



DIFUSÃO DO CONHECIMENTO SOB A PERSPECTIVA DA TEORIA DE REDES: MAPEAMENTO DA PRODUÇÃO CIENTÍFICA A PARTIR DE UMA BASE DE PERIÓDICOS DA FÍSICA

Patricia Freitas Braga

Doutoranda em Modelagem Computacional pelo Serviço Nacional de Aprendizagem Industrial, Brasil.

Professora da Universidade Jorge Amado, Brasil.

E-mail: patyfb04@gmail.com

Hernane Borges de Barros Pereira

Doutor em Engenharia Multimídia pela *Universitat Politècnica de Catalunya*, Espanha.

Professor do Serviço Nacional de Aprendizagem Industrial, Brasil.

E-mail: hbbpereira@gmail.com

Marcelo Albano Moret Simões Gonçalves

Doutor em Ciências Biológicas (Biofísica) pela Universidade Federal do Rio de Janeiro, Brasil. Professor da Universidade Estadual de Feira de Santana, Brasil.

E-mail: mamoret@gmail.com

Resumo

A busca por inovações decorrentes de demandas sociais é a base para o progresso socioeconômico e está estreitamente relacionada com a produtividade científica emergente. Nesse contexto, a difusão do conhecimento científico retrata os impactos e influências de ideias decorrentes dessa produtividade por meio de referências bibliográficas entre as publicações. Objetivando promover uma discussão sobre esse tema, esta pesquisa realizou um estudo sobre a difusão do conhecimento científico à luz da teoria de redes sociais e complexas, a partir das redes de citações, considerando alguns temas do domínio da física. As redes foram construídas a partir da base de periódicos da *American Physical Society* que contém artigos publicados de 1893 a 2009. As redes permitiram reflexões mais aprofundadas, principalmente no que concerne à identificação de grupos de influências. Foi elaborado um modelo computacional que possibilitou a construção das redes. Os resultados obtidos do cálculo de propriedades de redes permitiram-nos analisar a dinâmica da difusão do conhecimento, e observar as influências científicas e a distribuição geoespacial da produção científica mapeada. A distribuição de publicações possibilitou concluir que não há um padrão comum a todos os temas no comportamento da difusão do conhecimento. Ademais, observou-se que a produtividade científica tem sido influenciada por interesses sociais, econômicos e tecnológicos.

Palavras-chave: Redes de citações. Redes complexas. Difusão do conhecimento.

DIFFUSION OF KNOWLEDGE FROM THE PERSPECTIVE OF THE NETWORK THEORY: MAPPING THE SCIENTIFIC PRODUCTION FROM A PHYSICS DATABASE

Abstract

The search for innovations arising from social demands is the basis for socio-economic progress and is closely related to the emerging scientific productivity. In this context, the diffusion of scientific knowledge reflects the impacts and influences of ideas resulting from this productivity by means of references among publications. With the purpose of promoting a discussion on this subject, this research carried out a study on the diffusion of scientific knowledge based on the network theory, starting from the networks of citations of topics in Physics. The networks were built from the American Physical Society base of journals that contains papers published from 1893 to 2009. The networks allowed us to accomplish further reflection, especially with regard to the identification of groups of influences. The steps of the computational model that made possible the construction of the networks are presented. Network properties allowed us to analyze the dynamics of the diffusion of knowledge, and to observe the influences scientific and geospatial distribution of the scientific production analyzed. The distribution of publications allowed us to conclude that there is no common pattern to all subjects in the behavior of the diffusion of knowledge. Moreover, it was observed that the scientific productivity has been influenced by social, economic and technological interests.

Keywords: Citation networks. Complex networks. Diffusion of knowledge.

1 INTRODUÇÃO

A produção e difusão do conhecimento nas ciências assumem um papel relevante na articulação da pesquisa científica, tendo como base propulsora as demandas sociais emergentes. A busca constante por avanços tecnológicos para o desenvolvimento socioeconômico de um país reflete também nos incentivos provenientes de políticas públicas focadas na produtividade científica. Em vista disso, entende-se que a compreensão e mensuração da produtividade científica se tornam essenciais à tomada de decisão estratégica no gerenciamento de recursos alocados para pesquisa e à crescente necessidade por inovação tecnológica em diversos setores. Assim, é possível, segundo Vanti (2002), realizar diagnósticos das potencialidades de determinados grupos e/ou instituições considerando sua produtividade.

Neste trabalho, a análise da difusão do conhecimento de alguns tópicos da física é guiada pela hipótese de que redes sociais e complexas representam a estrutura topológica por onde a informação se difunde, tendo seus vértices mais preeminentes aqueles cujo tema expressado é mais relevante para a comunidade científica. Assim, como podemos observar a influência científica? Uma das evidências para observação da trajetória e influência científica é a quantidade de citações de um artigo. Garfield (1983) define citações como marco da origem de conceitos e ideias disseminadas no texto que as incorpora.

As referências encontradas em publicações formam redes de citações e são marcadores significantes para a observação e o estudo da difusão e influência do conhecimento científico na literatura. As citações, portanto, representam a influência científica originada da literatura antecessora, traçando assim rotas de conhecimento científico. Por exemplo, um determinado conhecimento pode reconfigurar conhecimentos já consolidados, evoluindo-os ou tornando-os obsoletos. No contexto da difusão do conhecimento, as citações de uma publicação viabilizam a estruturação de uma rede pela qual conceitos e teorias são usados consolidando as ideias ali propostas.

O objetivo deste trabalho é analisar diacrônica e geograficamente o impacto de artigos em alguns tópicos da física na reconfiguração desse conhecimento científico. A análise da

difusão do conhecimento científico nesta pesquisa se baseou na base de periódicos da *American Physical Society* (APS) que contém publicações entre os anos de 1893 a 2009 e suas citações. A partir do tratamento dos dados da APS, foram construídas as redes de citações, redes de coautoria, redes de instituições e redes de periódicos. Com essas redes, informações sobre a difusão do conhecimento dos tópicos selecionados foram obtidas. Aspectos como a distribuição de frequências de citações, identificação de grupos de influência, ciclo de vida de um tema foram analisados diacrônico e geograficamente.

Este artigo está organizado da forma que segue. A Seção 2 apresenta o referencial teórico. A Seção 3 apresenta os procedimentos metodológicos, desde o escopo, origem e seleção dos dados até o modelo de construção das redes de citações. A Seção 4 traz os resultados e as discussões. A Seção 5 apresenta as considerações finais do trabalho.

2 REFERENCIAL TEÓRICO

2.1 Difusão científica com base em redes de citações

A discussão sobre a difusão do conhecimento e a produtividade científica já foi anteriormente debatida por vários autores (CHEN; HICKS, 2004; CHEN et al., 2008; PRICE, 1965). Price (1965) realizou uma pesquisa onde foram analisadas as topologias das redes de citações e usados métodos quantitativos para verificação do ciclo de vida útil de publicações científicas. Chen e Hicks (2004) fizeram uma abordagem da difusão do conhecimento sobre a base teórica de redes complexas para visualizar e analisar as redes de citações dentro do campo da engenharia, enfatizando os impactos decorrentes deste estudo para a realização de planejamento estratégico e definição de políticas científicas. Chen et al. (2008) realizaram uma análise das citações oriundas do periódico *Data and Knowledge Engineering*, identificando tendências temáticas e colaboração científica.

Sabendo-se que hoje as citações representam um dos meios de aferição de impacto científico, recentemente, trabalhos que tratam de modelo de comportamento de citações foram propostos. Por exemplo, Wang, Song e Barabási (2013) discutiram a previsão do comportamento das citações a longo prazo. Nesse trabalho, os autores sugeriram um modelo onde as distribuições temporais de citações seguem um mesmo padrão temporal universal. Em outro, Naumis e Phillips (2012) propuseram que o desenvolvimento da pesquisa científica segue uma estrutura interna universal. Contudo, com a melhora e o avanço dos meios de comunicação, que se tornaram globalizados, a partir da década de 60, as citações a publicações científicas passaram a seguir um modelo de difusão anômala, similar ao modelo usado para explicar a transição vítrea molecular.

Apesar dos trabalhos referentes à quantificação da ciência por meio de dados estatísticos obtidos das citações, há ainda que se considerar o contexto e conteúdo dessas citações. Nesse sentido, Romancini (2010) propôs que sejam observadas as características do texto em que as citações estão presentes. Essa visão contextualizada permitirá compreender melhor os rumos da ciência em virtude da influência presente nas citações. A questão da influência científica está fortemente ligada ao momento vivenciado. Historicamente, observa-se isso em demandas (sociais, econômicas, políticas, etc.) que influenciam o foco dado a pesquisas e nesse contexto uma nova cadeia de influência científica pode se iniciar.

2.2 Teoria de redes

Assim como diversos autores, Pereira (2013) propôs três marcos históricos. O primeiro marco aparece no século XVIII com a publicação da solução para o problema do passeio nas pontes de Königsberg (atual Kaliningrado), que consistia em determinar um passeio por todas

as pontes sem passar por uma mesma ponte mais de uma vez. A demonstração de Euler (1736) solucionou o problema ao chegar à conclusão que é impossível um passeio em uma rede que tenha mais de 2 vértices com um número ímpar de ligações. Surgiu a **teoria dos grafos** com a qual é possível representar matematicamente as redes. Dois séculos depois, as **redes sociais** ganharam destaque quando Moreno (1934) introduziu o conceito de sociograma, marcando o início da sociometria, precursora da análise de redes sociais e da psicologia social (WASSERMAN; FAUST, 1994). Sobre as **redes complexas**, “podem ser situadas na interseção entre a teoria dos grafos e a mecânica estatística, envolvendo diversas áreas do conhecimento e, portanto, seu estudo tem característica multidisciplinar” (COSTA et al., 2007). Pereira (2013, p. 42) argumentou que:

Embora seus fundamentos remontem à teoria dos grafos, as redes complexas possuem propriedades que as diferenciam dos grafos não complexos, a exemplo da distribuição dos graus, coeficiente de aglomeração, estrutura de comunidade em qualquer escala, evidência de estrutura hierárquica, etc. Assim, a emergência de algumas propriedades de redes complexas não existentes em redes não complexas e em redes regulares favorece a diferenciação supracitada (PEREIRA, 2013, p. 42).

A teoria de redes, que inclui a teoria dos grafos, a análise de redes sociais e a teoria das redes complexas, deve ser considerada de modo a facilitar a caracterização da rede analisada, com base em suas propriedades. Pereira (2013) argumentou que a topologia de uma rede sugere interpretações específicas para os problemas estudados, e.g. análise dos chamados arranjos produtivos locais, concepção de estruturas organizacionais de informação para melhorar o desempenho de mecanismos de busca de dados e informação, configuração de comunidades em prol de uma melhor difusão e compartilhamento do conhecimento, etc.

Dentro do contexto desta pesquisa, diversos estudos sobre colaboração e difusão científica (e.g. redes de citação e redes de coautoria) foram baseados na teoria de redes, tornando-a uma importante aliada metodológica (ANDRADE et al., 2014; CHEN; HICKS, 2004; PRICE, 1965; WANG; SONG; BARANÁSI, 2013). Assim, a teoria de redes estabelece um ambiente propício para modelar problemas oriundos de diversas áreas do conhecimento.

3 PROCEDIMENTOS METODOLÓGICOS: MATERIAIS E MÉTODOS

A pesquisa realizada uniu as abordagens qualitativa e quantitativa. Por um lado, lançou-se mão a métodos quantitativos sob uma assunção filosófica positivista, por outro, recorreu-se a métodos de pesquisa como o estudo de caso (YIN, 1994), para, sob a assunção filosófica interpretativa, explicar os fenômenos que surgem das redes de citações construídas.

Os procedimentos metodológicos, de acordo com as abordagens qualitativa e quantitativa, consideraram as tarefas de elaboração de métodos de construção de redes de citações (método quantitativo), desenvolvimento de algoritmos que implementem os modelos propostos (método quantitativo), estudo dos comportamentos das redes estudadas (método qualitativo e quantitativo) e análise e interpretação das redes de citações (método qualitativo). Essas tarefas foram contextualmente organizadas nas próximas subseções.

3.1 Escopo, origem e seleção dos dados

A *American Physical Society* (APS), fundada em 1899, é uma associação reconhecida internacionalmente por profissionais da área da física por contribuir com o desenvolvimento e difusão do conhecimento da física. A análise das redes de citações desta pesquisa levou em

consideração apenas os artigos dos periódicos *Physical Review Letters*, *Physical Review* e *Reviews of Modern Physics*, publicados de 1893 até 2009, e as citações às publicações contidas na base de periódicos da APS, limitando a projeção das redes a apenas uma amostragem parcial das redes reais. Dessa forma, não foram analisadas nesta pesquisa as referências a publicações de bases mantidas outras editoras ou sociedades.

Considerando a limitação supracitada, a seleção dos temas e publicações teve como critérios a quantidade e a qualidade das informações (e.g. quantidade de publicações, conectividade, etc.) que a rede forneceria, de forma a subsidiar análises mais ampliadas. Foram selecionados os seguintes temas: *Conductivity*, *Fractals*, *Microwave*, *Networks* e *Photoelectrons*. De cada tema, foi selecionada uma publicação (i.e. um artigo) para analisar a sua difusão do conhecimento, em termos de citações.

As informações provenientes da base de periódicos da APS contêm os seguintes dados sobre as publicações: título, ano, autores e suas afiliações e periódico onde foi publicado o trabalho. Com essas informações, foi possível criar estruturas de grafos (i.e. a tarefa de elaboração de métodos de construção de redes de citações), formando redes de citações e outras redes derivadas destas, como as redes de coautoria, redes de instituições e redes de periódicos que apesar de importantes, estão fora do escopo deste artigo.

3.2 Propriedades de redes usadas

Uma rede é um grafo que é representado, segundo Gross e Yellen (2004), como um par de conjuntos $G = \{V, E\}$, onde V é o conjunto de vértices (i.e. entidades da rede) e E o conjunto de arestas (i.e. representação do relacionamento entre um par de vértice qualquer). A rede de citações são grafos dirigidos representados por $G = \{V, A\}$, onde A é o conjunto de arcos. Isso significa que a direção do arco diz respeito às influências das publicações. O estudo proposto se baseia em algumas propriedades de redes que ajudaram a caracterizar topologicamente as redes de citações, assim como respaldaram as análises e interpretações das redes:

- **Número de vértices** é dado pela cardinalidade do conjunto V ($n = |V|$) e indica a ordem do grafo. No contexto das redes de citações, uma publicação é um vértice;
- **Número de Arestas** é dado pela cardinalidade do conjunto A ($m = |A|$) e indica o tamanho do grafo. Nas redes de citações as arestas conectam as publicações pela existência de referências entre elas;
- **Grau de um vértice** representa a quantidade de conexões em cada vértice. No presente contexto, a quantidade média de conexões pode ser usada para observar a importância das publicações com respeito às referências recebidas ou feitas. Uma publicação com um grau alto é importante por ter influenciado significativamente outras publicações;
- **Grau Médio** $\langle k \rangle$ representa a quantidade média de conexões considerando todos os vértices da rede;
- **Densidade** é a razão entre o número de aresta (i.e. relacionamento) existentes e o número máximo possível. A densidade de uma rede é uma propriedade de coesão;
- **Caminho Mínimo Médio** representa a média das distâncias geodésicas entre todos os pares de vértices de uma rede. Em se tratando de redes de citações, quanto maior for a distância entre as publicações, menor a influência direta entre elas.
- **Coefficiente de Aglomeração** mede até que ponto os vizinhos de um vértice qualquer são vizinhos entre si. Em redes de citações, esta propriedade indica uma tendência à formação de agrupamentos entre as publicações.

As tarefas de estudar os comportamentos das redes de citação e de analisá-las e interpretá-las foram possíveis devido à escolha das propriedades de redes.

3.3 Construção das redes de citações

Os dados contidos na base de periódicos da APS necessitaram de um pré-tratamento para serem usados na construção e análise das redes de citações. Por essa razão, foi desenvolvido um modelo computacional para obtenção das informações que contemplasse a estruturação, construção e visualização das redes (i.e. a tarefa de desenvolvimento de algoritmos). A aplicação do modelo computacional possui três etapas:

- 1. Normalização e gravação dos dados:** Os arquivos XML e CSV originados da base de periódicos da APS são submetidos ao software RCC.NET (BRAGA et al., 2013) para obtenção das informações. Estes arquivos contêm dados de publicações e suas citações presentes nos periódicos *Physical Review Letters*, *Physical Review* e *Reviews of Modern Physics da APS*. As informações obtidas foram inseridas em um banco de dados estruturado e centralizado, assim a construção das redes tornou-se mais simples. Ademais, os dados inseridos são vinculados a coordenadas geográficas, utilizando a ferramenta do GMaps¹. Esta etapa é necessária para caracterizar a distribuição geográfica das publicações;
- 2. Construção das redes:** A partir do banco de dados supracitado, as redes de citações foram criadas considerando dois escopos diferentes: (a) redes criadas a partir de um tema geral e (b) redes criadas a partir de uma publicação de origem. Ao final da geração das redes, as mesmas poderão ser gravadas em formato NET (nativo do *Pajek*² e compatível com muitos *softwares* que trabalham com redes sociais e complexas), TXT ou RCC (essencial para a visualização destas redes no mapa do modelo);
- 3. Visualização das Redes:** Os arquivos gravados no formato RCC poderão ser usados nesta etapa pelo software RCC.NET que está preparado para ler e apresentar as redes com informações cartográficas das publicações. Ao selecionar um arquivo RCC, deve-se definir qual rede (no caso deste trabalho, selecionam-se as redes de citações; não obstante, as redes de coautoria e as redes de instituições também podem ser usadas neste módulo) será desenhada no mapa. Também é necessária a definição do autor da publicação (no caso de publicação com dois ou mais autores) que será considerado; com essa definição o software usará as coordenadas geográficas da afiliação do autor escolhido.

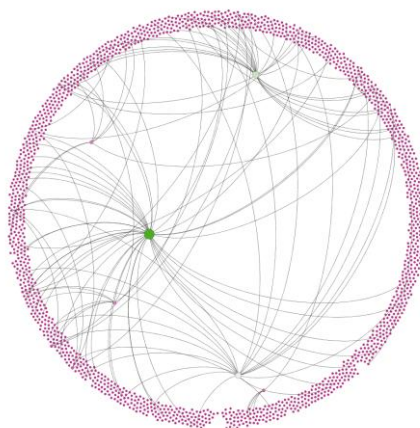
4 RESULTADOS E DISCUSSÃO

As redes de citações para o escopo 'tema geral' (definido na etapa ii do modelo computacional) foram construídas a partir da definição de temas da física; para este estudo, usamos os temas: *Networks*, *Conductivity*, *Fractal*, *Microwaves*, *Photoelectron*. Para cada rede de citação, o critério de seleção consistiu em identificar nos títulos das publicações a palavra do tema definido. Na Figura 1, apresentamos a rede de citações para o escopo 'tema geral' *Networks* e na Figura 2, a rede de citação da 'publicação de origem' *Statistical mechanics of complex networks* (ALBERT; BARABASI, 2002).

¹ Disponível em: <http://maps.google.com>.

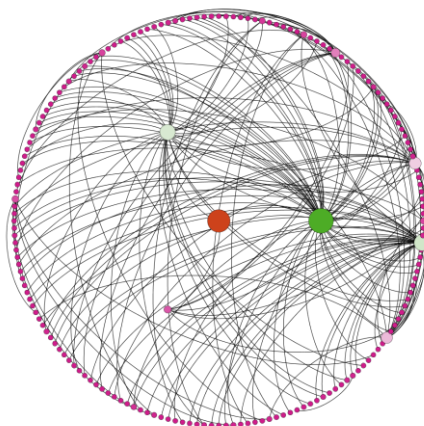
² Disponível em: <http://pajek.imfm.si/doku.php>.

Figura 1 - Rede de citações para o escopo 'tema geral' *Networks* ($n = 1564$, $m = 117$)



Fonte: Dados da pesquisa, 2013

Figura 2 - Redes de citações para o escopo 'publicação de origem' *Statistical mechanics of complex networks* ($n = 180$ e $m = 187$)



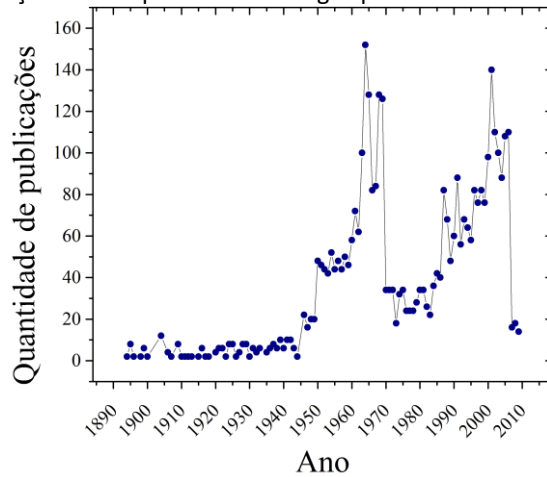
Fonte: Dados da pesquisa, 2013

A partir de cada tema, foi selecionada uma 'publicação de origem', para observação do seu ciclo de vida produtivo e seu impacto para a comunidade científica. A análise da difusão do conhecimento a partir das redes de citações considerou, por um lado, a distribuição das publicações por tempo e por localização geográfica e, por outro, a caracterização topológica dessas redes.

4.1 Distribuição de Citações por Intervalo de Tempo

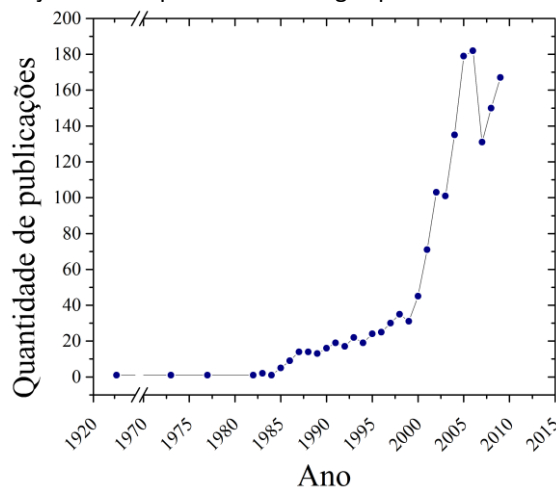
A distribuição de citações ao longo dos anos para cada tema mostrou comportamentos distintos. A depender do tema, houve momentos onde foram observados aumentos significativos na quantidade de publicações, assim como períodos onde esta quantidade é menor (Gráfico 1). Em outros casos, houve um pico na quantidade de publicações acerca daquele conhecimento específico, antes de entrar em declínio produtivo. Ainda há casos onde a produtividade é ascendente, por exemplo o tema *Networks* (Gráfico 2).

Gráfico 1 - Distribuição das frequências dos artigos publicados com o tema *Conductivity*



Fonte: Dados da pesquisa, 2013

Gráfico 2 - Distribuição das frequências dos artigos publicados com o tema *Networks*.



Fonte: Dados da pesquisa, 2013

Durante as pesquisas sobre os temas propostos, foram encontradas algumas informações (e.g. proposições de temas inovadores, concessão de prêmios importantes a cientistas por seus trabalhos relevantes, longevidade de uma teoria, modelo, método, etc.) que podem explicar o comportamento das redes de citações nos referidos domínios. A análise dessas informações e seu momento histórico permitiu-nos a entender o comportamento das redes de citações.

A ascensão de alguns temas foi marcada pelo surgimento de referências contendo nomes influentes na comunidade científica, como foi o caso de Benoit Mandelbrot, cujo reconhecimento advém da proposição da geometria fractal. Conseqüentemente, suas publicações influenciaram este campo do conhecimento científico. Em 1977, Mandelbrot publicou o livro *Fractals: Form, Chance and Dimension*, considerado pela *American Scientist Magazine* em 1999, um dos dez ensaios científicos mais influentes do século 20. A curva de publicações neste tema passa por uma brusca ascensão após 1983, um ano após seu livro, *The Fractal Geometry of Nature*, ser publicado, pois popularizou o tema Fractais.

Outro aspecto que influenciou a quantidade de publicações foi a concessão do Prêmio Nobel em Física de 1978 aos pesquisadores Arno Allan Penzias e Robert Woodrow Wilson pela descoberta da radiação cósmica de fundo em micro-ondas. Após essa data, a distribuição de citações no tema *Microwaves* começou a crescer. Esses fatos apontaram para um efeito de causalidade propiciado por eventos reais que impactaram deterministicamente na produção científica.

Ainda há a questão da longevidade da publicação. Networks é um tema relativamente recente, se comparado a outros temas da física. A quantidade de publicações encontradas na base de periódicos da APS mostrou uma curva ascendente (Figura 4). A partir da análise das redes de citações oriundas deste domínio do conhecimento, parece razoável afirmar que o tema ainda tem muitos problemas em aberto o que tem chamado a atenção de pesquisadores.

As redes de citações, a partir de uma publicação de origem, apresentaram distribuições similares. Todas as redes, exceto a rede formada a partir do artigo *Statistical mechanics of complex networks* (ALBERT; BARABÁSI, 2002), apresentaram um aumento na quantidade de citações por ano, atingindo um ápice seguido de uma tendência de declínio. Algumas publicações como o artigo *intitulado Transition from Resonant to Nonresonant Line Shape in Microwave Absorption* (BEN-REUVEN, 1965), tiveram seu ciclo de vida produtivo por um período longo, neste caso por quase 20 anos, sendo citado por diversas publicações, antes de cair quase em desuso. Price (1965) observou que uma publicação perdura em circulação por período médio de 10 anos.

Parece razoável afirmar que a primeira publicação sobre um tema não é necessariamente a mais produtiva quando comparada com as posteriores, pois as publicações podem ter surgido em momentos históricos diferentes (social, econômico, político ou tecnológico). Por exemplo, em 1965 existiram demandas específicas, diferentes das demandas emergentes em 2009. Em linhas gerais, do ponto de vista cronológico, percebeu-se que momentos históricos impactaram de forma determinante na difusão do conhecimento e consequentemente na consolidação das ciências.

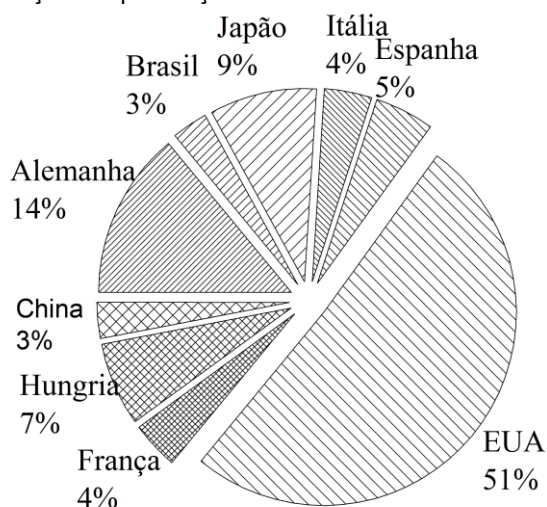
Cabe ressaltar que as citações apresentam comportamentos distintos. Naumis e Phillips (2012) observaram que o comportamento das citações segue modelos de difusão distintos. Até a década de 60, o comportamento seguia um modelo linear de difusão. A partir da década de 60, com a globalização dos meios de comunicação, as citações a publicações científicas passaram a seguir um modelo de difusão anômala (modelo não linear).

4.2 Distribuição de citações por países

Na busca por informações de quais países eram mais produtivos cientificamente, considerou-se a distribuição de frequências de publicações considerando a afiliação dos autores. Em quase todas as redes foi encontrada uma concentração muito alta de publicações provenientes dos Estados Unidos, seguidos pela Alemanha e Japão. Ainda assim, esses países juntos não representam uma percentagem significativa comparada à produtividade estadunidense em publicações da APS (Gráfico 3).

Este resultado é condizente com as estatísticas apresentadas pelo *Essential Science Indicators*. Para o período de 2002 a 2006, os Estados Unidos lideraram o ranking mundial de publicações científicas, com ampla vantagem em relação a outros países. Abaixo dos Estados Unidos há um grupo de cinco países (Japão, Alemanha, Reino Unido – representado pela Inglaterra – China e França) com produção científica mais expressiva (Figura 5).

Gráfico 3 - Distribuição das publicações considerando totais de todos os temas por países.



Fonte: Dados da pesquisa, 2013

Existem algumas justificativas para a alta produtividade apresentada por esses países. Fatores de ordem tecnológica e econômica, por exemplo, refletem diretamente na produtividade científica sobre certos temas. Em adição a isso, um dos fatores correlacionado à produção científica é produto interno bruto (PIB). O PIB é um dos indicadores mais relevantes para a mensuração de atividade econômica de um país. Por exemplo, o PIB dos Estados Unidos cresceu vertiginosamente entre 1970 e 2010, passando de U\$ 1,025 trilhões a U\$ 15,68 trilhões. O salto do PIB coincidiu com o aumento da quantidade de artigos de periódicos científicos e técnicos. Entre os países que tiveram aumento mais significativo em número de publicações estão os Estados Unidos e China. Entre 1982 e 2009, número de artigos nos Estados Unidos passou de aproximadamente 137.000 a cerca de 210.000, enquanto a China passou de quase zero a próximo de 70.000. Em relação a pedidos de patentes, entre 1985 e 2012, a China apresentou o salto mais significativo, passando de menos de 50.000 a mais de 1 milhão e meio de pedidos de patentes, seguida pelos Estados Unidos, que passou de aproximadamente de 64.000 a cerca de 314.000 pedidos de patentes, nesse mesmo período.

4.3 Análise das propriedades de redes

Com as redes de citações construídas para o escopo tema geral, iniciou-se o processo de caracterização topológica dessas redes. As propriedades das redes revelaram algumas características sobre sua estrutura. Um fenômeno marcante observado foi um número elevado de publicações com poucas referências bibliográficas entre si, em especial para as redes construídas a partir do escopo temas gerais. São redes esparsas, pois as publicações possuem poucas citações entre si por estarem restritas à base de periódicos da APS.

Nas redes construídas a partir do escopo publicação de origem, notou-se, em geral, uma equivalência no número de vértices e no número de arestas. Entretanto, essas redes apresentaram propriedades com valores baixos; em média, o grau médio é igual a 1,0208, o coeficiente de aglomeração é igual a 0,0034 e a densidade é igual a 0,0132. As redes construídas a partir do escopo temas gerais também apresentaram valores baixos.

Alguns fatores contribuíram para esses resultados, sendo o principal o limite de escopo das informações. A fonte de dados utilizada é restrita a artigos publicados nos periódicos da APS, assim como as citações entre eles. Cabe ressaltar que boa parte das referências usadas nas publicações da APS são de artigos publicados em periódicos externos à base de periódicos

da APS. A percepção da baixa influência que os artigos publicados na base de periódicos da APS exercem ou sofrem de outros artigos publicados na mesma base contribui para a credibilidade da sociedade em questão, pois referências cruzadas não são determinantes para os índices bibliométricos dos periódicos da APS.

Nas redes analisadas, tanto no escopo tema geral, quanto no escopo publicação de origem, identificou-se a existência de vértices altamente conectados (*hubs*) e de um número alto de vértices com poucas conexões incidentes. Em outras palavras, há um baixo número de publicações com muitas referências (*hubs*) e um grande número de publicações nas redes com poucas referências. Não obstante, observou-se que as publicações muito referenciadas (*hubs*) influenciaram a produção científica neste meio.

5 CONSIDERAÇÕES FINAIS

Este trabalho analisou a difusão do conhecimento, dentro do domínio da física, especificamente a partir da produção científica publicada na base de periódicos da *American Physical Society* (APS). A análise realizada considerou as redes de citações construídas para os escopos tema geral e publicação de origem, a partir das quais os padrões de comportamento foram discutidos.

Não obstante, alguns pontos críticos surgiram durante o processo de análise. O uso da base de periódicos da APS delimitou a abrangência das redes construídas, de modo que uma representação parcial foi analisada. Por exemplo, apesar de haver um número grande de publicações na base de periódicos da APS que aborda a teoria de Bose-Einstein, há pouca ocorrência de citações entre essas publicações na base de periódicos da APS, visto que os trabalhos foram publicados originalmente em periódicos alemães (BOSE, 1924; EINSTEIN, 1924; EINSTEIN, 1925).

Outro ponto crítico consistiu na incompletude dos dados necessários para a análise diacrônica e geográfica proposta. Por exemplo, nas publicações não existe georreferenciamento, senão as afiliações, a partir das quais obtivemos com o auxílio de um algoritmo as coordenadas geográficas. Entretanto, observou-se que em algumas publicações as afiliações eram inexistentes. A solução desse problema consistiu em situar as publicações em nenhum país.

Em virtude da delimitação à base de periódicos da APS, os temas selecionados para análises da difusão do conhecimento científico foram aqueles que apresentaram maior frequência na base de periódicos usada. Esse procedimento contribuiu para a construção de redes de citações com quantidades altas de vértices ($632 \leq n \leq 3976$), permitindo a observação de propriedades emergentes e a análise da topologia das redes.

Um dos aspectos mais relevantes observados é relativo à distribuição das publicações ao longo de um período e por concentração em países. A distribuição de publicações possibilitou a visualização das tendências temáticas como foco de pesquisas e produções científicas. Observou-se, por exemplo, que certos temas apresentaram um comportamento de altas e baixas na frequência de publicações associadas. Em outros temas, observou-se um pico inicial na quantidade de publicações seguido de uma queda gradativa. De modo geral, verificou-se que não há um padrão comum a todos os temas no comportamento da difusão do conhecimento que explique a distribuição da frequência de publicações científicas.

Na busca de evidências que explicassem o comportamento das distribuições estudadas, encontraram-se algumas informações e fatos relevantes, a exemplo do surgimento de demandas em determinado contexto socioeconômico ou de publicações de alguns ganhadores do Prêmio Nobel da Física. Por exemplo, o surgimento do laser na década de 60, impulsionou a alta na produtividade em temas relacionados à fotodetecção e das estatísticas de luz, sendo identificado este comportamento na distribuição de publicações ao longo dos

anos neste tema. Em linhas gerais, pôde-se observar que a produtividade científica tem sido influenciada por interesses sociais, econômicos e tecnológicos.

AGRADECIMENTOS

Este trabalho recebeu apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq.

REFERÊNCIAS

- ALBERT, Reka; BARABASI, Albert L. Statistical mechanics of complex networks. **Reviews of modern physics**, v. 74, p. 47-97, 2002.
- ANDRADE, Maria Teresinha Tamanini; et al. Contextualized Analysis of Social Networks: Collaboration in Scientific Communities. **Social Networking**, v. 3, p. 71-79, 2014.
- BEN-REUVEN, A. Transition from Resonant to Nonresonant Line Shape in Microwave Absorption. **Physical Review Letters**, v. 14, n. 10, p. 349, 1965.
- BRAGA, Patrícia F. et al. **RCC.NET**. 2013. Patente: Programa de Computador. Número do registro: BR5120130012619, Instituição de registro: INPI - Instituto Nacional da Propriedade Industrial.
- CHEN, C.; HICKS, D. Tracing knowledge diffusion. **Scientometrics**, v. 59, n. 2, p. 199-211, 2004.
- CHEN, C. et al. The thematic and citation landscape of Data and Knowledge Engineering (1985–2007). **Data & Knowledge Engineering**, v. 67, n. 2, p. 234-259, 2008.
- COSTA, Luciano da F. et al. Characterization of complex networks: A survey of measurements. **Advances in Physics**, v. 56, n. 1, p. 167-242, 2007.
- EINSTEIN, Albert. Quantentheorie des einatomigen idealen Gases. **Sitzungsberichte der Preußischen Akademie der Wissenschaften**, v. Physikalisch-mathematische Klasse, p. 261-267, 1924.
- _____. Quantentheorie des einatomigen idealen Gases: Zweite Abhandlung. **Sitzungsberichte der Preußischen Akademie der Wissenschaften**, v. 1, p. 3-14, 1925.
- EULER, Leonard. Solutio problematis ad geometriam situs pertinentis. **Commentarii Academiae Scientiarum Imperialis Petropolitanae**, v. 8, p. 128-140, 1736.
- GARFIELD, Eugene. **Citation Indexing – its theory and application in science, technology and humanities**. Philadelphia: ISI Press, 1983.
- GROSS, Jonathan L.; YELLEN, Jay (Ed.). **Handbook of Graph Theory**. Boca Raton: CRC PRESS, 2004.
- MORENO, Jacob L. **Who Shall Survive?** Washington DC: Nervous and Mental Disease Publishing Company, 1934.
- NAUMIS, Geraldo G.; PHILLIPS, James C. Diffusion of knowledge and globalization in the web of twentieth century science. **Physica A**, v. 391, n. 15, p. 3995-4003, 2012.
- PEREIRA, Hernane Borges de Barros. Redes Sociais e Complexas: Aplicações em Difusão do Conhecimento. **Academia de Ciências da Bahia: Memória III**, v. 3, p. 39-47, 2013.
- PRICE, Derek J. de Solla. Networks of Scientific Papers. **Science**, v. 149, n. 3683, p. 510-515, 1965.

ROMANCINI, Richard. O que é uma citação? A análise de citações na ciência. **Intexto**, v. 2, n. 23, p. 20-35, julho/dezembro 2010.

VANTI, Nadia Aurora Peres. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. **Ciência e Informação**, v. 31, n. 2, p. 152-162, maio/agosto 2002.

WANG, Dashung; SONG, Chaoming; BARABÁSI, Albert-László. Quantifying Long-Term Scientific Impact. **Science**, v. 342, n. 6154, p. 127-132, outubro 2013.

WASSERMAN, Stanley; FAUST, Katherine. **Social Network Analysis: methods and applications**. Cambridge: Cambridge University Press, 1994.

YIN, Robert K. **Case Study Research, Design and Methods**. Thousand Oaks: Sage Publications, Inc., 1994.

Artigo recebido em 07/02/2014 e aceito para publicação em 15/08/2014
