Review of: "Determining Affinity of Social Network using Graph Semirings"

Nicola Apollonio[1]

1 Italian National Research Council

Potential competing interests: No potential competing interests to declare.

I hope the authors may find my comments useful, even if they are severe.

In a nutshell, the authors choose a small concrete special graph, give a (questionable) definition without any reference to the state of the art or their modeling choices over other established ones, apply the definition to this example and compute some numbers that evaluates the affinity and stability. They claim to have done this within a semiring structure.

First of all, even ignoring the depth of the result, not motivating the article by discussing the state of the art and not placing it in a specific context sounds completely arbitrary. Second, this article has nothing to do with semirings (no more than saying "seven is prime" has to do with number theory).

The algebraic structure of semiring does not enter anywhere in the article . So please remove semiring from the title, as it is simply misleading.

The article is also flawed, as I show below (comment 2).

Comments

1. Section 1 looks obscure to me. There is a significant lack of motivation (see comment 5) and the paper suffers of an inappropriate use of terminology. The main example is the notion of connectivity. The authors seem to confuse connectivity with density. Connectivity (vertex- or edge-connectivity) is a global property of the graph (governed by an integer parameter) that has nothing to do with density or average degree. A disconnected graph can have a much higher average degree than a connected graph with the same number of vertices. The union of a small co-complete graph (discrete in the authors' terminology) and a large complete graph can have a much higher \beta parameter than a tree or a cycle: the former is disconnected while the latter are 1- and 2-connected, respectively.

2. The authors state: "Clearly, the \beta_G of a graph G will always be greater than or equal to that of its subgraphs". This statement is false (unless we restrict ourselves to spanning subgraphs, which is clearly not the case in the article): take the graph obtained from the union of a triangle and a stable set with n-3 vertices. For any n>3, the average degree of the graph is less than 2 (and tends to zero as n grows), while the average degree of the triangle is 2. Obviously, since the average degree and \beta_G differ by 1, the statement remains false for \beta_G.
3. I see no reason to claim that \beta_G is more expressive than 2\beta: they differ by the constant 1, so whatever

property of the graph is true for one is also true for the other.

4.An algorithm is given when the input is uniquely described, the rules are clear, and the output is also clear. A few computations from an example cannot be considered an algorithm (rather an illustration of it, provided the algorithm is described beforehand). It would also be desirable to have some idea of the complexity.Your example uses the paths of a reduced graph of an input graph, which is given along with a collection of sets covering the vertices of the input graph. Is the collection given? It probably is. But note that since a graph can have an exponential number (in the number of vertices) of paths, to make the "algorithm" effective you need to know something about the structure of the collection: if you take the discrete partition of the set of vertices as the input collection, the reduced graph is the co-complete graph while if you take the collection whose sets are the end-vertices of the edges, then the reduced graph coincides with the line graph of the input graph, and there may be exponentially many paths.

5. The r parameter is not adequately discussed or motivated: what property does r have, why r and not some other parameter; there are a plethora of intergroup affinity measures in network analysis, why should your r be interesting compared to those? This lack of discussion makes difficult to appreciate the notion of stability and affinity.

Minor comments.

The paper suffers from other limitations e.g.

(1) the Traveling Salesman Problem is indeed a problem, not an algorithm;

(2) discrete graph for the edge-less graph while correct is confusing;

(3) edges formed by... is ambiguous;

(4) using the same symbol, though in bold, for two different notions is very confusing;

(5) p_i (for the sample paths in the running example) sometimes lowercase sometimes uppercase.

I'm not saying that the parameter r introduced by the authors is trivial! After all, it can be of some interest. But not in the form presented here. In summary, I cannot recommend acceptance.