

# Contribuições da Análise de Redes Sociais para o estudo das redes sociais na Internet: o caso da *hashtag* #Tamojuntodilma e #CalaabocaDilma

## Contributions of Social Network Analysis for the study of online social networks: a case study of #Tamojuntodilma and #CalaabocaDilma hashtags

Raquel Recuero<sup>1</sup>

### RESUMO

Este trabalho busca discutir as contribuições da perspectiva de Análise de Redes Sociais (ARS) para o estudo das redes sociais na Internet. Na primeira parte do artigo, são apresentadas e problematizadas algumas das bases da ARS e suas métricas, bem como outras questões envolvidas. Na segunda parte, fazemos uma análise de duas *hashtags* relacionadas aos protestos que aconteceram no Brasil em Junho de 2013, que apareceram no Twitter durante o pronunciamento à nação da presidenta Dilma Rousseff. A primeira representou um grupo de apoio (#tamojuntodilma), e a segunda, um grupo de críticos (#calaabocadilma). Utilizando a ARS, demonstramos através de medidas os diferentes aspectos das redes sociais constituídas em torno de cada rede, notadamente a articulação de um grupo em torno da primeira *hashtag* e a organicidade da segunda, que emerge como uma resposta.

**Palavras-chave:** análise de redes sociais, redes sociais na internet, conversação em rede.

### ABSTRACT

The following work brings to the discussion the contributions of Social Network Analysis (SNA) to the study of online social networks. In the first part of the paper, we discuss the basics of SNA, its approach and metrics. In the second part, we focus on the analysis of two hashtags connected to the protests that happened in Brazil in June 2013 that appeared in Twitter during the President Dilma Rousseff's pronouncement to the Nation. The first hashtag (#tamojuntodilma - we are together Dilma) was a supportive one. The second (#calaabocadilma - shut up Dilma) was a critical one. Using SNA we show how the first one was organized by a group and the second one emerges as a response to the first.

**Keywords:** social network analysis, online social networks, networked conversation.

---

<sup>1</sup> Universidade Católica de Pelotas. Rua Félix da Cunha, 412, Centro, 96010-000, Pelotas, RS, Brasil. E-mail: raquel@raquelrecuero.com

## Introdução

As redes sociais na Internet representam um novo e complexo universo de fenômenos comunicativos, sociais e discursivos. Não porque a existência de “redes sociais” ou sua análise sejam algo novo, mas porque sua reinscrição no ciberespaço apresenta novos elementos e novas dinâmicas para seu estudo (Recuero, 2009; Benevenuto, 2010; Malini e Antoun, 2013). E essas dinâmicas tornam-se mais relevantes também com o advento dos *sites* de rede social<sup>2</sup> (boyd e Ellison, 2007), que provêm um novo contexto para as redes, permitindo o registro de parte dessas dinâmicas sociais e seu acesso pelos pesquisadores. É tal registro que permite, pela primeira vez, que interações e conversações sejam mapeadas e estudadas em larga escala. Não sem surpresas, o foco renovado pela disponibilidade de dados empíricos trouxe novo fôlego para os estudos de redes sociais e suas apropriações no ciberespaço (vide, por exemplo, Amaral, 2010; Zago 2010; Zago e Rebs, 2011; Pereira e Araújo, 2013).

Entretanto, o estudo das redes sociais ainda carece de ferramentas metodológicas e focos específicos que permitam lidar em larga e pequena escala com os dados relacionais que são coletados. É a partir desse ponto que nasceu este artigo. Pretendemos, assim, resgatar o paradigma (Freeman, 2004) de estudos da Análise de Redes Sociais (ARS), buscando situá-la conceitual e empiricamente para o estudo das redes sociais na Internet e suas conversações. Para tanto, discutiremos inicialmente os elementos principais da reinscrição das redes sociais no ciberespaço, depois levando-os para a discussão de como é constituída e quais são os elementos de análise da ARS. Para tornar as contribuições dessa perspectiva ainda mais evidentes, apresentamos um estudo de caso pontual, focado em duas *hashtags*<sup>3</sup> que tomaram os Trending Topics<sup>4</sup> do

Twitter (<http://www.twitter.com>) no dia 21 de junho de 2013, logo após o pronunciamento da presidenta Dilma Rousseff à nação a respeito dos protestos que aconteceram durante os dias anteriores em várias cidades do Brasil. Esses protestos, que foram marcados pela maior manifestação popular do País dos últimos anos, iniciaram-se pontualmente como uma manifestação contra o aumento das tarifas de transporte coletivo organizadas principalmente pelo Movimento Passe Livre. Com a violenta repressão policial que se seguiu a essas manifestações, notadamente em São Paulo no dia 13 de junho, os protestos espalharam-se pelo País, ocorrendo em mais de 400 cidades diferentes e levando mais de 2 milhões de pessoas às ruas<sup>5</sup>. No dia 21 de junho, após uma semana turbulenta, a presidenta finalmente faz um pronunciamento à nação, respondendo aos protestos. Durante o pronunciamento, surgem no Twitter as *hashtags* que escolhemos para este estudo. Elas foram #TamojuntoDilma e #CalaaBocaDilma, refletindo o apoio e a crítica ao pronunciamento. Neste estudo, procuraremos mostrar como a ARS permite que se identifiquem padrões nos dados que podem elucidar os contextos e as dinâmicas dos grupos que se manifestaram.

## O contexto das redes sociais na Internet

O conceito de rede social está associado pelos estudos de sociólogos, psicólogos e outros estudiosos ao trabalho de Simmel (Wasserman e Faust, 1994) e Moreno<sup>6</sup> (Degenne e Forsé, 1999), que, no início do século XX, desenvolvem trabalhos que focam as estruturas sociais e as redes de filiação dos atores sociais. Entretanto, a análise de redes sociais tem antecedentes ainda mais longínquos. Scott (2004) traça parte desses precursores aos trabalhos

<sup>2</sup> As autoras definem *site* de rede social como aquela ferramenta *online* que provê os usuários com três elementos: a construção de um perfil público ou semipúblico, a publicização das suas redes sociais e a navegação por dentro dessas redes por outras pessoas.

<sup>3</sup> Uma *hashtag* constitui-se em uma etiqueta de “contexto” no Twitter, que aponta de forma específica um termo que não apenas constrói contexto, mas igualmente permite que o tweet seja buscado e recuperado também pela etiqueta. Em geral, é representada pelo sinal “#”.

<sup>4</sup> Trending Topics constituem-se em uma lista de 10 termos e palavras que estão sendo “mais comentados” no Twitter durante um determinado período. No caso, para este trabalho, foram coletadas as *hashtags* porque estavam na lista dos Trending Topics do Brasil na data (21/06/2013).

<sup>5</sup> Os dados aqui apresentados foram coletados pela pesquisadora em conjunto com outros colegas durante estudo sobre os protestos no Brasil e serão futuramente publicados de forma detalhada. Também há outras informações aqui: [http://pt.wikipedia.org/wiki/Protestos\\_no\\_Brasil\\_em\\_2013](http://pt.wikipedia.org/wiki/Protestos_no_Brasil_em_2013)

<sup>6</sup> Moreno, por exemplo, é um precursor na aplicação dos chamados sociogramas (mapas da rede de atores) para as redes sociais, na tentativa de analisar de forma empírica as estruturas de conexões entre as pessoas.

da antropologia estrutural-funcionalista (principalmente através de Radcliffe-Brown) e à própria Gestalt, como um dos fundamentos da análise das dinâmicas dos grupos sociais. O conceito de rede social, desse modo, não surge com os estudos do ciberespaço. Ao contrário, está atrelado a uma perspectiva de estudo que é demarcada fortemente pelo conjunto de trabalhos que vai fundamentar o próprio paradigma da Análise de Redes Sociais. Trata-se, assim, de uma forma de observar a estrutura social (Wasserman e Faust, 1994), construída a partir dos dados relacionais dos atores sociais (indivíduos ou instituições e grupos) e suas interrelações. Scott (2004) explica que “em ciência social, a abordagem estrutural que é baseada no estudo da interação entre atores sociais é chamada análise de redes sociais. As relações que os analistas de redes sociais examinam são, geralmente, aquelas que conectam indivíduos humanos”<sup>7</sup> (p. 2). Entretanto, ressalta o autor, os estudiosos também podem focar redes onde os atores são grupos ou organizações. O que interessa ao estudioso de ARS são os “padrões de laços sociais” nos quais os atores estão envolvidos.

As redes sociais em sua representação no ciberespaço são um pouco diferentes das redes sociais no espaço *offline*, primeiramente, porque as conversações e as trocas sociais deixam rastro no *online* (boyd, 2010; Recuero, 2012). Esses “rastros” são publicados, arquivados, e, portanto, são recuperáveis e buscáveis. Segundo, são diferentes porque a própria representação do grupo social no ciberespaço altera o grupo em si. Essas redes são representadas principalmente através dos *sites* de rede social e daquelas outras ferramentas que permitiram sua apropriação desse modo. As redes representadas nessas ferramentas, assim, sofrem menos com a temporalidade das relações *offline*. Não têm, por exemplo, seus laços desgastados pela falta de contato. Constituem-se em redes mais estáveis e, com isso, mais complexas, maiores e compreendendo uma pluralidade de relações mais ampla que aquela das redes *offline*. Ellison *et al.* (2007) por exemplo, fizeram um estudo amplo indicando como o Facebook (<http://www.facebook.com>), enquanto suporte das relações sociais, modificou os processos sociais das pessoas, permitindo que atores que não tinham mais contato devido à distância pudessem investir ainda nesses laços sociais. No caso, o estudo mostrou que a ferramenta permitia aos atores a manutenção de uma rede social

com a qual ele não teria mais contato. Do mesmo modo, West *et al.* (2009) também têm um estudo parecido, discutindo como as “amizades” são vistas no universo do Facebook e nas esferas das relações públicas e privadas. Os *sites* de rede social publicizaram as conexões, mas também proporcionaram que os laços sociais (e as interações e relações) representados nos mapas se tornassem mais permanentes, menos fluidos, mais estáveis.

É preciso que se diga, ainda, que o *site* não é a rede. Trata-se de um suporte, que é apropriado de diferentes formas pelos grupos sociais e cujos efeitos são construídos pelo complexo universo de negociação de normas e formas de interação. O trabalho de Malini e Antoun (2013) demonstra isso quando discute os movimentos que emergem das apropriações desses suportes nas novas possibilidades de compartilhamento e cooperação que também são construídas pelas mudanças do digital. Esses exemplos trazem contribuições importantes a respeito das mudanças geradas pelos suportes nas redes sociais na internet.

Entretanto, talvez o conceito que melhor sintetiza essa mudança é aquele dos “públicos em rede” (boyd, 2010). A autora define o conceito como os públicos reestruturados na rede, que compreende, ao mesmo tempo, o espaço construído pelas tecnologias e o coletivo que emerge da apropriação desse espaço. Por isso, os públicos em rede têm características relacionadas a esse espaço e suas apropriações. Dentre as características do espaço, estão: (i) a permanência das interações, ou seja, o fato de que as interações tendem a ficar inscritas na rede e ali permanecerem; (ii) a “buscabilidade” dessas interações, que são recuperáveis; (iii) a replicabilidade dessas interações que podem ser reproduzidas facilmente e; (iv) a escalabilidade, ou seja, o potencial de alcance e multiplicação desses registros. Essas características nos mostram que há uma mudança no suporte da interação, que vai permitir que as conversações, antes mapeadas unicamente pela observação, tornem-se mais facilmente registradas. É aí que vemos a importante demarcação: Pela primeira vez, graças ao advento e à apropriação dos *sites* de rede social, os atores passam a registrar seus passos, suas conversas, suas interações e redes. E, com isso, o mapeamento dessas redes ganha novo potencial, com ares de “big data”, no sentido de que, pela primeira vez, é possível mapear gostos, atos, ideias e conexões de milhares de pessoas, procurar e estabelecer padrões entre essas múltiplas redes,

<sup>7</sup> Tradução da autora para: “In social science, the structural approach that is based on the study of interaction among social actors is called social network analysis. The relationships that social network analysts study are usually those that link individual human beings” (p. 2).

principalmente através das interações que são mediadas por essas ferramentas.

O estudo das redes é, portanto, o estudo dos padrões sociais. Como na Internet esses padrões tornam-se mais evidentes, há a possibilidade de estudá-los de uma forma mais abrangente e em maior escala. É nisso que as medidas e perspectivas construídas pelo paradigma da ARS podem ajudar o pesquisador. E é também a partir daí que discutiremos a abordagem e suas contribuições para os estudos das redes.

## A Análise de Redes Sociais em redes sociais na Internet

A Análise de Redes Sociais não é uma abordagem nova. Suas fundações mais antigas estão na Teoria dos Grafos e em parte da Sociometria, através do trabalho de Moreno, ainda na primeira metade do século XX (Scott, 2004). Entretanto, é só na segunda metade do século que a abordagem emerge a partir de um paradigma efetivo de pesquisa, através do trabalho de pesquisadores de diferentes disciplinas em algumas linhas centrais. Freeman (2004) aponta essas linhas que são desenvolvidas paralelamente: (i) a estrutura do grupo social; (ii) os dados das relações dos atores com atores; (iii) desenvolvimento de ferramentas para a visualização dos padrões de relações; ou ainda (iv) as propriedades matemáticas dos padrões sociais. Entretanto, é só com a junção desses focos que a ARS surge e desponta como uma perspectiva interdisciplinar, que congrega estudiosos de várias áreas interessados no estudo dos padrões sociais. Por isso, fala-se em “medidas”, cujos conceitos são, em sua maioria, explicitados matematicamente. O uso de elementos da estatística também é comum, bem como o foco quantitativo, o que não significa que os dados não possam ser obtidos através de entrevistas (veja Hogan *et al.*, 2007) ou que análises mais qualitativas estejam completamente excluídas do trabalho.

Lemieux e Ouimet (2004) explicam que a abordagem é constituída de três processos, o descritivo, a abordagem explicativa no primeiro nível (aquela que “limitam-se a tratar de regularidades nas redes e outras

formas de organização”, p. 13) e a de segundo nível (onde há a postulação de “princípios subjacentes às redes sociais e a outras formas de organização social”, p. 13). A análise de redes sociais, assim, poderia trazer análises mais qualitativas nesse terceiro processo, embora necessariamente ancoradas nos dados empíricos. Desse modo, a ARS é uma abordagem que traz um conjunto de métodos de coleta e análise, bem como uma perspectiva que é extremamente interessante para o estudo das redes sociais *online*, pois foca, exatamente, nas estruturas que podem ser percebidas através dos dados empíricos que são coletados dessas redes.

### Coleta de dados relacionais

O primeiro desafio do pesquisador é identificar que tipo de rede será necessária para o seu estudo. Como explicam Lemieux e Ouimet (2004), dados relacionais, em geral, não funcionam bem quando coletados através de amostragem. Além disso, redes sociais na Internet são escaláveis (boyd, 2010) e, por isso, a coleta de dados contempla pode tornar-se difícil. É preciso, assim, ou coletar toda a rede ou optar por um estudo de caso de um determinado grupo. Assim, coletam-se os dados:

(i) Através da rede ego centrada e seus graus de separação - Aqui coletam-se os dados a partir de um ator-ego e suas conexões (grau). O limite desta coleta, assim, é dado pela distância deste “ego”. A distância entre dois nós na rede é denominada “grau” ou “grau de separação”<sup>8</sup>. Assim, ao determinar essa distância, determina-se também a rede onde serão coletados os dados. Em termos de rede na Internet, podemos, por exemplo, coletar os dados de uma rede no Twitter, onde coletaremos 2 (dois) graus de ego. Isso significa que serão coletados todos os amigos do ator-ego e todos os amigos dos amigos desse ator. Ou poderíamos coletar apenas um grau (ego e amigos de ego).

(ii) Através de uma rede inteira - Coletam-se todos os dados de uma determinada rede que está limitada de alguma forma no ciberespaço. Por exemplo, poderíamos coletar todos os dados de um determinado grupo no Facebook (rede limitada por quem está no grupo e não pelas conexões). Essa seria uma coleta de rede inteira, pois o limite da coleta de dados se dá pela escolha do limite externo à rede (grupo do Facebook).

<sup>8</sup> O grau de separação, assim, é a distância entre quaisquer atores da rede social. Essa distância é calculada em termos de arestas (conexões) e não de atores.

Inicialmente, para realizar a coleta, também é preciso eleger o que serão considerados nós e conexões. Coletar dados relacionais na Internet é também um grande desafio pela pluralidade de representações das redes. Nesse sentido, há redes que são denominadas associativas ou de filiação e redes que são denominadas emergentes (Recuero, 2009). Enquanto as primeiras determinam relações mais ou menos estáveis, através de conexões mantidas pelas ferramentas, as segundas focam relações mais dinâmicas de fluxos conversacionais. Assim, é preciso decidir se as conexões mapeadas serão, por exemplo, relações de seguidos/seguidores ou de menções e respostas (Twitter), ou ainda de “amigos” ou de conversações (Facebook). Cada uma dessas redes vai informar elementos diferentes da estrutura social da rede onde “A” está inserido (vide Huberman *et al.*, 2009, por exemplo).

Um outro elemento é identificar a direção das conexões. É preciso compreendê-las para saber se os dados serão coletados por um grafo direcionado ou não direcionado. Os grafos direcionados têm suas conexões representadas por flechas que indicam a direção da conexão. Os não direcionados, por outro lado, não utilizam essa representação. E há implicações diferentes em cada uma delas. Por exemplo, se estamos diante de uma rede emergente no Facebook, ou seja, estamos mapeando quem cita quem em uma conversação, há, obviamente uma rede direcionada. Quando A cita B, há uma conexão direta entre A → B. Já quando mapeamos uma rede de amigos no Facebook, é natural que seja uma rede não direcionada, pois para que A seja amigo de B, este precisa aceitar o pedido. A conexão estabelecida, assim, é mútua e dispensa sua representação por flechas, tornando-se não direcionada. Assim, é importante delimitar muito bem o que será considerado uma conexão e como, pois conexões direcionadas e não direcionadas não podem coexistir no mesmo grafo (embora possamos ter conexões direcionadas nos dois sentidos, por exemplo A ↔ B se ambos citaram-se).

O último desafio da coleta de dados refere-se a sua “fasiabilidade”. Justamente por conta de os rastros deixados *online* constituírem-se em uma significativa quantidade de dados, não é uma tarefa simples coletá-los. Por isso, a maioria dos estudiosos prefere usar elementos automatizadores para essas coletas, notadamente *crawlers* (ou robôs), que utilizam as APIs dos *sites* (*application*

*programming interface*). Alguns *crawlers* são simples de utilizar, mas há outros que exigem um certo conhecimento de programação. Além disso, dependendo da API que o *crawler* utiliza, há limitações importantes para a coleta de dados que o pesquisador precisa conhecer e cujas informações normalmente estão disponibilizadas nas próprias descrições. Dentre as ferramentas mais utilizadas hoje, está o YourTwrapperKeeper (<https://github.com/540co/yourTwrapperKeeper>) (yTK), usado para retirar dados do Twitter, por exemplo<sup>9</sup>. Outras ferramentas, como o NodeXL (<http://nodexl.codeplex.com/>) (que funciona como uma aba no Windows Excel, portanto, é preciso tê-lo instalado) e o NetVizz (<https://apps.facebook.com/netvizz/>), vêm com *crawlers* embutidos, que facilitam o trabalho de coleta do pesquisador, embora trabalhem com mais limitações que a coleta realizada pelo yTK<sup>10</sup>.

Os dados coletados são geralmente apresentados sob a forma de uma matriz. Essa matriz ou sociograma (Scott, 2004) é que vai servir de base para que seja gerado o grafo, ou a visualização dos dados. Essa visualização, entretanto, também depende de medidas construídas anteriormente, novamente, de acordo com o problema de pesquisa eleito pelo pesquisador, conforme elencaremos a seguir.

## Primeiro nível de análise

Conforme Lemieux e Ouimet (2004), o processo de análise dos dados relacionais possui dois níveis. O primeiro, descritivo, é aquele onde está presente a descrição dos dados, suas medidas. Neste trabalho, focaremos as métricas mais comumente utilizadas em ARS e observaremos seu sentido. No segundo momento, abarcaremos a análise mais abrangente.

## Posição do nó

Uma das medidas mais importantes em ARS é a medida de posição do nó na rede, denominada centralidade

<sup>9</sup> O próprio sistema, entretanto, alerta que “fere” os termos de uso da ferramenta e que as contas que o utilizam podem ser suspensas.

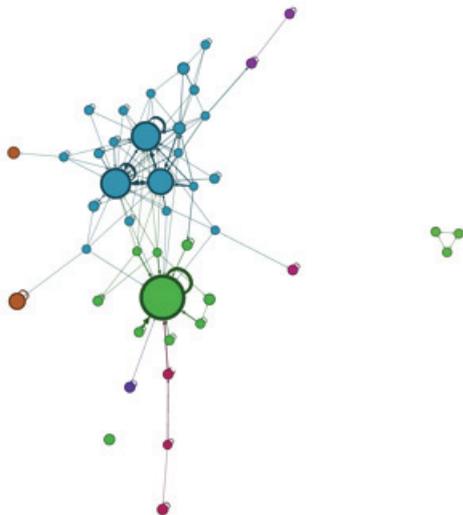
<sup>10</sup> Por exemplo, o NodeXL no Twitter foca a API de *search*, que limita os resultados temporalmente e por pedido. Com isso, embora seja possível obter resultados, não há uma completude neles. O NetVizz, por outro lado, para corroborar com as normas do Facebook, só permite que os dados sejam coletados de forma anônima.

(Scott, 2004; Degenne e Forsé, 1999; Wasserman e Faust, 1994). Ela trabalha com formas de entender o quão importante um determinado nó é para a rede ou o quão centralizada está a rede em torno de determinados nós. Entretanto, a medida de centralidade não é única. Há várias formas de observar a centralidade, conforme discutiremos a seguir.

(a) *Grau do Nó* - Essa é a primeira e mais simples medida. Ela representa o número de conexões que um determinado nó possui. Quanto mais conexões, mais central o nó é para a rede. Um grafo direcionado tem dois graus: o *indegree*, que representa a quantidade de conexões que um determinado nó recebe, e o *outdegree*, que representa a quantidade de conexões que o nó faz (Figuras 1 e 2 mostram isso). O grafo não direcionado, por outro lado, tem apenas um grau, que é o número de conexões. Assim, o grau de centralidade de um nó A no Twitter poderia ser representada pela quantidade de conexões que esse nó possui (*indegree* ou *outdegree*, seguidores e seguidos). No

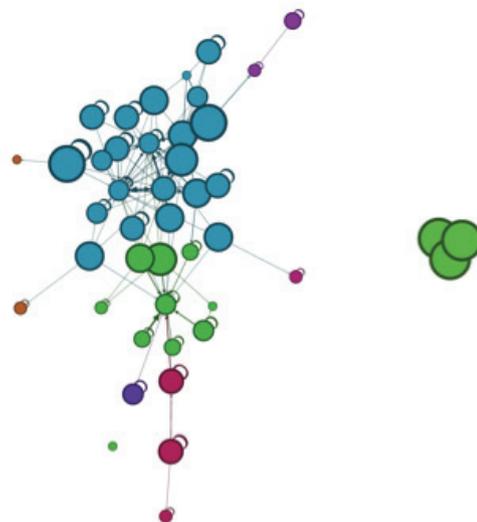
caso, os graus do nó variam entre 0 e o número máximo de conexões possíveis.

(b) *Grau de Intermediação (Betweenness)* - O grau de intermediação é outra medida de centralidade. Entretanto, ele mede o quão “entre” grupos no grafo um determinado nó está, ou o número de vezes em que o nó é “ponte” entre vários grupos de nós. Nós com altos graus de intermediação são aqueles que conectam grupos diferentes (e desconectados entre si) dentro do grafo (Figura 3). Assim, imaginemos dois grupos de usuários do Twitter: fãs do Luan Santana e do Restart<sup>11</sup>. Eles são muito conectados a seus pares (outros fãs do mesmo grupo/cantor). Entretanto, não são muito conectados entre si, porque quem curte um não necessariamente curte o outro. Ainda assim, há alguns usuários que gostam dos dois. Esses poucos são usuários que estão conectados aos dois grupos. São pontes, usuários com alto grau de centralidade *betweenness*.



**Figura 1.** Indegree no grafo. É um grafo direcionado, onde os maiores nós são aqueles que recebem mais conexões (nós maiores). Os nós, portanto, são proporcionais ao seu *indegree*.

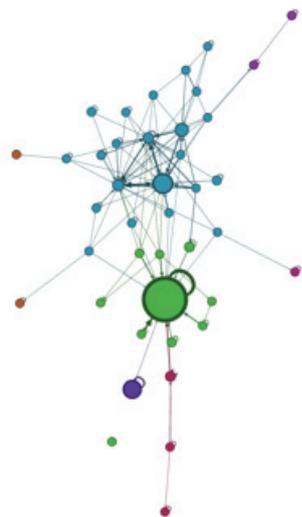
**Figure 1.** Indegree graph. This is a directed graph, where the biggest nodes are the ones with more connections.



**Figura 2.** Outdegree no grafo: Mesmo grafo direcionado. Desta vez, o tamanho do nó é proporcional a seu número de citações ou de conexões recebidas. Ou seja, os nós são proporcionais ao seu *outdegree*.

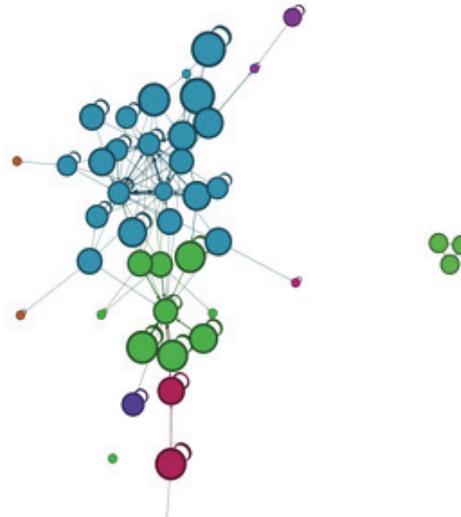
**Figure 2.** Outdegree graph. This is the same directed graph. The node size is proportional to its number of citations or received connections.

<sup>11</sup> Exemplo semelhante pode ser encontrado em Recuero *et al.* (2012). Tanto a banda Restart quanto o cantor Luan Santana têm fãs ativamente engajados e conectados no Twitter.



**Figura 3.** Grafos de conversação no Twitter. O tamanho do nó é proporcional a sua centralidade “betweenness” (ou grau de intermediação). Observe-se que o nó que conecta mais “grupos” de nó, ou seja, que constitui-se em maior ponte, é o maior.

**Figure 3.** Betweenness centrality. The node size is proportional to its betweenness centrality. The node with more bridges is the biggest.



**Figura 4.** Mesmo grafo de conversação. O tamanho do nó, agora é proporcional a sua centralidade “closeness” (ou grau de proximidade). Observe-se que há uma distribuição maior do grau entre os nós porque o grafo é pequeno.

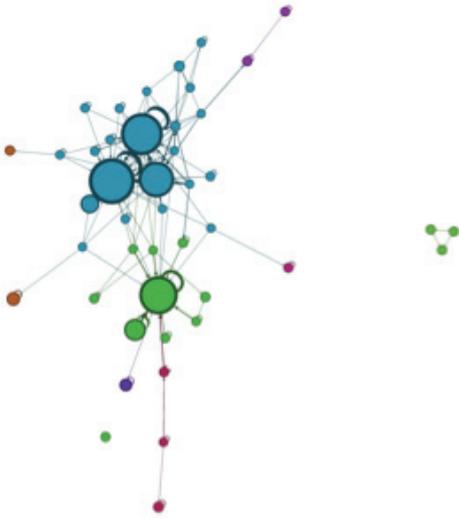
**Figure 4.** Closeness centrality. The node size is proportional to its closeness centrality. We see a more distributed pattern because it is a small graph.

(c) *Grau de Proximidade (Closeness)* - O grau de proximidade é uma medida do quanto um determinado nó está próximo dos demais na rede, ou ainda, qual é a distância desse nó dos demais (Figura 4). A medida dá-se pelo número de pontes, ou seja, de conexões-ponte, de nós que conectam vários grupos que não são interconectados. Esses nós acabam reduzindo as distâncias entre todos os nós do grafo. Os nós com maior grau proximidade, portanto, tendem a ser aqueles com menor distância média de todos os demais nós no grafo. Assim, no mesmo exemplo anterior, os fãs do grupo do Restart terão maior grau de proximidade entre si, mas menor grau de proximidade dos fãs do Luan Santana. Quanto mais “pontes” tivermos no grafo, portanto, maior será o grau de proximidade geral.

(d) *Centralidade Eigenvector* - Essa medida foca na influência do nó na rede. A ideia aqui é quantificar essa influência através da avaliação dos nós. Basicamente, essa medida valoriza mais os nós que recebem conexões de nós que também são valorizados do que os nós que recebem conexões de outros nós com um valor baixo (Figura 5).

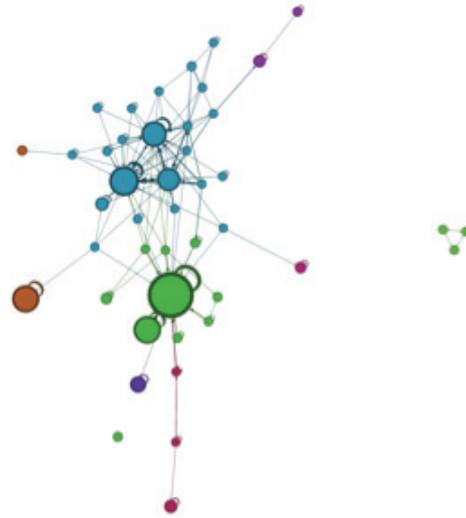
Assim, em um determinado grafo, um nó que receba 3 conexões de outros nós com alto grau vai ter uma centralidade *eigenvector* maior do que um outro nó com outras 3 conexões de nós com graus mais baixos. Em termos simples, poderíamos imaginar novamente o exemplo do Twitter: uma conta que tenha seguidores mais populares, ou seja, outras contas com um alto número de seguidores (maior *indegree*) tende a ter uma centralidade *eigenvector* maior do que outra conta com um mesmo número de seguidores que, na média, tem menos seguidores (menor *indegree*).

(d) *PageRank* - O Pagerank é uma forma também de medir influência dos nós pelas suas conexões (Figura 6). Como no item anterior, há uma medida atribuída aos nós dentro do grafo com base na qualidade de suas conexões. A ideia é diferente da medida anterior, pois aqui observa-se quais as chances de um ator seguindo *links* de forma aleatória terá de chegar na “página” do nó, ranqueando esses nós com conexões “melhores” de forma mais alta (de forma análoga ao algoritmo de Bin e Page, que é utilizado pelo sistema de busca do Google).



**Figura 5.** Centralidade Eigenvector: Tamanho dos nós proporcionais à centralidade *eigenvector*. Os nós mais populares são aqueles que têm conexões de outros nós populares.

**Figure 5.** Eigenvector Centrality: The node size is proportional to its eigenvector centrality. The more popular nodes are the ones with more connections from other popular nodes.



**Figura 6.** Pagerank: Tamanho dos nós proporcionais ao *pagerank*. Os nós com maior *pagerank* são um pouco diferentes da medida de centralidade *eigenvector*.

**Figure 6.** Pagerank: The node size is proportional to its pagerank. They are different than the ones with higher eigenvector centrality.

## Medidas de rede

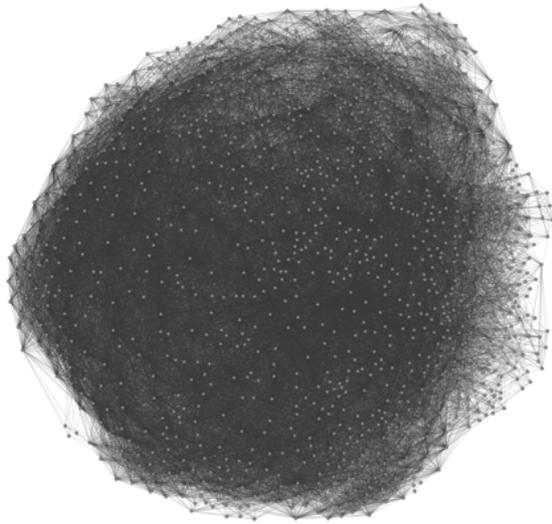
Além das medidas de localização dos nós no grafo, também são importantes as medidas que qualificam as conexões nesses grafos. Assim, temos uma série de conceitos que são utilizados para esse estudo, notadamente referentes não mais ao nó, mas à rede como um todo. Por isso, são também chamadas “medidas de rede” e focam, geralmente, nas conexões presentes no grafo. Dentre essas medidas:

(a) *Densidade* - A densidade do grafo refere-se à quantidade de conexões em relação ao número total de conexões possíveis no grafo. Assim, quanto maior a densidade, mais interconectado o grafo está. Quando uma rede tem o número máximo de conexões possíveis, diz-se que é um clique. Redes mais densas são, assim, aquelas onde há mais conexões entre os nós. Em geral, a densidade também é associada, por exemplo, à presença de “comunidades”, ou “clusters”.

(b) *Coefficiente de Clusterização* - Refere-se também ao quão interconectada é uma rede. Um *cluster* é um conjunto de nós mais densamente conectado que o resto da rede. Assim, o coeficiente de clusterização é uma medida do quanto os nós estão interconectados e próximos na rede (Figuras 7 e 8). É através do coeficiente que podemos também medir o quão densa é uma rede determinada.

(c) *Centralização* - A centralização é uma medida focada no quão centralizado um grafo está em torno de determinados nós. Scott (2004) define a centralização em relação à densidade. Para o autor, enquanto a densidade foca o “nível de coesão geral” de um grafo (aqui entendido como o quão interconectado ele está), a centralização foca em torno de em quais pontos focais essa coesão está organizada (p. 89). Por isso, diz-se que essas medidas são complementares. Assim, quando uma rede tem uma alta centralização, isso significa dizer que há poucos nós “segurando” a rede como um todo, ou seja, centralizando a rede.

(d) *Fechamento (Closure)* - O fechamento de uma rede refere-se a sua interconexão. Quanto mais triádes



**Figura 7.** Alta Densidade: No exemplo, temos um grafo de conversações no Facebook, com alta densidade ( $cc = 0.751$ ).

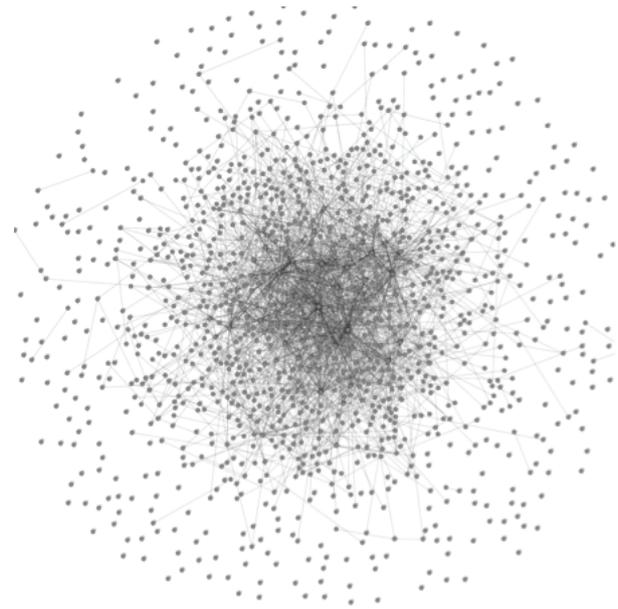
**Figure 7.** High Density: In this example, we see a graph of Facebook conversations with high density ( $cc=0.751$ ).

completas (ou seja, quanto mais fechadas são as conexões entre amigos dos amigos), mais conectada é a rede.

(e) *Coesão* - A coesão é também uma medida de conexão. Refere-se ao grau em que cada nó está conectado aos demais e também ajuda a compreender quais nós, uma vez retirados do grupo, desconectariam a rede. Quanto mais coesa, maior o número de nós necessários para desconectá-la.

(f) *Modularidade* - A modularidade é uma medida de agrupamento de nós. Ela divide os nós de um grafo em grupos, de acordo com suas conexões. Quanto mais densas as interconexões entre um determinado grupo de nós, maiores as chances de eles constituírem um módulo na rede. Ela tende a separar *clusters* de nós dentro da rede. Quanto menos clusterizada, maior a quantidade de módulos. Essa medida, portanto, nos mostra os grupos existentes de nós.

A análise, através dessas métricas, é geralmente realizada através de *softwares* de análise, tais como o Pajek (<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>), Gephi (<https://gephi.org/>) ou NodeXL (já citado anteriormente)<sup>12</sup>.



**Figura 8.** Baixa Densidade: No exemplo, temos um grafo de conversações no Twitter, com baixa densidade ( $cc = 0.028$ ).

**Figure 8.** Low density: In this example, we see a graph of Twitter conversations with low density ( $cc=0.028$ ).

São esses *softwares* que, através de algoritmos específicos, calculam as medidas e desenham a rede, também levando em conta determinadas formas de visualização. Há diversos algoritmos de visualização, através dos quais a rede pode ser montada. Além disso, é importante também salientar que a análise da rede deve dar-se através das medidas e não do seu “desenho”, uma vez que esse pode não indicar elementos que sejam relevantes para o pesquisador.

## Segundo nível de análise

Lemieux e Ouimet (2004) explicam que, no segundo nível de análise, temos a postulação de princípios. Aí focamos não apenas as medidas, mas o que elas querem dizer em qual contexto. A análise de redes, assim, vai muito além da mera constatação de medidas. É preciso discutir o que essas medidas querem dizer no contexto da pesquisa. É nesse nível que situo as métricas mais qualitativas e

<sup>12</sup> Há muitos outros programas que podem ser utilizados. Citamos alguns dos mais conhecidos atualmente.

teóricas, que, conjuntamente com as métricas oferecidas pela teoria dos grafos, auxiliam a compreender a rede da qual se coletam os dados. É nesse nível que está também o estudo da natureza dos laços sociais e do capital social e sua influência na estrutura social, por exemplo. Nesse sentido, o livro de Wellman e Berkowitz (2007), sobre a abordagem estrutural, é muito elucidativo com relação às variáveis qualitativas que podem ser abordadas. Seremos breves nessas contribuições, fazendo uma lista exemplificativa, mas não exaustiva, de como as medidas de rede podem ser abordadas como variáveis qualitativas.

(a) *Laços Sociais* - A abordagem de redes permite que se trate das conexões entre os atores sociais de forma particular. São essas conexões que formam, assim, a estrutura social. Avaliar o que essas conexões significam no *corpus* analisado é essencial. Assim, por exemplo, as medidas de modularidade e clusterização que direcionam os grupos na rede mapeada avaliam como os nós estão conectados. Entretanto, não avaliam a qualidade dessas conexões. Uma medida qualitativa pode auxiliar a compreender qual é natureza desses laços, como, por exemplo, fortes ou fracos (Granovetter, 1973). Além disso, é possível avaliar também as diferenças da representação desses laços *online*, como a questão da facilidade em estabelecer esse laço e os valores que advêm dele (Ellison *et al.*, 2007, por exemplo, fazem um estudo neste sentido no Facebook). Levando-se em conta que os laços são estabelecidos através da comunicação mediada nos *sites* de rede social, trata-se de um evidente foco de estudos para a Comunicação.

(b) *Capital Social* - O capital social é um conceito que foca os valores construídos na estrutura social. Embora seu conceito e sua aplicação sejam variados (Wellman, 2001), é um elemento de análise fundamental nas redes sociais. Burt (1995) é um dos pesquisadores que procurou compreender a influência do capital social na estrutura das redes sociais criando o conceito de “buraco estrutural”. O autor foca as vantagens que advêm a um ator que funciona como “ponte” entre vários grupos não conectados (poderíamos dizer, que tem grau de intermediação - *betweenness* - alto) quando esse ator está em um ambiente externo, ou seja, fora dos grupos (ou entre eles), especialmente em situações de competição ou conflito (Lemieux e Ouimet, 2004). O valor desse grau de intermediação (ou buraco estrutural) no ambiente interno a uma rede, por outro lado, seriam prejudiciais, pois “isolam” o ator. Trata-se, assim, de um estudo sobre o valor das conexões no grupo e sua influência, portanto, seu capital social. Há vários estudos semelhantes focando o conceito nas redes

sociais *online*, notadamente os de Ellison *et al.* (2007), Steinfeld *et al.* (2008) e os trabalhos de Wellman e seu grupo (Hamptom e Wellman, 2003 e Wellman *et al.*, 2001, por exemplo). Esses trabalhos também trazem o foco para os valores gerados pelas estruturas sociais que são formadas pela comunicação mediada, também, portanto, um foco a ser explorado pelos estudos da Comunicação.

(c) *Estrutura Social* - Outro foco bastante frequente dos trabalhos é aquele da identificação de padrões na estrutura social. Essa identificação é capaz, por exemplo, de trazer elementos para a análise dos canais de comunicação nas organizações (vide, por exemplo, a discussão estabelecida por Mizruchi, 2006). Além disso, a percepção de comunidades nessas estruturas e sua avaliação também pode ser um objeto da comunicação (vide o trabalho de Wellman, 2001). Os efeitos comunicativos dessas estruturas também podem ser objeto, como mostra o foco do trabalho de Passerino *et al.* (2007). Assim, os estudos aqui poderiam trazer contribuições a respeito de quais são as estruturas que emergem da comunicação mediada pelo computador no registro dos *sites* de rede social.

## Problemas éticos e limitações

Outra questão importante que precisa ser observada nos estudos de ARS nas redes sociais na Internet são as limitações dessa abordagem e suas questões éticas. Um dos primeiros problemas diz respeito ao anonimato dos dados, por exemplo (vide Kadushin, 2005). Embora, na prática, os *sites* de rede social explicitem que seu conteúdo está sendo (em sua maioria) publicado, ou seja, tornado público e que, naqueles onde opções de privacidade são oferecidas (como o Facebook e o Twitter, por exemplo), o único conteúdo que poderá ser buscado nas APIs é o público, há que se argumentar a dificuldade de compreensão dos públicos em geral das consequências desse ato. Assim, informações que tenham sido publicadas, especialmente, aquelas pessoais ou talvez, publicadas impensadamente pelos sujeitos de pesquisa podem trazer efeitos para estes se utilizadas em pesquisas. Kleinberg (2007), por outro lado, alerta que, devido ao formato das informações nessas redes sociais, muitas vezes, é muito difícil manter a privacidade dos indivíduos, mesmo com os dados anonimizados. Entretanto, é um problema que precisa ser focado pelos pesquisadores se não na coleta, ao menos na apresentação dos dados de pesquisa. Outra questão ética é o uso de dados de redes fechadas, às quais

o pesquisador tem acesso por fazer parte do grupo. Esse é outro problema. Assim, por exemplo, é possível mapear um determinado grupo fechado do qual se faz parte. Mas as implicações éticas desse mapeamento (uma vez que o grupo é fechado) também precisam ser discutidas.

A ARS também é uma abordagem que oferece algumas limitações. A primeira delas diz respeito a seu foco estrutural, nos padrões emergentes das interações. Assim, não se presta para estudos cujo foco não seja estrutural. Se o objetivo de uma pesquisa é trabalhar com um determinado discurso nas redes sociais, por exemplo, mas não se deseja observar os padrões de reprodução desse discurso, a abordagem da ARS não oferece contribuições. Além disso, seu foco em medidas e dados também pode ser enfraquecido quando o objetivo é construir uma análise mais teórica em cima dos fenômenos. A ARS tem um foco extremamente empírico, e o método e o estudo dos dados são extremamente relevantes. Outras limitações dizem respeito aos limites impostos pelas ferramentas, tanto na coleta quanto na análise de dados. A própria API da qual os dados são extraídos, por exemplo, pode oferecer limitações que precisam ser claramente compreendidas pelo pesquisador, sob pena de uma análise parcial ser tomada como uma análise completa. O conhecimento das métricas e das formas de visualização das redes também é essencial para que se realmente compreenda o que os dados querem dizer.

## Estudo de caso: as *hashtags* #Tamojuntodilma e #CalaabocaDilma

Para demonstrar as contribuições da ARS para o estudo das redes sociais na Internet, trouxemos, a título exemplificativo, um estudo de caso específico. Nosso caso específico refere-se a duas *hashtags* que tomaram o Twitter brasileiro na noite do dia 21 de junho de 2013. Nesse dia, a presidenta da República, Dilma Rousseff pronunciou-se à nação a respeito dos protestos que, durante as semanas anteriores e posteriores, assolaram o País. O pronunciamento à nação rapidamente repercutiu nas redes sociais, de modo especial no Twitter. Logo após a fala da presidenta, duas *hashtags* tomaram os Trending Topics, #tamojunto-

dilma, que manifestava diretamente apoio à presidenta e #calaabocadilma, que, ao contrário, pedia sua saída. Durante a noite do dia 21, foram mapeadas as duas *hashtags*, utilizando-se a ferramenta de crawl do NodeXL<sup>13</sup>. As duas *hashtags* foram buscadas no mesmo horário (às 22h) e foram coletados 4.368 *tweets* de 3.182 atores com a tag #tamojuntodilma e 5.597 *tweets* de 4.804 atores com a tag #calaabocadilma. Foram coletados apenas os *tweets*, ou seja, a rede que será construída é baseada apenas nestes e não na relação de seguidores. Assim, cada aresta ou conexão será construída por uma citação ou *retweet* que cite um ou mais outros atores. Se não há citações, menções ou *retweets*, não há conexão entre os nós. Optamos por esta coleta porque, em uma conversação, há claramente a participação de vários atores quando há referência aos participantes (vide boyd *et al.*, 2010). Nossa questão de pesquisa, nesses casos é, portanto, simples: Quais padrões podemos extrair das redes de conversação que utilizaram as *hashtags* e como esses padrões refletem o apoio/a crítica ao pronunciamento da presidenta?

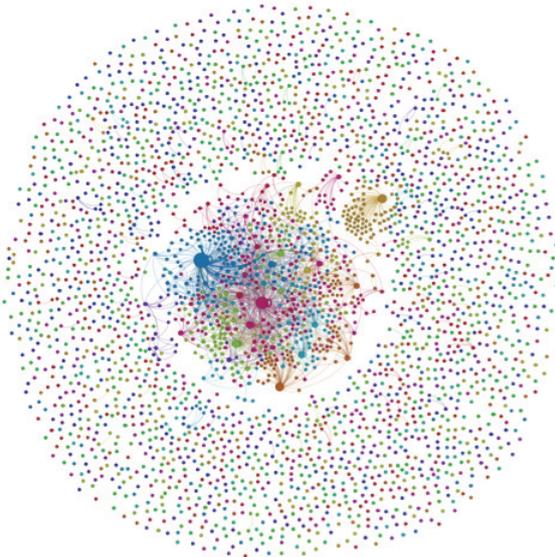
### #TamoJuntoDilma

Nesta coleta, foram obtidos 4.368 *tweets* de 3.182 atores (contas individuais). Observa-se, portanto, que mais de um *tweet* está relacionado ao mesmo ator. Ao construir-se o grafo dessa *hashtag*, temos as primeiras informações. Nota-se, no grafo da *hashtag* (Figura 9), claramente, um *cluster* de citações. Observe-se, no mapa, como os nós que não se mencionam ou citam, ou que usaram a *hashtag* de forma individual aparecem isolados.

Filtrando-se apenas os atores que fizeram, no período, mais de 4 *tweets* com a *hashtag*, ou seja, nos atores mais engajados em utilizá-la, podemos observar mais claramente o *cluster* da rede e seus participantes (Figura 10). No mapa, os nós com maior *indegree* estão demarcados como nós maiores, enquanto aqueles com menor *indegree* são menores. Quando observamos os *tweets* dos nós mais citados, aparece claramente apoio à presidenta.

O coeficiente de clusterização do grafo é de 0,007, enquanto o grau de conexão médio é 1,371. O máximo *indegree* (nó com número máximo de citações) é o @PT-Nacional, com 133 menções. A seguir, vêm partidários e apoiadores. Quando observamos o *outdegree*, a distribuição

<sup>13</sup> Entretanto, as análises e os grafos apresentados a seguir foram construídos com o auxílio da ferramenta Gephi.

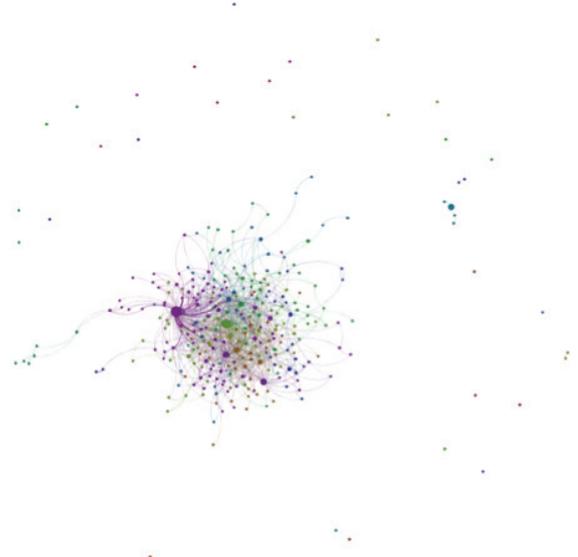


**Figura 9.** Imagem da rede inteira. Observe-se que, no algoritmo utilizado para “desenhar” a rede, os nós com mais conexões são atraídos para o centro, enquanto que aqueles sem conexões vão sendo “empurrados” para a periferia.

**Figure 9.** Whole network. We can see here that the nodes with more connections are attracted to the center, while the ones with fewer are pushed to the periphery.

é um pouco mais uniforme, com um topo de 24 *tweets* em um único nó. Observe-se, nas imagens a seguir (Figuras 11 e 12), como, na distribuição do *indegree*, há uma grande quantidade de nós que cita apenas uma vez outro nó, mas, logo a seguir, há uma distribuição de nós que cita várias vezes outro nó. Observe-se, nas Figuras 13 e 14, por outro lado, como a distribuição é menos uniforme no *cluster*, onde há mais citações por nó. Do mesmo modo, no *outdegree*, o número máximo de citações é menor, mas há também uma cauda longa<sup>14</sup> de nós que ainda têm muitos *tweets* e citações. Quando observamos o *cluster*, a quantidade de citações por nó aumenta, o que mostra que esses nós não apenas são mais ativos (tuitaram mais) e, igualmente, citaram-se mais.

Quando mapeamos apenas o *cluster*, a partir dos nós com maior número de *tweets*, novamente observa-se que o grafo é bastante conectado. O coeficiente de clusterização sobe para 0,033 e a média do grau de conexão para 3,88.



**Figura 10.** *Cluster* da rede. Aqui aparecem apenas os nós com mais de quatro *tweets*. Observe-se como estes estão com frequência interconectados e fazem menção uns aos outros (*cluster*).

**Figure 10.** Network Cluster. Here we see the nodes with more than four tweets. They are more frequently interconnected and make more citations to each other.

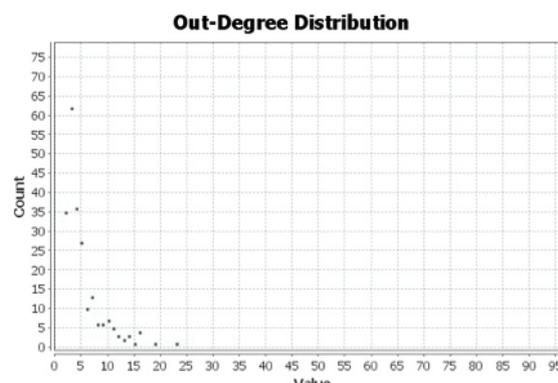
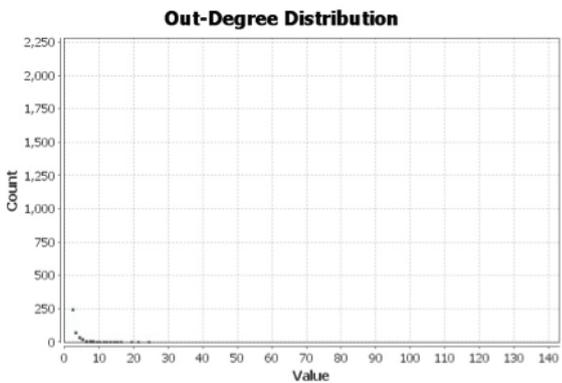
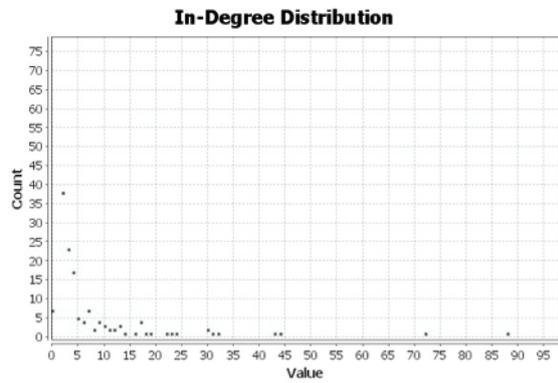
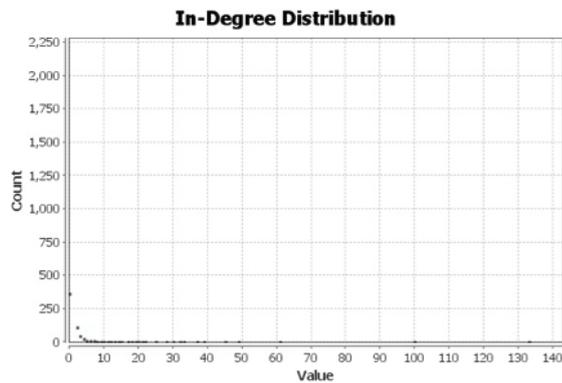
Isso significa que, entre os nós que fizeram mais do que quatro *tweets* com menções, há uma maior interconexão, tanto no *indegree* quanto no *outdegree*. Basicamente, isso quer dizer que esses nós tendem a se citar mais do que os demais, constituindo-se em um grupo que está ativamente “conversando”, ou seja, tuitando com a *hashtag* e citando-se entre si. Há uma ação no sentido de tuitar mais usando a *hashtag*, seja para aumentar sua visibilidade ou mesmo para participar da conversação mais fortemente.

Para entender um pouco melhor a distribuição dos *tweets* como apoio à presidentia, observamos também as *hashtags* presentes nos dados. É comum, no Twitter, a *hashtag* representar o contexto da fala, indicando, pela sua presença, as pistas de interpretação daquilo que é dito (Recuero, 2012). No caso, pudemos observar a presença de mais de 100 conjuntos de *hashtags* nos *tweets* (*tweets* com mais de uma *hashtag*). As *hashtags* encontradas com maior frequência juntas foram (Tabela 1):

<sup>14</sup> Referência ao termo “power-law”, uma distribuição onde há um pequeno número de nós com um alto *indegree/outdegree* e um grande número com baixo *indegree/outdegree* (Lei de Pareto).

**Tabela 1.** Dados das *hashtags* presentes nos *tweets* coletados.  
**Table 1.** Data from *hashtags* from the collected *tweets*.

<b>Hashtags Rede Inteira</b>	<b>%</b>	<b>Hashtags Cluster</b>	<b>%</b>
#tamojuntodilma	71,66%	#tamojuntodilma	86,47%
#calabocadilma #tamojuntodilma	11,72%	#tamojuntodilma #tamojuntodilma	2,59%
#tamojuntodilma #calabocadilma	2,15%	#tamojuntodilma #comdilhacontragolpe #ondavermelha	0,81%
#tamojuntodilma #calaabocadilma	2,01%	#tamojuntodilma #golpenao	0,81%
#tamojuntodilma #tamojuntodilma	0,89%	#violencianao #tamojuntodilma	0,65%
#tamojuntodilma #dilma	0,78%	Outros	8,67%
#violencianao #tamojuntodilma	0,41%		
Outros	10,38%		



**Figuras 11 e 12.** Gráficos de distribuição do *indegree* e *outdegree* na rede inteira.

**Figures 11 and 12.** Indegree and outdegree distribution in the whole network.

O uso massivo da *hashtag* de apoio indica que a maioria da rede estava construindo *tweets* de apoio. A presença de *hashtags* contraditórias indica usuários que desejam obter visibilidade de um grupo (apoiadores) ou

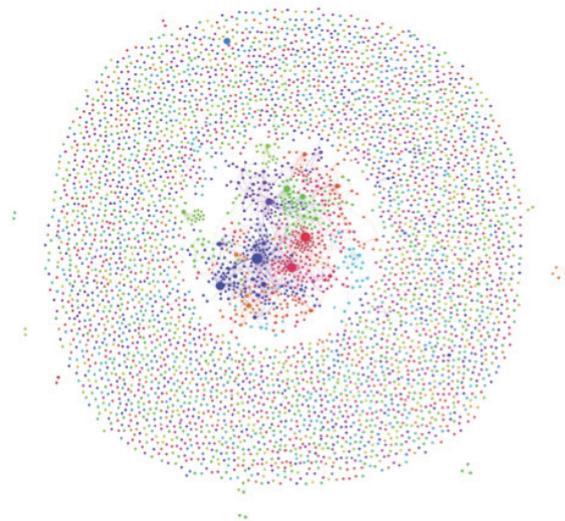
**Figuras 13 e 14.** Gráfico de distribuição do *indegree* e *outdegree* no *cluster* do grafo.

**Figures 13 and 14.** Indegree and outdegree distribution in the cluster.

do outro (não apoiadores) para seu *tweet*. Essas *hashtags*, entretanto, foram encontradas em maior quantidade em *tweets* de não apoio. Também é importante notar que há uma imensa massa de conjuntos de *tags* criando e

subvertendo sentidos. Observe-se também o conjunto de *hashtags* observadas com maior frequência nos *tweets* do *cluster* da rede. O que se vê é uma quantidade maior de *hashtags* com contexto de apoio ao pronunciamento.

O que isso tudo quer dizer? Primeiramente, o fato de ser possível observar um *cluster* que está interconectado mostra que há uma articulação da conversação entre os participantes, ou seja, não estamos falando de pessoas simplesmente falando do assunto, mas de pessoas citando-se, claramente envolvidas. O alto *indegree* e o alto *outdegree* parecem indicar que há uma ação coletiva de usuários falando sobre o assunto e usando a *tag*. Observe-se como o número do grau de conexão é mais alto dentro do *cluster*. Do mesmo modo, a presença de um grande número de *hashtags* de apoio como contexto, bem como a referência a apoiadores e partidários faz aparecer com mais força a hipótese de que há uma organização de um grupo em torno da *hashtag*, no sentido de obter visibilidade para seu apoio. Não é simplesmente o contexto de um *tweet*, mas uma manifestação dirigida, buscando-se tornar público o apoio à presidenta.



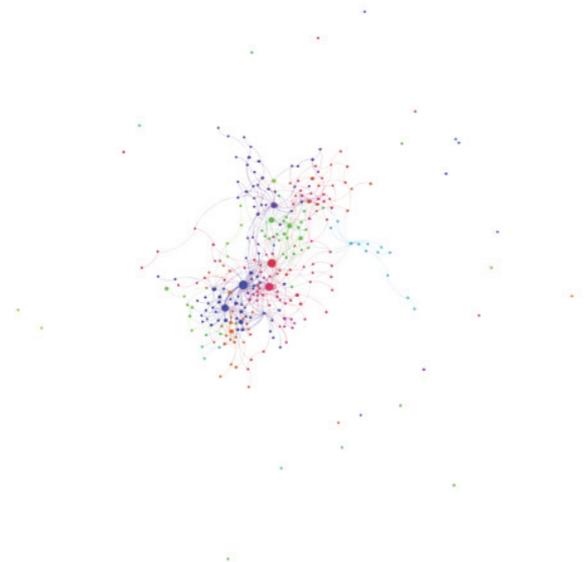
**Figura 15.** Imagem da rede inteira. Observe-se que, no algoritmo utilizado para “desenhar” a rede, os nós com mais conexões são atraídos para o centro, enquanto que aqueles sem conexões vão sendo “empurrados” para a periferia.

**Figure 15.** Whole network. Nodes with more connections are attracted to the center, while the ones with fewer are pushed to the periphery.

## #CalaabocaDilma

A coleta de dados por essa *hashtag* retornou mais resultados que a anterior. Foram 5.597 *tweets* de 4.804 atores, novamente indicando que mais de um *tweet* com a *hashtag* foi publicado pelo mesmo ator. A rede que observamos, entretanto, é menos clusterizada que a anterior (Figura 15). Mesmo no *cluster*, filtrado com o mesmo critério da anterior, há uma menor presença de nós e conexões (Figura 16). O coeficiente de clusterização está em 0,002 - bem inferior ao anterior. A média do grau de conexão é 1,165, também inferior. Entretanto, a rede tem mais *tweets* e mais atores, o que indica que, apesar do *cluster* ser menor e de as pessoas citarem-se menos, há mais gente participando. No *cluster*, é 2,811 o grau de conexão médio. O coeficiente de clusterização sobe para 0,017.

Esses dados já começam mostrando que se trata de uma rede menos fortemente conectada que a anterior. Ou seja, há mais nós, mas há menos participação em geral. Enquando que na rede anterior há muitos nós no *cluster*



**Figura 16.** Imagem do *cluster* da rede. Aqui aparecem apenas os nós com mais de quatro *tweets*. Observe-se como estes estão com frequência interconectados e fazem menção uns aos outros.

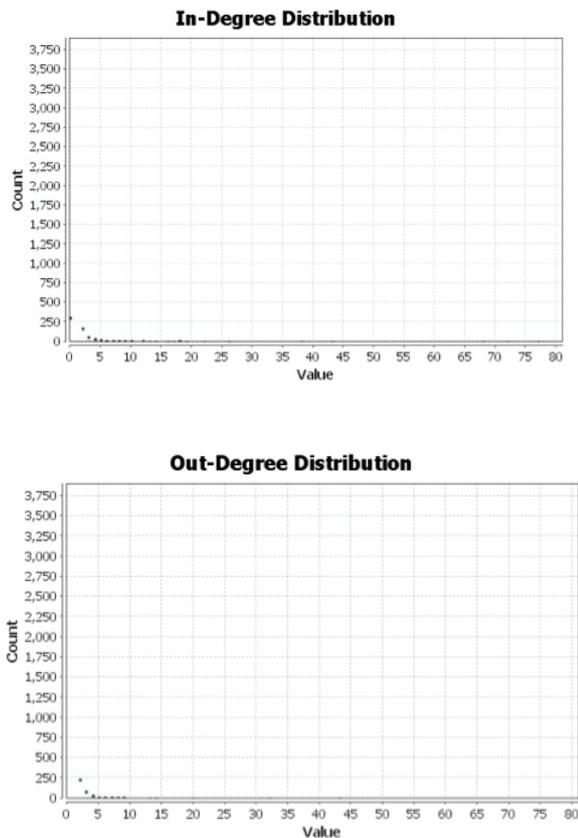
**Figure 16.** Network cluster. Nodes with more than four tweets. They are more frequently interconnected and make more citations to each other.

e muitas conexões de vários *tweets*, nesta, há menos nós e as conexões são mais fracas. Isso poderia indicar, por exemplo, que houve menor envolvimento dos atores na conversação com essa *hashtag*, menor articulação em torno dela. Isso talvez possa indicar que ela não surge de uma grande articulação de atores em conversar utilizando-a, mas de forma mais pontual em *tweets* esparsos de um maior número de atores.

Outro dado importante refere-se ao *indegree* máximo (71) e ao *outdegree* máximo (43) observados na rede. Os números máximos são altos (especialmente com relação ao *outdegree*), mas há menos nós envolvidos com a ação de tuitar usando a *hashtag*, conforme as figuras a seguir, que mostram a distribuição desses graus (Figuras 17 e 18). Observe-se que há menos nós com alto *indegree*, bem como menos nós com alto *outdegree* nas curvas totais dos *tweets*. No *cluster*, conforme esperado, há um aumento da participação dos nós, mas em relação

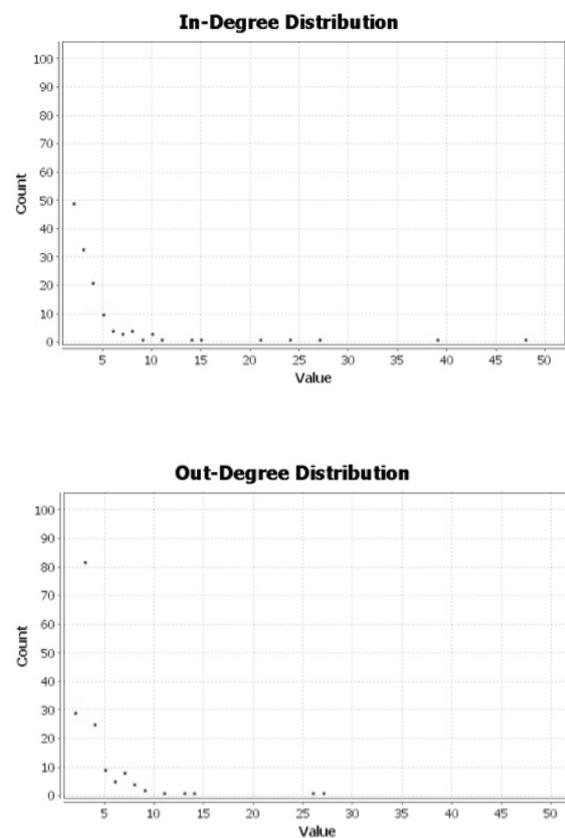
aos dados anteriores, não há tantos nós com alto nível de participação (Figuras 19 e 20). Isso poderia indicar que a rede, apesar de envolver mais participantes que a anterior, tem menos atores engajados em tuitar com a *hashtag* coletada. Novamente, temos indícios de um menor envolvimento dos atores.

Quando observamos as *hashtags* nos *tweets* (Tabela 2), novamente, há também indícios semelhantes aos da rede anterior. Há uma predominância do uso do contexto negativo, ou de *hashtags* negativas associadas. Vemos, como no outro exemplo, um cluster associado também na *hashtag* coletada com a “#tamojuntodilma”. Esses *tweets* talvez sejam aqueles com contexto mais dúbio, porque circulam tanto em um grupo quanto no outro. Como os *tweets* não foram analisados de modo qualitativo em seu conteúdo, podemos apenas indicar que focam visibilidade tanto de um grupo quanto de outro. Note-se, entretanto, que tanto na Tabela 1 quanto na Tabela 2, esses *tweets*



**Figuras 17 e 18.** Gráficos de distribuição do *indegree* e do *outdegree* na rede inteira.

**Figures 17 and 18.** Whole network *indegree* and *outdegree* distribution.



**Figuras 19 e 20.** Gráficos de distribuição do *indegree* e do *outdegree* no *cluster*.

**Figures 19 and 20.** Network *cluster* *indegree* and *outdegree* distribution.

Tabela 2. *Hashtags* dos *tweets*.

Table 2. Hashtags.

<b>Hashtags na Rede Inteira</b>	<b>%</b>	<b>Hashtags no Cluster</b>	<b>%</b>
#calaabocadilma	81,92%	#calaabocadilma	81,07%
#calabocadilma #tamojuntodilma	4,98%	#calabocadilma #foradilma	2,03%
#calabocadilma #foradilma	1,43%	#tamojuntodilma #calabocadilma	1,52%
#tamojuntodilma #calabocadilma	0,88%	#calaabocadilma #vemprarua	0,89%
Outros	10,79%	Outros	14,49%

aparecem em um percentual bem menor que aqueles de apoio ou de crítica.

Novamente, o que a análise das redes das duas *hashtags* parece mostrar? Primeiro, que observamos um maior *indegree* e *outdegree* na rede do #tamojuntodilma, o que parece indicar um envolvimento maior dos atores na conversação. O grau de clusterização mais alto, tanto no *cluster* quanto no grafo também indica uma maior participação dos atores na conversação com essa *hashtag*. Na #Calaabocadilma, ao contrário, há um grau de clusterização bem menor, apesar de uma quantidade maior de atores e *tweets*. Isso parece também indicar que houve um envolvimento menor na conversação, com uma maior organicidade e menor organização nas críticas. A *tag* de apoio, ao contrário, parece emergir de um movimento ativo da rede, especialmente entre os apoiadores. Ao mesmo tempo, a análise das *hashtags* também nos mostra um contexto específico onde tanto *tags* de crítica quanto de apoio foram utilizadas juntas. Isso também poderia indicar que tanto um grupo quanto o outro deseja fazer públicas (nas duas conversações) suas opiniões. E mostra também que a presença de uma *tag* influenciou a outra. Ou seja, ambas as redes posicionam-se em conversações, parcialmente interferindo-se quando usam as duas *hashtags* juntas. Há uma movimentação mais forte junto à *hashtag* de apoio, mas uma crítica mais orgânica e com maior número de participantes na *hashtag* contrária. O uso da ARS nestes casos, portanto, auxiliou a medir e a observar os dados dos padrões de interação da rede. Pudemos observar as medidas como indicativos desses padrões, bem como discutir possíveis implicações.

## Conclusões

O presente artigo buscou apresentar e discutir as contribuições da Análise de Redes Sociais para o es-

tudo das redes sociais na Internet. De modo específico, apresentamos a abordagem, discutimos suas métricas e contribuições, bem como suas limitações, e procuramos demonstrá-la em um estudo de caso pontual. No caso específico, procuramos demonstrar como algumas medidas oferecidas pela ARS poderiam ajudar a perceber as estruturas das conversações em torno de duas *hashtags* (#tamojuntodilma e #calaabocadilma) que surgem durante o primeiro pronunciamento à nação da presidenta Dilma durante os protestos que acontecem no Brasil em junho de 2013. O caso mostrou que enquanto uma das *hashtags* foi bastante articulada por um grupo, a outra foi mais orgânica e teve a participação de um maior número de pessoas. Embora limitado, o estudo traz uma breve aplicação da ARS apontando alguns indícios a partir dos dados recolhidos do Twitter.

A partir dessas discussões, podemos apontar que as principais contribuições da ARS para o estudo das redes sociais na Internet dão-se, principalmente, a partir de três grandes pontos:

(a) *Foco empírico* - A ARS traz, novamente, os dados empíricos para o centro da pesquisa. Resolvendo parte do problema de trabalhar com muitos dados, a ARS auxilia a mapear e observar as estruturas construídas através das interações de centenas ou milhares de atores, oferecendo ferramentas que auxiliam tanto pequenos casos quanto casos onde há uma grande quantidade de dados. Sistematização de coleta e análise de dados- A ARS oferece mecanismos de sistematização de coleta e análise de dados, proporcionando uma abordagem completa de estudo dos rastros que são construídos pelos atores nas ferramentas de comunicação mediada pelo computador.

(b) *Abordagem interdisciplinar* - permite ao pesquisador analisar a estrutura dos grupos sociais a partir de uma pluralidade de pontos de vista. A ARS é uma abordagem de fundo interdisciplinar que serve tanto às ciências sociais e humanas quanto às ciências exatas. Desde que o foco que se objetiva seja na estrutura da rede, há uma pluralidade de focos e objetos possíveis. Além disso, o

foco permite o diálogo com várias perspectivas científicas, fundamentando uma pesquisa interdisciplinar real que é tão necessária no estudo das redes sociais *online*.

(c) *Foco nos padrões e na estrutura* - Um dos grandes aspectos dos dados específicos das redes sociais, principalmente pelo seu aspecto de “rede” está no foco nos padrões de estrutura e nos padrões emergentes que os dados salientam. E é justamente este o viés da ARS que, portanto, auxilia a perceber e estabelecer melhor as inter-relações entre os conceitos teóricos trabalhados (tais como capital social, comunidades virtuais e etc.) com os dados empíricos coletados dos públicos em rede.

Essas contribuições, sistematizadas aqui, foram as principais que procuramos demonstrar no artigo. Outros estudos poderão explorar com maior profundidade esses elementos propostos, bem como seus efeitos e impactos nas redes sociais na Internet.

## Referências

- AMARAL, A. 2010. Redes sociais de música: Segmentação, Apropriações e Práticas de Consumo. *Revista ComCiência*. Disponível em: <http://www.comciencia.br/comciencia/handler.php?section=8&edicao=59&cid=748>. Acesso em 20/08/2013.
- BENEVENUTO, F. 2010. Redes Sociais Online: Técnicas de Coleta, Abordagens de Medição e Desafios Futuros. In: A.C.M. PEREIRA; G.L. PAPPÀ; M. WINCKLER; R.L. GOMES (orgs.), *Tópicos em Sistemas Colaborativos, Interativos, Multimídia, Web e Banco de Dados*. Belo Horizonte, Sociedade Brasileira de Computação, p. 41-70. Disponível em: <http://homepages.dcc.ufmg.br/~fabricio/download/mini-curso-swib10.pdf>. Acesso em: 20/08/2013.
- BOYD, D. 2010. Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications. In: Z. PAPACHARRISSI (ed.), *Networked Self: Identity, Community, and Culture on Social Network Sites*, New York, Futurelab, p. 39-58.
- BOYD, D.; GOLDER, S.; LOTAN, G. 2010. Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter. Disponível em: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5428313>. Acesso em: 20/08/2013.
- BOYD, D.; ELLISON, N. 2007. Social Network Sites: Definition, history and Scholarship. *Journal of Computer-Mediated Communication* **13**(1):210-230. <http://dx.doi.org/10.1111/j.1083-6101.2007.00393.x>
- BURT, R. 1995. *Structural Holes: The Social Structure of Competition*. Harvard, Harvard University Press, 324 p.
- DEGENNE, A.; FORSÉ, M. 1999. *Introducing Social Networks*. London, Sage, 256 p.
- ELLISON, N.; C. STEINFELD; LAMPE, C. 2007. The Benefits of Facebook “Friends:” Social Capital and College Students’ Use of Online Social Network Sites. *Journal of Computer Mediated Communication*, **12**(4):1143-1168. <http://dx.doi.org/10.1111/j.1083-6101.2007.00367.x>
- FREEMAN, L. 2004. *The Development of Social Network Analysis: A Study in the Sociology of Science*. Vancouver, Empirical Press, 218 p.
- GRANOVETTER, M. 1973. The Strength of Weak Ties. *American Journal of Sociology*, **78**(6):1360-1380. <http://dx.doi.org/10.1111/j.1083-6101.2007.00367.x>
- HAMPTON, K.; WELLMAN, B. 2003. Neighboring in Newville: How the Internet supports community and social capital in a wired suburb. *City & Community*, **2**(4):277-311. <http://dx.doi.org/10.1046/j.1535-6841.2003.00057.x>
- HOGAN, B.; CARRASCO, A.; WELLMAN, B. 2007. Visualizing Personal Networks: Working with Participant-Aided Sociograms. *Field Methods* **19**(2):116-144. Disponível em [http://homes.chass.utoronto.ca/~wellman/cgi-bin/counter.php?url=http://chass.utoronto.ca/~wellman/publications/visualizing\\_personal\\_networks/visualizing\\_personal\\_networks.html&f=visualizing\\_personal\\_networks&mode=1](http://homes.chass.utoronto.ca/~wellman/cgi-bin/counter.php?url=http://chass.utoronto.ca/~wellman/publications/visualizing_personal_networks/visualizing_personal_networks.html&f=visualizing_personal_networks&mode=1). Acesso em: 20/08/2013.
- HUBERMAN, B.; ROMERO, D.; WU, F. 2009. Social Networks that Matter: Twitter under the Microscope. *First Monday*, **14**(1). Disponível em: <http://firstmonday.org/ojs/index.php/fm/article/view/2317/2063>. Acesso em: 20/08/2013.
- KADUSHIN, C. 2005. Who benefits from network analysis: ethics of social network research. *Social Networks*, **27**(2):139-153. Disponível em: <http://www.sciencedirect.com/science/article/pii/S037887330500016X>. Acesso em: 20/08/2013.
- KLEINBERG, J. 2007. Challenges in mining social network data: processes, privacy, and paradoxes. *Proceedings of the 13<sup>th</sup> ACM SIGKDD International Conference on Knowledge discovery and data mining*, New York, **1**:4-5. Disponível em: <http://dl.acm.org/citation.cfm?id=1281195>. Acesso em 16/08/2014. <http://dx.doi.org/10.1145/1281192.1281195>
- LEMIEUX, V.; OUMET M. 2004. *Análise Estrutural das Redes Sociais*. Lisboa, Instituto Piaget, 132 p.
- MALINI, F.; ANTOUN, H. 2013. *@Internet e #Rua. Ciberativismo e Mobilização nas Redes Sociais*. Porto Alegre, Sulina, 278 p.
- MIZRUCHI, M.S. 2006. Análise de redes sociais: avanços recentes e controvérsias atuais. *Revista de Administração de Empresas - RAE*, **46**(3):72-86. Disponível em: <http://www.scielo.br/pdf/rae/v46n3/v46n3a13.pdf>. Acesso em: 20/08/2013.
- PASSERINO, L.; MONTARDO, S.; BENKENSTEIN, A. 2007. Análise de Redes Sociais em Blogs de Pessoas com

- Necessidades Especiais (PNE). *Novas tecnologias na educação*, 5(2):1-15. Disponível em: <http://seer.ufrgs.br/index.php/renote/article/view/14383/8279>. Acesso em 16/08/14.
- PEREIRA, B.C.; ARAÚJO, R.F. 2013. *Revista Anagrama*, 7(1):1-15 . Disponível em: <http://www.revistas.univerciencia.org/index.php/anagrama/article/view/8622/7918>. Acesso em: 20/08/2013.
- RECUERO, R. 2009. *Redes Sociais na Internet*. Porto Alegre, Sulina, 191 p.
- RECUERO, R. 2012. *A Conversação em Rede*. Porto Alegre, Sulina, 238 p.
- RECUERO, R.; AMARAL, A.; MONTEIRO, C. 2012. Fandoms, Trending Topics and Social Capital in Twitter. *Selected Papers of Internet Research*, 1:1-24.
- SCOTT, J. 2004. *Social Network Analysis: A handbook*. 2ª ed., Londres, Sage, 208 p.
- STEINFELD, C.; ELLISON, N.; LAMPE, C. 2008. Social capital, self-esteem, and use of online social network sites: A longitudinal analysis. *Journal of Applied Developmental Psychology*, 29(6):434-445. Disponível em: [https://www.msu.edu/~steinfie/Steinfeld\\_Ellison\\_Lampe\(2008\).pdf](https://www.msu.edu/~steinfie/Steinfeld_Ellison_Lampe(2008).pdf). Acesso em: 20/08/2013.
- WASSERMAN, S.; FAUST, K. 1994. *Social Network Analysis. Methods and Applications*. Cambridge, Cambridge University Press, 857p. <http://dx.doi.org/10.1017/CBO9780511815478>
- WELLMAN, B. 2001. Computer Networks As Social Networks. *Science*, 293(5537):2031-2034. <http://dx.doi.org/10.1126/science.1065547>
- WELLMAN, B.; BERKOWITZ, S. 2007. *Social Structures: A Network Approach*. USA, Emerald Group Publishing Limited, 528 p. (Contemporary Studies in Sociology, Vol. 15).
- WELLMAN, B.; HAASE, A.Q.; WITTE, J.; HAMPTON, K. 2001. Does the Internet increase, decrease, or supplement social capital? Social networks, participation, and community commitment. *American Behavioral Scientist*, 45(3):436. <http://dx.doi.org/10.1177/00027640121957286>
- WEST, A.; LEWIS, J.; CURRIE, P. 2009. Students' Facebook 'friends': public and private spheres. *Journal of Youth Studies*, 12(6). Disponível em: <http://www.tandfonline.com/doi/abs/10.1080/13676260902960752#.UiXz4mQqfpw>. Acesso em: 20/08/2013.
- ZAGO, G. 2010. Do Boato à Notícia: Considerações sobre a Circulação de Informações entre Twitter e Mídia Online de Referência. In: A. AMARAL; M.C. AQUINO; S. MONTARDO, *INTERCOM SUL 2010: Perspectivas da Pesquisa em Comunicação Digital*. São Paulo, Intercom, p. 174-206. Disponível em: <http://galaxy.intercom.org.br/ebooks/arquivos/ad5879b9c4f8d22f17340e630f1bed60.pdf#page=174>. Acesso em: 20/08/2013.
- ZAGO, G.; REBS, R. 2011. Usos e Apropriações do Foursquare no Brasil: Apontamentos para Discussão. *Revista Nexi*, 1:1-15 Disponível em: <http://revistas.pucsp.br/index.php/nexi/article/view/3312>. Acesso em: 20/08/2013.

Submissão: 03/09/2013

Aceite: 17/11/2013