Research Article

Determining Affinity of Social Network using Graph Semirings

Gete Umbrey[1], **Saifur Rahman**[2,3]

1. Department of Mathematics, Rajiv Gandhi University, India; 2. Rajiv Gandhi University, India; 3. Department of Mathematics, Jamia Millia Islamia, New Delhi, India

In this article, we take an example of the Facebook friendship network to propose an algorithm for determining the stability or affinity of connections between different social groups within a complex social network by decomposing the given network into certain components of predefined categories. We take the graph as the principal tool, and its operations, namely, *union* and *intersection* that form semirings on the set of graphs as the primary operations.

Corresponding authors: Gete Umbrey, gete.umbrey@rgu.ac.in; Saifur Rahman, srahman@jmi.ac.in

1. Introduction

The graph theory has been a favorite platform for describing and analyzing the networks in a more abstract and general way [1], where some of the popularly known algorithms, like the Dijkstra algorithm and the Traveling Salesman Problem, are presented in the context of studying network analysis and routing problems. A graph G is a connected graph if there is a path between every pair of vertices. A graph is disconnected if there exists at least a pair of vertices that is disconnected. Various computer and mathematical sciences problems essentially involve the study of graph connectivity theories namely, network applications, routing transportation networks, network tolerance, etc. are few to be named. The *Beta index* gives the simplest measure of the degree of connectivity of a graph. It measures the level of connections and is defined as $\beta = \frac{|E|}{|V|}$, where $|E|$ is the total number of edges, and $|V|$ is the total number of vertices in the graph or network. Trees or simple networks (without loops) have a Beta value of less than one. A connected network with one cycle has a Beta value of 1, and complex networks have a high Beta value. Twice the Beta index of a graph is the average vertex degree of that graph. The average vertex degree of a graph is the ratio of the summation of all node's degrees to the total number of nodes.

 doi.org/10.32388/XAVXHO

For ease of notation, let the average vertex degree of $G=(V,E)$ be given by $A_{d(G)}=2\times\frac{|E|}{|V|}=2\beta$. Noteworthily, the notion of beta index and average vertex degree is used to determine the decision graph using algebraic graph operations like graph union and intersection [2].

A graph is a suitable tool for expressing real-life problems, and it can represent a variety of information. For basic background and fundamental notions of graph theory, we refer to a textbook by Deo [3]. There is no general way to compare the average vertex degree of a graph $G=(V,E)$ and its subgraphs. Thus, to establish a numerical (or connectivity) comparison of a graph and its subgraphs, a ration is defined, which is a function of an average vertex degree, denoted and expressed as $\boldsymbol{\beta}_G=1+A_{d(G)}=1+2\beta=1+2\frac{|E|}{|V|}=\frac{|V|+2|E|}{|V|}$ [2]. For a discrete graph, $\boldsymbol{\beta}_G=1$, while for a complete graph G with n vertices, $\boldsymbol{\beta}_G=n$. Therefore, $\boldsymbol{\beta}_G\in[1,n]$. Clearly, $\boldsymbol{\beta}_G$ of a graph G will always be greater than or equal to that of its subgraphs. Note that in considering the graph connectivity, the empty graph $(\emptyset,\emptyset)$ is excluded. Both $A_{d(G)}$ and $\boldsymbol{\beta}_G$ are the measures of connectivity of a graph G. The main difference is that the least value of $A_{d(G)}$ and $\boldsymbol{\beta}_G$ are 0 and 1, respectively, while the greatest value of $A_{d(G))}$ and $\boldsymbol{\beta}_G$ of a graph G with n vertices are $n-1$ and n, respectively. Unlike the average vertex degree $A_{d(G)}$, its function $\boldsymbol{\beta}_G$ is a consistent rule to compare a graph's connectivity with its subgraphs.

Many decision problems in real life may involve the interplay of various factors. The problems that we are considering here are presumed to be complex, which may also involve vagueness. To deal with such varied problems, merely assigning the average vertex degree as a graph's weight may not be sufficient. Or, so to say, an average vertex degree of a graph merely tells us the intensity of the connectivity of the graph and may fail to address other parameters. Suppose the problem is intended to determine the maximum degree of agreement or, conformity of participants without undermining the number of participants, then we may need more than an average vertex degree to arrive at an efficient conclusion. In such contexts, $\boldsymbol{\beta}_G$ will be a preferred choice over $A_{d(G)}$. An interesting property of $\boldsymbol{\beta}_G$ is that it preserves the importance of order and size besides measuring its connectivity. Its value keeps on decreasing with the subsequent subgraphs, and increases or decreases proportionally with that of $A_{d(G)}$; this property will also help us to decide the stability of complex networks.

In thesis [4], the authors use the graphs (mostly simple graphs) as algebraic elements and the graph operations like *union, join, intersection,* etc. as algebraic operations to study various properties of graphs and algebraic structures, particularly the properties of semirings. Consequently, the authors have proposed and discussed various properties of the semiring structure (of graphs) like completely, regular semiring, conditions for regularity of a semiring of graphs, etc. [5]. Rajkumar et al. [6] have started studying the properties of semiring by assigning vertices of graphs as semiring elements, and also the weights of the graphs are taken from the semiring structure. Graph theory has important applications in finding shortest paths, using algebraic properties (for example, see {m/1/}). There is an interesting initiative of connecting the graph elements and operations to model biological networks like the food web and food chain, and study their features using graph energy [7]. Praprotnik, et al. [8] nicely proposed the combined use of semiring axioms and the notions of graph theory in network analysis by combining weights on parallel edges using semiring addition and the weights on the sequential edges using semiring multiplication.

2. Facebook: An Example of Network Analysis using Algebraic Graph Operations

One of the largest social networking sites in the world today is Facebook. People use Facebook for different reasons or motives. One of the scientific research projects conducted by Robinson, et al. [9] suggests that there are four types of Facebook users. However, there may be some overlapping behavior exhibited by the users. The examples are summarized as follows.

1. Relationship builders

This type of Facebook user primarily focuses on fortifying real-life friendships. They will usually respond to other's posts and use Facebook to connect with family and friends.

2. Town criers

They use Facebook as a platform to inform about the events and happenings in the world and community. They are generally not actively engaged in posting status updates or uploading pictures of themselves.

3. Selfies

They use the site primarily for self-promotion or self-validation with status updates, pictures, and videos. They often post to collect likes and comments.

doi.org/10.32388/XAVXHO

4. Window shoppers

This group of Facebook users feels to have some sort of social obligation to be on the site. They are rarely interested in sharing details of their own lives, and nor do they do much liking or commenting, although they see Facebook as an inescapable part of modern life.

Note that different investigations suggest different categories of Facebook users. As highlighted, the categories are purely based on specific research or particular regions, that don't guarantee any rigid boundaries among different categories. Nor do we intend to comment on the validity of such categorizations. Without considering any statistical proof, we just take an intuition of such social categorizations within a bigger social network for our decision problems, where Facebook has been chosen as an example. We intuitively consider that there are no rigid boundaries among the categories such that a Facebook user preferably belongs to more than one category, and there may also be some other Facebook users who do not fit any of the given categories.

In this section, we take an example of a Facebook friendship network to propose an algorithm for determining the stability or affinity of connections between different social groups within a complex social network by decomposing the given network into certain components of predefined categories. We take the graph as the principal tool, and its operations, namely, $union$ and $intersection$ that form semirings on the set of graphs as the primary operations.

2.1. Algorithm

The following table shows the Facebook users belonging to five different intuitive categories.

Relationship builders	Window shoppers	Town criers	Selfies	others
$f, a, e, q, i, m, p, h, j$	b, d, c, h, f, n, r, s	a, j, b, g, k, s	c, d, g, k, l, o	l, o

The following graph represents the network of Facebook friendships.

doi.org/10.32388/XAVXHO

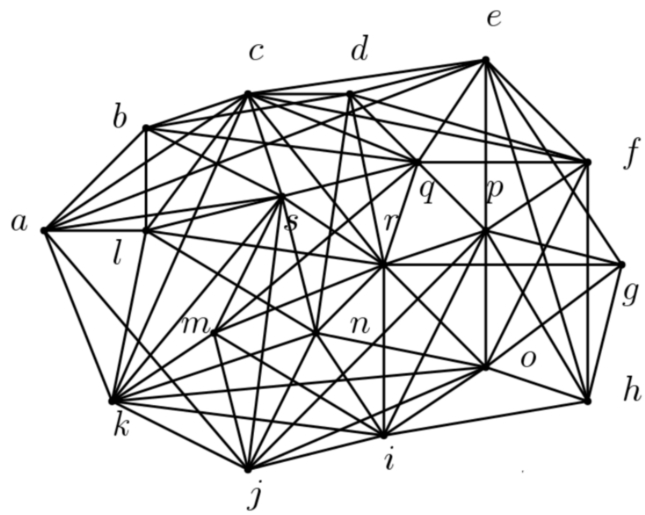

Figure 1. G

The given graph G is comparatively a complex network. Decompose the vertex set of G into five respective categories as listed above, forming five different components of G such that in each component, the respective edge set consists of all the edges in $E(G)$ that have both ends in the respective components. That is, components G_1, G_2, G_3, G_4 and G_5 are induced subgraphs of G. In this process, the original edges of G that connect across the components will be momentarily removed. Any two components will be connected by a path instead only if their intersection is a non-empty graph. The following is the reduced network of G decomposed into five connected components.

doi.org/10.32388/XAVXHO

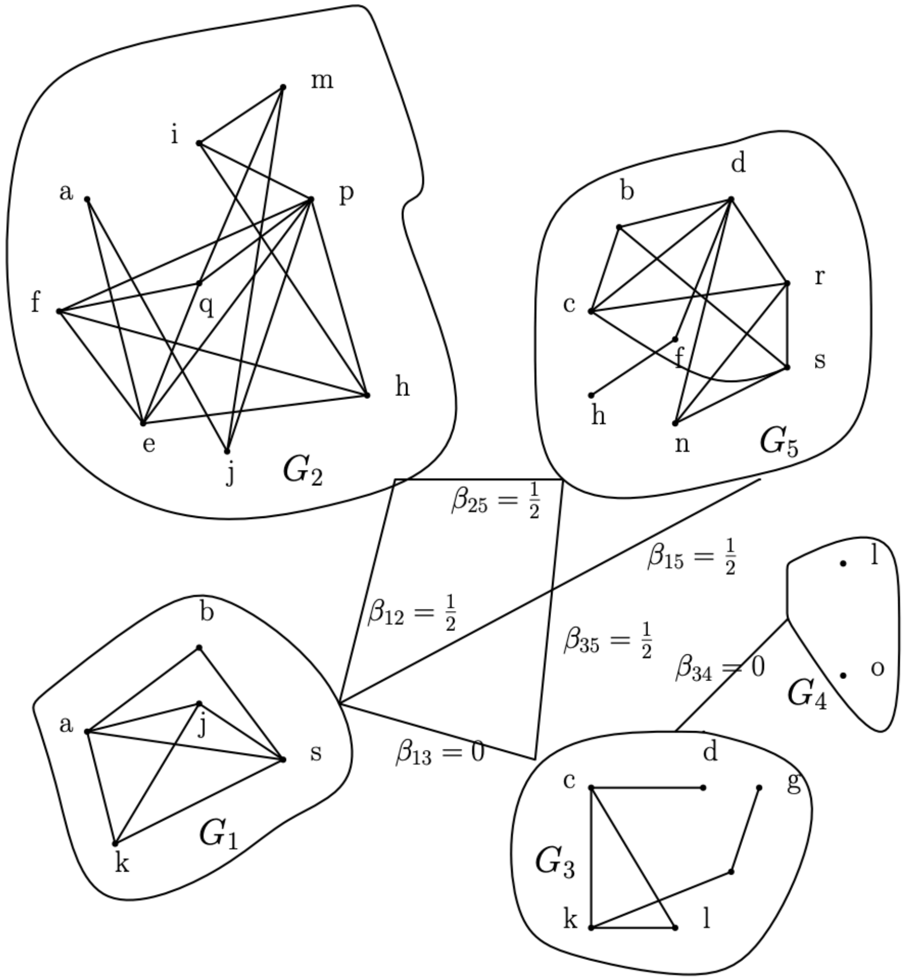

Figure 2. G'

The network G' can be further represented by the following simplified network, where each edge is an intersection graph of the corresponding end vertices. Subsequently, each edge is assigned its beta index.

 doi.org/10.32388/XAVXHO

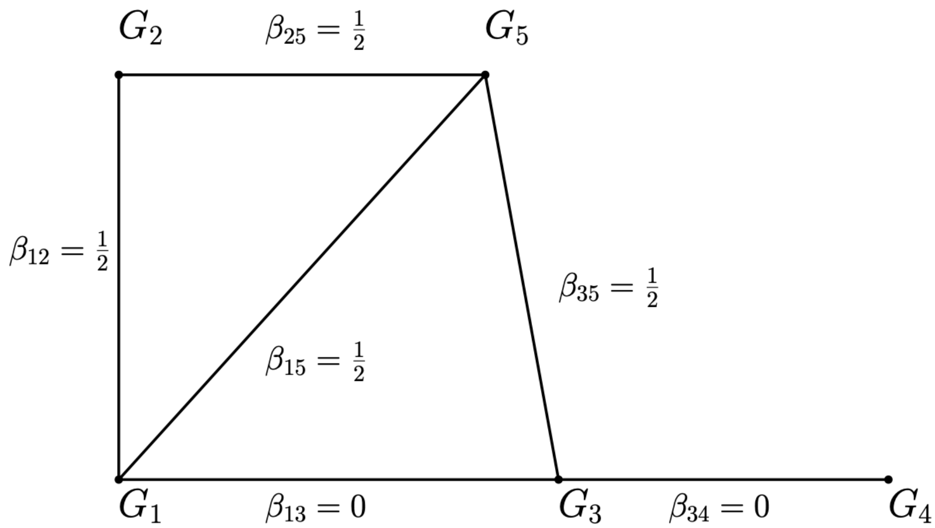

Figure 3. G''

While reducing the graph G to G' or G'', we have temporarily lost the original edges of G, connecting across different components. But this loss incurred will be compensated in the subsequent calculations as follows.

Let p_j be a path connecting G_2 and G_3 in the network G'', and C_{p_j} be the graph formed by the union of all the edges on the path p_j, whose edge set and vertex set are denoted by $E(C_{p_j})$ and $V(C_{p_j})$, respectively. Let $\beta_{C_{p_j}}$ be the beta index of the graph C_{p_j}. Then we express the weight of the graph C_{p_j} in terms of average vertex degree as $\boldsymbol{\beta}_{C_{p_j}} = 1 + 2\beta_{C_{p_j}} = 1 + A_{d(C_{p_j})}$, where $A_{d(C_{p_j})}$ is the average vertex degree of C_{p_j}, and the stability of the path p_j is

$$S_{p_j} = r_j + \boldsymbol{\beta}_{C_{p_j}},$$

where $r_j = 2 \times \frac{|\text{ Number of edges in } G \text{ formed by the vertex set } V\left(C_{p_j}\right) - \left|E\left(C_{p_j}\right)\right||}{\left|V\left(C_{p_j}\right)\right|}$. For instance, in this problem, four paths connecting G_2 and G_3, namely, $p_1 : G_2 - G_5 - G_3$; $P_2 : G_2 - G_5 - G_1 - G_3$; $P_3 : G_2 - G_1 - G_3$, and $p_4 : G_2 - G_1 - G_5 - G_3$. The following are different graphs formed by the union of all the edges on the respective paths connecting G_2 and G_3.

doi.org/10.32388/XAVXHO

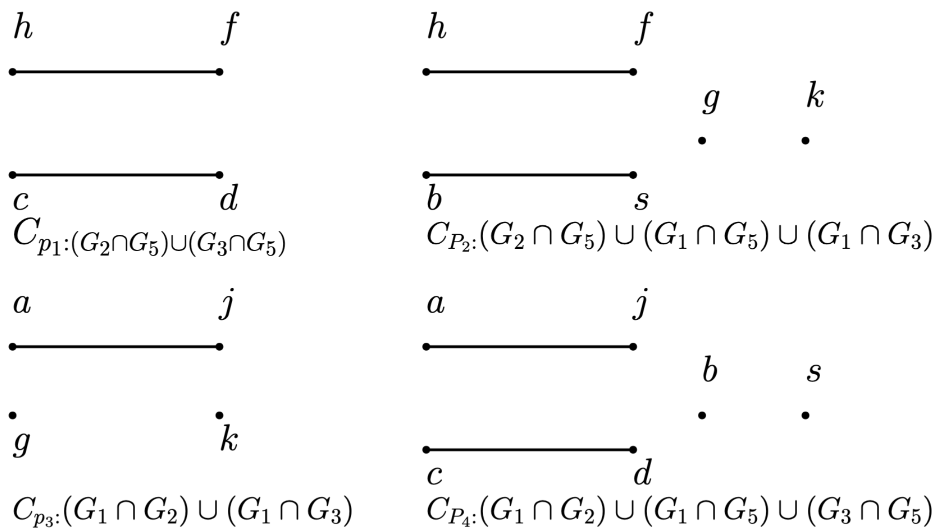

Figure 4. Graphs $C_{p_1}, \ldots, C_{p_4}$ are formed by the union of all the edges on the paths $p_1, \ldots, p_4$, respectively.

The stabilities of the paths are calculated in the following.

Calculation of the stability of p_1**:** From the above figure, the beta index of C_{p_1} is given by $\beta_{C_{p_1}} = \frac{2}{4} = \frac{1}{2}$. The number of edges in G formed by the vertex set $V(C_{p_1})$ is 4, namely, $(h,f),(c,d),(f,c)$ and (f,d), and the number of edges in C_{p_1} is 2, i.e., $|E(C_{p_1})| = 2$, namely, (c,d) and (h,f). Therefore, $r_1 = 2 \times \frac{|4-2|}{4} = 2 \times \frac{2}{4} = 1$. Hence the stability of the path p_1 is $S_{p_1} = r_1 + \boldsymbol{\beta}_{C_{p_1}} = r_1 + 1 + 2\beta_{C_{p_1}} = 1 + 1 + 2 \times \frac{1}{2} = 3$.

Similarly, the stability of the remaining paths is obtained as $S_{p_2} = 2.33$; $S_{p_3} = 2.5$, and $S_{p_4} = 4.33$. Thus, we conclude that p_4 is the most stable path connecting the relationship builders and the selfies in the given network.

Geometrical significance. One of Facebook's clever features is its friend suggestion, "People You May Know (PYMK)." It is for sure that Facebook doesn't suggest friends at random, but there are almost endless ways in which Facebook can suggest friends. The most common reason for PYMK pop-ups seems to be due to friendship networks. In line with this argument, we refer to an instance of the PYMK history of A's Facebook page. For example, at an instance we came across 150 individuals in the list of the PYMK; most of them have mutual friends with A, while 2-3 of them have no mutual friends with A.

doi.org/10.32388/XAVXHO

Interestingly, one individual among them, x (say), has only three friends in total, and none of her friends is A's mutual friend. The individual x is completely stranger to A; there is no way they would have shared their contacts, etc. They are supposedly neither on the same Facebook page nor do they have been to the same organization. We also presume that none of them would have visited one another's profile before. Despite this, how does she appear in A's PYMK list? To explore a probable reason, we browsed the list of her only three friends, and what we found is a likely convincing reason. One among her three friends, y (say), has one mutual friend z (say) with A. Again, z has 16 mutual friends with A, and each of those 16 friends has an average of 146 mutual friends with A and so forth. Apart from that, z also has some non-mutual friends, namely, p, q and r, etc., having a high number of mutual friends with A. Particularly, p, q and r have $658, 336$, and 539, respective mutual friends with A. Therefore, in this case, this friendship network is an important reason for x being included in A's PYMK list.

Our algorithm can most appropriately be linked with a way in which Facebook recommends friends. For instance, if Y is a friend of X, and Z is Y's friend, then Facebook may suggest Z as a friend of X and vice versa. Here, the point is that Facebook may suggest an unknown friend if you have some mutual friends. Another instance is when Y and Z are friends of X and P, where X and P are not friends, and Q is a friend of P. Then, Facebook may recommend Q as a friend of X and vice-versa. Here, X and Q have no mutual friends, but the PYMK may traverse the friendship networks either $X-Y-P-Q$ or $X-Z-P-Q$ or, both. Although such an algorithm may not always work. In line with this algorithm, we propose a geometrical significance of our algorithm. As per our calculation, the path p_4 is the most stable or, so to say, it has the greatest value of social affinity connecting two distinct groups of Facebook users. We note that C_{p_4} is the graph corresponding to the path p_4, obtained by combining all the edges on that path, whose vertex set is $\{a, b, c, d, j, s\}$. Therefore, we conclude that Facebook users a, b, c, d, j, and s play the most vital role in connecting the relationship builders and the selfies in the given Facebook network. Consequently, Facebook may recommend members in relationship builders as potential friends of members in the selfies and vice-versa, because of the members a, b, c, d, j, and s in the network. In this problem, we see that every member of the relationship builders is either friends or has some mutual friends with the corresponding member of selfies except h and l. That is, h and l are neither friends nor have mutual friends, but o and n being a friend of h and l, respectively are friends or have mutual friends, namely, i, j, k and r, and so forth. So, we conclude that every member of the relationship builders is a potential friend of selfies (except those who are already friends), and vice-versa.

doi.org/10.32388/XAVXHO

REMARK 1. If all the members in the respective categories are all strangers or so to say, all the components in G' are discrete graphs, then the graph C_{p_i} will be discrete, hence $\beta_{C_{p_j}} = 0$ or $\boldsymbol{\beta}_{C_{p_j}} = 1$. Therefore, the stability of the path p_j is given by

$$
\begin{aligned}
S_{p_j} &= r_i + 1 \\
&= \frac{1}{2} \frac{|\text{ Number of edges in } G \text{ formed by the vertex set } V\left(C_{p_j}\right) - \left|E\left(C_{p_j}\right)\right||}{\left|V\left(C_{p_j}\right)\right|} + 1 \\
&= \frac{1}{2} \frac{|\text{ Number of edges in } G \text{ formed by the vertex set } V\left(C_{p_j}\right) - 0|}{\left|V\left(C_{p_j}\right)\right|} + 1 \\
&= \frac{1}{2} \frac{\text{Number of edges in } G \text{ formed by the vertex set } V\left(C_{p_j}\right)}{\left|V\left(C_{p_j}\right)\right|} + 1
\end{aligned}
$$

Clearly, in this case, if $S_{p_j} = 1$, then the number of edges in G formed by the vertex set $V(C_{p_j})$ must also be zero, showing that none of the Facebook users in the set $V(C_{p_j})$ are friends.

Conflict of Interest

The authors declare that there is no conflict of interest.

Other References

- Boulmakoul, A., Generalized path-finding algorithms on semirings and fuzzy shortest path problem, Journal of Computational and Applied Mathematics, 162 (2004), 263-272.

References

1. ^*Flitter, H., and Grossmann, T, Accessibility Network Analysis, Geographic In- formation Technology Training Alliance (GITTA). Retrieved from www.gitta.info - Version from 28.4.2016.*
2. a, b*Umbrey, G. and Rahman, S., Application of Graph Semirings in Decision Net- works, Mathematical Forum, 28(1) (2020), 40-51.*
3. ^*Deo, N., (1974), Graph Theory with Applications to Engineering and Computer Science. Prentice-Hall, Inc., Englewood Cliffs, N. J.*
4. ^*Umbrey, G., Study of semirings (Doctoral thesis, Rajiv Gandhi University, Rono hills), Shodhganga@INFLIBNET, http://hdl.handle.net/10603/343412, (2021).*

5. ^*Rahman, S. and Umbrey, G., On Some properties of Semirings of Graphs, South- east Asian Bulletin of Mathematics, 46(4) (2022), 553-563.*
6. ^*Rajkumar, M., Jeyalakshmi, S., and Chanmdramouleeswaran, M., (2015), Semiring-Valued Graphs International J. of Math. Sci. and Engg. Appls. (IJM- SEA), Vol. 9 No. III., 141-152.*
7. ^*Umbrey, G. and Rahman, S., Determining Paths Energy of a complex network, Advances in Mathematics: Scientific Journal, 9(10) (2020), 8761-8770.*
8. ^*Praprotnik, S., and Batagelj. V., (2016) Semirings for temporal network analysis. Retrieved on 25, April 2020 from https://arxiv.org/abs/1603.08261.*
9. ^*Robinson, Tom, et al."I LOVE FB: A Q-Methodology Analysis of Why People 'Like' Facebook." IJVCSN 9.2 (2017): 46-61. Web. 16 Aug. 2020. doi:10.4018/IJVCSN.2017040103.*

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.