UNIVERSITY OF SÃO PAULO
SÃO CARLOS PHYSICS INSTITUTE

RENATO FABBRI

Complex networks for the participant

São Carlos

2015

RENATO FABBRI

Complex networks for the participant

Monograph presented to the Physics Graduate Program of the São Carlos Institute of Physics, University of São Paulo, for the qualification exam, as part of the requirements for obtaining the title of Doctor in Sciences.

Concentration area: Applied Physics
Option: Computational Physics
Advisor: Prof. Dr. Osvaldo Novais de Oliveira Jr.

São Carlos

2015

ABSTRACT

FABBRI, C. *Complex networks for the participant*. São Carlos Institute of Physics. University of São Paulo, São Carlos, 2015.

Complex networks form one of the most active fields of recent physics. With respectable efforts for exhibiting advances to the general audience, it seems, however, that few or none of these are targeted to the benefit of the individual that constitute such systems. That is, with a core knowledge about the field, and recipes for harnessing, provide means for the participant to interact and understand the networks he/she is in. This work aims to accomplish such task by means of the social networks of the participants. We verified that such networks exhibit time stability of topological measures and of basic connective sector sizes and exhibit differentiation of the textual production in each basic connective sector. We also formalized conceptualizations of these networks as OWL were they were possible, specially in relation to the social participation instances provided by law. Finally, software and data have been put available and used, as means to enable integrated analysis of different provenance and public benefit. Conceptual consequences have been documented and requires anthropological considerations. Furthermore, software, ontological and data contributions can be better documented and developed while a typological consideration of the physical properties observed in human interaction networks should bridge complex networks and the more traditional legacy of human sciences on the subject.

Keywords: Complex networks. Social networks. Complexity. Anthropological physics. Linked data. Semantic web. Social participation. Text mining. Natural language processing.

Contents

1
Introduction

Studies on human interaction networks have started long before modern computers, dating back to the nineteenth century, while the foundation of social network analysis is generally attributed to the psychiatrist Jacob Moreno in mid twentieth century (**?**). With the increasing availability of data related to human interactions, research on these networks has grown continuously. Contributions can now be found in a variety of fields, from social sciences and humanities (**?**) to computer science (**?**) and physics (**?**,**?**), given the multidisciplinary nature of the topic. One of the approaches from an exact science perspective is to represent interaction networks as complex networks (**?**, **?**), with which several features of human interaction have been revealed. For example, the topology of human interaction networks exhibits a scale-free trace, which points to the existence of a small number of highly connected hubs and a large number of poorly connected nodes.

There is gap of knowledge and technology between the complex networks legacy and the usufruct of the participant. This hiatus is reactive, and events suggest that it will maintain itself as an ecosystem of knowledge, technology and undertake of society in all scales. It should ease for example: entrepreneur goals, elaboration and preparation of documents, rapid achievement of knowledge. In general: processes that refer to collection and diffusion of goods and information.

This work presets a confirmation of such scenario and advances. Some strategies were selected to verify the complex networks to the advantage of the participant. Specially, very simple experiments seems able to modify social structures. In this context, we verified temporal stabilities in human interaction networks, and exposed that primitive sectors of the networks (hubs, intermediary and periphery) produce texts that are very different. This is useful for a not stigmatizing typology of participants in interaction networks via quantitative criteria. This yielded audiovisualization and interconnection of data as art and technological gadgets, as support for scientific research. Applications were complemented with the (Brazilian) Presidency of the Republic and UNDP.

Next section presents directions on related work. Section **??** is dedicated to the data analyzed. Section **??** holds the methods used to reach results. Results are given in Section **??**. The chronogram and outline of finished, ongoing and planed tasks are in Section **??**. Conclusions are in Section **??** followed by acknowledgments and the bibliography.

1.1 Literature review

The field of complex networks is relatively new (≈ 25 years) and literature presents diverging definitions. One definition that is having increasing acceptance considers a complex network a "large graph with non-trivial topological features". This definition is misleading in at least two points: there are networks of interest with trivial topological features, such as the paradigmatic Erdös-Rényi network and the lattice network (**?**). Second, it fails to deliver the fundamental message that the complex network is not an isolated mathematical graph structure. Complex networks of interest are real networks or idealized models for understanding them. Besides that, not only large graphs are of interest, but small graphs are very often used as toy examples and measured as extension of larger structures. A definition, still far from perfect, but preferred in this work, is: "often large and non-trivial graphs considered in, or for the consideration of, the environment they reside". This definition resolves both issues.

Books in general present a common and powerful repertory for characterization of complex systems through graphs. Maybe most importantly are the arsenal of measures: degree, strength, betweenness centrality, clustering coefficient, etc.; the basic network paradigms: Erdös Rényi, geographical, small-world, scale-free; and the transdisciplinary approach to the environments that yield the networks. The literature on social network analysis,foe example, can often be understood as dealing with complex networks in human social systems.

A careful consideration of the books and articles read for this research is given in Section **??**.

2
Materials

2.1 The Gmane public database of email lists (benchmarks)

2.2 Facebook, Twitter, Participa.br, Cidade Democrática, AA

2.3 My own social networks

Considerations about the right to annotate.

3
Methods

3.1 Circular statistics

3.2 Erdös Sectioning

3.3 PCA of measures along time

3.4 Kolmogorov-Smirnoff test for texts produced by sectors

3.5 Audiovisualization of data

3.6 Typological considerations

3.7 Semantic web

3.7.1 OWL ontology construction

4
Results

4.1 Time stability in human interaction networks

4.2 Semantic web

4.2.1 Linked data

4.2.2 RDF data conversion of data into linked data

4.2.3 Published linked data and OWL ontologies

4.3 Harnessing

4.3.1 Social percolation procedures

4.3.2 Recommendation systems for the enrichment of semantic navigation

4.3.3 Understanding the social being

5
Finished and planed tasks, chronogram

5.1 Documents

5.1.1 To be finished

Anthropological physics

The study of human systems raises conceptual and ethical issues that require anthropological considerations. There are two immediate routes to this concepts:

- What data should or can be used?
- Can one experiment in a network of humans? In which context?

The short answer is that ethics committees and procedures are dedicated to dealing with those issues. Even so, there is a key-concept from the anthropological legacy: the study of the self as exposed to the interested culture or context. In this sense, it is reasonable (if not a suggestion) that a researcher do reflexive consideration, i.e. that he/she observe and make assumptions about its own sampling of the world. Within this same framework, many social networks (email, Facebook, Twitter, Participa.br, AA) were openly mined, with feedback to and from the studied communities. The term "anthropological physics" started being used in Brazil around 2014 and can be thought as a subfield of Social Physics.

Gradus

Fazer o $2^(1002)$

Consider a idealized constitution of these networks:

- the resources of the environment are the persons, each with an amount of time available.
- The amount of resource employed by the environment to the network is constant through all connective sectors

5.1.2 Finished

5.2 Chronogram

Ano	Semestre	Atividade
2012	II	Revisão e estudo da bibliografia.
2013	I	Implementação computacional.
	II	Cursar disciplinas.
2014	I	Implementação computacional: refinamento do código e corpus.
	II	Apresentação de resultados.
2015	I	Exame de qualificação.
		Escrita de artigo.
	II	Cursar disciplina.
		Monitoria PAE.
2016	I	Defesa do doutorado.
		Escrita de artigo.

Table 5.1 – Cronograma de atividades

6
Conclusions