Bull. Korean Math. Soc. **33** (1996), No. 1, pp. 119–126

ON THE EDGE INDEPENDENCE NUMBER OF A RANDOM (N,N)–TREE

J. H. Cho and Moo Ha Woo

1. Introduction

In this paper we study the asymptotic behavior of the edge independence number of a random (n, n)-tree. The tools we use include the matrix-tree theorem, the probabilistic method and Hall's theorem. We begin with some definitions. An (n, n)-tree T is a connected, acyclic, bipartite graph with n light and n dark vertices (see [Pa92]). A subset M of edges of a graph is called *independent(or matching)* if no two edges of M are adjacent. A subset S of vertices of a graph is called *independent* if no two vertices of S are adjacent. The *edge independence number* of a graph T is the number $\beta_1(T)$ of edges in any largest independent subset of edges of T. Let $\Gamma(n, n)$ denote the set of all (n, n)-trees with n light vertices labeled 1, ..., n and n dark vertices labeled 1, ..., n. We give $\Gamma(n, n)$ the uniform probability distribution. Our aim in this paper is to find bounds on $\beta_1(T)$ for a random (n, n)-tree T in $\Gamma(n, n)$.

The matrix-tree theorem originated in the work of Kirkhoff (see Moon [M70] p.42) and relates the number of spanning trees of a labeled graph to the adjacency matrix. We now apply the theorem to a simple but important family of graphs. Consider a graph, denoted $G(V_1, V_2, V_3, V_4)$, whose vertex set V is partitioned into four nonempty sets:

$$V = V_1 \dot\cup V_2 \dot\cup V_3 \dot\cup V_4.$$

The edge set of $G(V_1, V_2, V_3, V_4)$ consists of all edges joining vertices of V_1 to vertices of V_2, vertices of V_2 to vertices of V_3 and vertices of V_3 to vertices of V_4.

Received September 24, 1994. Revised May 1, 1995.
1991 AMS Subject Classification: 05C80.
Key words and phrases: Independence Number, Probabilistic Method.
The present studies were supported in part by TGRC-KOSEF and BSRI 1409.

COROLLARY 1.1. *The number of spanning trees of* $G(V_1, V_2, V_3, V_4)$ *is*

$$(n_1+n_3)^{n_2-1}(n_2+n_4)^{n_3-1}n_2^{n_1}n_3^{n_4},$$

where $n_i = |V_i|$ *for* $i = 1$ *to 4.*

This formula can be established by applying row and column operations to calculate a cofactor of the required matrix. It can also be realized as a corollary of a very broad result of Knuth [Kn68] on generalized Prüfer codes.

The probabilistic method was first applied to graphs by Erdös [Er47], who pioneered its use with so many innovations that it may be more proper to call it *the Erdös method* (see [AlSE92]). Here we sketch only the portion of the method that we require. For background on probability theory for a discrete sample space, one can consult [AlSE92], [Bo85] or the appendix of [Pa85].

THEOREM 1.1 (MARKOV'S INEQUALITY). *Let* $X \geq 0$ *be a random variable and let* $t > 0$. *Then*

$$P(X \geq t) \leq \frac{E(X)}{t}. \tag{1.1}$$

On setting $t = 1$ in inequality (1.1), we have

$$P(X \geq 1) \leq E(X). \tag{1.2}$$

If X is non negative and integer valued, we also have

$$P(X = 0) + P(X \geq 1) = 1. \tag{1.3}$$

In our applications, the sample space always consists of graphs with at least n vertices and the random variable X counts certain types of subgraphs. It follows from (1.2) and (1.3) that if

$$E(X) \to 0 \text{ as } n \to \infty,$$

then

$$P(X \geq 1) \to 0$$

and

$$P(X = 0) \to 1$$

and we say "almost all graphs have no such subgraph". Suppose the sample space consists of (n, n)-trees and X counts sets of vertices which are undesirable, i.e. "bad sets". If we show $E(X) \to 0$, then we say "almost all trees have no bad sets".

In a graph T, a nonempty subset U_1 of $V(T)$, the vertex set of T, is said to be *matched* to a subset U_2 of $V(T)$, which is disjoint from U_1, if there exists an independent edge set M of T such that each edge of M is incident with a vertex of U_1 and a vertex of U_2 and every vertex of U_1 is incident with an edge of M, as is every vertex in U_2. Let U be a nonempty subset of $V(T)$ and its neighborhood $N(U)$ denote the set of all vertices of T adjacent with at least one element of U. Then the set U is said to be *nondeficient* if

$$|N(S)| \geq |S|,$$

for every nonempty subset S of U. The next theorem attributed to Hall[Ha35] allows us a useful corollary.

THEOREM 1.2. *Let T be a bipartite graph with partite sets V_1 and V_2. The set V_1 can be matched to a subset of V_2 if and only if V_1 is nondeficient.*

COROLLARY 1.2. *Let T be a bipartite graph with partite sets V_1 and V_2 and let d be a positive integer. Then*

$$\beta_1(T) \geq |V_1| - d \tag{1.4}$$

if and only if

$$|N(S)| \geq |S| - d, \tag{1.5}$$

for any subset S of V_1.

Proof. Construct a bipartite graph T_d from T by adding d new vertices to V_2 and all possible edges between V_1 and the set of d new vertices. Then the theorem is applied to T_d. □

2. Edge independence number for (n,n)-trees

Our aim is to determine bounds for β_1 for almost all (n,n)-trees in $\Gamma(n,n)$. Of course, we always have the upper bound

$$\beta_1 \leq n,$$

for all trees in $\Gamma(n,n)$. To find a lower bound we go back to Corollary 1.2. Let S be a subset of V_1 of cardinality k. We call S a bad $k-set$ if

$$|N(S)| \leq k-d-1,$$

for a given positive integer d. Let X_d be the number of bad sets in $\Gamma(n,n)$. To show there is a matching of cardinality $|V_1|-d$ for almost all trees in $\Gamma(n,n)$, we may show

$$E(X_d) \to 0 \text{ as } n \to \infty.$$

Now let A_d be the set of all trees in $\Gamma(n,n)$ for which a specified set of k light and $n-(k-d-1)$ dark vertices are independent. Then $|A_d|$ is the number of spanning trees of $G(V_1, V_2, V_3, V_4)$ in Corollary 1.1, with $|V_1| = n-(k-d-1)$, $|V_2| = n-k$, $|V_3| = k-d-1$ and $|V_4| = k$. By Corollary 1.1,

$$|A_d| = n^{(n-k)-1} n^{(k-d-1)-1} (n-k)^{n-(k-d-1)} (k-d-1)^k. \tag{2.6}$$

Since $\Gamma(n,n)$ is a sample space with the uniform probability distribution, $P(A_d)$ is the ratio of $|A_d|$ to $|\Gamma(n,n)|$. Hence

$$P(A_d) = n^{n-d-3}(n-k)^{n-k+d+1}(k-d-1)^k / n^{2n-2} \tag{2.7}$$

and

$$E(X_d) = \sum_{d+1<k<n} \binom{n}{k}\binom{n}{k-d-1} P(A_d), \tag{2.8}$$

where $\binom{n}{k}$ is the number of ways of choosing k light vertices and $\binom{n}{k-d-1}$ is the number of ways of choosing $k-d-1$ dark ones.

Also because of symmetry in $n-k$ and $k-d-1$ in the formula (2.7) for $P(A_d)$ and in (2.8) for $E(X_d)$, we have the following upper bound for $E(X_d)$.

$$E(X_d) \leq 2 \sum_{d+1<k<(n+d+1)/2} \binom{n}{k}\binom{n}{k-d-1} P(A_d). \tag{2.9}$$

Next we simplify (2.9) using Stirling's formula for $n!$, and the result is

$$E(X_k) = O(1)n^{n-d} \sum_{d+1<k<(n+d+1)/2} a_k, \tag{2.10}$$

where

$$a_k = (n-k)^{d+1/2}(k-d-1)^{d+1/2}/(n-k+d+1)^{n-k+d+3/2}k^{k+1/2}. \tag{2.11}$$

To see the behavior of the series, we investigate the ratio, a_{t+1}/a_t and find that the series increases all the way to the very last term, which is a_{k_0}, where

$$k_0 = \lfloor (n+d+1)/2 \rfloor. \tag{2.12}$$

Now the sum in (2.11) is bounded by the product of the last term and the length of the sum. Hence we have

$$E(X_d) = O(1)n^{n-d}(n-d)a_{k_0}, \tag{2.13}$$

where the factor $(n-d)$ is contributed by the length of the sum. On the other hand,

$$a_{k_0} = O(1)a_{(n+d)/2}. \tag{2.14}$$

Next we want to determine d as a function of n, so that,

$$E(X_d) \to 0 \text{ as } n \to \infty. \tag{2.15}$$

Then we can say "almost all (n,n)-trees have an independent set of order at least $n-d$". This idea is also expressed by saying *almost all (n,n)-trees T have*

$$\beta_1(T) \geq n-d. \tag{2.16}$$

Naturally the function just mentioned should be as small as possible. We set

$$d = \alpha n, \tag{2.17}$$

where α is a positive constant, and we will try to determine α as small as possible so that (2.16) holds. Note that the right side of (2.17) is not necessarily an integer. We often use non-integral quantities where we ought to round up or down. It should be clear that such deviations do not affect the validity of the results.

And a bit of algebra shows that

$$a_{(n+d)/2} = O(1)n^{-(n-\alpha n+2)}(2^{1-\alpha}(1-\alpha)^{2\alpha}/(1+\alpha)^{1+\alpha})^n. \tag{2.18}$$

Substituting (2.14) and (2.18) in the equation (2.13), we find that the upper bound for the expectation takes the following simple form:

$$E(X_k) = O(1)(2^{1-\alpha}(1-\alpha)^{2\alpha}/(1+\alpha)^{1+\alpha})^n. \tag{2.19}$$

Since we want the right side of (2.19) to approach to zero as $n \to \infty$, we just need to solve the inequality

$$2^{1-\alpha}(1-\alpha)^{2\alpha} < (1+\alpha)^{1+\alpha}. \tag{2.20}$$

A simple numerical calculation shows that it is sufficient to choose

$$\alpha = .27974. \tag{2.21}$$

These observations are summerized in the following theorem.

THEOREM 2.1. *For almost all (n,n)-trees T, the edge independence number $\beta_1(T)$ satisfies the inequality:*

$$.72026n \leq \beta_1(T) \leq n. \tag{2.22}$$

Our exact calculations in [ChP95] indicate that the expected value of β_1 is about

$$(.4385\cdots)2n = (.8770\cdots)n.$$

The latter number is in the middle of the interval described in Theorem 2.1. We suspect that the limit of the expected value of β_1/n exists and is approximately .875 ... and that the value of β_1 for almost all (n,n)-trees is even more closely concentrated about the asymptotic value of the mean than the interval of Theorem 2.1. The next step requires improvement of both upper and lower bounds in both equations (2.22).

References

[AlSE92] N. Alon, J. H. Spencer and P. Erdös, *The Probabilistic Method*, Wiley-Interscience, New York, 1992.

[Be74] E. A. Bender, *Asymptotic methods in enumeration*, SIAM Rev. **16** (1974), 485-515.

[ChP95] J. H. Cho and E. M. Palmer, *On the expected number of edges in a maximum matching of an (r,s)-tree*, Intern. J. Computer Math. **56** (1995), 39-50.

[Er47] P. Erdös, *Some remarks on the theory of graphs*, Bull. Amer. Math. Soc. **53** (1947), 292-294..

[Ga59] T. Gallai, *Über extreme Punkt- und Kantenmengen*, Ann. Univ. Sci. Budapest, Eötvös Sect. Math. **2** (959), 133-138.

[Ha35] P. Hall, *On representatives of subsets*, J. London Math. Soc. **10** (1935), 26-30.

[H69] F. Harary, *Graph Theory*, Addison-Wesley, Reading, 1969.

[Kn68] D. E. Knuth, Another enumeration of trees, Canad. J. Math. **20** (1968), 1077-1086.

[M70] J. W. Moon, *Counting Labelled Trees*, Canad. Math. Congress, Montreal, 1970.

[Pa85] E. M. Palmer, *Graphical Evolution*, Wiley-Interscience, New York, 1985.

[Pa92] E. M. Palmer, *Matchings in random superpositions of bipartite trees*, J. Comput. Appl. Math **41** (1992).

[Sch94] E. Schmutz, *Matchings in Superpositions of (n,n)-Bipartite Trees*, J. Random Structures & Algorithms **5** (1994), 235-241.

J. H. Cho and Moo Ha Woo

DEPARTMENT OF MATHEMATICS EDUCATION, KOREA UNIVERSITY, 5-KA ANAM-DONG, SEOUL 136-701, KOREA