

Uma aproximação às redes sociais para avanços técnico-científicos

José Lamartine de Andrade LIMA NETO^{2,6}, Dirceu de Freitas Piedade MELO^{2,6},
Claudia Silva SANTANA^{1,6}, Gillian Leandro de Queiroga de LIMA^{3,6}, Patrícia
Nicolau MAGRIS^{1,6}, Trazíbulo Henrique Pardo CASAS^{5,6}, Hernane Borges de
Barros PEREIRA^{1,4,6}

¹UNEB. Av. Silveira Martins, nº 2555, Cabula, Salvador - BA

²IFBA. R. Emídio Santos s/n Barbalho, Salvador - BA

³UFBA. Av. Reitor Miguel Calmon, Vale do Canela, Salvador - BA

⁴SENAI CIMATEC, Av. Orlando Gomes, 1845 - Piatã, Salvador - BA

⁵UEFS, Av. Milton Santos, módulo 5, Campus Universitário - Feira de Santana - BA

⁶Programa de Doutorado Multi-institucional e Multidisciplinar em Difusão do
Conhecimento (DMMDC) - UFBA, IFBA, SENAI, UNEB, LNCC, UEFS

{claudiasisan, gillianqueiroga, joselamartineneto, patmagris,
hbbpereira}@gmail.com, dirceumelo@ymail.com, henrique@uefs.br

Abstract. *This paper presents a method of network construction based on analysis of a historic social network, composed of scientists who are part of the book's narrative "The Information" - History, a Theory, a Flood - (2011), by James Gleick. The obtained data permit to evaluate the network structure and the influence of certain characters throughout the book. The characteristics that the network acquired after being modified by the removal of the greater degree of vertices were analyzed. As a result found a network with small world phenomenon.*

Resumo. *O presente trabalho apresenta um método de construção de redes baseado em análise de uma rede social histórica, constituída de cientistas que fazem parte da narrativa do livro "The Information" - History, a Theory, a Flood - (2011), de James Gleick. Os dados obtidos possibilitaram avaliar a estrutura da rede e a influência de determinados personagens ao longo do livro. Foram analisadas as características que a rede adquirira após ser modificada com a retirada dos vértices de maior grau. Como resultado foi encontrada uma rede com o fenômeno mundo pequeno.*

1. Introdução

O interesse pelo estudo das redes de influência científica tem sido estimulado pela necessidade de compreender a dinâmica de criação e difusão do conhecimento, por meio do mapeamento de sujeitos e organizações que fazem parte desse processo, buscando entender, entre outros aspectos, como ocorre a colaboração e o compartilhamento de conhecimento entre essas pessoas e grupos.

O estudo se localiza no campo de Redes Sociais, uma estrutura composta por pessoas ou organizações (i.e. nós ou vértices), conectadas por um ou mais tipos específicos de relacionamentos (e.g. amizade, parentesco, interesses, crenças, conhecimento ou prestígio, entre tantas outras). Objetivamente, uma rede social é um mapa de relações específicas entre aqueles que a compõem (WASSERMAN; FAUST, 1994, p. 04).

No contexto da ciência, observamos que as redes de influência científicas são estudadas como instrumentos para analisar a difusão do conhecimento construído por cientistas que, em colaboração, realizam pesquisas. Esse ambiente é propício para a socialização e difusão do conhecimento, pois promove o processo de produção e disseminação de novos avanços técnico-científicos, colaborando com o progresso da ciência e da tecnologia em todas as áreas.

Do ponto de vista de sua representação, uma rede consiste em um grafo definido por $G = (V, E)$ que é uma estrutura matemática constituída por dois conjuntos V (finito e não vazio) e E (relação binária sobre V). Os elementos de V são chamados de vértices e os elementos de E são chamados de arestas (GROSS; YELLEN, 1999). Cabe uma ressalva, nessa pesquisa não consideramos laços (i.e. arestas ou arcos que incidem em apenas um único vértice), nem pares de vértices que possuem arestas múltiplas.

Pretendemos estudar como uma rede social instituída por descoberta de conhecimentos se forma ao longo da história e suas influências epistemológicas. Usamos o livro *“The Information – History, a Theory, a Flood”* (GLEICK, 2001). Esse livro é o resultado de uma arqueologia da informação propiciando o entendimento de que, a partir dos tambores na África, das bibliotecas ao redor do mundo, criação do neologismo *bit* (unidade de medida da informação), da entropia de Claude Shannon, da internet e do código genético, toda e qualquer informação poderá ser medida.

Neste artigo será apresentado um método de construção e análise de uma rede social histórica. A fim de traçar a história da origem do que se conheceu como Era da Informação, esta obra segue um caminho que gera especulações sobre o destino dos avanços (tecnologias) que apareceram ao longo da história, à medida que o cenário se descortina, revelando também pessoas e suas inter-relações. Dessa forma consideramos ainda como objetivos deste trabalho: (1) verificar a apropriação e interpretação de fluxos de informação e conhecimento e (2) identificar como uma dada rede social (ARS) dos personagens na obra de Gleick se relacionaram com a produção de avanços técnico-científicos.

Após a coleta e tratamento de dados, analisamos os resultados obtidos com o uso de técnicas da teoria de redes (e.g. cálculo de propriedades, caracterização topológica, simulações, etc.). Assim, foi possível avaliar a influência de determinados personagens ao longo da narrativa do livro, a partir da estrutura da rede.

2. Trabalhos Correlatos

Elson, Dames e McKeown (2010) desenvolveram um método para extrair redes sociais a partir de romances britânicos do século XIX e seriados de TV, obtendo redes de interações dialógicas. O método depende da capacidade de determinar quando dois personagens estão em conversação. O método apresentado considera que os personagens estão no mesmo espaço-tempo, falam a mesma língua e são "mutuamente cientes um do outro".

Dentro dessa mesma linha de abordagem de ARS, Gil, Kuenzel, Suen (2011) apresentaram uma combinação de duas abordagens (de um estudioso literário e de um cientista da computação) para realizar análise literária. O trabalho desses autores consiste em três componentes principais: (1) extração de redes de interação entre personagens, (2) cálculo das propriedades das redes, e (3) procuraram responder questões sobre as obras (e.g. os diferentes gêneros ou tipos de mídia são caracterizados por tipos distintos de redes de interação?). Com isso conseguiram construir classificadores para distinguir uma infinidade de características que definem filmes e peças de teatro.

Martins (2014) esclarece que o conhecimento e as experimentações são ampliados a partir das redes sociais que vêm se mostrando um importante objeto de pesquisa em diversas áreas nas últimas décadas (i.e. co-participação em atividades que gerem conhecimento).

Balancieri *et al* (2005) sinalizaram que as análises de redes de pesquisa podem ser revisitadas à luz das possibilidades recentes surgidas com as novas Tecnologias de Informação e de Comunicação (TICs). Para os autores, as redes de pesquisa impulsionam a criação do conhecimento e o processo de inovação resultantes do intercâmbio de informações e, sobretudo, da junção de competências de grupos que unem esforços na busca de metas comuns.

Indo além das histórias contrafactuais da ciência como um exercício de probabilidade das histórias possíveis mas, que não se realizaram, segundo Pessoa Jr. (2000), a teoria da ciência é baseada na noção de "avanço", em que unidades passadas de cientista para cientista se conservariam na passagem de uma história possível para outra, por estarem ligadas entre si por redes de influência causal.

3. Cenário

O autor conduz os leitores por um percurso que se inicia com a atenção dada ao uso dos tambores na África, passa pela construção de locais específicos de guarda de informações como as bibliotecas, o telégrafo e, continua com o processamento de grandes volumes dados a partir da era do computador e as novas possibilidades geradas pela mecânica quântica, o mapeamento do código genético e o compartilhamento da informação em larga escala por meio da internet.

Nossa pesquisa consistiu em investigar as questões relevantes para construção de redes, tendo como objetivo à apropriação e interpretação de fluxos de informação e conhecimento. Assim, foi necessária a definição de redes sociais existentes no livro. Para este trabalho, usaremos a rede Personagem x Personagem que se trata de uma rede não dirigida composta por todos os indivíduos que são citados ao longo dos capítulos. As ligações entre esses podem se dar de diversas formas: citações diretas, contatos pessoais, profissionais e/ou acadêmicos.

4. Método

A elaboração deste trabalho foi possível a partir do desenho de um método composto dos seguintes passos:

- 1º- Selecionar a obra ou o conjunto de documentos a ser usado como fonte de dados.
- 2º- Estabelecer as relações que serão estudadas. As relações escolhidas passam a ser a

condição de escolha dos elementos que serão coletados para procedermos à construção da(s) rede(s). Consideramos a leitura da obra como condição *sine qua non* para aplicar o método proposto. O artefato de entrada deste passo é a obra ou o conjunto de documentos usado como fonte de dados.

3º- Coletar os vértices que comporão a rede (e.g. nomes dos personagens, avanços, instituições, etc.) e as relações que os conectam (e.g. circunstâncias, avanço, data, etc.). Recomendamos que o procedimento de coleta seja o mais minucioso possível. O artefato de saída é um arquivo em formato de planilha ou texto (idem para passos 4 e 5).

4º- Elaborar o vocabulário de controle para correção dos dados oriundos da coleta. Este passo é importante para eliminar as ambiguidades, identificar corretamente os vértices para não haver replicação de dados (e.g. homônimos no caso de indivíduos distintos, nomes diferentes para o mesmo indivíduo). O artefato de saída é um arquivo com a lista de etiquetas para os nomes, avanços e outros elementos que se constituirão vértices; é uma espécie de glossário.

5º- Preparar os dados para construir a(s) rede(s) a ser(em) analisada(s). Os dados serão organizados de acordo com as ferramentas computacionais com as quais as redes serão criadas.

6º- Construir a(s) rede(s) de interesse. O artefato de saída é um arquivo no formato NET (formato típico do Pajek).

7º- Analisar a(s) rede(s) de interesse à luz da teoria de redes. Não obstante a análise quantitativa, recomendamos que os pesquisadores tenham auxílio de pesquisadores com competências específicas (e.g. caso seja um estudo sobre relacionamentos políticos, recomendamos a participação de um cientista político).

5. Resultados e Discussão

Com os personagens listados ao longo da história na obra de Gleick foi construída uma rede não-dirigida com 239 vértices e 444 arestas. Esta rede (Figura 01) apresentou o fenômeno mundo pequeno (WATTS; STROGATZ, 1998).

A rede encontrada quando comparada a uma rede aleatória similar (i.e. mesmo número de vértices e grau médio) teve seu coeficiente de aglomeração muito maior ($C_{WSReal} = 0,34$ e $C_{WSAleat} = 0,015$). Ademais, consideramos o caminho mínimo médio encontrado ($L_{Real} = 5,30$) próximo do valor desta propriedade da rede aleatória equivalente ($L_{Aleat} = 4,32$).

Usando a rede social (Figura 1) como ponto de partida, foram retirados sucessivamente os vértices com os maiores graus (i.e. com as maiores centralidades de graus), especificamente são os vértices que representam os personagens Charles Babbage ($k = 36$), Alan Turing ($k = 25$), Claude Shannon ($k = 22$), Michael Faraday ($k = 21$), e calculadas as propriedades de cada rede resultante. Com a eliminação dos quatro maiores vértices os valores de grau médio e Coeficiente de Aglomeração Médio diminuem respectivamente 13,2% e 10,2% e para o Caminho Mínimo ocorre aumento de 36,2%. Essas variações sugerem que possíveis mudanças (consideremos um exercício contrafactual) na estrutura da rede podem ocasionar perda na eficiência do fluxo de informação.

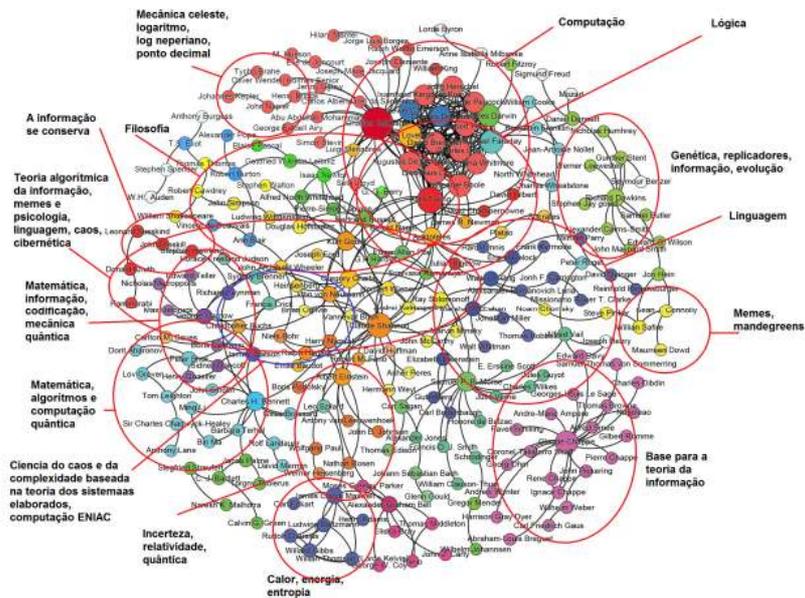


Figura 01. Rede social entre personagens citados na obra de Gleick com algumas áreas em cores.

Quanto à conectividade dos vértices, a rede de relacionamento entre personagens ao longo da história não apresentou em sua distribuição de graus uma curva característica.

6. Conclusão

O uso da teoria de redes permite que as informações sejam visualizadas e significadas de formas cognitivas diferentes, nas quais as interpretações sugerem novas possibilidades de análise sobre as relações de afinidades entre os personagens do livro de Gleick e, por outro, evidenciam correlações importantes com alguns fatos já registrados na história que foram identificados no trabalho e aqui referenciados como avanços técnico-científicos.

Neste sentido, este trabalho nos permitiu:

- A elaboração de um método baseado em análise de uma rede social histórica, constituída de cientistas que fazem parte da narrativa do livro *“The Information” - History, a Theory, a Flood* de James Gleick (2011);
- A identificação das relações de afinidades entre personagens baseado no livro;
- A criação da(s) rede(s);
- O cálculo das propriedades da(s) rede(s);
- O reconhecimento dos personagens mais importantes da(s) rede(s);
- A observação, identificação e caracterização topológica da rede;

As relações entre os personagens se configuraram pela existência de elos, ou "intermediários" comuns que ligam pelo menos duas pessoas pela convivência simultânea ou não (i.e. com ou sem sincronismo diacrônico-espacial). Isso propiciou a formação de rede mundo pequeno, com cada vez mais pessoas conectando grupos de pessoas, local ou globalmente (no espaço e no tempo).

A partir do método sugerido neste artigo, algumas constatações/reflexões resumem as consequências da análise com base nos resultados encontrados. Do ponto de vista da estrutura da rede, a retirada de um vértice impacta na rede (e.g. ponto de corte). Por outro lado, o estudo sobre o fluxo de informações extraídas de uma rede social de conhecimento histórica nos permite especular possíveis impactos: (a) a ausência da prensa de tipos móveis de Gutenberg implicaria em atraso na disseminação de conhecimento comprometendo “o Renascimento, a Reforma Protestante e o nascimento da ciência” (GLEICK, 2011, p. 327); (b) a ausência de Charles Babbage e sua máquina analítica impactaria no surgimento dos computadores; (c) sem Aristóteles, tudo decorrente da lógica estaria comprometido.

Referências

- Balancieri, R.; Bovo, A. B.; Kern, V. M.; Pacheco, R. C. S.; & Barcia, R. M. (2005). "A análise de redes de colaboração científica sob as novas tecnologias de informação e comunicação: um estudo na Plataforma Lattes", In: *Ciência da Informação*, Vol. 34(1), p.64.
- Eades, P. (1984). "A heuristic for graph drawing". *Congressus Numerantium* 42 (11): 149–160.
- Elson, D. K.; Dames, N.; & McKeown, K. R. (2010). "Extracting social networks from literary fiction". *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*. ACM Publications. pp 138-147
- Fruchterman, T. M. J. & Reingold, E. M. (1991), "Graph Drawing by Force-Directed Placement". *Software – Practice & Experience (Wiley)* 21 (11): 1129–1164, doi:10.1002/spe.4380211102.
- Gil, S.; Kuenzel, L.; & Suen, C. (2011). "Extraction and Analysis of Character Interaction Networks-From Plays and Movies". CS 224W Final Project Report Sunday, December 11, 2011
- Gleick, J. (2011). "A informação - Uma história, uma teoria, uma enxurrada". (Título original: *The Information: A History, A Theory, A Flood*). Companhia das Letras, São Paulo. ISBN 978-85-8086-711-4
- Gross, J. T. & Yellen, J. (1999). "Graph Theory and its Applications". Boca Raton: CRC Press.
- Martins, D. (2014). "Análise dinâmica de redes sociais de coparticipação em bancas de defesa de teses e dissertações: um estudo de caso a partir de múltiplos indicadores na área da Ciências da Comunicação", In: *Encontros Bibli: Revista Eletrônica De Biblioteconomia E Ciência Da Informação*, 19(40), 99-116. doi:<http://dx.doi.org/10.5007/1518-2924.2014v19n40p99>.
- Pessoa Jr., O. (2000). "Histórias contrafactuais: o surgimento da Física Quântica". *Estudos Avançados* 14 (39).
- Wasserman, S. & Faust, K. (1994). "Social network analysis: methods and applications". Cambridge: Cambridge University Press.
- Watts, D. J. & Strogatz, S. H. (1998). "Collective dynamics of 'small-world' networks" [on-line]. *Nature* 393, 440–442.