INPE – National Institute for Space Research
São José dos Campos – SP – Brazil – May 16-20, 2016

The postgraduate Brazilian studies in Physics Teaching using Complex Network

Jefferson Oliveira do Nascimento[1], *Camila de Sousa Pereira-Guizzo*[2], *Roberto Luiz Souza Monteiro*[3], *Davidson Martins Moreira*[4], *Marcelo Albano Moret Simões Gonçalves*[2], *Hernane Borges de Barros Pereira*[6].

[1]Programa de Modelagem Computational - SENAI CIMATEC, Salvador, Bahia, Brazil, jeffersonascimento@gmail.com
[2]Programa de Modelagem Computational - SENAI CIMATEC, Salvador, Bahia, Brazil, camila.pereira@fieb.org.br
[3]Programa de Modelagem Computational - SENAI CIMATEC, Salvador, Bahia, Brazil, roberto@souzamonteiro.com
[4]Programa de Modelagem Computational - SENAI CIMATEC, Salvador, Bahia, Brazil, davidson.moreira@gmail.com
[5]Programa de Modelagem Computational - SENAI CIMATEC, Salvador, Bahia, Brazil, mamoret@gmail.com
[6]Programa de Modelagem Computational - SENAI CIMATEC, Salvador, Bahia, Brazil, hernanebbpereira@gmail.com.

Abstract: In this paper we study the topological structure of a semantic network based on keywords from thesis and dissertations on Brazil's Physics Teaching in the decade of 1996-2006. The results obtained were through some indices of social and complex networks. We analyzed the vertices with higher degree of centrality (hubs) we have inferred that the work carried out showed a pattern in terms of chosen keywords, indicating a preference for words associated with the training of physics teachers. The found topology of studied semantic networks are small-world and scale-free.

keywords: Complex Networks, Control in Complex Systems, Modeling, Numerical Simulation and Optimization, Social Network, Physics Teaching.

1. INTRODUCTION

The network theory, as a possibility to study connected entities, contains the graph theory, the nuances of social networks and complex networks, which enables the characterization and investigation of the behavior in question, through of emergent properties. In this paper we studied a semantic network based on keywords dissertations, theses and postdoctoral reports in the postgraduate Brazilian studies in Physics Teaching, in the years between 1996 and 2006. This paper is organized in four sections. In the second section, we present the proposal of this paper and the construction techniques and study of the semantic network. In the third section the results and discussion of the network created from keywords. The fourth and final section presents the conclusions of this paper.

2. PURPOSE AND METHODS

Este artigo tem como finalidade realizar a construção de uma rede semântica das palavras-chave pertencentes as dissertações, teses e relatórios de pós-doutorado especificamente na temática de Ensino de Física, entre os anos de 1996 a 2006. Também, verificar as propriedades emergentes nas redes e caracterizá-las conforme as suas topologias.

To create a semantic network, we use the graph theory. A graph G = (V, E) is a mathematical and consistent structure in two settings: V (finite and not empty) and E (binary relations on V) [1]. The elements of V are called vertices and the elements of E are the edges. The first set (V) is finite and not empty and the second (E) corresponds to the possible binary interactions of V [1].

The properties that must be observed in the analysis of semantic networks proposed in this paper correspond to those related to basic statistical indexes of the theory of complex networks. With this information, it is possible to characterize the topology of a network, such as random networks, small-world and scale free. Therefore, to characterize the proposed semantic network, it must be observed the following properties: number of vertices, number of edges, average degree, density, average clustering coefficient and average minimal path length:

Number of vertices (n): Total number of vertices or cardinality of set V:

$$n = |V| \tag{1}$$

Number of edges (m): Total number of edges or cardinality of set E:

$$m = |E| \tag{2}$$

Average degree ($\langle k \rangle$): The average degree of an undirected network is the average of k_i

$$\langle k \rangle = \frac{1}{n}\sum_i^n k_i \tag{3}$$

Density (Δ): Is the total of existing edges, divided by the maximum possible number of edges:

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

Average clustering coefficient (C_{ws}): Mean clustering coefficient of the vertices of the network:

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

The clustering coefficient of a vertex v, denoted by C_v, measures the proportion of existing edges between neighbors of vertex *v*, E_v and the maximum possible number of edges:

$$C_V = \frac{2E_v}{k_v(k_v-1)} \quad (6)$$

Average minimal path length (L):

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

dij is the geodesic distance, in terms of number of edges, between the vertices *i* and *j*.

Diâmetro (D): It is the largest geodesic distance, between the vertices *i* and *j*:

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

In order to build up the semantic network based on keywords, perform a manual treatment on the keywords, as a second procedure, a treatment by means of software. For manual processing, we have the titles in a text file, where each line received the keyword of each Brazilian postgraduate studies in Physics Teaching. After this phase, we conducted a formatting obeying the rules established by Pereira et al. [2]

For the treatment of words by means of computational resources for the construction of the semantic network, the procedures were used, as described in [2,3-4]. It notes that the ambiguities are eliminated following the algorithm developed in [5].

3. RESULTS AND DISCUSSION

Figure 1 depicts the resulting semantic network based on keywords of the Brazilian studies in Physics Teaching:

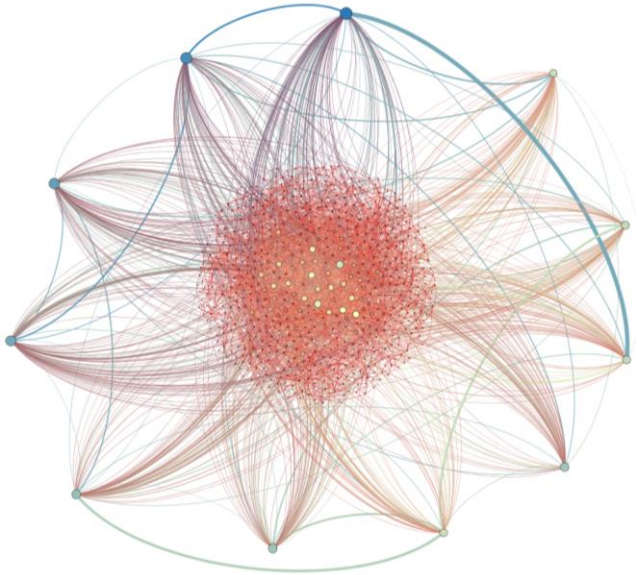

Figure 1 – Semantic network based on keywords

Table 1 shows the properties obtained for the semantic network based on keywords (Figure 1):

Table 1 - Properties of the semantic network based on keyword

Index	Value	Index	Value
$n = \|V\|$	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

Table 1 shows the value of average minimal path length is three edges to connect two words in the network. The value of the diameter (D) is six for the semantic network. This information indicates which words belong the keywords of the Postgraduate Brazilian studies are connected to each other at most by six other words.

The density (Δ) is a measure that is related to the level of connectivity within the network or indicative of the degree of cohesion of the vertices [3]. In Pereira et al. [2] this information in a given semantic network of titles of papers, indicating the amount of connections between words, i.e., "a tendency of the titles to link through a large number of words". It also may be applied to semantic network based on keywords. The found value it was of 1.3%. Thus it is possible that the displayed semantic network, there are a large number of words connecting cliques, indicating low density. Thus, the found density may correspond to less cohesion between the connections in the network.

The value found for the clustering coefficient (C_{ws}) was approximately 75%, indicating a high connectivity between words present the semantic network. We realize that there is a preference in it for some vertices (hubs) (Table 2):

Table 2 – Vertices with high degree of connection

Vertices (V)	Degree of vertex (k)
History of Science	102
Textbook	97
Experimental activity	89
Language	85
Continuing Teacher training	76
Teacher training	74
Didactics proposal	74
Philosophy of science	67
Teacher Knowledge	66
Mechanics	64

Table 2 shows the top ten vertices with higher degrees in the semantic network. This allows us to infer the importance (degree centrality) that a particular node has, through its connections with the nodes connected to it immediately (neighboring vertices) [3]. The information portray that to the keyword network, the vertex with greater centrality degree was "history of science". The words "textbook", "experimental activity" and "language", follow as the second, third and fourth vertices, respectively, in terms of importance in the semantic network. Thus, Table 3 shows the 10 vertices selected by the respective degrees of numbers that may represent important issues related to physical contents.

Table 3 – Some vertices with its degrees of connections

Vertices (V)	Degree of vertex (k)
Mechanics	64
Modern physics and contemporary	53
Astronomy	47
Electromagnetism	44
Optical	33
Universal Gravitation of Newton	26
Thermodynamics	25
Quantum mechanics	14
Special Relativity of Einstein	12
Cosmology	9

The vertices contained in Table 3 indicate that subjects of importance to the teaching of physical discipline were just has chosen as keywords. The vertex "Mechanics" (k = 64) is the first subject that arises in the semantic network and displays degree of connection k = 64, i.e., they have a low degree centrality compared to the first nine vertices contained in Table 2. This information may indicate that new postgraduate Brazilian studies in Physics Teaching may have, low probability of occurring with important subjects in Physics Teaching.

For the characterization of the semantic network topology, the degree distribution was carried out, is displayed in Figure 2. We perform a linear fit and realized the pattern observed follows a power law of the form $P(k) \sim k^{-\gamma}$:

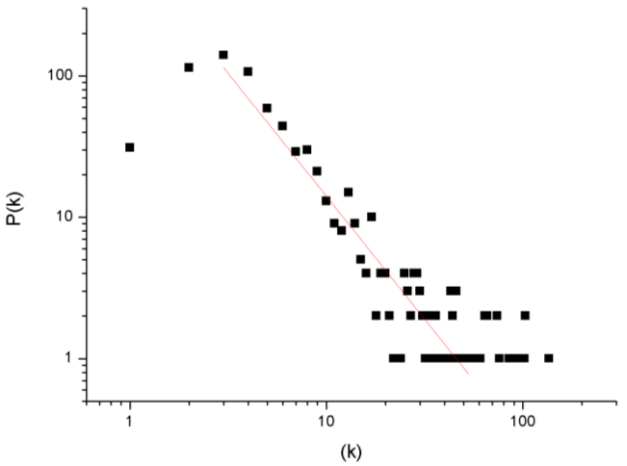

Figure 2 - Degree distributions of the semantic networks based on keywords (with γ = 1,73784 and R^2 = 0,85309)

According to the graph in Figure 2, the distribution of degrees, gave us indications that the semantic network can present a behavior of a free scale network. According to Barabási [6], it is clear then that possibly there was for the decade investigated a preferential adherence by some vertices in the network (i.e., words), it is indicating the presence of hubs. Thus, the vertices highly connected have a greater chance to acquire new connections than those who are less connected. Table 2 shows a comparison that we did between the semantic network of keywords and it equivalent random network (network with the same number of vertex and average degree):

Table 2 – The semantic network of keywords and it equivalent random network

Network	Index	Value	Index	Value
Keywords	$n = \|V\|$	669	The major component. (%)	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	The major component. (%)	100
	$m = \|E\|$	3002	$\langle k \rangle$	8,8
	Δ	0,013	C_{WS}	0,012
	L	3,02	D	6

We found that the networks had the same approximate value for the average minimal path length (L). We also note that the diameter (D) of the semantic network and the equivalent network are the same. The clustering coefficients (Cws) of semantic network showed a high value compared to its equivalent random network. So before these results observed in comparison, it is in accordance with the method proposed by Watts and

Strogatz [7] and we realized that the semantic network (Figure 1) shows the Small-World phenomenon.

According to Pereira et al. [2], Caldeira et al. [4] Santos and Junior et al. [8] the data of Table 2 also indicators for the pattern have seen in Figure 2: a semantic network characterized as scale free.

4. CONCLUSIONS

Through the analysis of the semantic network based on keywords of studies in Brazilian Physical Education, through its indices of the theory of complex networks and social networks, it was possible to characterize the topology of the semantic network built. It was also possible to study which words had higher degrees to the Brazilian Physics Teaching in the decade 1996-2006.

The properties calculated and indicated in this paper about the semantic network studded it shows the small-world phenomenon and a scale free behavior. Both features are possible to the same network, given that one does not exclude the other. Due to the characteristics such as growth and the preferential adhesion of the semantic network built and studied, there is more likely to occur new keywords with the words in Table 2 of the Table 3 on new postgraduate Brazilian studies in Physics Teaching.

The high connectivity in networks evidenced by clustering coefficient, the average minimal path length with the same values of the equivalent random networks and a degree distribution indicating a power law also indicates this characterization. Thus it is resistant the withdrawal of vertices occurring randomly but it presents vulnerability in the event of withdrawal of words with high degree of centrality. So if get up of the semantic network (Figure 1) the vertices belonging to Table 2, possibly present would be divided into a larger number of components which could hinder the spread of certain information. According to Fadigas et al. [3] This difficulty may occur in terms of not favor readers, search engines of the Brazilians studies in physics teaching by their keywords.

To be observed vertices with higher centralities (hubs), it is possible to suggest that the semantic network based on keywords aimed at training of physics teacher. So the minorities analyzed research explored methods for focused teaching education for students. However, these statements require more detailed research, which was not objective in this article.

ACKNOWLEDGMENTS

JON thanks to FAPESB (Fundação de Amparo à Pesquisa do Estado da Bahia) for partial financial support due to the Ph.D. scholarship (BOL170/2015) and DMM and MAM thank CNPq (Brazilian federal grant agency) for partial financial support derived from its Research Productivity (N. 301591/2009-1) and (N. 304454/2014-1).

References

[1] J. L. Gross & j. Yellen, Graph theory and its applications. CRC press, Boca Raton FL USA 2005.

[2] H. B. B. Pereira, I. S. Fadigas, R. L. S. Monteiro, A. J. A. Cordeiro & M. A. Moret. Semantic networks based on titles of scientific papers. Physica A **390**, 1192 (2011).

[3] I. S. Fadigas, T. H. P. Casas, V. Senna, M. A. Moret & H. B. B. Análise de redes semânticas baseada em títulos de artigos de periódicos científicos: o caso dos periódicos de divulgação em educação matemática Pereira, Educação Matemática Pesquisa, **11**, 167 (2009).

[4] S. M. G. CALDEIRA, T. Petit Lobão, R. F. S. Andrade, A. Neme & J. G. V. Miranda. The network of concepts in written texts. Eur. Phys. J. B 49, 523 (2006).

[5] CALDEIRA, S. M. G. (2005). Caracterização da Rede de Signos Lingüísticos: Um modelo baseado no aparelho psíquico de Freud. Dissertação de mestrado em Modelagem Computacional, Centro de Pósgraduação e Pesquisa da Fundação Visconde de Cairu. Salvador, 2005.

[6] A. L. Barabási. Scale-Free Networks: A Decade and Beyond. Science **325**, 412 (2009).

[7] D. J. Watts & S. H. Strogatz. Collective dynamics of 'small-world' networks. 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. Análise de diferentes conceitos de educação por meio das redes semânticas. DataGramaZero **15**(3), 3 (2014).