



Social networks and genetic algorithms to choose committees with independent members



Eduardo Zamudio*, Luis S. Berdún, Analía A. Amandi

ISISTAN Research Institute (UNICEN/CONICET), Campus Universitario, Paraje Arroyo Seco, Tandil, Buenos Aires B7001BBO, Argentina

ARTICLE INFO

Keywords:
Committee
Group selection
Independence
Social network
Genetic algorithm

ABSTRACT

Choosing committees with independent members in social networks can be regarded as a group selection problem where independence, as the main selection criterion, can be measured by the social distance between group members. Although there are many solutions for the group selection problem in social networks, such as target set selection or community detection, none of them have proposed an approach to select committee members based on independence as group performance measure. In this work, we propose a novel approach for independent node group selection in social networks. This approach defines an independence group function and a genetic algorithm in order to optimize it. We present a case study where we build a real social network with on-line available data extracted from a Research and Development (R&D) public agency, and then we compare selected groups with existing committees of the same agency. Results show that the proposed approach can generate committees that improve group independence compared with existing committees.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Organizations need representative individuals to make decisions about particular concerns. These representative individuals are appointed in committees, and we expect from his members to make decisions based on the benefit of the whole community they are representing, avoiding bias that could arise from closeness between them. In this context, the best committees are those which show the greatest independence between his members. How to choose these members based on objective criteria could be a difficult task, either because of the definition of the criteria or because of the analysis of the community from where members are chosen. Therefore, a committee in which some of its members are closely related is an unbalanced committee.

Fig. 1 shows a graphical example of difference between balanced and unbalanced committees that allow us to appreciate the distribution of selected nodes within a graph. A balanced distribution is essential to improve desirable features, such as independence. For instance, a committee to discuss about budget allocation must avoid biased decisions by ensuring that committee members are not closely related.

As mentioned before, Fig. 1 shows a simple example of individuals and their dispersal. Fig. 2 shows a graphical representation of the community used to evaluate this approach. This graph allows us to understand the problem complexity and critical importance of choosing the best committee members to maximize independence.

Initially, the committee member selection problem can be solved by a mathematical combination, but the computational cost associated to this approach could be very high. For instance, given a community with n members, the maximal number of groups is given by $2^n - 1$, and complexity is $O(2^n)$. In addition, if committees are r size groups, the number of possible solutions is given by applying binomial coefficient ${}_nC_r$ and complexity is $O(n!)$.

If there is no polynomial function to solve the problem, an alternative could be to adopt a non deterministic approach to approximate optimal solutions. For instance, a stochastic approach could produce random solutions, and then apply an independence function to rank these solutions. This approach is subjective because of the probability in selecting random committee members, and because of the joint probability of the committee.

However, the problem can be addressed by implementing some optimization strategy to approximate optimal solutions, such as genetic algorithms. A genetic algorithm could be implemented to search for the greatest independence between committee members, but not necessarily to guarantee the best solution. In other words, could be enough to approximate an optimal solution. For committee selection problem, the best solutions will be determined by the maximal independence between his members.

* Corresponding author. Tel.: +54 249 4439882/3764206786.

E-mail addresses: eduardo.zamudio@isistan.unicen.edu.ar, eduardozamudio@gmail.com (E. Zamudio), lberdun@exa.unicen.edu.ar (L.S. Berdún), amandi@exa.unicen.edu.ar (A.A. Amandi).

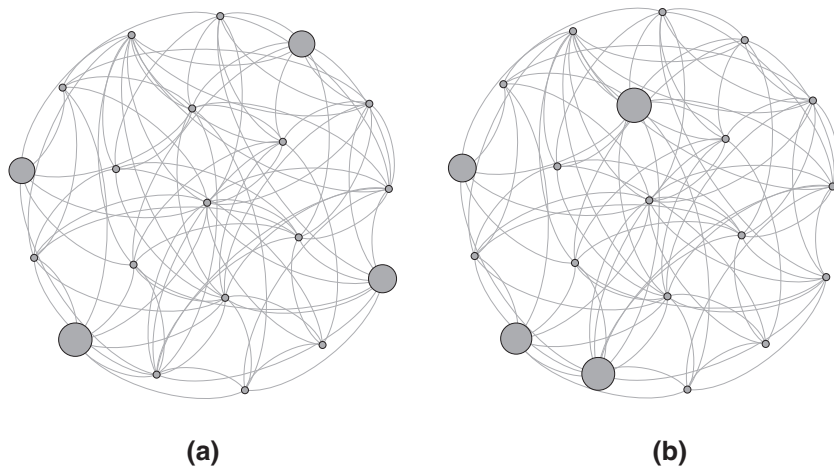


Fig. 1. Difference between balanced (a) and unbalanced (b) committees, where selected members are the largest 4.

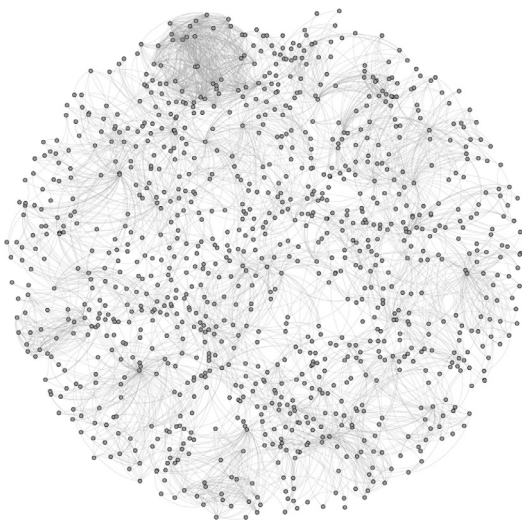


Fig. 2. Graphical representation of the community used to evaluate the approach.

If we consider committee candidates as individuals connected with each other through ties, it is possible to determine which of these ties could be relevant to analyze independence. These individuals and their relationship represent the basic elements of a social network; therefore, we can apply social network analysis to select committee members with the greatest independence. However, a social network approach requires a social network, and data to represent its elements, such as actors, ties, kind of network, and analysis object.

The current social network analysis techniques aim to identify the value or number of relations, roles or prominence of nodes, and to discover hidden groups or cohesive groups. The aim of this work is to present an alternative to the committee selection problem by choosing a set with maximal independence between members. To do this, we build a social network and then we define an independence group performance function and a genetic algorithm, to obtain n member committees with the greatest independence between members.

The main contributions of this work are summarized as follows. (1) We propose an approach for the committee selection problem with independent members as a group selection problem in social networks. (2) We define a novel group independence performance function to assess group fitness in social networks. Then, such a measure was optimized by means of a genetic algorithm. (3) We build a social network from a Research and Development (R&D) public agency with on-line available data. (4) We use such a social network

to evaluate the proposed approach. Then, we compare results with current committees of the same public agency.

This document is organized as follows. Section 2 describes the construction process of the social network. Section 3 describes the implementation of the genetic algorithm and the function to evaluate group independence. Section 4 describes a case study and the configurations of the genetic algorithm, along with a discussion of the experiment results. Section 5 presents a discussion of the current literature. Finally, Section 6 presents conclusions as well as future work.

2. Social network construction

In order to choose committee members, we propose to build a social network to calculate distances between candidates, and then apply a genetic algorithm to get potential committees with the greatest distances between their members.

A social network is a set of individuals (actors) and relations (ties) between them; the social network analysis is used to study structures created by these relations and individuals.

We are particularly interested in the construction of a social network for its ability to represent analysis criteria based on ties. To clarify this concept, we built a network of researchers related through co-authorship and workplace. In this network, actors are the researchers, and ties are the criteria for calculating distance between each pair of researchers.

As mentioned above, relations between actors define what can be analyzed in the network. The aim of this analysis is to calculate distances between a set of actors. In order to do this, we built a consolidated graph. This graph contains every kind of relation proposed as analysis criterion. Fig. 3 shows a unified graph from two kinds of relations (*coauthor* and *same workplace*) of five researchers (A, B, C, D, and E) where relations are binary (relation is present or not), undirected (direction is meaningless), and irreflexive (a researcher does not publish with himself or does not work with himself).

Our proposal is to establish the greatest independence between committee members based on their distances. Thus, we need to calculate distances between committee members, for which we use the *shortest path* and *geodesic distance* (length of the shortest path) (Freeman, 1977) over the unified graph.

The graph must be connected to apply this metrics, which means that every actor must be reachable from every other actor in the network. This can be determined through a reachability matrix, which can be obtained through matrix multiplication (Wasserman & Faust, 1994).

Distances between each pair of actors is represented by a proximity matrix, obtained by applying power to the matrix

Download English Version:

<https://daneshyari.com/en/article/382654>

Download Persian Version:

<https://daneshyari.com/article/382654>

[Daneshyari.com](https://daneshyari.com)