

COLEÇÃO

Cibercultura

LAB
404

INTRODUÇÃO À ANÁLISE DE REDES SOCIAIS ONLINE

Raquel Recuero



INTRODUÇÃO À ANÁLISE DE REDES SOCIAIS ONLINE

Raquel Recuero

UNIVERSIDADE FEDERAL DA BAHIA

Reitor

João Carlos Salles Pires da Silva

Vice-reitor

Paulo Cesar Miguez de Oliveira

Assistente do Vice-reitor

Paulo Costa Lima



EDITORA DA UNIVERSIDADE
FEDERAL DA BAHIA

Diretora

Flávia Goulart Mota Garcia Rosa

CONSELHO EDITORIAL

Alberto Brum Novaes

Ângelo Szaniecki Perret Serpa

Caiuby Alves da Costa

Charbel Niño El Hani

Cleise Furtado Mendes

Evelina de Carvalho Sá Hoisel

José Teixeira Cavalcante Filho

Maria do Carmo Soares Freitas

Maria Vidal de Negreiros Camargo

COLEÇÃO 
Ciberultura 

EDITORES

Editor

Prof. Dr. André Lemos

Editor Associado

Prof. Dr. José Carlos Ribeiro

Editor Científico

Prof. Dr. Edson Dalmonte

COMISSÃO EDITORIAL

Adriana Amaral, UNISINOS

Alex Primo, UFRGS

Eduardo de Jesus, PUC-MG

Eduardo Pellanda, PUC-RS

Eduardo Vizer, Unisinos

Fabio Duarte, PUC-PR

Fabio Fernandes, PUC-SP

Fabio Malini, UFES

Fatima Regis, UERJ

Fernanda Bruno, UFRJ

Fernando Firmino, UEPB

Gisele Beiguelman, USP

Jamil Marques, UFC

Lidia Oliveira, UA - PT

Lucia Santaella, PUC-SP

Luis Adolfo Andrade, UNEB

Lynn Alves, UNEB

Macello Medeiros, UFRB

Marco Silva - UERJ

Marco Toledo Bastos, USP

Marcos Palacios, UFBA

Massimo di Felice, USP

Nelson Pretto, UFBA

Paulo Serra - Beira Interior - PT

Raquel Recuero, UCPEL

Rob Shields, University of Alberta, Canadá

Rodrigo Firmino, PUC-PR

Sandra Montardo, FEEVALE

Sandra Rubia, UFSM

Sergio Amadeu - UFBA

Simone Pereira de Sá, UFF

Sueli Fragoso, UNISINOS

Vinicius Andrade Pereira, UERJ/

INTRODUÇÃO À ANÁLISE DE REDES SOCIAIS ONLINE

Raquel Recuero

2017, Raquel Recuero

Direitos para esta edição cedidos à [Edufba](#).

Feito o Depósito Legal.

Grafia atualizada conforme o Acordo Ortográfico da Língua Portuguesa de 1990, em vigor no Brasil desde 2009.

Projeto Gráfico

Amanda Lauton Carrilho

Editoração e finalização

Josias Almeida Jr.

Ilustrações

Raquel Recuero

Revisão

Lia Nery

Normalização

Jade Santos

Sistema de Bibliotecas – UFBA

R311 Recuero, Raquel

Introdução à análise de redes sociais / Raquel Recuero. – Salvador: EDUFBA, 2017.

80p.: (Coleção Cibercultura)

ISBN 978-85-232-1669-6

1. Redes sociais on line x análise. 2. Mídia social x pesquisa. 3. Comunicação x internet. 4. Comunicação x pesquisa. 5. Computadores e civilização. I. Título.

CDU 316.472.4: 004.738.5

Editora filiada à



Editora da UFBA

Rua Barão de Jeremoabo

s/n - Campus de Ondina

40170-115 - Salvador - Bahia

Tel.: +55 71 3283-6164

Fax: +55 71 3283-6160

<http://www.edufba.ufba.br/>

edufba@ufba.br

Sumário

Introdução: sobre o que é este livro e para quem se destina?	7
O que é a Análise de Redes?	9
Quais são os principais conceitos de ARS?	21
Quais são as principais métricas da Análise de Redes?	39
Como coletar, analisar e visualizar dados para Análise de Redes?	57
Conclusão: e então, o que vimos neste livro?	75
Referências	77

Introdução: sobre o que é este livro e para quem se destina?

A Análise de Redes Sociais (ARS) constitui-se em uma abordagem de pesquisa cuja popularidade tem aumentado nos últimos anos, de modo particular, entre os pesquisadores da área de Comunicação. Com a disponibilidade cada vez maior de dados de interações sociais nas ferramentas de mídia social e, de modo particular, com a publicação das estruturas sociais nessas ferramentas, a ARS (juntamente com a representação dessas redes por grafos) e o estudo dessas estruturas passaram a receber mais atenção. É nesse âmbito que várias obras, entre artigos e livros, vêm surgindo e introduzindo o estudo dessas estruturas a partir da análise de redes e da compreensão da representação dessas redes sociais na internet.¹

Este livro é uma pequena compilação dos principais conceitos e elementos para a compreensão e a aplicação da ARS. Destina-se, principalmente, a estudantes que desejam começar seus trabalhos e, de modo particular (embora não exclusivo), àqueles que desejam fazer pesquisas com dados de mídia social. É um livro que busca ser um guia introdutório e simplificado de conceitos, práticas e formas de análise. Também por conta disso, junto aos conceitos, há muitos exemplos, buscando tornar mais compreensível um assunto que é deveras complexo. Fórmulas matemáticas e conceitos teóricos densos foram simplificados e explicitados de modo a permitir o entendimento da proposta. Assim, o norte desta obra está na busca das aplicações empíricas e no aprendizado pela prática. Boa parte dos fundamentos do que está exposto neste material encontra-se em obras anteriores, como Recuero (2009), Recuero (2012) e Recuero, Bastos e Zago (2015). O que se faz aqui, portanto, é uma revisão de alguns conceitos básicos já defendidos e sua discussão de forma mais direta e mais relacionada a aplicações empíricas.

[7]

1 Vide, por exemplo, Recuero, Bastos e Zago (2015); Bruns e Burgess (2012); Bruns e colaboradores (2012) dentre outros.

Dentro desta ideia, a estrutura deste livro está baseada em perguntas e respostas. Cada capítulo inicia-se com uma questão-chave que será respondida no decorrer do texto. O encadeamento do texto, assim, está estruturado de modo a entrar aos poucos nas questões-chave, fornecendo as bases para a compreensão de como fazer análise de redes e a seguir, trazendo elementos complementares. Nos últimos capítulos, buscamos trazer os conceitos teóricos e os dados empíricos trabalhados nos exemplos. O ideal, para que se acompanhe este livro, é que o estudante tenha instalado em seu computador algum *software* de análise de redes e que tenha, também, dados que possam ser analisados. Se este não for o caso, ferramentas específicas podem ser conferidas no capítulo 6. Dados de mídia social também podem ser utilizados como teste porque são mais facilmente disponíveis e podem ser coletados com várias das ferramentas apresentadas aqui. Outros dados coletados pelos pesquisadores podem ser também trazidos para a análise.

A proposta deste livro, de modo simplificado, é baseada em uma breve apresentação e histórico do paradigma, os principais conceitos, suas métricas e, finalmente, suas formas de representação e visualização. Optamos também por, ao invés de apresentar aplicações em casos em um capítulo final, apresentar exemplos resumidos e focados nos conceitos durante o texto. Esses exemplos estão demarcados por quadros dentro do texto e podem ser lidos de modo independente dentro do conjunto do livro

O que é a Análise de Redes?

A Análise de Redes Sociais (ARS) é uma das perspectivas de estudo de grupos sociais que permite sua análise sistemática a partir de sua estrutura, através de medidas específicas para esta. É uma abordagem que tem suas raízes na Sociometria e na Teoria dos Grafos, de viés matemático, para analisar relações sociais. (RECUERO; BASTOS; ZAGO, 2015) Por não ser exclusivamente um método, mas uma abordagem na visão dos grupos sociais, a opção pela ARS pressupõe a percepção do grupo social como uma rede e de sua análise através de determinadas premissas metodológicas, que trataremos a seguir.

A ideia que embasa os estudos das estruturas sociais é aquela de que os indivíduos, os atores sociais, estão inseridos em estruturas complexas de relações com outros atores. Essas estruturas têm um papel fundamental no comportamento e na visão de mundo desses indivíduos, mais do que outras categorias muitas vezes tomadas *a priori* como mais importantes, como classe social, sexo ou idade. (DEGENNE; FORSÉ, 1999; WASSERMAN; FAUST, 1994) Cada indivíduo, desse modo, está inserido em grupos sociais que são constituídos a partir de suas relações nos mais variados ambientes (por exemplo, família, escola, amigos ou trabalho). Essas relações são estabelecidas por interações e associações e vão conferir aos atores determinadas posições nas suas redes sociais, que vão sendo modificadas por essas mesmas ações. A posição desses atores é, ao mesmo tempo, produto e produtora de interações, ou seja, a rede influencia e é influenciada pela posição de seus usuários.

Assim, por exemplo, a circulação de uma proposta de emprego dentro de um determinado grupo social está diretamente relacionada às decisões individuais em repassar a informação a outros atores ou não. Temos um determinado ator A que está em busca de um emprego. A posição desse ator A na rede, ou seja, o lugar onde ele está com relação às suas conexões a outros e às conexões desses outros, vai determinar se ele receberá ou não essa proposta. Assim, podemos perceber que as decisões individuais têm um reflexo na rede como um todo, porque permitem que determinadas informações circulem ou não. Do mesmo modo, as estruturas dessa rede são constituintes de e constituídas por essas interações e pelas decisões dos atores sobre elas.

A rede dentro da qual qualquer indivíduo está inserido (ou seu grupo social) é também a responsável por uma grande parcela de influência sobre esse indivíduo. O lugar de alguém na estrutura social advém de uma série complexa de relações, da qual emergem normas, oportunidades e, inclusive, limitações. Um determinado ator que está inserido na estrutura de uma empresa em uma posição subalterna, por exemplo, tem determinadas expectativas para si oriundas dessa posição. Dependendo dela, ele pode ter acesso a informações importantes para “crescer” na empresa ou mesmo ter limitações extremas nesse acesso, impedindo que consiga ter oportunidades ali dentro. Ou seja, a percepção da estrutura em torno dos atores é fundamental para que possamos compreender também seu comportamento. Além disso, o comportamento individual dos atores reflete-se na rede como um todo, moldando-a e adaptando-a, sendo também, portanto, fundamental para que possamos compreender a estrutura em si.

São exemplos como esses que estão estabelecidos nas premissas da percepção dos grupos sociais pela análise de redes. Ela vê, assim, no estudo das estruturas sociais, uma oportunidade para compreender os fenômenos decorrentes dessa estrutura e da posição dos atores nela.

De onde vem a Análise de Redes Sociais?

A análise de redes sociais nasce de um ramo interdisciplinar de pesquisa, cujas bases podem ser encontradas nas mais variadas ciências, principalmente no início do século XX, particularmente, a partir da década de 1930. A discussão sobre os possíveis antecedentes da abordagem é intensa. De modo geral, a maioria dos autores cita como antecedentes da ARS sistematizações mais completas, como a Sociometria e a Teoria dos Grafos (WASSERMAN; FAUST, 1994; DEGENNE; FORSÉ, 1999; SCOTT, 2001), embora traços dos conceitos possam ser observados em trabalhos muito anteriores. (FREEMAN, 2004) Especificamente, para fins deste livro, focaremos nestes, embora elementos variados possam ser observados nas mais variadas disciplinas, notadamente, na Antropologia, na Psicologia, na Matemática, na Física e na própria Sociologia. Scott (2001) credita o “nascimento” da ARS como abordagem ao

desenvolvimento da Sociometria, que trouxe sistematização analítica a partir de fundamentos da teoria dos grafos. Já o desenvolvimento desse método, o autor atribui aos pesquisadores que, na década de 1930, passaram a estudar os padrões de relações e a formação de grupos sociais como cliques e, finalmente, aos antropólogos que a partir desses elementos começaram a estudar os conceitos de “comunidade”. Para o autor, são essas tradições que vão formar aquilo que, na década de 1960, vai se constituir na tradição dos estudos de análise de redes.

De modo geral, um dos principais fundamentos da ARS está na abordagem de Jacob Moreno e na invenção do sociograma, no início da década de 1930. (WASSERMAN; FAUST, 1994) O sociograma é a representação da rede, no qual os atores sociais são apresentados como nós, e suas conexões, representadas por linhas que unem esses nós. O principal interesse de Moreno era medir as relações dos grupos, compreendendo, a partir dessa medida, como esses conjuntos de atores eram estruturados. Sua abordagem foi denominada Sociometria. Em seu livro, *Who shall survive*, ele utiliza, pela primeira vez, o termo “redes”, referindo-se às redes de relações interpessoais entre indivíduos. Scott (2001) argumenta ainda que o trabalho de Moreno, fortemente ancorado na Psicologia, trazia em si traços importantes dos trabalhos de Tönnies, Weber, Simmel e outros pesquisadores alemães. Freeman (2004) credita parte do trabalho de Moreno a Helen Jennings, uma de suas primeiras colaboradoras. Apesar das ideias de Moreno, a seu tempo, não terem tido tanta receptividade como após a década de 1950, quando foram reabordadas por outros grupos, é na perspectiva sociométrica, de medida, que estão a maioria dos principais conceitos da análise de redes.

Outro fundamento importante é a chamada Teoria dos Grafos. (RECUERO, 2009) Se a Sociometria trouxe para a análise de redes o foco na estrutura social para que se compreendesse a dinâmica dos grupos, foi na teoria dos grafos que a análise de rede encontrou formas mais sistemáticas de medida e compreensão dessa estrutura. Os grafos são representações das redes, e em sua teorização estão as principais métricas utilizadas para a compreensão de sua estrutura e das posições de seus nós.

A teoria dos grafos compreende uma parte da matemática que estuda conjuntos de objetos e suas conexões. Suas origens estariam no trabalho de Euler e na solução que ele propôs para o enigma das Pontes de Königsberg. A

história relata que a cidade de Königsberg seria atravessada por sete pontes e que popularmente havia um desafio de desenhar um caminho por ela onde cada uma das pontes seria atravessada uma única vez. Euler teria demonstrado que tal desafio era impossível de ser resolvido utilizando um grafo, dando assim origem à teoria.

Scott (2001) credita a Cartwright e Harary a primeira aplicação dos grafos para os sociogramas. Essa aplicação permitiu que novas perspectivas fossem compreendidas dentro da dinâmica dos grupos sociais. De modo mais específico, os autores contribuem para a compreensão das subdivisões do grafo.

Recentemente, uma mudança importante na análise de redes sociais é a ampliação do foco do estudo de grupos pequenos para grupos em larga escala. Essa transformação está relacionada primeiramente à disponibilização de dados sociais, especialmente pelas ferramentas digitais de comunicação. Em um segundo momento, também está relacionada ao uso de métodos computacionais, que permitiram a coleta e a análise desses dados sociais. Desse modo, disciplinas como a Sociologia Computacional, Ciências Sociais Computacionais e outras passaram a aparecer nas áreas das Ciências Exatas, por exemplo, enquanto outras disciplinas, relacionando Métodos Computacionais e Métodos Digitais, passaram a aparecer nas Ciências Humanas e Sociais. É nesse contexto que se desenha um novo momento da análise de redes sociais.

Redes sociais e *sites* de rede social são a mesma coisa?

É preciso diferenciar as redes sociais dos *sites* de rede social. Embora seja senso comum referirmo-nos às ferramentas sociais digitais, tais como Facebook, Twitter, Orkut etc. como “redes sociais”, o conceito de rede social, tal como explicitamos na seção anterior, não é sinônimo delas. (RECUERO, 2009) Enquanto uma rede social está relacionada à percepção de um grupo social determinado pela sua estrutura (a “rede”), que é geralmente oculta, pois só está manifesta nas interações, as ferramentas sociais na internet são capazes de publicizar e influenciar essas estruturas sociais. (BOYD; ELLISON, 2007) Ou seja, o Facebook, por si só, não apresenta redes sociais. É o modo de apropriação que as pessoas fazem dele que é capaz de desvelar redes que existem ou

que estão baseadas em estruturas sociais construídas por essas pessoas (muitas vezes, de modo diferente daquele previsto pela própria ferramenta). Uma vez que passem a usar o Facebook, os atores criarão ali redes sociais que passarão a ser exibidas por ele.

É por conta desta e de outras características que Boyd e Ellison (2007) criaram o conceito de “*site* de rede social”. Para as autoras, algumas ferramentas *online* apresentam modos de representação de grupos sociais baseados nas relações entre os atores. Essas ferramentas teriam as características de (1) permitir que os atores construam um perfil público ou semipúblico; (2) permitir que esses atores construam conexões com outros atores; e (3) permitir que esses atores possam visualizar ou navegar por essas conexões. O *site* de rede social é, assim, diferente da rede social, pois aquele representa esta. Além disso, o *site*, enquanto ferramenta que é apropriada² pelos usuários, não é uma tradução das conexões sociais existentes no espaço *offline*. Ao contrário, eles amplificam conexões sociais, permitem que estas apareçam em larga escala (RECUERO, 2009) e também atuam de modo a auxiliar na sua manutenção.

Nas “redes sociais na internet”, assim, é possível ter centenas de conexões sociais ativas, pelas quais se está constantemente recebendo informações. O *news feed* do Facebook é um exemplo disso: quando um determinado ator conecta-se à ferramenta, recebe ali mensagens publicadas ou valorizadas pelas suas conexões na ferramenta. Dentre essas mensagens estão, inclusive, aquelas publicadas por outros atores com os quais ele não tem uma proximidade no espaço *offline* e que, sem esta ferramenta, dificilmente receberia. Ao mesmo tempo, o Facebook também permite que conexões sociais que normalmente seriam distantes e tenderiam a desaparecer no tempo (por exemplo, “conhecidos” de algum ambiente social) sejam mantidas. O processo de construção das redes ali também é diferente. Basta uma solicitação, não exigindo a interação social que construiria relações de amizade como aquelas características do espaço *offline*. As redes sociais na internet não são iguais, assim, em sua constituição, manutenção e publicização, às redes sociais *offline*. São outro fenômeno, característico da apropriação dos *sites* de rede social.

2 “Apropriada” é usada aqui no sentido de apropriação técnica e simbólica, conforme defendido por Lemos (2003). As ferramentas utilizadas precisam ser compreendidas em seus aspectos técnicos e, posteriormente, ressignificadas para que possam ser utilizadas pelas pessoas.

O estudo das redes sociais através da análise de redes pode ser feito tanto para as redes sociais *online* quanto para as redes *offline*. O que é preciso levar em conta é como esses fenômenos são diferentes entre si.

Outros elementos típicos dos *sites* de rede social que também são importantes para quem os analisa foram explicitados por Danah Boyd (2010) através do conceito de “públicos em rede” (*networked publics*). Embora esse conceito não esteja diretamente relacionado com análise de redes, ele auxilia a compreensão de como os *sites* de rede social influenciam os processos de representação dos grupos. A autora utiliza a noção de *affordances* para explicitar elementos que emergem das características técnicas dessas ferramentas e suportam suas apropriações. Essas *affordances* seriam: (1) persistência; (2) replicabilidade; (3) escalabilidade; e (4) buscabilidade.

A persistência refere-se ao fato de que as interações que são constituídas nos meios *online* tendem a permanecer no tempo. Ao contrário, por exemplo, de interações orais, que tipicamente desaparecem após sua locução, as interações digitais tendem a permanecer, constituindo conversações que podem ser recuperadas em outros momentos. (RECUERO, 2012) A persistência permite que as interações nas redes sociais *online* possam acontecer em momentos diversos, inclusive quando os atores envolvidos não estão presentes ao mesmo tempo. Com isso, há uma ampliação das possibilidades de manutenção e recuperação de conexões e valores sociais, conceitos que veremos adiante. A replicabilidade, por sua vez, permite que a informação transite de modo mais rápido e com menos ruído na estrutura das redes sociais *online*. Como as interações permanecem no espaço *online*, elas são mais facilmente replicadas, e a informação pode circular mais rapidamente e com maior fidedignidade. Esses dois elementos, persistência e replicabilidade, são chave para que as informações publicadas nessas redes também sejam facilmente escaláveis, ou seja, possam rapidamente percorrer toda a estrutura das redes, de modo “viral”. Como é fácil replicar as interações, e uma vez que estas estão permanentes no espaço *online*, é igualmente fácil escalar a transmissão de uma determinada informação que está contida nessas interações. Por fim, também graças à persistência, as interações podem ser buscáveis, bem como as informações nelas contidas.

Em parte, são essas características estruturais que permitem, por exemplo, que consigamos analisar as redes sociais na internet. Sem permanência das

interações e das conexões sociais, por exemplo, não conseguiríamos reconstruir as redes e analisar suas estruturas. Sem a buscabilidade, outro exemplo, não conseguiríamos coletar dados dessas conversações. Entretanto, essas mesmas características também alteram essas estruturas sociais, tornando-as diferentes das redes sociais *offline*.

Além desses elementos, Boyd (2010) aponta para o que chama de “dinâmicas” desses públicos em rede, características emergentes da estrutura dos grupos sociais *online*. Essas dinâmicas, embora muito mais importantes para o processo de conversação nos *sites* (RECUERO, 2012), também caracterizam e influenciam as redes sociais que emergem desse processo. Elas seriam: (1) audiências invisíveis; (2) colapso dos contextos; e (3) borramento das fronteiras entre o público e o privado. As redes expressas nesses *sites* de rede social tendem a ser, conforme dissemos, muito maiores e mais interconectadas do que as redes *offline* (o estudo de Backstrom e colaboradores, 2012, sobre o Facebook, por exemplo, demonstra isso). Entretanto, essa estrutura interconectada de atores que estão aproximados na mediação constitui audiências que não são imediatamente discerníveis para quem está interagindo. Por isso, Boyd diz que as audiências que rodeiam as interações no espaço *online* não são imediatamente discerníveis, logo, “invisíveis”. Além disso, como as interações permanecem e são reproduzidas nessas ferramentas, elas podem facilmente ser deslocadas de seus contextos originais e replicadas em outros. Esse descolamento do contexto poderia, por exemplo, aumentar o potencial para conflitos das interações desses grupos. Finalmente, o apagamento ou borramento das fronteiras entre o público e o privado traduz-se na dificuldade em demarcar espaços que são tipicamente dados nos grupos sociais *offline*, tais como os grupos familiares, com os quais se tem intimidade, e aqueles com os quais não se tem. Nos *sites* de rede social é muito mais difícil compreender elementos de natureza privada, pois, por princípio, o que se coloca nessas ferramentas é público. Essa confusão acaba por expor os atores, aumentando a percepção de intimidade e sua participação pela rede.

São esses elementos que proporcionam os contextos nos quais podemos perceber como as redes sociais na internet podem ser diferentes em suas apropriações e práticas sociais, e na circulação de informações das redes sociais *offline*.

Para que serve a ARS?

Podemos usar análise de redes sociais hoje para estudar diversos fenômenos. Notadamente, a ARS tem sido usada em áreas emergentes, como as áreas da Comunicação Social e da Sociologia Computacional, para compreender fenômenos associados à estrutura das redes sociais, principalmente, *online*. Graças ao aumento da quantidade de dados sociais disponibilizados por conta dos usos das ferramentas de comunicação mediada pelo computador, e por ser uma abordagem bastante propícia para o estudo e a visualização de grandes quantidades de dados, a ARS tem ganho projeção e tem sido utilizada com vieses mais quantitativos. Desse modo, essa abordagem é interessante, por exemplo, para analisar comportamentos de uma grande quantidade de atores sobre um evento ou tópico, bem como a influência desses atores nos processos de comunicação sobre esse tópico. Mas isso não quer dizer que a ARS não possa ser utilizada em casos específicos, com viés mais qualitativo. Além disso, uma grande parte do escopo está também no desenvolvimento de visualizações de grafos que permitam a compreensão dessas estruturas.

O que importa, para um estudo de análise de redes, é que o objeto do trabalho tenha uma estrutura mapeável e que esse mapeamento seja útil para a compreensão do fenômeno que o pesquisador visa investigar (vide o exemplo a seguir, na Figura 1). E o que pode ser estudado a partir de sua estrutura social?

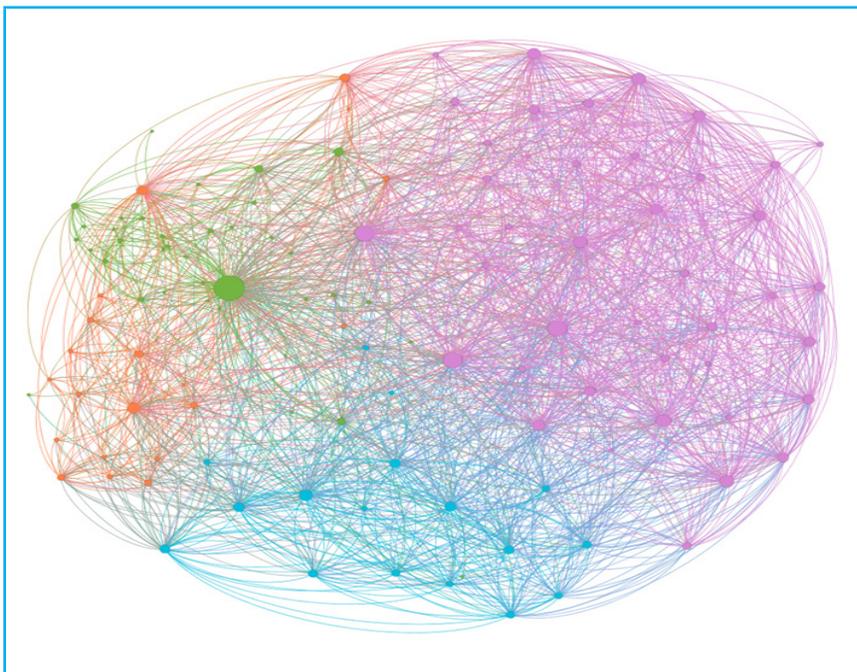
1) Estudos das relações entre os elementos da estrutura do fenômeno: questões para ARS são aquelas que focam fenômenos sociais cuja estrutura é central para o problema de pesquisa. Desse modo, por exemplo, saber como o capital social é constituído em um determinado grupo é um estudo possível de ser abordado com ARS. Outro exemplo é o estudo de como a informação circula num determinado grupo social, quais são os “gargalos”, quais são os atores mais influentes e importantes no grupo para que a informação circule. Outro caso poderia ser um estudo em que se busca identificar subgrupos (*clusters*) dentro de um grande grupo (rede), tal qual a identificação de grupos ativistas dentro do Facebook. Esses são problemas que permitem a abordagem estrutural, pois permitem que sejam mapeados aqueles nós que estão envolvidos na replicação de informação e na própria estrutura. A partir deste mapeamento estrutural e da aplicação de métricas, podemos compreender

como o capital social está sendo gerado pelas conexões, *onde* se concentra e *quem* é mais importante na estrutura. Podemos identificar como a estrutura influencia a difusão de informação e ainda mais, quais grupos existem nela.

2) Estudos nos quais o objeto possa ser estruturalmente mapeado: nem todo objeto é passível de ser mapeado em sua estrutura, por várias razões. Em alguns casos, por exemplo, uma parte da estrutura está oculta, é inacessível ao mapeamento. Em outros, a estrutura é grande demais, não é passível de ser mapeada no tempo de que se dispõe para o trabalho. E também há aqueles nos quais simplesmente não se consegue observar a estrutura. Ou seja, é preciso certificar-se de que o objeto pode ser mapeado e compreender como essa estrutura vai ser “vista” (ou seja, quais serão os nós e as conexões da rede). Além disso, para que o objeto seja mapeado, é preciso que os seus dados sejam acessíveis ao pesquisador, que é o nosso próximo ponto.

3) Estudos nos quais o problema de pesquisa foque um conjunto de dados passível de ser coletado e mapeado com os recursos disponíveis: outra questão central para o trabalho com análise de redes é que o estudo em questão foque em dados que possam ser acessados pelo pesquisador. Embora muitos problemas de pesquisa possam se beneficiar de uma abordagem de redes, nem todo tipo de dado é passível de ser acessado e coletado pelo pesquisador. Também há casos em que os dados existem, mas não há recursos para sua coleta e análise no intervalo de tempo disponível para fazer a pesquisa. Em alguns casos em que o pesquisador se depara com esse problema, é possível fazer algum tipo de recorte nos dados. Nesses casos, é importante avaliar como o recorte foi feito e se esse recorte permite que a análise seja válida para os objetivos. Também é importante observar se há recursos computacionais e humanos para a coleta de dados, sua armazenagem, análise e posterior construção dos grafos.

Figura 1 - Rede de interações no Flickr



Vemos aqui (Figura 1) uma rede formada em torno de pessoas que interagem (comentam e recebem comentários) na rede Flickr. Cada ator está representado por um ponto ou esfera (nó, como veremos adiante). Cada conexão entre os nós representa várias interações no período de dois anos (por exemplo, comentários nas fotos). A estrutura da rede é formada pelos atores e suas interações e pode nos mostrar, por exemplo, quais atores mais recebem comentários/interações da rede (tamanho maior dos pontos), apontando, assim, quem são os atores mais “populares”. Também poderíamos analisar quais comentam mais (será que são os mesmos que recebem mais comentários?). A própria estrutura da imagem mostra que todos tendem a comentar/receber comentários entre si, pois não há nenhum nó “solto”. Será que isso indica que as pessoas que não se conhecem nos vários grupos tendem a se conhecer pelas interações? Essas ideias são exemplos de questionamentos e estudos que são possíveis fazer com análise de redes. Nos próximos capítulos vamos discutir como fazer esse tipo de estudo e como criar e compreender grafos de redes sociais.

Dentre exemplos de trabalhos atuais que utilizam ARS para o estudo dos fenômenos, no âmbito da área da Comunicação, temos trabalhos focados: 1) nas conversações mediadas por computador, que analisam principalmente como a estrutura da rede influencia os processos de difusão de informações em grupos *online* (BRUNS et al, 2012); 2) no modo como a estrutura desvela os processos de conversação nesses grupos (BRUNS et al, 2012; RECUERO, 2014); 3) nos processos de colaboração emergentes (RECUERO; BASTOS; ZAGO, 2015); e até mesmo 4) no modo como aplicar essas outras premissas em cima das métricas construídas pela ARS. (MALINI, 2016)

Quais são os principais conceitos de ARS?

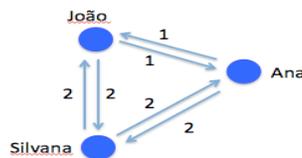
A ideia de “redes sociais” é uma metáfora estrutural para que se observem grupos de indivíduos, compreendendo os atores e suas relações. Ou seja, observam-se os atores e suas interações, que por sua vez, vão constituir relações e laços sociais que originam o “tecido” dos grupos. Essas interações proporcionam aos atores posições no grupo social que podem ser mais ou menos vantajosas e lhes dar acesso a valores diferentes. São esses elementos que a ARS busca desvelar, conforme explicamos no capítulo anterior.

A análise de redes sociais, desse modo, trabalha com a representação dos grupos como sociogramas (grafos sociais), que são analisados a partir das medidas de suas propriedades estruturais. Nessa metáfora, os laços ou relações sociais constituem as conexões entre os nós (ou nodos) que são os atores sociais. Dependendo do objeto estudado, assim, as conexões podem ser observadas como interações, relações informais, ou laços sociais mais estruturados.

A teoria dos grafos provê o suporte matemático de análise. Um grafo consiste em um conjunto de nós e suas conexões (ou arcos). O grafo é, desse modo, uma representação de dois conjuntos de variáveis (nós e conexões). Concebendo uma rede social como uma dessas representações, os nós seriam os atores sociais (compreendendo esses atores como organizações sociais, grupos ou mesmo indivíduos no conjunto analisado) e suas conexões (aqui entendidas como os elementos que serão considerados parte da estrutura social, como interações formais ou informais, conversações etc.). A representação dos dados para a formação da rede, na teoria dos grafos, dá-se através de matrizes. Nessas matrizes, são estipuladas as relações entre os atores do grupo analisado. Elas são essenciais para a construção dos grafos, pois demonstram os dados que servem de base para a estrutura geral da rede (Figura 2).

Figura 2 – Matriz e grafo que a representa

	João	Ana	Silvana
João	0	1	2
Ana	1	0	2
Silvana	2	2	0



A matriz acima demonstra as relações entre três indivíduos, colegas de trabalho, João, Ana e Silvana. Imaginemos que cada relação é constituída do número de vezes que cada um dos atores interagiu com o outro em uma data específica. Assim, no dia escolhido, vemos que Ana interagiu com João apenas uma vez (cumprimentou-o ao chegar ao trabalho), mas interagiu com Silvana duas vezes (além de cumprimentá-la na chegada, conversou com ela durante o intervalo para o café). Silvana, por sua vez, interagiu duas vezes com Ana e duas vezes com João (cumprimentou e conversou com ambos em momentos posteriores distintos). Assim, em nossa matriz, temos conexões de João para Ana com peso 1 (apenas uma interação) e de João para Silvana com peso 2 (duas interações). Do mesmo modo, temos Ana para João com peso 1, e Ana para Silvana com peso 2 (duas interações). Já Silvana interagiu duas vezes tanto com João como com Ana. É isso que vemos na matriz. Já na representação de grafo, vemos três nós, cada um representando cada ator e as conexões com seu valor entre cada dupla. Observa-se que Silvana é o nó com maior “peso” no grafo, pois interagiu duas vezes com os demais nós. O grafo, assim, representa a matriz de modo topográfico. Os nós representam os atores, e as conexões, as interações entre eles.³ A seguir, veremos com maior detalhamento o que constituem cada um desses elementos.

[22]

3 Note-se que no grafo optamos por manter a direcionalidade das interações com os valores, embora esse grafo pudesse ser representado de modo não direcionado. (Conforme veremos a seguir.)

A rede social, seus nós e representações

A rede social é constituída de nós (ou nodos) que representam os atores sociais na estrutura. O nó pode representar tanto um indivíduo quanto uma categoria ou grupo. Assim, em um determinado estudo, os nós podem representar pessoas individuais, e em outro, grupos de pessoas. Tudo depende do que o pesquisador está trabalhando.

Os nós geralmente pertencem todos à mesma categoria (por exemplo, todos os nós representam indivíduos; ou todos os nós representam organizações; ou todos os nós representam perfis etc.). Esse tipo de rede é chamada “rede de um modo” (*one mode networks*). Há casos, entretanto, em que em uma mesma rede temos dois tipos diferentes de nós (por exemplo, indivíduos e organizações). Nesse caso, essas redes são denominadas redes de dois modos (*two mode networks*). Um dos exemplos mais típicos de uma rede de dois modos é aquela em que são analisados indivíduos e eventos como os dois tipos de nós. Entretanto, é preciso ter cuidado quando se trabalhar com redes de dois modos, pois a maior parte das métricas utilizadas em ARS foi criada para as redes de um modo.⁴

As conexões: laços sociais

[23]

As redes sociais também são compostas de conexões. Essas conexões, ou arcos, indicam normalmente algum tipo de relação social (interação, conversação, relação de amizade ou pertencimento etc) entre diferentes nós.

Compreender como os indivíduos estão conectados na estrutura social é um desafio importantíssimo e sobre o qual muitos autores tentaram lançar indicativos. Um dos trabalhos mais conhecidos nesse sentido é o de Mark Granovetter sobre a “força dos laços fracos” (1973). Granovetter propôs uma classificação dentro da qual os laços sociais poderiam ser “fortes”, “fracos” ou “ausentes”. Um laço forte seria associado a relações sociais que indicam maior

4 Vide Borgatti e Everett, 1997.

proximidade entre dois nós, denotando algum tipo de intimidade (como por exemplo, amizade); enquanto os laços fracos representariam associações mais fluidas e pontuais entre os atores (por exemplo, relações com “conhecidos”). Já os laços “ausentes”, para o autor, seriam insignificantes em termos de importância estrutural ou completamente ausentes entre dois determinados nós.

Granovetter (1973, p. 1361) define a “força” do laço como “uma combinação (provavelmente linear) da quantidade de tempo, da intensidade emocional, da intimidade (confiança mútua) e da reciprocidade que caracterizam o laço”.⁵ Desse modo, presume-se que a quantidade de interações entre dois atores pode representar, assim, a força da conexão, uma vez que conexões fortes requerem maior investimento do que conexões fracas. A partir dessa discussão, podemos observar que os tipos de laço social representam conexões que são qualitativamente diferentes, e é um desafio importante compreender como a estrutura construída através dos dados e da análise de redes pode, efetivamente, representar esse conceito complexo.⁶

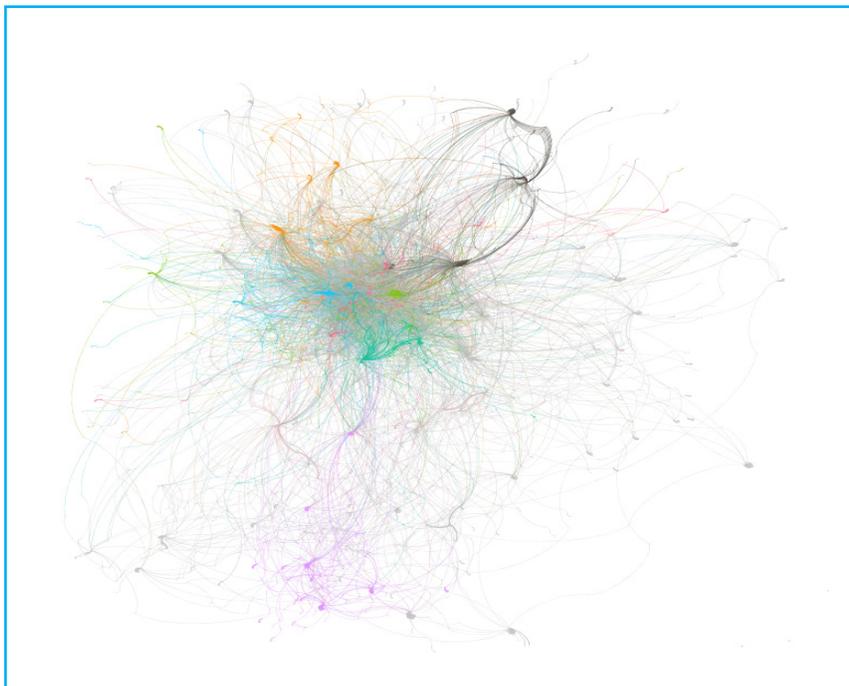
Tradicionalmente, na análise de redes, as conexões são representadas de modo numérico e direcional, indicando um valor que é relacionado ao “peso” da conexão. Esse “peso” pode dar pistas a respeito do tipo de conexão entre os vários nós da rede. Desse modo, as conexões mais fracas (com valor menor) ou mais fortes (com valor maior) podem trazer indícios do tipo de laço social que existe entre um determinado par de atores. Ao mesmo tempo, as conexões também podem ser direcionadas (quando A envia uma mensagem a B, podendo ou não receber resposta) ou não direcionadas (quando se presume que a conexão tem a mesma força de ambos os lados, como por exemplo, com uma relação de “amizade” num *site* de rede social – não é possível indicar a direção dessa conexão, diferentemente de quando alguém envia uma mensagem a outrem). Veremos mais sobre o direcionamento das conexões a seguir neste capítulo.

5 Tradução da autora para: “a (*probably linear*) combination of the amount of time, the emotional intensity, the intimacy (*mutual confiding*) and the reciprocal services which characterize the tie”

6 Embora muitos autores (como Granovetter) tenham trabalhado a construção da rede a partir da percepção dos atores sobre suas relações, outros trabalhos demonstraram que este sistema é, no mínimo, não confiável quando comparado à observação sistemática dessas declarações. (EAGLE; PENTLAND; LAZER, 2009)

Geralmente, entretanto, onde há conexões mais fortes, há também conexões mais multiplexas, ou seja, conexões provenientes de relações sociais mais variadas (por exemplo, interações em diferentes ambientes sociais). Conexões mais fortes também são mais recíprocas e tendem a constituir, na estrutura da rede, o que chamamos de *cluster* ou agrupamento. Um *cluster* é um conjunto de nós mais densamente conectados (ou mais interconectados) do que os demais na rede (Figura 3).

Figura 3 – *Cluster* em grafo de rede social



Na Figura 3 vemos um grafo que representa uma rede social. Há diversos nós, marcados pelas esferas, alguns mais distantes e alguns mais conectados ao centro da imagem. Esses nós mais conectados, no espaço mais denso, constituem um *cluster*. Estruturalmente, esses nós estão mais próximos, ou porque interagem mais (e suas arestas têm um peso maior) ou porque possuem mais conexões entre si do que com os demais nós da rede. Podemos observar que há menos conexões entre os nós mais distantes e mais conexões nesses nós centrais. *Clusters* também são chamados de “comunidades”. Em termos sociológicos, a estrutura de *cluster* demonstra nós mais conectados na rede e, nessas conexões, maior número de interações. Portanto, laços mais fortes estariam mais presentes nesses grupos do que entre os demais, com características mais próximas das definições de comunidade. (RECUERO, 2009)

Capital social e os valores das conexões

O conceito de laço é também associado à noção de capital social. Como vimos, a noção de laço social de Granovetter (1973) implica no fato de que esse laço pode constituir determinadas vantagens estruturais aos atores (informação, intimidade, reciprocidade etc.). Essas “vantagens” ou benefícios são formas de capital. (RECUERO, 2009) Assim, por exemplo, a informação é um valor social. Alguém só recebe determinada informação (por exemplo, “a prova será suspensa”) dentro de um determinado grupo (no mesmo exemplo, um grupo de alunos). O acesso ao grupo é determinado pelas interações e pela participação ou não nesta estrutura social. Assim, fazer parte do grupo constrói um benefício social, na medida em que dá aos atores oportunidades de acesso a determinadas informações. Mas como o conceito de capital social é percebido na literatura?

De um modo geral, a maioria dos autores que trabalha com o conceito concorda com o fato de que capital social representa um valor associado à estrutura social. (BOURDIEU, 1972; PUTNAM, 2000; BURT, 1992; COLEMAN, 1988) Burt (1992) argumenta que o conceito é uma metáfora para as transações que caracterizam as interações sociais. Putnam (2000) divide o capital social em dois tipos: *bridging* (relacionado às conexões, como pontes ou laços

fracos entre os atores de diferentes grupos) e *bonding* (relacionado à qualidade dessas conexões, ou laços fortes, dentro de um mesmo grupo). O capital social para o autor, portanto, estaria associado ao tipo e qualidade das conexões em uma rede. Estar em uma rede social, assim, permite a construção de valores para os atores.

Desse modo, as relações sociais são constituídas de trocas através das quais os atores buscam atingir objetivos e interesses, como um sistema econômico. É preciso investir (interagir) na estrutura social para colher os benefícios. Os valores de capital social são, desse modo, associados a normas de comportamento, participação e às próprias conexões que alguém possui, além de vantagens competitivas advindas desses valores. (BURT, 1992, p. 348)

Poderíamos argumentar, assim, que a interação social é guiada pela percepção (ainda que inconsciente) de valores, posição e vantagens sociais. Embora Bourdieu (1972) apresente uma visão mais crítica nesse sentido (para o autor, o capital social é uma fonte de desigualdade social), para outros autores, o capital social é a própria “cola” da estrutura do grupo social, facilitando a coordenação da ação dos grupos. (PUTNAM, 2000) O próprio Putnam enumera alguns desses valores, apontando a confiança, as normas e as redes como elementos de capital social. Ou seja, a partir desse ponto de vista, as trocas sociais implicam na construção de valores cuja percepção por parte do grupo também atua na construção de relações de confiança, resultado dos investimentos individuais na estrutura.

Diferentes tipos de laço social geram vantagens e elementos para os indivíduos na rede social, como veremos no decorrer deste livro. Laços fracos podem ser mais eficientes para circular informação nova e diferente na rede (GRANOVETTER, 1973), enquanto laços fortes são mais associados a valores como confiança e intimidade. (PUTNAM, 2000) Dessa maneira, a posição do ator na estrutura social está associada a valores aos quais ele tem ou não acesso. A existência de grupos mais conectados do que o resto da rede (*clusters*) também pode estar associada a outros tipos de capital social. Além disso, a estrutura possui valores não apenas para os atores individualmente, mas para os grupos sociais como um todo. Compreender a estrutura e a posição dos atores na rede pode auxiliar também a compreender quais valores existem nessas redes e como esses valores circulam.

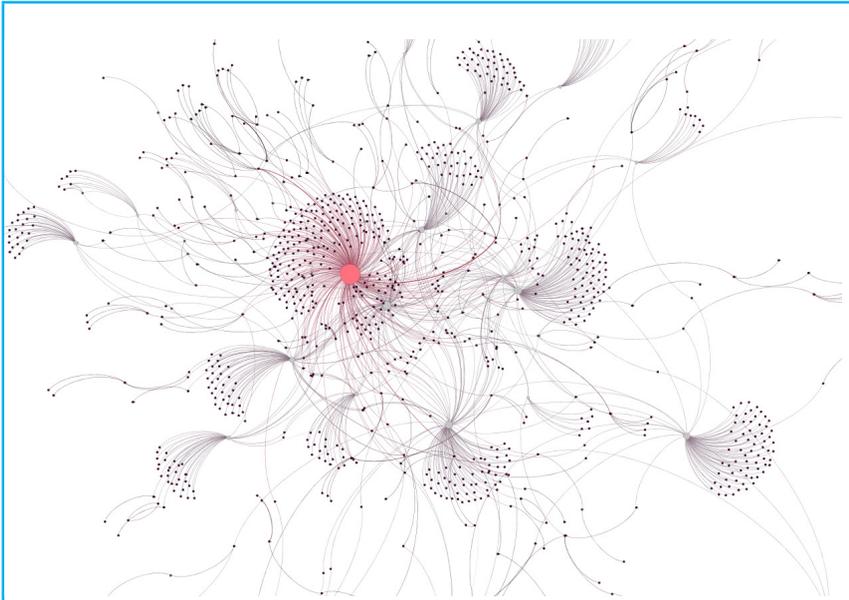
[27]

Há vários outros conceitos que nos ajudam a compreender os vários tipos de valores que estão presentes em uma determinada rede social. De modo geral, os estudiosos das redes associam o conceito de capital social à circulação de informação. Entretanto, outros autores já relacionaram o conceito de capital social a outros valores mais concretos. Bertolini e Bravo (2004), nesse contexto, relacionaram o conceito de capital social a diferentes níveis na rede. Em um primeiro nível, o capital social pode ser apropriado de modo individual, e está relacionado a elementos provenientes dos laços fracos. Nesse nível, há três tipos de capital social: as conexões sociais que os atores possuem, as informações às quais têm acesso, e o conhecimento das normas associadas ao grupo/rede social a que pertencem. No segundo nível, o capital social está associado aos laços fortes e ao coletivo, ou seja, benefícios que só podem ser usufruídos coletivamente. Aqui, os tipos de capital social são dois: a confiança no ambiente social, no sentido do ambiente ser favorável aos atores pela confiança construída entre eles, e a institucionalização, relacionada ao reconhecimento do grupo enquanto grupo. Essa tipologia oferece um parâmetro interessante para observar a emergência de valores na estrutura social e a influência dessa estrutura em retorno, nesses valores. Também auxilia a compreender como os valores podem ser usufruídos.

Há ainda outros valores que podem ser percebidos como decorrentes de investimento e captura de recursos na rede. Popularidade, por exemplo, pode ser discutida como um tipo de valor social. (RECUERO, 2009) Alguém só consegue ser popular diante de outras pessoas. A popularidade é uma concessão, no sentido de que o ator popular concentra mais capital social, em termos de atenção e visibilidade de seus pares, do que outras pessoas não populares. Para que alguém seja popular, portanto, é preciso que outros não sejam. O capital social, nas redes sociais, tem aspectos complexos, de apropriação e de escassez. Autores como Barabási (2003) demonstraram, por exemplo, a presença de conectores, ou seja, um grupo pequeno de nós que teriam muito mais conexões que os demais na mesma rede, de modo análogo à popularidade. Embora conectores tenham sido percebidos pela literatura subsequente como uma estrutura característica das redes, sua presença indica que o capital social é um recurso escasso e que pode ser apropriado por alguns, mas sua apropriação pressupõe a redução das chances dos demais também dele se apropriarem. Ou seja, temos canais, pessoas e grupos muito populares em quaisquer

redes, mas a chance de algum de nós conseguir sê-lo é escassa, embora, em princípio, todos tenhamos acesso aos mesmos recursos.

Figura 4 - *Hub* ou conector na rede (em vermelho, no centro)



Na Figura 4 vemos um conector. O nó vermelho, ao centro, tem um tamanho maior porque tem mais conexões que os demais. Este tem 387 conexões, enquanto que os demais têm menos de 200. Esta rede representa uma conversa, na qual o nó vermelho foi citado 387 vezes, enquanto que os demais foram citados em um número bastante inferior. Poderíamos argumentar, assim, que o nó central tem mais popularidade, sendo citado na conversa por vários outros nós, do que os demais.

Homofilia, Buracos Estruturais e Fechamento

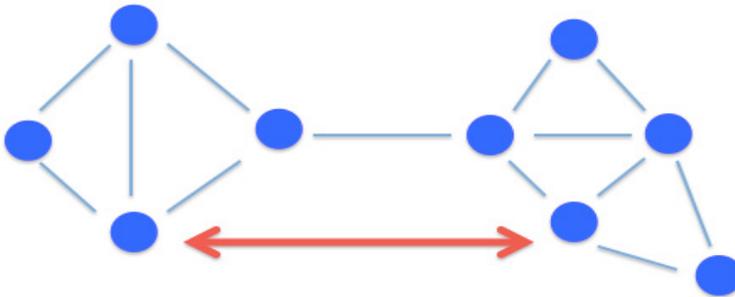
Pessoas mais próximas tendem a ter interesses comuns e padrões de comportamento semelhantes, seja porque se aproximam em virtude destes, seja porque passam a tê-los em razão de conectarem-se a grupos que os valorizam. Isso significa que grupos sociais tendem a agregar pessoas “parecidas”, como, por exemplo, grupos religiosos, que compartilham da mesma religião; grupos de dançarinos, que compartilham da mesma atividade; grupos de colegas que trabalham ou estudam na mesma localidade e compartilham disciplinas etc. Essa característica é chamada “homofilia” e está relacionada à similaridade dos atores em um mesmo *cluster* social, resultando no fato de que esses atores tendem a ter acesso às mesmas fontes e a circular as mesmas informações. Assim, a homofilia está relacionada também ao capital social, uma vez que pode auxiliar na construção e no fortalecimento dos laços sociais que vão gerá-lo.

Entretanto, há atores que se conectam a vários grupos, pois transitam em círculos variados. Essas conexões são geralmente referidas como “pontes” em redes sociais, pois aproximam, justamente, grupos distantes e diferentes entre si. Essas conexões são geralmente constituídas de laços fracos e são fundamentais para a estrutura da rede e a interconexão entre os diferentes grupos, conforme demonstrou Granovetter (1973). Conexões-pontes são importantes porque aproximam esses grupos e permitem que as informações circulem também entre eles.

A partir da ideia de pontes, Burt (1992) propõe o conceito de “buraco estrutural”. Esses buracos representam a ausência de conexões entre dois nós que possuem grupos/fontes informativas complementares ou não redundantes (Figura 5).

[30]

Figura 5 – Buraco estrutural (em vermelho)



Burt (1992) associa, assim, a ideia de buraco estrutural ao conceito de redundância.⁷ Um *cluster* geralmente tem conexões redundantes, ou seja, conexões que interligam o mesmo conjunto de nós. Já conexões não redundantes são aquelas que interligam os atores de diferentes grupos. Conexões redundantes o são porque nelas circulam as mesmas informações. Já as não redundantes são aquelas capazes de trazer informações novas para o grupo. Assim, como dissemos, as conexões transmitem informação e estão relacionadas ao conceito de capital social. O buraco estrutural, portanto, representa uma falha no caminho de transmissão de informações, que poderia dar acesso a fontes de informação diferentes para os dois grupos em questão. Desse modo, aqueles atores que fazem a “ponte” (ou mediação) entre diferentes grupos possuem uma vantagem em relação aos demais, pois têm acesso a tipos diferentes de informação, enquanto que os buracos estruturais representam uma desvantagem para os grupos.

Os *clusters* são associados ao “fechamento da rede” e, por isso, opostos aos buracos estruturais. O fechamento da rede é a qualidade associada a todos os nós de uma determinada rede estarem interconectados. Quanto mais fechada a rede, maior é o indicativo de que todos os atores que ali estão conectados compartilham de conexões. Ou seja, *clusters* tendem a compreender grupos mais fechados dentro de uma rede social. Quando um *cluster* é inteiramente

7 Observa-se em que medida um nó A possui conexões redundantes com os demais nós do grupo (ou seja, conexões com atores que estão conectados entre si).

conectado, é chamado de clique. Um clique é um grupo em que todas as conexões possíveis existem.

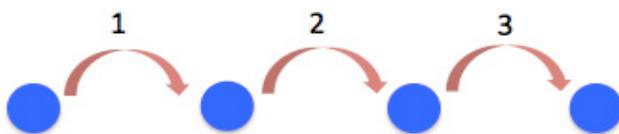
Tanto o conceito de fechamento quanto o conceito de buraco estrutural estão associados também à noção de capital social. Um clique é um indicativo de que um determinado grupo possui mais conexões redundantes e, portanto, níveis mais altos de capital social relacionados a conexões mais fortes.

Graus de separação

As conexões entre os atores também indicam a distância entre dois nós na rede social. Elas são representadas por arcos ou linhas. Cada arco é, assim, denominado “grau de separação” entre dois nós A e B. Desse modo, quando um nó A está conectado por um laço social a um nó B, ele está indiretamente conectado também às conexões de B e dos amigos dos amigos de B. A figura a seguir mostra como esses arcos funcionam na rede. Ali vemos os graus de separação do primeiro nó (ator) da rede e os demais. Alguém que está, assim, a três graus de separação, está conectado aos amigos dos amigos do primeiro autor (ego), sem estar diretamente conectado a ele.

[32]

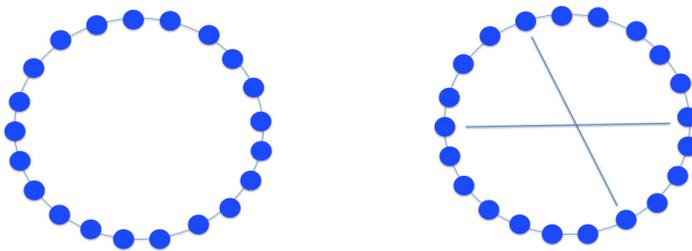
Figura 6 – Graus de separação



Essa ideia de grau de separação é bastante conhecida popularmente dentro da chamada teoria de “mundos pequenos”. Uma rede mundo pequeno é aquela em que um conjunto de nós é aproximado na rede por algumas conexões, que terminam por reduzir a distância (grau de separação) entre todos

os nós na estrutura. Isso significa que grupos de indivíduos que não se conhecem necessariamente estão bastante próximos na estrutura geral da rede por conta de alguns indivíduos que estão mais conectados (conhecem mais pessoas) do que a média dos demais.⁸

Figura 7 – Bastam algumas conexões para reduzir consideravelmente a distância entre os atores no grafo



A ideia de mundos pequenos é bastante presente no imaginário popular, relacionadas às coincidências dos encontros fortuitos e da descoberta dessas conexões entre os atores. Watts e Strogatz (1998) demonstraram que esse modelo é possível graças à presença dos laços fracos que interligam os diversos *clusters* na rede. Com isso, bastam algumas conexões do tipo fraco para reduzir de modo significativo a distância entre quaisquer dois atores na rede. A presença desse fenômeno foi apresentada e relatada de modo extensivo pela literatura. A estrutura de mundos pequenos, assim, é encontrada em redes sociais e está relacionada à presença de “pontes” entre os vários nós da rede e à redução do “caminho” (*path*) entre dois nós quaisquer da rede pela presença dessas conexões, que são apresentadas por autores como Granovetter (1973) e Watts e Strogatz (1998) como conexões fracas ou associadas aos laços fracos, e portanto, ao capital social do tipo *bridging* de Putnam (2000), pois permitem que os atores tenham acesso a fontes diferentes de informações. Uma das

8 Esta rede tem, assim, duas propriedades: uma distância (grau de separação) menor entre os atores da rede e um coeficiente de clusterização (associação dos nós à *clusters* ou grupos) alto.

métricas que pode ser usada no estudo das redes sociais é, assim, aquela de “caminho médio” ou “distância média” (*average path*), que calcula o caminho médio mais curto entre todos os nós da rede. Essa métrica pode indicar o quão interconectada está a rede pelos diversos laços existentes.

MUNDOS PEQUENOS, SEIS GRAUS DE SEPARAÇÃO E O FACEBOOK

A chamada “teoria dos seis graus de separação” parte da ideia de que todas as pessoas estão, em maior ou menor grau, conectadas no planeta. Essas “cadeias” de pessoas seriam representadas por grupos interconectados por conhecidos, o que reduziria significativamente as distâncias entre os atores sociais (teoria conhecida como “o problema dos mundos pequenos”). Essa ideia foi testada por vários pesquisadores, incluindo matemáticos como De Sola Pool e também pelo psicólogo Stanley Milgram na década de 1960. (RECUERO, 2009) No experimento de Milgram, este enviou uma série de correspondências a pessoas aleatoriamente escolhidas, solicitando que reenviassem a mesma a quem pudesse conhecer um determinado comerciante de Boston. Ao final, Milgram calculou a distância média entre as pessoas em uma rede social e concluiu que as pessoas estariam conectadas por seis graus (entre quaisquer duas pessoas, na média, teríamos cinco outras pessoas). (RECUERO, 2009) O fenômeno dos “mundos pequenos” foi testado várias outras vezes, por exemplo, por Duncan Watts (2003), sempre com resultados próximos de seis, apontando que não somente a distância média entre as redes sociais seria pequena, mas igualmente, que a capacidade para conexões seria limitada. Entretanto, um estudo sobre a interconexão no Facebook, publicado por Backstrom e colaboradores (2012) mostrou que a média da distância social no Facebook era de apenas quatro graus, ou seja, uma média de três atores, que é muito menor que a dos experimentos anteriores. Ou seja, poderíamos discutir que as redes sociais nessa ferramenta tenderiam a estar mais conectadas do que as redes sociais *offline* nas quais seria baseada.

Para saber mais:

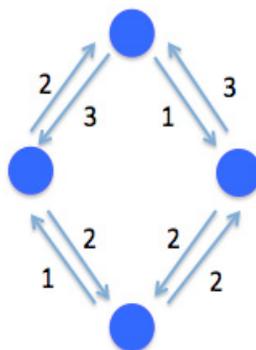
(<https://research.fb.com/three-and-a-half-degrees-of-separation/>)

Rede direcionada e não direcionada

Um grafo pode representar uma rede direcionada ou não direcionada. Uma rede direcionada (grafo direcionado) é aquela na qual as conexões possuem direção (geralmente representadas por uma seta). A opção por um grafo direcionado ocorre quando as conexões estabelecidas na rede não são iguais (ou não têm o mesmo peso) entre os atores. Uma rede que analise, por exemplo, interações entre os atores em um determinado contexto, é frequentemente direcionada. Imaginemos, por exemplo, uma rede em um canal de mídia social, como o Twitter, que é frequentemente utilizado em estudos dessa natureza. A rede de seguidores, por exemplo, é direcionada. A pode seguir B sem que este último o siga de volta. Ou A pode enviar 10 mensagens a B, e este apenas responder a cinco destas. As conexões ou não ocorrem nos dois sentidos ou possuem valores diferentes em cada direção, não são iguais entre os atores A e B. Assim, a matriz que representa essas relações é assimétrica, ou seja, com valores diferentes para $A \rightarrow B$ e $B \rightarrow A$.

No exemplo a seguir vemos a representação de um grafo direcionado, com a separação das conexões de entrada e saída de cada nó. Como podemos ver, cada nó tem conexões direcionadas aos demais nós, e em alguns casos, os valores desses arcos (conexões) são diferentes, representando a assimetria da rede. A rede é direcionada por ser representada pelas duas linhas ou apenas pela linha que representa o dado que se está levando em conta no momento do cálculo da métrica. Por exemplo, no caso de grau de entrada - *indegree* -, veremos apenas a contabilização das conexões que cada nó recebe dos demais. Já se contabilizarmos apenas o grau de saída - *outdegree* -, veremos o número de conexões que cada nó faz para os demais.

Figura 8- Grafo direcionado

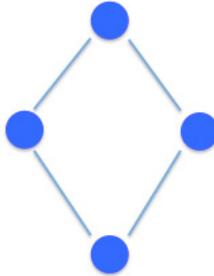


Em alguns casos, as relações assimétricas em uma matriz social também podem ser chamadas de relações valoradas e podem ser constituídas de sinais (+ ou -) para indicar se as interações são amigáveis (positivas) ou hostis (negativas) entre os atores.

Finalmente, uma rede também pode ser não direcionada, ou seja, suas conexões no grafo não possuem direção ou sua direção não importa, são representadas apenas por uma linha, mostram uma matriz na qual os dados dessas conexões são exatamente iguais entre os dois atores. Por exemplo, no Facebook, é preciso que um determinado ator A aceite se conectar a B para que a conexão se estabeleça. Se A aceitar B, a conexão não terá direção, pois ambos estarão conectados do mesmo modo tanto no sentido A-> B quanto B-> A. Desse modo, a matriz resultante dessas relações é simétrica, ou seja, com valores iguais entre A e B.

Na imagem a seguir, vemos a representação de um grafo não direcionado, no qual as conexões entre os nós são apenas linhas e não mais setas, pois o valor da conexão é o mesmo, independentemente da direção.

Figura 9 – Grafo **não direcionado**



Quando mapeamos uma rede social, a presença de uma comunidade é geralmente percebida em cima de um grupo de atores mais densamente conectados, uma vez que essas conexões representem relações sociais. *Clusters*, assim, podem representar comunidades (e usualmente são referidos assim pela literatura). Entretanto, é importante salientar que *clusters* também podem representar outras coisas, dependendo do que se está mapeando e do que o pesquisador considerará como conexão em um determinado grupo social.

Redes egocentradas e redes inteiras

[37]

Quando estudamos uma rede social também é preciso definir como sua estrutura será abordada. Há, geralmente, dois focos: rede ego ou rede inteira. A rede ego é aquela cuja estrutura é desenhada a partir de um indivíduo central, determinando, a partir desse “ego”, um número de graus de separação.⁹ Essa opção é utilizada quando se deseja avaliar a rede em torno de um determinado indivíduo ou quando não há um modo de limitar a rede que se deseja mapear. A rede inteira é aquela que é mapeada na sua integridade, normalmente

9 Embora o número de graus de separação represente a distância entre os nós, e ego seja apontado como um número inteiro (1, 2, 3 etc.), em alguns casos vemos esses números “quebrados” (1,5; 2,5). Essa quebra significa que estamos analisando todos os nós a um grau X de ego e mais as conexões entre esses nós.

quando é possível limitar essa rede de modo externo (por exemplo, todos os indivíduos que fazem parte de uma determinada instituição, ou todos os indivíduos que fazem parte de um grupo X). Nesse caso, não utilizamos como medida de coleta de dados o grau de separação.

Os objetivos de um trabalho que foque a análise de redes, portanto, estão relacionados à análise da estrutura de um determinado grupo social e à observação dos efeitos dessa estrutura. A seguir, veremos quais elementos podem ser analisados.

Quais são as principais métricas da Análise de Redes?

Vimos nos capítulos anteriores que análise de redes trabalha em cima da teoria dos grafos e tem um forte viés matemático. Desse modo, a ARS utiliza, principalmente, medidas que auxiliam o pesquisador a compreender a estrutura da rede e o papel de suas várias partes no fenômeno que deseja observar. Neste capítulo, vamos estudar brevemente quais são essas métricas e como utilizá-las.

Essas medidas podem ser divididas em métricas de rede inteira e métricas de nó. Enquanto as primeiras são usadas para se obter perspectivas sobre a rede como um todo (por exemplo, o quão conectada é a rede analisada), as segundas visam compreender características dos nós individualmente e de sua posição na rede (por exemplo, quais desses nós são mais conectados na rede).

Métricas de Rede

[39]

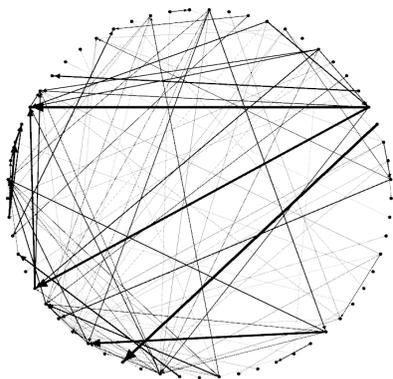
As métricas de rede são aquelas que indicam características da rede que está sendo analisada como um todo, e não apenas dos nós. Assim, são usadas para extrair características gerais da estrutura observada.

Densidade

A densidade está relacionada ao número de conexões presentes no grafo em relação ao número total possível. Ela reflete, assim, a proporção de conexões para o grafo completo, ou seja, aquele no qual existem conexões para

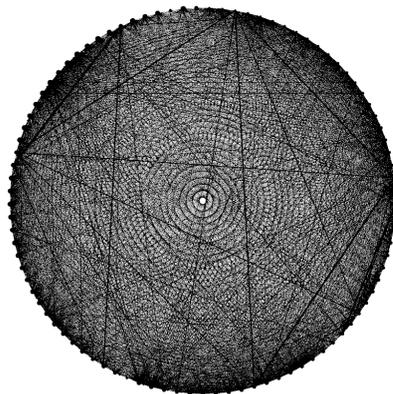
todos os pares de nós. Em outras palavras, a densidade demonstra o quão conectado um determinado grafo é/está. Assim, se examinarmos um grupo de nós no tempo, poderemos, por exemplo, compreender se eles estão se tornando mais conectados (isto é, se a rede está mais densa) ou menos. A medida de densidade máxima de um grafo é 1 (um).

Figura 10 - Rede pouco densa



Densidade = 0,003

Figura 11 - Rede muito densa



Densidade = 0,412

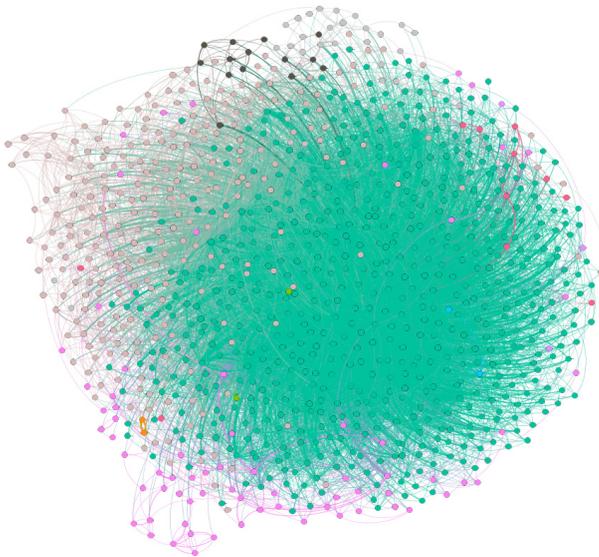
[40]

Nas figuras, vemos o exemplo de uma rede bastante densa (Figura 11) e uma rede pouco densa (Figura 10). Observamos essa diferença visualmente antes de vê-la numericamente. Essa diferença de densidade tem impacto na rede. No grafo mais denso, por exemplo, há mais chances de uma determinada informação circular, enquanto no grafo menos denso, essa chance é menor. A densidade também pode indicar outras características da rede, dependendo do que o pesquisador delimitou, no âmbito do seu trabalho, como conexão em sua coleta de dados.

Em termos de grafos sociais, a densidade da rede pode estar relacionada com a presença de um ou mais *clusters*. Ou seja, pode indicar redes nas quais os atores estão muito conectados ou nos quais há muita participação (positiva ou negativa). Normalmente, para que o pesquisador consiga avaliar a densidade de uma rede, é preciso compará-la com outras redes semelhantes.

A modularidade é uma métrica de grupabilidade, ou seja, está relacionada à tendência de determinados nós se conectarem entre si. Desse modo, observamos a tendência de um determinado nó estar mais conectado a nós de um subconjunto A do que aos de um subconjunto B.

Figura 12 – Grafo com diferentes módulos.

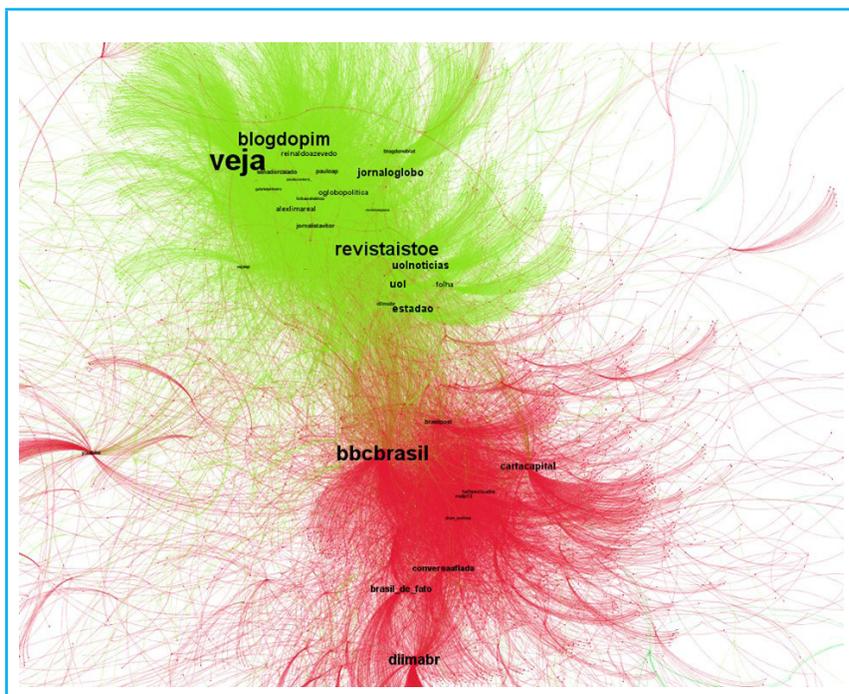


Na Figura 12 vemos um exemplo de um grafo modular. Os vários grupos coloridos representam os diferentes módulos da rede e mostram conjuntos de atores que tendem a se conectar mais entre si do que com os demais (*clusters*). Há um grande grupo verde, que representa um módulo, e outros grupos menores (principalmente rosa e cinza) que representam os demais módulos. Isso significa que há um grupo maior que tende a conectar mais nós.

Em termos de redes sociais, a modularidade é uma métrica que auxilia na identificação de subgrupos de atores dentro de um grande grupo. Com isso,

podemos observar grupos cujas conexões são mais frequentes entre si do que com o restante da rede.

Figura 13 - Exemplo de modularidade



No exemplo da Figura 13, vemos uma rede coletada em torno de uma temática específica constituída em conversações no Twitter. Neste caso, foram coletados os *tweets* que continham a palavra “protesto” entre os dias 16 e 21 de agosto de 2015. Na época, o Brasil vivia uma grande explosão de protestos contra a presidenta Dilma e uma insatisfação generalizada com a crise econômica do país. Junto aos protestos, existia uma clara divisão entre direita e esquerda que vinha se agravando desde o final das eleições de 2014. O grafo representa essa divisão. Nesta rede, as conexões representam citações nos *tweets*. As contas mais citadas são aquelas que estão presentes em tamanho maior no grafo.

O que é interessante aqui é que a aplicação da métrica de modularidade demonstra uma divisão da conversação em termos de citações. Há “buracos estruturais” entre os grupos verde e vermelho, o que demonstra que a informação circula menos entre esses dois grupos do que dentro de cada um deles. Os atores que citam as contas do grupo vermelho não citam as contas do grupo verde e vice-versa. Temos claramente dois *clusters* de citação. Com isso, há indícios da presença do fenômeno de “bolhas informativas” (PARISER, 2011) dentro de cada grupo. A separação entre os *clusters* indica que as diferentes informações que circulam em cada módulo dificilmente os transcendem, circulando entre eles. Ou seja, dependendo da posição na rede de cada ator, ele tende a receber mais informações com um tema A ou B. O conteúdo das mensagens também indica que esses dois *clusters* possuem posições políticas antagônicas, com o *cluster* vermelho defendendo o governo Dilma, e o *cluster* verde condenando-o. Vemos, neste contexto, quais veículos são utilizados por cada grupo para propagar informações, e vemos que as citações dos veículos tendem a refletir a percepção de sua posição política pela audiência. As bolhas, assim, parecem compreender posicionamentos políticos.

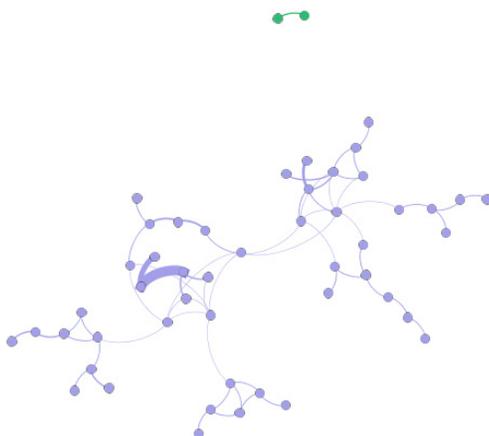
[43]

Componentes conectados

A rede também pode ser analisada pelo número de componentes conectados. Um componente conectado é um conjunto de nós dentro da rede no qual existe um caminho possível entre quaisquer dois desses nós. Ou seja, um componente conectado é um grupo de nós que está conectado de modo

integral, sem nós “flutuantes”. O número de componentes conectados indica o número de subconjuntos dentro de uma rede ou, em outras palavras, o número de redes desconectadas em um grafo.¹⁰

Figura 14 – Componentes conectados



[44]

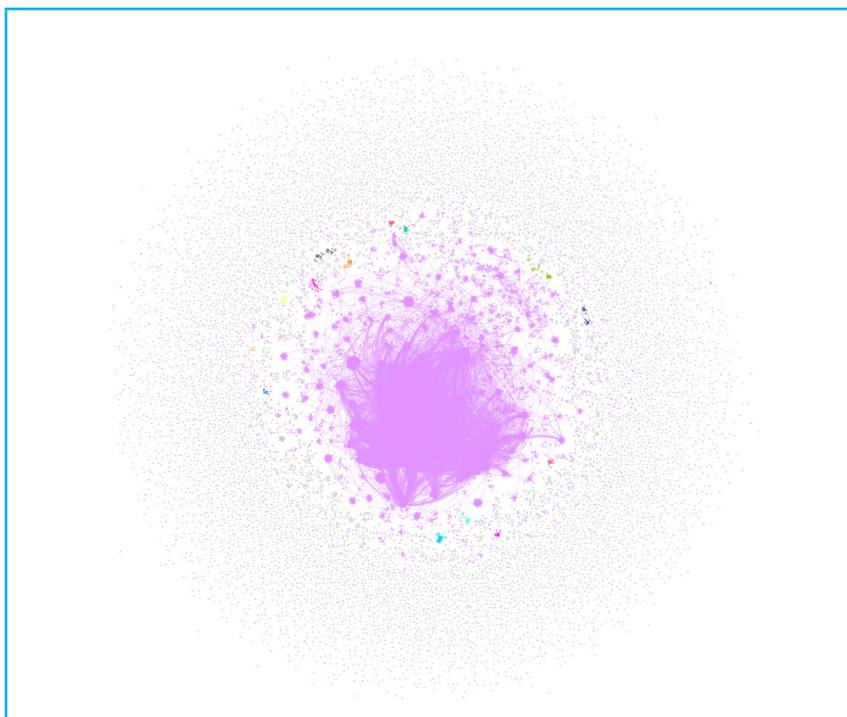


Na Figura 14 vemos um grafo com três componentes conectados, delimitados por cores. É importante observar que cada componente está inteiramente separado (ou seja, não possui arestas para os demais). Há um componente principal, ao centro, com um maior número de nós, e outros dois componentes acima e abaixo.

10 Dentro do mapeamento de uma rede inteira podemos ter subconjuntos “desconectados” da rede principal, como no caso. Esses conjuntos fazem parte da mesma rede porque estão unidos pelo evento ou conceito analisado, mesmo que não estejam diretamente conectados aos demais nós.

Em grafos direcionados, os componentes conectados podem ser fortemente conectados ou fracamente conectados. Nos componentes fortemente conectados, todos os nós chegam a todos os nós seguindo a direção da conexão. Já no componente fracamente conectado, todos os nós chegam a todos, mas, ignorando-se a direção da conexão, ou seja, utilizando-se arestas com qualquer direcionalidade. Essa diferenciação importa porque, na prática, conexões direcionadas podem indicar que há apenas um caminho de interação entre dois atores, significando que, embora a informação de A chegue a B, a de B não necessariamente chegue a A.

Figura 15 - Componentes conectados



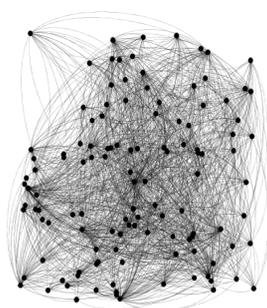
O uso da métrica de identificação dos componentes também permite a identificação de *clusters* na rede social. Na figura 15, construída em torno das conversações sobre a temática “*impeachment*” no Twitter no período de agosto de 2016, pouco antes do julgamento da então presidenta Dilma Rousseff, vemos um grão colorido pelos componentes conectados. Ao centro, em rosa, está um grande *cluster* de pessoas que se citam e citam as mesmas contas na conversação. Essa identificação de *cluster* traz uma ação ativista em torno da temática, na qual um grupo busca influenciar a conversação através de repetidos *retweets* e menções. Este grupo tinha posição fortemente favorável ao *impeachment*. Vemos também a presença de outros componentes, mas não conectados entre si. Dentro desses grupos também podemos observar a estrutura de circulação de diferentes tipos de informação.

Diâmetro da rede

O diâmetro da rede também é uma métrica de conexão. Ele mostra a maior distância entre os nós dentro do grafo. Essa medida pode indicar também o quão conectado é o grafo e o número de arcos que compõem o caminho mais distante entre dois nós.

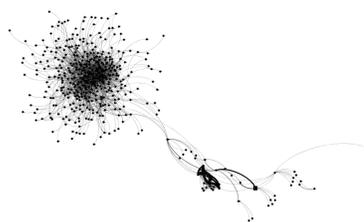
[46]

Figura 16 - Grafo com diâmetro 5



Diâmetro=5

Figura 17 - Grafo com diâmetro 7



Diâmetro=7

No exemplo, vemos dois grafos com diâmetros diferentes. Enquanto no primeiro caso (Figura 16), a distância maior são 5 arcos, no segundo, são 7. O diâmetro das redes indica a distância média que a informação precisa “navegar” para chegar a todos os atores. Vemos no formato das redes que aquela com diâmetro menor tem maior facilidade para que a informação circule, pois o caminho médio (*path*) é menor. Na outra rede (Figura 17), o caminho médio é maior, pois um grupo de nós está bastante distante do maior *cluster*. Ou seja, embora os nós do *cluster* estejam mais diretamente conectados ao centro, há um grupo de nós mais distante, que aumenta a distância média. Neste caso, a circulação de informação é mais prejudicada, pois há grande poder para os *gatekeepers* (nós que fazem a ponte entre os dois grupos).

Fechamento (*closure*)

O fechamento refere-se à redundância das conexões na rede. Conforme já explicamos, dizemos que conexões redundantes são aquelas que conectam o mesmo conjunto de nós (uma rede muito densa, portanto, tende a ter muitas conexões redundantes). Já conexões não redundantes são aquelas que servem como “pontes”, conectando grupos diferentes de nós que, de outra forma, não estariam conectados. É, portanto, uma medida também de interconexão. Quanto mais fechado o grafo, mais denso, mais redundante e possivelmente maior sua associação com capital social de segundo nível e laços sociais fortes. Comunidades, por exemplo, têm um fechamento mais alto que a média da rede

[47]

Métricas de nó

As métricas de nó são aquelas relacionadas aos atores, que vão caracterizar os nós ou de modo coletivo ou de modo individual. Elas calculam a posição desse nó na rede a partir de várias premissas. Compreender o posicionamento

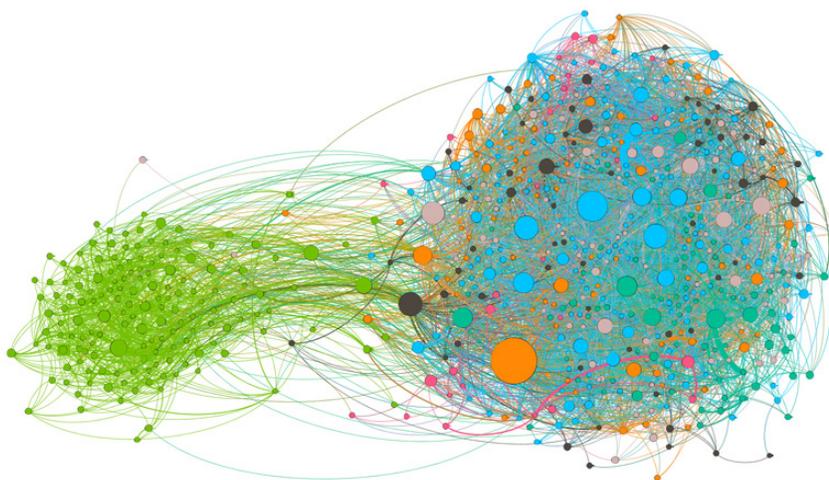
dos nós na rede é muito importante para também compreender seus papéis na rede.

Grau de entrada (*Indegree*)

Constitui-se do número de conexões que um determinado nó recebe. É uma métrica de grafo direcionado, ou seja, para sabermos quantas conexões um nó recebe, precisamos saber a direção de cada conexão. Grafos não direcionados ou simétricos não possuem *indegree*.

O grau de entrada é um indicativo de centralidade do nó na rede, mas relacionando-a a características específicas de acordo com o que se definiu como conexão na coleta de dados. Um ator que recebe muitas conexões sociais, por exemplo, pode ser alguém altamente popular nessa rede, o que também pode indicar que esse ator tem uma probabilidade maior, em relação a outros nós, de receber informações que circulam na rede. Além disso, em uma rede de citações, o grau de entrada pode indicar um ator que seja mais citado do que outros, também evidenciando uma posição privilegiada na rede. Relações negativas também podem indicar que um determinado nó com alto *indegree* tem um papel importante na rede. O grau de entrada também pode ser analisado de modo comparativo, mostrando o quão mais conectado é um determinado ator em relação aos demais na mesma rede.

Figura 18 - Grafo com tamanho dos nós indicando grau de entrada



No exemplo (Figura 18), podemos observar uma rede de conversação cujo tamanho dos nós está demarcado pelo grau de entrada. Quanto maior o número de conexões recebidas pelo nó, maior o seu tamanho na rede. Vemos, nesta rede, que alguns nós que estão em uma posição mais central na rede possuem maior número de conexões. Os nós maiores são aqueles que foram mais citados.

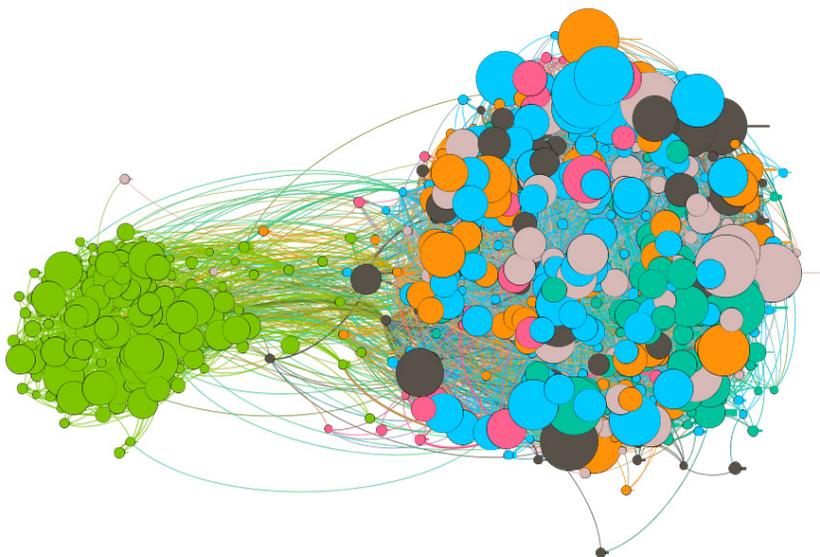
[49]

Grau de saída (*Outdegree*)

Constitui-se do número de conexões que um determinado nó faz. É também uma métrica de grafo direcionado, ou seja, grafos não direcionados ou simétricos não possuem *outdegree*. O grau de saída também é uma medida de centralidade de um nó, mas que mostra outras características comportamentais desses nós. Em uma rede social na qual as conexões são relações sociais, um alto grau de saída pode indicar atores que tentam se aproximar mais de outros na rede, buscando constituir relações. Não necessariamente essas conexões são recíprocas. Além disso, em uma rede de conversações, um nó com

alto *outdegree* pode indicar um ator que faz muitas citações, mais do que os demais, e que é possivelmente mais ativo na conversação. Diferentemente de um ator muito popular, pode representar um ator muito investido na conversação, ou seja, muito participativo.

Figura 19 – Grafo com tamanho dos nós indicando grau de saída

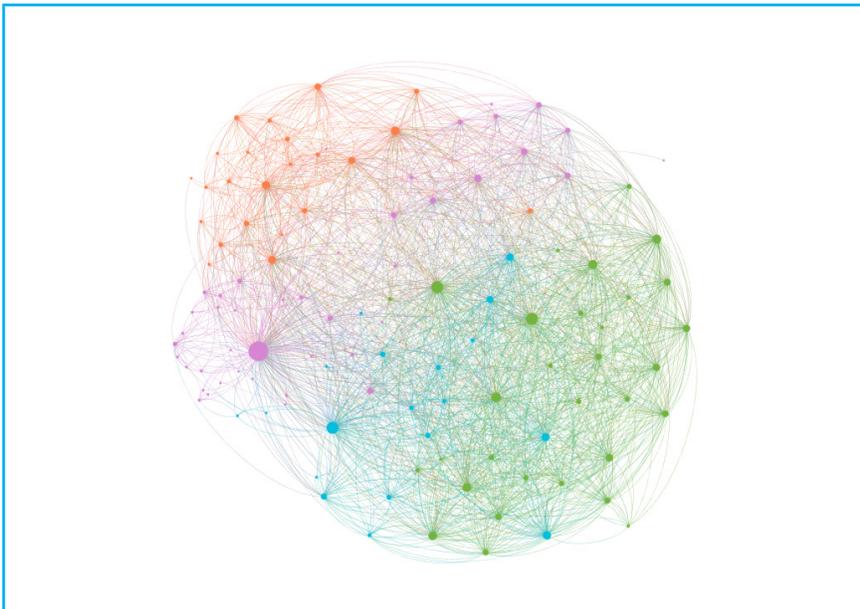


[50]

Na Figura 19, vemos o mesmo grafo, mas desta vez, o tamanho dos nós representa o número de conexões feitas por cada nó (*outdegree*). Observamos que os nós mais citados não são aqueles que mais fazem citações. Neste mapa, vemos que, ao contrário do grau de entrada, há uma maior proximidade de tamanho dos nós. Isso indica que a participação é mais uniformemente distribuída do que a citação. Ou seja, embora as interações estejam mais distribuídas pelos atores da rede, poucos deles são efetivamente citados pela maioria. Esses dados indicam que quem é muito citado tem uma posição mais privilegiada na rede, provavelmente com um nível maior de capital social (uma vez que as citações são positivas).

Em grafos simétricos ou não direcionados, há apenas a métrica de grau. Nesse caso, como o número de conexões é só um, não importa ou não há direcionamento. Assim, o grau representa o número total de conexões de um nó. É o caso, por exemplo, de redes de amigos no Facebook, em que a existência de um laço social depende totalmente da reciprocidade entre os atores. Um ator não pode ser amigo de outrem sem que a conexão seja bilateral, logo, não importa a direção dela (quem solicitou a “amizade”). Do mesmo modo que as métricas anteriores, o grau também é uma métrica de centralidade que indica a posição do nó na rede. Nesse caso, entretanto, como não se sabe o direcionamento da conexão, não é possível fazer tantas inferências como no caso dos graus de entrada e saída.

Figura 20 – Rede do Flickr



A rede da figura 20 está estabelecida em torno de conexões (amizades) na ferramenta de imagens Flickr. Neste caso, os nós (que representam os atores/contas) possuem apenas grau, pois a conexão não tem direcionamento. Esta é uma rede egocentrada, ou seja, mapeada a partir de um ator-ego (o maior nó, em rosa, à esquerda). Foram mapeados os amigos de ego e suas interconexões. Sobre esta rede foi aplicada uma métrica de modularidade, que avaliou a tendência entre os amigos do ator central de se conectarem entre si. Observamos que ego tem, pelo menos, quatro grupos associados a si. Em uma posterior análise qualitativa, feita em cima da modularidade, foi possível perceber a que espaço social cada módulo estava relacionado: em laranja estavam seus amigos da faculdade, em rosa, sua família, em verde, os amigos do trabalho, e em azul, outras conexões mais esparsas. Esta é uma rede bastante conectada, o que também pode indicar formas de capital social mais relacionadas às conexões mais fortes, pois é uma rede que apresenta conexões que estão abertas à circulação de informação. Seu formato mostra que há conexões redundantes (fechamento) e densidade (embora essa métrica deva ser analisada sempre em relação a outros casos).

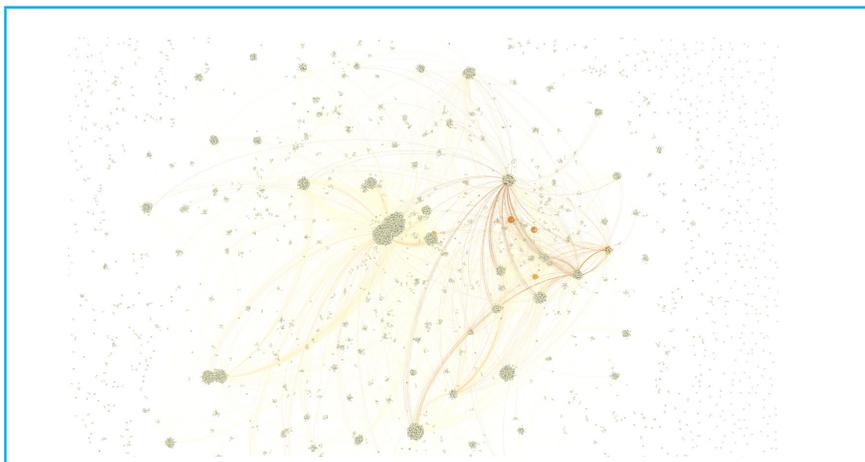
Grau de intermediação (*Betweenness*)

[52]

O grau de intermediação é uma das medidas mais interessantes e, ao mesmo tempo, mais complexas em redes sociais. Basicamente, a métrica indica o quão “ponte” algum nó é, ou seja, o quanto um determinado ator conecta diferentes grupos, indicando, assim quais atores conectam diferentes *clusters* na rede.

Um ator que possui um alto grau de intermediação não necessariamente é aquele com maior grau de entrada ou saída. O grau de intermediação calcula outro tipo de centralidade. Aqui, a investigação é a respeito de quais atores são capazes de manter a estrutura da rede, cujo papel conectivo é mais central. Em uma estrutura de comunicação (por exemplo, uma empresa), esses atores são aqueles que conectam diferentes grupos/setores no processo. A função de “ponte” de um determinado ator é importante, pois mostra que aquele nó pode atuar como filtro de informação entre dois grupos, e indica um posicionamento fundamental na circulação de informação na rede.

Figura 21 – Rede de conversação sobre a nova vinheta da Globeleza para o Carnaval de 2017



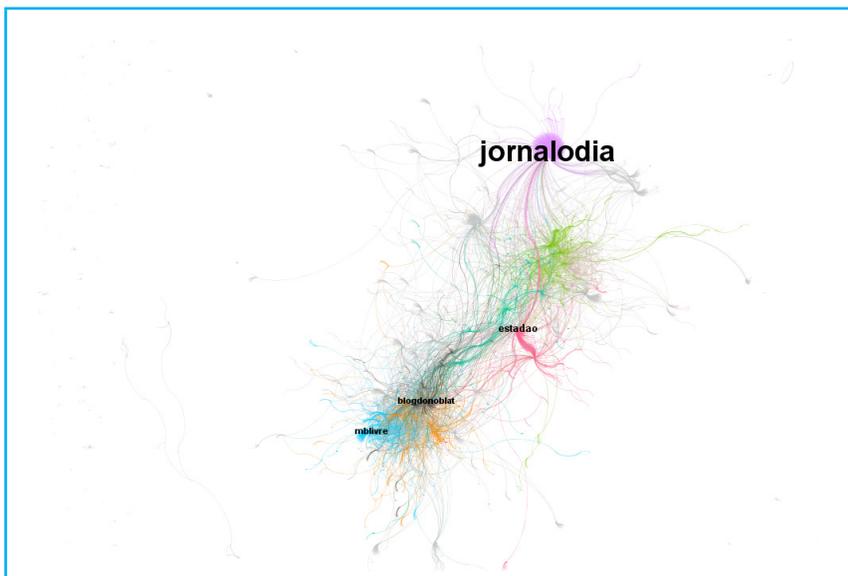
Na figura 21 vemos um exemplo de uma rede de conversação em torno da vinheta da nova Globeleza, da Rede Globo, para o Carnaval de 2017, com dados de 29.727 *tweets* e 23.141 contas individuais. Essa vinheta foi bastante comentada na internet justamente porque quebrava o paradigma das anteriores, cujo foco era a nudez de uma mulher negra. Agora, o filme trazia os vários tipos de Carnaval do País, com dançarinos vestidos. No exemplo, os nós maiores e mais escuros (em tom de laranja) são aqueles com maior grau de intermediação. É importante observar que esta é uma rede pouco densa, com *clusters* dispersos, o que aumenta o grau de intermediação da maioria dos nós, pois muitos fazem ponte entre os vários clusters. Os mais escuros têm um grau de intermediação de duas a três vezes maior que a média dos demais.

Observando esta métrica, podemos identificar quais usuários conseguiram “falar” para grupos mais diferentes sobre a vinheta (ou seja, sendo “pontes” entre os diferentes grupos). Sem nenhuma surpresa, essas contas são todas de humor e fizeram piada com a mudança. É bastante comum observarmos que aquelas mensagens que têm maior poder para “cruzar” fronteiras em grupos sociais parecem ser, justamente, aquelas humorísticas. Vemos, entretanto, que há bastante desconexão na rede, indicando a circulação de conteúdos diferentes e a dificuldade de determinadas informações transporem fronteiras.

Centralidade *eigenvector* (centralidade de autovetor)

Este tipo de centralidade avalia a importância de um nó na rede em cima da, digamos, qualidade de suas conexões. Nós com maior centralidade *eigenvector* são aqueles mais conectados a outros nós cujas conexões são relevantes. Em uma rede social, poderíamos considerar que essa centralidade adiciona a influência indireta (conexões das conexões) à influência direta (conexões diretas) de um determinado ator. No Twitter, por exemplo, um nó com alta centralidade *eigenvector* em uma rede de seguidores seria aquele cujos seguidores também possuem muitos seguidores. Ou seja, quanto maior o número de seguidores das pessoas que seguem uma determinada conta A, maior será a centralidade *eigenvector* desta conta. Essa métrica é muito utilizada para apontar influenciadores nas conversações, pois pode indicar nós que, pela sua participação, embora modesta, consigam dar visibilidade a uma mensagem para outros nós importantes, que podem aumentar essa visibilidade de modo exponencial pela sua quantidade de contatos, por exemplo.

Figura 22 – Grafo de conversações em torno de *impeachment*



No grafo da figura 22 vemos as conversações em um dado momento no Twitter sobre o *impeachment* da então presidenta Dilma Rousseff. Os nós que estão marcados são aqueles que possuem maior centralidade *eigenvector*. Observe-se que são o “Jornal O Dia”, o “Estadão”, o “blog do Noblat” (contas de mídia) e o “MBLivre” (conta do movimento que pregava o *impeachment*). Essas contas, desse modo, possuem um papel central na construção da narrativa desse fato, no momento em que os dados foram coletados, pois suas conexões as fazem particularmente importantes e tendem a tornar suas mensagens mais visíveis na rede. Podemos dizer, também, que esses atores, neste momento particular, tinham maior capital social na rede, valor este construído pelas suas conexões e seus valores subsequentes. Ou seja, são os demais atores que dão a estas contas sua visibilidade e que as tornam importantes no quadro geral.

[55]

Coefficiente de clusterização

O coeficiente de clusterização indica o quão conectado a um *cluster* está um determinado nó em sua vizinhança. Assim, um alto coeficiente de clusterização indica que um determinado nó está muito conectado a seus vizinhos. Alguns programas calculam o coeficiente médio de clusterização da rede, que pode indicar o quão conectada é a rede como um todo. O coeficiente médio é um indicativo relevante do grau de interconexão dos nós daquela estrutura, o que auxilia na percepção de *clusters*, por exemplo.

Closeness (Proximidade)

A proximidade representa o quanto um determinado nó está “mais próximo” dos demais. Ela é determinada pelo comprimento dos caminhos de cada nó. Por conta disso, também é uma medida de centralidade do nó, que mostra sua posição na rede e sua capacidade de receber, por exemplo, informações por esse posicionamento. Essa é uma métrica interessante para avaliar pequenos grupos dentro de um grande conjunto de nós.

Como coletar, analisar e visualizar dados para Análise de Redes?

Neste capítulo discutiremos modos de coleta, análise e visualização de dados de redes sociais. O objetivo aqui é explicitar modos de abordar metodologicamente esses procedimentos e suas implicações, bem como apresentar ferramentas que possam auxiliar nesse tipo de trabalho.

Métodos de coleta de dados

Há vários modos de coleta de dados para estudos de análise de redes. Entretanto, em primeiro lugar, é preciso decidir com qual tipo de rede se está trabalhando: se a rede é egocentrada ou rede inteira. Caso a rede seja egocentrada, é preciso também observar qual a distância de ego (grau de separação) que utilizaremos para a coleta de dados (é preciso ainda levar em conta que, quanto mais graus, maior a quantidade de dados).

Podemos coletar dados de modo qualitativo (através de entrevistas com os atores da rede, por exemplo) ou quantitativo (através da coleta de dados de redes existentes). Cada uma dessas formas de coleta tem vantagens e desvantagens. Formas qualitativas, por exemplo, permitem ao pesquisador um *insight* maior na percepção dos atores sobre suas relações. Já formas mais automáticas retiram essa subjetividade das relações. E qual seria a melhor forma? Depende dos objetivos da pesquisa e do tamanho da rede a ser coletada. Redes muito grandes necessariamente precisarão de algum tipo de coleta automática de dados. Já redes menores podem ser coletadas pelo próprio pesquisador.

Coletas qualitativas

A coleta de dados de modo qualitativo pode ser feita através de entrevistas, questionários ou mesmo da observação do pesquisador de um determinado ambiente ou grupo. As técnicas de coleta podem ser diversas e, inclusive, ser combinadas, dependendo do caso. Assim, por exemplo, o pesquisador pode analisar as interações de alunos em sala de aula, determinando que os alunos serão os nós da rede, e as interações entre eles, suas conexões. Esses dados poderão ser coletados através da observação *in loco* (o pesquisador anota as interações entre os alunos de acordo com a sua observação do grupo) e através de entrevistas com os alunos. Nesse caso, os dois conjuntos de dados poderão ser utilizados de modo comparativo, para compreender os pontos comuns e os pontos divergentes entre a percepção que os alunos têm de suas conexões e aquela do pesquisador.

Dados de entrevistas e questionários sobre relações sociais são relevantes, mas extremamente subjetivos. Caso sejam utilizados como modo de coleta de dados, é muito importante que o pesquisador tenha consciência desta subjetividade e a analise no trabalho.

A coleta de dados deverá resultar em uma matriz, que indicará as relações entre os nós. Cada nó, assim, será um dos atores ou suas representações, enquanto as conexões serão alguma forma de relação entre esses atores definida pelo pesquisador. A matriz, assim, pode representar quantas vezes um ator A conversou com seus colegas de grupo. Nesse caso, o pesquisador vai numerar cada uma dessas vezes de modo a dar “peso” para as diferentes conexões de A e constituir um mapa que permita compreender as relações entre esses atores.

Coletas quantitativas

As coletas quantitativas de dados geralmente focam em bases de dados preexistentes (por exemplo, dados de interação de ferramentas de mídia social). Nesses casos, é bastante comum que se utilizem ferramentas construídas pelos pesquisadores (é comum que essas bases de dados específicas exijam

algum tipo de ferramenta também específica para elas). Também é possível utilizar ferramentas de coleta de dados já existentes, desde que o pesquisador conheça o funcionamento da ferramenta, da base de dados e suas limitações. Geralmente, essas ferramentas de coleta automática de dados são denominadas *crawlers*.

É cada vez mais raro que bases de dados possuam dados completamente acessíveis aos pesquisadores. Mesmo o Currículo Lattes, por exemplo, ferramenta de dados de pesquisadores das universidades de todo o país e objeto de vários estudos de colaboração e produção científica das diversas áreas,¹¹ apesar de conter dados públicos, hoje conta com um *captcha* para impedir a coleta automatizada de dados. Em outros casos, onde há banco de dados privado (como o Facebook, em que os dados são de propriedade da empresa), os dados disponíveis para o público são cada vez mais reduzidos.¹²

Há alguns tipos de *crawlers* disponíveis publicamente. A maioria deles é focado em ferramentas específicas ou banco de dados específicos cuja estrutura é mais ou menos estável (como o Twitter e o Facebook, por exemplo). Entretanto, não é incomum que pequenas mudanças na estrutura dos bancos de dados dessas ferramentas-repositórios (cujas bases de dados são privadas) “quebrem” esses *crawlers*, exigindo que ajustes sejam feitos. Muitas vezes, os ajustes são bastante simples, mas necessitam de alguém que compreenda a estrutura e os modos de programação da ferramenta.

11 Vide, por exemplo, Oliveira, Santarem e Santarem, 2009; Balancieri et al., 2005; Digiampietri, 2015; Bastos, Zago e Recuero, 2016.

12 No caso específico do Facebook, a equipe de pesquisa da própria empresa costuma divulgar estudos realizados com os dados. Entretanto, como o banco de dados não é acessível a ninguém, exceto a este grupo, há uma série de fatores complicadores com relação à fidedignidade dos mesmos. Várias dessas análises são divulgadas de forma anedótica na própria ferramenta, na página Facebook Data Science - <<https://www.facebook.com/data/?fref=ts>>.

ALGUNS CRAWLERS POPULARES PARA MÍDIA SOCIAL

Neste quadro, apresentamos alguns *crawlers* disponíveis publicamente que podem ser utilizados pelos pesquisadores para coletar dados de mídia social. É uma boa forma de começar a coletar dados para analisar.

yTK (<https://github.com/540co/yourTwapperKeeper>): *crawler* para Twitter. O YouTwapperKeeper está um pouco desatualizado e é bem complicado de instalar. Precisa de um servidor (porque fica rodando 24 horas as buscas que foram solicitadas) e uma conexão bastante estável para não sair do ar toda hora. Essencialmente, há também uma série de riscos ao utilizá-lo, porque tecnicamente a API do Twitter tornou esse tipo de coleta ilegal. A qualquer momento, portanto, pode ser barrado. As vantagens: ele coleta *tweets* do momento em que se iniciou a busca em diante, e tende a coletar uma quase totalidade dos *tweets* no período em que estiver rodando.

NodeXL (<https://nodexl.codeplex.com/>): um dos melhores sistemas para quem está começando. Não precisa servidor, você pode rodar de casa mesmo. Tem uma interface bastante amigável e funciona como um *layer* para o Excel (é preciso ter Excel). O Node também apresenta junto um *crawler* para Twitter, Facebook, Youtube e Flickr, que é bastante útil. Na versão gratuita, tem uma série de limitações, mas dá para começar. Na versão paga, pode auxiliar em coleta de dados para análise de sentimento, além de permitir o uso de uma série de métricas mais complexas.

NetVizz (<https://apps.facebook.com/netvizz/>): Funciona via Facebook e para o Facebook. Coleta dados de grupos e de páginas de busca e permite que sejam exportados. Permite busca de conexões entre as páginas e outras ações que o NodeXL, por exemplo, não faz. Seu uso é relativamente simples, mas a coleta de dados pode ser bastante demorada.

Gephi (<https://gephi.org/>): O Gephi é um aplicativo popular para a visualização e manipulação dos dados, mas recentemente acrescentou um *plugin* que permite também a coleta de dados do Twitter. Entretanto, exige que se tenha uma autorização para o acesso à API.

Uma outra possibilidade é aquela de utilizar bancos de dados públicos, que foram recolhidos por outros grupos. Com a cobrança de transparência nos dados de pesquisa, eventos, grupos e instituições costumam disponibilizar conjuntos de dados (geralmente anonimizados) para o público em geral. Embora sejam dados “de segunda mão”, ou dados secundários, também podem ser utilizados, desde que o pesquisador tenha conhecimento de seus

métodos de coleta e das possíveis limitações que possam ter incorrido nessas coletas.

A questão da privacidade

Uma questão fundamental para ser considerada na coleta de dados é a da privacidade dos atores. Atualmente, comitês de ética de todo o mundo têm questionado o uso de bases de dados sem o consentimento dos envolvidos nela. São questões muito importantes, pois a análise de dados pode expor, ainda que sem intenção, indivíduos em situação de risco ou mesmo que não compreendem a dimensão do uso desses dados ou não sabem que seus dados estão sendo usados. Por conta disso, é sempre importante anonimizar os dados coletados, mesmo que estes sejam de bancos de dados públicos, dificultando o reconhecimento de atores individuais de modo que possam ser prejudicados pela pesquisa. A disponibilização dos bancos de dados utilizados também é bastante comum, mas esses bancos de dados são geralmente disponibilizados de modo anônimo. Finalmente, figuras públicas e entidades têm essa questão mais relativizada, pois não têm expectativa de privacidade quando utilizam essas ferramentas para fazer pronunciamentos públicos. É o caso, por exemplo, do uso do Twitter por presidentes ou mesmo veículos de mídia. Há vários guias que podem ser utilizados para auxiliar na discussão sobre a ética com os dados que podem ser consultados *online*.¹³

[61]

Visualizando dados com ARS

Uma vez coletados os dados de mídia social, inicia-se então o processo de análise. Em alguns casos, será preciso ainda fazer alguma limpeza para

13 Ver: <https://plato.stanford.edu/entries/ethics-internet-research/> e <https://aoir.org/reports/ethics.pdf>.

adaptar os dados ao que se precisa. Após, é preciso representar as relações entre os atores através de uma matriz que possa ser objeto da análise de redes. Basicamente, a análise dá-se através da aplicação das métricas de rede para os dados que foram coletados. Assim, investiga-se a estrutura da rede e/ou a posição dos nós dentro dessa rede, de acordo com as questões de pesquisa postas. A visualização dos dados pode também ser utilizada de modo paralelo, buscando compreender essas estruturas.

Para a visualização de dados, há diferentes algoritmos, cada qual com premissas próprias. É importante salientar, entretanto, que a visualização é uma ferramenta da análise de dados, mas não pode ser utilizada sem algum tipo de centrimento nas métricas. A visualização é uma forma de mostrar aquilo que as métricas calculam, e não uma justificativa *per se*. Ela deve ser, portanto, visualmente informativa daquilo que as métricas demonstram. Por conta disso, é sempre importante descrever quais métricas e algoritmos foram utilizados para a visualização, detalhando os processos metodológicos. Dependendo do foco do trabalho, também pode ser importante apresentar os valores das métricas, para fins comparativos. Trabalhos com análise de redes sociais, pelo seu caráter quali-quantitativo ou quanti-qualitativo, portanto, precisam de um capítulo metodológico bastante detalhado.

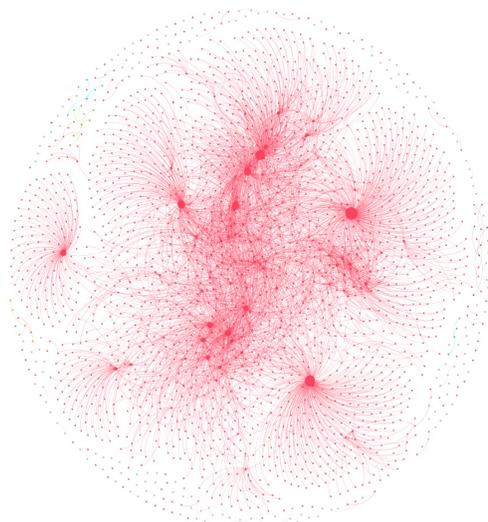
Sobre a visualização, há diversos algoritmos que são utilizados e muitos que são desenvolvidos pela comunidade. Em geral, a maioria deles é apresentada em algum artigo pelos autores do algoritmo, e o pesquisador pode investigar quais são as premissas da visualização para compreender qual delas é a mais adequada para seu trabalho. Discutiremos a seguir alguns dos algoritmos mais comuns utilizados para visualização e suas vantagens e desvantagens.

Há vários tipos de algoritmos utilizados para desenhar os grafos. Alguns dos mais utilizados são aqueles “direcionados pela força” (*force-directed*). Esses algoritmos atribuem valores para os nós baseados em sua posição na rede e, utilizando um conjunto de forças de repulsão e atração (de modo mais ou menos análogo a um sistema gravitacional), posicionam esses nós buscando um equilíbrio no sistema. A partir desse princípio, os nós que estão conectados tendem a se atrair, enquanto que aqueles que não estão, tendem a se repelir. Esses algoritmos trabalham com um conjunto de iterações através das quais o grafo vai sendo moldado, iterações estas sustentadas por premissas específicas. Dentre os algoritmos dessa classe, estão:

Fruchterman-Reingold: apresentado pelos autores (FRUCHTERMAN; REINGOLD, 1991) em um artigo de 1991. Atribui vetores de força para os nós, cuja soma determina a direção do nó, utilizando um conceito denominado “temperatura global”. A ideia é que essa temperatura controla os movimentos dos nós (quando a temperatura está mais alta, os nós movem-se mais rápido). O grafo vai sendo resfriado a cada iteração, até que o sistema atinja um equilíbrio. O algoritmo minimiza a interseção de conexões e baseia-se na premissa de que nós interconectados são atraídos entre si (como os autores explicitam).

Esse algoritmo apresenta grafos em formato circular e tende a colocar os nós que não têm conexões ou têm menos conexões para a periferia do grafo e centralizar os *clusters*. Há uma certa uniformidade ou simetria na construção do grafo, o que o torna visualmente interessante em casos de grafos pequenos ou com um conjunto de clusters. É um algoritmo que não funciona bem para grafos muito grandes, tendendo a ser lento e a nem sempre apresentar um resultado visualmente informativo.

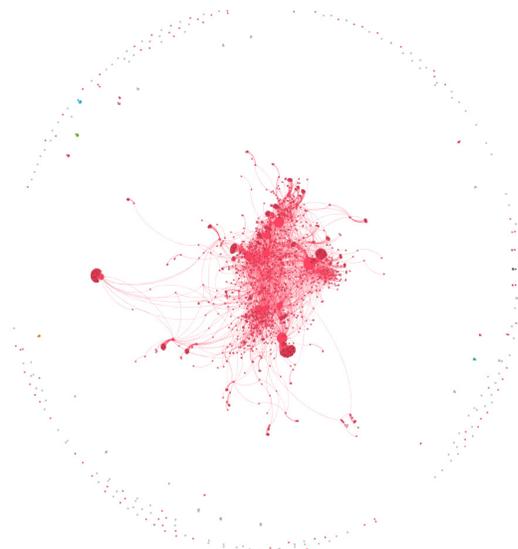
Figura 23 – Visualização de grafo com o algoritmo FR.



Na Figura 23 temos uma visualização de um grafo de 4.794 arestas e 2.239 nós através do algoritmo. Note-se que a visualização não tem seus parâmetros modificados. Vemos aqueles nós com maior *indegree* em tamanho maior. Na periferia, vemos os nós menos conectados, e, ao centro, aqueles com maior quantidade de conexões entre si (não necessariamente maior *indegree*).

Force Atlas: Force Atlas é um algoritmo desenvolvido pelo mesmo time que desenvolveu o Gephi. Atua também de modo a aproximar os nós conectados e afastar os que não estão conectados, mas foca em redes maiores que o algoritmo anterior. Assim, tende a trazer para o centro os *clusters* e afastar os nós que não estão conectados ou que estão menos conectados. O objetivo é espacializar redes de modo mais claro, permitindo sua interpretação. Permite, no Gephi, que o pesquisador modifique suas interações, por exemplo, modificando a força de atração ou repulsão dos nós ou modificando quais nós (dependendo da qualidade) ficarão mais centrais (terão maior força de atração). O Force Atlas também utiliza o peso dos nós para montar o grafo. É um algoritmo que exige que o pesquisador saiba o momento de parar quando achar que o grafo está legível o suficiente.

Figura 24 – Visualização de grafo com o algoritmo Force Atlas

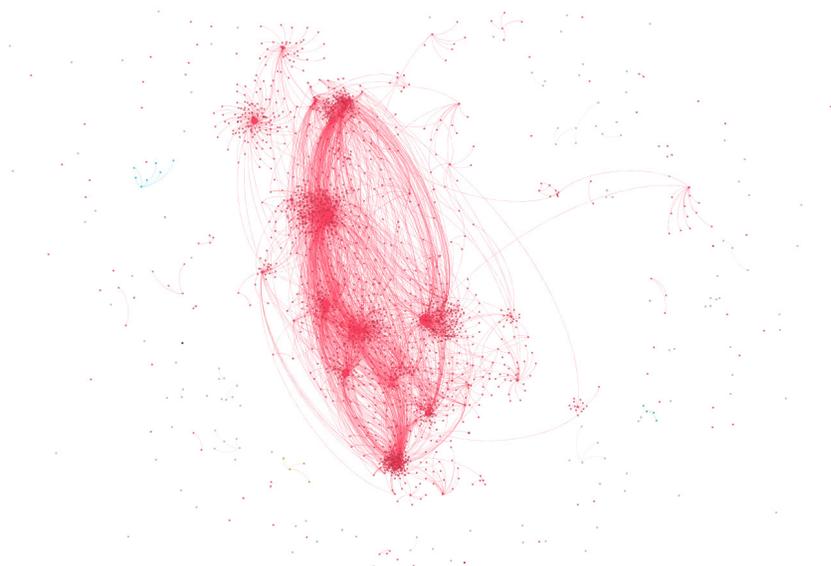


Na Figura 24, vemos o mesmo conjunto de dados anterior, mas desta vez, construído com o Force Atlas padrão, sem distribuição pela atração ou ajuste de tamanhos. Vemos que o algoritmo tende a “juntar” mais os nós e as conexões ligadas ao componente central, enquanto afasta aqueles nós e conexões desconectados ou com poucas conexões. Em torno dos maiores nós, os grupos de nós conectados apenas a ele ficam também mais próximos pela força de atração.

Force Atlas 2: é um algoritmo construído em cima do Force Atlas (JACOMY et al, 2014). Funciona em cima das mesmas premissas, mas é mais rápido (e por isso, mais adequado para grafos maiores), pois trabalha com parâmetros de escalabilidade. Assim, ele funciona com os nós se afastando, e as conexões se atraindo (o que faz com que os nós com maior número de conexões terminem por ficar mais centrais no grafo). Assim como seu antecessor, requer que o pesquisador pare quando achar que o grafo está legível o suficiente. O algoritmo também permite uma série de manipulações no Gephi, no qual o pesquisador pode fazer com que os nós fiquem mais afastados ou mais juntos, mexer no espaço total do grafo etc.

OpenOrd (MARTIN et al, 2011): trabalha principalmente com grafos não direcionados, auxiliando a visualizar os *clusters*. É um desenvolvimento em cima do algoritmo de FR, trabalhando com um número fixo de iterações ordenadas, cujas fases podem também ser selecionadas pelo pesquisador no Gephi. O foco aqui é tentar separar os *clusters* do grafo. Este algoritmo pode trabalhar com grafos muito grandes. O OpenOrd tende a trazer os nós com maior grau para o centro do grafo. O algoritmo termina sua atuação sozinho, não precisa da interferência do pesquisador para parar.

Figura 25 – Visualização de grafo com o algoritmo OpenOrd

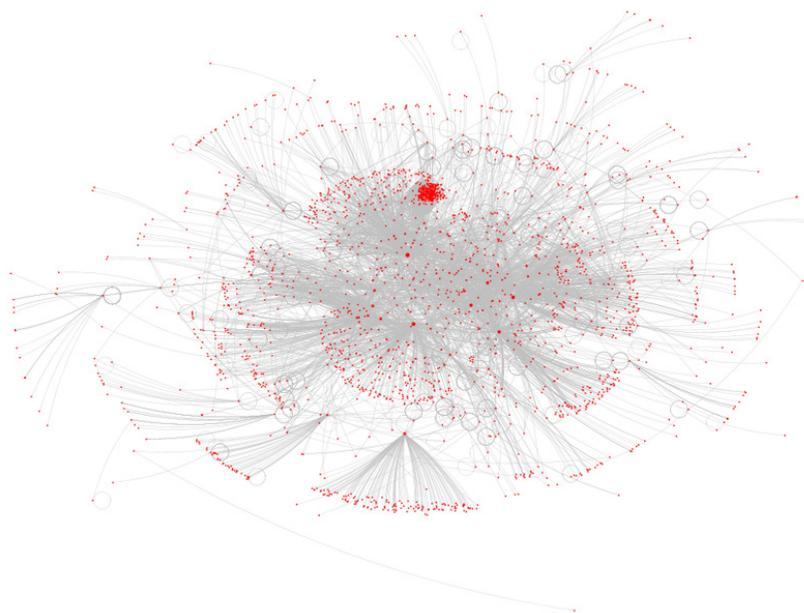


Na Figura 25 vemos os mesmos dados já plotados anteriormente, mas desta vez, com o algoritmo OpenOrd. Notamos que os *clusters* ficam mais visíveis e destacados, diferentemente das visualizações anteriores. Os nós e grupos mais desconectados são afastados do centro, enquanto os *clusters* são separados entre si, permitindo identificá-los mais claramente.

[66]

Harel-Koren: também é um algoritmo multiescala e direcionado pela força. Tende a juntar nós próximos, sobrepondo-os como se fossem um único nó. Em algoritmos multiescala, a estrutura do grafo é simplificada para que se tenha rapidez, enquanto propriedades topológicas características são preservadas (dentro do possível)

Figura 26 – Visualização de grafo com o algoritmo Harel-Koren

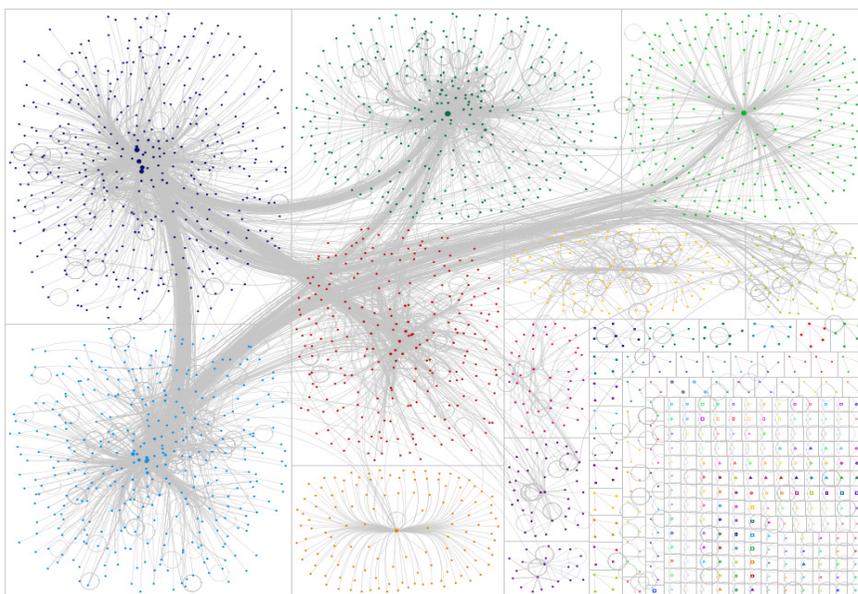


Na Figura 26 vemos a visualização dos dados de acordo com este algoritmo. De modo semelhante ao OpenOrd, ele separa os *clusters*, mantendo os nós de cada um mais próximos entre si.

Alguns *softwares* ainda permitem modos de visualização diferentes, baseados em algoritmos de grupabilidade. O NodeXL, por exemplo, permite que os grupos sejam atribuídos com base nos *clusters* presentes nos dados e que esses *clusters* sejam visualizados em caixas separadas, cada qual com as conexões entre os grupos evidentes. É um modo diferente de visualização, que permite identificar de modo mais claro os diferentes grupos de um mesmo conjunto de dados, a partir do cálculo das métricas.

[67]

Figura 27 – Nós agrupados em caixas, *layout* do NodeXL



Na Figura 27, vemos um grafo feito no NodeXL, no qual os nós foram separados por *clusters*, e os *clusters*, visualizados em “caixas” separadas, que permitem divisar quais são os grupos. Os dados referem-se à participação do público em um debate com os candidatos a prefeito de Porto Alegre, em 2016. As conexões são agrupadas para permitir uma melhor visualização do grupo. Cada grupo de nós aqui representa os apoiadores de um determinado candidato ou de outro. Os dois grupos azuis possuem maior clusterização e são os grupos que apoiavam os candidatos que efetivamente foram para o segundo turno.

Há ainda outros *softwares* que auxiliam neste processo e na criação dos grafos que permitem a visualização dos dados. Veja no quadro a seguir.

SOFTWARES DE ANÁLISE E VISUALIZAÇÃO DE DADOS

NodeXL (<https://nodexl.codeplex.com/>): como vimos no quadro anterior, é uma ferramenta que coleta e analisa dados de mídia social. Mas pode ser bastante útil para outros tipos de dados, bastando que se importe a matriz ou que esta seja construída ali. Possui várias métricas, mas a visualização dos dados é um pouco limitada, com poucos algoritmos. É útil para trabalhar com grafos pequenos (até 20 mil nós). Acima disso, começa a ficar muito lento. Funciona apenas em Windows.

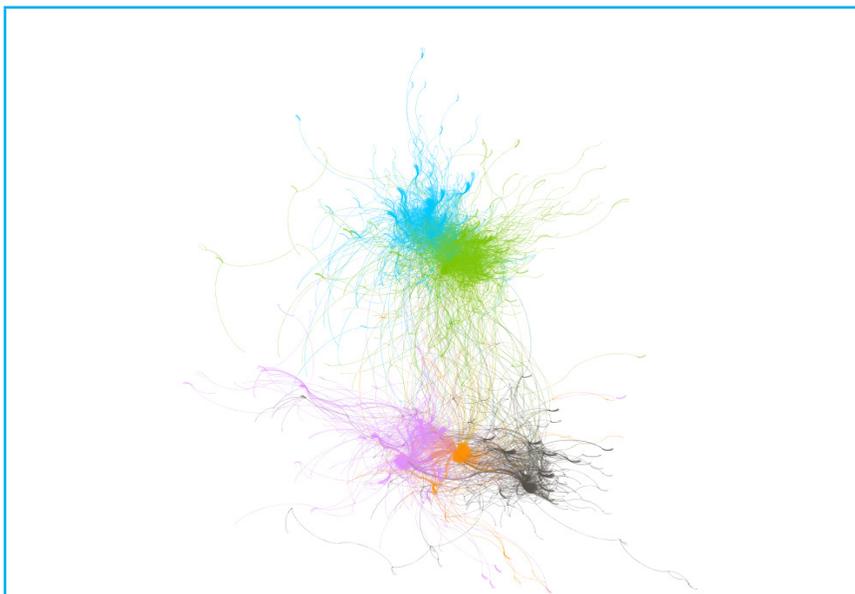
Gephi (<https://gephi.org/>): é uma das ferramentas mais utilizadas, não apenas por ser gratuita, mas igualmente por ser aberta e incluir diferentes *plugins* feitos por usuários. Entretanto, possui uma interface bastante difícil e uma curva de aprendizagem maior. Entre suas partes mais fortes está a visualização de dados e a inclusão frequente de novas métricas. Entretanto, exige que o usuário acompanhe sempre as atualizações. Também funciona bem com matrizes maiores que o NodeXL, mas tende a ficar lento com matrizes muito grandes. É preciso aumentar a memória dedicada a ele manualmente para conseguir trabalhar com dados maiores.

NetDraw (<https://sites.google.com/site/netdrawsoftware/home>): é uma ferramenta relativamente limitada, mas gratuita. É antiga, e tem também uma curva de aprendizado maior que o NodeXL, mas funciona também em Windows. Como o NodeXL, é uma das ferramentas mais simples de ser compreendida por um usuário novato.

Pajek (<http://mrvar.fdv.uni-lj.si/pajek/>): o Pajek é uma das ferramentas mais conhecidas para análise de redes sociais. Também tem uma interface um pouco difícil, mas há várias dezenas de tutoriais e livros de ajuda para quem quer começar. Os grafos não são tão bonitos quanto os do Gephi, mas possui uma quantidade bem útil de métricas e formas de visualização.

É importante salientar que geralmente é possível importar e exportar arquivos entre os vários *softwares* de análise, o que permite criar modos mais personalizados de combinações de métricas e visualização. Assim, por exemplo, é possível coletar dados com o NodeXL e visualizá-los no Gephi. Arquivos de rede em geral têm formato NET ou GRAPHML, que são abertos pela maioria dos programas.

Figura 28 – Exemplo de visualização



Como escolher a melhor forma de visualização?

A melhor forma de visualizar dados depende do que se deseja mostrar com esses dados. Assim, se o objetivo é identificar, por exemplo, atores que influenciaram mais a rede, a visualização deve destacar esses nós. Se o objetivo é mostrar os principais grupos relacionados, outros tipos de visualização são mais adequados. Neste exemplo, coletamos os *tweets* que mencionavam o deputado Jair Bolsonaro (PSC-RJ) durante a eleição para a presidência da Câmara dos Deputados, para a qual este era candidato. Por se tratar de uma figura bastante polêmica, pretendia-se ver se era possível identificar a rede de apoiadores e contraditores do deputado. A partir de uma rede de 23.165 nós que mencionaram o parlamentar no período de coleta (dia 02 de fevereiro de 2017) e os 32.460 *tweets* coletados, primeiramente calculamos o grau dos nós (onde o grau de entrada indica o número de vezes que um dado nó foi citado ou retuitado, e o grau de saída, a quantidade de vezes que este nó citou ou retuitou outro). Após o cálculo dos graus, aplicamos um algoritmo de modularidade, com o objetivo de separar os nós por vizinhança, uma vez que a hipótese seria de que os apoiadores de Bolsonaro tendem a se citar entre si, assim como seus detratores, gerando dois núcleos claros de atores.

Uma vez incluídos os cálculos, criamos uma visualização da rede no qual os nós com maior *indegree* seriam maiores (de modo proporcional), enquanto aqueles com menor *indegree* seriam menores. A hipótese aqui é de que poderíamos observar melhor as redes de apoio e as redes contrárias a partir dos nós mais citados por elas. Aplicamos, assim, o algoritmo Force Atlas 2 (pelo tamanho da rede) até que os dois *clusters* fossem separados e pudéssemos verificar qualitativamente suas citações. O resultado é o que vemos na Figura 27: há dois grandes conjuntos de nós, um superior, com dois *clusters* (verde e azul), e um inferior, com várias cores, mas principalmente, laranja, rosa e preto. Os *clusters* verde e azul estão muito associados, o que indica que houve uma tendência a serem citados pelas mesmas pessoas. Já entre eles e os *clusters* da parte inferior do grafo há uma grande distância, mostrando que aqueles grupos não estão próximos dos superiores. Esses *clusters* existem em meio a buracos estruturais da rede, mostrando que há limitação na circulação das informações.

Ao examinarmos os nós mais citados, observamos que os *clusters* superiores estão agregados em torno da conta oficial do deputado no Twitter, além de contas de apoiadores. Esses *clusters* estão colocados por *indegree*, ou seja, citações e *retweets*. Isso indica que, possivelmente, o grupo está armado em torno dessas contas, o que indicaria que é composto, principalmente, de apoiadores. Já os *clusters* inferiores compreendem nós que estão agregados em torno de contas de não apoiadores do deputado, ou seja, os *retweets* e menções referem-se a críticas. Neste caso, os nós mais citados são mais influentes na rede, porque suas mensagens são compartilhadas com mais frequência do que outras. Entretanto, é importante notar que são influentes nos seus respectivos *clusters*, mas não fora dele. Poderíamos, assim, indicar ainda que a influência dos nós mais citados é limitada a uma audiência específica.

Deste modo, conforme imaginávamos, há uma grande separação entre quem apoia e quem condena a candidatura do deputado, isto é, um grupo não tem contato com o outro. Além disso, pudemos também observar nas métricas que os *clusters* dos apoiadores são mais densos e possuíam maior *outdegree* médio, indicando uma ação mais ativista em torno de seu nome (atores que citavam mais contas de apoiadores).

A análise de redes não é uma abordagem sem debate. Há muitas críticas e controvérsias, que abordaremos brevemente neste capítulo, de modo a problematizar um pouco mais a proposta.

Uma das críticas mais comuns feitas à ARS é que seu forte caráter empírico não deixaria espaço para uma reflexão teórica mais aprofundada. Assim, as análises resultantes do uso da abordagem seriam principalmente descritivas. Entretanto, autores como Borgatti e colaboradores (2014) argumentam que muitos trabalhos teóricos foram realizados em cima de bases de ARS, como por exemplo, boa parte do corpo teórico de capital social, e de modo específico, as discussões sobre buracos estruturais. (BURT, 1992) De fato, muitos trabalhos que utilizam ARS são descritivos e, em muitos casos, se limitam a descrever sob a forma de grafos as conexões entre os autores. Mas mais do que um problema da abordagem, este é um problema específico do foco dos trabalhos. Deste modo, é sempre importante preocupar-se com a discussão em torno do trabalho de ARS, observar suas regularidades e irregularidades e se propôr teorizações e implicações às questões levantadas.

Juntamente à crítica sobre a falta de aprofundamento teórico, há também a percepção de que a análise de redes é apenas um método. (BORGATTI et al, 2014) Entretanto, a própria metáfora de rede é utilizada como base para uma série de abordagens atualmente, como demonstram Watts (1999) e mesmo o trabalho de Barabási (2003). Embora, em seu âmago, a análise de redes tenha um forte viés empírico, ela está estabelecida sobre bases teóricas importantes, conforme procuramos demonstrar nos capítulos deste livro. Desse modo, o uso dos elementos metodológicos da análise de redes não pode ser completamente desconectado de suas premissas teóricas e é exatamente por isso que falamos em “abordagem” e não simplesmente em “método”.

A análise de redes sociais também é frequentemente criticada por prover apenas quadros estáticos das relações entre os atores em um dado momento. (BORGATTI et al, 2014) Não há, assim, uma preocupação para com a dinâmica da estrutura social. (RECUERO, 2009) Embora a análise de redes realmente construa para o pesquisador retratos estáticos das redes, a construção de vários desses retratos é capaz de promover, junto às redes, um viés capaz de

compreender a dinâmica dessas estruturas. Algumas ferramentas de análise, inclusive, são capazes de adicionar camadas de grafos dinâmicos aos dados coletados temporalmente.

Outro tensionamento é o uso dos métodos de análise de redes diante de outras abordagens teóricas, como por exemplo, a Teoria Ator-Rede (TAR). Embora no Brasil tenhamos já tradições de trabalhos com fenômenos relacionais voltadas tanto para a TAR quanto para a ARS, é preciso ainda delimitar que não se trata da mesma coisa. Bastos, Recuero e Zago (2014) trouxeram essas questões de modo bastante pontual, diferenciando os elementos teóricos e as premissas das duas abordagens. Como a TAR é uma abordagem em cuja origem não temos definições e mesmas propostas metodológicas, pode parecer sedutor simplesmente aplicar as métricas de ARS aos seus elementos. No entanto, conforme argumentamos no artigo, as imprecisões da TAR não dão espaço para a aplicação de medidas de ARS sem uma reflexão anterior, pois o conceito de rede, na primeira, é “invariavelmente interminável e potencialmente indefinível”. (BASTOS; RECUERO; ZAGO, 2014, p. 584)

Essas críticas apontam fortemente para a principal vantagem de usar abordagens de redes para o estudo de fenômenos sociais atualmente: seu viés metodológico permite às ciências sociais um maior contato empírico, e seu viés teórico pode ser amplamente construído e debatido nessas áreas para a construção das análises das estruturas sociais.

Conclusão: e então, o que vimos neste livro?

Neste livro, buscamos apresentar de forma simples e direta alguns dos principais conceitos de análise de redes sociais, de forma a munir rapidamente estudantes e interessados de referências e informações que possibilitem o uso da perspectiva para seus trabalhos. Nosso foco principal foi discutir as métricas e os elementos utilizados na ARS e o modo como estes podem ser apropriados pelos interessados.

Iniciamos pela apresentação da perspectiva e um breve histórico da ARS. A seguir, discutimos os principais conceitos que interessam aos sociólogos e comunicólogos que estão se apropriando da perspectiva para seus estudos de redes. Neste capítulo, discutimos conceitos mais qualitativos, como capital social e buracos estruturais. A seguir, apresentamos as métricas principais da análise de redes e suas aplicações, através de exemplos que demonstrassem como colocar essas métricas em prática. Finalmente, discutimos ainda os estudos para os quais serve a abordagem, o modo de coletar dados e de estruturar uma análise de redes e *softwares* que podem ser utilizados.

É importante compreender que essa perspectiva tem um viés fundamental na questão estrutural. De modo diferente de outras abordagens (como a Teoria Ator Rede, por exemplo), a ARS parte de premissas metodológicas bem estruturadas e constantemente discutidas e amplificadas pela comunidade. Além disso, seu centro está principalmente em métricas e abordagens quantitativas, mais do que análises qualitativas. A ARS, como perspectiva, pode ser agregada e misturada com outras abordagens metodológicas, desde que seja realizada uma discussão a respeito dessa integração ou não, ressaltando suas vantagens e limitações. Há várias limitações no uso de ARS, a principal delas relacionada ao fato de que sistemas complexos são reduzidos à sua estrutura e analisados a partir dela. Apesar disso, há modos de buscar analisar dinamicamente essas estruturas (por exemplo, através de sucessões temporais de grafos). O que é relevante aqui é perceber que a ARS existe dentro do arcabouço

de métodos possíveis para que pesquisadores das áreas que tradicionalmente não trabalham de modo tão empírico e com tantos dados (como é o caso da Comunicação) possam se munir de modos de análise e compreender elementos mais amplos em seus dados.

Nosso objetivo, portanto, é ampliar a compreensão desses modos e permitir que a ARS seja mais popularizada entre os pesquisadores, de modo especial, entre aqueles que trabalham com dados provenientes do ciberespaço. Esperamos que este livro sirva, assim, a seu propósito.

Referências

BACKSTROM, L. et al. Four Degrees of Separation, *Cornell University Library*, Ithaca, NY, 5 jan. 2012. Disponível em: <<https://arxiv.org/abs/1111.4570>>. Acesso em: 17 out. 2004.

BALANCIERI, R. et al. A análise de redes de colaboração científica sob as novas tecnologias de informação e comunicação: um estudo da Plataforma lattes, *Ciência da Informação*, Brasília, DF, v. 34, n. 1, jan./abr. 2005.

BARABÁSI, A. *Linked: How everything is connected to to Everything Else and What It Means for Business, Science, and Everyday Life*. New York: Basic Books, 2003.

BASTOS, M. T.; ZAGO, G. S.; RECUERO, R. A Endogamia da Comunicação: redes de colaboração na CSAI. *Revista da Famecos*, Rio Grande do Sul, v. 23, n. 2, 2016.

BASTOS, M. T.; ZAGO, G. S.; RECUERO, R. Encontros e Desencontros entre TAR e ARS: o laço fraco entre teoria e método. *Contemporanea*, Salvador, v. 12, n. 3, p. 1-15, 2014.

BERTOLINI, S.; BRAVO, G. *Social Capital, a Multidimensional Concept*. [S.l.:s.n.], [2004]. Disponível em: <<http://www.ex.ac.uk/shipss/politics/research/socialcapital/other/bertolini.pdf>>. Acesso em: 17 out. 2004.

BORGATTI, S. P.; EVERETT, M. G. Network analysis of 2-mode data. *Social networks*, [S.l.], v. 19, p. 243-269, 1997.

BORGATTI, S. P.; BRASS, D. J.; HALGIN, D. S. Social network research: Confusions, criticisms, and controversies. In: BRASS, D. J.; LABIANCA, G. et al. (Ed.). *Research in the Sociology of Organizations*. Bradford, UK: Emerald publishing, 2014. v. 40

BOURDIEU, P. *Outline of a Theory of Practice*. Cambridge: Cambridge University Press, 1972.

BOURDIEU, P. The forms of Capital. In: RICHARDSON, J. G. *Handbook of theory and research for the sociology of education*. Translated by Richard

Nice. New York: Greenwood, 1986. p. 41-58. Disponível em: <<http://www.pontomidia.com.br/raquel/resources/03.html>>. Acesso em: 23 fev. 2004.

BOYD, D.; ELLISON, N. Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, [S.l.], v. 13, n. 1, p. 210-230, 2007.

BOYD, D. Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications. In: PAPACHARISSI, Z. (Ed.). *A Networked Self: Identity, Community, and Culture on Social Network Sites*. New York: Routledge, 2010. p. 39-58

BRUNS, A. et al. *#qldfloods and @QPSMedia: Crisis Communication on Twitter in the 2011 South East Queensland Floods*. Brisbane, Qld: ARC Centre of Excellence for Creative Industries and Innovation, 2012.

BRUNS, A.; BURGESS, J. E. Researching news discussion on Twitter: New methodologies. *Journalism Studies*, Florida, v. 13, 2012.

BURT, R. The Social Structure of Competition. In: BURT, R. *Structural Holes: the social structure of competition*. Cambridge: Harverd University Press, 1992.

COLEMAN, J. S. Social Capital and the Creation of Human Capital. *American Journal of Sociology*, Chicago, v. 94, p. S95-S120, 1988.

DIGIAMPIETRI, L. A. Análise da Rede Social Acadêmica Brasileira. 2015. 160 f. Tese de conclusão de curso (Livre-Docência, Informação e tecnologia) – Escola de Artes. Ciências e Humanidades, Universidade de São Paulo, São Paulo, 2015.

DEGENNE, A.; FORSÉ, M. *Introducing Social Networks*. London: SAGE, 1999.

EAGLE, N.; PENTLAND, A.; LAZER, D. Inferring friendship network structure using mobile phone data. *Proceedings of the National Academy of Sciences of the United States of America*, Washington, 2009.

FREEMAN, L. Centrality in social networks: Conceptual clarification. *Social Networks*, Amsterdam, n. 1, p. 215-239, 1978.

FREEMAN, L. *The development of social network analysis: a study in the sociology of science*. Vancouver: Empirical Press, 2004.

FRUCHTERMAN, T.; REINGOLD, E. Graph Drawing by Force-directed Placement. *Software: Practice and Experience*, v. 21, n. 11, p. 1129-1164, 1991.

GRANOVETTER, M. S. The Strength of Weak Ties. *American Journal of Sociology*, Chicago, v. 78, n. 6, p. 1360 - 1380, 1973.

HU, Y. Efficient, High-Quality Force-Directed Graph Drawing. *The Mathematica Journal*, San Francisco, v. 10, n. 1, 2008. Disponível em: <http://www.mathematica-journal.com/issue/v10i1/graph_draw.html>. Acesso em: 08 jun. 2015.

JACOMY, M. et al. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLOS One*, San Francisco, v. 9, n. 6, 2014. Disponível em: <<http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0098679>>. Acesso em: 8 jan. 2015.

KRECKEL, R. *Soziale Ungleichheiten*. Goettingen, DE: Otto Schartz & Co, 1983. (Soziale Welt, Sonderheft 2).

LE MOS, A. *Cibercultura: Tecnologia e vida social na cultura contemporânea*. Porto Alegre: Sulina, 2003.

MALINI, F. Um método perspectivista de análise de redes sociais: Cartografando topologias e temporalidades em rede. In: ENCONTRO DA COMPÓS, 25., 2016, Goiânia. *Anais...* Campós: Goiânia, 2016. Disponível em: <http://www.compos.org.br/biblioteca/compos_malini_2016_3269.pdf>. Acesso em: 23 de maio de 2017.

MARTIN, S. et al. *OpenOrd: An Open-Source Toolbox for Large Graph Layout*, SPIE Conference on Visualization and Data Analysis. [S.l.], 2011. Disponível em: <<http://proceedings.spiedigitallibrary.org/proceeding.aspx?ArticleId=731088>>. Acesso em: 8 jan. 2015.

MILGRAM, S. The Small World Experiment. *Psychology Today*, New York, v. 1, n. 1, p. 61-67, maio 1967. Disponível em: <<http://snap.stanford.edu/class/cs224w-readings/milgram67smallworld.pdf>>. Acesso em: 8 jan. 2015.

NEWMAN, M.; WATTS, D.; STROGATZ, S. Random Graph Models of Social Networks. *Proceedings of the National Academy of Sciences of the United States of America*, Washington, v. 99, p. 2566-2572, 2002.

OLIVEIRA, E. F.; SANTAREM, L. G. S.; SANTAREM SEGUNDO, J. E. Análise das redes de colaboração científica através do estudo das co-autorias nos cursos de pós graduação do Brasil no tema tratamento temático da informação. In: CONGRESSO ISKO-ESPANA, 9., 2009, Valencia. *Anais...* Valencia: ISKO, 2009.

PARISER, E. *The Filter Bubble: what the Internet is hiding from you*. New York: The Penguin Press, 2011.

PUTNAM, R. D. *Bowling Alone: The collapse and Revival of American Community*. New York: Simon e Schuster, 2000.

RECUERO, R. *Redes Sociais na Internet*. Porto Alegre: Sulina, 2009.

RECUERO, R. *A Conversação em Rede: comunicação mediada pelo computador*. Porto Alegre: Sulina, 2012.

RECUERO, R. Contribuições da Análise de Redes Sociais para o estudo das redes sociais na Internet: o caso da hashtag #Tamojuntodilma e #CalaabocaDilma. *Revista Fronteiras*, São Leopoldo, v. 16, p. 60-77, 2014.

RECUERO, R.; BASTOS, M. T.; ZAGO, G. *Análise de Redes para Mídia Social*. Porto Alegre: Sulina, 2015.

WATTS, D. J. *Six Degrees: The Science of a Connected Age*. New York: Norton, 2003.

WATTS, D.; STROGATZ, S. Collective dynamics of 'small-world' networks. *Nature*, [S.l.], v. 393, p. 440-442, 1998.

SCOTT, J.; CARRINGTON, P. J. (Org.). *The SAGE handbook of social network analysis*. London: SAGE, 2011.

SCOTT, J. *Social Network Analysis: A handbook*. 2. ed. New York: SAGE, 2001.

WASSERMAN, S.; FAUST, K. *Social Network Analysis: methods and aplications*. Cambridge: Cambridge University Press, 1994.