

Incidência Fidelidade aplicada a rede semântica de títulos

Marcelo do Vale Cunha^{1,4}, José Garcia Vivas Miranda², Hernane Borges de Barros Pereira^{1,3}

¹Programa de Modelagem Computacional, SENAI-Cimatec, Salvador BA – Brasil

²Universidade Federal da Bahia (UFBA), Salvador, BA – Brasil

³Universidade Estadual da Bahia (UNEB), Salvador, BA – Brasil

⁴Instituto Federal da Bahia (IFBA), Barreiras, BA – Brasil

{celaocunha, vivasm, hbbpereira}@gmail.com

Abstract. *The archived publications in scientific journals of high impact scientific configure the memory of mankind of a period of time. Therefore, it is necessary studies to investigate the relationships between publications, authors and the editorial line of journals, from the printed content in your articles. This work seeks to distinguish periodic from semantic networks of their titles, using an inherent semantic networks for memory access mechanism: a critical network. The method is based on the construction of networks from the incidence fidelity index. The results show the presence of the phenomenon in networks titles and suggest ways to differentiate journals from the indices of networks in critical networks.*

Resumo. *As publicações arquivadas em revistas científicas de alto impacto configuram a memória científica da humanidade de um referido período de tempo. Diante disso, faz-se necessário estudos que investiguem as relações entre as publicações, os autores e a linha editorial dos periódicos, a partir dos conteúdos impressos em seus artigos. Este trabalho busca diferenciar periódicos a partir das redes semânticas de seus títulos (RST), utilizando um mecanismo inerente à redes semânticas para acesso à memória: a rede crítica. O método se baseia na construção das redes a partir do índice incidência-fidelidade. Os resultados mostram a presença do fenômeno em redes de títulos e sugerem maneiras de diferenciar periódicos a partir dos índices de redes nas redes críticas.*

1. Introdução

O conjunto de trabalhos que compõe uma revista de publicação científica faz parte de um sistema formal de comunicação: o periódico científico. Este sistema expressa – em palavras, diagramas, imagens e equações – o conhecimento de atividades de pesquisa, não só para contribuir com o avanço científico da humanidade, mas também para reforçar os laços de comunicação entre cientistas e da ciência com a sociedade em geral.

Diversos autores vêm estudando o periódico científico, a partir da perspectiva da colaboração científica entre os indivíduos que lêem, publicam e avaliam os trabalhos das revistas científicas [Garvey 1979] Ziman 1979 [Miranda and Pereira 1996] [Vanz and Stumpf 2010]. Outros estudos focam nas redes sociais e semânticas que este sistema fomenta [Fadigas et al. 2009, Pereira et al. 2011, Fadigas et al. 2013, Cunha et al. 2013].

A técnica da modelagem de um texto através de redes semânticas oferece um meio empírico de acesso à organização mental do conhecimento de quem proferiu o discurso [Albuquerque and Pimentel 2004, Sternberg 2011]. Teixeira et al. 2010 propuseram um método para determinar a rede semântica que melhor representa o acesso à memória declarativa de um indivíduo, baseado em um indicador denominado *incidência-fidelidade IF*. Este índice pondera as arestas de uma rede semântica, a partir da frequência de aparição de cada aresta (par de palavras) nas cliques das redes (sentenças).

No contexto desta pesquisa, a rede social formada tem como relacionamento entre seus atores o vocabulário comum entre os autores utilizado para compor os títulos das publicações. Fadigas et al. 2009, Pereira et al. 2011 propuseram esta modelagem para comparar periódicos. Cunha et al. 2013 mostram como evoluem os índices de redes ao longo do tempo, para *RST* do periódico Nature. Segundo Fadigas et al. 2013, o uso da teoria de redes sociais em redes de títulos contribui de forma relevante para mapear a colaboração entre pesquisadores pertencentes a uma mesma comunidade científica.

Para a construção das redes neste trabalho, utiliza-se a mesma premissa dos trabalhos supracitados: as palavras de um título são vértices de uma clique¹. Cada clique conecta-se com outra se houver um vértice comum. Apresenta-se aqui a aplicação do índice *incidência-fidelidade* nas *RST* de periódicos de alto impacto, a fim de verificar o fenômeno da *rede crítica* nas redes que constituem a base de dados de Pereira et al. 2011. De acordo com Teixeira et al. 2010, uma rede semântica de cliques na configuração *crítica* possui o máximo de informação com mínimo de resíduo.

Sendo assim, nesta pesquisa propõe-se a aplicação do filtro *IF* nas redes semânticas dos periódicos, para compará-los a partir de seus índices de rede. Algumas interpretações são sugeridas para classificação de periódicos partir das *RST* de seus artigos, na configuração de máxima informação, i.e. *rede crítica*.

2. Incidência-fidelidade

O índice *incidência-fidelidade* (*IF*) é composto por dois conceitos: (a) o de *incidência* *I*, oriundo do trabalho de Nelson et al. 1998, que mede a probabilidade de ocorrência de um par de palavras nas sentenças de um texto e (b) o de *fidelidade* *F*, proposto por Teixeira et al. 2010, que mede a probabilidade de um par de palavras ocorrer juntas em uma mesma sentença, dentro do universo de sentenças em que elas foram evocadas, ou seja, dentro do conjunto de sentenças que contém pelo menos uma destas palavras.

A *incidência-fidelidade* $IF_{(\Psi, \Omega)}$, para um par de palavras Ψ e Ω , é dada pelo produto dos Índices $I_{(\Psi, \Omega)}$ e $F_{(\Psi, \Omega)}$ (Equação 1). Para cada par de palavras, seus valores atuam como peso nas arestas das redes de cada texto.

$$IF_{(\Psi, \Omega)} = I_{(\Psi, \Omega)} \cdot F_{(\Psi, \Omega)} = \frac{S_p}{N_S} \times \frac{S_p}{S_\Psi + S_\Omega - S_p} = \frac{(S_p)^2}{N_S(S_\Psi + S_\Omega - S_p)} \quad (1)$$

O termo N_S na Equação 1 representa o número de sentenças do texto, o que sugere que o índice *IF* depende do tamanho do texto. Para as redes de Teixeira et al. 2010 este

¹Clique é um conjunto de n vértices mutuamente conectados. O elemento básico de uma rede de cliques não é o vértice e sim a clique [Fadigas and Pereira 2013].

fato não comprometeu os resultados, visto que os textos analisados diferem muito pouco entre si do número total de sentenças.

Para minimizar o efeito do número de sentenças, Aguiar 2009 propôs um ajuste no índice, Equação 2. Na Equação, I_{Max} e I_{Min} representam respectivamente o maior e o menor valor de *incidência* que ocorreu no texto. Da mesma forma, F_{Max} e F_{Min} representam respectivamente o maior e o menor valor de *fidelidade* ocorridos no referido texto.

$$IF_{N(\Psi,\Omega)} = (I_N) \times (F_N) = \frac{I_{(\Psi,\Omega)} - I_{Min}}{I_{Max} - I_{Min}} \times \frac{F_{(\Psi,\Omega)} - F_{Min}}{I_{Max} - I_{Min}} \quad (2)$$

O índice *incidência-fidelidade* atua como um filtro que elimina informação da rede original (i.e. pares de palavras com IF mais baixos que um valor limite IF_L). O IF proposto por Teixeira et al. 2010 é mais adequado para filtrar e comparar textos quando não existe diferença considerável de tamanhos entre eles (quantidade de sentenças). O IF proposto por Aguiar 2009, por sua vez, é mais adequado para comparar textos que possuem uma diferença considerável de tamanho.

3. Metodologia

3.1. Dados da pesquisa

Os títulos dos artigos científicos foram obtidos em periódicos de circulação internacional (idioma em inglês), das mais abrangentes áreas do conhecimento. A Tabela 1 mostra informações básicas sobre os periódicos usados na pesquisa ². O propósito deste trabalho é diferenciar estes periódicos a partir de suas redes críticas, utilizando os dois índices IF supracitados.

<i>Periódico</i>	<i>Nome nesta Pesquisa</i>	<i>Frequência de Publicações</i>	<i>Período coletado</i>	<i>Número de títulos</i>
<i>Agricultural Forest Entomology</i>	AFE	Anual	1999 a 2008	371
<i>Antipode: A Radical Journal of Geography</i>	ARJG	5 por ano	1969 a 2008	971
<i>Applied Psycholinguistics</i>	APPL	Trimestral	1980 a 2009	658
<i>Chemistry & Biology</i>	CB	Mensal	1994 a 2008	1643
<i>Human Relations</i>	HR	Anual	1997 a 2008	738
<i>Nature</i>	NAT	Semanal	1999 a 2008	30490
<i>Physical Review A</i>	PRA	Mensal	2007 e 2008	3089
<i>Physical Review B</i>	PRB	Mensal	2007 e 2008	7847
<i>Physical Review C</i>	PRC	Mensal	2007 e 2008	1572
<i>Physical Review D</i>	PRD	Mensal	2007 e 2008	5527
<i>Physical Review E</i>	PRE	Mensal	2007 e 2008	4157
<i>Physical Review Letters</i>	PRL	Mensal	2007 e 2008	5929
<i>Probabilistic Engineering Mechanics</i>	PEM	Anual	1986 a 2009	703
<i>Science</i>	SCI	Semanal	1999 a 2008	11798
<i>Sociology Health and Illness</i>	SHI	Bimestral	1979 a 2008	845

Tab. 1. Principais informações sobre os dados adotados nesta pesquisa.

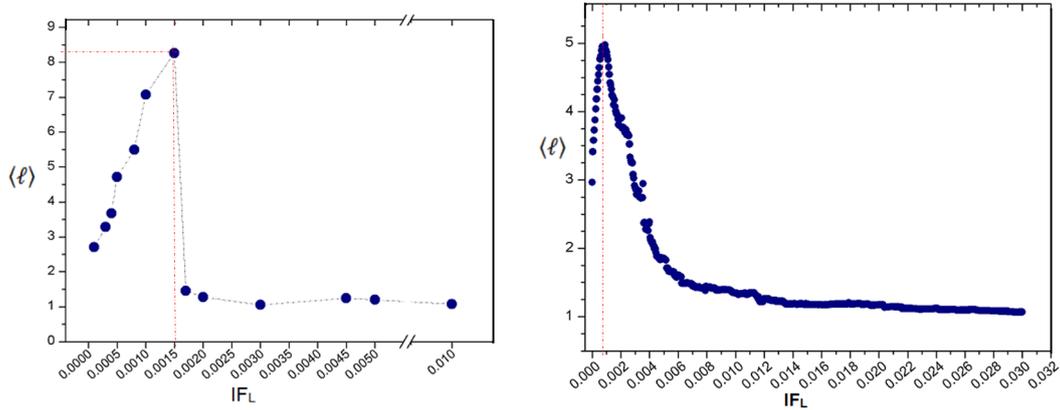
²A base de dados desta pesquisa é a mesma utilizada por Pereira et al. 2011.

3.2. Determinação da rede crítica em RST

Redes críticas foram utilizadas para investigar mecanismos inerentes à linguagem humana, tanto em discursos orais, e.g. estudantes universitários [Teixeira et al. 2010], quanto em discursos escritos, e.g. autores de romances Aguiar 2009. Neste trabalho faremos o mesmo para rede de títulos utilizando as duas abordagens para IF .

Para encontrar o valor de IF que representa a rede crítica do discurso de um periódico, optou-se³ por usar o índice *caminho mínimo médio* ($\langle \ell \rangle$). A análise consiste em verificar o que acontece com o valor do $\langle \ell \rangle$ a medida que se aumenta o valor de IF_L . Em redes de títulos, o comportamento é o mesmo para redes de discursos escritos e orais: a medida que se aumenta o valor de IF , o valor de $\langle \ell \rangle$ aumenta, atinge um valor máximo e em seguida cai bruscamente. Para rede de títulos, dois exemplos são dados na Figura 1.

O índice IF atua como um filtro que “limpa” o texto à medida que seu valor cresce. Inicialmente, o valor de $\langle \ell \rangle$ aumenta - já que a rede perde atalhos⁴ entre os vértices - até um valor máximo. Neste ponto, um pequeno incremento no valor de IF , faz a rede se quebrar em subredes menores.



(a) Caminho mínimo médio em função do índice IF_L (proposto por Teixeira et al. 2010) da rede de títulos da revista AFE, com 370 títulos.

(b) Caminho mínimo médio em função de IF_L (proposto por Aguiar 2009) da rede de títulos da revista Science, com 11798 títulos.

Fig. 1. Localização do ponto crítico, pelas duas abordagens metodológicas que usam a incidência-fidelidade.

Na Figura 1, para valores de $IF \simeq 10^{-3}$, temos a configuração da *rede crítica* para estas RST's. Este mesmo valor também é encontrado nas redes de discurso oral do trabalho de Teixeira et al. 2010. Nesta configuração, a rede possui informação relevante muito bem conectada em uma estrutura que merece ser investigada do ponto de vista topológico. A Figura 2 mostra um exemplo de uma rede crítica da RST do periódico Physical review E.

3.3. Redes de títulos de tamanho fixo

Nesta etapa, as redes foram geradas utilizando o índice IF , proposto por Teixeira et al. 2010. Para esta análise, o número de títulos de cada periódico foi fixado em 300, já

³Este comportamento também ocorre com outros índices, como o Diâmetro D e a diferença normalizada entre vértices e arestas [Teixeira et al. 2010].

⁴Um atalho é uma aresta que quando retirada faz com que o menor caminho entre dois vértices que estavam conectados previamente, tenha caminho maior que 2.

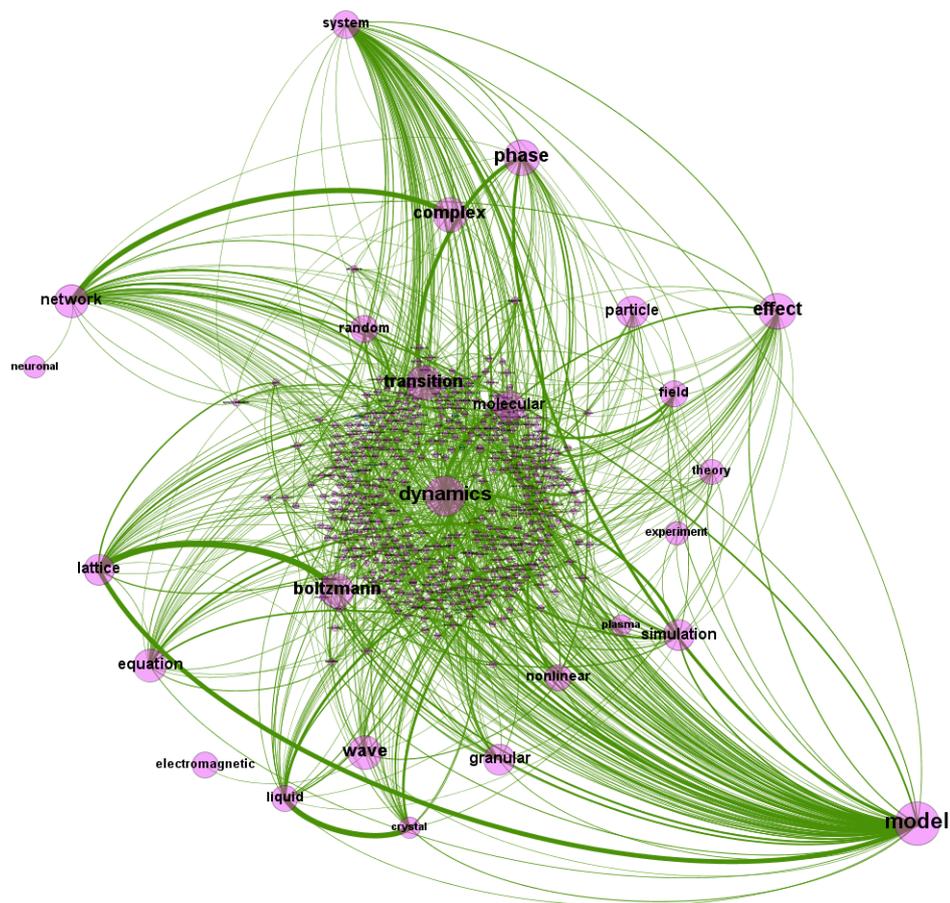


Fig. 2. Rede crítica da RST do periódico Physical review E. A espessura das arestas está proporcional ao valor de IF proposto por Aguiar 2009.

que para o método de Teixeira et al. 2010 as redes não podem diferir consideravelmente em quantidade de cliques. Este número foi escolhido porque a revista *AFE*, de 370 títulos, é a que possui a menor quantidade de títulos da base⁵.

Foram arbitrados em torno de 20 valores⁶ de IF , que variaram de $1 \cdot 10^{-4}$ até $5 \cdot 10^{-2}$, para cada rede em todos os periódicos da base.

Os 300 títulos de cada revista foram escolhidos de forma aleatória dentro do universo de títulos de cada uma. Para realizar este procedimento, foi usado o software livre *guash*, produzido por Monteiro et al. 2010. Foram gerados 100 arquivos de texto contendo 300 títulos aleatórios de cada periódico. Cada arquivo passou pelo processo de tratamento das palavras e em seguida foram geradas as redes e calculados alguns índices de redes para cada valor de IF . Feito isso, foram calculados os valores médios destes índices para as redes de mesmo valor de IF .

3.4. Periódicos com toda sua base de dados

Nesta etapa, as redes foram geradas utilizando o índice IF , proposto por Aguiar 2009. Neste caso, não foi necessário fixar o número de títulos dos periódicos, ou seja

⁵A *Nature*, de 35163 títulos, é a revista que possui a maior quantidade de títulos da base. A *Science* vem em 2º lugar, com 11799 títulos.

⁶20 é satisfatório para perceber o fenômeno da rede crítica. Para alguns periódicos foi preciso 22 ou 23 valores, para melhor detalhamento da região crítica. Outros periódicos, com apenas 15 valores já foram suficientes.

cada periódico da base foi analisado por todos os seus títulos.

Foram escolhidos 600 pontos igualmente espaçados dentro do intervalo de $0 \leq IF_L < 3 \cdot 10^{-2}$. Com este intervalo é possível gerar não só a rede crítica, como também a rede canônica⁷, i.e. rede sem arestas removidas $IF_L = 0$. Para cada periódico, os resultados das próximas seções mostram que a rede de um dado conjunto de títulos possui configuração crítica.

4. Resultados e Discussões para rede de títulos de tamanho fixo (Seção 3.3)

A Figura 3 mostra os valores de $\langle \ell \rangle$ de todos os periódicos em função de IF_L , inclusive a média de seus valores para todos os periódicos. Os dois gráficos revelam o valor da *incidência-fidelidade* crítica (IF_C) para as redes estudadas como sendo $IF_C = IF_L \simeq 10^{-3}$.

A Figura 3 exibe dois gráficos: à esquerda, cada valor de IF_L corresponde a 1500 valores de $\langle \ell \rangle$ (i.e 100 valores para cada um dos 15 periódicos). Muitos destes pontos se sobrepõem, devido a proximidade entre os valores. Na mesma figura, à direita, temos o valor médio de $\langle \ell \rangle$ para cada valor de IF_L .

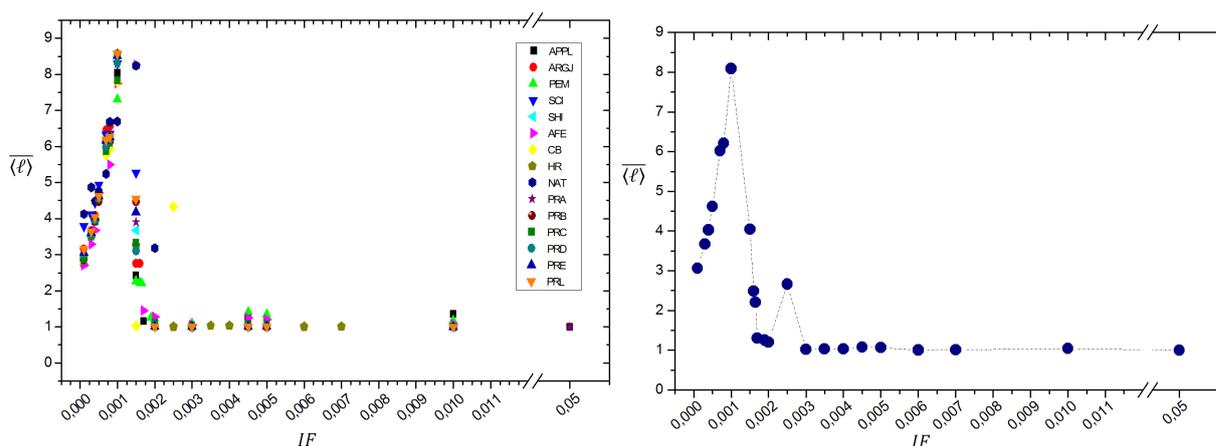


Fig. 3. Na esquerda: Valores de $\langle \ell \rangle$ em função de IF_L para todos os periódicos. Na direita: Média dos valores de $\langle \ell \rangle$ dos periódicos em função de IF_L

O valor $IF_C = 10^{-3}$ é o mesmo encontrado nas redes de discursos orais do trabalho de Teixeira et al. 2010. Isto pode indicar que a escolha de uma publicação pelo periódico tem alguma semelhança com a escolha das frases que um indivíduo faz para expressar suas ideias em um discurso.

O trabalho de Pereira et al. 2011 sugere que periódicos podem ser diferenciados a partir dos valores de *densidade* (Δ), *caminho mínimo médio* ($\langle \ell \rangle$) e pelo coeficiente angular da distribuição de graus (γ) de suas redes.

A Figura 4 mostra o gráfico box-plot⁸ para a distribuição da *densidade*. Percebe-se claramente que existem sobreposições dos valores deste índice para periódicos diferentes. Uma análise mais detalhada, permite-nos perceber que os valores da *densidade* das 100 redes de cada periódico compõe um intervalo relativamente grande de valores, que so-

⁷Esta condição ocorre se I for mínimo ou se F for mínimo.

⁸Boxplot é uma ferramenta útil para a comparação gráfica de várias amostras. seu gráfico informa importantes aspectos do conjunto de dados através de cinco números: valor mínimo, 1º quartil, 2º quartil, 3º quartil e valor máximo [Tukey 1977].

brepõe faixas de valores de *densidade* de outras revistas. Assim, fica inviável diferenciar periódicos utilizando títulos aleatórios em sua base.

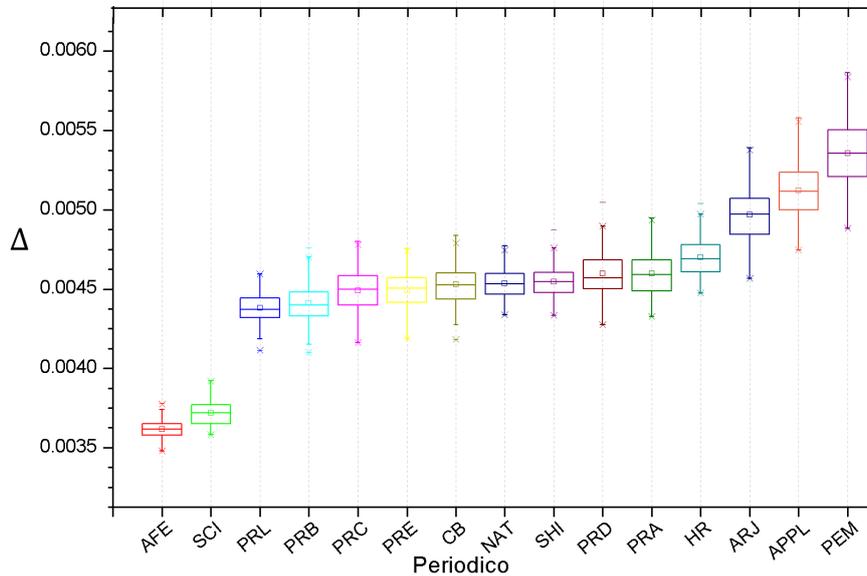


Fig. 4. Média, Máximo, Mínimo e Mediana dos valores de densidade das 100 redes de 300 títulos geradas para cada periódico.

Estes títulos aleatórios configuram um discurso semelhante ao discurso oral de um indivíduo, pois $IF_C = 10^{-3}$ [Teixeira et al. 2010]. Isto sugere que a adição de um grupo de títulos em uma rede de títulos dependa de títulos anteriores. No trabalho de [Cunha et al. 2013], os autores mostram que as publicações do periódico *Nature* possuem correlação de persistência em um intervalo de 5 semanas para alguns índices de redes, inclusive a *densidade*, quando submetidos ao método *DF A* proposto por Peng et al. 1994.

4.1. Distância entre palavras em redes críticas.

De acordo com a hipótese de Cunha et al. 2013, baseado em Caldeira 2005, um título representa a menor unidade de significado do texto formado pelo conjunto de títulos de uma revista, i.e. do “discurso da revista”. Cada título contém um significado, que representa seu artigo. Assim, a partir de sua rede de cliques e de suas métricas, pode-se investigar como temáticas de diferentes publicações relacionam-se entre si.

O *diâmetro* D e o *caminho mínimo médio* $\langle \ell \rangle$ medem, em termos de arestas, respectivamente, a maior distância geodésica e a média das distâncias geodésicas entre dois vértices de uma rede [Newman 2003]. Tomando como exemplo o *diâmetro*, nas redes canônicas $D \simeq 5$. Ou seja, para este valor de D , pelo menos duas palavras possuem distância geodésica de no máximo 5 em uma rede.

Como as redes em questão são redes de cliques, isto significa que dois vértices que possuem $\ell = 5$ pertencem uma subrede de 5 cliques de configuração em formato de linha, ou próximo disto⁹. Portanto, para as redes canônicas - em que praticamente toda a rede está conectada - dois títulos quaisquer de um dado periódico estão relacionados por no máximo três outros entre eles.

⁹Em uma rede de cliques, esta configuração maximiza o *diâmetro*. Para mais informações sobre as diferentes classificações de uma rede de cliques, ver Fadigas and Pereira 2013.

Sabe-se que a rede crítica exibe em seu maior componente, D e $\langle \ell \rangle$ maiores que na rede canônica. Possivelmente, as revistas mais multidisciplinares terão o valor desses índices com maiores aumentos em relação à rede canônica correspondente. É como se, nesta configuração, as menores unidades de significado (i.e. títulos) pudessem ser reduzidas a pares de palavras.

O gráfico da Figura 5 exibe o aumento de cada índice (D e $\langle \ell \rangle$) em relação à rede canônica. Pode-se perceber que, revistas que se dizem “multidisciplinares” possuem os maiores aumentos, chegando até aproximadamente 4 vezes o tamanho da rede original. É o que acontece com a *Nature* e com a *Science*.

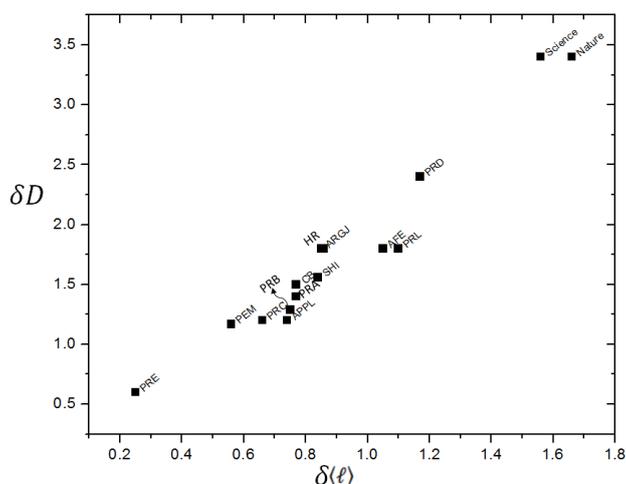


Fig. 5. Variação do D em função da variação do $\langle \ell \rangle$.

Entre as redes críticas, a diferença no diâmetro está associada com sua conectividade (porcentagem do seu maior componente). Pode-se perceber que quanto mais conectada é a *rede crítica*, menor é o valor do seu *diâmetro*. A Figura 6 ilustra este fenômeno.

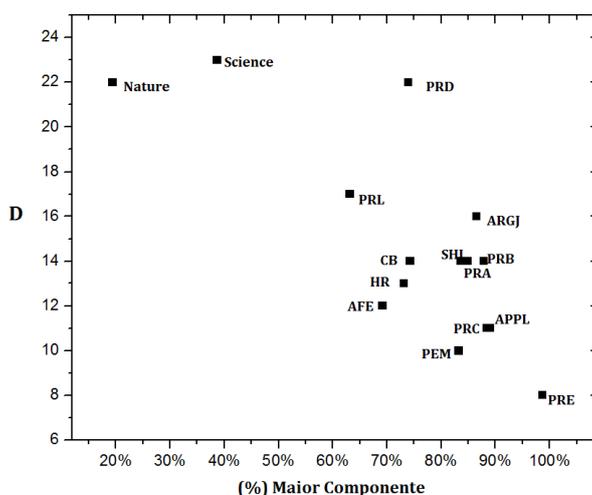


Fig. 6. Porcentagem do maior componente em função do diâmetro das redes críticas.

Este resultado pode nos dar indícios sobre a interdisciplinariedade de um periódico. Percebe-se que a *Nature* e a *Science* com baixa conectividade e alto *diâmetro*

estão no topo, à esquerda, enquanto que *PRC*, *APPL* e *PEM* com alta conectividade e baixo *diâmetro* estão no canto inferior direito, ou seja, grupos diametralmente opostos no gráfico.

4.1.1. Grau médio

Nas redes críticas, os valores de *grau médio* ($\langle K \rangle$) são menores que os da rede canônica. A evidência de que estas redes são livres de escala evidencia a existência de Hubs. Pode-se analisar aqui (Figura 7) a relação entre a *variação do grau médio* $\delta\langle k \rangle$ (Equação 3) com a *variação da densidade* $\delta\Delta$ (Equação 4) das redes críticas em relação às configurações iniciais (i.e. redes canônicas).

$$\delta\langle k \rangle = \left| \frac{\langle k \rangle_{critica} - \langle k_{canonica} \rangle}{\langle k_{canonica} \rangle} \right| \quad (3)$$

$$\delta\Delta = \left| \frac{\Delta_{critica} - \Delta_{canonica}}{\Delta_{canonica}} \right| \quad (4)$$

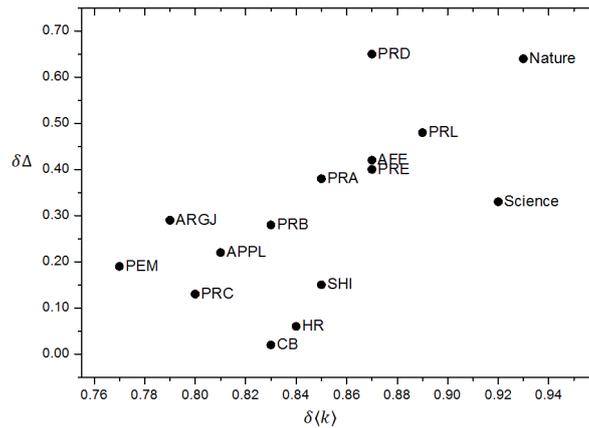


Fig. 7. *Variação da densidade em função da variação do grau médio das redes críticas em relação às redes canônicas.*

5. Veracidade do fenômeno crítico

Foi demonstrado aqui que é possível encontrar rede crítica para as redes de títulos. Esta rede apresenta o máximo de informações com o mínimo de resíduos. Ela pode evidenciar a maneira como um grupo de cientistas escolhem palavras importantes para compor títulos de seus trabalhos para submissão em um periódico e, também, como estes trabalhos são escolhidos pelo cientistas avaliadores.

Para mostrar que isso é intrínseco de uma comunidade de cientistas, propõe-se o seguinte experimento: para um certo conjunto de títulos que apresenta o fenômeno da rede crítica, o tamanho de cada título é mantido (número de palavras em cada título) e as palavras deste conjunto são redistribuídas aleatoriamente nos títulos, produzindo um texto embaralhado.

Curiosamente, a rede gerada a partir do texto embaralhado não apresenta o fenômeno da *rede crítica*. Os gráficos da Figura 8 mostram como varia o $\langle \ell \rangle$ em função

de IF_L em redes com palavras embaralhadas para os periódicos *Science* e *Chemistry and Biology*. Observe que não há um ponto crítico, como foi observado nestes mesmos periódicos sem o processo da embaralhamento.

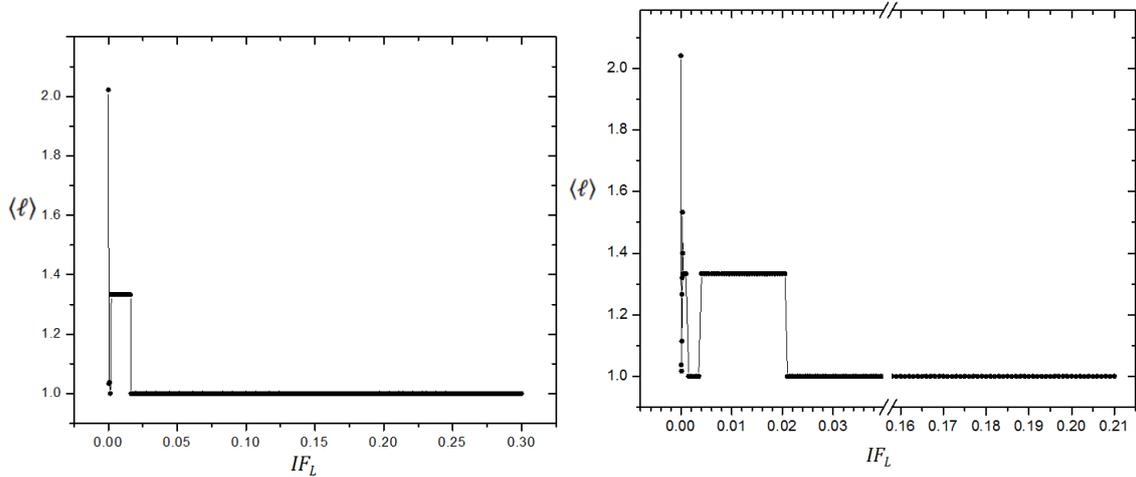


Fig. 8. Não existem pontos críticos para redes de títulos com palavras embaralhadas. À esquerda o periódico *Chemistry and Biology*. À direita, o periódico *Science*.

A Figura 9 mostra os valores dos índices destas redes para alguns valores de IF_L . Pode-se observar que a primeira filtragem ($IF_L = 5 \cdot 10^{-5}$) faz a rede se desmontar, mesmo com IF_L muito baixo. Percebe-se que a rede nessa configuração, apesar de conter 499 vértices, apresenta poucas arestas, configurando-se em pares de palavras isolados.

No início do processo de filtragem a rede não apresentou o fenômeno da perda de atalhos, comum quando se aumenta IF_L , a partir da rede canônica. O valor de $\langle \ell \rangle$ cai bruscamente, ao invés de aumentar, para depois ter a queda brusca. Fica evidente com este experimento que a o fenômeno da rede crítica é algo intrínseco ao comportamento da comunidade de cientistas que publicam em um dado periódico.

IF	n	m	$\langle \ell \rangle$	$\langle C \rangle$	$\langle k \rangle$	Δ
0.00000	15039	261405	2.04	0.34	34.77	0.0023
0.00005	499	263	1.04	0.01	1.05	0.0021
0.00010	221	116	1.02	0.02	1.05	0.0047
0.00015	28	18	1.11	0.10	1.28	0.0476
0.00020	12	10	1.27	0.24	1.67	0.1515
0.00025	10	9	1.32	0.29	1.80	0.2000
0.00030	6	7	1.53	0.478	2.33	0.4667
0.00035	5	6	1.40	0.60	2.40	0.6000
0.00040	4	4	1.33	0.58	2.00	0.6667
0.00100	4	4	1.33	0.58	2.00	0.6667
0.00150	3	3	1.00	1.00	2.00	1.0000
0.00350	3	3	1.00	1.00	2.00	1.0000
0.00350	3	3	1.00	1.00	2.00	1.0000

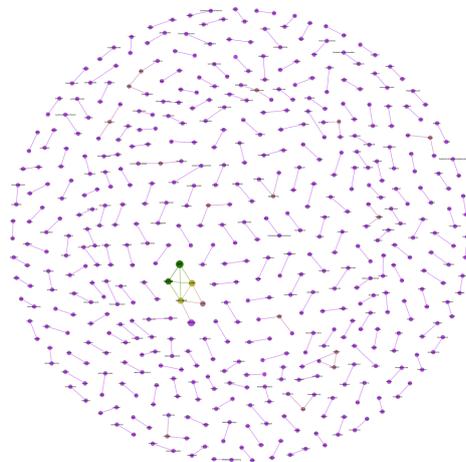


Fig. 9. Índices de redes para cada valor de IF_L para rede de títulos embaralhados da *Science*, à esquerda e Rede para $IF_L = 5 \cdot 10^{-5}$ para rede de títulos embaralhados da *Science*, à direita.

6. Conclusões

As redes complexas oferecem muitas possibilidades de análise de fenômenos naturais. Neste sentido, a comunicação científica é fruto de uma complexa rede de pesquisadores que interagem de diversas formas. Vimos aqui, que a partir do vocabulário comum dos cientistas nos títulos de suas publicações, é possível modelar uma maneira de visualizar localmente como funciona o sistema formal de comunicação científica, i.e. o periódico científico.

Como se sabe, os periódicos tornaram-se estruturas de representação do conhecimento científico da humanidade. Conforme vimos, as redes dos periódicos possuem um ponto crítico bem definido, assim como em redes de discursos escritos e de discursos orais, de trabalhos anteriores a este. Isto nos leva a supor que existem muitas semelhanças entre discursos de pessoas comuns e de conjuntos de publicações em um dado periódico. Este conjunto precisa passar rigorosamente por um comitê avaliador, composto por mais de uma pessoa, diferentemente das redes de discursos orais e escritos mencionados, que são oriundos de uma só pessoa.

Com a rede crítica foi possível indicar algumas maneiras de diferenciar periódicos, cruzando os índices de suas redes críticas. A partir deste olhar, pesquisadores que estudam colaboração e cooperação científica poderão contribuir com interpretações mais precisas para diferenciar periódicos com o uso de redes semânticas. Assim, com a verificação deste fenômeno em *RST*, novas possibilidades se abrem para o estudo da colaboração científica.

Referências

- [Aguiar 2009] Aguiar, M. S. (2009). Redes de palavras em textos escritos: Uma análise da linguagem verbal utilizando redes complexas. Programa de pós-graduação em física, Universidade Federal da Bahia, Salvador.
- [Albuquerque and Pimentel 2004] Albuquerque, F. J. B. and Pimentel, C. E. (2004). Uma aproximação semântica aos conceitos de urbano, rural e cooperativa. *Psicologia: Teoria e Pesquisa*, 20(2):175–182.
- [Caldeira 2005] Caldeira, S. (2005). Caracterização da rede de signos linguísticos: Um modelo baseado no aparelho psíquico de freud. Mestrado interdisciplinar em modelagem computacional, Fundação Visconde de Cairu, Salvador.
- [Cunha et al. 2013] Cunha, M., Rosa, M. G., de Sousa Fadigas, I., Miranda, J. G. V., and de Barros Pereira, H. B. (2013). Redes de títulos de artigos científicos variáveis no tempo. In *BraSNAM - II Brazilian Workshop on Social Network Analysis and Mining*, pages 1744–1755.
- [Fadigas et al. 2009] Fadigas, I., Henrique, T., Pereira, H., Senna, V., and Moret, M. (2009). Análise de redes semânticas baseada em títulos de artigos de periódicos científicos: o caso dos periódicos de divulgação em educação matemática. *Educação Matemática Pesquisa*, 11(1):167–193.
- [Fadigas and Pereira 2013] Fadigas, I. and Pereira, H. (2013). A network approach based on cliques. *Physica A: Statistical Mechanics and its Applications*, 392(10):2576 – 2587.

- [Fadigas et al. 2013] Fadigas, I. S., do Vale Cunha, M., Rosa, M. G., and de Barros Pereira, H. B. (2013). Análise de redes de coautoria por meio de redes semânticas uniformes. In *BraSNAM - II Brazilian Workshop on Social Network Analysis and Mining*, pages 1553–1564, Maceió.
- [Garvey 1979] Garvey, W. D. (1979). *Communication: The Essence of Science*. Oxford, NY: Pergamon Press, Inc.
- [Miranda and Pereira 1996] Miranda, D. and Pereira, M. (1996). O periódico científico como veículo de comunicação: uma revisão de literatura. *Ibict*, 25(3):375–382.
- [Monteiro et al. 2010] Monteiro, R. L. S., de Barros Pereira, H. B., Moret, M. A., and Fadigas, I. S. (2010). Scntools - social and complex network tools. RPI 2083 de 07/12/2010.
- [Nelson et al. 1998] Nelson, D. L., McEvoy, C. L., and Schreiber, T. A. (1998). The university of south florida word association, rhyme and word fragment norms. *Behavior Research Methods, Instruments, & Computers*.
- [Newman 2003] Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review*, 45(2):167–256.
- [Peng et al. 1994] Peng, C.-K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., and Goldberger, A. L. (1994). Mosaic organization of dna nucleotides. *Phys. Rev. E*, 49(2):1685–1689.
- [Pereira et al. 2011] Pereira, H., Fadigas, I., Senna, V., and Moret, M. (2011). Semantic networks based on titles of scientific papers. *Physica A: Statistical Mechanics and its Applications*, 390(6):1192–1197.
- [Sternberg 2011] Sternberg, R. (2011). *Psicologia cognitiva*. Artes Médicas Sul, Porto Alegre, RJ.
- [Teixeira et al. 2010] Teixeira, G., Aguiar, M., Pereira, H., Miranda, J., Cunha, M., Morais, J., Carvalho, C., and Dantas, D. (2010). Complex semantics networks. *International Journal of Modern Physics C*, 21(3):333–347.
- [Tukey 1977] Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley.
- [Vanz and Stumpf 2010] Vanz, S. and Stumpf, I. (2010). Colaboração científica: revisão teórico-conceitual. *Perspectivas em Ciência da Informação*, 15:42–55.
- [Ziman 1979] Ziman, J. (1979). *Comunidade e comunicação*, pages 115–138. Conhecimento público. São Paulo: EDUSP.