Horton–Strahler number and centrality measures

Carlos Olguin ✉
Pontificia Universidad Católica de Chile

Cristian Riveros ✉
Pontificia Universidad Católica de Chile

Abstract

Centrality measures are widely used to assign importance to graph-structured data. This way of analyzing a graph has given so much interpretation to different graphs, in such ways that, we can formalize concepts such as importance in a graph by using this measure. Previously in a different paper, we present the formalization of this concepts, using potential functions as a way of viewing the rooting proper tie of different centralities, in this paper we analyze this concept in a centrality we define using Strahler number, a concept that has has been in the records for a while.

2012 ACM Subject Classification Theory of computation → Database theory

Keywords and phrases Streams, query evaluation, enumeration algorithms.

1 Introduction

The Strahler number can be a new concept apparently but we can observe that this concept and his uses has been with us since a long time, in different fields such as hydrology, applied mathematics eo. The usage of this number can be defined and used as a centrality in a graph, this means, it can present different properties such as a potential function, the capacity to roots trees, eo. In this paper we proves different approaches and uses of this centrality,

2 Preliminaries

Graphs. A graph is a pair $G = (V, E)$ where V is a finite set and $E \subseteq V \times V$ is a finite relation. Each pair of vertex let be (v_1, v_2) , $v_1, v_2 \in V$ this set of vertex its called edges and can be directed or not, so the edges $\in E$. A path in a graph is a sequence of non-repeated nodes connected through edges present in a graph, if there exist a path between two vertex, this vertex are called connected. An example, let be $x, y \in V$ a path can be expressed so $\{x, x_1, x_2, \cdots, y_1, y_2, \cdots, y_n\}$ with $x, x_i, y_i, y \in V$

Trees. A tree is an undirected graph in which any two vertices are connected by only and only one path. A rooted tree is refereed as a tree with a vertex who serves as the "root" of the tree, being a references to the others vertices in the tree. A neighborhood of a vertex v is denoted as $N_G(v)$, and contains all vertex (u) if $u, v \in V$ then $u \in N_G(v)$.

Centrality. The centrality of a graph is defined as a assignation of a number to nodes within a graph corresponding to their network position, the application of a centrality to a graph indicates the importance of the vertex in the graph. There exist a different number of centrality measures, so the importance of a vertex changes from centrality measure to another, also, a centrality measure has to be any function $C : VG \to R$ that assigns a score $C(v, G)$ to v depending on its graph G.

All-subgraphs centrality. . Given a graph $G = (V, E)$ and a vertex $v \in V$, we denote by $A(v, G)$the set of all connected subgraphs of G that contain v, formally, $A(v, G) = \{S \subseteq G | v \in V(S)$ and S is connected$\}$. Then all-subgraphs centrality of v in G is defined as: $AllSubgraphs(v, G) = \log_2 |A(v, G)|$.

Cristian: Agrega estas referencias [2] y [3]. También agrega esta sobre Strahler number [1] y esta otra [4]

Strahler number. In mathematics, the Strahler number or Horton–Strahler number of a mathematical tree is a numerical measure of its branching complexity. One may assign a Strahler number to all nodes of a tree, in bottom-up order, as follows:
If the node is a leaf (has no children), its Strahler number is one.
If the node has one child with Strahler number i, and all other children have Strahler numbers less than i, then the Strahler number of the node is i again.
If the node has two or more children with Strahler number i, and no children with greater number, then the Strahler number of the node is i + 1.

Potential function. A potential function is a function that measures the "potential" of every rooted tree, i.e., a tree with one node selected and the assessment depends on the selection. Now, a centrality measure admits some potential function if the comparison between two adjacent vertices is determined by the potential of their corresponding subtrees. Some centralities measures admits a potential function but it can not root trees, but, all centralities that root trees has to have a potential function.

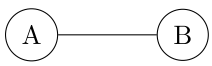
3 Main results

We can observe that the Strahler number is similar to a centrality measure, in a sense which, is a function who takes a node an assigns a score (with certain conditions), that's why the co-relation between Strahler number is visible. We can define the centrality measure using Strahler number in which sense, a higher Strahler number means higher centrality, it's important to remember that a potential function tell us if a centrality measure is able to root trees, in other words, if a centrality measure admits a potential function, it will roots trees if and only if the potential function is simetric (not necessarily works backwards). It is valid to inquire if this centrality roots trees, we can approach this by defining a a potential function, taking the concept of Strahler number we will take a potential function :

$$f_{S_N}(v,T) = \begin{cases} 1 & \text{If v is a leaf (does not have childs)} \\ i & \text{if v has one child with } f_{S_N} = i\text{, and all other children have } f_{S_N} < i \\ i+1 & \text{If v has two or more children with } f_{S_N} = i \text{ and no other children with } f_{S_N} > i \end{cases} \tag{1}$$

We can see that this potential function makes sense because the potential of a node it depends of their childs (by definition of Strahler number).

We can see now if this function is symmetric knowing that the concept of the Strahler number, is to represent a directed graph and measure the centrality to different contexts. We can make a counter-example for this and observe that if we consider:



In this example we can observe that the centrality goes to the center in the way of A to B and D to C, such that, $f(A,T_{A,B}) = f(B,T) = 0$ in such way that this is not symmetrical and in this sense does not root trees.

Knowing that this admits a potential function f and f is not symmetrical, by theorem it will not root trees, in other words, we can not find the root of tree in all cases with this centrality. If this centrality does not roots trees we can explore other things like if this potential function in a monoids version, such thing can't because we can observe that, this potencial function implies a monoid that is not associtative (which is not a monoid by definition), such leaf function can be drafted like:

$$\max\{x,y\} = \begin{cases} \max\{x,y\} & if \quad x \neq y \\ \\ \max\{x,y\}+1 & if \quad x = y \end{cases} \tag{2}$$

This function can represent correctly the centrality of a single vertex , but, it can not represent a potential function (recursively) because is not associative (and thus is not a monoid), let an example be

$$\max\{1,\max\{1,2\}\} = 2 \qquad \max\{\max\{1,1\},2\} = 3 \qquad \rightarrow 2 \neq 3 \tag{3}$$

If this centrality can not root trees this means, we can not find a root for a tree T, which can be a problem in different aplications.

Possible Aplications

Knowing this centrality does not root tress, in the problem of finding the root of a vector river in such way that, we can not find the origin of the river in all cases using this centrality.

Another application is in binarie trees which has a specific use of the Strahler number in the way that, knowing that this centrality is not symmetric we can not assure that using the Balance number in this binary tree we will find the root in all cases.

4 Conclusions

The Strahler number is a concept that as we saw, we can associate with a potential function and in such way a centrality. Knowing this and using the concept and properties of the potential function, we can assure that it will no root trees in all cases, in such way that we establish a cases in which this centrality does not root and by extend, it will not root in all cases. We formalize this by using the concepts of monoids and his relation with potential function, finding that we can not find a monoid in such way that it can form a potential function.

References

1 Javier Esparza, Michael Luttenberger, and Maximilian Schlund. A brief history of strahler numbers. In *Language and Automata Theory and Applications - 8th International Conference, LATA 2014, Madrid, Spain, March 10-14, 2014. Proceedings*, volume 8370, pages 1–13. Springer, 2014.

2 Cristian Riveros and Jorge Salas. A family of centrality measures for graph data based on subgraphs. In *23rd International Conference on Database Theory, ICDT 2020, March 30-April 2, 2020, Copenhagen, Denmark*, volume 155 of *LIPIcs*, pages 23:1–23:18, 2020.

3 Cristian Riveros, Jorge Salas, and Oskar Skibski. How do centrality measures choose the root of trees? In *26th International Conference on Database Theory, ICDT 2023, March 28-31, 2023, Ioannina, Greece*, volume 255 of *LIPIcs*, pages 12:1–12:17.

4 A Stahler. Hypsometric (area altitude) analysis of erosional topology. *Geol Soc Am Bull*, 63:1117–1142, 1952.