

Redes Sociais e Complexas: um modelo computacional para a investigação da pós-graduação Brasileira em Ensino de Física

Nascimento, J. O. do.^{1*}; Pereira-Guizzo, C. S.; Moreira, D. M.; Monteiro, R. L. S.;
Pereira, H. B. B.; Moret, M. A.

1 Programa de Modelagem Computacional, Faculdade de Tecnologia SENAI CIMATEC, Salvador, BA, Brasil.

2 Universidade do Estado da Bahia, Salvador, Bahia, Brasil.

* e-mail: jeffersonascimento@gmail.com

Resumo

Por meio da utilização dos índices pertencentes a teoria de redes sociais e complexas é possível a análise de propriedades emergentes em redes semânticas. Este artigo tem como objetivo descrever e analisar uma rede semântica formada pelas palavras-chave pertencentes aos trabalhos de mestrado, doutorado e livre docência realizados na área de Ensino de Física, entre os anos de 1996 à 2006, no Brasil. Para alicerçar o estudo realizamos os cálculos e analisamos os índices pertencentes às redes sociais e complexas. Também apresentamos a metodologia utilizada à construção da rede, em que as palavras-chave receberam um tratamento prévio criterioso antes de um segundo passo, que é por meio da utilização de *softwares*. Percebemos indicações de que a rede apresentou a topologia livre de escala e o fenômeno pequeno mundo. Por fim, as análises indicaram também que a maioria das temáticas dos hubs emergentes na rede, estão relacionados para formação do Físico e não em metodologias para o Ensino de Física nos diversos níveis de ensino.

Abstract

With the indexes belonging to theory and complex social networks, it is possible to analyze emergent properties in semantic networks. This article aims to describe and analyze a semantic network formed by keywords, belonging to the master's works, doctoral and free teaching made in the area of Physical Education, between 1996-2006, in Brazil. In support of the study, we conducted calculations and analyze the belongings indices of social and complex networks. We also present the methodology used to construct the network, where the keywords given a careful pretreatment before a second step is with software. We noticed indications that the network presented the free topology of scale and the small world phenomenon. Finally, the analysis also indicated that most of the themes of the emerging hubs in the network are related to formation of the physics' teacher and not on methodologies for Physical Education in different educational levels.

Keywords (Palavras chaves): Complex Networks, Social Networks, Physics Education.

1. Introdução

A teoria de redes como possibilidade de estudo de entidades conectadas, ocorre por meio da teoria dos grafos e as nuances pertencentes às redes sociais e complexas. Com a utilização das métricas da teoria de redes é possível a caracterização e a investigação do comportamento das supracitadas entidades, observando as suas propriedades emergentes. Diante deste cenário o estudo de padrões e/ou modelos de previsões, por exemplo, em estruturas compostas por atores conectados têm apresentado uma frequência

maior de realização. Nesse sentido, a utilização de ferramentas computacionais cada vez mais robustas e sofisticadas, têm sido ampliadas. Verificam-se, como exemplos dessas estruturas, indivíduos, organizações, páginas *webs*, computadores, palavras em títulos, neurônios, redes de citações, redes de coautoria, etc. [1].

Neste artigo, apresentamos uma rede semântica construída pelas palavras-chave de dissertações, teses de doutorado e livre docência em Ensino de Física no Brasil, entre os anos de 1996 e 2006. Indicamos a distribuição de graus realizada e as possíveis

contribuições da aplicação da Teoria de Redes Sociais e Complexas aos padrões estruturais semânticos, verificados nas referidas palavras-chave, para o Ensino de Física nacional.

Este trabalho está organizado em 4 seções. Na segunda, apresentaremos os materiais e métodos utilizados para a construção da rede. Na terceira, os resultados e as discussões consequentes à rede semântica originada são discutidos. Na quarta seção, as considerações finais do presente artigo.

2. Materiais e métodos

Este artigo tem como finalidade realizar a construção de uma rede semântica das palavras-chave pertencentes as dissertações, teses de doutorado e livre docência especificamente na temática de Ensino de Física. O período de análise correspondeu de 1996 à 2006, conforme informações disponíveis em Salem e Kawamura [2]. Consequentemente, também verificar as propriedades emergentes na rede e, caracterizá-la conforme a sua topologia.

Para alcançarmos os objetivos, iniciamos com a abordagem de grafo $G = (V, E)$, que corresponde a uma estrutura matemática que consiste em dois conjuntos: V (finito e não vazio) e E (relações binárias sobre V) [3]. Os elementos de V são denominados de vértices e os elementos de E , são as arestas. O primeiro conjunto (V) é finito e não vazio e o segundo (E), corresponde as possíveis interações binárias sobre V .

As propriedades observadas das redes complexas, para análise da rede semântica de palavras-chave deste artigo, correspondem àquelas relacionadas aos índices de estatística básica. Com estas informações, é possível caracterizar a topologia da rede, como sendo, por exemplo, aleatória (*Random Network*), mundo pequeno (*Small-World*) ou livres de escala (*Scale Free*). Portanto, para o estudo da rede semântica proposta, observar-se-ão as seguintes propriedades: número de vértices (n), número de arestas (E), grau médio ($\langle k \rangle$), densidade (Δ), coeficiente de aglomeração (C_v) e coeficiente de aglomeração médio (C_{ws}) caminho mínimo médio (L) e diâmetro (D), especificados da seguinte maneira:

Número de vértice (n): Corresponde a cardinalidade do conjunto de vértices da rede

$$n = |V| \quad (1)$$

Número de vértices (m): Corresponde a cardinalidade do conjunto de arestas da rede

$$m = |E| \quad (2)$$

Grado médio ($\langle k \rangle$): Representa a quantidade média de conexões em cada vértice em uma rede e, é obtido da seguinte maneira:

$$\langle k \rangle = \frac{1}{n} \sum_i^n k_i \quad (3)$$

Densidade (Δ): É o total de arestas divididas pelo número máximo possível de arestas:

$$\Delta = \frac{m}{n(n-1)/2} \quad (4)$$

Esta propriedade está relacionada com a coesão na rede.

Coeficiente de aglomeração Médio (C_{ws}): O coeficiente de aglomeração médio dos vértices pertencentes a uma rede:

$$C_{ws} = \frac{1}{N} \sum_{V=1}^N C_V \quad (5)$$

O coeficiente de aglomeração de um vértice V , denominado de C_V , mede a proporção de arestas entre os vizinhos do vértice V , E_V , e o máximo número de arestas possíveis,

$$C_V = \frac{2E_V}{k_V(k_V - 1)} \quad (6)$$

Este índice apresenta a extensão em que os amigos de um indivíduo são amigos uns com os outros.

Caminho mínimo médio ou distância geodésica (L):

$$L = \frac{1}{n(n-1)} \sum_{i \neq j} d_{ij} \quad (7)$$

O termo d_{ij} corresponde a distância geodésica, em termos do número de arestas existentes entre os vértices i e j ;

Diâmetro (D): É a maior distância geodésica da rede, calculada da seguinte maneira:

$$D = \max_{(i,j)} d_{ij} \quad (8)$$

Em relação às medidas de redes sociais, as análises realizadas foram por meio das observações das centralidades de grau. Esse índice indica a importância que um determinado vértice tem, por meio de suas conexões, com vértices vizinhos conectados a ele [1].

Para que a rede pudesse ser construída, realizamos um tratamento manual nas palavras-chave. Em um arquivo de texto, cada linha recebeu um único conjunto das palavras-chave pertencentes a cada trabalho *Stricto Sensu* (correspondendo a uma sentença). Na sequência, utilizamos as regras de formatação, descritas no Quadro 1.

Quadro 1: Regras para a formatação das palavras-chave. Fonte: Adaptado de Pereira et al. [3].

1	Cada conjunto de palavras-chave corresponde a uma sentença;
2	Os sinais gráficos de cada sentença como o ponto e vírgula, ponto de interrogação, ponto de exclamação e reticências são eliminados;
3	Os nomes devem formar uma palavra simples. Por exemplo, "Aprendizagem Significativa" e "Mecânica Estatística", respectivamente, tornar-

	se-ão: “aprendizagem significativa” e “mecânica estatística”.
4	Números ordinais devem ser escritos da seguinte forma, por exemplo, “8ª” e “1º”, tornar-se-ão, respectivamente: “oitava” e “primeiro”.
5	Os números devem ser escritos de forma textual, por exemplo, “1972” e “1995”, tornar-se-ão “umnovetedois” e “umnovenovecinco”.
6	Palavras compostas devem ser consideradas como uma só palavra, por exemplo, “ensino de ciências”, tornar-se-á, “ensinodeciências”.
7	Palavras escritas de formas incorretas, devem ser corrigidas.
8	As palavras que se repetem no conjunto de palavras-chave, devem ser retiradas, permanecendo apenas uma.
9	Linguagem especializada deve ser mantida, sempre que possível.
10	Palavras que apresentam mútua importância, devem se tornar apenas uma, por exemplo, “Física Quântica”, tornar-se-á, “Física Quântica”.
11	Palavras escritas em linguagem diferente do Português, devem ser traduzidas.

No que se refere ao tratamento por meio de recursos computacionais, foram utilizados os procedimentos em conformidade com [1,4-5]. Cabe ressaltar que as ambiguidades eliminadas, seguindo o algoritmo desenvolvido em [5].

3. Resultados e Discussão

Por meio da metodologia referenciada no tópico anterior, foi possível a construção da rede semântica apresentada na Figura 1:

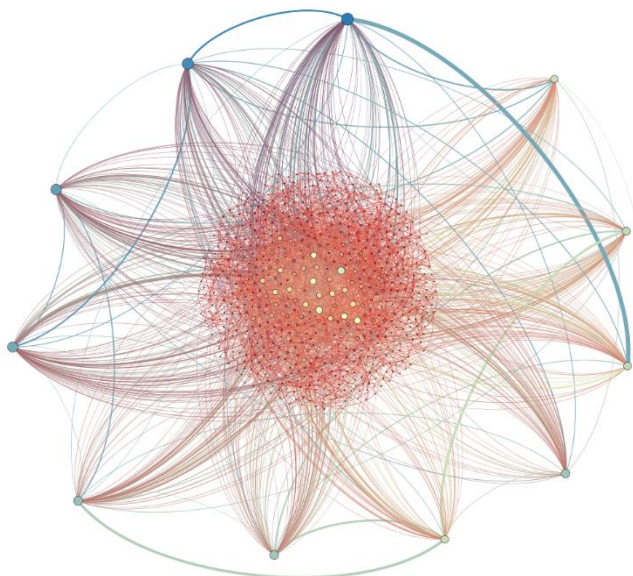


Figura 1: Rede semântica das palavras-chave dos trabalhos em Ensino de Física. Fonte: Dos autores, 2016.

Na Tabela 1, apresentamos as propriedades obtidas em relação a rede semântica das palavras-chave (Figura 1), relativo aos índices de redes complexas:

Tabela 1: Propriedades pertencentes à Rede Semântica

Propriedade	Valor	Propriedade	Valor
$n = V $	669	Maior Componente (%)	98,95
$m = E $	5910	$\langle k \rangle$	8,8
Δ	0,013	C_{ws}	0,74
L	3,06	D	6

Conforme a Tabela 1, observamos que o valor do caminho mínimo médio (L) é de aproximadamente três arestas para conectarem duas palavras na rede. O diâmetro (D), apresentou valor de seis, na rede. Tais informações refletem o fato de que as palavras pertencentes as palavras-chave de trabalhos distintos, estão conectados entre si, no máximo, por meio de seis outras palavras.

A densidade (Δ) é uma medida que está relacionada com o nível de conectividade dentro da rede, ou seja, um indicativo do nível de coesão dos vértices [4]. Conforme Pereira et al. [4] esta informação em uma dada rede semântica de títulos, indica a quantidade de ligações entre as palavras, ou seja, uma “tendência dos títulos estarem ligados através de um número grande de palavras”, pensamento este que pode ser aplicado também para a rede semântica de palavras-chave (Figura 1). O valor verificado é da ordem de 1,3%. Assim é possível que na rede semântica apresentada, haja um grande número de palavras conectando as cliques, tornando a rede espaça, com baixa densidade. Desta forma, a densidade encontrada pode corresponder a uma menor coesão entre as conexões na rede.

O valor verificado para o coeficiente de aglomeração (C_{ws}), foi da ordem de 75%, indicando uma alta conectividade entre as palavras presentes da rede semântica.

Em relação a caracterização da topologia da rede semântica, a distribuição de graus realizada, é a exposta na Figura 2. Ao se realizar um ajuste linear, o padrão verificado segue uma lei de potência, da forma $P(k) \sim k^{-\gamma}$.

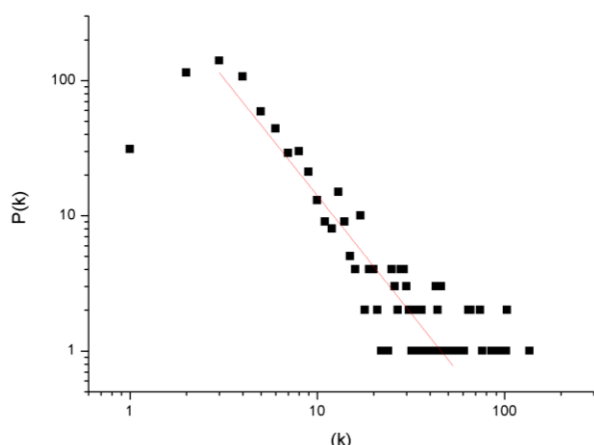


Figura 2 - Distribuição de graus da rede semântica (com $\gamma = 1,73784$ e o ajuste $R^2 = 0,85309$)

De acordo com o gráfico da Figura 2, a distribuição de graus realizada, fornece indícios de uma rede livre de escala. Conforme Barabási [6], percebe-se então que possivelmente há, para a década investigada, uma adesão preferencial por certos vértices na rede (i.e. palavras-chave), indicando a presença de *hubs*. Assim, os vértices altamente conectados têm uma probabilidade maior em adquirir novas ligações do que aqueles que se apresentam menos conectados. Na Tabela 2, realizamos uma comparação entre as redes semânticas de palavras-chave e a rede aleatória equivalentes dela (rede com o mesmo número de vértice e grau médio):

Tabela 2 – Propriedades: rede semântica de palavras-chave e a rede aleatória equivalente (*random network*)

Rede	Propriedade	Valor	Propriedade	Valor
Keywords	$n = V $	669	Maior Componente (%)	98,95
	$m = E $	5910	$\langle k \rangle$	8,8
	Δ	0,013	C_{ws}	0,74
	L	3,06	D	6
Random Network	$n = V $	669	Maior Componente (%)	98,95
	$m = E $	3002	$\langle k \rangle$	8,8
	Δ	0,013	C_{ws}	0,012
	L	3,02	D	6

Percebemos que as redes apresentaram o mesmo valor aproximado para o caminho mínimo médio (L). Observamos também que o diâmetro da rede semântica e a rede equivalente, são os mesmos. Os coeficientes de aglomeração (C_{ws}) da rede semântica de palavras-chave, apresentou valor elevado em comparação com a rede aleatórias equivalentes. Então, diante destes resultados verificados na comparação,

que estão com de acordo com o método proposto por Watts e Strogatz [7], percebemos que a rede semântica (Figura 1), apresenta o fenômeno mundo pequeno (*Small-World*).

De acordo com Pereira et al. [4], Caldeira et al. [5] e Santos Junior et al. [8], os dados da Tabela 2, para a rede das palavras-chave, também são indicadores para o padrão verificado na Figura 2: uma rede semântica caracterizada como livre de escala. Em relação ao coeficiente γ verificado na Figura 2, é menor do que normalmente é indicado na literatura científica e o proposto por Barabási e Albert [9]: $2,1 \leq \gamma \leq 4$. Contudo, as análises de redes semânticas realizadas por Fadigas et al. [1] apresentaram $\gamma \leq 2,2$ e $\gamma < 2$, na pesquisa de Seyed-allaei et al. [10]. Há uma preferência na rede por determinados vértices (*hubs*), conforme Quadro 2:

Quadro 2: Centralidades de grau dos 10 primeiros hubs da rede.

Vértices (V)	Grau (k)
História da Ciência	102
Livro didático	97
Atividade experimental	89
Linguagem	85
Formação continuada de professores	76
Formação de professores	74
Proposta didática	74
Filosofia da ciência	67
Concepção do professor	66
Mecânica	64

O Quadro 2 traz como exemplo, os dez primeiros vértices com maiores medidas de graus na rede. Este fato nos permite inferir sobre a importância (centralidade de grau) que um determinado nó tem, por meio de suas conexões, com os nós imediatamente conectados a ele (vértices vizinhos) [1]. As informações retratam, que para a rede de palavras-chave, o vértice com maior centralidade de grau foi “história da ciência”. A palavras “livro didático”, “atividade experimental” e “linguagem”, seguem como segundo, terceiro e quarto vértices, respectivamente, em termos de importâncias na rede. Seguindo este raciocínio, no Quadro 3, estão relacionados 10 vértices, selecionados pelos seus respectivos números de graus, que podem representar tópicos importantes relacionados aos conteúdos de física.

Quadro 3: Centralidades de grau de vértices que representam temáticas importantes em Física.

Vértices (V)	Grau (k)
Mecânica	64

Física moderna e contemporânea	53
Astronomia	47
Eletromagnetismo	44
Óptica	33
Gravitação	26
Termodinâmica	25
Mecânica Quântica	14
Relatividade Restrita	12
Cosmologia	9

Os vértices contidos no Quadro 3 indicam que, tópicos considerados importantes para o ensino da disciplina física, foram pouco escolhidos como palavras-chave. O vértice “Mecânica” ($k=64$) é o primeiro assunto que surge na rede semântica e apresenta grau de conexão $k=64$, ou seja, apresentam baixa centralidade de grau se comparados com os nove primeiros vértices contidos no Quadro 2. Estas informações podem indicar que novos trabalhos poderão ter, baixa probabilidade de ocorrerem com tópicos importantes no Ensino de Física.

4. Considerações finais

Por meio da análise da rede semântica formada pelas palavras-chave pertencentes aos trabalhos em Ensino de Física no Brasil (1996-2006), foram possíveis os cálculos dos índices da teoria de redes complexas e sociais. Também caracterizamos a topologia da rede e observamos vértices importante, que se destacaram nela. As propriedades obtidas indicaram que a rede semântica apresentou o fenômeno mundo pequeno (*small-world*) e também se caracterizou como uma rede livre de escala (*scale free*). Ambas as características são possíveis, pois, não são excludentes entre si.

A observação das palavras-chave com maiores centralidades de grau nos faz sugerir que, devido as características como o crescimento e a adesão preferencial da rede semântica construída e estudada, há uma maior probabilidade de ocorrerem novas palavras-chave com as palavras do Quadro 2 do que as demais palavras do quadro 3 (com exceção ao vértice Mecânica).

A alta conectividade evidenciada nas redes, pelo coeficiente de aglomeração, os caminhos mínimos médios com os mesmos valores das redes aleatórias equivalentes e, uma distribuição de graus indicando uma lei de potência, norteiam também essa caracterização. Pelo fato delas serem resistentes à retirada de vértices de forma aleatória, apresentam vulnerabilidade em caso de retirada das palavras com

maior medida de centralidade (Quadro 1). Assim, caso retirássemos da rede semântica (Figura 1) os vértices pertencentes ao Quadro 2, possivelmente apresentar-se-ia dívida em componentes o que poderia dificultar a propagação de uma determinada informação. De acordo com a Ref. [1] esta dificuldade pode ocorrer em termos de não favorecer aos leitores, mecanismos de busca dos trabalhos, por palavras-chave. Ao serem observados os vértices com maiores centralidades, foi possível inferir que, a rede semântica de palavras-chave, objetivaram conteúdos voltados para a formação de professores. Sugere-se também que a minorias dos trabalhos exploraram metodologias de Ensino de Física voltadas para os discentes.

4. Agradecimentos

JON agradece à FAPESB, referente ao apoio financeiro parcial recebido por meio da bolsa de doutorado (BOL170/2015). DMM e MAM agradecem ao CNPq pelo suporte financeiro parcial oriundo de suas bolsas de Produtividade em Pesquisa (No. 301591/2009-1) e (No. 304454/2014-1).

5. Referências

- [1] I. S. Fadigas, T. H. P. Casas, V. Senna, M. A. Moret & H. B. B. Pereira, Educação Matemática Pesquisa, 11, 167 (2009).
- [2] S. Salem & M. R. D. Kawamura, Ensino de Física no Brasil: catálogo analítico de dissertações e teses (1996-2006). São Paulo: Instituto de Física da USP / PROFIS, 2009.
- [3] J. L. Gross & J. Yellen, Graph theory and its applications. CRC press, Boca Raton FL USA 2005.
- [4] H. B. B. Pereira, I. S. Fadigas, R. L. S. Monteiro, A. J. A. Cordeiro & M. A. Moret, Physica A 390, 1192 (2011).
- [5] S. M. G. CALDEIRA, T. Petit Lobão, R. F. S. Andrade, A. Neme & J. G. V. Miranda, Eur. Phys. J. B 49, 523 (2006).
- [6] A. L. Barabási Science 325, 412 (2009).
- [7] D. J. Watts & S. H. Strogatz, Nature 393, 440 (1998).
- [8] R. P. Santos Júnior, A. A. A. O. Rodrigues, C. R. Lopes, I. P. Déjardin, J. L. B. Peixoto & M. V. Cunha, DataGramaZero 15(3), 3 (2014).
- [9] A. L. Barabási & R. Albert, Science, 286, 509 (1999).
- [10] H. Seyed-allaei, G. Bianconi & M. Marsili, Phys. Rev. E 73, 046113 (2006).