

UNIVERSIDADE DE SÃO PAULO  
INSTITUTO DE FÍSICA DE SÃO CARLOS

RENATO FABBRI

Redes complexas para o participante

São Carlos

2015



RENATO FABBRI

## Complex networks for the participant

Monografia apresentada ao Programa de Pós-Graduação em Física do Instituto de Física de São Carlos da Universidade de São Paulo, para o Exame de Qualificação como parte dos requisitos para obtenção do título de Doutor em Ciências.

Área de concentração: Física Básica  
Orientador: Prof. Dr. Osvaldo Novais de Oliveira Jr.

São Carlos

2015



# RESUMO

FABBRI, C. *Redes complexas para o participante*. Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2015.

As redes complexas formam uma das áreas mais ativas da física redente. Há esforços consideráveis para apresentar estes avanços ao público geral, mas tudo indica que poucos ou nenhum são voltados para o benefício do indivíduo que constitui estes sistemas. Ou seja, com um núcleo de conhecimento da área, e receitas para aproveitamento, fornece meios para o participante interagir e entender as redes nas quais ele se encontra. Este trabalho objetiva realizar tal tarefa por meio das redes sociais do participante. Verificamos que tais redes exibem uma estabilidade temporal de medidas temporais e dos tamanhos relativos dos setores conectivos básicos (hubs, intermediários, periféricos). Observamos uma diferenciação da produção de texto de cada setor básico. Também formalizamos as conceitualizações destas redes como OWL onde foi possível, principalmente as instâncias de participação dosial previstas por lei e implementadas como software. Software e dados foram disponibilizados e usados. Protocolos escolhidos para facilitar a integração de estruturas de diferentes procedências, para reutilização dos dados em outros trabalhos e pesquisas, e para o benefício público. Consequências conceituais requerem considerações antropológicas e estão sendo documentadas. Próximos passos são: melhor documentação e desenvolvimento do aparato em software, ontologias e dados; considerações tipológicas das propriedades físicas observadas nas redes de interação humana, com atenção aos outliers, às relações entre topologia do agente e texto produzido, e à ponte com a bagagem mais tradicional das ciências humanas no assunto.

**Palavras-chave:** Redes complexas. Redes sociais. Complexidade. Física antropológica. Dados ligados. Web semântica. Participação social. Mineração de texto. Processamento de linguagem natural.



# Sumário

---

<b>1</b>	<b>Introdução</b>	<b>7</b>
1.1	Literature review . . . . .	8
<b>2</b>	<b>Materials</b>	<b>9</b>
2.1	The Gmane public database of email lists (benchmarks) . . . . .	9
2.2	Facebook, Twitter, Participa.br, Cidade Democrática, AA . . . . .	9
2.3	My own social networks? . . . . .	9
<b>3</b>	<b>Methods</b>	<b>11</b>
3.1	Circular statistics . . . . .	12
3.2	Erdős Sectioning . . . . .	12
3.3	PCA of measures along time . . . . .	12
3.4	Kolmogorov-Smirnoff test for texts produced by sectors . . . . .	12
3.5	Audiovisualization of data . . . . .	12
3.6	Typological considerations . . . . .	12
3.7	Semantic web . . . . .	12
3.7.1	OWL ontology construction . . . . .	12
<b>4</b>	<b>Results</b>	<b>13</b>
4.1	Time stability in human interaction networks . . . . .	14
4.2	Semantic web . . . . .	14
4.2.1	Linked data . . . . .	14
4.2.2	RDF data conversion of data into linked data . . . . .	14
4.2.3	Published linked data and OWL ontologies . . . . .	14
4.3	Harnessing . . . . .	14
4.3.1	Social percolation procedures . . . . .	14
4.3.2	Recommendation systems for the enrichment of semantic navigation . .	14
4.3.3	Understanding the social being . . . . .	14

**5 Finished and planed tasks, chronogram 15**

5.1 Documents . . . . . 15

5.1.1 To be finished . . . . . 15

5.1.2 Finished . . . . . 16

5.2 Chronogram . . . . . 16

**6 Conclusions 17**



# 1

## Introdução

---

Estudos sobre redes de interação humana foram iniciados bem antes dos computadores modernos, datam do século XIX, enquanto a fundação da “Análise de Redes Sociais”/ARS (*Social Network Analysis*/SNA) é geralmente atribuída ao psiquiatra Jacob Moreno na metade do século vinte (?). Com a crescente disponibilidade de dados relacionados à interação humana, a pesquisa destas redes tem aumentado continuamente. Contribuições podem ser encontradas em uma variedade de áreas, de ciências sociais e humanidades (?) a ciências sociais (?) e física (?, ?), dada a natureza multidisciplinar do assunto. Uma das abordagens da perspectiva de uma ciência exata é representar a rede de interação como uma rede complexa (?, ?), com a qual algumas características foram reveladas. Por exemplo, a topologia das redes de interação humana exibem um traço livre de escala, o que aponta para a existência de um pequeno número de hubs super conectados e um grande número de vértices pouco conectados.

Há um hiato de conhecimento e tecnologia entre o legado de redes complexas e o usufruto do participante. Este hiato é reativo, e há evidência de que conseguirá se manter como um ecossistema de conhecimento, tecnologia e empreendimento da sociedade em todas as suas escalas. Deve facilitar, por exemplo: elaboração e preparação de documentos, aquisição rápida de conhecimento, realização de empreitadas coletivas. Em geral: processos de coleta e difusão de informação (e bens).

Este trabalho apresenta uma confirmação deste cenário e avanços. Algumas estratégias foram selecionadas para verificar a aplicabilidade de conceitos de redes complexas para o benefício do participante. Em especial, experimentos muito simples parecem capazes de modificar estruturas sociais. Neste contexto, verificamos estabilidades temporais nas redes de interação humana, e expomos que os setores primitivos das redes (hubs, intermediários e periféricos) produzem textos bastante diferentes entre si. Este conhecimento é útil para uma tipologia não estigmatizante de participantes em redes de interação. A audiovisualização e interconexão de dados com arte e engenhocas em software deram suporte contínuo à pesquisa científica.

Aplicações foram complementadas com a Presidência da República e o PNUD/ONU.

A próxima seção apresenta considerações gerais sobre a literatura. A Seção 2 é dedicada aos dados analisados. A Seção ?? contém os métodos usados para atingir os resultados, que são explicitados na Seção ?. O cronograma de atividades e uma comparação entre afazeres planejados, em andamento e finalizados estão na Seção ?. A monografia termina com as conclusões na Seção ?, seguida de agradecimentos e bibliografia.

## 1.1 Literature review

The field of complex networks is relatively new ( $\approx 25$  years) and literature presents diverging definitions for the field itself. 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 Erdős-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: Erdős Rényi, geographical, small-world, scale-free; and the transdisciplinary approach to the environments that yield the networks. The literature on social network analysis, for 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 Erdős 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<sup>(1002)</sup>

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.

Tabela 5.1 – Cronograma de atividades

# 6

## Conclusions

---