The Topological Properties of the Local Clustering Coefficient

Leigh Barnes Metcalf, PhD[a,*]

[a] *Software Engineering Institute, c/o Carnegie Mellon University,4500 Fifth Ave, Pittsburgh, PA, USA 15217*

Abstract

The local clustering coefficient is examined for its topological implications and a new formula to generate the local clustering coefficient is given. The formula is based on the homology of the local induced subgraph.

---

1. Introduction

The local clustering coefficient at a vertex of a graph, first introduced by Watts and Strogatz in [2], is defined as the number of edges that exist between the neighbors of the vertex divided by the total number of edges that could exist. The local clustering coefficient is then used to calculate the clustering coefficient of the graph, which is used to determine if the graph is a small world network.

In this paper, we consider the topological implications of the local clustering coefficient at a vertex and generate a new formula for calculating the local clustering coefficient. The new formula is based on the induced subgraph at a vertex and the homology of the subgraph. The local clustering coefficient is shown to be the ratio of the rank of the homology group of the local induced subgraph divided by the rank of the homology group of a regular graph with degree d.

2. Topological Discussion

Let $\mathcal{G}$ be a graph with vertices $\mathcal{V}$ and edges $\mathcal{E}$. Choose a vertex $v \in \mathcal{V}$. The local clustering coefficient of a vertex $v \in \mathcal{V}$ is defined as the number of edges that exist between the neighbors of v divided by the number of edges that could exist. Let G_v be the induced subgraph on the neighborhood of v and consider the following proposition.

Proposition 1. *The number of edges that exist between the neighbors of a vertex is equal to the rank of the homology group of the subgraph generated by the closed neighborhood of v, and this subgraph is connected.*

Proof. We begin by showing G_v is connected. Assume it is not connected. Then there exists a vertex v_k in the neighborhood of v which is not connected to v, contradicting the definition of the neighborhood of v.

The tree generated by the closed neighborhood of v, that is, v and its adjacent vertices, is a maximal subtree of the subgraph generated by the closed neighborhood of v. Denote this tree by T. We know from [1] that:

$$H_1(G_v) \cong H_1(G_v \backslash T)$$

---

* Principle Corresponding Author

Email address: `lbmetcalf@sei.cmu.edu` (Leigh Barnes Metcalf, PhD)

The rank of $H_1(G_v)$ is then the number of cycles generated by $H_1(G_v \backslash T)$. Note that an edge is in $H_1(G_v \backslash T)$ if and only if it is an edge between two neighbors of v. Therefore, the number of cycles in $H_1(G_v)$ is equal to the number of edges between two neighbors of v and is equivalent to the rank of $H_1(G_v)$. $\square$

The rank of the first homology group is also called the first Betti number. From [1] we know that the first Betti number, denoted β_1, is given by:

$$\beta_1 = |\mathcal{E}| - |\mathcal{V}| + |\mathcal{C}|$$

where $\mathcal{C}$ is the number of components of the graph. For a connected graph, $|\mathcal{C}|$ is 1. Thus, the following lemma is true.

Lemma 1. *The local clustering coefficient at a vertex v is given by the Betti number of G_v divided by deg(v)(deg$(v) - 1)/2$.*

An examination of the denominator of the local clustering coefficient also yields a direct connection to the Betti number of a regular graph. If we consider our induced graph on the closed neighborhood a regular graph with degree k, then it has $k+1$ vertices. The Betti number of that graph is then $k(k-1)/2$.

The local clustering coefficient is then a measure of the closeness of the induced subgraph to regularity. It is a weight associated with the vertex that measures the vertex's association with its neighbors.

3. Conclusion

The local clustering coefficient at a vertex v is shown to be related directly to the homology of the induced subgraph at the vertex and is in fact a ratio of Betti numbers. Utilizing the Betti number, we were then able to generate a direct equation for the local clustering coefficient utilizing only the edges and vertices of the induced subgraph. This also yields a faster computation of the local clustering coefficient.

References

[1] Peter Giblin, *Graphs, Surfaces and Homology*, 3 ed., Cambridge University Press, September 2010.

[2] D. J. Watts and S. H. Strogatz, *Collective dynamics of 'small-world' networks.*, Nature **393** (1998), no. 6684, 440–442.