

Artigo Original

Proposta de metodologia para análise de redes sociais aplicadas a sites de saúde

Ricardo Barros Sampaio ^a, Luana P. Silva ^a, Elmo Fagner Sampaio Esteves ^a, Wagner de Jesus Martins ^a

^a Fundação Oswaldo Cruz, Diretoria Regional de Brasília (DIREB), Coordenação de Gestão e Integração Estratégica (CGIE), Núcleo de Redes, Brasília(DF), Brasil.

DOI: 10.3395/reciis.v8i1.789pt

Submetido: 15/Mar/2013

Aceito: 02/Out/2013

Resumo

O presente trabalho teve como foco principal a avaliação empírica de métodos aplicáveis no estudo de redes sociotécnicas para sites do governo e outras mídias digitais que discutem ou apresentam informações sobre a saúde. O estudo foi realizado em decorrência de atividades de pesquisa sobre redes sociais realizadas na Fiocruz Brasília. A proposta inicial era entender como estavam relacionados os diferentes ambientes informacionais da Web e se havia uma metodologia que poderia ser aplicada a esse tipo de estudo. O entendimento sobre as relações que se configuram nesses ambientes tecnológicos pôde revelar uma estrutura de contatos e fluxos informacionais diferente das que são percebidas nas redes e esferas não tecnológicas. Como conclusão do trabalho, acreditamos que a utilização de métodos de análise de redes é pertinente a esse tipo de estudo e que pode ser aplicada de forma mais extensiva e contínua com o objetivo de proporcionar uma comunicação mais eficaz entre governo e população. Entre os resultados encontrados foi notada uma clara estrutura de relacionamentos entre entidades de áreas similares de saúde, além de sites com alta centralidade ou destaque entre os demais.

Palavras-chave: Modularidade; Centralidade; Sites; Gestão, Redes Sociais, Navegador, Internet, Mídias Sociais, Indicadores de Gestão

Conflito de interesse: Os autores declararam não haver conflitos de interesse.

Fontes de financiamento: Não houve.

Introdução

A sociedade, hoje, vive em uma era em que as informações já estão disponíveis em grandes quantidades. No entanto, a tarefa com a qual nos deparamos é a de obter a capacidade de tratar e disponibilizar a informação que venha atender a necessidade informacional do usuário, neste caso a sociedade. A saúde, a cada dia que passa, vem sendo tratada com maior interdisciplinaridade. Não adianta tratar apenas a saúde pontual do cidadão, mas todo o contexto em que o mesmo se encontra inserido, como cultura, meio ambiente e uma diversidade de outros fatores.

Para que seja possível possa tratar as informações acerca da saúde e do Sistema Único de Saúde (SUS) é necessário a utilização de novas ferramentas e métodos que facilitem o planejamento de políticas públicas e promovam a gestão inovadora nos sistemas públicos de saúde¹.

Os métodos e as possibilidades de se estudar as relações existentes entre indivíduos e entidades teve uma mudança significativa nos últimos anos, com a crescente utilização dos meios tecnológicos para a comunicação e difusão do conhecimento. A cada dia mais organizações se valem de diferentes mecanismos para divulgar as informações que são importantes para o seu público. Ser e estar acessível é uma questão importante para os órgãos do governo que devem se comunicar com a população.

O presente trabalho teve como objetivo inicial o estudo de uma metodologia para avaliação de ambientes da Web e seus respectivos relacionamentos. Essa demanda de pesquisa foi decorrente de atividades realizadas na Fiocruz Brasília, na Coordenação de Gestão e Integração Estratégica, durante o período entre 2012 e 2013. Foram analisadas redes que utilizam as tecnologias de informação e comunicação para veiculação e discussão de conteúdos relacionados à saúde, além de sites de informação e de serviço do governo na área de saúde. O segundo objetivo do trabalho foi entender como a atual conformação dos diferentes ambientes informacionais da saúde estão dispostos e os resultados empíricos desta análise. Não foi objeto do presente trabalho a análise qualitativa dos diferentes ambientes, mas sim a estrutura de relacionamentos baseada nos conceitos de análise de redes.

A justificativa para um trabalho com esse viés se deve ao grande número de fóruns, blogs e sites que circulam informação, e que são associados ao governo, sem um controle formal sobre o que está sendo ou não discutido e por quais entidades. O termo controle não é empregado, aqui, no sentido de controlar o fluxo informacional, mas sim de propiciar uma melhor interação entre os diferentes ambientes informacionais, tornando-os mais eficazes na sua proposta de atender à sociedade.

Entender a forma como as entidades se relacionam nos meios tecnológicos, utilizando as mídias sociais, através dos métodos de Análise de Redes Sociais (ARS), permite trazer informações que nos meios clássicos de pesquisa não seriam possíveis. A ARS compreende um extenso conjunto de métodos para avaliação da estrutura de grupos sociais^{2,3}. A utilização da ARS possibilita percepção diferenciada de um ambiente social. Com isso, novos fenômenos podem ser investigados e novas hipóteses podem ser estabelecidas.

Fundamentação teórica

Os organismos ou sistemas nos seus diversos tipos podem ser representados por redes tecnológicas, biológicas, sociais ou informacionais e, em geral, possuem um alto grau de complexidade no que diz respeito aos seus relacionamentos. As redes sociotécnicas são assim definidas por não se tratar de uma rede de computadores ou de apenas um conjunto de pessoas, mas de uma interconexão de seres humanos por meio da tecnologia⁴. A complexidade faz parte da arquitetura da maior parte das redes estudadas nos dias de hoje e é inerente às redes sociotécnicas que, em sua grande maioria, possuem um comportamento de redes complexas^{5,6}. As redes complexas têm como uma de suas características a formação de pequenos grupos ou de atores que possuem um alto número de relações com os demais, enquanto a grande maioria dos demais elementos da rede social possuem um número reduzido de relacionamentos.

Na área de saúde, vários são os atores ou entidades que podem ser analisados utilizando o conceito de redes. Entender os padrões e os relacionamentos existentes e em formação entre esses organismos é um dos objetivos deste estudo.

Segundo Fleury e Ouverney:

Não há concordância entre os estudiosos em relação ao poder teórico do conceito de redes: alguns o utilizam como metáfora para demonstrar que as políticas públicas envolvem multiplicidade de atores diversos; outros o vêem como uma ferramenta analítica valiosa para o estudo das relações entre atores e poder público; outros, ainda, consideram-no um método de análise da estrutura social⁷.

Uma rede social consiste em um ou mais conjuntos finitos de atores (vértices) e das relações definidas entre eles. Esses atores podem ser tanto pessoas ou empresas, analisadas como unidades individuais, como podem ser unidades sociais coletivas. O

conceito de ator é flexível, permitindo diferentes níveis de agregação, o que possibilita sua adequação a diferentes problemas de pesquisa³.

Uma rede é uma representação simplificada que reduz um sistema a uma estrutura abstrata composta, em grande parte, por padrões de relacionamentos existentes entre os atores. Os vértices que representam os atores e as linhas representando os relacionamentos, podem ser caracterizados por meio da utilização de diferentes tipos de informações, como nomes, tamanho ou espessura para que detalhes sobre o sistema sejam avaliados. No entanto, mesmo com essas informações extras, muito ainda pode ser perdido no processo de reduzir um sistema completo a uma representação de rede⁸. Ao analisarmos uma rede, podemos ver os relacionamentos existentes de forma mais clara; no entanto, nem todos os aspectos envolvidos na relação podem ser observados.

De acordo com Marteleto, “estudar a informação através das redes sociais significa considerar as relações de poder que advêm de uma organização não-hierárquica e espontânea e procurar entender até que ponto a dinâmica do conhecimento e da informação interfere nesse processo”⁹.

Para o presente estudo foram utilizadas algumas métricas na análise da rede. Entre essas, temos a centralidade que se baseia na definição do vértice mais “central” ou “importante” da rede. A centralidade de um vértice pode ser definida como o relacionamento deste com outros vértices, sendo que os mais proeminentes são aqueles extensivamente envolvidos com os demais. Estudos sobre a centralidade tiveram início em 1948 mas, em 1978, foi realizado um dos primeiros para o entendimento da estrutura de grupos^{10,11}.

Outro método utilizado para análise da rede foi o PageRank, que recebeu esse nome graças ao cientista que o criou, Lawrence Page. O PageRank pode ser descrito como um modelo de representação do comportamento humano de busca na internet¹². O cálculo do PageRank leva em consideração a estrutura de links da Web, ou seja, identifica quais deles fazem referência a outros e quais são esses outros, bem como o peso do link que os referencia, o tráfego de navegação e a estimativa de busca. As páginas da Web com maior número de links apontando para elas teriam, por exemplo, um maior PageRank do que aquelas com número reduzido.

Por último, foi utilizado o método de modularidade, que tem como objetivo separar os vértices de uma rede em módulos com base nos tipos de relacionamentos existentes. Esse método nos permite avaliar os agrupamentos existentes em uma rede muito densa, e tem recebido considerável atenção que se traduz na detecção e caracterização da estrutura de comunidades em redes¹³.

Metodologia

A metodologia de pesquisa utilizada para a busca dos ambientes informacionais sobre saúde se deu com base na necessidade de um melhor entendimento acerca do tipo de informação e de discussões que estavam em curso naquele momento. A princípio se queria entender em quais locais e como se davam as trocas de informações sobre saúde na Web.

Foram escolhidas três plataformas para realizar a busca de grupos que discutem saúde na internet. A pesquisa foi realizada nos três ambientes virtuais das redes sociais mais acessadas no Brasil em setembro de 2012, de acordo com o Ranking Alexaⁱ. Neste período, o Facebook, plataforma fundada em 2004, se tornou não apenas a rede social mais popular, mas também o site mais visitado do Brasil. O Twitter foi a segunda rede social mais popular no país, alcançando a 11ª posição no ranking geral. A terceira plataforma com maior popularidade era o Orkut¹⁴.

No caso da rede de relacionamentos Orkut, apesar de terem sido encontrados mil grupos relacionados à saúde, foram descartados, pois não possuíam, em sua maioria, espaços em que estivessem sendo discutidas questões de saúde, mas ambientes tomados por spams com caráter de marketing. Foram encontrados 2.289 perfis do Twitter que também não foram utilizados em nossa análise por não apresentarem fóruns que fossem passíveis de análise sobre os assuntos discutidos, nem mesmo grupos com representantes fixos.

O Facebook poderia ter sido a ferramenta utilizada, caso possuísse maior abertura em relação aos dados armazenados no site. No entanto, além da dificuldade com o motor de busca da plataforma, que está em processo de modificação, diversos grupos seriam excluídos da análise por serem fechados a agentes externos. Foram encontrados, nessa plataforma, 62 fanpages (páginas de empresas e instituições utilizadas para divulgação de produtos, serviços e informações úteis a seus seguidores) e 334 grupos de discussão com o termo “saúde”.

ⁱ Alexa Internet, Inc. é um serviço pertencente à empresa Amazon que mede as visitas diárias em um site e a quantidade de visualizações mensais desse mesmo site e produz um Ranking global, por país e por categoria.

Após a busca inicial, optamos por mapear uma rede de sites que abordam a temática saúde na Web. A coleta dos dados foi realizada manualmente a partir de dois: o site do Fórum dos Conselhos Federais da Área de Saúde (FCFAS) e o do Portal do Ministério da Saúde. A raiz das buscas se deu a partir dos sites oficiais de cada órgão presente no organograma do Ministério e dos links externos encontrados no site do FCFAS. Para cada novo link encontrado, uma nova raiz foi criada, formando um banco de dados de sites da saúde e suas referências. Os sites que não abordavam assuntos específicos de saúde ou que não fossem de conteúdo relacionado ao Brasil não foram utilizados como raízes.

A pesquisa resultou em um total de 1.455 sites que de alguma forma se relacionavam com saúde, seja por serem sites de conselhos profissionais da área, sites governamentais da saúde, ou páginas de faculdades para ofertas de cursos na área. A busca não automatizada gerou um trabalho manual de grande volume, no entanto teve como vantagem a possibilidade de avaliar os links que estavam sendo referenciados e a sua consistência com assuntos relacionados à saúde. Mecanismos de busca automáticos não conseguiriam fazer essa distinção que nos foi necessária.

Para fins de análise, foram utilizados todos os sites que tinham o domínio “.gov”, bem como as relações existentes entre eles. O motivo desta escolha se deve ao fato da grande quantidade de sites encontrados e da necessidade de se fazer um recorte para entender o método de pesquisa que estava sendo empregado para a análise. Neste momento, decidimos por nos concentrar na aplicabilidade da pesquisa e na descrição dos métodos utilizados.

O passo seguinte da análise foi o tratamento dos dados utilizando planilha Excel. Apesar do grande volume de sites pesquisados e suas respectivas linhas de referenciamento, esta nos serviu para o agrupamento, limpeza e codificação dos sites. Para um número maior de sites se faz necessária a utilização de uma ferramenta mais robusta de tratamento de dados.

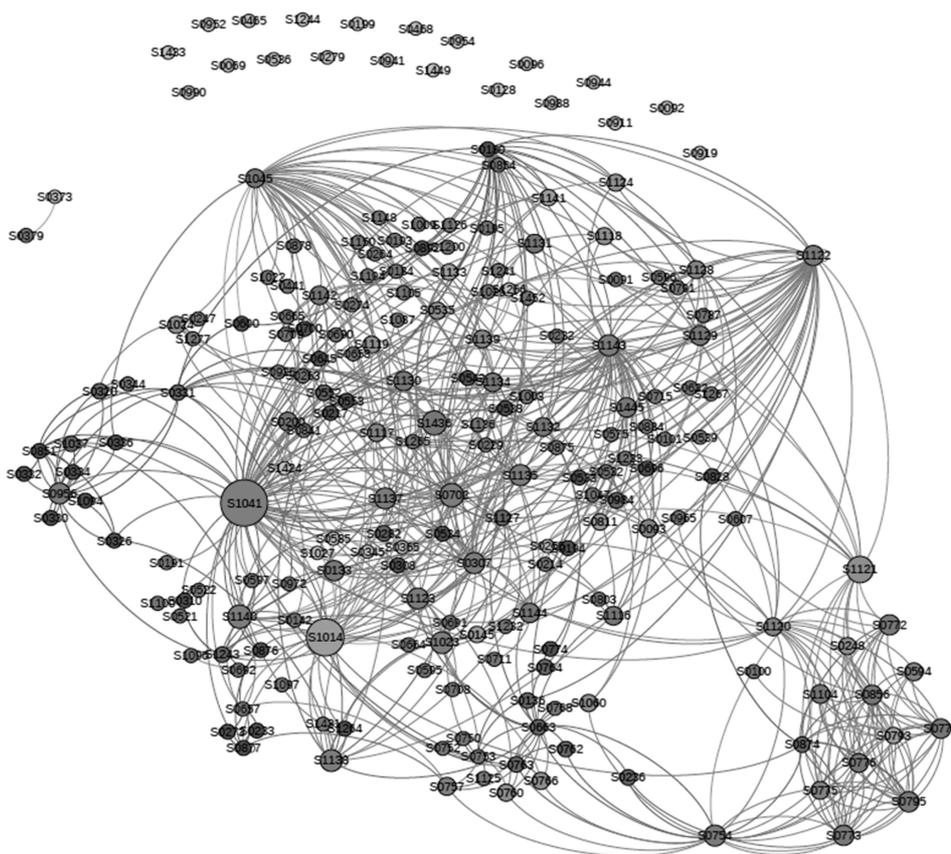
A análise dos relacionamentos e redes foi realizada através do software Gephi. Com a ferramenta foi possível a manipulação dos vértices na rede para melhor visualização destes. Além da distribuição, a ferramenta proporciona uma avaliação de várias métricas, o que nos permitiu obter um maior entendimento da rede e dos vértices que estavam presentes. Conceitos como centralidade com foco no grau de entrada, grau de saída, *closeness* e *betweness* foram avaliados. Outras medidas, como PageRank e Modularidade, também foram utilizadas.

Resultados

Entre os sites pesquisados, apenas os que continham .gov em seu domínio é que foram objeto de demonstração na discussão dos resultados. A primeira análise feita na rede tinha como objeto a centralidade com base no grau de entrada do vértice. Conforme mencionado anteriormente, a centralidade demonstra o melhor “posicionamento” do vértice na rede. Para uma melhor visualização, foram utilizados dois parâmetros para o mesmo método: o primeiro, a divisão por cores (graus de cinza) e o outro a divisão por tamanho. Ou seja, os vértices mais centrais no que diz respeito ao grau de entrada tinham o maior tamanho e uma cor diferenciada das demais. Para a visualização, foi utilizada uma distribuição denominada Force Atlas 2 do software Gephi, que é adaptado para uma avaliação mais qualitativa dos vértices. Portanto, a Figura 1 demonstra quais os sites que recebem mais referências de outros sites. Como foram analisados somente os sites .gov, alguns dos vértices representados como os da parte superior da figura não tiveram nenhuma referência de outros sites .gov ou mesmo referenciaram outros sites. Já o vértice S1041 teve o maior número de referências de outros sites, o que faz com que ele seja um dos mais “citados” ou “importantes” nessa perspectiva.

Outro ponto importante na visualização desse grafo é a direcionalidade das linhas. Pelo fato de as setas não estarem evidentes, alguns vértices com muitas linhas acabam não tendo um grau de entrada muito grande, como é o caso do S1045 e do S1122, nos cantos superiores esquerdo e direito respectivamente. Se forem verificados esses mesmos vértices na Figura 3 é notada a sua diferença em tamanho, demonstrando o grau de saída.

Figura 1
Grau de Entrada



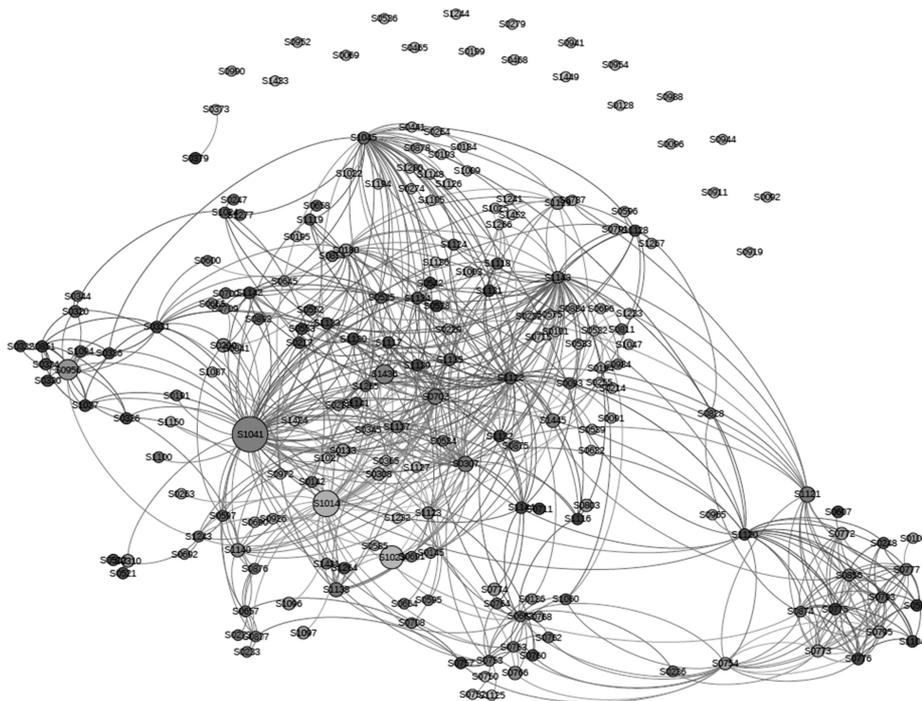
A Tabela 1 demonstra os nove sites com maior grau de entrada, além dos respectivos valores para o grau de saída, PageRank e *Modularity* que serão discutidos adiante. Pela tabela, podemos ver que o vértice S1041 é o site do Ministério da Saúde e que possui 43 referências dos demais vértices ou sites pesquisados.

Tabela 1

Id	Rótulo	Grau de Entrada	Grau de Saída	PageRank	ModularityClass
S1041	portalsaude.saude.gov.br	43	35	0,0470	3
S1014	portal.anvisa.gov.br	31	6	0,0311	4
S1121	saude.ce.gov.br	17	0	0,0119	2
S1436	wdatasus.gov.br	14	0	0,0181	4
S0702	funasa.gov.br	13	27	0,0117	4
S1140	saude.rs.gov.br	13	5	0,0085	4
S1138	saude.rn.gov.br	11	7	0,0098	5
S1023	portal.saude.gov.br	11	0	0,0254	5
S1143	saude.sp.gov.br	10	46	0,0059	6

Na Figura 2, é avaliado o PageRank dos vértices sob a mesma distribuição Force Atlas 2. É possível notar que a disposição e tamanho dos vértices são bastante próximos da medida de centralidade por grau de entrada. Como o PageRank se baseia no comportamento de busca dos usuários, é de se esperar que os sites mais referenciados sejam também os sites mais procurados pelos usuários e por isso a sua similaridade.

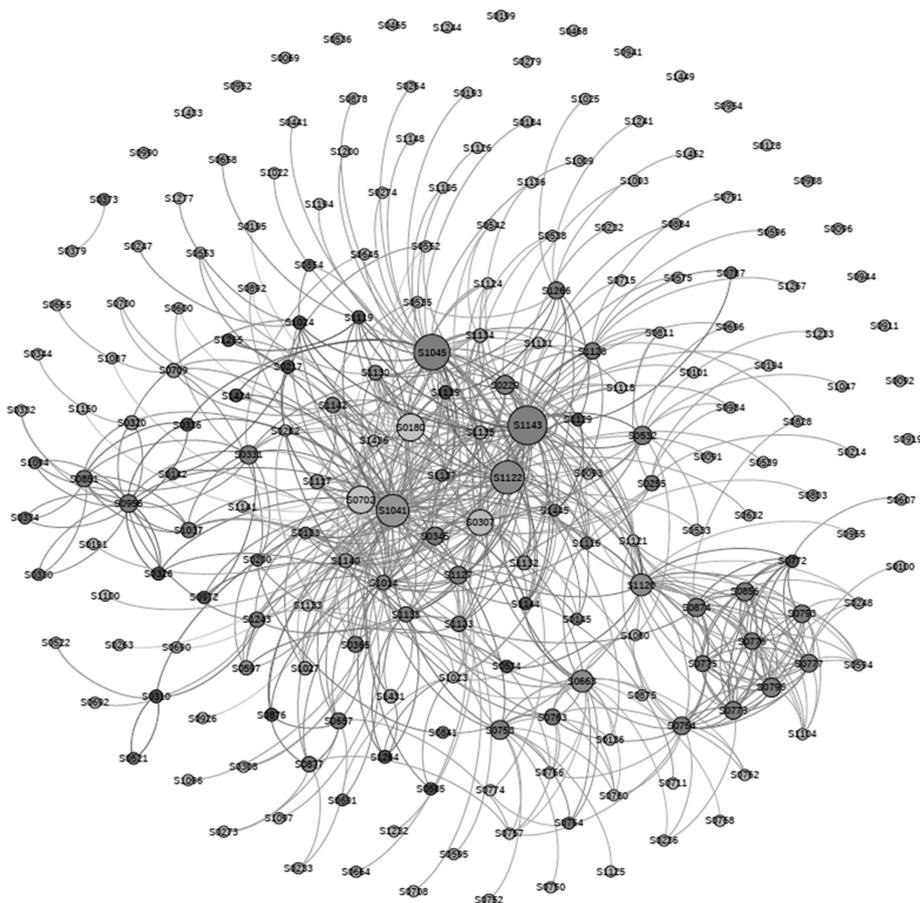
Figura 2
Page Rank



No entanto, vale ressaltar que alguns sites com alto grau de entrada não necessariamente têm um alto PageRank, como é o caso do vértice S1140 e S1138 que estão no canto inferior esquerdo. Para entender essa diferença se faz necessário um estudo mais aprofundado dos tipos de sites que os referenciam e quais as suas proximidades com sites ditos “importantes” na rede.

Na Figura 3, temos a análise da rede, levando em consideração o grau de saída dos vértices. Ou seja, os sites que mais referenciam outros sites. Veja que a diferença para os sites mais referenciados é muito grande. Para essa visualização, foi utilizada a distribuição Fruchterman Rheingold, da ferramenta Gephi, que simula o grafo como um sistema de partículas. Pelo grafo podemos notar uma grande quantidade de vértices que não fazem nenhuma referência a outros, são vértices mais periféricos. O que pode ser uma característica dos sites do governo, que têm como objetivo o repasse de informações e não necessariamente a referência a outros sites. Pela figura, podemos notar também um forte referenciamento entre sites “bem relacionados”, que pode ser visto na parte inferior direita do grafo.

Figura 3
Grau de Saída



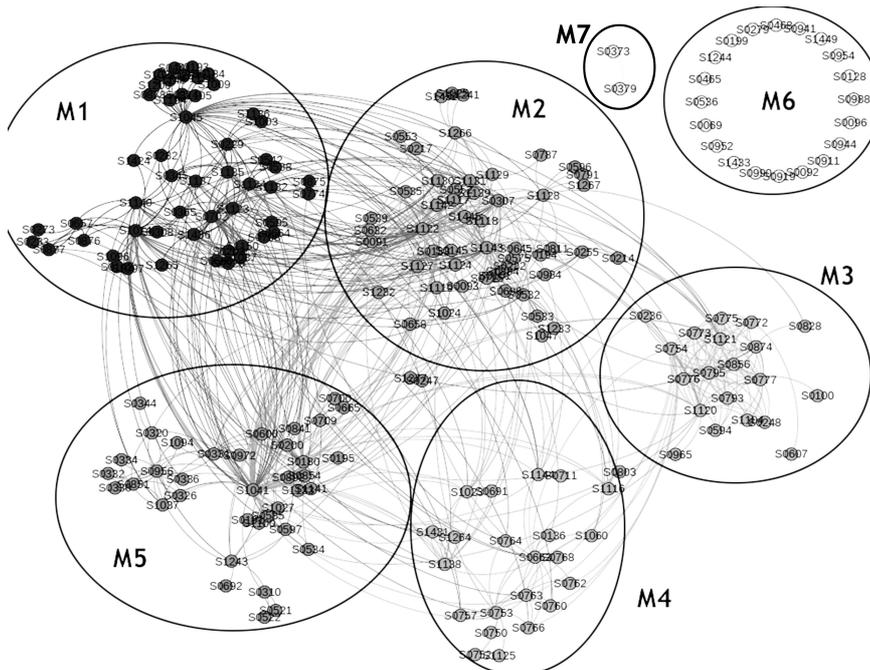
A Tabela 2 nos mostra os nove sites que mais referenciam outros e podemos notar uma participação de sites de governos estaduais de São Paulo, Santa Catarina, Distrito Federal e Bahia, identificados como S1143, S1045, S1122 e S1120 respectivamente.

Tabela 2

Id	Rótulo	Grau de Entrada	Grau de Saída	PageRank	ModularityClass
S1143	saude.sp.gov.br	10	46	0,0059	6
S1045	portales.saude.sc.gov.br	7	41	0,0059	4
S1122	saude.df.gov.br	9	37	0,0056	6
S1041	portalsaude.saude.gov.br	43	35	0,0470	3
S0702	funasa.gov.br	13	27	0,0117	4
S0180	bibliosus.saude.gov.br	2	27	0,0064	3
S0307	conselho.saude.gov.br	10	24	0,0119	6
S1120	saude.ba.gov.br	6	18	0,0049	2
S0663	fhb.df.gov.br	3	18	0,0039	5

No quarto e último grafo, representado pela Figura 4, foi trabalhado o método de Modularity ou modularidade. Para um maior aprofundamento neste método, é necessária uma análise qualitativa dos sites concentrados em cada módulo, de modo a entender o motivo pelo qual esses vértices se posicionaram dessa maneira. Para a distribuição, foi trabalhado individualmente cada módulo, utilizando o Force Atlas 2 para que ficassem claros os seus agrupamentos.

Figura 4
Modularidade



Uma das avaliações mais claras que temos sobre os módulos é a concentração em módulos de conselhos, bibliotecas, e secretarias. No entanto, mesmo dentro dos módulos, existe um número maior de sites e se faz necessária uma análise mais qualitativa do seu conteúdo e a sua contextualização.

Discussão

O foco principal deste trabalho é o estudo dos métodos aplicados na análise e não necessariamente o entendimento dos resultados obtidos com base nos sites pesquisados. Para que ele tivesse o viés de análise dos sites de forma pontual, um trabalho mais qualitativo sobre o conteúdo dos mesmos deveria ser realizado e também ser objeto de discussão neste espaço. O que buscamos, aqui, é demonstrar as possibilidades de uma análise de sites da Web utilizando os métodos de ARS.

Apesar de existirem bilhões de documentos na Web e da ideia de que podemos acessar qualquer um desses documentos facilmente, nem todas as páginas são interconectadas, dificultando o alcance a determinados conteúdos. Partindo de uma página qualquer, poderemos acessar apenas 24% do repertório existente na Web.

Essa limitação existe porque os links são direcionados, ou seja, podemos trafegar apenas em uma direção. Se temos um trajeto $A > B > C > D$, podemos sair de A e chegar a D, porém não é possível percorrer o caminho inverso. Este caminho seria possível se a rede não fosse direcionada, ou se as relações de direção fossem recíprocas.

As consequências de uma rede direcionada no contexto da Web foram estudadas pela primeira vez por Andrei Broder, do Alta Vista e seus colaboradores da IBM e da Compaq¹⁵. O estudo foi realizado com uma amostra de 200 milhões de nós, o que representava, na época, quase metade de todas as Webpages existentes. E a principal avaliação da consequência de direcionalidade dos links foi que a Web não constitui uma rede homogênea, mas é composta por quatro grupos, que os autores denominaram “grandes continentes”, com características diferentes quanto à navegabilidade.

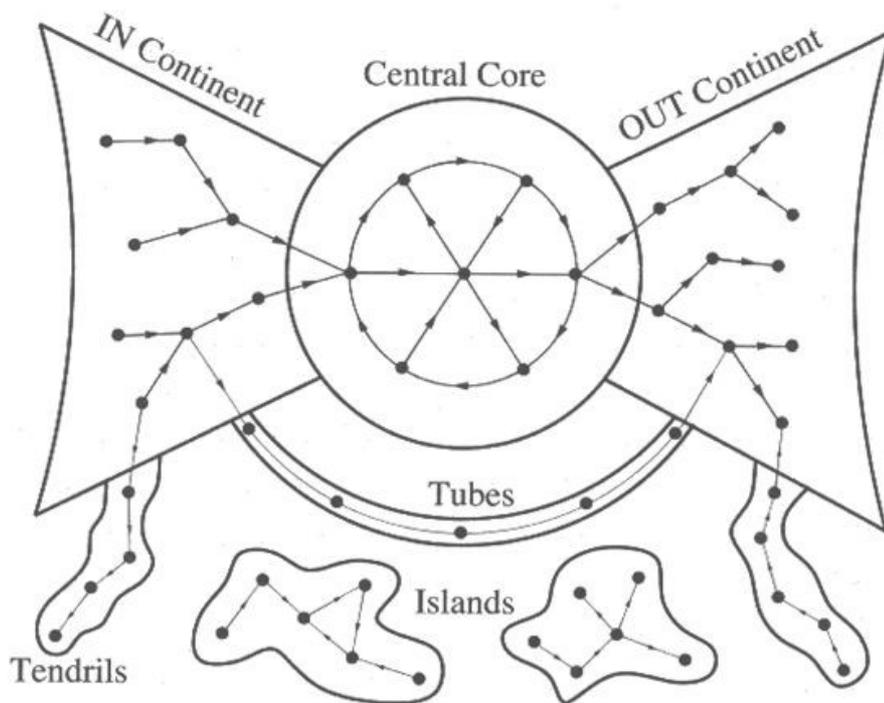
Barabási, em seu livro *Linked*, descreve a Web e fala de núcleos com base no trabalho de Broder, e os apresenta conforme a Figura 5¹⁶. O núcleo central, contém cerca de um quarto de todas as páginas da Web. Esse continente é facilmente navegável, pois todos os nós são conectados entre si. O segundo e terceiro continentes, interior e exterior, também possuem grande extensão, porém têm atributos que dificultam a navegação. Iniciando a navegação por páginas do continente interior, pode-se chegar facilmente ao núcleo central, mas não há como retornar à origem. Ao contrário do que ocorre com os nós pertencentes ao continente exterior, alcançados a partir do núcleo central. Mas, uma vez fora deste núcleo, não há links que o tragam de volta.

O último quarto dos documentos encontra-se em filamentos ou ilhas desconectadas, que são grupos interconectados, mas que não possuem ligação com o núcleo central, são inacessíveis a partir do núcleo e não possuem links de volta para ele.

Esse modelo nos mostra que a limitação de nossa navegabilidade está relacionada com a página escolhida como ponto de partida. Se partirmos do núcleo central, não poderemos acessar metade das páginas existentes na Web. Iniciando pelo continente interior poderemos acessar grande parte das páginas existentes, e iniciando pelo território exterior chegaremos rapidamente a um fim.

Figura 5

Os continentes de uma rede direcionada



Fonte: Barabási, 2003

No que diz respeito ao conceito de centralidade outras medidas também são possíveis, como o conceito de proximidade ou de intermediação (*closeness e betweenness*); no entanto, apenas o grau de entrada e saída estão sendo discutidos. No caso dessas centralidades, o que podemos observar é que o número de referências recebidas ou realizadas pode estar relacionado com a característica e o contexto do órgão responsável pelo site. Esse tipo de estudo pode influenciar a montagem de um site informacional ou de serviços dentro de um órgão, ou até mesmo as políticas públicas para confecção de um site, uma vez que as informações ali contidas têm como objetivo atender a necessidade que a população tem de informações ou de serviços. Os sites que obtiveram maior referenciamento o conseguiram pelo seu grau de importância dentro do governo ou pelo tipo de serviço que proporcionam? Esta seria uma pergunta a ser analisada.

Quanto ao fator PageRank, é importante que essa medida faça parte das análises realizadas para que se possa observar qual o comportamento dos usuários dentro de uma rede. Com base nos relacionamentos, podemos observar onde os usuários acabariam chegando se estivessem fazendo buscas com base nos links existentes entre as páginas da Web. O entendimento desse caminho pode fazer com que sites que tenham certo tipo de informação possam ser ajustados. Essa é uma análise e um ajuste com base nos relacionamentos em formação que podem modificar muito rapidamente com o passar do tempo.

Por último, temos o conceito de modularidade, que traz um visão diferenciada sobre o comportamento dos sites e seus relacionamentos. Em alguns casos esses relacionamentos podem ser bastante óbvios como, por exemplo, quando notamos que todos os hemocentros ou os conselhos federais tinham tendência a se agrupar. No entanto, existem outros relacionamentos que são formados e que podem trazer informações que não são conhecidas, como, por exemplo, saber que as bibliotecas virtuais têm um forte relacionamento com sites mais informacionais e que não oferecem serviços.

O estudo dos relacionamentos entre os sites é apenas uma forma de se avaliar o fluxo ou acesso informacional, mas um trabalho bem estruturado nessa área pode trazer enormes benefícios para a população. Outro fator importante deste estudo é a realidade que ele transmite quanto ao seus contatos e objetivos. Quem está referenciando quem e por quê? Quais as informações mais relevantes e quem as fornece? É objetivo desta ou daquela instituição disponibilizar essas informações ou as de outro órgão?

Conclusões

Este trabalho de pesquisa teve como ponto focal a utilização e descrição de uma metodologia para avaliação dos relacionamentos entre sites do governo com o objetivo inicial de apoiar gestores da saúde no que diz respeito aos ambientes informacionais. Apesar da aplicação prática desta pesquisa no ambiente da Fiocruz Brasília ter gerado uma série de desdobramentos, para este artigo tivemos como premissa descrever os elementos e métodos utilizados para que fosse possível a sua replicação em uma escala maior de ambientes informacionais.

Acerca dos resultados obtidos, podemos concluir que a aplicação de métodos sobre análise de redes em sites da Web é positiva e pode servir de baliza para tomadas de decisão acerca de processos de comunicação e interação do governo com a população de maneira mais eficiente. Existem grupos claramente definidos (homocentros, atenção básica e outros) que abordam os diferentes tópicos da saúde, e saber tratar esses grupos de acordo com as necessidades informacionais da população é um grande avanço. O entendimento sobre a centralidade de certos sites da saúde também é visto como positivo, pois pode apoiar decisões sobre maior ou menor concentração de informações.

No que diz respeito aos sites pesquisados e sua conformação em rede, podemos dizer que existe uma coesão entre os diferentes atores, vista por meio da análise dos relacionamentos, o que demonstra a realização de trabalhos de forma conjunta. No entanto, existe claramente uma duplicação de informações nos diferentes sites do governo, mas não sabemos dizer se isso é ou não positivo para a população.

A pesquisa teve alguns problemas que pretendemos relatar como forma de trazer um registro para futuros trabalhos. O mapeamento dos sites foi realizado manualmente, o que impediu a apresentação de resultados quantitativos que correspondessem à realidade da dinâmica dos sites de saúde na Web. Apesar de terem sido encontrados 1.455 sites, a busca foi realizada apenas nas páginas iniciais ou na página de links. Portais com número muito grande de páginas podem ter links em outras páginas e não só nas que foram avaliadas. A utilização de um software para automatizar a coleta de informações, adaptado às demandas dessa pesquisa, pode fornecer dados mais precisos para análises futuras.

Outro aspecto para futuros trabalhos é a análise qualitativa dos sites. Apesar deles possuírem muitos links a eles direcionados ou eles próprios direcionarem a muitas outras páginas, o seu conteúdo não foi avaliado, o que se fosse feito poderia acrescentar informações relevantes e subsidiar um melhor entendimento sobre os relacionamentos existentes.

O mapeamento e a análise das redes são instrumentos para a gestão de política pública e inteligência, que auxiliam na tomada de decisão e organização de serviço. Para que os serviços de saúde sejam mais acessíveis a todos, o planejamento deve conhecer a distribuição de serviços e dos usuários¹⁷.

A rede mapeada para análise pode ser entendida como uma foto em um determinado momento, mas a aplicação da análise de redes deve transpor essa visão. O desenvolvimento de tecnologias que apresentem ferramentas que possam disponibilizar em tempo real a situação de sites, fóruns e outras redes é uma possibilidade de fácil implementação e de grande impacto, caso se faça um bom uso das análises e dos resultados que podem ser obtidos.

Referências

1. Vasconcellos AG, Morel CM. Enabling policy planning and innovation management through patent information and co-authorship network analyses: a study of tuberculosis in Brazil. *PLoS ONE*. 2012; 7 (10). doi:10.1371/journal.pone.0045569.
2. Nooy W, Mrvar A, Batagelj V. *Exploratory social network analysis with pajek*. Expanded edition. Cambridge (UK): Cambridge University Press; 2011.
3. Wasserman S, Faust K. *Social network analysis: methods and applications*. Cambridge (UK): Cambridge University Press; 1994.
4. Cebrián JL. *A rede: como nossas vidas serão transformadas pelos novos meios de comunicação*. São Paulo: Summus Editoria; 1999.
5. Dorogovtsev SN, Goltsev AV, Mendes JFF. *Critical phenomena in complex networks*. Ithaca, NY: Cornell University Library, 2008. doi:10.1103/RevModPhys.80.1275. Disponível em: <<http://arxiv.org/abs/0705.0010>>.
6. Barabási AL, Réka, A. Emergence of Scaling in Random Networks. *Science* 1999. Oct, 15; 286(5439): 509–512. doi:10.1126/science.286.5439.509.
7. Fleury S, Ouverney AM. *Gestão de redes: a estratégia de regionalização da política de saúde*. Rio de Janeiro: Editora FGV; 2007.

8. Newman MEJ. Networks: an introduction. Oxford (UK): Oxford University Press; 2010.
9. Marteleto RM. Análise de redes sociais – aplicação nos estudos de transferência da informação. Ci. Info. 2001; 30(1). Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/226>>.
10. Bavelas A. A Mathematical model for group structures. Human Organization. 1948; 7: 16–30.
11. Freeman LC. Centrality in social networks conceptual clarification. Social networks. 1979; 1(3): 215–239.
12. Page L, Brin S, Motwani R, Winograd T. The PageRank citation ranking: bringing order to the Web. Technical Report. 1999. Disponível em: <<http://ilpubs.stanford.edu:8090/422/>>.
13. Newman MEJ. Modularity and community structure in networks. Proc. Natl. Acad. Sci. U.S.A. 2006; 103(23): 8577–8582. doi:10.1073/pnas.0601602103.
14. Matsuki E. Saiba quais são as cinco redes sociais mais acessadas do Brasil. [2012 set 06]. In: Empresa Brasileira de Comunicação [Internet]. Brasília: EBC, Disponível em: <<http://www.ebc.com.br/tecnologia/2012/09/saiba-quais-sao-as-cinco-redes-sociais-mais-acessadas-do-brasil>>. Acesso em: mar. 2014.
15. Broder A, Kumar R, Maghoul F, Raghavan P, Rajagopalan S, Stata R, Tomkis A, Wiener J. Graph structure in the Web. Comput. Netw. 2000. Jun; 33(1-6): 309–320. doi:10.1016/S1389-1286(00)00083-9.
16. Barabási AL. Linked: how everything is connected to everything else and what it means. New York: Plume; 2003.
17. Oliveira EXG, Carvalho MS, Travassos C. The territorial basis of the Brazilian National Health System: mapping hospital networks. Cad Saúde Pública. 2004; 20(2): 386–402. doi:10.1590/S0102-311X2004000200006.