AddDel Model for the Evolution of Social Networks

Shikha Sarathe
Research Scholar (Computer Sci. Engg),
Vaishnavi Group of Institutions Bhopal- 462001 Madhya Pradesh, India)

Sumit Sharma
Asst. Professor (Computer Sci. Engg),
Vaishnavi Group of Institutions Bhopal- 462001 Madhya Pradesh, India)

Abstract
In the past decade we have seen that researchers have focused on the study of networks which are complex, dynamic and hierarchical structured. With the evolution, the network become more dense and dense, hence we say that the network shows a dynamic nature. Researchers have spot that the evolution process for Social or Complex network is not an arbitrary process rather there subsist some fundamental reasons behind their evolution process. For this particular reason we have designed the network as an undirected graph where vertex represents a member in the network and edges represents the relation among them. We create a model in which the edges appear as well disappear based on some parameters. In this work we consider deletion of edges apart from the addition. The deletion strategies consider the newly made links and the existing links. In this paper we present a model and also discuss whether this model predicts real social networks.
Keywords: Social Networks, Score, Factors, Graph, Communities, Gephi, Networkx.

1. Introduction

A network is a set of objects, which we call vertices or sometimes nodes, with connections between them, called edges shown in Fig 1. A network is a mathematical object consisting of a set of nodes that are connected to each other by edges. Social networks easily make connections between the individuals based on common interests, groups (i.e., school, friends, society, etc). With the help of social networks, it is easier for people to search and communicate with particular person who are in their networks. In past few years, due to the exceptionally large growth of social networks, researchers have turned their interest towards the evolution of social networks. Many models have been advanced which gives the insight into the evolution process of these networks. Most of these models shows the growth of network with respect to time [1]. In this paper, we represented a social network as a network having members which are linked to each other on the basis of relation so called friendship. Social networks evolves with the time, so with the mechanism of evolution we have designed a model which is somehow related to real-world networks. Degree distribution of node plays a vital role in networks which is the probability P(k) of a node to have degree k. Some important laws of degree distribution are: power law with $P(k) \sim k^{-\gamma}$ where γ represents the degree exponent, and exponential law with $P(k) \sim e^{-\frac{k}{c}}$ where c is constant.

Suppose, if A and B are two members connected in a network then they must have share common interests, interactions, characteristics etc. In our work, we termed these characteristics as factors. The more is the factor the more characteristics, interaction or interests are similar. By considering more numbers of factors, we can have more similarities among the individuals in order to have a more precise understanding of the model. For example, in an institute specialization plays the most significant role in determining the degree of friendship among the network members.
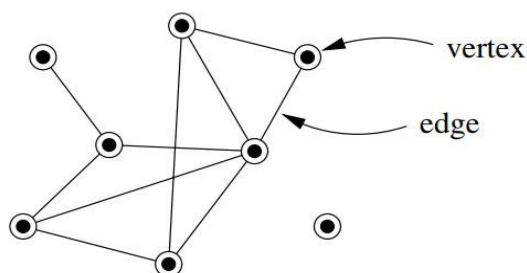

Figure 1. Example of network with eight vertices and ten edges.

The solidity of the companionship is dependent on the number of familiar factors. The term "quality" can be compared with these relationships in order to know how strong is the bonding among the individuals. In the upcoming section, we will deal with the works related to the evolution process.

1.1 Related Work

Till date, many advancement models have been proposed related to Social Network. However, most of the paper focuses on addition of nodes or links. Networks in which vertex and edges can be added and/or deleted is called as dynamic networks [2]. In such types of network, edges are generally created with the help of triad closure and focal closure [3]. If any node is deleted then we can say that the adjacent links gets deleted too. There are various social network prediction methods which focuses on predicting new links. Apart from this researchers have made some efforts to know what are the parameters accountable for the evolution process of the network. After studying network evolution strategies researchers have concluded that the outcome of evolution process is not due to centrality but are the results of the simultaneous choices of network members. Several network formation strategies were probe showing that edge locality plays a major role in evolution process [4]. One of the most famous random graph model is AB model [5] and that are frequently used to model social networks because it is very specific to the real-world network but it mainly focuses on the complex networks like communication network, transportation network, financial network, biological network etc.

2. Our contribution

Here, we will discuss the model proposed by us. We have proposed a non-deterministic approximate based model on the role played by nodes in a social network in terms of their interactions with other nodes. The basic idea used for the evolution process is the transitive property and focal closure. The transitive property is the property among three nodes say X, Y , and Z, such that if their exists a tie between X-Y and X-Z, than there is high chance to be a tie between Y-Z. For example if more friends are common, then more chances of forming a new links among themselves. The focal closure means that nodes say Y and Z that must share a common interest as well as interaction if they are connected to each other. It has been observed that these network evolves with time and their evolution is not randomly processed. Some properties are accountable for the evolution process like preferential attachment, degree-distributions etc.

2.1 Preliminaries

The network is modelled as an undirected graph G= (V,E) where the nodes (V) represent the network members and an edge e = (u,v) ε E represents the relationship between them. There must be some common characteristics (factor) for any type of social relation, so we have attached factor with each of the edge. If any nodes gets attached by an edge than they must have some common factors. In order to decide the quality of relationship, we have also attached score value with these factors. To every factor in set F_n a score value s is assigned to it. For example if $f_1, f_2 , f_3, f_4 ....., f_n$ is a set of factors for an edge en then $s_1, s_2 , s_3, s_4 ...., s_n$ are the score values associated with these factors. In this model, a cumulative score value S_n is defined for n= 0,1,2,3,...k where k is total number of edges in graph G. Fig 2 shows the initial graph of social network. Let $P_d^t \approx f_t(S,S_e)$ where P_d^t be the probability of deleting an edge (d) at time t, S is the existing edges in network and S_e is the newly added edge in the network with the vary in time and **α** is some exponent, **t** is the threshold value for the addition of a new edge and ʈ is the threshold value for the deletion of an edge, **α** = ʈ = 0,1,2,3... In the following section we will describe the evolution process for our model.

2.2 Evolution Process

Assume, the initial graph be G_0 and set of score values be $score_0$ which is the association of the all the scores of the edges present in G_0 at time i = 0. Periodically an edge will be added and=or deleted between the nodes according to the evolution rules. Assume at time stamp i, $G_i = (V, E_i)$ is the evolved graph from the base graph then for all i≥ 0, $G_i = (V, E_i)$ is subgraph of $G_{i+1} = (V, E_{i+1})$ i.e. $E_i \subseteq E_{i+1}$.
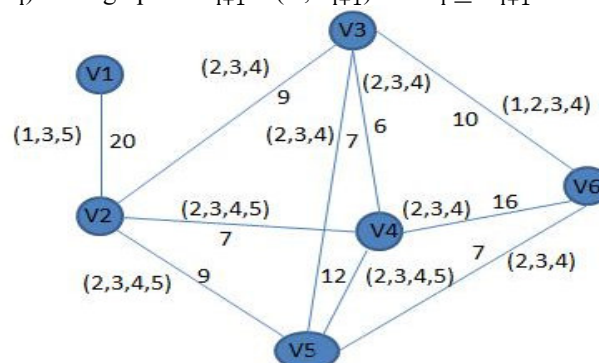

Figure 2. A graph with six nodes.

According to our algorithm, once a new edge is linked during the evolution process in the graph, their is a

very high chance that the newly added edge will not disappear in the coming time-stamp but at the same time if an edge is deleted their is a very high chance that it will never be added. We have considered in our evolution process that the set of factors as well as the score value will remain the same throughout the algorithm. In the next part, we will discuss the evolution rules of the model

2.3 Rules of Evolution

In this section we define the procedure of our algorithm, how it works with the time period. This model mainly focuses on the addition as well as deletion of nodes based on real life observations that for two strangers to become friends they use friend-of-a-friend technique which is as same as transitive property. A new node is added if the cumulative score is more than a specific threshold t and an existing node is deleted if the cumulative score (including newly added edge) is more than a particular threshold $Ţ$. The cumulative score is the intersection of the sets F_{uw} and F_{wv} and then computing the maximum of $\frac{|F_{uw} \cap F_{wv}|}{2} \cdot (\frac{S_{uw}}{|F_{uw}|} + \frac{S_{wv}}{|F_{wv}|})$ is applied over all possible $w \in V$.

The reason behind the use of cumulative function is as follows:

- The fraction $\frac{S_{uw}}{|F_{uw}|}$ means the *score* of a factor between the *edge (u, w)* and by taking the average of $\frac{|F_{uw} \cap F_{wv}|}{2}$ gets the mean of a *factor* which will be common in both.
- Mean is multiplied by the $|F_{uw} \cap F_{wv}|$ results in new score value, so on the basis of new cumulative score value it is decided whether the new edge will be added or not.

Algorithm 1: Evolution Process

1    $G_0 = (V_0, E_0), t, \alpha, Ţ$;

2   i ← 0;

3   **while** *no more edges can be added* **do**

4   for each *(u, v)* is not added, and $\forall w$ such that
         (u,w), (w, v) εE_i,   find,

5    $z = arg\ max_w\ (\frac{k}{2} \cdot (\frac{S_{uw}}{|k_{uw}|} + \frac{S_{wv}}{|k_{wv}|}))$ where
     $k = F_{uw} \cap F_{wv}$.

6   **if** $(\frac{k}{2} \cdot (\frac{S_{uz}}{|k_{uz}|} + \frac{S_{zv}}{|k_{zv}|}))_w > t$ **then**

7   add *(u, v)*   with probability *p* and discard with
          probability *1-p*;

8   **end**

9   **for** *already exisiting edges e* **do**

10   **if** $(\frac{Z}{S_e})^\alpha > Ţ$ **then**

11   delete e ;

12   **else**

13   don't delete ;

14   **end**

15   **end**

16   **end**

17   i ←   i + 1

and that are frequently used to model social networks because it is very specific to the real-world network but it

mainly focuses on the complex networks like communication network, transportation network, financial network, biological network etc.

On new cumulative score value, the evolution rule checks whether their will be a new link or not by comparing it with the threshold value t. If the new score is greater than the threshold then a new edge will be formed in the network with probability *p* else rejected with probability *1-p*. Suppose an edge is added in the network then we will check that due to the newly added edge can any node show adverse effect i.e. if a node is in level 1 then it has strong connections with the nodes of level 1, if the same node goes to level 2 then it has strong connections with the nodes of level 2, at the same time their is a very high possibility that the same node has weak connections with the level 1, so we can infer that as we go far the connections have adverse effect on the previously added nodes. Due to this reason some edges gets deleted in the network. The fraction $\frac{z}{s_e}$ represents the cumulative score value to the score value of any random edge *e*. If the value of $(\frac{z}{s_e})^\alpha$ is greater than the threshold value $\mathcal{T}$ of for any random exponent it means that a particular edge has adverse effect due to the addition of new edge. So delete all the edges *e* with the same score value on which we have checked for and if it is less than the threshold $\mathcal{T}$ the graph will remain the same i.e. no effect on the existing edge due to the addition of a new edge.

After the completion of evolution process for time stamp i, it gets increased by 1 and follow the same rules for evolution of social network. Here we have described the evolution rules for the network evolution. In Fig 2 we will apply our algorithm to see the evolution. This figure consists of six nodes having an edges among them along with the *factors* as well as *score value* on the edges. For example V_2 is linked with V_3 with the set of *factors* $F_{v2,v3}$ as {2,3,4} and *score value* $S_{v2,v3}$ as 9 and V_3 is linked with V_6 with the set of *factors* $F_{v3,v6}$ as {1,2,3,4} and *score value* $S_{v3,v6}$ as 10. So on applying the score function we get the common *factors* $F_{v2,v6}$ as {2,3,4} and *score value* $S_{v2,v6}$ as 8. In view of the fact, the calculated score value is greater than the threshold value *t* (say 6), so there will be a new link formed between V_2 to V_6. Fig 3 shows the modified graph of social network after addition of an edge with the evolution rules.
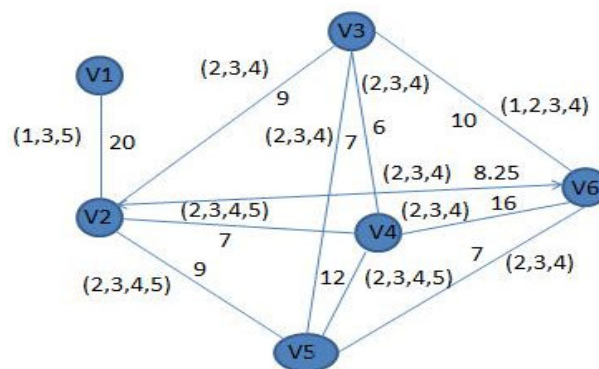

Figure.3. A graph with the addition of an edge

At the same time, a graph is to be checked for the deletion of an edge as well. From Fig 3 randomly choose an edge with any *score* value as 7 (say) between V_2 to V_4 and let the threshold value be $\mathcal{T}$ 5. Here we will check for two cases over different α i.e. for (α = 10) and (α = 11). So after applying the score function on that particular edge V_2 to V_4 we observe that

- for α = 10, we get the score value $s_{v2,v4}$ as 4.80 which is less than the threshold value. So the graph will remain the same.
- for α = 11, we get the score value $s_{v2,v4}$ as 5.62 which is greater than the threshold value so all the edges with the *score value 7* will get deleted and the edges are $(V_2 - V_4)$, $(V_5 - V_6)$ and $(V_3 - V_5)$.

A modified graph of social network after deletion of an edge is shown in Fig 4. In this model we have added and deleted the edges with some proximity. This algorithm will iterate until unless no further edges is either added or deleted.

3. Experiments and results

For the implementation of our evolution model we have assumed some parameters. This model is tested out experimentally by the use of *Networkx* [6] which is a module in python and *Gephi* [7] as the visualization tool for mapping out the graphs so formed and observing the change in its properties with the change in different parameters in the rules of evolution. After giving the input, the program runs to have a graph as output. The graph shown in Fig 2 called as "initial graph" where we will apply the algorithm for the evolution process. On this initial graph we apply our algorithm which results in the formation of a new graph with the appearance of a new edge or disappearance of an existing edge(s) randomly. Set of *factors* is defined over all the edges randomly differs from a person to person in the network which are responsible for the addition and deletion of the links. In order to know

how strong the *friendship* is we have defined score values on each edge randomly as well. For example, *a* and *b* are two node and we want to form a connection between them with the function *f(a,b)* based on set of *factors F* and *cumulative score value S* then we say
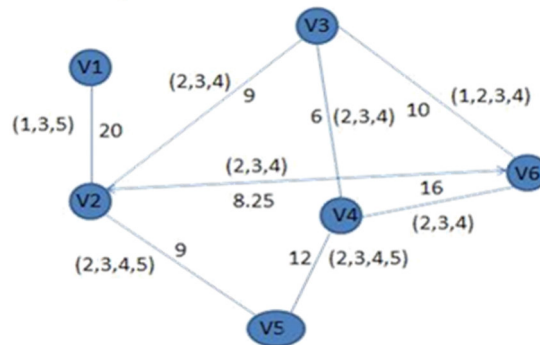

Figure.4. A graph with the deletion of an edge.

$f(a,b) = \alpha(S_{a,b}) + \beta(F_{a,b})$, where $F_{a,b} = f_a \cap f_b$, $S_{a,b} = \frac{|S_a \cup S_b|}{2}$ and α, are the balance factor to control the balance. We consider the experiment on an undirected graph with different number of nodes which will show different structure for the evolution network. In this experiment we took finite number of factors as well as score value not to have complex networks. The score value on an individual edges is taken randomly from the set of all score values S = (1, 2, 3,.....,8) and in similar way factors were also taken randomly from the set of all factors F = (1, 2, 3,.....,20). Repeatedly we have to run this program for different number of nodes.

3.1 Experiments on the vary nodes

Here, we will notice the results for different number of nodes (i.e. 400 and 500).

1) *Evolved graph for 400 nodes:* The evolved graph is shown in Fig 5.
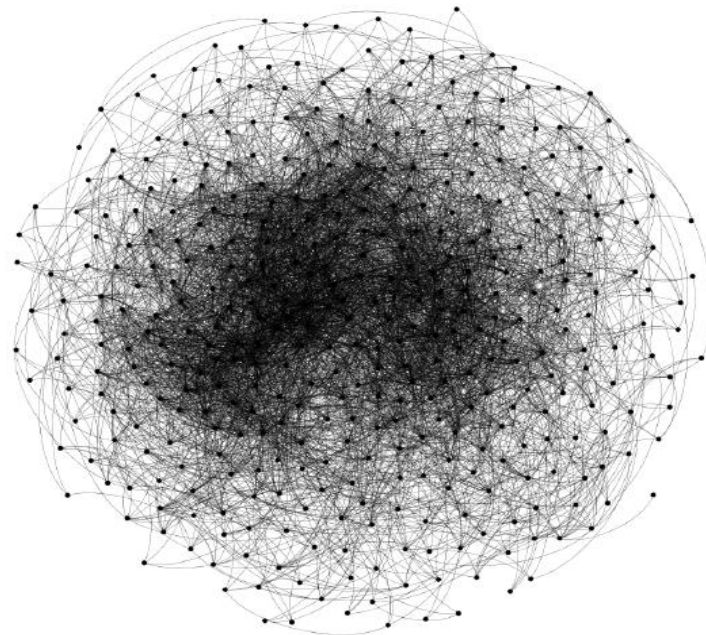

Figure.5. Evolved graph for 400 nodes.

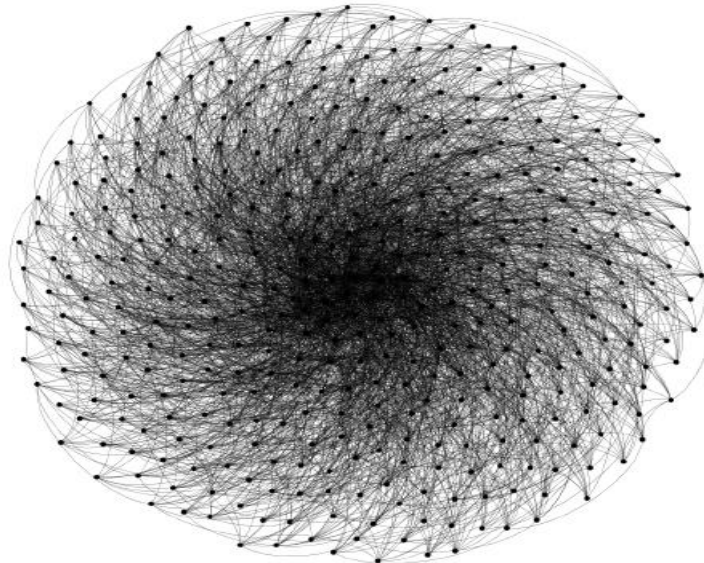2) *Evolved graph for 500 nodes:* The evolved graph is shown in Fig 6.

Figure. 6. Evolved graph for 500 nodes.

We noticed in both the evolved graph shown in Fig 5and Fig 6 that their is a formation of community structures (i.e. common location, occupations, interests, interactions etc) which are quite common in real networks. A member in a graph is not only a part of just a community but also may be a member of various communities existing in the network.

4. Conclusion

In our work we have designed the social network model on the basis of relation between the two person who satisfies the transitive property (i.e. friend-of-a-friend). In this paper we have elaborated the model of social networks of [8] for the evolution process. In this work we also consider deletion of edges apart from looking at the addition of edges. Our improvisation is based on the real life scenarios i.e., in a social network nodes not only get added but also deleted over a period of time. Since the process of removal of these links seems to be dependent on the scores of the newly made connections, our deletion strategy considers the scores of the edges (newly added edge and the already existing ones) while deleting an already existing edge. From our experiments result we can conclude that their is a high resemblance with the real data sets. In future, we will try to understand the consequences on nodes after the deletion of edge(s) and see whether our model predicts accurately the real life social networks

References

S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, "Structure of growing networks: Exact solution of the barabasi′–albert's model," arXiv preprint cond-mat/0004434, 2000.

E. Ben-Naim and P. Krapivsky, "Addition–deletion networks," Journal of Physics A: Mathematical and theoretical, vol. 40, no. 30, p. 8607, 2007.

R. Toivonen, L. Kovanen, M. Kivela,¨ J.-P. Onnela, J. Saramaki,¨ and K. Kaski, "A comparative study of social network models: Network evolution models and nodal attribute models," Social Networks, vol. 31, no. 4, pp. 240–254, 2009.

P. Doreian and F. N. Stokman, Evolution of social networks, vol. 1. Psychology Press, 1997.

A.-L. Barabasi,ˆ H. Jeong, Z. Neda,´ E. Ravasz, A. Schubert, and T. Vicsek, "Evolution of the social network of scientific collaborations," Physica A: Statistical mechanics and its applications, vol. 311, no. 3-4, pp. 590–614, 2002.

A. Hagberg, P. Swart, and D. S Chult, "Exploring network structure, dynamics, and function using networkx," tech. rep., Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008.

M. Bastian, S. Heymann, M. Jacomy, et al., "Gephi: an open source software for exploring and manipulating networks.," Icwsm, vol. 8, pp. 361–362, 2009.

A. K. Verma and M. Pal, "Evolving social networks via friend recommen-dations," in Signal-Image Technology & Internet-Based Systems (SITIS), 2015 11th International Conference on, pp. 379–383, IEEE, 2015.