

# REDES COMPLEXAS TEMPORAIS

## Ferramentas para a análise de narrativas históricas e literárias

Giovanna Lazzari Miotto<sup>1</sup>

Orientador: Prof. Dr. Silvio Renato Dahmen<sup>2</sup>

<sup>1</sup>Instituto de Informática, <sup>2</sup>Instituto de Física, Universidade Federal do Rio Grande do Sul

### INTRODUÇÃO

A abstração de sistemas reais como grafos é uma prática eficaz para compreendê-los. Entretanto, a estrutura estática de um grafo não reproduz o **caráter mutável de redes interpessoais**, com elementos inconstantes ao longo do tempo.

Redes sociais e outros sistemas dinâmicos, embora ubíquos, carecem de técnicas consagradas para analisá-los. O campo de **redes complexas temporais** tem emergido a fim de entender e prever o comportamento desses sistemas tempo-dependentes.

Uma narrativa literária é um tal sistema: **conflitos eclodem, alianças e inimizades germinam, e personagens entram e saem de cena**. Submetida aos métodos que processam redes reais, os resultados podem elucidar o que define um conto fictício e o distingue de uma história real.

### OBJETIVOS

- **Desenvolver** ferramentas computacionais capazes de ler, representar e visualizar narrativas como redes complexas
- **Estudar e testar** métodos recentes de medição de centralidade para redes complexas tempo-dependentes
- **Comparar** os dados estáticos e dinâmicos obtidos para uma mesma narrativa e para narrativas diferentes

### MÉTODOS

Foi implementado o método de medições de centralidades de autovetor recentemente proposto por Taylor *et al*<sup>3</sup> - uma narrativa é reduzida a uma **sequência de snapshots estáticos** da história em determinados instantes. Generalizando-se o conceito de matriz de adjacência para uma **matriz de super-adjacência** e aplicando um **fator de acoplamento temporal  $\epsilon$**  entre os *snapshots*, é calculada a chamada **joint centrality** da narrativa: a centralidade de autovetor da matriz acoplada, obtida através da extração e interpretação de seu **autovetor dominante**.

$$\mathbf{C}_{adjac}^{(t)} = \begin{bmatrix} a_{11}^{(t)} & a_{12}^{(t)} & \dots & a_{1n}^{(t)} \\ a_{21}^{(t)} & a_{22}^{(t)} & \dots & a_{2n}^{(t)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}^{(t)} & a_{n2}^{(t)} & \dots & a_{nn}^{(t)} \end{bmatrix}; \mathbf{C}(\epsilon) = \begin{bmatrix} \epsilon \mathbf{C}^{(1)} & \mathbf{I} & 0 & \dots \\ \mathbf{I} & \epsilon \mathbf{C}^{(2)} & \mathbf{I} & \ddots \\ 0 & \mathbf{I} & \epsilon \mathbf{C}^{(3)} & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}; c_{joint}(\lambda_{max}) = \begin{bmatrix} c_1^{(1)} \\ c_2^{(1)} \\ \vdots \\ c_n^{(1)} \\ c_1^{(2)} \\ c_1^{(2)} \\ \vdots \\ c_n^{(t)} \end{bmatrix}$$

Fig 1 Cada matriz de adjacência  $\mathbf{C}^{(t)}$  em um tempo  $t$  é multiplicada pelo acoplamento  $\epsilon$  e bloco-diagonalmente instalada na matriz de super-adjacência  $\mathbf{C}(\epsilon)$ . O autovetor dominante  $\mathbf{c}(\lambda)$  conterá as *joint centralities* para cada tempo  $t \in [1, T]$  para cada personagem  $n \in [1, N]$ , nesta ordem.

Conforme definidas por Prado *et al*<sup>4</sup>, foram calculadas outras centralidades, como a **vitalidade**, através da comparação do **Freeman index**, que mede o grau de concentração de uma rede ao redor do personagem mais central, antes e após a reavaliação exaustiva da narrativa com a remoção de cada personagem.

$$\mathbf{F}_{idx} = \frac{\sqrt{N(N-1)}}{(N-1)(N-2)} \sum_{n=1}^N (w_{max} - w(n)); \mathbf{V}(n) = \mathbf{F}(\mathbb{C} - \{n\}) - \mathbf{F}(\mathbb{C})$$

Fig 2 O índice de Freeman  $\mathbf{F}$  da história é medido a partir de suas *time-averaged conditional centralities*  $w$ . A vitalidade  $\mathbf{V}$  de um personagem é a diferença entre o  $\mathbf{F}$  da história sem ele e o  $\mathbf{F}$  da história intacta.

O algoritmo, implementado para **Python** com as bibliotecas **iGraph**, **NumPy** e **Matplotlib** para síntese de grafos, operações matriciais e visualização, foi então aplicado a dois objetos de estudo: a minissérie fictícia *"Anos Rebeldes"* (Globo, 1992) e a rebelião histórica do príncipe Heinrich ante seu pai, o *kaiser* Friedrich II Hohenstaufen do Sacro-Império Romano, no séc. XIII.

<sup>3</sup> Taylor, D., Myers, S. A., Clauset, A., Porter, M. A., & Mucha, P. J., **Eigenvector-based centrality measures for temporal networks**, CoRR (2015), arXiv:1507.01266v2

<sup>4</sup> Prado, S. D, Dahmen, S. R., Bazzan, A. L. C., Mac Carron, P., Kenna, R., **Temporal Network Analysis of Literary Texts**, Adv. Complex Syst. 19, 1650005 (2016), arXiv:1602.07275

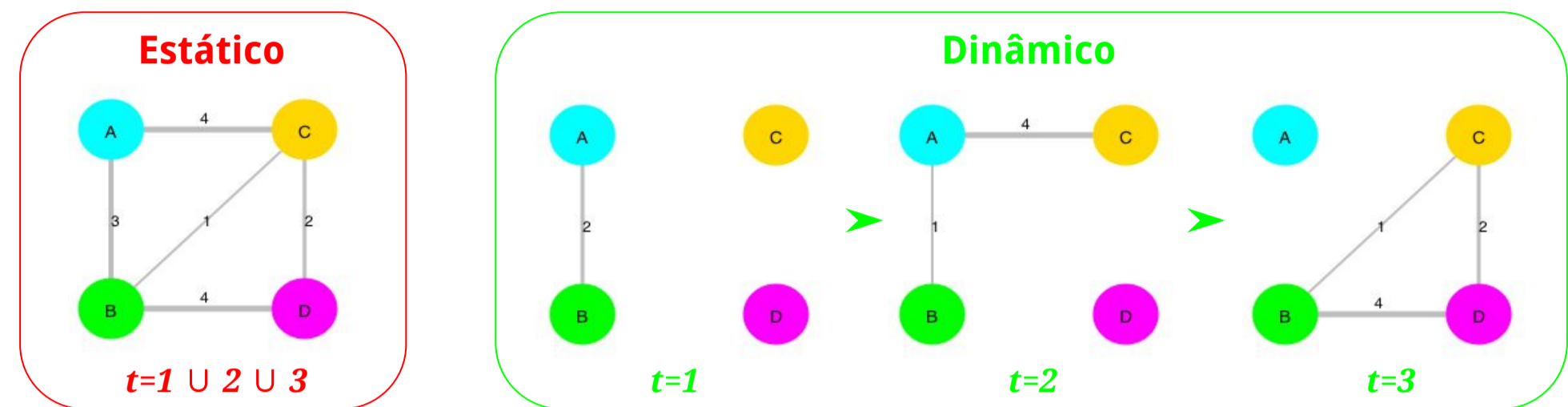


Fig 3 Um mesmo sistema de quatro atores e três *snapshots* descrito como redes agregada ( $E$ ) e acoplada ( $D$ ). Observa-se que o modelo estático parece contar uma história diferente do modelo dinâmico.

### RESULTADOS

O algoritmo foi aplicado aos objetos de estudo para análise de centralidades tradicionais (**degree**, **betweenness** e **closeness**) e comparação da centralidade de autovetor estático (**eigenvector centrality**) com as **joint**, **conditional**, **marginal** e **time-averaged centralities** propostas, em especial sob o contexto da vitalidade. Observou-se que nodos "abafados" no modelo agregado se tornaram mais salientes no acoplado, e que, entre casos dinâmicos, o histórico revelou **Freeman index** uma ordem de magnitude menor que o fictício, demonstrando que **uma obra literária tende a ser distintivamente mais centralizada, enquanto sistemas reais tendem a ser mais difusos**.

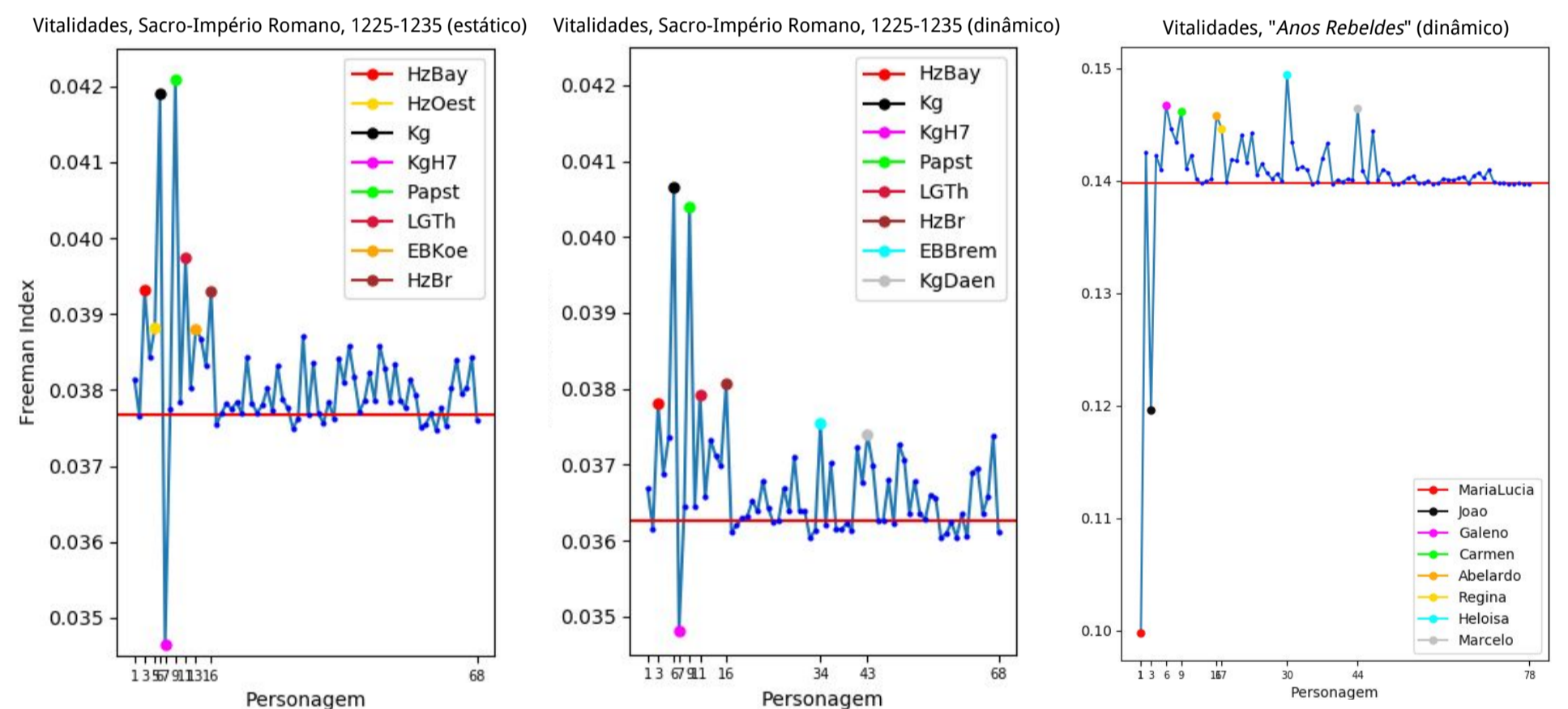


Fig 4 Vitalidades, determinadas pelo **Freeman index** da narrativa pós-remoção de cada nodo: caso histórico estático ( $E$ ), histórico dinâmico ( $C$ ) e fictício dinâmico ( $D$ ), estes dois com acoplamento  $\epsilon = 10^{-5}$ . A linha vermelha aponta o **Freeman index** original. Nota-se que o da minissérie é uma ordem de magnitude maior, indicando um sistema mais concentrado.

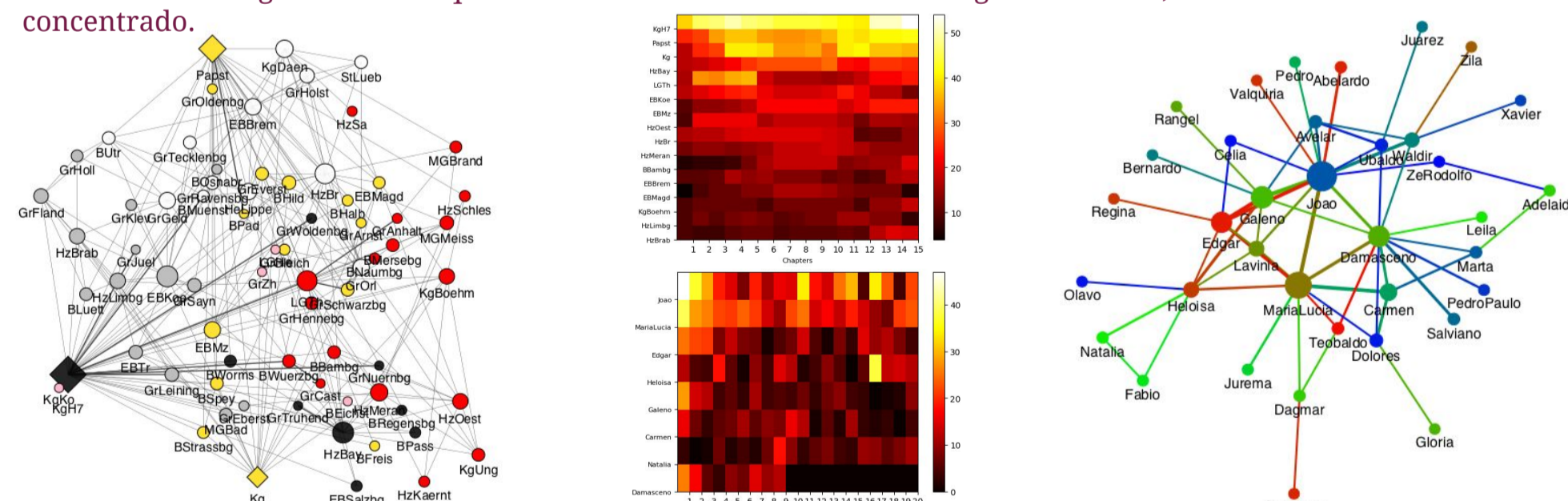


Fig 5 Sacro-Império Romano: grafo ( $E$ ) e heatmap de **degree centrality** ( $C, sup.$ ) na eclosão do conflito, em 1225. *"Anos Rebeldes"*: grafo ( $D$ ) e heatmap de **degree centrality** ( $C, inf.$ ) para o primeiro capítulo, exibido pela Globo em 1992.

### DISCUSSÃO

As ferramentas desenvolvidas poderão ser aplicadas a novos objetos de estudo e empregadas em futuras pesquisas sobre redes complexas, visando a análise e comparação quantitativas de narrativas de caráter verídico, semi-fictício e fictício, a fim de revelar as diferenças e semelhanças nos sistemas tempo-dependentes que as representam - ou seja, tanto a **estrutura comum** subjacente a narrativas distintas quanto o que **distingue um sistema real e formado espontaneamente na sociedade de um fictício**. Em particular, o método poderá ser utilizado para tentar determinar, objetivamente, a veracidade de narrativas lendárias que alegam descrever eventos reais.