(v_2)$. Notice also that we obtain $c(p) \neq c(q)$.

Case D: $|W| = 1, W = \{v_0\}, deg(p) \geq 3$ and there exists a leaf vertex adjacent to v_0.

This case is similar to Case B. Since $|W| = 1$ and $deg(p) \leq deg(q)$, we have $p \neq v_0$. If $q \neq v_0$, we let v_1, v_2, v_3 be leaves adjacent to p, q, v_0 respectively; and if $q = v_0$, we let v_1, v_2 be leaves adjacent to p, q respectively, and v_3 be a leaf not adjacent to either of the vertices p, q. We define G' be the full subgraph on $V \backslash \{v_1, v_2, v_3\}$. Then $\Delta(G') \leq \frac{|G'|}{3}$ hence G' admits a strongly 3-balanced coloring $c' : V \backslash \{v_1, v_2, v_3\} \rightarrow \{1, 2, 3\}$ such that $c'(p) \neq c'(q)$. We extend c' to a strongly 3-balanced coloring to $c : V \rightarrow \{1, 2, 3\}$ as in Case B.

Case E: $|W| = 1, W = \{v_0\}$, p is special and there exists a leaf vertex adjacent to v_0.

This case is similar to Case C. Let v_1 be the only leaf adjacent to p, u be the unique non-leaf vertex adjacent to p, v_2 be a leaf vertex adjacent to v_0. We let G' be the full subgraph on $V \backslash \{v_1, v_2, p\}$. Then $|G'| = 3(k-1)$ and $\Delta(G') \leq k-1$. By inductive hypothesis, there exists a strongly 3-balanced coloring $c_0 : V \backslash \{v_1, v_2, p\} \rightarrow \{1, 2, 3\}$. Then we extend c_0 to a strongly 3-balanced coloring $c : V \rightarrow \{1, 2, 3\}$ as follows: we let $c(p) \in \{1, 2, 3\}$ such that $c(p)$ is distinct from $c_0(u)$ and $c_0(q)$. Then we define $c(v_2) \in \{1, 2, 3\}$ such that $c(v_2)$ is distinct from $c_0(v_0)$ and $c(p)$. Finally we let $c(v_1) \in \{1, 2, 3\}$ such that $c(v_1)$ is distinct from $c(p)$ and $c(v_2)$.

Case F: $|W| = 1, W = \{v_0\}$, and there is no leaf vertex adjacent to v_0.

Since $|V| = 3k$, there exists a special vertex v adjacent to v_0. Let v_1 be the unique leaf adjacent to v. Let also v_2 be a leaf not adjacent to any of the vertices p, q, v (such a leaf exists because $k \geq 4$), and let w be the unique vertex adjacent to v_2.

We define G' to be the full subgraph on $V \backslash \{v, v_1, v_2\}$. By inductive assumption, there exists a strongly 3-balanced coloring $c_0 : V \backslash \{v, v_1, v_2\} \rightarrow \{1, 2, 3\}$, moreover, if $p, q \in V \backslash \{v, v_1, v_2\}$ then $c_0(p) \neq c_0(q)$.

If $\{p, q\} \subset V \backslash \{v, v_1, v_2\}$, then we let $c(v_2)$ be any element of $\{1, 2, 3\}$ distinct from $c_0(w)$. Then we let $c(v)$ be any element of $\{1, 2, 3\}$ distinct from $c_0(v_0)$ and $c(v_2)$. Finally, we let $c(v_1)$ be any element of $\{1, 2, 3\}$ distinct from $c(v)$ and $c(v_2)$. Thus we have extended c_0 to a strongly 3-balanced coloring $c : V \rightarrow \{1, 2, 3\}$ such that $c(p) \neq c(q)$.

If $\{p, q\} \cap \{v, v_1, v_2\} \neq \emptyset$ then $\{p, q\} \cap \{v, v_1, v_2\} = \{v\}$ and we may assume that $p = v$. Then we let $c(v)$ be any element of $\{1, 2, 3\}$ distinct from $c_0(v_0)$ and $c_0(q)$; then we let $c(v_2)$ be any element of $\{1, 2, 3\}$

distinct from $c_0(w)$ and $c(v)$; finally we let $c(v_1)$ be any element of $\{1, 2, 3\}$ distinct from $c(v)$ and $c(v_2)$.

Case G: $|W| \geq 2$.

In this case the claim follows immediately from Lemma 5.1. ■

Now we can prove an analogous result for k-balanced graphs.

Proposition 5.4. *Let $G = (V, E)$ be a tree with n vertices where $\Delta(G) \leq \frac{n}{k}$ and $k \geq 3$. Then G is strongly k-balanced.*

Proof. The proof is by induction on k. For $k = 3$, the claim is true by Proposition 5.2.

Assume now $k \geq 4$. The tree G has $m = \lfloor \frac{n}{k} \rfloor$ vertices $v_1, \ldots, v_m$ such that $d(v_i) \leq 2, 1 \leq i \leq m$. Moreover, for all distinct $i, j \in \{1, \ldots, m\}$, the vertices v_i and v_j are not connected by an edge. Let also $V_0 = \{v_1, \ldots, v_m\}$, and G_1 be a full subgraph on the subset $V \backslash V_0$. Then G_1 is a forest with $n - m$ vertices but with $\Delta(G_1) \leq \Delta(G)$. This implies that G_1 is a subgraph of a tree G_2 with $n - m$ vertices where $\Delta(G_2) \leq \Delta(G)$.

Then $\Delta(G_2) \leq \Delta(G) \leq \frac{n}{k} = \frac{1}{k-1}(n - \frac{n}{k}) \leq \frac{1}{k-1}(n - m) \leq \frac{|G_2|}{k-1}$. By inductive hypothesis, we obtain that G_2 is strongly $(k-1)$-balanced, hence G_1 is strongly $(k-1)$-balanced. Since no two elements of V_0 are adjacent, we obtain that G is strongly k-balanced. ■

Now, for random labeled trees, Theorem B follows immediately from Theorem 4.1 and Proposition 5.4; and for random unlabeled trees, it follows immediately from Theorem 4.5 and Proposition 5.4.

R e f e r e n c e s:

[BK] Ben-Eliezer, I. and Krivelevich, M. Perfectly balanced partitions of smoothed graphs. *Electronic Journal of Combinatorics*, **16** (1) (2009), note N14.

[B] Bollobás, B. Random Graphs. Academic Press Inc. (1985)

[BG] Bollobás, B. and Guy, R. Equitable and proportional coloring of trees, *Journal of Combinatorial Theory, Series B* **34**, (2): (1983) 177186

[Cah1] Cahit, I. Cordial graphs: a weaker version of graceful and harmonious graphs, *Ars Combinatoria*, **23** (1987) 201-207.

[Cah2] Cahit, I. On cordial and 3-equitable graphs, *Utilitas Mathematica*, **37** (1990) 189-198.

[Cay] Cayley, A. A Theorem on Trees, *Quart.J.Math*, **23**, (1889) 376-378.

[DG] Drmota, M. and Gittenberger, B. Distribution of nodes of given degree in random trees. *Journal of Graph Theory,* **31/3** (1999), 227-253.

[Ga] Gallian, J.A. A dynamical survey of graph labeling. *The Electronics Journal of Combinatorics*, **16** (2009)

[GoSh] Goh, W. and Schmutz, E. Unlabeled trees: distribution of the maximum degree. *Random Structures and Algorithms*, **5(1)**, (1994), 13-24.

[GrSl] Graham, R. and Sloane, N. On additive bases and harmonious graphs. *SIAM Journal of Algebraic and Discrete Mathematics*, **1**, (1980) 382-404.

[KLST] Kong, M.C, Lee, S-M, Seah, E, and Tang, A. A complete characterization of balanced graphs. (English summary) *J. Combin. Math. Combin. Comput.* **66** (2008), 225-236.

[LLT] Lee, S-M, Liu, A, and Tan, S.K. On balanced graphs, *Congr. Numer.*, **87** (1992), 59-64.

[M] Moon, J.W. On the maximum degree in a random tree. *Michigan Math. Journal*, Volume 15, Issue 4, (1968), 429-432.

[MM] Meir, A. and Moon, J.W. On nodes of degree two in random trees. *Mathematika*, iss.2, **vol. 15** (1968), 188-192.

[O] Otter, R. The number of trees. *Annals of Mathematics*, Second Series, **49** (3), 583-599.

[W] West, D. B. *Introduction to Graph Theory*, 2nd ed. Prentice-Hall, Inc., Upper Saddle River NJ (2001), 82-83.

Azer Akhmedov, Department of Mathematics, North Dakota State University, Fargo, ND, 58102, USA
E-mail address: azer.akhmedov@ndsu.edu

Warren Shreve, Department of Mathematics, North Dakota State University, Fargo, ND, 58102, USA
E-mail address: warren.shreve@ndsu.edu