Multi-objective Restructuring in Social Networks

R. Chulaka Gunasekara, Kishan Mehrotra, Chilukuri K. Mohan
Department of Electrical Engineering and Computer Science
Syracuse University, NY 13210-4100
{rgunasek, mehrotra, mohan}@syr.edu

Abstract—In most social networks that are observed over time, we find that some individuals leave and others join the network. It is often important to modify the connections in the resulting network to satisfy desired properties associated with the network as well as individual nodes. We formulate this as a multi-objective optimization problem that requires maximization of two measures: the network Information Flow Quality (IFQ) and the Personal Satisfaction Quality (PSQ). Algorithms are developed to accomplish these optimization tasks, and shown to result in satisfactory network reconfiguration.

Keywords—Social network analysis; Information flow; Network re-structuring; Centrality; User satisfaction

I. INTRODUCTION

Social networks are extremely dynamic in nature. Dynamic behavior of networks can occur in different forms such as people leaving the network, new people joining the network, and changing the existing connections in the network. An important research problem is to determine how the network should restructure to cope with such dynamic behavior.

Restructuring the network following a node deletion has been addressed by [1, 2, 3]. In these studies, restructuring is accomplished by using heuristics such as exploring the role of the deleted node (see [1]) and adding just enough links to maintain the quality of information flow (IF) (see [2]). Game theoretic models have also been studied [4] to capture multiple conflicting objectives of self-interested members in network formation [5].

Diffusion of information is influenced by the structure of social networks [6, 7]. Members in a social network would like to obtain new information at a quick rate, hence the network structures which enable quick dissemination of information have been studied [8]. Measurements to capture the properties of IF in social networks have also been proposed [2].

Different members in a social network play different roles when it comes to IF in the network. These members can be classified based on their placement within the structure of the network [9]. Recent research has examined the problem of node classification in dynamic networks with both text content of IF and the structure of the network [10]. Proper identification of the key players in a network has an important role in IF and other dynamics of social networks [11].

This paper proposes a new IF policy for social networks, based on the greedy behavior of members that attempt to send information quickly to more significant members in their neighborhood. We also propose two new network quality measures which capture the IF rate and satisfaction of a member based on the placement. These measures are loosely coupled with the IF policy, so that applicability extends to cases where different IF policies govern the IF. The proposed measures are then used to identify the best placement for a new member and the necessary changes to make when members are removed from the network.

The rest of the paper is structured as follows. Section II discusses the proposed new IF policy. Section III presents the new IF quality measure and Section IV presents the measure to capture an individual’s satisfaction. Sections V and VI illustrate how these network measures are used to obtain the best structure in the cases of addition and deletion of a member in a social network, and compares the results of the proposed methods with previous approaches to solve the problem. Section VII presents concluding remarks and future work.

II. NODE SIGNIFICANCE BASED IF POLICY

Geodesic distance [12] can be used to determine how quickly a node can disseminate information to all the other nodes in the network. Therefore it is highly used to measure IF in social networks. Geodesic distance based measures use the following assumptions:

a. Each node in the network will disseminate information to all its neighbors at the same time, and
b. Times taken by B to receive information sent by A and vice versa are equal.

But these assumptions do not necessarily hold true in real world social networks. Instead, we assume that a person is likely to share information first with more significant people and then with less significant people. Based on this expected behavior, we propose a new IF policy for which we evaluate the maximum IF delay between two members.

Example: According to the network in Figure 1, node 12 has nodes 2, 4, 6 and 7 as neighbors. Based on degree centrality node 4 has the highest significance among the neighbors. Node 12 gives priority to node 4 and will send information to node 4 in 1 unit time. But, the maximum delay for the node 2 in receiving information sent by node 12 will be 2 time units. Once node 2 receives the information sent by node 12, node 2 can forward the received information to other nodes in its neighborhood, if need be. If node 2 has some other information to be sent to node 12 or node 14 then node 2 will give preference to those two nodes. So the maximum delay from node 2 to node 1 will be 3. If node 12 needs to send some information to node 1, the information should flow through node 2 and the maximum delay for the IF will be the sum of...
maximum delay of IF from node 12 to node 2 and node 2 to node 1.

We define the IF measure as the above described maximum IF delay between any two nodes. To differentiate, the geodesic distance approach assumes that nodes 2, 4, 6 and 7 will receive information sent by node 12 at a delay of 1 unit time whereas under the proposed IF policy, they will receive information in 2, 1, 3 and 3 units of time respectively. We argue that the node significance based IF policy is more reasonable in several real-world applications.

III. IFQ OF A SOCIAL NETWORK

The IFQ of a social network determines the rate at which information can be propagated through the network under various conditions. Sarr and Missaoui defined the IFQ of a network in terms of number of nodes in the network and the eccentricity of the node with the highest closeness centrality [2]. This measure has two deficiencies:

a. IFQ measures based on the geodesic distance do not capture real world information flow patterns.

b. This measure only considers the eccentricity of the node with the highest closeness centrality, which depicts the least time taken by the network to send information to all the nodes in the network. Since information can enter into the network from any node, determining the IFQ based only on the eccentricity of the node with the highest closeness centrality does not capture the whole picture.

The proposed information flow quality measure (IFQM) adheres to the greedy behavior of members as described in section II. In addition, it considers the best case, worst case and the average case and is based on four elements:

1) Size of the biggest connected component in the network ($n_c$) - Higher value of $n_c$ implies benefit of more members, thus the IFQ is high.

2) Maximum information flow delay between any two nodes in the network ($\alpha_h$) - This measures the worst case time of information flow in the network. Lower the $\alpha_h$, the IFQ is better.

3) Eccentricity of the node with the highest closeness centrality ($\alpha_t$) - This determines the maximum time taken to disseminate information to the whole network in the best case. Since the information will be disseminated to all the neighbors at the same time, the geodesic distance will be considered here as the distance measure. Lower the $\alpha_t$, the IFQ is better.

4) Average information flow measure in the network. ($\alpha_a$) - This is the average delay taken by any node in the network to send information to any other node in the network. This is the average case of the information flow in the network. Lower the $\alpha_a$, the IFQ is better.

So, the proposed IFQM ($\lambda$) is:

$$\lambda = < n_c, \alpha_h, \alpha_t, \alpha_a > \quad (1)$$

Table I illustrates the maximum and average IF delays of all nodes of figure 1. Note that $\alpha_h$, $\alpha_t$ and $\alpha_a$ values shown are calculated assuming the node significance based IF policy defined in the previous section.

For the network structure of Figure 1 the IFQM is

$$\lambda = < 14, 8, 3, 3.63 >$$

IV. PLACEMENT BASED SATISFACTION OF A MEMBER

A person’s satisfaction can change according to his placement in the network [13]. Empirical studies on social networks reveal that the satisfaction of a person who is a part of a social network depends on factors such as workload, easy access to higher authorities and responsibility levels [14]. We propose a new measure to determine the PSQ of a member by determining some of the above factors using network centrality measures. For this we assume the following personal characteristics of a person are a priori known: outgoings/ability of influence, ability of mediation and ability of effective communication.

These properties of the new node can be determined using additional sources (such as the interview process for a new candidate who is joining an organization).

a. Degree Centrality - Measures the number of direct connections of a node and can be considered as a measure of outgoings of a person.

b. Betweenness Centrality - This is the number of shortest paths between two nodes that pass through the particular node [12]. A high betweenness centrality suggests that the individual is connecting various different parts of the network and can be considered as a measure of mediation.

c. Closeness Centrality - This indicates how close a node is to all the other nodes in a network. A high closeness centrality means that it has a small average distance to other nodes in the network. This can be considered as a measure of effective communication.

PSQ is measured as:

1) Amount of potential personal qualities utilized -

We assume that a person is unsatisfied if his personal

---

**TABLE I: Maximum and Average information flow measures of nodes in Figure 1**

<table>
<thead>
<tr>
<th>Node</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max delay to any node</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>6</td>
<td>8</td>
<td>6</td>
<td>8</td>
<td>6</td>
<td>8</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Average delay to all nodes</td>
<td>4.1</td>
<td>3.4</td>
<td>3.1</td>
<td>3.3</td>
<td>3.9</td>
<td>3.9</td>
<td>3.5</td>
<td>4</td>
<td>4</td>
<td>3.5</td>
<td>3.9</td>
<td>4.2</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>Overall Average</td>
<td>3.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
qualities are not fully utilized. We also assume that the maximum (potential) personal quality values are known before a person is added to the network. The utilization of the personal qualities can be measured by the following.

a. Percentage of potential outgoingness utilized in the network \( (\phi_d) \) - Measured as the ratio between the actual degree centrality of a node and the maximum potential degree centrality.

b. Percentage of potential mediation ability utilized in the network \( (\phi_b) \) - Measured as the ratio between the actual betweenness centrality of a node and the maximum potential betweenness centrality.

c. Percentage of potential communication ability utilized in the network \( (\phi_c) \) - Measured as the ratio between the actual closeness centrality of a node and the maximum potential closeness centrality.

2) Closeness to important people -
We assume a person will be more satisfied if he is able to communicate with the leader of the network quickly. So PSQ depends on information flow delay between a node and the leader of the network \( (\tau_n) \).

So, the proposed PSQ of a member \( (\mu) \) is:

\[
\mu = \langle \phi_d, \phi_b, \phi_c, \tau_n \rangle
\]

A person is more satisfied if \( \phi_d, \phi_b, \) and \( \phi_c \) are large, and \( \tau_n \) is low; i.e., when natural outgoingness, mediation ability, and communication abilities are more utilized and he communicates with the leader of the network quickly.

V. ADDING A NEW MEMBER TO THE NETWORK

Addition of a new member to social networks should be done in a manner which would be most beneficial to all the parties in the network. So, when a new member is added to the network we would like to optimize \( \lambda \) as well as \( \mu \). Algorithm I describes how this is to be accomplished.

**Algorithm I**: Node Addition

**Input:**
1. (a) Max outgoingness ability of the person - \( O_p \)
2. (b) Max mediation ability of the person - \( M_p \)
3. (c) Max communication ability of the person - \( C_p \)
4. (d) Emphasis on IFQ - \( W_i \)
5. (e) Emphasis on PSQ - \( W_s \)

**Output**: Placement for the new person

6. Calculate the maximum number of edges possible for the new node without exceeding \( O_p \).
7. Determine the set of edge combinations that do not exceed \( M_p \) and \( C_p \) - let the set of combinations be \( C_e \).
8. for all \( c \in C_e \) do
9. Evaluate Network IFQ - \( \lambda_c \)
10. Evaluate PSQ - \( \mu_c \)
11. end for
12. Rank \( \lambda_c \) and \( \mu_c \) values in increasing order.
13. Calculate the combined rank for each combination:
   \[ R_{Combined} = w_i \times R_{Info-flow} + w_s \times R_{Satis} \]
14. Return - The combination that gives the lowest \( R_{Combined} \).

Since different node placements deliver different values of \( \lambda \) and \( \mu \), we need to find the place where the best combination of above measures is achieved within the small community to which the new member is going to be added.

Example: Consider an organization which is modeled as a network shown by Figure 1. A new member\( (15) \) needs to be added to the network with each \( O_p, M_p \) and \( C_p \) not exceeding the top 50% of the existing network.

More than half of the members of the existing network should have higher degree, betweenness and closeness centrality values than the newly added member. So, the maximum number of edges that could be connected to the new member will be 2. Following steps are followed to determine the best place to add the new node.

1) Rank the placements of the new member in the order of the Network IFQ. Assume for some placement \( p \) the rank obtained is \( R_{Info-flow} \).
2) Rank the placements of the new member in the order of the PSQ. Assume for some placement \( p \) the rank obtained is \( R_{Satis} \).
3) Obtain the combined rank \( R_{Combined} \) for each placement \( p \) by using the Eq 3.

\[
R_{Combined} = w_i \times R_{Info-flow} + w_s \times R_{Satis} \quad (3)
\]

where, \( w_i \) and \( w_s \) are the weights given to Network IFQ and PSQ respectively, and \( w_i + w_s = 1 \). If more preference is given to maximize the network information flow quality than the PSQ, \( w_s \) should be kept higher than \( w_i \). The place which gives the lowest combined rank will be selected as the place to add the new member.

Table II and III show the best places to add a new node \( (15) \) considering the Network IFQ and PSQ respectively. The best placement obtained for the new node depends on the values chosen for \( w_i \) and \( w_s \). Fig 2 shows the different placements obtained for the new node 15 for different choices of \( w_i \) and \( w_s \).

If the intention is to maximize the user’s satisfaction without considering the information flow of the network, the
new user can be placed as shown in Figure 2a. If both the objectives are given the same priority, his placement should be done as shown in Figure 2b. If the goal is to maximize the IFQ of the organization without considering the new user’s satisfaction, he could be placed as shown by Fig 2c.

**Computational complexity of adding a node** - Let \( k \) be the number of edges to be added to the new node. Then the maximum number of connection combinations possible for the new node is \( O(\|V^k\|) \). Computing the centrality measures for each combination takes \( O(\|V\|E) \) \{15\}. So, the complexity of the node addition is \( O(\|V\|^k+1\|E\|) \).

VI. REMOVING A MEMBER FROM THE NETWORK

Restructuring of the network after removing members from the network should be done in a manner that minimizes the impact. To re-adjust the network after a removal of a member which disconnects the network, we maximize IFQ while minimizing the change in the qualities of the existing members. Solutions in [1] and [2] are based on identification of the significance of the deleted node. Our solution is different in following ways:

a. We do not require pre-identification of roles of significant and insignificant members of the network.

b. Our solution guarantees that existing members of the network are not required to alter their characteristics drastically.

There can be several cases to consider depending on the placement of the deleted node. For example consider the network in shown in Figure 4 of [2].

Case 1 - **Degree of the deleted node is one**

Since node \( A \) had only one neighbor, neither information flow of the network will not get affected heavily nor will the network gets disconnected by its removal. So, no adjustments are needed in this case.

Case 2 - **Deleted node does not disconnect the network**

If node \( F \) is removed, all of its existing connections will be removed. But all the neighbors of node \( F \) remain connected to the network. So the network will remain connected, thus no adjustments are required in this case.

Case 3 - **Deleted node disconnects some member from the network**

If node \( B \) is removed from the network all the connections from node \( B \) will also be removed, thus disconnecting node \( A \), whose only connection was with node \( B \). Now node \( A \) can be considered as a new member to be added to the network as discussed in Section V.

Case 4 - **Deleted node creates multiple connected components of the network**

Removal of node \( W \) from the network produces the network consisting of three connected components. Algorithm 2 is used to restructure the network.

**Algorithm 2**: Network re-structuring when multiple connected components are available

**Input**: Connected components after the node removal

**Output**: Reconnected network

1. Convert all connected components but the largest connected component to ‘Pseudo-nodes’.
2. Sort the ‘Pseudo-nodes’ by the decreasing order of size - \( S^p \).
3. for all \( P_n \) in \( S^p \) do
   4. Link \( P_n \) to the largest connected component by using the Algorithm 1
   5. For each node in \( P_n \) select the best node to be connected to the largest connected component
5. end for

Our objective is to connect these components in a manner which maximizes the IFQ of the network while maintaining the characteristics of each node within a specific limit. Towards this goal, all the connected components, except for the largest one, will be considered as pseudo-nodes (See Figure 3a). Then, the largest pseudo-node is added to the largest connected component using the method in Section V. The qualities of the pseudo-node will be taken as the average of the qualities of all the members in the pseudo-node. The addition of the largest pseudo-node to the largest connected component, without exceeding characteristics of any node by more than 50% is shown in Figure 3b. Next, each member of the pseudo-node is considered separately and the node which gives the best Network IFQ when connected to the largest connected component is found. The result of this step is shown in Figure 3c. This process is applied for each of the remaining pseudo nodes. The final network is shown Figure 3d.

The quality of the networks obtained by previous network restructuring methods (see [1, 2]) and our approach (with no restriction for characteristic changes) for deletion of node \( W \) from the network shown in Figure 4 in [2] is summarized in...
In this paper, we have proposed a new information flow policy for social networks based on the significance of members in the network. Then two new quality measures for social networks are proposed based on the proposed information flow policy. The proposed quality measures can be modified to accommodate other alternative information flow policies. These proposed measures were then used to determine the best network restructuring during member additions and member removals from social networks. During the network restructuring process for member additions, both the network’s IFQ and the new member’s satisfaction quality are optimized to obtain the best placement for the new member. During the network restructuring process for member removals, the network’s IFQ is optimized while minimizing the changes in the qualities of existing members.

Our future work involves including the content of the information flow policy proposed. We believe the significance of the content of the messages transferred in the network will also be a key factor in determining the information flow in the network. We also intend to model the network restructuring using a game theoretic approach, which will enable us to optimize multiple criteria during the restructuring process.

TABLE IV: Comparison of results for deletion of node W in Figure 4 of [2]

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Network</td>
<td><img src="image1" alt="Image" /></td>
<td><img src="image2" alt="Image" /></td>
<td><img src="image3" alt="Image" /></td>
</tr>
<tr>
<td>Diameter</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Eccentricity of center node</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Ave. distance</td>
<td>2.4</td>
<td>2.4</td>
<td>2.22</td>
</tr>
</tbody>
</table>

In this paper, we have proposed a new information flow policy for social networks based on the significance of members in the network. Then two new quality measures for social networks are proposed based on the proposed information flow policy. The proposed quality measures can be modified to accommodate other alternative information flow policies. These proposed measures were then used to determine the best network restructuring during member additions and member removals from social networks. During the network restructuring process for member additions, both the network’s IFQ and the new member’s satisfaction quality are optimized to obtain the best placement for the new member. During the network restructuring process for member removals, the network’s IFQ is optimized while minimizing the changes in the qualities of existing members.

Our future work involves including the content of the information flow policy proposed. We believe the significance of the content of the messages transferred in the network will also be a key factor in determining the information flow in the network. We also intend to model the network restructuring using a game theoretic approach, which will enable us to optimize multiple criteria during the restructuring process.

**REFERENCES**


