

Sistemas complexos e redes semânticas: conceitos correlacionados à orientação CTS em pesquisas sobre o Ensino de Ciência no Brasil

Nascimento, J. O. do*¹, Brito, L. P.², Moret, M. A.^{1,3}

1 Departamento de Supercomputação/Modelagem Computacional, CIMATEC, Salvador/BA, Brasil.

2 ICEN e IEMCI – Universidade Federal do Pará, UFPA, Belém/PA, Brasil.

3 Universidade do Estado da Bahia, UNEB, Salvador/BA, Brasil.

* e-mail: jeffersonascimento@gmail.com, licurgo.brito@gmail.com

Resumo

Os sistemas complexos podem ser estudados por meio de um formalismo matemático de sistemas dinâmicos, com equações diferenciais ou com equações por meio de diferenças. Também podem ser investigados por meio de mapas logísticos, autômatos celulares e a Teoria e Ciência de Redes, dentre outras possibilidades. Este artigo tem por objetivo construir e analisar redes semânticas complexas baseadas em palavras-chave para investigar em que medida a orientação Ciência, Tecnologia e Sociedade (CTS) se apresenta como alicerce temático e a correlação dela com as demais pesquisas científicas na área de Ensino de Ciências. Utilizamos em caráter inicial de pesquisa, os artigos aceitos, apresentado e publicados nos anais no Encontro Nacional de Pesquisa em Educação em Ciências (ENPEC), edição 2017. Ao fim, apresentamos os resultados dos índices estatísticos pertencentes às redes complexas e a importância da orientação CTS à pesquisa em Ensino de Ciências, conforme o ENPEC de 2017.

Palavras-chave: Sistemas Complexos, Redes semânticas, Modelagem Computacional, CTS, Ensino de Ciências.

Abstract

Complex systems can be studied through a mathematical formalism of dynamic systems, with differential equations or with equations through differences. They can also be investigated using logistic maps, cellular automata and Network Theory and Science, among other possibilities. This article aims to build and analyze complex semantic networks based on keywords to investigate the extent to which the Science, Technology and Society (CTS) orientation presents itself as the thematic foundation and its correlation with other scientific research in the field of Sciences. We initially used the articles accepted, presented and published in the annals at the National Research Meeting in Science Education (ENPEC), 2017 edition. At the end, we present the results of the statistical indexes belonging to complex networks and the importance of guidance CTS to research in Science Education, according to ENPEC 2017.

Keywords: Complex Systems, Semantic Networks, Computational Modeling, CTS, Science Teaching.

1. Introdução

Conforme Nascimento, Pereira e Moret [1] os sistemas complexos não dispõem de uma definição previamente definida. Porém, conforme a literatura científica, estes sistemas correspondem àqueles que são dinâmicos, em constante evolução, formado de um

grande número de unidades que interagem localmente, por meio de uma dinâmica não-linear. Desta forma, um texto escrito corresponde a um sistema complexo, em que não necessariamente devemos pensar no texto contido no maior livro que pudermos lembrar (em quantidade de palavras), mas a todos os textos escritos.

Assim, os títulos de obras literárias e/ou técnicas, com suas palavras-chave e os resumos associados em uma produção científica ou não, apresentam as características de sistemas complexos. Diante do exposto, este trabalho de pesquisa tem como objetivo investigar a utilização de palavras-chave como elementos auxiliares para identificar as principais temáticas no Encontro Nacional de Pesquisa em Educação em Ciências (ENPEC), cuja última edição ocorreu em julho de 2017. Desta forma, verificar como à orientação de Ciência, Tecnologia e Sociedade (CTS) se apresenta como destaque, por meio das conexões com as demais temáticas. Para tanto, buscamos um sistema complexo que contemplasse a interação entre palavras-chave dos artigos publicados nos anais do evento e, para tal, optamos pelas redes semânticas complexas.

A presente pesquisa está alicerçada pela teoria dos grafos e na teoria e ciência de redes. Assim, construímos a rede semântica baseada em palavras-chave presentes nos artigos aceitos e contidos nos anais do evento; calculamos os índices das redes complexas e a centralidade de grau pertence às redes sociais. Realizamos também a caracterização topológica da rede e verificamos às implicações fornecidas por esse modelo computacional para o contexto principal da fonte de dados, que é a Pesquisa em Educação em Ciências. Destacamos que rede semântica “[...] é o nome dado a rede de relacionamentos entre palavras ou conceitos e sua análise quantitativa torna-se mais uma contribuição para o estudo da linguagem” [2].

A teoria dos grafos, as análises das medidas de importância dos vértices pertencentes às redes sociais e a teoria das redes complexas, têm sido utilizadas na literatura científica com o objetivo de investigar variados tipos de redes. Em relação às redes semânticas, destacamos algumas obras da literatura científica realizados por Fadigas et al. [3], Pereira et al. [4], Cunha [2], Fadigas e Pereira [5], Pereira et al. [6], Rosa et al. [7], Nascimento et al. [8], Nascimento, Pereira e Moret [9] e Nascimento et al. [10], Nascimento et al. [11]. Este trabalho está organizado em 5 seções. Na segunda seção, apresentaremos a contextualização teórica. Na terceira seção, a metodologia da presente pesquisa será apresentada. Na quarta, os resultados, com as discussões pertinentes e os agradecimentos, na sequência. Ao fim, as referências que alicerçaram esta pesquisa.

2. Metodologia

A base de dados utilizada correspondeu em 1211 artigos publicados nos anais do ENPEC 2017. Por meio deles, foi possível construirmos as redes semânticas complexas para que as análises, em propriedades emergentes, pudessem ser verificadas e analisadas. Para representarmos uma rede semântica

estática, utilizamos um grafo $G = (V, E)$ que consiste em uma estrutura matemática composta por dois conjuntos: V (finito e não vazio) e E (relações binárias sobre V). Os elementos de V são denominados de vértices e os elementos de E são as arestas. As propriedades e os recursos computacionais que foram utilizados na construção e análises das redes semânticas propostas nesta pesquisa correspondem àquelas relacionadas aos índices estatísticos da teoria de redes sociais e complexas, apresentados em Nascimento et al. (2018, 2019). Com estas informações, foi possível caracterizar a topologia da rede de palavras-chave. Portanto, para a caracterização da rede semântica proposta, observamos as seguintes propriedades: número de vértices (n), número de arestas (m), grau médio ($\langle k \rangle$), densidade (Δ), coeficiente de aglomeração médio (C_{ws}), caminho mínimo médio (L) e diâmetro (D).

Estes índices escolhidos seguem também as abordagens presentes em Pereira et al. [4], pois, da mesma forma como ainda não há uma definição do que sejam os sistemas complexos, e os estudos específicos sobre redes sociais e complexas “[...] estejam amadurecendo, ainda há uma falta de padronização no uso e formalização de alguns conceitos relacionados com redes”.

Tabela 1: Índices pertencentes às redes complexas. Fonte: Adaptado de Nascimento, Pereira e Moret [1].

n	Número de Vértices	$n = V $	(1)
m	Número de Arestas	$m = E $	(2)
$\langle k \rangle$	Grau Médio do Vértice	$\langle k \rangle = \frac{1}{n} \sum_i k_i$	(3)
Δ	Densidade da Rede	$\Delta = \frac{m}{n(n-1)/2}$	(4)
C_v	Coeficiente de Aglomeração do Vértice	$C_v = \frac{2E_v}{k_v(k_v-1)}$	(5)
C_{ws}	Coeficiente de Aglomeração da Rede	$C_{ws} = \frac{1}{N} \sum_{v=1}^N C_v$	(6)
L	Caminho mínimo médio	$L = \frac{1}{n(n-1)} \sum_{i \neq j} d_{ij}$	(7)
D	Diâmetro da Rede	$D = \max(d_{ij})$	(8)

Para o pré-processamento das palavras-chave, cálculos e utilização de algoritmos para construção das redes semânticas, devem ser seguidas as seguintes regras:

Quadro 1: Regras para o pré-processamento das palavras-chave. Fonte: Adaptado de Nascimento et al. [8].

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 "umnovecinco".
6	Palavras compostas devem ser consideradas como uma só palavra, por exemplo, "ensino de ciências", tornar-se-á, "ensinociências".
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-chave escritas em linguagem diferente do Português, devem ser traduzidas para o português.

A seguir, um exemplo do método de construção das redes semânticas, cuja base de dados é o ENPEC 2017:

- Conjunto 1:** cts, clube de ciências, iniciação científica;
Conjunto 2: ensino de astronomia, abordagem temática, cts;
Conjunto 3: ctsa, água, livro didático, ensino de ciências.

Com base nos conjuntos de palavras-chave anteriores, origina-se a rede semântica representada na Figura 1:

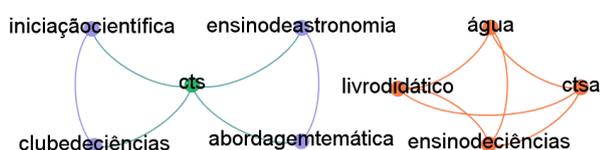


Figura 1. Exemplo de construção da rede semântica de cliques baseada em palavras-chave. Fonte: Dos autores.

Em relação aos algoritmos utilizados, sugerimos as obras Pereira et al. [4], Pereira et al. [6], Nascimento, Pereira e Moret [1], e Nascimento et al. [10], Nascimento et al. [11] e Nascimento [12].

3. Discussão dos resultados

Por meio da metodologia proposta de construção de redes disposta na sessão anterior, apresentamos a

rede semântica baseada em palavras-chave dos artigos publicados nos anais do ENPEC 2017 (Figura 2):

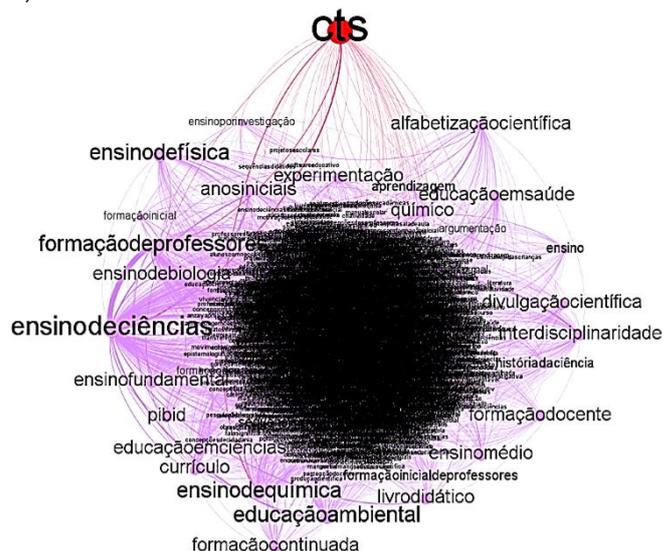


Figura 2: Rede semântica baseada em Palavras-Chave do ENPEC 2017. Fonte: Dos autores

O tamanho dos vértices mais periféricos está proporcional ao seu respectivo grau de conexão na rede. A maior componente apresenta vértices com as tonalidades da cor lilás. O vértice em destaque, CTS (vermelho), está em posição de destaque por corresponder a um *hub* e ser também o foco da presente pesquisa. Na tabela 1 apresentamos os índices de Rede Semântica de Palavras-chave (RSPC) do ENPEC 2017, da Figura 2:

Tabela 1: Índices pertencentes à RSPC do ENPEC 2017. Fonte: Dos autores.

Rede	Propriedade	Valor	Propriedade	Valor
RSPC	$n = V $	2077	Maior Componente (%)	89,41
	$m = E $	5830	$\langle k \rangle$	5,614
	Δ	0,003	C_{ws}	0,848
	L	3,399	D	8
RSPC (Maior Comp.)	$n = V $	1857	Maior Componente (%)	100
	$m = E $	5524	$\langle k \rangle$	5,949
	Δ	0,003	C_{ws}	0,832
	L	3,399	D	8

Conforme a Tabela 1, observamos que o valor do caminho mínimo médio (L) é de aproximadamente três arestas, para conectarem duas palavras quaisquer na rede. Ao observarmos o diâmetro (D), que corresponde ao maior menor caminho entre dois vértices. Tais informações refletem o fato de que as palavras pertencentes as palavras-chave de trabalhos distintos estão conectadas entre si, por meio de oito

Assim, retirando das redes semânticas baseadas em palavras-chave do ENPEC 2017 o *hub* CTS (Figura 2 e Figura 3), a rede apresentaria um colapso, desfragmentando-se de tal forma que o sistema complexo resultante da interação das palavras-chave, não haveria mais como ser estudado/analísado, pois, não mais existiria. As propriedades que emergem da dinâmica de evolução do sistema, como o coeficiente de aglomeração, por exemplo, não teriam mais a sua observância.

Utilizando o método proposto por Watts e Strogatz [13], realizamos uma comparação dos índices estatísticos entre a RSPC que representa o sistema (maior componente) com a sua rede aleatória equivalente, conforme Tabela 3:

Tabela 3: índices pertencentes à maior componente da RSPC e a sua equivalente aleatória do ENPEC 2017. Fonte: Dos autores.

Rede	Propriedade	Valor	Propriedade	Valor
RSPC (Maior Comp.)	$n = V $	1857	Maior Componente (%)	100
	$m = E $	5524	$\langle k \rangle$	5,949
	Δ	0,003	C_{ws}	0,832
	L	3,399	D	8
RSPC Aleatória equivalente	$n = V $	1857	Maior Componente (%)	100
	$m = E $	5493	$\langle k \rangle$	5,949
	Δ	0,003	C_{ws}	0,002
	L	4,436	D	8

Diante de tal comparação, pelo método proposto, foi possível verificarmos que a RSPC do ENPEC 2017 apresenta o fenômeno *Small-World* (mundo pequeno), ou seja, alta aglomeração e caminhos curtos entre seus vértices (i.e. palavras).

Ainda em busca de caracterizarmos a RSPC do ENPEC 2017, ratificamos que pela Tabela 1, Tabela 2 e Tabela 3 ratificamos o seguinte fato dito anteriormente: há uma preferência na rede por de terminados vértices, os *hubs*, especificamente os contidos na Tabela 2. Desta forma, realizamos a distribuição de graus da RSPC do ENPEC 2017, conforme Figura 4. O padrão estatístico observado sugere que a rede apresenta um comportamento de rede *Scale Free*, com a presença de uma Lei de Potência na forma $P(k) \sim k^{-\gamma}$:

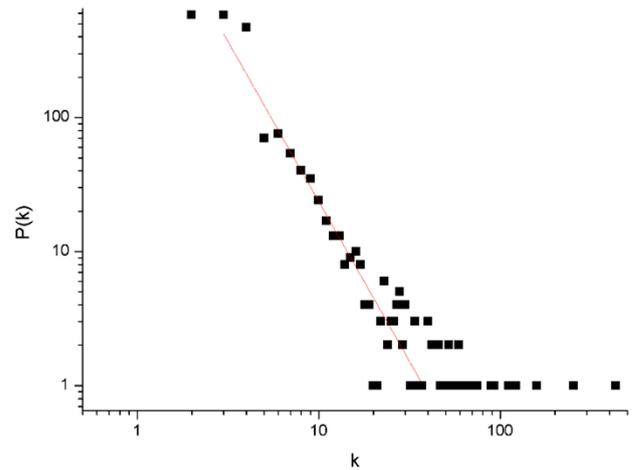


Figura 4 – Distribuição de graus da rede semântica baseada em palavras-chave do ENPEC 2017 (com $\gamma = 2,3869$ e o ajuste $R^2 = 0,89627$). Fonte: Dos autores.

Considerando a estatisticamente a Lei de Potência encontrada, conforme a figura anterior, em que há total coerência com a proposta de Barabási e Albert [14] e Barabási [15], percebemos então que há, uma adesão preferencial pelos vértices (*hubs*). Assim, os vértices com elevados valores de arestas apresentam uma probabilidade maior em adquirir novas ligações do que aqueles que são menos conectados. Este fenômeno é percebido em natureza como “[...] os ricos ficam cada vez mais ricos” [14]. O valor do coeficiente angular (γ), encontrado no ajuste linear, Figura 4, pertence ao intervalo proposto por Barabási e Albert [14]: $2,1 \leq \gamma \leq 4$. Concluimos então em relação a caracterização da topologia da rede semântica, aplicado o método de Watts e Strogatz [13] e calculada a distribuição de graus da rede, argumentamos que a mesma é caracterizada como *Small-World* (mundo pequeno) e *Scale Free* (livre de escala).

4. Considerações finais

Realizamos as análises no sistema complexo emergente do ENPEC 2017, observando as propriedades suscitadas pela forma singular na qual a interação entre as diferentes palavras-chave em 1211 artigos aceitos, apresentados e publicados, suscitaram.

As redes semânticas baseadas em palavras-chave com base em alguns índices da teoria e ciência de redes (redes complexas e sociais), foi possível a caracterização das mesmas. Conforme o método apresentado por Watts e Strogatz [13], realizamos uma comparação entre os seus coeficientes de aglomeração e o caminho mínimo médio das redes semânticas, entre as componentes gigantes delas e as suas respectivas redes aleatórias equivalentes. Seguimos os nortes fornecido pelo método Watts e Strogatz [13] e consideramos as restrições propostas por Watts [16]: a rede deve ser esparsa, não dirigida, não ponderada, conectada e não conter laços. Desta forma, pudemos argumentar que as redes semânticas apresentaram o

fenômeno mundo pequeno (*Small-World*). As redes mundo pequeno são eficientes no trânsito de informações. O que demonstra a qualidade existente nela, por meio das conexões realizadas, na rede semântica do ENPEC 2017. Conforme Barabási e Albert [14], realizamos a distribuição de graus das redes semânticas e encontramos, um comportamento de que seguem uma lei de potência do tipo $P(k) \sim k^{-\gamma}$. Logo, as redes também apresentaram características de serem livres de escala (*Scale Free*). Estes comportamentos não são antagônicos entre si.

Como as redes *Scale Free* são resistentes à retirada de vértices de forma aleatória e apresentam vulnerabilidade em caso de retirada dos vértices que correspondem aos *hubs* na rede da Figura 2, contidos na Tabela 2, isso sugere que, sem elas, os “discursos” do ENPEC 2017 não existiria. Isso significa que, caso extraíssemos da rede semântica da Figura 2, vértices altamente conectados (*hubs*) da Tabela 2 e, também, aqueles vértices que representam os pontos de cortes, a rede passaria a ter muitos componentes o que dificultaria a difusão da informação em pesquisas realizadas no ENPEC 2017. Diante destes fatos, a própria adesão de futuros trabalhos nos próximos eventos estaria comprometida, pela ausência da adesão preferencial (característica da rede livre de escala)

Também haveria uma impossibilidade de encontrarmos trabalhos semelhantes, ou pertencentes à base de dados do evento, pelos instrumentos de buscas científicas, como pelo Google Acadêmico, Página de Periódicos da CAPES ou no próprio evento. Então, pelas características das redes, pelos índices de redes complexas podemos argumentar que houve a difusão da informação, por meio da temática de CTS, no ENPEC 2017. Tais fatos expostos ratificam neste artigo, em caráter inicial de pesquisa, uma importância *sine qua non*, não apenas para o Encontro Nacional de Pesquisa em Educação em Ciências, mas para o cenário de Ensino de Ciências para o Brasil, quiçá, o mundo científico na supracitada área.

5. Agradecimentos

Marcelo A. Moret agradece ao CNPq pelo suporte financeiro parcial oriundo de sua bolsa de Produtividade em Pesquisa (No. 305291/2018-1). Jefferson Nascimento agradece a oportunidade de ter realizado pós-doutoramento no departamento de supercomputação/modelagem computacional no Cimatec/BA.

6. Referências

[1] NASCIMENTO, J. O.; PEREIRA, H. B. B.; MORET, M. A. Grafos e teoria de redes: uma análise do ensino de física brasileiro no período 1972-2006 por meio de cliques de

palavras-chave. **REVISTA CEREUS**, Gurupi, v. 10, n. 2, p. 315-339, 2018.

[2] CUNHA, M.V. **Redes semânticas baseadas em títulos de artigos científicos**. 127f. Dissertação (Mestrado em modelagem computacional e tecnologia industrial) – Faculdade de Tecnologia Senai CIMATEC, Salvador, 27 nov. 2013.

[3] FADIGAS, I. S.; CASAS, T. H. P.; SENNA, V.; MORET, M. A.; PEREIRA, H. B. B. Análise de redes semânticas baseada em títulos de artigos de periódicos científicos: o caso dos periódicos de divulgação em educação matemática. **Educação Matemática Pesquisa**, v. 11, n. 1, 2009.

[4] PEREIRA, H. B. B. et al. Semantic networks based on titles of scientific papers. **Physica A: Statistical Mechanics and its Applications**, v. 390, n. 6, p. 1192-1197, 2011.

[5] FADIGAS, I. S.; PEREIRA, H. B. B. A network approach based on cliques. **Physica A: Statistical Mechanics and its Applications**, v. 392, n. 10, p. 2576-2587, 2013.

[6] PEREIRA, H. B. B. et al. Density: A measure of the diversity of concepts addressed in semantic networks. **Physica A: Statistical Mechanics and its Applications**, v. 441, p. 81-84, 2016.

[7] ROSA, M. G.; FADIGAS, I.; MIRANDA, J. G. V.; CUNHA, M. V.; MONTEIRO, R.L.S.; PEREIRA, H. B. B. Robustness in semantic networks based on cliques. **Physica. A (Print)**, v. 472, p. 94-102, 2017.

[8] NASCIMENTO, J. O. do; MONTEIRO, R. L. S.; MOREIRA, D. M.; MORET, Marcelo A.; PEREIRA, H. B. B. Semantic networks of keywords from Brazilian dissertations on physics teaching. **Discontinuity, Nonlinearity and Complexity**, 2018.

[9] NASCIMENTO, Jefferson Oliveira do; PEREIRA, Hernane Borges de Barros; MORET, Marcelo Albano. Redes de afiliação no ensino de física brasileiro: uma análise dos programas de pós-graduações por meio de redes sociais e complexas. **Revista Mundi Engenharia, Tecnologia e Gestão** (ISSN: 2525-4782), v. 3, p. 58-1-58-15, 2018.

[10] NASCIMENTO, J. O. do; PEREIRA, H. B. B.; MORET, M. A.; BARBOSA, L. P.; TAKIISHI, H. The Network Theory as computational modeling tools for studies in CBECiMat. **MATERIALS SCIENCE FORUM** (ONLINE), v. 930, p. v-xi, 2018.

[11] NASCIMENTO, J. O. do; PEREIRA, H. B. B.; CUNHA, M. V.; MORET, M. A. Sistemas complexos e Ciências das redes: redes semânticas baseadas em abstracts e keywords do Ensino de Física Nacional. **REVISTA CEREUS**, v. 11, p. 161-183, 2019.

[12] NASCIMENTO, J. O. do. Redes sociais e complexas: redes semânticas do Ensino de Física Brasileiro. 123f. Tese (Doutorado em modelagem computacional e tecnologia industrial) – CIMATEC, Salvador, 28 jun. 2019.

[13] WATTS, Duncan J.; STROGATZ, Steven H. Collective dynamics of small-world networks. **Nature**, v. 393, n. 6684, p. 440-442, 1998.

[14] BARABÁSI, A. L. e ALBERT, R. Emergence of Scaling in Random Networks. **Science**, n. 286, pp. 509-512, 1999.

[15] BARABASI, A.L. Scale-free networks: a decade and beyond. **Science**. 2009; 325:412– 413.

[16] WATTS, D. J. Small words: The dynamics of networks between order and randomness. Princeton University Press, Princeton, 1999.