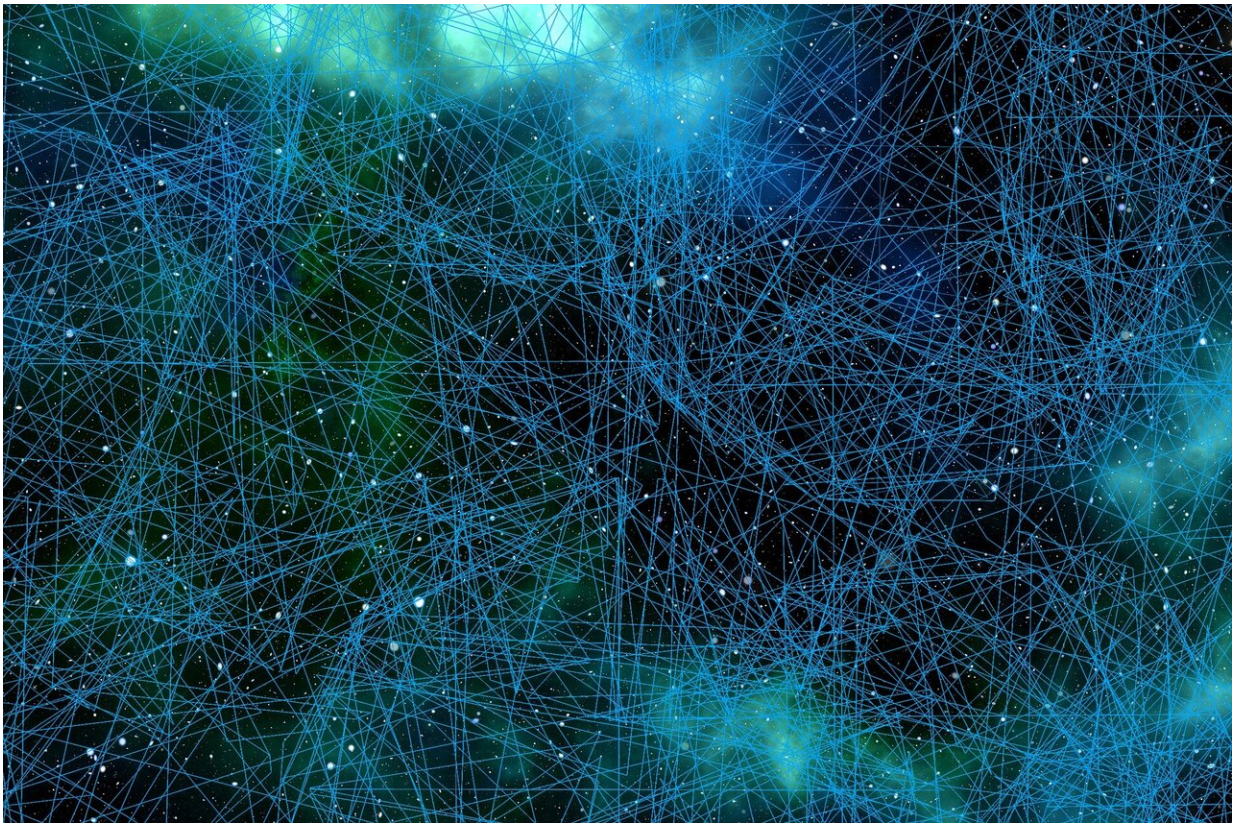# A new tool for data scientists and biologists and more

July 16 2019

Credit: CC0 Public Domain

The social network Linkedin will tell a user how he/she is connected to another. In real life, points of connection are not always that evident. However, identifying patterns or relationships and commonalities among

entities is a task that is critically important advantage for businesses, biologists, doctors, patients and more.

A new computational tool developed in the lab of USC Viterbi School Ming Hsieh Department of Electrical and Computer Engineering professor Paul Bodgan in collaboration with Ming Hsieh professor Edmond Jonckheere, is able to quickly identify the hidden affiliations and interrelationships among groups/items/persons with greater accuracy than existing tools.

The researchers in Bogdan's lab are sort of like detectives and the puzzle they are trying to figure out is how one clue, person, item or action is connected and related to another entity. Imagine a lab dedicated to a scientific "Six degrees of ..." to discover hidden interrelationships. The problem they are tackling is known by researchers who study complex networks as the "community detection problem"—identifying and mapping out which individuals or items have in common and how they are connected.

Such a computational tool could be leveraged by various groups: political strategists trying to find voters' overlapping values or shared attributes; or biologists who want to predict the potential of a drug's side effects or interactions —without running years' worth of live experiments. Their research is also being deployed to identify which parts of the brain are working on the same functions—a key piece of information for neuroscientists and individuals suffering from brain damage to anticipate if certain areas of the brain might take over functionality for injured tissue. One can also imagine this lab's algorithm working on finding points of contact on seemingly unrelated information.

Their recent paper, titled "Ollivier-Ricci Curvature-Based Method to Community Detection in Complex Networks", in the journal Nature *Scientific Reports*, documents the method the group has developed to

create this improved tool.

Methodology/Proof of Concept:

Ph.D. candidate Jayson Sia who worked on the research indicates that the algorithm they developed, the Ollivier-Ricci curvature (ORC)-based community identification, was tested and validated on four known real-world data sets the field for which the goal is to find the point of connection among the "nodes" or individuals/ individual items in a group by looking at the links between them or what is known in technical jargon as "edges." The data sets include a drug-drug interaction network, the Zachary's Karate Club; a college football conference affiliations; and a set of over 1000 political blogs.

Says lead author Sia, "In this paper, we utilized a novel geometric approach via the Ollivier-Ricci curvature which offers a natural method to discover inherent network community structures.."

Curvature in the geometric context, explains Sia, "essentially measures how a surface deviates from being flat (or how a surface 'curves'). The geometry of surfaces is related to the study of map projections and how distances are measured in a curved surface such as the Earth. The Ollivier-Ricci curvature extends this concept of 'curvature' to networks with positively curved edges being 'well connected' and naturally forming a 'community.' Negatively curved edges on the other hand are interpreted as 'bridges' between communities and cutting such edges would isolate information flow between communities."

  More information: Jayson Sia et al, Ollivier-Ricci Curvature-Based Method to Community Detection in Complex Networks, *Scientific Reports* (2019). DOI: 10.1038/s41598-019-46079-x