Probabilistic Study of Some Topological Indices in Bucket Tree Structures

Ramin Kazemi
Department of Statistics
Imam Khomeini International University
Qazvin
Iran
r.kazemi@SCI.ikiu.ac.ir http://www.ikiu.ac.ir/members/?id=231&lang=1

Abstract: In the fields of chemical graph theory, molecular topology, and mathematical chemistry, a topological index also known as a connectivity index is a type of a molecular descriptor that is calculated based on the molecular graph of a chemical compound. These parameters also have applications in drug structures. In this paper, we give some new probabilistic results on the first Zagreb, the Platt, Narumi-Katayama and Gordon-Scantlebury indices in two bucket tree structures.

Key-Words: Topological indices, tree structures.

1 Introduction

Topological indices are numerical parameters of a graph which characterize its topology and are usually graph invariant. The research and application of the first Zagreb index appears mainly in mathematical chemistry. 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 or activity relationships. For a tree T of size n, Li *et al.* [4] studied the extreme values of the first Zagreb index of T.

Definition 1. The first Zagreb index $Z(G)$ of G is defined as

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

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

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 reflects the extent of branching of the molecular carbon-atom skeleton, and can thus be viewed as molecular structure-descriptor. Nikolic' *et al.* [7] studied the mathematical properties of this quantity.

A path in a graph is a sequence of adjacent edges, which do not pass through the same vertex more than once, and the length of the path is the number of edges in it.

Definition 2. For a simple graph G, the Godon-Scantlebury index of G is equal to the number of paths of length two in G [2], and the Platt index is equal to the total sum of the degrees of all edges in G [9].

The first Zagreb index is related to the Gordon-Scantlebury and Platt indices. Let $S(G)$, and $P(G)$ be the Gordon-Scantlebury index and the Platt index of the graph G, respectively. Nikolic' *et al.* [6] showed $Z(G) = 2\big(S(G) + E(G)\big)$ and $P(G) = 2S(G)$, where $E(G)$ is the number of edges of G.

Definition 3. The Narumi-Katayama [8] index $N(G)$ of G is defined as

$$N(G) = \prod_{v \in \mathrm{V(G)}} d(v).$$

2 Bucket tree structures

Trees are defined as connected graphs without cycles. 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. They are one of the most natural combinatorial tree models with applications in several fields, e.g., it has been introduced as a model for the spread of epidemics, for pyramid schemes, for the family trees of preserved copies of ancient texts and furthermore it is related to the Bolthausen-Sznitman coalescence model. It is easy to show by induction that there are $(n-1)!$ different recursive trees with n nodes. It is of particular interest in applications to assume the random recursive tree model and to speak about a random recursive tree with n nodes, which means that one of the $(n-1)!$ possible recursive trees with n nodes is chosen with equal probability, i.e., the probability that a particular tree with n nodes is chosen is always $1/(n-1)!$.

An interesting and natural generalization of random recursive trees has been introduced by Mahmoud and Smythe [5], which are called bucket recursive trees. In this model the nodes of a bucket recursive tree are buckets, which can contain up to a fixed integer amount of $b \geq 1$ labels. A probabilistic description of random bucket recursive trees is given by a generalization of the stochastic growth rule for ordinary random recursive trees (which are the special instance $b=1$), where a tree grows by progressive attraction of increasing integer labels: when inserting label $n+1$ into an existing bucket recursive tree containing n labels (i.e., containing the labels $\{1,2,...,n\}$) all n existing labels in the tree compete to attract the label $n+1$, where all existing labels have equal chance to recruit the new label. If the label winning this competition is contained in a node with less than b labels (an unsaturated bucket or node), label $n+1$ is added to this node, otherwise if the winning label is contained in a node with already b labels (a saturated bucket or node), label $n+1$ is attached to this node as a new bucket containing only the label $n+1$. Starting with a single bucket as root node containing only label 1 leads after $n-1$ insertion steps, where the labels $2,...,n$ are successively inserted according to this growth rule, to a so called random bucket recursive tree with n labels and maximal bucket size b [5]. Of course, the above growth rule for inserting the label $n+1$ could also be formulated by saying that, for an existing bucket recursive tree T with n labels, the probability that a certain node $v \in T$ with capacity $1 \leq c(v) \leq b$ attracts the new label $n+1$ is proportional to the number of labels contained in v, i.e., $c(v)/n$.

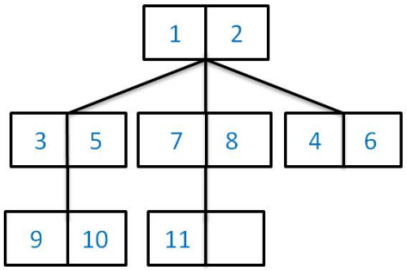

Fig. 1. A bucket recursive tree of size $n=11$ with maximal bucket size $b=2$ (fixed capacities).

Figure 1 illustrates a bucket recursive tree of size $n=11$ with maximal bucket size $b=2$.

Let $g(b)$ be a real valued function of b such taht $g(1)=0$ and $g(b) \geq 1$ for all $b \geq 2$. It is obvious that the size of buckets is lesser than n for $b \geq 1$. We can write this number as $n - g(b)$. i.e., $|V(T)| = n - g(b)$. Since $\sum_v \mathrm{d}(v) = 2|E(T)|$, thus for a bucket recursive tree of size n,

$$\sum_{v \in V(G)} \mathrm{d}(v) = 2(n - 1 - g(b)). \quad (1)$$

Kazemi [3] introduced a new version of bucket recursive trees where the nodes are buckets with variable capacities labelled with integers $1,2,...,n$. In fact, the capacity of buckets is a random variable in these models. A size-m bucket recursive tree T_m with variable bucket capacities and maximal bucket size b starts with the root labeled by 1. The tree grows by progressive attraction of increasing integer labels: when inserting label $j+1$ into an existing bucket recursive tree T_j, except the labels in the non-leaf nodes with capacity$<b$ all labels in the tree (containing label 1) compete to attract the label $j+1$. For the root node and nodes with capacity b, we always produce a new node $j+1$. But for a leaf with capacity $c < b$, either the label $j+1$is attached to this leaf as a new bucket containing only the label $j+1$or is added to that leaf and make a node with capacity $c+1$. This process ends with inserting the label m (i.e., the largest label) in the tree. Figure 2 illustrates such a tree of size 11 with $b=2$. This

model can be considered as another generalization of random recursive trees ($b=1$).

The probability that a certain node v attracts the new label $n+1$ is proportional to the number of labels contained in v. I.e., the probability p, which gives the probability that label $n+1$ is attracted by

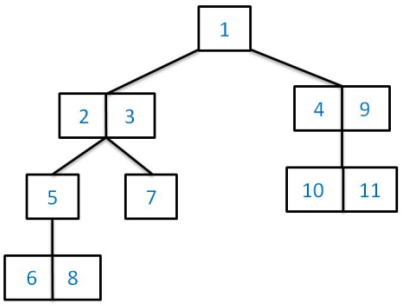

Fig. 2. A bucket recursive tree of size 11 with maximal bucket size $b=2$ (variable capacities).

node v in the tree of size n is $p=c(v)/(n-|\gamma|)$,

where $\gamma=\{v\in T; c(v)<b$ and v is a non-leaf$\}$.

The motivation for studying the topological indices of bucket recursive trees is multifold. For example, if n atoms in a branching molecular structure (such as dendrimer) are stochastically labelled with integers $1,2,...,n$, then atoms in different functional groups can be considered as the labels of different buckets of a bucket recursive tree.

3 The main results

Let $Z_{n,b}$ be the first Zagreb index of a bucket recursive tree of size n with maximal bucket size b (fixed capacities). Let $U_{n,b}$ be a randomly chosen bucket belong to $V(T)$. If label n attached to a leaf with capacity $c(v)<b$, then $Z_{n,b}=Z_{n-1,b}$. Otherwise, by stochastic growth rule of the tree discussed in Section 2 and definition of the first Zagreb index, $Z_{n,b}=Z_{n-1,b}+2d_{U_{n-1,b}}+2$, where $d_{U_{n-1,b}}$ is the degree of the randomly chosen bucket $U_{n-1,b}$.

Theorem 1. For $n\geq b+1$,

$$\mathrm{E}(Z_{n,b})=(4b+2)n+\mathrm{O}(\log n),$$

$$\mathrm{Var}(Z_{n,b})=8bn+\mathrm{O}(\log^2 n).$$

Theorem 2. As $n\to\infty$, $n^{-1}Z_{n,b}\xrightarrow{P}4b+2$ and

$$\frac{1}{n}\sum_{i=b}^{n}\frac{Z_{i,b}-E(Z_{i,b})}{i}\xrightarrow{P}0.$$

Corollary 1. There exists a random variable Z such that $Z_{n,b}-E(Z_{n,b})\xrightarrow{L^2}Z$, since

$$\sup_n E(Z_{n,b}-E(Z_{n,b}))^2=\sup_n \mathrm{Var}(Z_{n,b})<\infty.$$

Thus $\{(Z_{n,b}-E(Z_{n,b}))^2\}_{n\geq b}$, is uniformly integrable. Then $Z_{n,b}-E(Z_{n,b})\xrightarrow{L^2}Z$ and $\{Z_{n,b}-E(Z_{n,b})\}_{n\geq b}$ is uniformly integrable. Finally, $Z_{n,b}-E(Z_{n,b})\xrightarrow{D}Z$ [1].

Corollary 2. Let $S_{n,b}$ and $P_{n,b}$ be the Gordon-Scantlebury and Platt indices of a random bucket recursive tree, respectively. Then

$$E(S_{n,b})=2bn+\mathrm{O}(\log n),$$

$$\mathrm{Var}(S_{n,b})=(2b-1)n+\mathrm{O}(\log^2 n),$$

$$E(P_{n,b})=4bn+\mathrm{O}(\log n),$$
$$\mathrm{Var}(P_{n,b})=4(2b-1)n+\mathrm{O}(\log^2 n),$$

$$n^{-1}S_{n,b}\xrightarrow{P}2b,\ n^{-1}P_{n,b}\xrightarrow{P}4b.$$

Theorem 3. Let $N_{n,b}$ be the Narumi-Katayama index of a random bucket recursive tree T of size n with maximal bucket size b. Then for $n\geq b+1$,

$$E(N_{n,b})=\prod_{i=1}^{n-b-1}\left(2b\left(1-\frac{1+g(b)}{n-i}\right)+1\right).$$

Corollary 3. Let S_n, N_n and P_n be the Gordon-Scantlebury, Nurami-Katayama and Platt indices of a random recursive tree, respectively $(b=1)$. Then

$$E(Z_n)=6n+\mathrm{O}(\log n),\quad \mathrm{Var}(Z_n)=8n+\mathrm{O}(\log^2 n),$$

$$E(S_n)=2n+\mathrm{O}(\log n),\quad \mathrm{Var}(S_n)=n+\mathrm{O}(\log^2 n),$$

$$E(P_n)=4n+\mathrm{O}(\log n),\quad \mathrm{Var}(P_n)=4n+\mathrm{O}(\log^2 n),$$

$$E(N_n) = \prod_{i=1}^{n-2}\left(2\left(1-\frac{1}{n-i}\right)+1\right), \quad n^{-1}S_n \xrightarrow{P} 2, \quad n^{-1}P_n \xrightarrow{P} 4.$$

Let $Z'_{n,b}$, $S'_{n,b}$, and $P'_{n,b}$ be the first Zagreb, Gordon-Scantlebury and Platt indices of another bucket trees (variable capacities), respectively. Then

$$\mathrm{E}(Z'_{n,b}) = (2b(b+1)+2)n + \mathrm{O}(\log(n-|\gamma|)),$$
$$\mathrm{Var}(Z'_{n,b}) = (4b(b+1)+f(b))n + \mathrm{O}(\log^2(n-|\gamma|)).$$

where and $f(1)=0$ and for $b \geq 2$, $f(b)>0$. With the above approach we can obtain:

$$\mathrm{E}(S'_{n,b}) = b(b+1)n + \mathrm{O}(\log(n-|\gamma|)),$$

$$\mathrm{Var}(S'_{n,b}) = (b(b+1)+\frac{f(b)}{4}-1)n$$

$$+\mathrm{O}(\log^2(n-|\gamma|)),$$

$$\mathrm{E}(P'_{n,b}) = 2b(b+1)n + \mathrm{O}(\log(n-|\gamma|)),$$

$$\mathrm{Var}(P'_{n,b}) = 4(b(b+1)+\frac{f(b)}{4}-1)n + \mathrm{O}(\log^2(n-|\gamma|)),$$

$$n^{-1}S'_{n,b} \xrightarrow{P} b(b+1), \quad n^{-1}P'_{n,b} \xrightarrow{P} 2b(b+1).$$

If $b=1$, then $|\gamma|=0$ [3] and we can see all results are the same for two models.

References:

[1] P. Billingsley, *Probability and Measure.* John Wiley and Sons, New York, 1995.

[2] M. Gordon, and G. R. Scantlebury, *Non-random polycondensation: statistical theory of the substitution effect,* Trans. Faraday Soc. 60 (2), 1964, pp. 604-621.

[3] R. Kazemi, *Depth in bucket recursive trees with variable capacities of buckets,* Acta Math. Sin, English Series, 30(2), 2014, pp. 305-310.

[4] X. Li, Z. Li, and L. Wang, *The inverse problems for some topological indices in combinatorial chemistry,* J. Comput. Biol. 10, 2003, pp. 47-55.

[5] H. Mahmoud, and R. Smythe, *Probabilistic analysis of bucket recursive trees,* Theor. Comput. Sci. 144, 1995, pp. 221-249.

[6] Nikolić, R., Tolić, I. M., Trinajstić, I. M. and Baučić, I. *On the Zagreb indices as complexity indices,* Croatica Chemica Acta. 73, 2000, pp. 909-921.

[7] I. 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.

[8] H. Narumi, and M. Katayama, *Simple topological index. A newly devised index characterizing the topological nature of structural isomers of saturated hydrocarbons,* Mem. Fac. Engin. Hokkaido Univ. 16**,** 1984, pp. 209-214.

[9] J. R. Platt, *Inuence of neighbor bonds on additive bond properties in paraffins,* Chem. Phys. 15, 1947, pp. 419-420.