Physica A 392 (2013) 1547–1556

Contents lists available at SciVerse ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Skeleton of weighted social network

X. Zhang [a,*], J. Zhu [b]

[a] *School of Economics and Management, Beijing University of Posts and Telecommunications, Beijing 100876, China*
[b] *Department of Statistics, University of Michigan, Ann Arbor 48109, USA*

ARTICLE INFO

Article history:
Received 9 May 2012
Received in revised form 15 October 2012
Available online 7 December 2012

Keywords:
Social network
Skeleton
Percolation
Fractality
Information diffusion

ABSTRACT

In the literature of social networks, understanding topological structure is an important scientific issue. In this paper, we construct a network from mobile phone call records and use the cumulative number of calls as a measure of the weight of a social tie. We extract skeletons from the weighted social network on the basis of the weights of ties, and we study their properties. We find that strong ties can support the skeleton in the network by studying the percolation characters. We explore the centrality of w-skeletons based on the correlation between some centrality measures and the skeleton index w of a vertex, and we find that the average centrality of a w-skeleton increases as w increases. We also study the cumulative degree distribution of the successive w-skeletons and find that as w increases, the w-skeleton tends to become more self-similar. Furthermore, fractal characteristics appear in higher w-skeletons. We also explore the global information diffusion efficiency of w-skeletons using simulations, from which we can see that the ties in the high w-skeletons play important roles in information diffusion. Identifying such a simple structure of a w-skeleton is a step forward toward understanding and representing the topological structure of weighted social networks.

1. Introduction

Social networks have been the focus of scholars in various fields, such as psychology, economics, and physics. In this framework, social life consists of the flow and exchange of norms, values, ideas, and other social and cultural resources [1]; furthermore, the social actions of individuals can be affected by the structure of the underlying network [2]. Research on the structure of social networks is important from the perspective of not only individuals, but also society as a whole [3]. Many emerging concepts, such as small-world property [4], scale-free behavior [5], community structure [6], and fractality [7], form the basis of our understanding of complex network systems. Moreover, some scholars provide several conceptual models to represent the topology of networks in a way that can be understood easily and intuitively. Although most current research on topology attempts to maintain and describe the information in all its detail, a simple conceptual model is also important, especially when it captures graphically many fundamental properties [8]. For example, Carmi et al. [9] introduce and use the method of k-shell decomposition and the methods of percolation theory and fractal geometry to study a map of the Internet; in doing so, they find a model for the structure of the Internet. Their method of decomposition is robust and provides insight into the underlying structure of the Internet and its functional consequences. Siganos et al. [8] propose the Jellyfish as a model for the inter-domain Internet topology; this model captures and represents the most significant topological properties, and it can be easily drawn by hand. Goh et al. [10] propose a fractal skeleton, a special type of spanning tree based on the betweenness centrality of ties. The fractal skeleton has the property of being a critical branching tree. The

---

[*] Corresponding author.
E-mail addresses: zhangxiaohang@bupt.edu.cn (X. Zhang), jizhu@umich.edu (J. Zhu).

0378-4371/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2012.12.001

original fractal networks are viewed as a fractal skeleton dressed with local shortcuts. All these methods shed light on our understanding of the underlying structure of networks. However, they mainly focus on unweighted networks. In reality, many networks are weighted, a property that is associated with information diffusion and social actions. Some researchers provide several methods to analyze cluster or hierarchical structure of weighted networks. For example, Lambiotte and Ausloos [11] reveal the emergence of social communities and music genres by analyzing a weighted listener network and a weighted music network where the weights are computed by a symmetric correlation. Song et al. [12] introduce a graph-theoretic approach to extract clusters and hierarchies in complex data-sets. This is achieved by building topologically embedded weighted networks containing the subset of most significant links and analyzing the network structure. Serrano et al. [13] define a filtering method that can extract the relevant connection backbone in complex multi-scale networks, preserving the edges that represent statistically significant deviations with respect to a null model for the local assignment of weights to edges.

In this paper, we present a detailed analysis of the skeleton of a weighted social network constructed from mobile phone call records. The skeleton of the social network is extracted on the basis of the ties with the highest weights. The definition of the skeleton is similar but not identical to that of the asset graph proposed by Onnela et al. [14] (see Section 3.1). We describe the data and several basic characteristics of the original network in Section 2. The definition of the skeleton network and its characteristics, including susceptibility, centrality, self-similarity and fractality, are presented in Section 3. We also explore the roles of the skeleton network in information diffusion and compare the skeleton network with other conceptual networks, including the *k*-shell and Jellyfish models, in Section 3. Finally, we discuss our findings in Section 4.

2. Data and basic network characteristics

In this study, we use one dataset, consisting of mobile phone call records in one city over three months, to construct a social network. This social network covers approximately 72.7% of the city's entire population, which to some extent assures the representativeness of the social network from the view of coverage. For the purpose of retaining customer anonymity, each subscription was identified by a surrogate key, guaranteeing that the privacy of customers was respected. As in the work of Onnela et al. [3], we kept only voice calls, filtering out all other services, such as voice mail, data calls, text message, chat and operator calls. We filtered out calls involving other operators, incoming and outgoing, keeping only those transactions in which the calling and receiving subscription is governed by the same operator. This filtering was needed to eliminate the bias between this operator and other operators as we have full access to the call records of this operator, but only partial access to the calls made to subscriptions governed by other operators. Moreover, we also filter out the calls between cities, keeping only those within the same city, since long-distance calls occupy a low proportion and we attempt to assure relative closeness of the social network. Two users are connected with an undirected link if there was at least one reciprocated pair of phone calls between them (i.e., A called B, and B called A). We quantify the weight of the undirected link (i, j) by the total number of calls made between i and j over the studied period. This procedure eliminates a large number of one-way calls, most of which correspond to single events. The resulting social network contains 155,406 vertices and 1,074,578 links. The mobile call graph naturally captures only a subset of the underlying social network, which consists of all forms of social interactions, including face-to-face interactions, email and Internet communication. However, research on media multiplexity suggests that the use of one medium for communication between two people implies communications via other means as well [3,15].

Several basic network characteristics of the social network are shown in Fig. 1. We study the cumulative degree distribution $P_>(k)$, defined as $P_>(k) = \int_k^{+\infty} p(x)dx$, where $p(x)$ denotes degree probability density function. The degree distribution is characterized by wide variability and a heavy tail, representing a heterogeneous topology. As mentioned earlier, the weight of social interaction (tie) is measured by the aggregate number of calls made. The same applies to strength of a vertex, defined as $s_i = \sum_{j\in\mathcal{N}(i)} w_{ij}$, where $\mathcal{N}(i)$ denotes the neighborhood of vertex i and w_{ij} is the weight of tie (i, j). The strength distribution is broad so that while the majority of vertices made a couple of calls, a small fraction of users placed numerous calls with each other. The extend of coherence around a vertex i is quantified by clustering coefficient $C_i = 2t_i/[k_i(k_i - 1)]$, where t_i denotes the number of triangles around vertex i and k_i is the degree of vertex i [4]. The network has fairly high average clustering coefficients, which can be seen as a manifestation of the presence of triangular relationships.

3. The skeleton and its characteristics

3.1. Definition of the skeleton

Let us consider a graph $G = (V, E)$ of $|V|$ vertices and $|E|$ ties. Then, the definition of a w-skeleton is as follows.

A subgraph $G' = (V', E')$ induced by the set $V' \subseteq V, E' \subseteq E$ is a w-skeleton, or a skeleton of order w, if and only if the weight w_{ij} of any tie $(i, j) \in E'$ induced in G' is equal to or greater than w, all vertices in V' are connected through at least one path, and G' is the maximum subgraph with this property. A w-skeleton of G is the largest connected component (LCC) in the w-subgraph that is obtained from G by removing all the ties with weights less than w and the isolated vertices (i.e., vertices with which there are no ties). The sketch of the 150-skeleton of the social network is shown in Fig. 2.