

On the 5-Local Profiles of Trees

Efron Manik^{1,2}, Saib Suwilo¹, Tulus¹, Opim Salim Sitompul³

¹ Department of Mathematics, University of Sumatera Utara, Medan-Indonesia,

² Department of Mathematics Education, Nommensen HKBP University, Medan-Indonesia,

³ Department of Information Technology, University of Sumatera Utara, Medan-Indonesia

E-mail: efmanik@gmail.com

Abstract. For T and U trees, we denote $c(U, T)$ as the number of copies U in T , in other words: the number of homomorphism injective from U to T . The path, star, and y -shaped tree is a list of all the homomorphisms of the tree by 5-point. Furthermore, the 5-profile of the T_n tree is symbolized by P, S , and Y for the paths, star, and y -shaped trees respectively. So the set limit of 5-profile, $\Delta_T(5)$, is a subset of \mathbf{R}^3 . Notation $(p_1, p_2, p_3) \in \Delta_T(5)$ corresponds to P, S , and Y respectively. $\mathcal{P}(5)$ is a projection of $\Delta_T(5)$ on the first two coordinates. Determining the area boundary of $\mathcal{P}(5)$ is a challenging task. The d -millipede tree produces the points of $\mathcal{P}(5)$ whose sum of p_1 and p_2 are very small at some point. The study of whether the point generated by the d -millipede tree is the lower bound of the set $\mathcal{P}(5)$ is a question that requires investigation. The expanded d -millipede tree is defined as the tree produced by adding sides to the leaves of the d -millipede tree. This paper discusses the coordinates of point $\mathcal{P}(5)$ generated by the expanded d -millipede tree. The d -millipede tree is a tree with the least number of sub-trees from the family of trees with the same number of points and degrees. The d -millipede tree produces the lowest point in region $\mathcal{P}(5)$. The points generated by the optimum tree and points generated by the expanding d -millipede tree are above the curve connecting the points generated by the d -millipede tree. So we estimate that the lower boundary region $\mathcal{P}(5)$ is the curve connecting the points generated by the d -millipede tree.

1. Introduction

Image can be viewed as a graph and image segmentation can be generated using the methods in graph [1, 2]. Peter [3] uses a minimum spanning tree to segment an image named with the MSTSSCIMLRO algorithm. The algorithm is run where the first step is to write the image in graph. Determining the minimum spanning tree of the graph is the second step. The next step is to remove all sides that are larger than the threshold that form the tree into several sub-trees. A sub tree is seen as a segment in the image.

The minimum spanning tree of a graph is not unique. Following the Kruskal's algorithm, we can conclude that the difference occurs due to the different sequence of edges of equal weight in the graph. Although the minimum spanning tree is not unique to an image, but the image segmentation produced by the different minimum spanning tree is the same [4].

Determining the same object in a digital image is a challenging task [5]. If segmentation is obtained by cutting edges of a tree produced by the minimum spanning tree method, then determining the same sub-tree is the same case by searching for the same object or segment in an image.



Let U, T tree, we denote $c(U, T)$ as the number of copies U in T , or in other words: the number of homomorphism is injecting from U to T . Suppose $T_1^k, \dots, T_{N_k}^k$ are lists of all the homomorphisms of trees with k -points, where T_1^k is the path and T_2^k is the star. A k -profile of the T tree is the $p^{(k)}(T) \in R^{N_k}$ vector in which the i -th coordinate is:

$$(p^{(k)}(T))_i = \frac{c(T_i^k, T)}{Z_k(T)}, \text{ where } Z_k(T) = \sum_{j=1}^{N_k} c(T_j^k, T).$$

We are interested in understanding the set limit of k -profiles, namely:

$$\Delta_T(k) = \left\{ p \in R^{N_k} : \exists (T_n), |T_n| \xrightarrow{n \rightarrow \infty} \infty, \text{ and } p^{(k)}(T_n) \xrightarrow{n \rightarrow \infty} p \right\},$$

where $|T|$ is the number of points in T .

For the case of $k = 5$, we will obtain $N_k = 3$. The path, star, and y-shaped tree are lists of all homomorphism of the 5-point tree. Furthermore, the 5-profile of the T_n tree is symbolized by P, S , and Y for path, star, and y-shaped trees respectively. Thus the limit of the set of 5-profiles, $\Delta_T(5)$, is the subset of R^3 . The symbols $(p_1, p_2, p_3) \in \Delta_T(5)$ correspond to P, S , and Y respectively. $\mathcal{P}(5)$ is the projection of $\Delta_T(5)$ to the first two coordinates. Determining the lower bounds of $\mathcal{P}(5)$ is a challenging task. The d -millipede tree produces the points of $\mathcal{P}(5)$ whose sum of p_1 and p_2 from that point is very small at some point. Whether the points generated by the d -millipede tree is the lower bound of $\mathcal{P}(5)$ is a question that needs to be investigated. The expanded d -millipede tree is a d -millipede tree on which leaves are followed by several edges. This paper discusses the coordinates of the $\mathcal{P}(5)$ points generated by the expanded d -millipede tree. The problem to be discussed in this paper is to compare the coordinates of the points generated by the d -millipede tree with the coordinates of the points generated by the expanded d -millipede tree on $\mathcal{P}(5)$.

This paper discusses the same sub tree of a tree. Section 2 discusses the properties in the k -profile tree and the set of limits of the k -profile tree. Specifically for $k = 2$ will be discussed more deeply. Section 3 will define an expanded d -millipede. The point coordinates of the limit set of k -profile trees generated by d -millipede and the expanded d -millipede will be compared. Finally, this paper ends with Section of conclusions and suggestions.

2. The 5-Local Profil Tree

In contrast to the set of k -profile limits of the graph, the set of k -profile limits of the tree, $\Delta_T(k)$, has some interesting properties [6]. If two different points are members of $\Delta_T(k)$ then all the points between those points are also members of $\Delta_T(k)$. This is summarized in the following theorem.

Theorem 1 The set $\Delta_T(k)$ is convex.

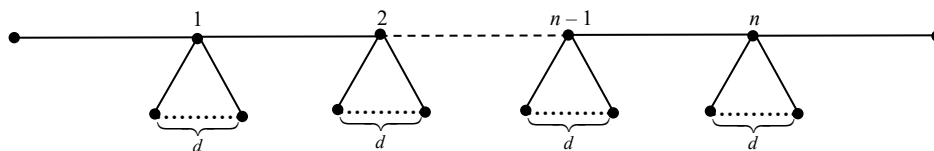


Figure 1. d -millipede

A d -millipede is a tree that all non-leaf points each have $d+2$ degrees, as shown in Figure 1. The number of non-leaf points, n , is called the length of d -millipede.

Notify $P_k(T) = c(T_1^k, T)$ and $S_k(T) = c(T_2^k, T)$, and T can also not be written for a clear reference. Suppose T_n is $(k-4)$ -millipede with length n , and R_k is $\frac{k-4}{2}$ -millipede with length 2 where k is even. It is easy to see that for $k \geq 6$ apply:

$$z_k(T_n) \geq c(R_k, T_n) \geq 2(n-2) \binom{k-3}{(k-2)/2} \geq (n-2)(3/2)^{k/2},$$

and

$$S_k(T_n) = 0, \quad P_k(T_n) \leq n(k-3)^2.$$

Since $\lim_{k \rightarrow \infty} \frac{(k-3)^2}{(3/2)^{k/2}} = 0$ then $p_1 + p_2$ can be made close to zero.

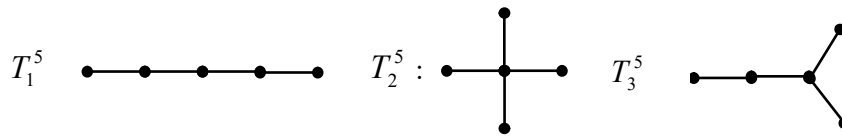


Figure 2. All Tree that is not Homorphic for 5-points

For $k = 5$, we have only three non-homomorphic trees as in Figure 2, so $N_k = 3$. We will denote $P = c(T_1^5, T)$ for path, $S = c(T_2^5, T)$ for star, and $Y = c(T_3^5, T)$ corresponding to its shape. Thus the set of 5-profile limits, $\Delta_T(5)$, is the subset of \mathbf{R}^3 corresponding to P , S , and Y respectively. Symbols $(p_1, p_2, p_3) \in \Delta_T(5)$ corresponding to P , S , and Y respectively. $\mathcal{P}(5)$ is a projection on the first two coordinates of $\Delta_T(5)$. This set can be written with $\mathcal{P}(5) = \{(p_1, p_2) \in \mathbf{R}^2 \mid (p_1, p_2, p_3) \in \Delta_T(5)\}$. By selecting T_n as a path of length n , it is easy to prove that point $(1, 0) \in \mathcal{P}(5)$. Likewise, choosing T_n as star n points then it is easy to show that point $(0, 1) \in \mathcal{P}(5)$. Using Theorem 1 it is concluded that the line segment connecting both points $(1, 0)$ and $(0, 1)$ also lies in $\mathcal{P}(5)$.

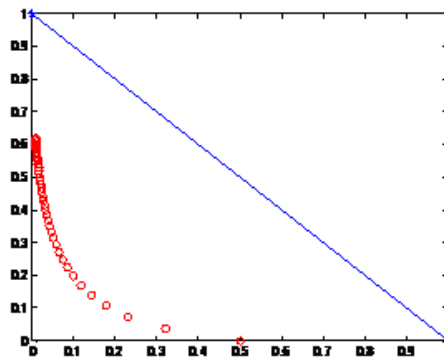


Figure 3. \circ : $\mathcal{P}(5)$ of d -millipede

Suppose the T_n tree is a d -millipede with length n . It is easy to check that if $d = 1$ then the point $(0.5, 0) \in \mathcal{P}(5)$ which corresponds to T_n . If $d = 2, 3, \dots$ then $P = (d+1)^2(n-2)$, $S = nC_4^{d+2}$, and $Y = 2(d+1)(n-2)C_2^{d+1} + 2(d+1)C_2^{d+1}$. Thus $Z_5(T) = P + S + Y = (d+1)^2(n-2) + nC_4^{d+2} + 2(d+1)(n-2)C_2^{d+1} + 2(d+1)C_2^{d+1}$. By taking the limit obtained:

$$\frac{1}{(d+1)^2 + C_4^{d+2} + 2(d+1)C_2^{d+1}} \left((d+1)^2, C_4^{d+2}, 2(d+1)C_2^{d+1} \right) \in \Delta_T(5)$$

and

$$\frac{1}{(d+1)^2 + C_4^{d+2} + 2(d+1)C_2^{d+1}} \left((d+1)^2, C_4^{d+2} \right) \in \mathcal{P}(5).$$

All points of $\mathcal{P}(5)$ can be seen in Figure 3 which is the same as [6]. Using Theorem 1 it is concluded that all regions between the points corresponding to d -millipede and the line connecting the points (0, 1) and (1, 0) are the points of $\mathcal{P}(5)$.

3. Results and Discussion

The expanded d -millipede tree is a d -millipede tree whose leaves are connected to several edges of the same amount (eg. a number of i) for each leaf, as Figure 4. The length of the expanded d -millipede tree is the same as the length of d -millipede that forms it. The addition of these sides will affect the values of P , S , and Y . Specifically for $i = 1$ or $i = 2$, the value of star S does not change, but the value of P and Y value change.

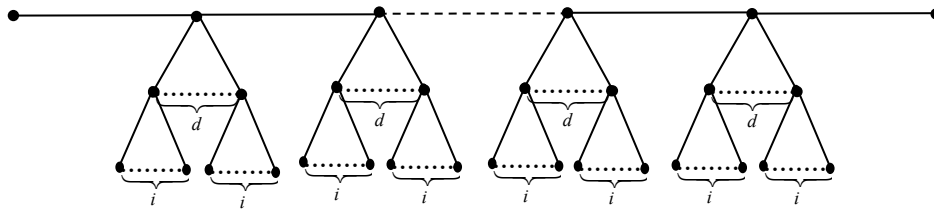


Figure 4. d -millipede with Expansion i

Suppose the T_n tree is a d -millipede with length n with an expansion i . It is easy to check that if $i = 1, 2$ then $P = (d+1)^2(n-2) + 2id(d+1)(n-1) + ni^2 C_2^d$, $S = n C_4^{d+2}$, and $Y = 2(d+1)(n-1) C_2^{d+1} + nd(d+1) C_2^i + nid C_2^{d+1}$. Thus $Z_5(T) = P + S + Y = (d+1)^2(n-2) + 2id(d+1)(n-1) + ni^2 C_2^d + n C_4^{d+2} + 2(d+1)(n-1) C_2^{d+1} + nd(d+1) C_2^i + nid C_2^{d+1}$. By taking the limit obtained:

$$\frac{1}{Q} \left((d+1)^2 + 2id(d+1) + \frac{1}{2} i^2 d(d-1) + \frac{1}{24} (d+2)(d+1)d(d-1) + (d+1)^2 d + \frac{1}{2} d(d+1)i(i-1) + \frac{1}{2} id^2(d+1) \right) \in \Delta_T(5)$$

Where $Q = (d+1)^2 + 2id(d+1) + \frac{1}{2} i^2 d(d-1) + \frac{1}{24} (d+2)(d+1)d(d-1) + (d+1)^2 d + \frac{1}{2} d(d+1)i(i-1) + \frac{1}{2} id^2(d+1)$, and

$$\frac{1}{Q} \left((d+1)^2 + 2id(d+1) + \frac{1}{2} i^2 d(d-1) + \frac{1}{24} (d+2)(d+1)d(d-1) \right) \in \mathcal{P}(5).$$

If $i = 3, 4, \dots$ then $P = (d+1)^2(n-2) + 2id(d+1)(n-1) + ni^2 C_2^d$, $S = n C_4^{d+2} + nd C_4^{i+1}$, and $Y = 2(d+1)(n-1) C_2^{d+1} + nd(d+1) C_2^i + nid C_2^{d+1}$. Thus $Z_5(T) = P + S + Y = (d+1)^2(n-2) + 2id(d+1)(n-1) + ni^2 C_2^d + n C_4^{d+2} + nd C_4^{i+1} + 2(d+1)(n-1) C_2^{d+1} + nd(d+1) C_2^i + nid C_2^{d+1}$. By taking the limit obtained:

$$\frac{1}{Q} \left((d+1)^2 + 2id(d+1) + \frac{1}{2} i^2 d(d-1) + \frac{1}{24} (d+2)(d+1)d(d-1) + \frac{1}{24} d(i+2)(i+1)i(i-1) + (d+1)^2 d + \frac{1}{2} d(d+1)i(i-1) + \frac{1}{2} id^2(d+1) \right) \in \Delta_T(5)$$

Where $Q = (d+1)^2 + 2id(d+1) + \frac{1}{2} i^2 d(d-1) + \frac{1}{24} (d+2)(d+1)d(d-1) + \frac{1}{24} d(i+2)(i+1)i(i-1) + (d+1)^2 d + \frac{1}{2} d(d+1)i(i-1) + \frac{1}{2} id^2(d+1)$, and

$$\frac{1}{Q} \left((d+1)^2 + 2id(d+1) + \frac{1}{2} i^2 d(d-1) + \frac{1}{24} (d+2)(d+1)d(d-1) + \frac{1}{24} d(i+2)(i+1)i(i-1) \right) \in \mathcal{P}(5).$$

For small d and i values, ie. $d, i = 1, 2, 3, \dots, 30$, numerical inference is presented as Figure 5. It was found that the point coordinates by the expanded d -millipede tree lie above the curve connecting the point generated by the d -millipede tree.

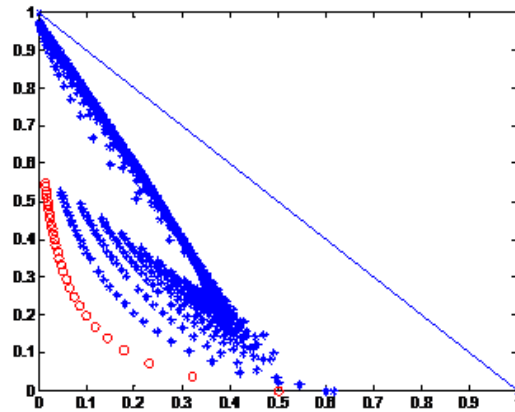


Figure 5. \circ : $\mathcal{P}(5)$ of d -millipede, $*$: $\mathcal{P}(5)$ of d -millipede with extension i

For the value of $i > 30$, we will get the value of $S = nC_4^{d+2}$ corresponding to the d -millipede tree much smaller than the value of $S = nC_4^{d+2} + ndC_4^{i+1}$ corresponding to the d -millipede tree with the expansion i . So it is easy to show that the point coordinates of $\mathcal{P}(5)$ generated by the expanded d -millipede tree lie above the curve connecting the point coordinates of $\mathcal{P}(5)$ generated by the d -millipede tree.

The optimal tree T is the tree that has the most number of subtrees from the family tree where the trees have the number of points and degrees from each point are the same [7]. The formation of such trees can be seen by looking at Figure 6. Note the 2-millipede tree with a length of 20. So this tree has 20 points of degree $(2 + 2)$, and has 20 times 2 and plus 2 (equal to 42) points of one degree. So that the sequence of degrees, d , consists of the first 20 terms of 4 and then the other is 1. The optimum tree can be seen in Figure 6.

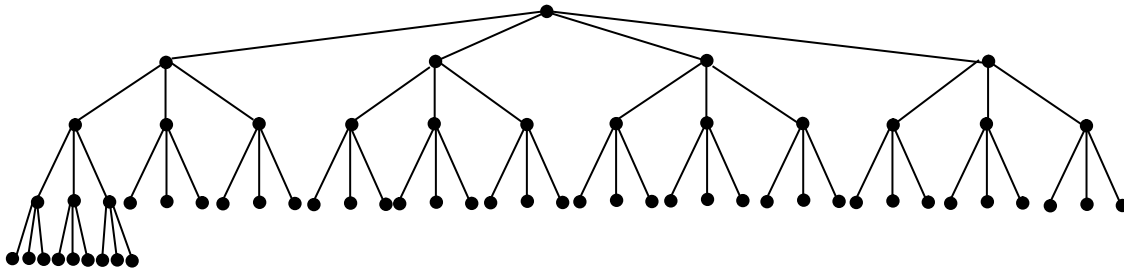


Figure 6. Optimal Tree with 20 Points of Degree 4 and Degree 1 for Other Points

In the optimum tree, each addition of points $(d+1)^2$, at the 5th and subsequent levels, will add the value of P , S , and Y as multiples of its own value. By adding a point of $n(d+1)^2$, the value increases to $P_n = n((d+1)^2 + d(d+1)^2 + (d+1)^2 C_2^{d+1}) = n((d+1)^2 + d(d+1)^2 + \frac{1}{2}(d+1)d(d+1)^2)$. The S_n value becomes $S_n = n(d+1)$, and Y_n becomes $Y_n = n((d+1) C_2^{d+1} + dC_2^{d+1}) = \frac{1}{2} n(d+1)d(2d+1)$. If a sequence is convergent, then the tail of the sequence or sub sequence will converge to the convergence point of the original sequence.

Figure 7 shows that the point generated by the optimum tree for various degrees d still remains within the region of $\mathcal{P}(5)$ in Figure 3. The projection area of $\Delta_T(5)$ is always between the segments connecting the points $(0, 1)$ and $(1, 0)$ with a curve connecting the profile points generated by the d -millipede tree.

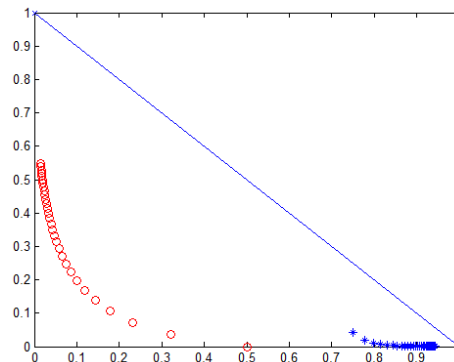


Figure 7. \circ : $\mathcal{P}(5)$ of d -millipede, $*$: $\mathcal{P}(5)$ of the Optimal Tree

The d -millipede tree is a tree with the least number of sub-trees from the family of trees with the same number of points and degrees [8]. The d -millipede tree produces the lowest point in region $\mathcal{P}(5)$. The points generated by the optimum tree and points generated by the expanding d -millipede tree are above the curve connecting the points generated by the d -millipede tree. So we estimate that the lower boundary region $\mathcal{P}(5)$ is the curve connecting the points generated by the d -millipede tree.

4. Conclusions and Suggestions

A d -millipede is a tree that all non-leaf points each have $d+2$ degrees. The expanded d -millipede tree is a d -millipede tree whose leaves are connected to several edges of the same amount (eg. a number of i) for each leaf. The addition of these sides will affect the values of P , S , and Y .

The d -millipede tree is a tree with the least number of sub-trees from the family of trees with the same number of points and degrees. The d -millipede tree produces the lowest point in region $\mathcal{P}(5)$. The points generated by the optimum tree and points generated by the expanding d -millipede tree are above the curve connecting the points generated by the d -millipede tree. So we estimate that the lower boundary region $\mathcal{P}(5)$ is the curve connecting the points generated by the d -millipede tree.

References

- [1] Dessai, R.S.G, S.D.C.S. Araujo, and C. Fernandes. Object Identification using Graph Theory. *International Journal of Engineering Science*, 2017, 10783.
- [2] Kim, S., S. Nowozin, P. Kohli, and CD Yoo. Higher-Order Correlation Clustering for Image Segmentation. In: *Advances in neural information processing systems*. 2011. p. 1530-1538.
- [3] Peter, S.J., "Minimum Spanning Tree-based Structural Similarity Clustering for Image Mining with Local Region Outliers," *IJCA*, 2010, 8(6), pp. 0975- 8887
- [4] Manik, E., S. Suwilo, Tulus, and O.S. Sitompul. The Uniqueness of Image Segmentation Generated by Different Minimum Spanning Tree. *GJPAM*, 2017, 13(7), pp. 2975–2982
- [5] Yuan, J. 2009. Image and Video Data Mining. Dissertation
- [6] Bubeck, S. and N. Linial. On the local profiles of trees. *Journal of Graph Theory*, 2016, 81.2: 109-119.
- [7] Zhang, X. M., Zhang, X. D., Gray, D., & Wang, H. The number of subtrees of trees with given degree sequence. *Journal of Graph Theory*, 2013, 73(3), 280-295.
- [8] Sills, A.V. and H. Wang. The Minimal Number of Subtrees of a Tree. *Graphs and Combinatorics*, 2015, 31:255–264.