



PROBABILISTIC ANALYSIS OF THE FIRST ZAGREB INDEX

RAMIN KAZEMI

Communicated by Ivan Gutman

ABSTRACT. In this paper we study the first Zagreb index in bucket recursive trees containing buckets with variable capacities. This model was introduced by Kazemi in 2012. We obtain the mean and variance of the first Zagreb index and introduce a martingale based on this quantity.

1. Introduction

A graph is a collection of points and lines connecting a subset of them. The points and lines of a graph are also called vertices and edges of the graph, respectively. The vertex and edge sets of a graph G are denoted by $V(G)$ and $E(G)$, respectively. Trees are defined as connected graphs without cycles, and their properties are basics of graph theory. Recursive trees are rooted labelled trees, where the root is labelled by 1 and the labels of all successors of any node v are larger than the label of v . Also, bucket trees are a generalization of the ordinary trees where buckets (or nodes) can hold up to $b \geq 1$ labels. Mahmoud and Smythe introduced bucket recursive trees as a generalization of ordinary recursive trees [5]. In this model the capacity of buckets is fixed. They applied a probabilistic analysis for studying the height and depth of the largest label in these trees and Kuba and Panholzer analyzed these trees as a special instance of bucket increasing trees which is a family of some combinatorial objects [4]. Recently, Kazemi [8] introduced a new version of bucket recursive trees where the nodes are buckets with variable capacities labelled with integers $1, 2, \dots, n$. In fact, the capacity of buckets is a random variable in these models. He has studied the quantity depth in these trees via a combinatorial approach. We introduce the model below for the reader's convenience.

MSC(2010): Primary: 05C05; Secondary: 60F05.

Keywords: Bucket recursive trees with variable capacities of buckets, the first Zagreb index, martingale.

Received: 20 April 2013, Accepted: 12 June 2013.

Definition 1.1. A bucket recursive tree with variable capacities of buckets starts with the root labelled by 1 that has $r \geq 0$ descendants each of them making a subtree. Every node v in the subtrees has capacity $c(v) < b$ or $c(v) = b$. The nodes v with capacities $c(v) < b$ are connected together with 1 edge and the nodes with capacities b have descendants ≥ 0 again each of them making a subtree such that the labels within these nodes are arranged in increasing order. The tree is completed when the label n is inserted in the tree.

The last nodes with $c(v) \leq b$ labels at the end of branches are called “leaves” and other nodes are called “non-leaves”. The stochastic growth rule for this model is as follows. For constructing a tree of size $n + 1$ (attracting label $n + 1$ to a tree of size n), if a leaf v has the capacity $c(v) < b$, then we add the label $n + 1$ to this node and make a node with capacity $c(v) + 1$. But for a node with capacity b , we produce a new node $n + 1$. Figure 1 illustrates such a tree of size 16 with $b = 3$. This model can be considered as a generalization of random recursive trees ($b = 1$).

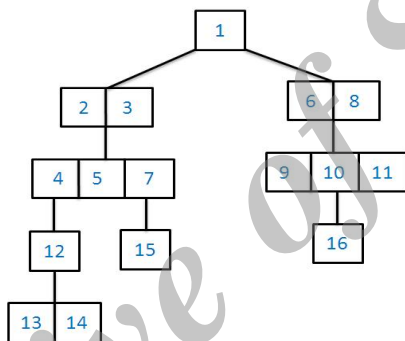


FIGURE 1. A tree of size 16 with $b = 3$.

Let $|\cdot|$ denotes the size of sets. The probability p , which gives the probability that label n is attracted by node v in the tree of size $n - 1$ is [8]

$$(1.1) \quad p = \frac{c(v)}{n - 1 - |\gamma|},$$

where $\gamma = \{v \in T; c(v) < b, \text{ and } v \text{ is a non-leaf}\}$. Let $g(b)$ be a function of b , where $g(1) = 0$ and $g(b) \geq 1$ for all $b \geq 2$. It is obvious that the size of buckets is lesser than n for $b > 1$. We can write this number as $n - |\gamma| - g(b)$. i.e., $|V(T)| = n - (|\gamma| + g(b))$. Since $\sum_v d(v) = 2|E(T)|$, thus for our tree of size n ,

$$(1.2) \quad \sum_v d(v) = 2(n - 1 - (|\gamma| + g(b))).$$

A molecular graph is a simple graph such that its vertices correspond to the atoms and the edges to the bonds. Note that hydrogen atoms are often omitted. Chemical graph theory is a branch of mathematical chemistry which has an important effect on the development of the chemical sciences. By IUPAC terminology, a topological index is a numerical value associated with chemical constitution purporting for correlation of chemical structure with various physical properties, chemical reactivity

or biological activity. In an exact phrase, if \mathcal{G} denotes the class of all finite graphs then a topological index is a function \mathcal{F} from \mathcal{G} into real numbers with this property that $\mathcal{F}(G) = \mathcal{F}(H)$, if G and H are isomorphic. One of the most important topological indices of a graph is the first Zagreb index.

Definition 1.2. [3] *The first Zagreb index $Z(G)$ of G is defined as*

$$Z(G) = \sum_{v \in V(G)} (d(v))^2,$$

where $d(v)$ denotes the degree of the vertex v in G .

Thus the first Zagreb index of a graph is defined as the sum of the squares of the degrees of all vertices in the graph. This index introduced by chemists Gutman and Trinajstić [3]. This index reflects the extent of branching of the molecular carbon-atom skeleton, and can thus be viewed as molecular structure-descriptors. Nikolić et al. [6] studied the mathematical properties of this quantity. The first Zagreb index and its variants have been used to study molecular complexity, chirality, ZE isomerism, and heterosystems, whilst the overall Zagreb indices exhibited a potential applicability for deriving multilinear regression models [7]. The first Zagreb index has also been used in the studies of quantitative structure-property/activity relationships (QSPR/QSAR) [2]. Also, for a tree T of size n , Li et al. [9] studied the extreme values of the Zagreb index of T in 2003. Suppose n atoms in a branching molecular structure are stochastically labelled with integers $1, 2, \dots, n$. In this case, atoms in different functional groups can be considered as the labels of different buckets of a bucket recursive tree (the size of the largest functional group is b). We study the first Zagreb index in bucket recursive trees with variable capacities of buckets since the structures of many molecules such as dendrimers may be considered as a bucket tree. Also, we show our results so that for $b = 1$ reduce to the results on random recursive trees. It is possible to show the results in the shorter forms.

2. The Main Results

The sequence $(X_n)_{n \geq 1}$ of random variables is said to be a *martingale* relative to the sigma-field \mathcal{F}_n if and only if for all $n = 1, 2, \dots$, $\mathbb{E}(X_{n+1} | \mathcal{F}_n) = X_n$ (a.e.) [1]. Let Z_n be the first Zagreb index of a bucket recursive tree of size n with variable capacities of buckets and \mathcal{F}_n be the sigma-field generated by the first n stages of these trees. Let U_n be a randomly chosen bucket belong to $\lambda_k \cup \mu_k$ of size- n tree, where [8]

$$\begin{aligned} \lambda_k &= \{v \in T; c(v) = k < b, \text{ and } v \text{ is a leaf}\}, \\ \mu_k &= \{v \in T; c(v) = b \text{ and } d(v) = k\} \cup \{\text{root node}\}. \end{aligned}$$

By stochastic growth rule of the tree,

$$\begin{aligned} Z_n &= Z_{n-1} + (d_{U_{n-1}} + 1)^2 - d_{U_{n-1}}^2 + 1 \\ (2.1) \quad &= Z_{n-1} + 2(d_{U_{n-1}} + 1), \end{aligned}$$

where $d_{U_{n-1}}$ is the degree of the randomly chosen bucket U_{n-1} .

Theorem 2.1. *The process $Z = (Z_n - \mathbb{E}(Z_n), n \geq 1)$ is a martingale.*

Proof. From (2.1),

$$\begin{aligned} \mathbb{E}(Z_n | \mathcal{F}_{n-1}) &= Z_{n-1} + 2 \frac{b(b+1)}{2} \frac{1}{n-1-|\gamma|} \left(2(n-2 - (|\gamma| + g(b))) \right) + 2 \\ (2.2) \quad &= Z_{n-1} + 2b(b+1) + 2 - 2b(b+1) \frac{1+|\gamma|+g(b)}{n-1-|\gamma|}, \end{aligned}$$

since label n can be attached to bucket with capacities $1, 2, \dots, b-1$ or b in a tree of size $n-1$.

Taking expectation of the relation (2.2):

$$(2.3) \quad \mathbb{E}(Z_n) = \mathbb{E}(Z_{n-1}) + 2b(b+1) + 2 - 2b(b+1) \frac{1+|\gamma|+g(b)}{n-1-|\gamma|},$$

with the initial conditions $Z_1 = 0$ and $Z_2 = 2$. The recurrence equation (2.3) leads to

$$(2.4) \quad \begin{aligned} \mathbb{E}(Z_n) &= n(2b(b+1) + 2) - 2b(b+1)(1 + |\gamma| + g(b))H_{n-1,\gamma} \\ &- (2 + 2b(b+1)), \quad n \geq 1 \end{aligned}$$

where $H_{n,\gamma} = \sum_{j=1}^n \frac{1}{j-|\gamma|}$. Also

$$\begin{aligned} \mathbb{E}(Z_n - \mathbb{E}(Z_n) | \mathcal{F}_{n-1}) &= Z_{n-1} + (1-n)(2b(b+1) + 2) \\ &+ 2b(b+1)(1 + |\gamma| + g(b))H_{n-2,\gamma} \\ &+ (2 + 2b(b+1)) \\ &= Z_{n-1} - \mathbb{E}(Z_{n-1}). \end{aligned}$$

That is, the process $Z = (Z_n - \mathbb{E}(Z_n), n \geq 1)$ is a martingale. □

Corollary 2.2. *We have*

$$\mathbb{E}(Z_n) = (2b(b+1) + 2)n + \mathcal{O}(\log(n - |\gamma|)).$$

Theorem 2.3. *We have*

$$\text{Var}(Z_n) = \left(4b(b+1) + f(b) \right) + \mathcal{O}(\log^2(n - |\gamma|)),$$

where $f(1) = 0$ and for $b \geq 2$, $f(b) > 0$.

Proof. Let $Y_1 = 0$ and for $n \geq 2$,

$$Y_n := Z_n - Z_{n-1} - 2b(b+1) - 2 + 2b(b+1) \frac{1+|\gamma|+g(b)}{n-1-|\gamma|}.$$

Then $\mathbb{E}(Y_n | \mathcal{F}_{n-1}) = 0$. From (2.1),

$$(2.5) \quad \mathbb{E}((Z_n - Z_{n-1} - 2)^2 | \mathcal{F}_{n-1}) = 4\mathbb{E}(d_{U_{n-1}}^2) = 2b(b+1) \frac{Z_{n-1}}{n-1-|\gamma|}.$$

Also

$$\begin{aligned}
 \mathbb{E}((Z_n - Z_{n-1} - 2)^2 | \mathcal{F}_{n-1}) &= \mathbb{E} \left(\left(Y_n + 2b(b+1) \left(1 - \frac{1 + |\gamma| + g(b)}{n-1-|\gamma|} \right) \right)^2 | \mathcal{F}_{n-1} \right) \\
 (2.6) \qquad \qquad \qquad &= \mathbb{E}(Y_n^2 | \mathcal{F}_{n-1}) + 4b^2(b+1)^2 \left(1 - \frac{1 + |\gamma| + g(b)}{n-1-|\gamma|} \right)^2.
 \end{aligned}$$

Now, from (2.4), (2.5) and (2.6),

$$\begin{aligned}
 \mathbb{E}(Y_n^2) &= 2b(b+1) \frac{\mathbb{E}(Z_{n-1})}{n-1-|\gamma|} - 4b^2(b+1)^2 \left(1 - \frac{1 + |\gamma| + g(b)}{n-1-|\gamma|} \right)^2 \\
 (2.7) \qquad \qquad &= \left(4b(b+1) + f(b) \right) + \mathcal{O} \left(\frac{\log(n-|\gamma|)}{n-|\gamma|} \right), \quad n \geq 2
 \end{aligned}$$

where $f(1) = 0$ and for $b \geq 2$, $f(b) > 0$. By definition of Y_n , we have

$$\text{Var}(Z_n) = \mathbb{E}(Z_n - \mathbb{E}(Z_n))^2 = \sum_{i=1}^n \mathbb{E}(Y_i^2) = \left(4b(b+1) + f(b) \right) + \mathcal{O}(\log^2(n-|\gamma|)),$$

since for any $1 \leq i \neq j \leq n$, $\mathbb{E}(Y_i Y_j) = 0$. □

Corollary 2.4. *By Chebyshev's inequality, $n^{-1} Z_n \xrightarrow{P} 2b(b+1) + 2$.*

Theorem 2.5. *As $n \rightarrow \infty$,*

$$\frac{1}{n} \sum_{i=1}^{n-|\gamma|} \frac{Z_i - \mathbb{E}(Z_i)}{i} \xrightarrow{P} 0.$$

Proof. We have

$$\begin{aligned}
 \mathbb{E} \left(\sum_{i=1}^{n-|\gamma|} \frac{Z_i - \mathbb{E}(Z_i)}{i} \right)^2 &= \mathbb{E} \left(\sum_{i=1}^{n-|\gamma|} \frac{1}{i} \sum_{j=1}^i Y_j \right)^2 \\
 &= \mathbb{E} \left(\sum_{i=1}^{n-|\gamma|} (H_{n,0} - H_{i-1,0}) Y_i \right)^2 \\
 &= \sum_{i=1}^{n-|\gamma|} (H_{n,0} - H_{i-1,0})^2 \mathbb{E}(Y_i^2) \\
 &\leq (H_{n,0} - \min_i H_{i-1,0})^2 \sum_{i=1}^n \mathbb{E}(Y_i^2) \\
 &= \left(4b(b+1) + f(b) \right) n \log^2(n-|\gamma|) + \mathcal{O}(\log^4(n-|\gamma|)).
 \end{aligned}$$

By Chebyshev's inequality proof is completed. □

Corollary 2.6. *Assume $b = 1$. Then $|\gamma| = 0$, $g(1) = 0$, $f(1) = 0$ and all results reduce to random recursive trees. i.e.,*

$$\begin{aligned}\mathbb{E}(Z_n) &= 6n + \mathcal{O}(\log n), \\ \text{Var}(Z_n) &= 8 + \mathcal{O}(\log^2 n), \\ n^{-1}Z_n &\xrightarrow{P} 6, \\ \frac{1}{n} \sum_{i=1}^n \frac{Z_i - \mathbb{E}(Z_i)}{i} &\xrightarrow{P} 0.\end{aligned}$$

Acknowledgments

The author expresses its gratitude to the referees for reading the paper carefully and giving valuable comments.

REFERENCES

- [1] P. Billingsley, *Probability and Measure*, Third edition, Wiley Series in Probability and Mathematical Statistics, A Wiley-Interscience Publication, John Wiley and Sons, Inc., New York, 1995.
- [2] J. Devillers and A. T. Balaban, *Topological indices and related descriptors in QSAR and QSPR*, Gordon and Breach, Amsterdam, 1999.
- [3] I. Gutman and N. Trinajstić, Graph theory and molecular orbitals. Total φ -electron energy of alternant hydrocarbons, *Chem. Phys. Lett.*, **17** (1972) 535-538.
- [4] M. Kuba and A. Panholzer, A Combinatorial approach to the analysis of bucket recursive trees, *Theoret. Comput. Sci.*, **411** (2010) no. 34-36 3255-3273.
- [5] H. Mahmoud and R. Smythe, Probabilistic analysis of bucket recursive trees, *Theoret. Comput. Sci.*, **144** (1995) 221-249.
- [6] S. Nikolić, I. M. Tolić, N. Trinajstić and I. Baučić, On the Zagreb indices as complexity indices, *Croatica Chemica Acta.*, **73** (2000) 909-921.
- [7] S. Nikolić, G. Kovačević, A. Miličević and N. Trinajstić, *On molecular complexity indices*, In *Complexity in Chemistry: Introduction and Fundamentals*, eds D. Bonchev and D. H. Rouvray, Taylor and Francis, London, 2003 29-89.
- [8] R. Kazemi, Depth in bucket recursive trees with variable capacities of buckets, *Acta Math. Sin. Engl. Ser.*, 2013. DOI: <http://dx.doi.org/10.1007/s10114-013-2115-2>.
- [9] X. Li, Z. Li and L. Wang, The inverse problems for some topological indices in combinatorial chemistry, *J. Comput. Biol.*, **10** (2003) 47-55.

Ramin Kazemi

Department of Statistics, Imam Khomeini International University, Qazvin, Iran

Email: kazemi@ikiu.ac.ir