

Sistemas complexos e Ciências das redes: redes semânticas baseadas em abstracts e keywords do Ensino de Física Nacional

Complex Systems and Network Science: semantic networks based on abstracts and keywords of Brazil's Physics Teaching

Jefferson Oliveira do Nascimento¹, Hernane Borges de Barros Pereira², Marcelo do Vale Cunha³, Marcelo Albano Moret⁴.

RESUMO

Ao ser elaborada uma obra científica, faz-se necessário a construção e apresentação de seu *abstract* (resumo) e da escolha de *keywords* (palavras-chave), que correspondem a sínteses sucintas do trabalho escrito, expondo de forma direta as nuances apresentadas no corpo do texto (*abstract*) e por palavras centrais do assunto abordado (*keywords*). Como possibilidades de buscas de resultados das interações semânticas em produções científicas, a fim de que possamos estudar as propriedades emergentes resultantes destas interações, podemos realizar análises utilizando os seus respectivos *abstracts* e *keywords*. Diante da diversidade de técnicas científicas para o estudo de possíveis padrões que possam ocorrer nos *abstracts* e *keywords*, utilizamos a Teoria dos Grafos, Teoria e Ciências das Redes, por meio de redes semânticas. Este artigo tem como objetivo apresentar o método de construção de um modelo computacional baseado em *abstracts* e *keywords* de dissertações e teses do Ensino de Física nacional, calcular os índices estatísticos das redes complexas e as medidas de centralidade de grau e de intermediação pertencentes às redes sociais. Verificar a topologia das redes semânticas e as implicações fornecidas pelo modelo computacional para o contexto principal da fonte de dados à área de conhecimento investigada. Verificamos que as redes apresentam o fenômeno *Small-World* e, pelas leis de potências encontradas, sugerem ser também *Scale Free*.

Palavras-chave: Redes Semânticas Complexas, Redes Sociais, Ensino de Física.

ABSTRACT

When a scientific work is elaborated, it is necessary to construct and present its abstract and the chosen keywords, which correspond to succinct syntheses of the written work, exposing in a direct way the nuances presented in the body of the text (*abstract*) and by the central words of the paper (*keywords*). As possibilities to search results of semantic interactions in scientific productions, so that we can study the emergent properties resulting from these interactions, we can perform analyzes using their respective abstracts and keywords. Faced with the diversity of scientific techniques for the study of possible patterns that may occur in abstracts and keywords, we use the Theory of Graphs, Network Theory and Sciences, through semantic networks. This article aims to present the method of constructing a computational model based on abstracts and keywords of dissertations and theses of the National Physics Teaching, calculate the statistical indexes of complex networks and the measures of centrality of degree and intermediation belonging to social networks. To verify the topology of the semantic networks and the implications provided by the computational model for the main context of the data source to the researched area of knowledge. We have verified that the networks present the phenomenon *Small-World* and, by the laws of powers found, also suggest being *Scale Free*.

Keywords: Complex Semantics Network, Social Networks, Physics Teaching.

¹ Mestre e Doutorando em Modelagem Computacional e Tecnologia Industrial. Departamento de Modelagem Computacional, CIMATEC – Salvador, BA, Brasil. E-mail: jeffersonascimento@gmail.com.

² Doutor em Engenharia Multimídia (Universidat Politécnicat de Cataluña). Universidade do Estado da Bahia; Programa de Modelagem Computacional - CIMATEC - Salvador/BA. E-mail: hernanebbpereira@gmail.com.

³ Mestre e Doutorando em Modelagem Computacional e Tecnologia Industrial. Departamento de Modelagem Computacional, CIMATEC – Salvador, BA, Brasil. E-mail: celaocunha@gmail.com

⁴ Doutor em Biofísica (UFRJ). Universidade do Estado da Bahia; Programa de Modelagem Computacional - CIMATEC - Salvador/BA. E-mail: mamoret@gmail.com

1. INTRODUÇÃO

Ao ser elaborada uma obra científica, faz-se necessário a construção e apresentação de seu resumo (*abstracts*), correspondendo a uma condensação sucinta do trabalho escrito, expondo de forma direta as nuances apresentadas no corpo do texto. A arte e a precaução ao se elaborar um resumo de uma obra científica se assemelha com a escolha dos títulos e das palavras-chave. Assim como os títulos e palavras-chave de trabalhos científicos correspondem a uma importante e inicial demonstração de um conteúdo técnico que se discutirá no texto de uma produção científica (CUNHA, 2013, NASCIMENTO et al., 2018), assim também é para a elaboração do resumo.

Como possibilidades de buscas de interações semânticas em produções científicas, a fim de que possamos estudar propriedades emergentes resultantes destas interações (características de sistemas complexos), podem ser realizados estudos e/ou análises, observando os seus respectivos resumos (*abstracts*) e seus conjuntos de palavras-chave (*keywords*). Uma dessas possibilidades é partirmos do alicerce fornecido pela Teoria de Redes (Redes Sociais e Complexas), por meio de redes semânticas. Desta forma, para Pereira (2013, p. 39-40), as Redes Sociais e Complexas, apresentam as seguintes características:

[...] possuem propriedades não triviais, inexistentes em modelos mais simples. Essas redes são sistemas complexos que, em geral, envolvem inúmeros elementos organizados em estruturas que podem existir, ou coexistir, em diferentes escalas. Seus processos de ação e de organização não são usualmente descritos por regras simples ou redutíveis a apenas um nível explanatório. Frequentemente, suas características principais emergem de interações entre suas partes constituintes e não podem ser previstas a partir de uma compreensão isolada de cada uma destas partes. Nesse sentido, deparamo-nos com a especialização de técnicas e métodos de cálculo para propriedades emergentes (e.g. densidade, coeficiente de aglomeração, caminho mínimo, distribuição de graus, centralidade e prestígio, etc.) de redes sociais e complexas. O estudo desses índices favorece a análise, interpretação e resolução de problemas em diversas áreas.

Especificamente sobre uma rede semântica, para Steyvers e Tenenbaum (2005), é uma representação abstrata de apenas um aspecto do conhecimento semântico, que corresponde às relações entre pares de palavras. Um outro pensamento é exposto em Cunha (2013, p. 4), em que uma 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”. Diante da diversidade de definições encontradas na literatura científica, enfatizamos que nesta pesquisa estamos alicerçados

em Rosa et al. (2017). Para estes autores uma rede semântica é um sistema utilizado para representar o conhecimento imbuído de intenção de funcionalidade e estabelecido pelo contexto. Uma rede semântica é composta de palavras, conceitos ou entidades com significado semântico e seus relacionamentos e, é representada pela Teoria dos Grafos (ibidem).

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 intuito de investigar variados tipos de redes. Nesse viés, destacam-se algumas obras da literatura científica de estudos relacionados às redes semânticas, realizados por Caldeira (2005), Fadigas et al. (2009), Pereira et al. (2011), Cunha (2013), Fadigas e Pereira (2013), Pereira et al. (2016), Rosa et al.(2017), Costa et al. (2017), Nascimento, Pereira e Moret (2017), Nascimento et al. (2018a, 2018b) e Nascimento, Pereira e Moret (2018a, 2018b).

Diante do exposto este artigo apresenta como o objetivo construir, analisar e investigar as redes semânticas baseadas em *abstracts* (resumos) e *keywords* (palavras-chave) de dissertações e teses do Ensino de Física Brasileiro apresentadas e defendidas no período entre 1972 e 1995. O presente trabalho está subdividido da seguinte maneira: na primeira parte da pesquisa, nossas considerações iniciais são apresentadas na introdução. No tópico seguinte, apresentamos os materiais e métodos utilizados. Na sequência os resultados são disponibilizados e, por fim, as considerações finais.

2. MATERIAIS E MÉTODOS

Como fonte de dados para mineração das informações referentes aos resumos e palavras-chave, utilizamos o primeiro e segundo volumes do Catálogo Analítico de Ensino de Física no Brasil: dissertações e teses. Para o período de 1972 - 1992 há 177 trabalhos *Stricto Sensu*, entre dissertações de mestrado e teses de doutorado (USP, 1992). Para o período posterior, 1992 – 1995, há 53 trabalhos que correspondem na íntegra às pesquisas realizadas neste período (USP, 1996). Há também 16 dissertações e teses remanescentes e catalogadas do período de 1972-1992 neste último catálogo. Então, para o período total (1972-1995), encontramos 246 dissertações e teses apresentadas e defendidas na área de Ensino de Física no Brasil. As redes semânticas apresentadas neste artigo, são baseadas em *abstracts* (resumos) e *keywords* (palavras-chave) dos trabalhos citados acima.

Para representarmos uma rede semântica, utilizamos um grafo $G = (V, E)$ que consiste em uma estrutura matemática composta por dois conjuntos: V (finito e não vazio) e (relações binárias sobre V) (GROSS; YELLEN, 2005). Os elementos de V são denominados de vértices e os elementos de E são as arestas. Inspirados em Pereira (2011) e Nascimento, Pereira e Moret (2018) as métricas de redes complexas e sociais observadas e calculadas nas redes semânticas desta pesquisa são: número de vértices, número de arestas, grau médio, densidade, coeficiente de aglomeração, caminho mínimo médio e diâmetro, especificados da seguinte maneira:

Número de vértice (n):

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

Número de arestas (m):

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

Grado médio ($\langle k \rangle$):

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

Densidade (Δ):

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

O coeficiente de aglomeração de um vértice V :

$$C_V = \frac{2E_v}{k_v(k_v - 1)} \quad (5)$$

Coeficiente de aglomeração Médio (C_{ws}):

$$C_{ws} = \frac{1}{N} \sum_{v=1}^N C_V \quad (6)$$

Caminho mínimo médio ou distância geodésica (L):

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

Diâmetro (D):

$$D = \max (d_{ij}) \quad (8)$$

Em relação à caracterização da rede semântica por meio do cálculo das medidas de centralidades dos vértices, utilizamos a centralidade de grau e a centralidade de proximidade. Centralidade de grau: Número de vértices adjacentes a um determinado vértice v . Depende do número de arestas incidentes:

$$C_D(v_i) = d(v_i) = \sum_j v_{ij} \quad (9)$$

Centralidade de Intermediação: Conforme Freeman (1978/9, p. 221), é a frequência de um vértice v ocorrer entre pares de outros vértices em caminhos mais curtos (geodésicos) que os conectam, ou ainda, é a probabilidade de um vértice k ocorrer aleatoriamente em geodésicas ligando v_i e v_j :

$$C_B(v_k) = \sum_{i,j} b_{ij}(v_k) = \frac{g_{ij}(v_k)}{g_{ij}} \quad (10)$$

Para que as redes semânticas pudessem ser construídas, realizamos inicialmente um processamento manual nas palavras que constituem os resumos e as palavras-chave das dissertações e teses em Ensino de Física. Posteriormente utilizamos *softwares* e algoritmos para visualização das redes. Estes procedimentos utilizados são os indicados pelas regras e nortes fornecidos em Pereira et al. (2011), Nascimento et al. (2018) e Nascimento, Pereira e Moret (2018). A título de exemplo à construção das redes semânticas propostas nesta pesquisa, baseadas em resumos, apresentamos no Quadro 1

dois resumos escolhidos de forma aleatória, pertencentes a base de dados utilizada nesta pesquisa, para construirmos a rede resultante:

Quadro 1 - Dois resumos de trabalhos *Stricto Sensu* em Ensino de Física utilizados para a construção da uma rede semântica. Fonte: USP (1992, 1996).

Resumo 1
Examina o debate entre Bohr e Einstein sobre a interpretação da teoria quântica no período 1927-1949, de forma a explicitar a articulação entre pontos de vista físicos e epistemológicos. Inicialmente estuda a formação da teoria quântica, desde os resultados de Planck (1900) até a fundamentação desenvolvida por Bohr (1927). Em seguida, faz uma análise das interpretações propostas por diversos outros autores (Fock, Taketani e Langevin) tomando como parâmetros a dualidade onda-partícula, a interação objeto/meio de investigação, a descrição probabilística da natureza, a completude da teoria, a adequação teoria/objeto e o conceito de não-separabilidade. Como principal resultado, atesta a importância dos aspectos epistemológicos contidos nesse debate e a riqueza de significados científicos: enquanto Bohr precisa o objetivo e clarifica a consistência da teoria quântica, Einstein preocupa-se com sua evolução numa perspectiva ampla. Ressalta a importância de estudos epistemológicos para a compreensão das teorias físicas e, por extensão, para o seu ensino.
Resumo 2
Discute importância e métodos para o uso de microcomputadores no ensino, em particular no caso da física. Apresenta algumas características e principais aplicações desenvolvidas com computadores na educação. Propõe uma classificação do software educacional baseada em três aspectos: a relação deste com o currículo, sua função instrucional e sua utilização. Sistematiza as etapas de produção de um software didático incluindo: fase de preparação pedagógica, de programação e de testes com os alunos, incluindo, também, uma proposta de avaliação do software. Sugere aplicações de alguns softwares no ensino de física, particularmente para o conteúdo de mecânica quântica. Defende a introdução da informática na escola como apoio ao professor e ao laboratório, apontando a utilização do computador nas categorias: tutor, tutorado e ferramenta.

A seguir a rede semântica baseada nos resumos contidos no Quadro 1:

53 com resumos e 49 com palavras-chave. Com estas informações, apresentamos na Figura 2 e Figura 3, as redes semânticas baseada em resumos (1972-1992 e 1992-1995, respectivamente) do Ensino de Física:

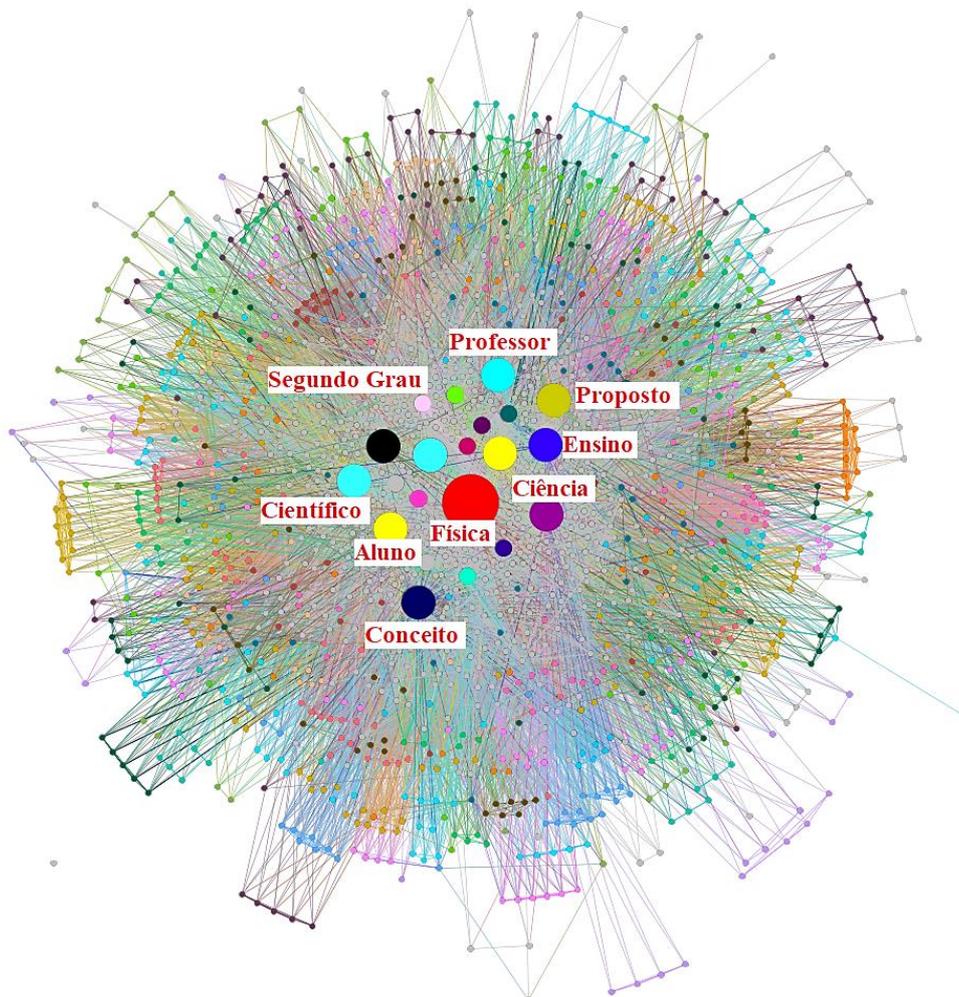


Figura 2 - Rede semântica baseada em resumos do Ensino de Física Brasileiro no período de 1972-1992.

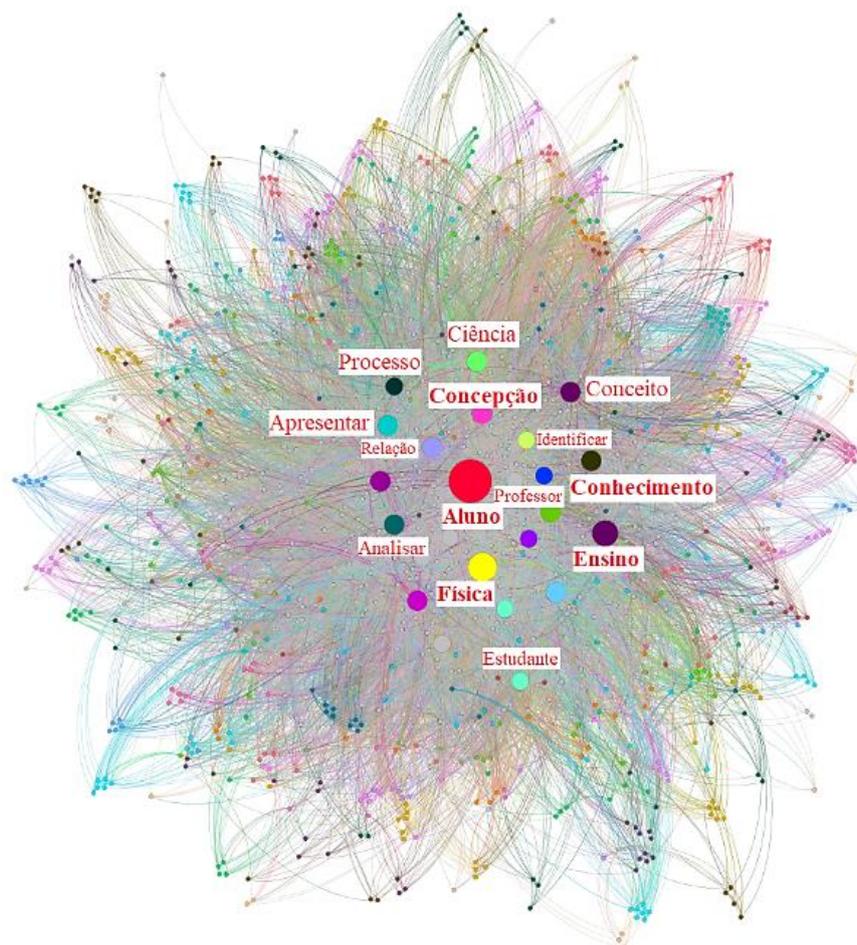


Figura 3 - Rede semântica baseada em resumos do Ensino de Física Brasileiro no período de 1992-1995.

Utilizamos basicamente quatro algoritmos para a visualização das redes das Figuras 2 e 3: *Yifan Hu proporcional*, Não sobrepor, Force Atlas2, e *Frunchterman Reingold* do *software Gephi*. Com a ação conjunta destes algoritmos conseguimos enfatizar os *hubs* nas redes, localizando-os mais na região central dos respectivos grafos. Nas figuras os tamanhos dos *hubs* estão em proporcionalidade com o grau de conexão de cada um. As cores foram baseadas conforme o espectro luz visível, as que apresentam maiores velocidades (vermelho, por exemplo), pertencem aos *hubs* com maiores valores de centralidade de grau. Apresentamos a seguir, as redes semânticas baseadas em Palavras-chave para os períodos de 1972-1992 e 1992-1995, respectivamente:

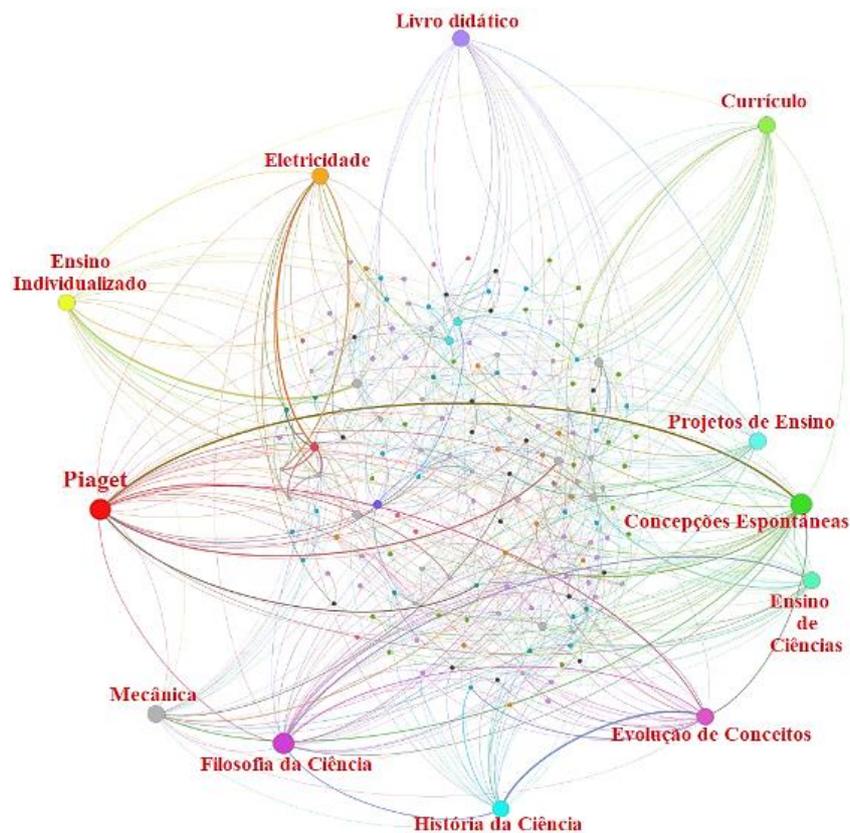


Figura 4 - Rede semântica baseada em palavra-chave do Ensino de Física Brasileiro no período de 1972-1992.

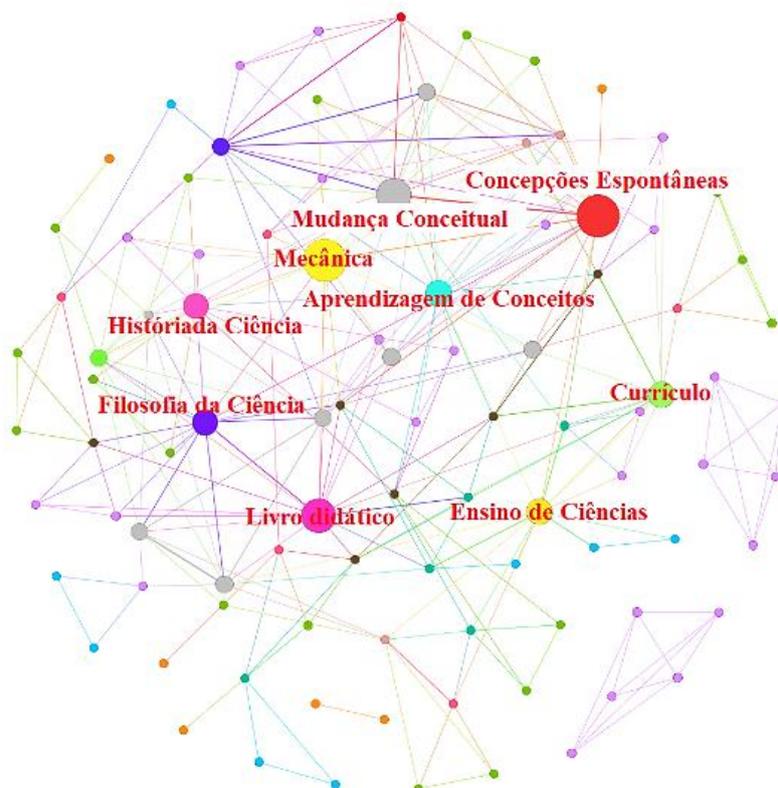


Figura 5 - Rede semântica baseada em palavra-chave do Ensino de Física Brasileiro no período de 1992-1995.

Seguindo o raciocínio para visualização das Figuras 2 e 3, utilizamos 2 algoritmos para a visualização das redes baseadas em palavras-chave: Não sobrepor e Force Atlas2, também pertencentes ao *software Gephi*. Com a utilização deles, Na Figura 4, destacamos os *hubs* para a periferia da rede e, na Figura 5, mais na região central.

Nas tabelas abaixo, apresentamos os índices de estatística básica da teoria de redes, devidamente calculados para as redes de *abstracts* (resumos) e *keywords* (palavras-chave):

Tabela 1 - Índices de redes complexas verificados para a rede semântica baseada em *abstracts* no período de **1972-1992**.

Rede Semântica	n	m	Qt. Componentes	Maior componente (%)	$\langle k \rangle$	Δ	C_{ws}	L	D
<i>Abstracts</i>	1565	31329	2	99,94%	40,037	0,026	0,752	2,335	4
Maior componente	1564	31329	1	100%	40,063	0,026	0,752	2,335	4
Aleatória equivalente	1564	31898	1	100%	40,508	0.0260	0.0262	2.310	3

Tabela 2 – Índices de redes complexas verificados para a rede semântica de *abstracts* de **1992-1995**.

Rede Semântica	n	m	Qt. Componentes	Maior componente (%)	$\langle k \rangle$	Δ	C_{ws}	L	D
<i>Abstracts</i>	1223	24086	4	99,75%	39,388	0,032	0,728	2,291	4
Maior componente	1220	24086	1	100%	39,485	0,032	0,728	2,291	4
Aleatória equivalente	1220	24132	1	100%	39.560	0.032	0.033	2.236	3

Tabela 3 – Índices de redes complexas verificados para a rede semântica baseada *keywords* no período de **1972-1992**.

Rede Semântica	$n = V $	$m = E $	Qt. Componentes	Maior componente (%)	$\langle k \rangle$	Δ	C_{ws}	L	D
<i>Keywords</i>	171	626	2	99,42	7,322	0,043	0,686	2,763	6
Componente Gigante	170	626	1	100	7,365	0,044	0,686	2,763	6
Aleatória equivalente	170	647	1	100	7,365	0,044	0,051	2,7478	5

Tabela 4 – Índices de redes complexas verificados para a rede semântica baseada em *keywords* de **1992-1995**.

Rede Semântica	$n = V $	$m = E $	Qt. Componentes	Maior componente (%)	$\langle k \rangle$	Δ	C_{ws}	L	D
<i>Keywords</i>	97	276	4	87,63	5,691	0,059	0,743	3,036	6
Componente Gigante	85	255	1	100%	6	0,071	0,711	3,048	6
Aleatória equivalente	85	266	1	100%	6	0.0736	0.08205525	2.60532	5

Conforme as informações nas Tabelas 1 e 2 observamos que o valor do número de vértices que compõem as redes baseadas em *abstracts* (resumos) 1972-1992 (177 resumos) e no período de 1992-1995 (53 resumos) apresentou a seguinte variação: $1220 \leq n \leq 1564$. A pouca diferença de 344 vértices entre os dois períodos indica possivelmente que o vocabulário utilizado para a elaboração dos resumos apresentou um conjunto de palavras em comum em ambos os períodos. O número de arestas (conexões), cuja variação ocorreu foi a correspondente: $24086 \leq m \leq 31329$, ratifica o elevado número de arestas que ocorreram no período 1972-1992, com 7243 conexões a mais do que no período consequente.

Nas Tabelas 3 e 4, para as redes semânticas de palavras-chave a variação do número de vértices correspondeu a $85 \leq n \leq 170$ e para as arestas $255 \leq m \leq 626$. Para as palavras-chave o valor dos vértices em 1972-1992 (20 anos) corresponde ao dobro utilizado entre 1992-1995. Por mais que já fosse esperado que o primeiro período de análises a quantidade palavras-chave utilizada fosse maior, contudo, no último período analisado, em apenas três anos utilizou-se metade da quantidade de palavras-chave utilizadas em 20 anos (1972-1992). O número de arestas indica que o primeiro período analisado apresentou, aproximadamente, o triplo de conexões em relação ao período de 1992-1995.

Os valores calculados para o grau médio ($\langle k \rangle$), em ambas as redes baseadas em resumos, sugerem que, em média, um vértice (palavra), conectou-se a 40 outros vértices, aproximadamente. O intervalo verificado para a medida da densidade (Δ) correspondente foi: $0,026 \leq \Delta \leq 0,032$. Para a rede semântica de palavras-chave, os valores verificados para o grau médio ($6 \leq \langle k \rangle \leq 7,365$), indicam que, em média, uma palavra conectou-se a outras 7, aproximadamente. Nas redes baseadas em resumos, o intervalo verificado para a medida da densidade (Δ) correspondente foi: $0,026 \leq \Delta \leq 0,032$. Em relação às redes de palavras-chave foi $0,044 \leq \Delta \leq 0,071$. Estes intervalos verificados, indicam que as redes semânticas das Figura 2, Figura 3, Figura 4 e 5, são esparsas. Como este índice está relacionado ao nível de coesão dos vértices em uma rede semântica, para a rede baseada em resumos, argumentamos que, é uma tendência destes resumos e as palavras-chave se conectarem por meio de um grande número de palavras (PEREIRA et al. 2011; NASCIMENTO; PEREIRA; MORET, 2018).

Em relação aos caminhos mínimos médios (L) verificados nas redes semânticas baseados em resumos, sugerem aproximadamente que duas arestas são necessárias para conectarem duas palavras nas redes e por três palavras, nas redes semânticas baseadas em palavras-chave.

O diâmetro (D) calculado para as redes semânticas baseadas em resumos, obteve valor 4; para as redes de palavras-chave, 6. Estes resultados indicam o fato que as palavras pertencentes às sentenças de um resumo ou de sentenças de resumos distintos estão conectadas entre si, por meio de 4 outras palavras. Para as redes de palavras-chave o valor encontrado indica que as palavras pertencentes as palavras-chave de dissertações e teses distintas estão conectadas entre si, por meio de seis outras palavras (NASCIMENTO et al. 2018; NASCIMENTO; PEREIRA; MORET, 2018). Em relação ao Coeficiente de Aglomeração (C_{ws}), tanto nas redes baseadas em resumos quanto em palavras-chave, o valor encontrado foi da ordem de 75%. Mesmo com a exceção, a rede semântica da Figura 4 com 68%, ainda assim, esses valores indicam uma alta conectividade entre as palavras presentes nas redes semânticas de resumos e palavras-chave.

Diante da alta conectividade verificada nas redes, percebemos que há uma preferência nelas por determinados vértices. As distribuições de graus realizada para as redes semânticas da Figura 2, Figura 3, Figura 4 e Figura 5, indicaram seguir uma lei de potência do tipo $P(k) \sim k^{-\gamma}$, sugerindo que sejam redes *Scale Free* (livres de escala), conforme as pesquisas de Barabási e Albert (1999):

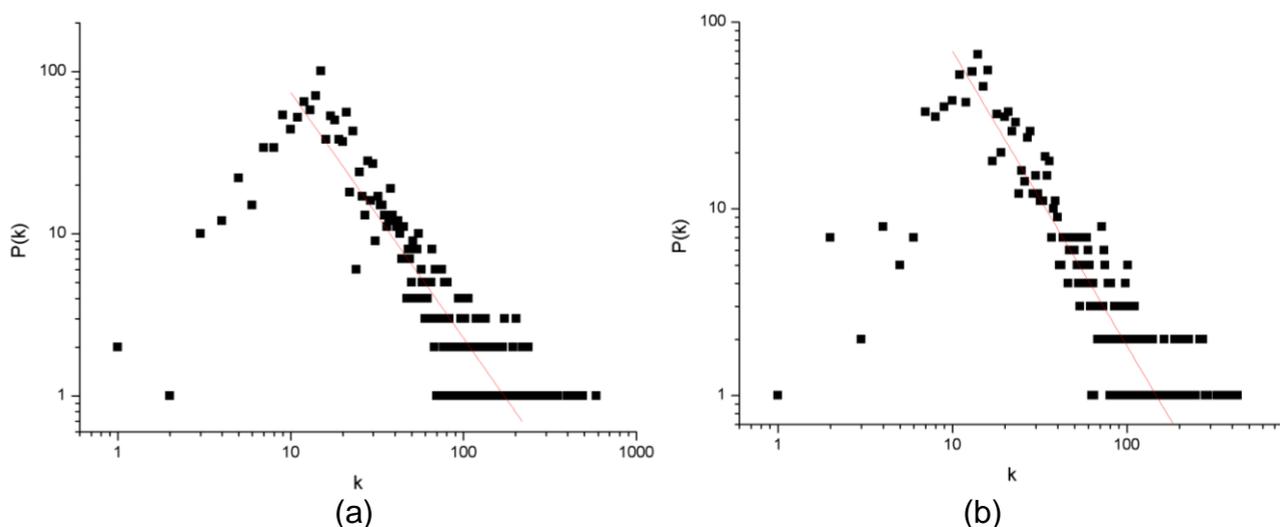


Figura 6 - Distribuição de graus das redes semânticas baseadas em resumos. Em (a): período de 1972-1992 (com $\gamma = 1,5147$ e o ajuste $R^2 = 0,82346$). Em (b): período de 1992-1995 (com $\gamma = 1,5787$ e o ajuste $R^2 = 0,869$).

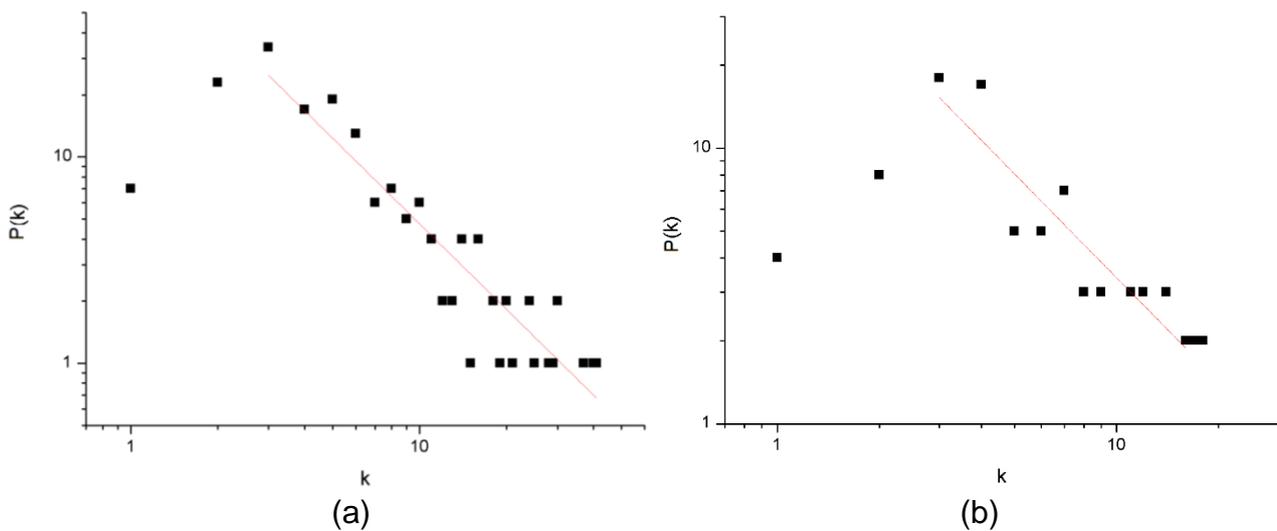


Figura 7 - Distribuição de graus das redes semânticas baseadas em palavras-chave. Em (a): período de 1972-1992 (com $\gamma = 1,58301$ e o ajuste $R^2 = 0,87372$). Em (b): período de 1992-1995 (com $\gamma = 1,25214$ e o ajuste $R^2 = 0,8045$).

Conforme Tabelas 1, 2, 3 e 4, ao compararmos as componentes gigantes ou maiores componentes das redes de *abstracts* e *keywords* com as suas respectivas redes aleatórias equivalentes, constatamos que ambas apresentaram o fenômeno *Small-World* (mundo pequeno). Para isto, utilizamos o método apresentado por Watts e Strogatz (1998). Devido aos valores dos coeficientes de aglomeração percebemos que havia indicativo de haver tendências, preferências por certos vértices/palavras que constituem as redes semânticas baseadas em resumos e palavras-chave: os *hubs*. Para corroborar este pensamento, realizamos a mineração dos vértices com maiores graus (centralidade de grau) nas redes semânticas. Para isso, propomos na presente pesquisa que, para ser um vértice altamente conectado (hub), a seguinte condição deve ser satisfeita: apresentar como grau k_i^{hub} valor maior ou igual a soma do grau médio com dobro do desvio padrão

k_{lim} :

$$k_i^{hub} \geq k_{lim} \quad (11)$$

Em que $k_{lim} = \langle k \rangle + 2 \sigma$. Desta forma, com as métricas de Teoria e Ciências das Redes, propomos que a condição acima pode ser rescrita da seguinte forma:

$$k_{lim} = \langle k \rangle + 2 \sqrt{\frac{\sum (k - \langle k \rangle)^2}{n-1}} \quad (12)$$

Em que: $\langle k \rangle$ é o grau médio da rede; k é o grau dos vértices e n é o número de vértices da rede.

Utilizando a Eq. 11 e Eq. 12, para as redes semânticas baseadas em resumos de 1972-1992 e 1992-1995, respectivamente, encontramos os seguintes valores: $k_i^{hub} \geq 153,938$ e $k_i^{hub} \geq 140,439$. Já em relação às redes baseadas em palavras-chave, os valores encontrados corresponderam a $k_i^{hub} \geq 22,274$ e $k_i^{hub} \geq 14,815$, respetivamente para os supracitados períodos. Com base nestas informações, realizamos a mineração dos 22 primeiros *hubs* das redes de resumos e 22 *vértices das redes* palavras-chave, conforme Tabela 5 e Tabela 6. Ressaltamos que para a rede baseada em palavras-chave, há 10 *hubs* para o primeiro período de análise e 6 *hubs* para o segundo:

Tabela 5 – Vértices e suas Centralidade de Grau verificados para a rede semântica baseadas em resumos. Fonte: Dos autores.

1972-1992						1992-1995					
Vértices		<i>k</i>	Vértices		<i>k</i>	Vértices		<i>k</i>	Vértices		<i>k</i>
1	Física	586	12	Conhecimento	328	1	Aluno	431	12	Através	262
2	Aluno	489	13	Conteúdo	310	2	Física	415	13	Ciência	262
3	Ensino	447	14	Utilizar	295	3	Ensino	407	14	Partir	248
4	Ser	438	15	Não	292	4	Concepção	365	15	Concluir	245
5	Apresentar	419	16	Analisar	286	5	Conhecimento	358	16	Escola	232
6	Professor	406	17	Estudo	283	6	Professor	349	17	Não	231
7	Ciência	399	18	Aspecto	277	7	Ser	328	18	Aprendizagem	226
8	Conceito	398	19	Segundo Grau	276	8	Analisar	292	19	Mudança	226
9	Proposta	350	20	Concepção	272	9	Conceito	281	20	Científico	225
10	Científico	345	21	Ensino de Física	264	10	Apresentar	272	21	Processo	224
11	Concluir	340	22	Atividade	260	11	Relação	272	22	Identificar	221

Tabela 6 – Vértices e suas Centralidade de Grau verificados para a rede semântica baseadas em Palavras-chave Fonte: Dos autores.

1972-1992						1992-1995					
Vértices		<i>k</i>	Vértices		<i>k</i>	Vértices		<i>k</i>	Vértices		<i>k</i>
1	Piaget	41	12	Livro Didático	20	1	Concepções Espontâneas	18	12	Calor	12
2	Filosofia da Ciência	40	13	Laboratório-atividades	20	2	Mecânica	18	13	Educação Dialógica	11
3	Concepções Espontâneas	37	14	Laboratório-análise	19	3	Mudança Conceitual	17	14	Gravitação	11
4	Currículo	30	15	Ciências de Primeiro Grau	18	4	Livro Didático	17	15	Evolução de Conceito	11
5	Ensino de Ciências	30	16	Instituição de Ensino-perfil	18	5	Aprendizagem de Conceitos	16	16	Mecânica Quântica	9
6	Evolução de Conceitos	29	17	Método Keller	16	6	Filosofia da Ciência	16	17	Paulo Freire	9
7	Projetos de	28	18	Paulo Freire	16	7	Currículo	14	18	Luz	9

	Ensino										
8	História da Ciência	25	19	Aprendizagem	16	8	História da Ciência	14	19	Museu de Ciências	8
9	Ensino Individualizado	24	20	Estrutura Conceitual	16	9	Ensino de Ciências	14	20	Temperatura	8
10	Mecânica	24	21	Termodinâmica	15	10	Historia da Ciência	12	21	Laboratório Atividades	8
11	Eletricidade	21	22	Desenvolvimento Cognitivo	14	11	Ensino Construtivista	12	22	Projetos de Ensino	7

Tabela 7 – Vértices com os maiores valores de Centralidade de Intermediação para o período de 1972-1992. Fonte: Dos autores.

Rede semântica de Palavras-chave					Rede semântica de Resumos						
Vértices	C.I.	Vértices	C.I.		Vértices	C.I.	Vértices	C.I.			
1	Filosofia da Ciência	2298.34	12	Eletricidade	733.08	1	Física	78613.60	12	Concluir	26572.59
2	Piaget	2168.21	13	Instituição de Ensino-perfil	730.02	2	Aluno	67132.47	13	Científico	26002.94
3	Currículo	2026.49	14	Livro Didático	701.82	3	Conceito	46904.41	14	Analisar	21772.16
4	Concepções Espontâneas	1959.37	15	Resolução de problemas	623.88	4	Ciência	45188.87	15	Atividade	19829.66
5	Ensino de Ciências	1201.30	16	História da Ciência	519.67	5	Ensino	44956.34	16	Concepção	19228.30
6	Projetos de Ensino	1051.86	17	Estrutura Conceitual	419.10	6	Ser	44389.39	17	Partir	18114.09
7	Laboratório-atividades	1048.62	18	Epistemologia	350.45	7	Apresentar	38250.25	18	Curso	18100.96
8	Mecânica	1024.01	19	Recursos Didáticos	344.90	8	Professor	33099.13	19	Aspecto	17580.23
9	Ensino Individualizado	765.30	20	Método Keller	317.50	9	Proposto	30009.03	20	Não	17391.31
10	Evolução de Conceitos	752.27	21	Aprendizagem	280.27	10	Utilizar	29978.53	21	Estudo	16918.52
11	Laboratório-análise	751.04	22	Ótica	264.95	11	Conhecimento	28300.50	22	Segundo Grau	16263.67

Tabela 8 – Vértices com os maiores valores de Centralidade de Intermediação para o período de 1992-1995. Fonte: Dos autores.

Rede semântica de Palavras-chave					Rede semântica de Resumos						
Vértices	C.I.	Vértices	C.I.		Vértices	C.I.	Vértices	C.I.			
1	Ensino de Ciências	574.39	12	Professor-Perfil	243.0	1	Física	42587.51	12	Através	14894.11
2	Livro Didático	540.88	13	História da Ciência	220.00	2	Aluno	40188.14	13	Relação	14701.05
3	Currículo	450.56	14	Ensino construtivista	203.96	3	Ensino	32515.45	14	Partir	13398.33
4	Projetos de Ensino	428.39	15	Mecânica quântica	200.47	4	Concepção	29308.52	15	Ideia	12801.34
5	Concepções Espontâneas	379.22	16	Paulo Freire	178.46	5	Conheciment o	28334.86	16	Escola	12724.68
6	Filosofia da	354.74	17	Ciências de	164.0	6	Professor	25561.82	17	Concluir	12484.

	Ciência			Primeiro Grau						19	
7	Mudança conceitual	348.69	18	Museu de Ciências	159.32	7	Ser	25109.27	18	Processo	11815.95
8	Mecânica	324.27	19	Eletromagnetismo	158.47	8	Ciência	19187.61	19	Identificar	10334.05
9	Aprendizagem de Conceitos	323.62	20	Gravitação	157.01	9	Conceito	16893.06	20	Científico	10180.97
10	Educação dialógica	285.09	21	Exposição científica	131.29	10	Analisar	16689.32	21	Ensino de Física	9965.18
11	Aperfeiçoamento de professores	245.0	22	Pesquisa e Ensino	127.57	11	Apresentar	15875.64	22	Não	9704.97

As Tabela de 5 e 6 apresentam cada uma delas os vinte e dois vértices elencados com maiores graus das redes, ou seja, as palavras que são mais importantes, devido ao fato de realizarem um elevado número de conexões, com os vértices imediatamente conectados a eles (centralidade de grau) (PEREIRA et al., 2011). Já nas Tabelas 7 e 8 são elencados os vértices com os maiores valores de centralidade de intermediação. Conforme Lopes et al. (2014) *Apud* Nascimento, Pereira e Moret (2018, p. 58-3 – 58-4) “[...] para que o valor desta centralidade seja elevado, em um vértice, o mesmo deve estar no caminho entre diversos outros nós”.

Assim, as informações retratam que, para a rede semântica baseada em resumos no período de 1972-1992, o vértice mais importante foi “física”. Os vértices “física”, “aluno” e “ensino” apresentaram os maiores valores de centralidade de grau, constituindo-se assim, as palavras mais conectadas nesta rede. Desta forma, há um indicativo nas redes semânticas da Figura 2, que as dissertações e teses objetivaram como foco principal, metodologias para o ensino de física voltadas para os discentes. Outros vértices que corroboram com este resultado são “analisar”, “proposta” e “estudos”. Seguindo este raciocínio, na Tabela 6, as palavras-chave relacionadas que correspondem a metodologias em Ensino de Física são: “ensino individualizado”, “método Keller”, “ensino de ciências” e “ensino Individualizado”.

Sejam metodologias para o Ensino de Física ou outras categorias investigadas no período analisado, há uma tendência de que o nível de ensino mais objetivado nos trabalhos, tenha correspondido à educação básica. O primeiro destaque nesse viés ocorre com o *hub* “Escola”, na rede semântica baseada em resumos. Outro fato a ser observado é o vértice “segundo-grau” como *hub* na rede baseada em resumos e na rede baseada em palavras-chave, que aparecem como referência ao nível de ensino. Especificando mais ainda, os *hubs* “ciências” na rede de resumos, “ensino de ciências”, na rede de

palavras-chave e o vértice “Ciências de Primeiro Grau” além de também sugerirem o foco das pesquisas na educação básica, ratificam que na própria educação básica, as dissertações e teses se concentraram em maior quantidade ao ensino fundamental (ou primeiro grau - nomenclatura utilizada nos períodos investigados).

Pelo exposto até aqui as análises das redes semânticas baseadas em resumos e palavras-chave sugerem que o foco principal das dissertações e teses correspondeu metodologias para o Ensino de Física. Quanto ao nível de ensino objetivaram a educação básica com o foco no ensino fundamental. Desta forma, pelas análises das redes semânticas, há indícios que as pesquisas em metodologias de ensino, para a educação básica, tenham ocorrido por atividades ligadas a laboratórios para o Ensino de Física no Ensino Fundamental. Pela observância dos vértices “Laboratório-atividades” e “Laboratório-análise”, há indícios na rede semântica da Figura 3 de indicadores de pesquisas realizadas por meio de atividades experimentais.

Outros fortes indícios que encontramos nas redes semânticas corresponde ao que chamaremos nesta pesquisa de “Psicologia Cognitiva”. Na rede semântica baseada em resumos da Figura 2 a presença dos *hubs* “conceito”, “concepção”; dos *hubs* presentes na rede baseada em Palavras-chave (Figura 3), “Piaget”, “concepções espontâneas” “Evolução de Conceitos” e dos vértices “Aprendizagem”, “Estrutura Conceitual”, “Desenvolvimento Cognitivo”; destacam o quão presente nas redes estão *vértices e/ou hubs* relacionados à Teoria da Aprendizagem Significativa de Ausubel e a Teoria de Piaget. Conforme descrito no parágrafo anterior nas redes semânticas de resumo (Figura 2) e a de palavras-chave (Figura 3) apresentam *hubs* com um forte destaque à temática de concepções espontâneas: “concepção” e “concepções espontâneas”.

De acordo com Moreira (2010) essas concepções correspondem a aprendizados incoerentes cientificamente em relação a um determinado fenômeno a ser estudado. Desta forma, percebemos então pelas redes estudadas que a temática de concepções espontâneas esteve associada, por vezes, à psicologia cognitiva da teoria de Jean Piaget. Por esta abrangência, assim como metodologia para o Ensino de Física é o primeiro destaque das análises realizadas, percebemos então, que a temática de concepções espontâneas corresponde ao segundo maior destaque do Ensino de Física no período analisado: 1972-1992. Enfatizamos também que a temática de currículo se fez presente nas redes semânticas, com o *hub* “currículo”. Assim como “currículo” esteve presente de forma discreta nas redes, a temática que se refere a formação do professor de física,

também. Os *hubs* “Magistério”, “formação” e o vértice “licenciatura” na rede de resumos sugerem este pensamento.

Conforme as redes semânticas de resumos (Figura 3) e palavras-chave (Figura 5) referentes ao período de 1992-1995, de forma semelhante ao período de 1972-1992, há fortes indícios de que tenham se destacado a temática, nas dissertações e teses objetivando metodologias para o Ensino de Física. Conforme as Tabelas 5, 6 e 8 os vértices “Aluno”, “Física” e “Ensino” corroboram estes resultados. Na rede de palavras-chave (Figura 5), cujos vértices estão elencados na Tabela 08, conforme os valores das medidas de Centralidade de Intermediação, temos os vértices “projetos de ensino” (C.I. = 428.39) e “pesquisa e ensino” (C.I. = 127.57), que também indicam a presença do referido tema.

Pelas análises realizadas na rede semântica de palavras-chave da Figura 5 e as informações nas Tabelas 6 e 8, em observância também aos vértices “evolução de conceitos”, “aprendizagem de conceitos” e “ensino construtivista”, por ambas as centralidades, corroboram com a marcante presença da temática de Psicologia Cognitiva e Concepções Espontâneas no período de 1992-1995. Os vértices “representações mentais”, “conceituação”, “concepção”, “aprendizagem”, “conceito”, “processo”, “relação” e “Identificar” presentes na rede baseada em resumos (Figura 3), conforme os dados da Tabela 6 e 8, ratificam a presença das supracitadas temáticas. Pela medida de centralidade de intermediação, os vértices relacionados ao tema de concepções espontâneas estão correlacionados com a temática de psicologia cognitiva.

Em termos do nível de ensino, a educação básica continuou apresentando forte presença e preferência, superando a do ensino superior nas redes semânticas de palavras-chave e resumos. Vértices que apresentam importâncias pelas conexões com seus vizinhos e pela intermediação entre outros vértices, como “História” (com acento), “Historia” (sem acento) e “Curso”, estão relacionados às pesquisas voltadas para o ensino superior, nas redes semânticas estudadas. O vértice “educação dialógica” (Tabelas 6 e 8) apresenta na rede semântica da Figura 5, centralidade de intermediação em temáticas de pesquisas relacionadas a Paulo Freire com o ensino fundamental, médio e superior. Fato este que se repete com o vértice “Educação dialógica” na rede semântica baseada em palavras-chave (Figura 5).

Nas Tabelas 6 e 8 os vértices “ensino de ciências”, “ciências de primeiro grau” e “museu de ciências” corresponderam em ambas as medidas de centralidade, para o Ensino de Física na educação básica, especificamente o ensino fundamental. Na rede

baseada em palavras-chave os vértices que ratificam tais pesquisas, por ambas as medidas de centralidade são “ciência” e “escola”. Semelhante ao que ocorreu no período anterior, pesquisas relacionadas a laboratórios de ciências estão fortemente correlacionadas nas redes semânticas com o ensino fundamental. Os vértices “História” “História da Ciência”, “Filosofia da Ciência” (Tabela 6 e 8) indicam a presença da temática de História e Filosofia da Ciência no presente período de análise.

No período de 1992-1995, as redes semânticas apresentaram em destaques, vértices que correspondem a assuntos pertencentes ao Ensino de Física, são eles: “calor”, “mecânica”, “gravitação”, “mecânica quântica”, “luz” e “eletricidade”. É válido ressaltar que apenas o vértice “mecânica” pertencente à rede semântica baseada em palavras-chave, possui centralidade de grau correspondente a um *hub*. Os vértices “Calor” e “Temperatura” estão presentes nas redes, correlacionados com as temáticas de psicologia cognitiva e concepções alternativas.

O vértice “Mecânica Quântica” apresenta na rede semântica de palavras-chave correlação com a educação básica, especificamente ao nível médio e, também com a temática de Currículo, mesmo esta última não se apresentando de forma expressiva no período analisado. O vértice “Gravitação” apresentou correlações nas redes com o vértice “Mecânica” e com as temáticas de História e Filosofia da Ciência e Educação Dialógica (Paulo Freire). O vértice “Luz” apresentou correlação com as temáticas de História e Filosofia da Ciência, Psicologia Cognitiva e Educação Dialógica (Paulo Freire). O vértice “eletricidade” está correlacionado com as temáticas de Concepções Espontâneas, História e Filosofia da Ciência e Educação Dialógica (Paulo Freire). Por fim, ressaltamos que os vértices “mecânica”, “eletricidade” e “termodinâmica” realizaram uma interseção entre os períodos de 1972-1992 e 1992-1995, apresentaram-se em destaques em ambos.

4. CONSIDERAÇÕES FINAIS

Por meio da metodologia apresentada e/ou sugerida para a construção das redes semânticas baseadas em *abstracts* (resumos) e *keywords* (palavras-chave) na presente pesquisa, foi possível alcançar resultados expressivos para o Ensino Física Brasileiro. Com o devido mapeamento dos *hubs* nas redes semânticas foi possível indicar as principais temáticas em que o Ensino de Física Nacional se alicerçou, tanto no período de 1972-1992 quanto para o de 1992-1995.

Em concomitância com as análises dos *hubs* das redes, os vértices que apresentaram elevadas conexões, em que os valores dos graus médios deles estavam próximos aos dos determinados pelas Eq. 11 e Eq.12 (porém, não constituindo *hubs*) e, os vértices com maiores valores de centralidade de intermediação; constituíram nortes fundamentais para verificarmos os supracitados alicerces do Ensino de Física Brasileiro.

Assim, para o período de 1972-1992, apresentaram-se como foco nas redes semânticas, as temáticas de metodologias para o Ensino de Física voltada para os discentes do Ensino Fundamental. Destacaram-se nas redes também, a Teoria da Aprendizagem Significativa de David Ausubel e a Teoria de Jean Piaget, sendo que esta última esteve fortemente correlacionada com pesquisas realizadas na temática de concepções espontâneas. Por fim, de forma discreta, as temáticas de currículo e de formação de professores também corresponderam aos destaques do Ensino de Física para o referido período de análise das redes semânticas.

Para o período de 1992-1995, metodologias direcionadas ao Ensino de Física continuaram com um forte destaque nas pesquisas realizadas, assim como as temáticas de psicologia cognitiva e concepções espontâneas. A educação básica continuou em maior evidência e preferência nas redes semânticas, quando comparadas ao ensino superior. Verificamos também nas redes, pesquisas relacionadas com a Educação Dialógica (Paulo Freire) para ambos os níveis de ensino, bem como, a temática de História e Filosofia da Ciência. Por fim, em termos de assuntos de Física, mecânica, eletricidade e termodinâmica apresentaram maior preferência nas redes semânticas em ambos os períodos analisados.

Conforme o método apresentado por Watts e Strogatz (1998), realizamos uma comparação nas redes semânticas de *abstracts* e *keywords* entre os seus coeficientes de aglomeração e o caminho mínimo médio, entre as componentes gigantes delas e as suas respectivas redes aleatórias equivalentes. Desta forma, podemos argumentar que as redes semânticas, em ambos os períodos, apresentaram o fenômeno *Small-World*. Conforme Barabási e Albert (1999), também realizamos a distribuição de graus das redes semânticas e encontramos, para cada uma das redes, um comportamento em que há indícios de seguirem uma lei de potência do tipo $P(k) \sim k^{-\gamma}$. Então, as redes também apresentaram características de serem *Scale Free*.

Os valores encontrados para o coeficiente angular γ , da lei de potência, estão fora do intervalo indicado por Barabási e Albert (1999), porém, apresentam coerência com as redes *Scale Free* encontradas nas obras de Seyed-allaei et al. (2006), Fadigas et al.

(2009), Nascimento et al. (2018) e Nascimento, Pereira e Moret (2018). Diante das evidências apresentadas, podemos argumentar que as redes são híbridas: redes *Scale Free* e *Small-World*. Por fim, pelas características das redes, tanto pelos índices de redes complexas quanto os de redes sociais, podemos inferir que houve a difusão da informação do Ensino de Física nos períodos analisados.

AGRADECIMENTOS

Jefferson Nascimento agradece à FAPESB pelo suporte financeiro parcial devido à bolsa de doutorado (BOL170/2015) e Marcelo A. Moret agradece ao CNPq pelo suporte financeiro parcial oriundo de sua bolsa de Produtividade em Pesquisa (No. 304454/2014-1).

REFERÊNCIAS

- BARABÁSI, A. L. e ALBERT, R. Emergence of Scaling in Random Networks. **Science**, n. 286, pp. 509-512, 1999.
- CALDEIRA, S. M. G. **Caracterização da Rede de Signos Linguísticos: Um modelo baseado no aparelho psíquico de Freud**. 130f. Dissertação (Mestrado Interdisciplinar em Modelagem Computacional) - Centro de Pós graduação e Pesquisa da Fundação Visconde de Cairu. Fundação Visconde de Cairu, Salvador, nov. 2005.
- COSTA, C. C. S.; NASCIMENTO, J. O. do; MORET, M. A.; PEREIRA, H. B. B. **Um modelo computacional para construção de redes de colaboração científica em Física**, p. 218-221 . In: . São Paulo: Blucher, 2017. ISSN 2358-2359, DOI 10.5151/phypro-viii-efa-46.
- CUNHA, M.V. **Redes semânticas baseadas em títulos de artigos científicos**. 127f. Dissertação (Mestrado em modelagem computacional e tecnologia industrial) – CIMATEC, Salvador, 27 nov. 2013.
- 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.
- 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.
- GROSS, J. L.; YELLEN, J. **Graph theory and its applications**. CRC press, Boca Raton FL USA 2005.
- MOREIRA, Marco Antônio. **O que é, afinal, aprendizagem significativa**. Material de apoio aula inaugural do Programa de Pós-Graduação em Ensino de Ciências Naturais. UFMG, Cuiabá, MT, 2010. Disponível em <<http://www.if.ufrgs.br/~moreira/oqueefinal.pdf>>. Acesso em: 19 mar. 2014.

NASCIMENTO, J. O. do.; PEREIRA, H. B. B.; MORET, M. A.; **Redes semânticas baseadas em palavras-chave do Ensino de Física Brasileiro**: uma comparação nos métodos de pré-processamento dos dados, p. 122-127. In: . São Paulo: Blucher, 2017. ISSN 2358-2359, DOI 10.5151/phypro-viii-efa-28

NASCIMENTO, J. O. do; MONTEIRO, R. L. S.; MOREIRA, D. M.; MORET, M. A.; PEREIRA, H. B. B. Semantic networks of keywords from Brazilian dissertations on physics teaching. **Discontinuity, Nonlinearity and Complexity**, v. 7, p. 173-183, 2018.

NASCIMENTO, J. O do; PEREIRA, H.B.B.; MORET, M. A .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**, v. 3, p. 58-1-58-15, 2018.

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.

NASCIMENTO, J. O. do; 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**, v. 10, p. 315-339, 2018.

ROSA, M. G.; FADIGAS, I. S.; 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.

UNIVERSIDADE DE SÃO PAULO, Instituto de Física. **Ensino de Física no Brasil: Catálogo Analítico de Dissertações e teses: (1972-1992)**. São Paulo, 1992.

UNIVERSIDADE DE SÃO PAULO, Instituto de Física. **Ensino de Física no Brasil: Catálogo Analítico de Dissertações e teses: (1992-1995)**. São Paulo, 1996.

SEYED-ALLAEI, H., Bianconi & M. Marsili, G. Scale-free networks with an exponent less than two. **Phys. Rev. E**. 73, 046113, 2006.

STEYVERS, M.; TENENBAUM, J. B. The Large-Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth. **Cognitive Science**, 29: 41-78, 2005.

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