International Journal of Progressive Sciences and Technologies (IJPSAT)
ISSN: 2509-0119 Vol. 3 No. 1 June 2016, pp.18-22

http://ijpsat.ijsht-journals.org

Profiling People in Social Graphs

Abhijit Kashyap[1], Anoop Aparajit[2]

[1]Computer Science and Engineering,
Vellore Institute of Technology,
Vellore, Tamil, Nadu, India

[2]Computer Science and Engineering,
Vellore Institute of Technology,
Vellore, Tamil, Nadu, India

ABSTRACT: Our world is now a connected place. Through social networking, the Internet has brought people closer together than they ever have been. The result is a complex, massive and global web or graph of people and the relationships between them. This is both useful and dangerous. In this paper, we aim to reveal the power of the social graph, showing how it can be used to gather information about a person and build an accurate profile about him. A social graph can also tell us the 'proximity' of a person to others, which in many cases can provide valuable insights to security agencies and governments.

KEYWORDS: Social graphs, communication, social media, connectedness, relationship classification

1. INTRODUCTION

A social graph is a representation of the interconnections of relationships in an online social network. [1] Analysing a social graph can reveal interesting patterns and conclusions about a person's social life. These conclusions can be used for profiling a person. Such profiling has uses in various fields related to human resources.

This paper aims to provide measures that indicate how close and connected two people are in a social graph, and to draw conclusions from them

2. THEORY

A weighted graph is a graph in which each edge or vertex is assigned some weight based on certain heuristic values that are predetermined. An example of a vertex-weighted graph [1] is given below.

Corresponding Author: Abhijit Kashyap

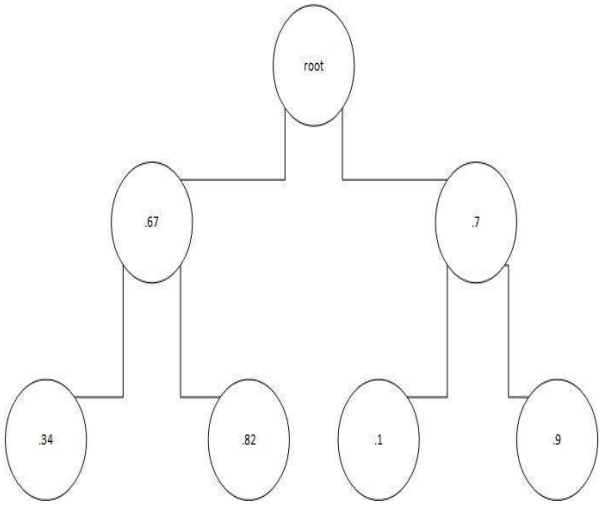

Fig. 1. Vertex-Weighted Graph

In this paper, we combine the concepts of weighted graphs and social graphs. An example of a weighted social graph is shown below.

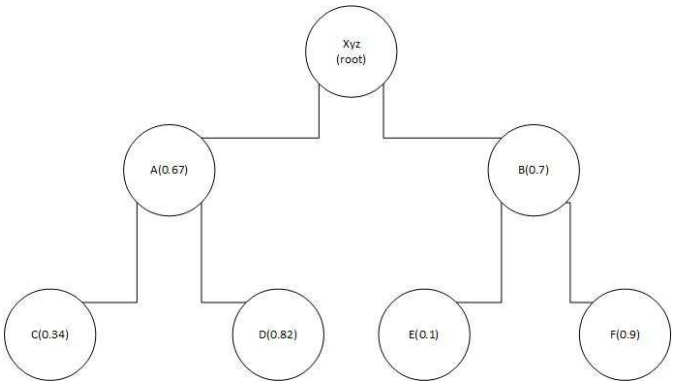

Fig. 2. Vertex-Weighted Social Graph

We define two terms at this stage – closeness and connectedness. Closeness between two people, or nodes, is a measure of the familiarity, intimacy and interaction between them. Connectedness is a measure of how easy it is for one person to 'reach' another.

These measurements are done considering one person in the social graph as the root node of the graph. Thus, the closeness value of an arbitrary node Y in the graph (that is not the root node) indicates the closeness that Y has with the root R, but not the other way around. The relation is not commutative, as we will soon see.

Generally, social networks allow their users to form close bonds with certain people and not with others. For example, Facebook uses the concept of friends, Twitter uses followers, and Google+ uses circles. For the sake of simplicity, let us call this kind of relationship as a 'friend relationship'.

2.1 CLOSENESS

Similarly, each of these networks allows users to 'like', 'share' and 'comment' others' posts. These actions may have different names on different social networks (for example, sharing on Twitter is called retweeting), but the underlying action is the same. Again, for simplicity, let us use the verbs 'like', 'share' and 'comment'.

Now, we determine the closeness value of a node Y with respect to a root node as follows.

1) For every friend relationship – add 1
2) For every like – add 2
3) For every share – add 3
4) For every comment – add 4

Friend relationships between nodes are only counted once on Facebook and Twitter. On Google+, however, a person Y and a root R may be in multiple circles together. Likes, shares and comments may occur any number of times. For example, if a person Y is friends with a root node R on Facebook, then the closeness value of Y is 1. If Y likes four of R's posts and shares one of them, then the total closeness value of Y with respect to R is

$$1 + 4*2 + 1*3 = 12$$

These values have been decided assuming that comments represent a higher intimacy level than shares, which are higher ranked than likes, and so on. A friend relationship is given the least importance because very often, people tend to make friend relationships with even acquaintances or celebrities. Friend relationships are thus not the best indicator of closeness.

2.2 CONNECTEDNESS

Of course, closeness, as we have defined it, alone cannot be an accurate measure of the real relationship between people. We need another attribute. By also considering the connectedness of a node Y with respect to a root R, we start to form a clearer picture of the relationship between Y and R. The connectedness between node Y and root R is determined as follows.

For every edge contained in a distinct path between Y and R, assign a value of 1. For example, consider the following graph.

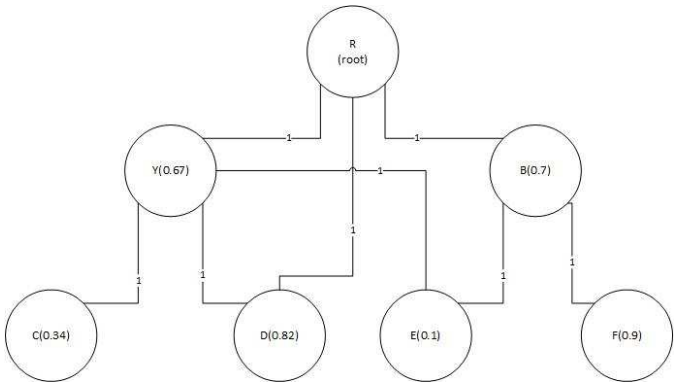

Fig. 3. Sample social graph to show connectedness

There are three distinct paths between node Y and root R. One of the paths is a direct edge between Y and R, indicating a friend relationship.

Now, the connectedness can be considered in two ways.

1) Adding the cost of all the paths between Y and R. Let us call this the total connectedness.
2) Taking the minimum weighted path between Y and R. Let us call this the minimum connectedness.

Note that the connectedness of Y with respect to R might not be the same as the connectedness of R with respect to Y. On Twitter, Y might follow R, but R might not follow Y. Thus, there would be a directed edge from Y to R, and no directed edge from R to Y. However, for simplicity, let us say that the friend relationship is always mutual. Thus, the connectedness relationship is commutative, and we can say that the connectedness of Y with respect to R is the same as the connectedness of R with respect to Y, which is the same as the connectedness between Y and R, or between R and Y.

The following results can be drawn from the connectedness:

1) If the minimum connectedness between Y and R is 1, then Y and R have a friend relationship.

2) The higher the value of the total connectedness between Y and R, the deeper Y and R are in the same social circle.
3) If the minimum connectedness between Y and R is high, then Y and R are in different social circles.

Note that when we use the term social graph, we always mean a finite social graph, which terminates at a certain depth from the root node R. If we were considering the entire social graph structured around a root R, the social graph would almost definitely cover the entire social network, and the total connectedness between Y and R would become meaningless. For example, even a school going teenager living on a remote island nation may be connected to the President of the United States of America through his friends, friends of friends and so on. The total connectedness of this teenager with the President would be a huge value, indicating that these individuals are in the same social circle, which is absolutely incorrect. Thus, we need to restrict the depth of our social graph based on our needs. Interestingly, the minimum connectedness may be smaller than it may seem. [2]

By considering a combination of the minimum connectedness and total connectedness, other conclusions may be drawn.

3. COMBINING CLOSENESS AND CONNECTEDNESS

We now have two parameters two measure relationships in a social graph – closeness and connectedness - which are useful on their own and potent when put together.

We can draw many conclusions by taking combinations of values of closeness and connectedness. They are summarized below.

1) If minimum connectedness = 1, we can make a list of closest friends based on the closeness value as shown for the graph below.

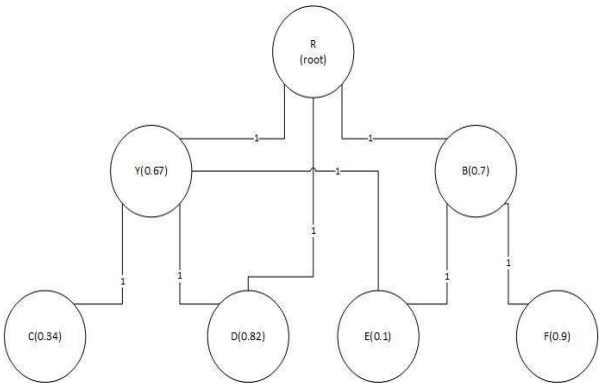

Fig. 4. Sample social graph

Table 1. Closeness of people to root R

Person	Closeness
Y	0.67
B	0.7
D	0.82

2) Minimum connectedness between **root** and a random node **y** is only a measure of the likeness of root knowing y.

3) No of Paths (connections) between root and y shows how likely it is for the root to receive any news about y.

4) Profiling based conclusion -

Table 2. Classifications drawn from connectedness and closeness

Connectedness	Closeness	Result
Low	Low	Acquaintance
Low	High	Genuine Friends
High	Low	Different Social Circles
High	High	Fans/Obsessed People

5) Another profiling based conclusion -

Table 3. Classifications drawn from number of connections and closeness

Number of connections	Closeness	Result
Low	Low	Unknown
Low	High	Personal
High	Low	Same circle/Common Friends
High	High	Personal, Same circle

4. CONCLUSIONS

The method described in this paper can be used as a reasonably accurate classification technique. It depicts the obvious relations between people and discovers the non-obvious ones. Practically, this can be used by social networking companies to make better predictions about people and provide better user experiences. For example, a person might only want to receive updates, news and notifications from his or her closest friends and not from acquaintances.

It could also be used to logically group news items together. For example, notifications received from people with similar closeness and connectedness levels may belong to the same social circle, or implicit social group. [4]

ACKNOWLEDGMENT

We thank our institution, VIT University, for providing us with the opportunity to research about this topic and make a paper on it.

REFERENCES

[1] [Online] Available: https://en.wikipedia.org/wiki/Glossary_of_graph_theory

[2] [Online] Available: https://research.facebook.com/blog/three-and-a-half-degrees-of-separation/

[3] Eric Gilbert and Karrie Karahalios ,"Predicting Tie Strength With Social Media," CHI 2009, April 4–9, 2009, Boston, Massachusetts, USA.

[4] Maayan Roth, Guy Flysher, Naty Leiser, Assaf Ben-David, Ilan Horn, Yossi Matias, David Deutscher, Ari Leichtberg, Ron Merom," Suggesting Friends Using the Implicit Social Graph," KDD'10, July 25–28, 2010, Washington, DC, USA

[5] Roth et al. - Suggesting (More) Friends Using the Implicit Social Graph [Online] Available http://static.googleusercontent.com/media/research.google.com/en//pubs/archive/37120.pdf