Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Brazilian Campaign Financial Network in 2014 Elections: Topological Properties

André Manhães Machado[1]
Universidade Federal do Espírito Santo, Av. Fernando Ferrari, 514, 29075-910 - Vitória/ES
Maria Claudia Silva Boeres[2]
Universidade Federal do Espírito Santo, Av. Fernando Ferrari, 514, 29075-910 - Vitória/ES

Abstract. General elections were held in Brazil on 5 October 2014 to elect the President of Republic, the Senators, the State Governors and State Legislatures, amounting to 1603 positions to be filled in by more than 18.000 running candidates. In order to meet this competition, candidates receive private financial contribution to make campaign among their electorate. Based on this, electoral justice requires candidates to provide a detailed accounting of revenue which is available to public just after elections is over. Using these reports, we model a new social network, until now unexplored, which represents the financial relation between the participants involved in Brazilian campaign of 2014 elections. Analyzing the structure of the network, we show that the major component, which represents more than 96% of nodes and edges within the original network, has a small-world behavior and it can be classified as a scale-free network, two well-studied network models found in different areas of social networks.

Keywords. Scale-free, Small-world, Network Analysis, Brazilian Elections

1 Introduction

Social Network Analysis (SNA) consists of a set of techniques used for investigating social structures with the purpose of studying the exchange of resources or the interactions among actors within a complex network. Complex networks are networks whose structure is irregular, complex and dynamically evolving in time. A complex network can be modeled as a graph with non-trivial topological features where a set of nodes, representing individuals or organizations, are joined together by links indicating interaction [1]. Usually, these networks have thousands or millions of nodes.

We point out three non-trivial topological features (graph properties) commonly used to categorize the topological structure of a complex network: (a) the average shortest path length L, computed as the average over all shortest paths lengths in the network; (b) the clustering coefficient C, which indicates the trend of nodes to form groups in the graph and (c) the degree distribution, which gives the probability of having a node with a positive integer number of k links [1].

[1]andre.manhaes@gmail.com
[2]boeres@inf.ufes.br

Complex networks typically have highly heterogeneous degree distributions with long tails. This topological feature motivates a major concern in network analysis focused in identifying the most central or relevant nodes in the system [2].

Many real systems are modeled as complex networks. Social networks, the Internet, food webs, metabolic and protein networks, scientific collaboration networks and epidemiological networks are some examples widely explored in the literature [3].

Last Brazilian general elections provided around 1600 positions to be filled in by more than 18.000 running candidates. In order to meet this competition, these candidates receive private financial contribution to make campaign among their electorate. All this data is publicly available in the web. As far as we know, a social network modeling the financing relations in the Brazilian electoral process is not yet investigated in the scientific literature.

In this paper, we propose to model the Brazilian Campaign Financing Network (henceforth designated as BCFN) of 2014 elections as a complex network. Using the information provided by the Superior Electoral Court (TSE), the BCFN graph is defined to represent the existence of monetary relations (graph edges) between the participants (graph nodes) involved in Brazilian electoral campaign financing. The BCFN graph is disconnected, but the major connected component represents more than 96% of the whole network.

This paper is organized as follows. The next section, describes the Brazilian electoral process main aspects and proposes a model for BCFN network. Section 3 studies the topological properties of the network. The results and conclusions are reported in the Section 4.

2 Dataset and Graph Definition

The 2014 Brazilian elections simultaneously appointed the president of the Republic (1 seat), one-third of senators (27 out of 81 seats), all members of the chamber of deputies (513 seats), all state governors (27 seats) and legislators (1035 seats), summing up to 1603 seats.

In the elections process, candidates for legislative office were required to submit a report with the registry of campaign contributions to their State's Electoral Court (TRE in Portuguese), which were subsequently sent to the Superior Electoral Court (TSE in Portuguese). These reports, available in [4], provide detailed information on every declared financial contribution to candidates campaigns, including the name of the contributor, the recipient, the contributions type, as well as the total amount and the date of the contribution. Parties and candidates may create committees and directories with the goal to rise funds to one or more candidates, in order to organize the providing accounting as defined in law.

The political contributions to candidates come from several sources: a) individuals (which are not running for public offices); b) corporations; c) other political candidates; d) committees, parties and directories (henceforth designed as PCDs). Moreover, political candidates can make donations to committees and directories, which in turn can donate to candidates and between them. Corporations can also make donations to committees and directories. In other words, all participants in the electoral process can be recipient/contributor, except individuals and business persons.

Using the reports available in [4], we propose in this paper the $BCFN$ network modeled as a graph $G = (V, E)$, where the set of nodes V consists of any contributor or recipient who appears receiving or giving any type of monetary good in 2014 Brazilian elections. This includes all candidates, individuals, corporations, parties, committees and directories; and the set of edges E, where an edge $\{v, w\} \in E$ if $v \in V$ is receiving or donating any kind of contribution from/to $w \in V$.

3 Topological Properties

The research on complex networks is concerned with the effort of defining new concepts and measures to characterize the topology of real networks. In last decades, the major outcome has been the identification of principles and statistical properties common to most of the real networks studied. Initially, the most basic property considered, for further analysis of topological structure, is the distribution of connected components in the graph [5].

A graph connected component is defined as a maximal connected subgraph. A subgraph is connected when all of its nodes are linked to one other through paths: all vertices in a connected subgraph can reach one other through at least on path, but they have no connection outside the component. A subgraph is maximal if is impossible to add any new node or edge without invalidating the property of connectedness [6].

The number of connected components of a graph, including their size and quantity, is taken as an indication of the opportunities and obstacles to communication or the transfer of resources within the network. Thus, the analysis of components is a mandatory step of structural description of a network [6].

The graph topology is also an important way to understand the efficiency, robustness or reliability of a network. In order to comprehend the structure or relation between individuals within the network, the commonly studied topological features are the clustering coefficient, the degree distribution of nodes and the average shortest path length [7], concepts that are briefly explained in the following.

First introduced by Watts and Strogatz [8], the clustering coefficient C measures the probability that two connected nodes $v_1 \in V$ and $v_2 \in V$ are connected to a common third node $w \in V$ in the graph $G = (V, E)$. Let h_v be the number of neighbors of v, t_v be the number of triangles observed for v and $h_v(h_v - 1)/2$ be the total number of possible triangles for v. Hence, the clustering coefficient $C(v)$ for node $v \in V$ is defined as ratio between the actual number of triangles of v and the maximum possible number of triangles for v:

$$C(v) = \frac{2t_v}{h_v(h_v - 1)} \tag{1}$$

The clustering coefficient C of the whole network $G = (V, E)$ is defined as the mean

DOI: 10.5540/03.2017.005.01.0398

of all clustering coefficient $C(v)$, $\forall v \in V$:

$$C = \frac{1}{|V|} \sum_{v \in V} C(v) \tag{2}$$

where $|V|$ is the number of nodes in the network.

The characteristic path length L or the average shortest path length L of a graph $G = (V, E)$ is defined as the path length between two nodes, averaged over all pairs of nodes [9]. L is a metric of how well connected a graph is and one of the most important parameters to measure the overall routing efficiency and functional integration of complex networks [9]. L is defined as:

$$L = \frac{1}{|V|(|V|-1)} \sum_{v,w \in V} d(v, w) \tag{3}$$

where $d(v, w)$ is the distance between v and w, i.e., the length of the shortest path between $v \in V$ and $w \in V$.

Real-life networks usually have a power-law degree distribution, that is, the node degree distribution follows the function $P(k) \propto k^{-\sigma}$, where k is the degree and σ is the power-exponent. Networks with power-law degree distribution are designated as scale-free networks and many networks in nature, ecology, economy and technology have been found to be scale-free [10].

These three properties are used to classify the network topology into three categories: scale-free networks, small-world and random networks [11]. Random networks have small shortest paths L and low clustering coefficient C. Small-worlds are networks with a large clustering coefficient C and small average distance L. Scale-free networks appear in the context of dynamic networks in which new nodes connect preferentially to highly connected nodes in the network, producing a degree distribution which follows as power-law function [12].

4 Results and Discussion

Table 1 shows the properties of Brazilian Campaign Finance Network ($BCFN$) of 2014 elections. The column Properties shows topological properties and distribution of participants in each network. The column $BCFN_M$ refers to the major component of $BCFN$ graph. The column $BCFN_S$ designates the set of components of $BCFN$, excluding the major component $BCFN_M$. The columns R_m and R_s indicate, respectively, the ratio of each property of $BCFN_M$ and $BCFN_S$ relative to $BCFN$ graph. The $BCFN$ graph consists of 1232 different connected components. The major component ($BCFN_M$) represents more than 96% of the set of nodes V and more than 97% of the set of edges E in the original graph. The remaining nodes and edges are distributed among the 1231 smaller components (grouped as the set $BCFN_S$). Since the connected component $BCFN_M$ has most of nodes and edges of $BCFN$ and each small component has around 5.9 nodes on average, for further analysis, we characterize only the topological properties of $BCFN_M$.

Table 1: A comparison between the major component $BCFN_M$ and the graph of small components $BCFN_S$.

Property	$BCFN_M$	$BCFN_S$	R_m	R_s
Components	1	1231	0.08%	99.92%
$\|V\|$	180601	7264	96.1%	3.9%
$\|E\|$	241509	6889	97.2%	2.8%
Candidates	17609	1538	92.0%	8.0%
Corporations	16326	397	97.6%	2.4%
Individuals	146037	5297	96.5%	3.5%
PCD	628	32	95.1%	4.9%

For the aim of classifying $BCFN_M$ network, the metrics L and C of $BCFN_M$, denoted from now on by L_M and C_M, can be compared to a random graph with the same number of nodes $|V|$ and the same number of edges $|E|$. Let $G_{rand} = (V_{rand}, E_{rand})$ be a random network with average degree μ and $|V_{rand}| \gg \mu$. The average shortest-path length L_{rand} is roughly estimated as [13]:

$$L_{rand} \approx \frac{ln(|V_{rand}|)}{ln(\mu)} \tag{4}$$

whereas for $|V_{rand}| \gg \mu$ the clustering coefficient for random network approaches the value of [13]:

$$C_{rand} \approx \frac{\mu}{|V_{rand}|}. \tag{5}$$

where average degree μ is defined as:

$$\mu = \frac{2|E|}{|V_{rand}|} \tag{6}$$

Using this method, the Table 2 shows a comparison between the major component $BCFN_M$ and a random network G_{rand} with the same size. The column Property refers to the topological properties, the column $BCFN_M$ designates the major component of $BCFN$ and $G_{rand}(V_{rand}, E_{rand})$ refers to random graph. When compared to a random network, the major component $BCFN_M$ shows a smaller average path length L_M and a higher clustering coefficient C_M. In fact, the clustering coefficient C_M of $BCFN_M$ is much greater than that clustering coefficient C_{rand} of a random network, $C_M \gg C_{rand}$. As the average path length is low and and the clustering coefficient is high compared to a random network, we can conclude that the $BCFN_M$ has a structure of small-world. Given that the $BCFN_M$ network has candidates from distinct political spectrum (far left to far right parties), the small-world behavior indicates the candidates within $BCFN_M$ are close in terms of campaign finance (average of 5.66 hops between any of the participants), which is unusual when we consider the differences between the proposed political projects of candidates. A possible explanation for that behavior in small-worlds is the presence of hubs within the network, which are nodes with a great number of others nodes connected to them. The presence of hubs can be inferred through the analyses of degree distribution.

Table 2: Comparison between the component $BCFN_M$ and a random network $G_{rand}(V, E)$.

Property	$BCFN_M$	$G_{rand}(V_{rand}, E_{rand})$
$\|V\|$	180601	180601
$\|E\|$	241509	241509
μ	2.67	2.67
C	1.75×10^{-2}	1.48×10^{-5}
L	5.66	12.30

Figure 1 shows the degree distribution of component $BCFN_M$. The red line corresponds to the fitted curve $P(k) \approx 4.7 \times 10^4 k^{-2.117}$ with $R^2 = 0.9608$. We can observe that the power-exponent σ is 2.117 which lies within the interval 2 and 3, typically found in real-life networks [14]. In addition, a long tail showed in the plot, confirms a characteristic of degree distribution. This indicates the presence of hubs (nodes) with degrees much higher than most nodes within the network, which is also a common property of small-world networks. Moreover, the network has a great number of terminal nodes (with unitary degree). These two facts indicate that the $BCFN_M$ network has a great number of participants (terminal nodes) taking little part or marginal involvement in political campaign finance, while others, in far fewer number, are taking huge engagement in the electoral process (hub nodes).

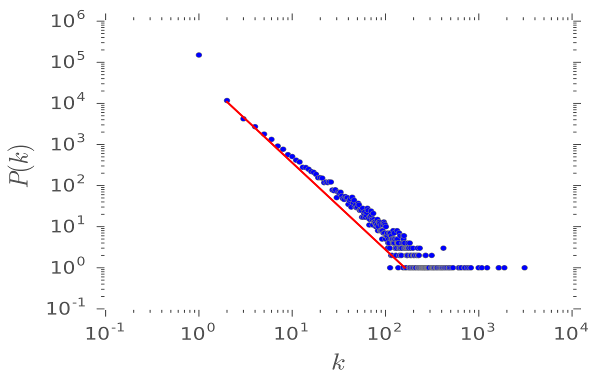

Figure 1: Degree distribution $P(k)$ of component $BCFN_M$.

We conclude that network modeled after the participants taking part in financial campaign in Brazilian elections of 2014 is a graph with a big component, which contains more than 96% of participants within this network. Moreover, we show that the major component has a behavior of small-world and scale-free network, indicating the presence of

DOI: 10.5540/03.2017.005.01.0398

a huge number of participants connected to a few players and a great proximity between the different participants in the elections.

References

[1] K. Klemm and V. M. Eguíluz. Growing scale-free networks with small-world behavior. *Physical Review E*, 65(5), may 2002.

[2] L. C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41, 1977.

[3] S. Wasserman and K. Faust. *Social network analysis: Methods and applications*, volume 8. Cambridge university press, 1994.

[4] Superior Electoral Court. Estatísticas eleitorais - eleições 2014. Available in http://www.tse.jus.br/eleicoes/estatisticas/estatisticas-eleitorais-2014, 2014. Accessed: 2015-02-07.

[5] B. Kang, K.I. Goh, D.S. Lee, and D. Kim. Complex networks: structure and dynamics. *Sae Mulli*, 48(2):115–141, 2004.

[6] John P. Scott and Peter J. Carrington. *The SAGE Handbook of Social Network Analysis*. Sage Publications Ltd., 2011.

[7] R. Pastor-Satorras, A. Vázquez, and A. Vespignani. Dynamical and correlation properties of the internet. *Physical review letters*, 87(25), nov 2001.

[8] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world'networks. *nature*, 393(6684):440–442, 1998.

[9] M. Rubinov and O. Sporns. Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, 52(3):1059–1069, Sep 2010.

[10] E. Ravasz and A. L. Barabási. Hierarchical organization in complex networks. *Physical Review E*, 67(2), feb 2003.

[11] A. E. Motter, A. P. S. de Moura, Y. C. Lai, and P. Dasgupta. Topology of the conceptual network of language. *Physical Review E*, 65(6), jun 2002.

[12] A. Barabási. Emergence of scaling in random networks. *Science*, 286(5439):509–512, oct 1999.

[13] A. P. S. de Moura, Y. C. Lai, and A. E. Motter. Signatures of small-world and scale-free properties in large computer programs. *Physical Review E*, 68(1), jul 2003.

[14] T. Zhou, G. Yan, and B.H. Wang. Maximal planar networks with large clustering coefficient and power-law degree distribution. *Physical Review E*, 71(4), Apr 2005.

DOI: 10.5540/03.2017.005.01.0398