

**INTERPRETAÇÃO DE FLUXO INFORMACIONAL ATRAVÉS DA  
FERRAMENTA NODEXL<sup>1</sup>  
um projeto open-source para redes sociais, de tecnologia e informação.**

**Jeferson Carvalho<sup>2</sup>**

**Resumo**

Cada vez torna-se mais comum os estudos sobre Redes Sociais, principalmente com o advento da internet que transfere e agiliza essas relações para um ambiente digital através de softwares e aplicativos. A internet é uma Rede de Tecnologia, mas quando dela abstraído a Web, torna-se uma Rede de Informação. Estudar, atualmente, as Redes Sociais Conectadas à internet é interpretar fluxos de informações derivados de relacionamento nesses ambientes digitais. O NodeXL é uma ferramenta *open-source* que permite interpretar dados dinâmicos, apresentando vértices, grafos, nós e demais métricas encontradas no estudo de redes em uma planilha.

**Palavras-chave:** Comunicação. Tecnologia. Redes Sociais. Fluxo de Informação. Open Source.

---

<sup>1</sup> Texto original, como recebido pela coordenação do Interprogramas.

<sup>2</sup> Faculdade Cásper Líbero. Mestrado em Processos Midiáticos. Email: [jefersondecarvalho@gmail.com](mailto:jefersondecarvalho@gmail.com)

## INTRODUÇÃO

A internet é uma Rede de Tecnologia, mas quando dela abstraído a Web, torna-se uma Rede de Informação. Estudar, atualmente, as Redes Sociais conectadas à internet é interpretar fluxos de informações derivados de relacionamento nesses ambientes digitais. O objeto desse estudo é a apresentar através de demonstração sobre os fundamentos epistemológicos de estudo de Redes de Tecnologia e Informação o alcance do projeto *open-source*<sup>1</sup> denominado NodeXL, ferramenta que permite interpretar dados dinâmicos, apresentando vértices, tríades, nós e demais métricas encontradas no estudo de redes em uma planilha do Microsoft Excel.

O estudo científico das redes é um campo interdisciplinar combinando ideias de diferentes áreas, como matemática, economia, biologia, ciência da computação, ciências sociais e outras. O interesse nos estudos de redes tem crescido nos últimos anos, principalmente a partir da década de 90. A evolução computacional tem ajudado o processamento de dados instituindo o estudo e facilitando o emprego de seus conceitos básicos.

As publicações em ciências sociais com títulos de “redes sociais” aceleraram exponencialmente nas últimas três décadas, e a tendência não mostra nenhum sinal de queda. Uma proliferação similar ocorreu nos estudos, pesquisas e aplicações dos fenômenos de redes sociais, desde os naturais aos mais complexos, em matemática e física. (Knoke: 2008:2)<sup>2</sup>

---

<sup>1</sup> A definição de *Open Source* foi criada pela Open Source Initiative (OSI) e determina que um programa de código aberto deve garantir distribuição livre, código fonte, trabalhos derivados, (a licença deve permitir modificações), e integridade do autor sobre o código fonte.

<sup>2</sup> Tadução livre do autor: “Social science publications with “social network” as key concept accelerated exponentially during the past three decades, a trends showing no signs of slackening. A similar proliferation of network research applications to natural phenomena as well as to complex social system occurred in mathematics and physics”.

## 7º Interprogramas de Mestrado

O avanço tecnológico, a facilidade de encurtar as distâncias, as viagens, a comunicação global e a exploração com finalidade econômica desses ambientes, contribuem para as crescentes interações digitais através da utilização de sites de Redes Sociais Conectadas (Lima, 2009) como Facebook, Orkut, Myspace.

A existência de três componentes nas Redes, independente da natureza e do recorte social, biológico, econômico ou físico, que deve ser estudados são: os indivíduos, as conexões ou interconexões e a interação entre os sistemas. Esse consenso foi encontrado nos autores pesquisados. Entretanto, esse terceiro componente, carrega um fator crucial para as Rede que é o padrão dessas conexões, (Barabási:2010).

A internet é exemplo primário de uma Rede Tecnológica, que tecnicamente é uma rede global conectada via cabos, fibras e satélites que une computadores e sistemas computacionais. Através da Rede Tecnológica da internet pacotes de informações são enviados e recebidos respeitando os protocolos da internet. A interação na internet não é burocrática e nem hierárquica, não é raro pacotes de informações se perderem por falhas em software ou hardware, “o que contrivui para o equívoco, que a internet é caótica e não altamente controlada.” (Galloway:2004).

O ambiente web está na camada *browser*, e o grande responsável pela sua criação foi Tim Berns-Lee, conseqüentemente desencadeando uma variedade de softwares e aplicações. A web é uma Rede de Informação. Os nós da web são as páginas de cada site, de cada portal de conteúdo e os vértices são os *hyperlinks* que são textos grifados que puxam ou levam a navegação entre conteúdos renderizados em outras páginas da web. Um *hyperlink* é um software.

Software, é claro, é o que organiza a Internet, direciona as mensagens de email, entrega as páginas da Web para seus servidores, comuta o tráfego da rede, associa endereços IP e encaminha as páginas da Web para o *browser*. (Manovich, 2008:3)<sup>3</sup>

---

<sup>3</sup> Tradução livre do autor: Software, of course, is what organizes the Internet, routing email messages, delivering Web pages from a server, switching network traffic, assigning IP addresses, and rendering Web pages in a browser.

## 7º Interprogramas de Mestrado

A partir do momento em que usuários começam a adicionar *hyperlinks* entre suas páginas de conteúdo a estrutura da web passa a revelar padrões desse sistema. Quanto mais *hyperlinks* relacionados com um conteúdo maior a relevância dessa página. Esse é um conceito básico utilizado pela engenharia de busca semântica na internet. “A operação realizada em 1996 e batizada de BackRub foi uma cópia da toda a World Wide Web para preparar a indexação de busca do Google” (Wu: 2010).

Com o advento tecnológico, a disponibilidade dos recursos computacionais barateados e o acesso à rede em larga escala têm massificado a utilização das estruturas de Redes Sociais, tornam-se conectadas e transferidas para o ambiente digital (Lima:2009), disponibilizando maior nível de detalhes sobre os estudos e aparecendo novos métodos para o estudo de métrica das redes apoiados no poder computacional.

### **FERRAMENTA NODEXL**

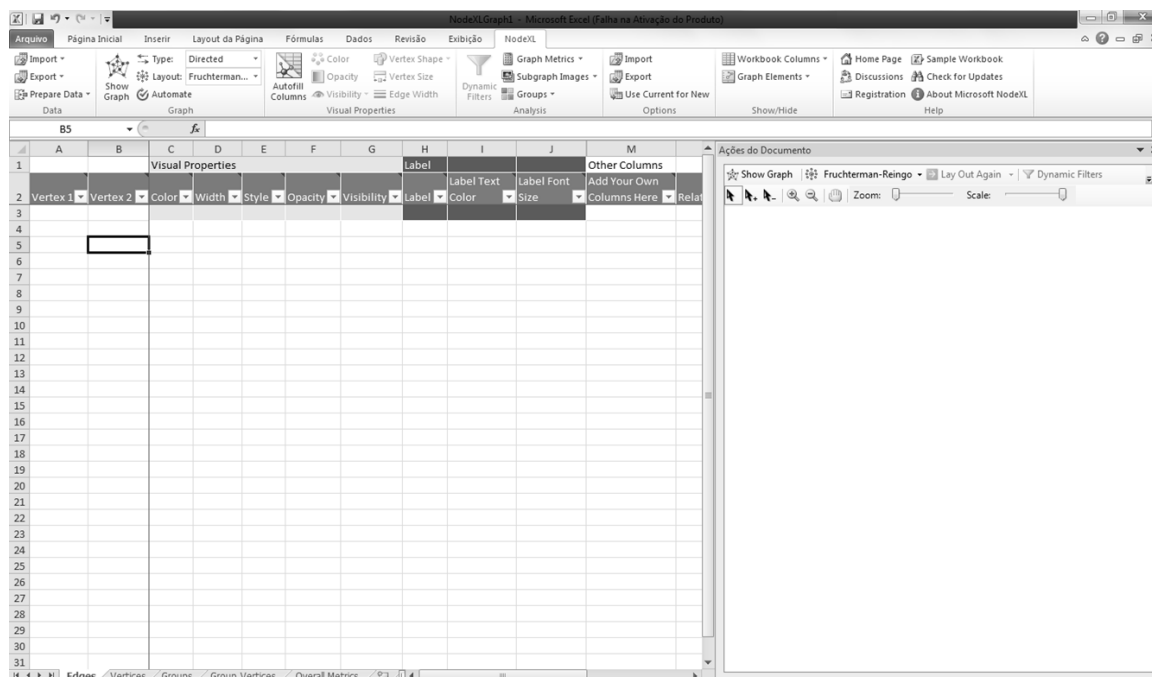
A ferramenta denominada NodeXL foi desenvolvida por um time da Fundação de Pesquisa e Projetos Externos da Microsoft, distribuído em equipes nos Estados Unidos e na Europa e contando com colaboradores ao redor do mundo. Durante os primeiros 2 anos do projeto foram levantados os requerimentos necessários e as devidas prioridades para disponibilizar as facilidades existentes em outras ferramentas de análise gráfica, integração estatística e visualização estratégica de informações evitando a complexidade e pensando na utilização em grande escala por perfis distintos de usuários.

O manuseio e a visualização de vértice, tamanho, cor, formato, imagens e rótulos, além de outras propriedades paralelas como o ajuste de espessura da linha são as características que enriquecem o controle de visualização do NodeXL. A ferramenta utiliza a plataforma do Microsoft Excel e se apropria de seus recursos matemáticos.

# 7º Interprogramas de Mestrado

O NodeXL é um complemento gratuito para o Microsoft Excel e com interfaces disponíveis no conceito de *open source*, de tempos em tempos novas atualizações são disponibilizadas, os testes são realizados nas redes das próprias instituições acadêmicas e adicionalmente com os profissionais que participam do projeto. As atualizações visam melhoras em usabilidade, esclarecimento dos processos de análises e nos aspectos da estratégia de visualização e informação.

Figura 1 – Após instalar o complemento do NodeXL para o Excel, disponível na Comunidade Codeplex, o usuário deve procurar a pasta do programa do “Microsoft NodeXL” e executar o arquivo “NodeXL Microsoft Excel Template”, que abrirá o programa Microsoft Excel com a extensão NodeXL já associada a ele.



As interfaces *open source* disponibilizadas permitem que programadores desenvolvam suas próprias rotinas e adaptem às necessidades requeridas. Para não programadores o NodeXL tem uma estrutura nativa em tabelas funcionais previamente configurada para apresentar múltiplas representações gráficas e interpretativas de relacionamentos dentro de uma rede social de uma forma simples pois manipula os gráficos de acordo com as variáveis disponíveis para as redes, como por exemplo, You Tube, Facebook, Twitter, Blogs e Emails.

## 7º Interprogramas de Mestrado

O NodeXL pode apoiar os pesquisadores que estão interessados em informações sobre análises de redes sociais e também a profissionais que desejam aplicar o estudo de redes em problemas práticos. A familiaridade apropriada das planilhas de Excel traz diversos filtros, diagramação, propriedades customizáveis e demais controles manuais integrados com métricas, métodos estatísticos e as visualizações gráficas de relatórios.

O arquivo para instalação do NodeXL está disponível no site da comunidade Codeplex da Microsoft <sup>4)</sup> e funciona para as versões a partir do Excel 2007, ele não está disponível para as versões do OpenOffice e tampouco para Mac OS. Ele também pode importar versões de outros programas como o Pajek, UCINET, GraphML e Vizster. E possui, alternativamente, a facilidade de importar arquivo no formato Excel para programas sem compatibilidade ou fora de utilização como o GRADAP e STRUCTURE utilizados pelos autores Wasserman e Faust e citado no destacável tratado sobre análise em Redes Sociais, *Social Network Analysis*, que comentam “uma pena que modelos estatísticos não estão incluídos em nenhum dos pacotes dos computadores para a análise de redes sociais” <sup>5</sup> (Wasserman:1994).

---

<sup>4</sup> Endereço para acesso a Comunidade Codeplex <http://nodexl.codeplex.com>

<sup>5</sup> Tradução livre do autor: “It is quite unfortunate that adequate statistical models are not included in any of the major social network analysis computer packages.

## FLUXO INFORMACIONAL

O fluxo informacional é derivada da relação existente em uma Rede. Há algumas definições de Redes que variam de acordo com suas aplicações. Uma definição básica é qualquer coleção de objetos nas quais alguns pares deles estejam conectados através de algo <sup>6</sup>. (Easley e Kleinberg, 2010). Outra definição mais estruturada e utilizada em comum por autores que estudam as Redes Biológicas, Tecnológicas, Sociais e de Informação, é a Rede como uma coleção de pontos agrupados juntos em pares por linhas <sup>7</sup>, (Newman: 2010). A relação direcionada, derivada dos pontos dicotômicos revela o fluxo existente, independente da frequência ou da intensidade entre eles. Tradução livre do autor: “A network is, in its simplest form, a collection of points joined together in pairs by lines. In the jargon of the field the points are referred to as vertices or nodes and the lines are referred to as edges.”

O diferencial do NodeXL é permitir a aplicação de filtros em seus componentes e a existência de relatórios customizados, a partir de uma importação de base de dados de e-mails, e a conexão de forma *on line* ao Twitter e a portais de fotos e vídeos como Flickr e Youtube.

A seguir veremos o objeto desse estudo que é a apresentação, através de demonstração sobre os fundamentos epistemológicos de estudo de Redes de Informação, o alcance da ferramenta *open-source* NodeXL na interpretação de fluxos informacionais nos ambientes de e-mail, Facebook e Twitter.

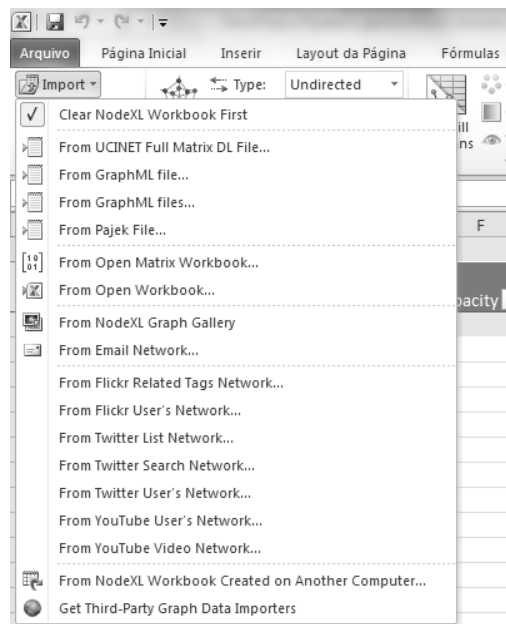
---

<sup>6</sup> Tradução livre do autor: “In the most basic sense, a network is any collection of objects in which some pairs of these objects are connected by links.”

<sup>7</sup> Tradução livre do autor: “A network is, in its simplest form, a collection of points joined together in pairs by lines. In the jargon of the field the points are referred to as vertices or nodes and the lines are referred to as edges.”

# 7º Interprogramas de Mestrado

Figura 2 – Na figura abaixo estão apresentadas as possibilidades de importação direta e indireta dos ambientes de Rede Sociais Conectadas como Flickr, Twitter e Youtube. Além da importação de dados no formato de outros programas de interpretação de fluxo informacional como UCINET, GraphML e Pajek. E, finalmente, a importação a partir de um cliente de email e de planilhas de dados.



## EMAIL

A forma mais simples de transformar mensagens de e-mails em uma rede de relacionamento é usando a facilidade de importação do NodeXL. Essa funcionalidade permite que o sistema operacional Windows crie arquivos indexados nos clientes de e-mails <sup>6</sup> Outlook Express e Outlook.

Ao escolher a importação de dados diretamente para o NodeXL os e-mails recolhidos podem ser organizados pelos filtros disponíveis como tamanho, remetente, assunto e outros. Recomenda-se uma prévia filtragem antes do início da interpretação pois a quantidade de mensagens pode comprometer a análise.

---

<sup>6</sup> Cliente de e-mail é um programa de correio eletrônico e notícias que acessa arquivos de mensagens eletrônicas de um provedor de mensagem, lê arquivos XML e RSS e têm filtros de anti-mensagem indesejável e mecanismo de prevenção a golpes.



# 7º Interprogramas de Mestrado

Figura 3 – A caixa de ajuda para importação de e-mails disponibiliza opções de filtros para reduzir a quantidade de mensagens a serem apresentadas graficamente como selecionar Remetente em Cópia, Remetente em Cópia Oculta, Destinatário, Endereço de Pastas Organizadas, Data de Envio ou Recebimento, com ou sem Arquivos anexados e pelo tamanho deles.

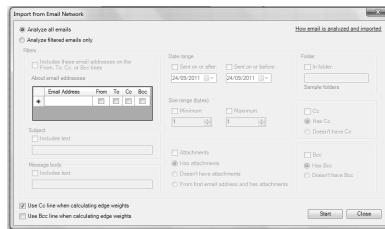


Figura 4 – Resultado da importação de e-mails recebidos pelo endereço eletrônico jefersondecarvalho@terra.com.br utilizando o complemento NodeXL.

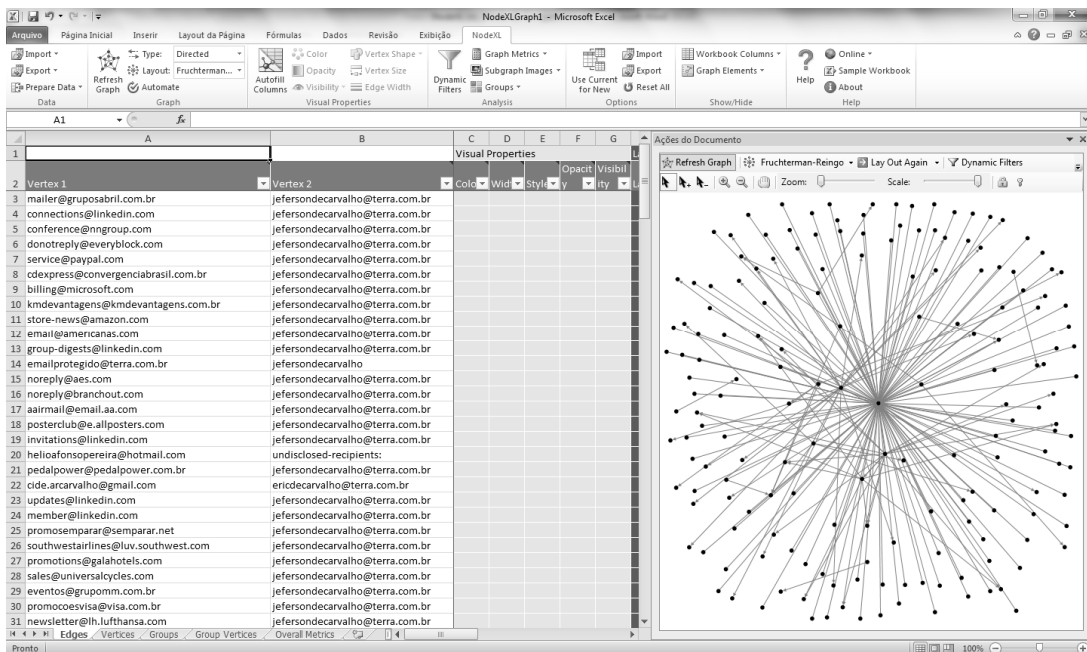


Figura 5 – A rede de e-mails pode ser uma rede egocêntrica, onde o objeto está no foco central, quando analisado somente o fluxo unidirecional. Os nós representam os interlocutores, remetentes ou destinatários, e o sentido das conexões representam o fluxo da informação: enviada ou recebida. Destaque abaixo de uma mensagem enviada pelo endereço eletrônico bol@bol.com.br e recebido pelo autor no e-mail jefersondecarvalho@terra.com.br.

## 7º Interprogramas de Mestrado

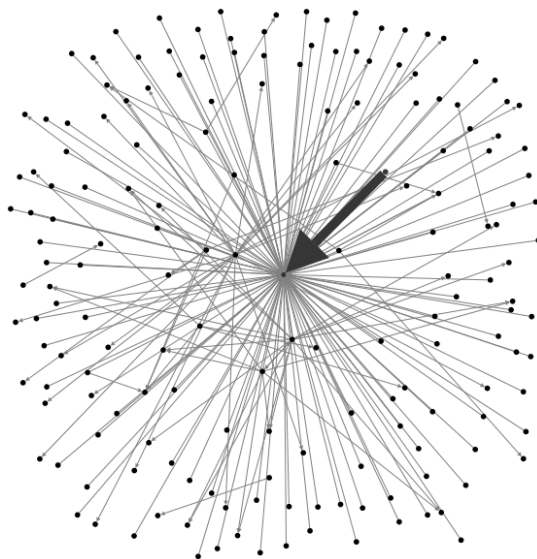
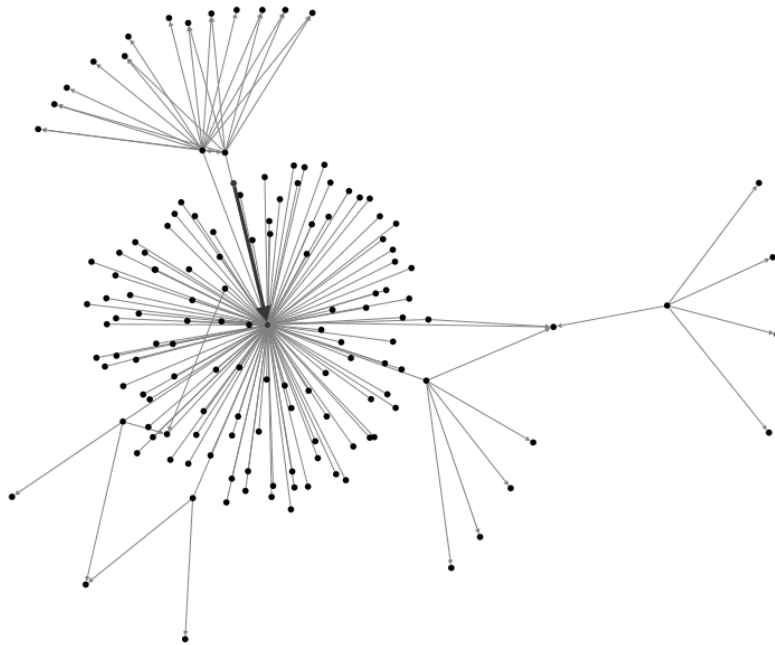


Figura 6 – O complemento NodeL permite ajustar a representação gráfica através de modelos pré-configurados adaptando a exibição para tipos variados de redes.



Figura 7 – Representação gráfica do recebimento de e-mails do autor com exibição gráfica ajustada para uma rede não egocêntrica.

## 7º Interprogramas de Mestrado



### FACEBOOK

O Facebook disponibiliza uma Interface de Aplicação de Programas (API) <sup>7</sup> para capturar os dados de sua rede, alternativamente seria possível capturar dados da tela do próprio Facebook, porém esse artifício não foi desenvolvido, e sim, o respeito a política de privacidade. A utilização dessa API permite a importação dos dados que são definidos pelo usuário, portanto é possível saber, a partir de um indivíduo, que ele obtenha dados de todos os seus amigos e, eventualmente, daqueles que são os amigos dos meus amigos.

O NodeXL não se conecta diretamente com o perfil do usuário no Facebook, mas sim, após importar um arquivo no formato GraphXL. Ao acessar o link do aplicativo Namegenweb, disponível no seguinte endereço: <http://apps.facebook.com/namegenweb>, o usuário deve escolher e baixar o arquivo que conterá as informações a serem importadas pelo NodeXL.

---

<sup>7</sup> API (Application Programmers Interfaces) é um conjunto de rotinas e padrões estabelecidos por um software para a utilização das suas funcionalidades por aplicativos que não pretendem envolver-se em detalhes da implementação mas apenas usar seus serviços.

## 7º Interprogramas de Mestrado

A importação de dados resultará numa rede egocêntrica onde todos os indivíduos serão amigos do usuário que está importando as informações. Sabendo como é a sua rede e quais as pessoas que estão conectadas pode ajudar a refinar a privacidade para a importação de informações, por exemplo, filtrando familiares de sua rede. O tempo médio para baixar as informações é de 1 minuto para cada 200 amigos.

Figura 8 – Representação gráfica da rede de relacionamento do autor com característica egocêntrica, todos os vértices são amigos, uma rede de primeiro grau, e amigos dos amigos, nesse caso uma rede de segundo grau, que é somente possível consegui-la se a parametrização de privacidade dos demais usuários esteja determinada para visualização pública.

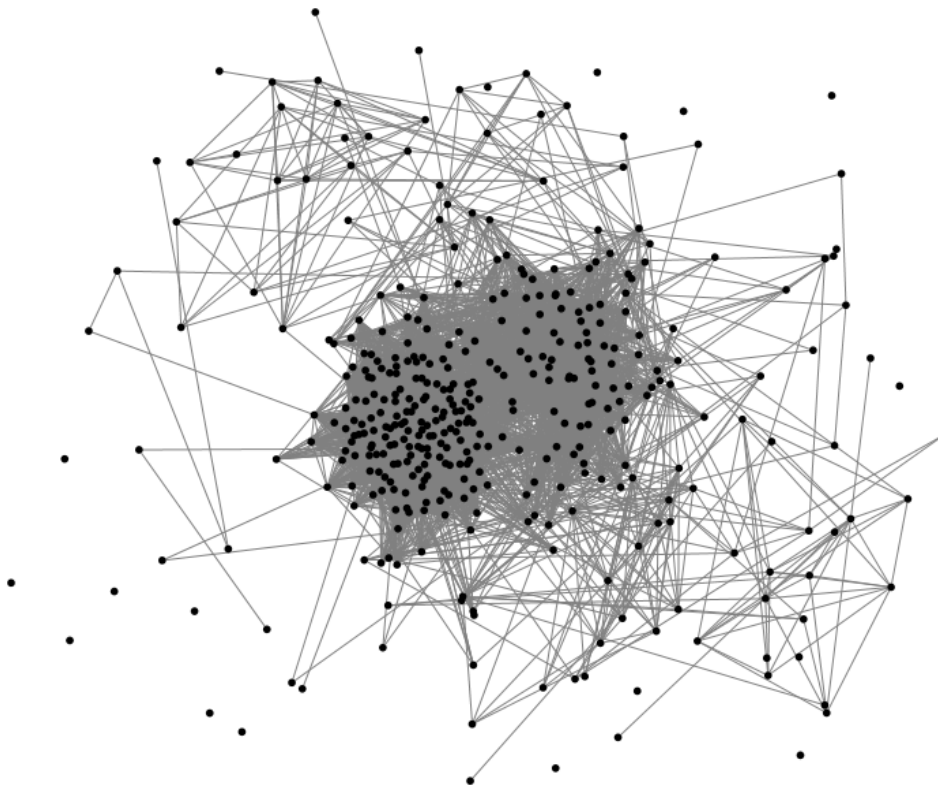
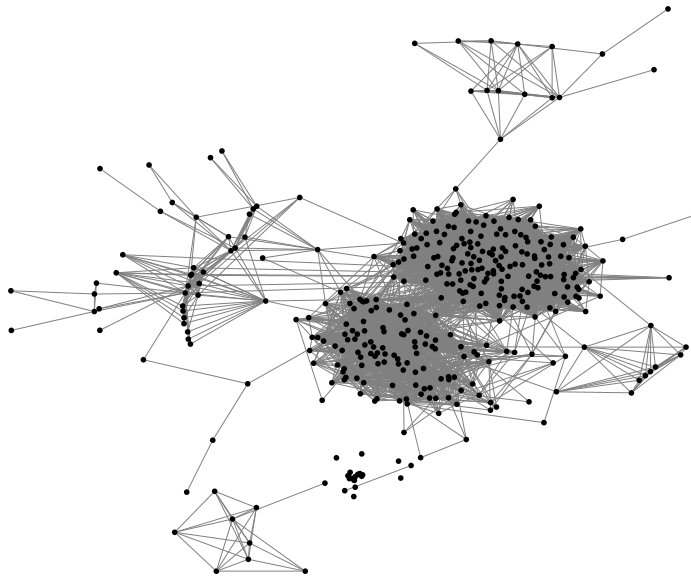


Figura 9 – A mesma representação acima porém com ajuste a partir dos gráficos pré-definidos, torna-se visível as redes de relacionamento de segundo grau dos amigos do autor fora da concentração no núcleo da rede.



## TWITTER

O Twitter concentra conversação, entretenimento e informação em uma única rede. O NodeXL recebe e interpreta esses dados de forma direta, porém esse processo é geralmente lento devido às taxas de transferência e a largura de banda da internet visto que o Twitter tem frequentes limitações de carga pois milhões de pessoas estão enviando informações para ele simultaneamente e também porque a rede cresce em larga escala de forma imprevisível.

Ao recolher informações do Twitter é necessário estar atento à limitação de dados que é imposto pelo ambiente onde se pode receber 150 requerimentos por hora, ou seja, dados de 150 pessoas por hora. A análise do ambiente do Twitter não é uma análise egocêntrica. Isso significa que é possível alcançar os amigos dos amigos dos amigos, esse montante de dados provavelmente ultrapasse o limite estipulado, o resultado é um tempo maior para receber todas as informações. O tempo médio para baixar as informações é de 30 segundos por usuário.

Figura 10 – Representação gráfica dos seguidores, em primeiro nível, do perfil @neofluxo no Twitter, - que tem a característica de não seguir nenhum outro perfil -, com destaque para o usuário @jefcarvalho.

## 7º Interprogramas de Mestrado

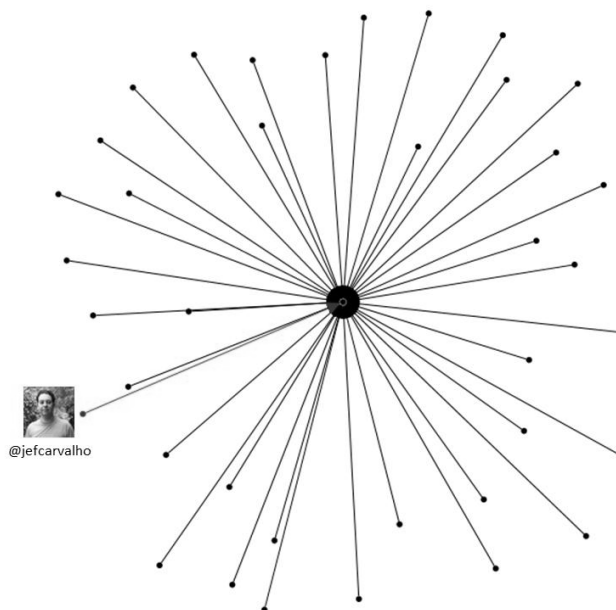
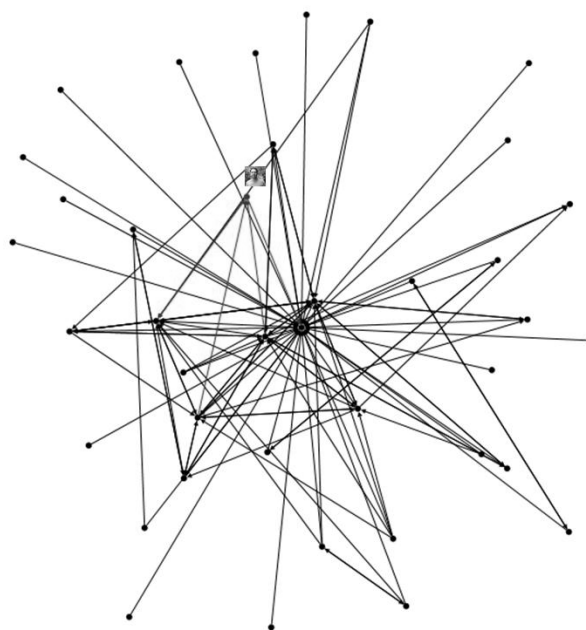
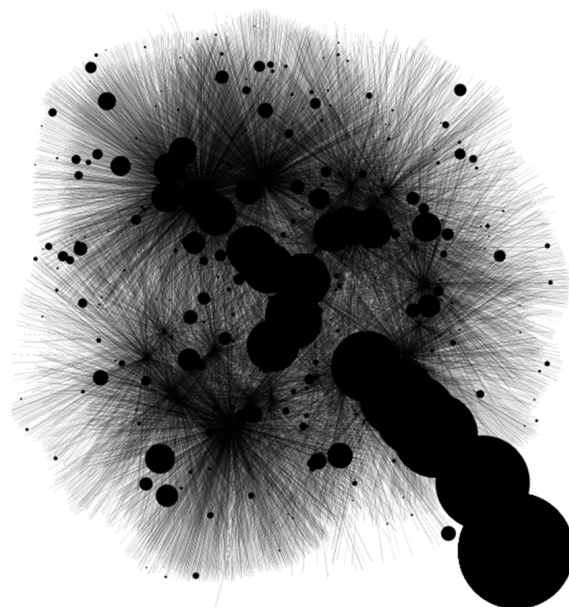


Figura 11 – Representação gráfica dos seguidores perfil @neofluxo que também seguem outros seguidores do @neofluxo no Twitter, com destaque para o usuário @jefcarvalho.



## 7º Interprogramas de Mestrado

Figura 12 – Representação gráfica dos seguidores dos seguidores dos seguidores, uma rede de relacionamento de segundo nível, do perfil @neofluxo no Twitter. A rede é densa e não é visível o destaque do usuário @jefcarvalho.





## **CONCLUSÃO**

O estudo referenciou autores no campo estimulológico de Redes Sociais como Stanley Wasserman pioneiro na aplicação computacional dos clássicos programas GRADAP e UCINET para interpretação de dados. No estudo epistemológico de Rede como Mark Newman e de Rede de Tecnologia e Informação como Jon Kleinberg. Todos enfrentaram em outra época, a mesma problemática, porém com menor volume de dados interpretáveis.

A apresentação prática dos modelos Kleinberg e Newman utilizados em estudos de “mundo pequenos” com baixa escala de atores pode ser expandida pelo intermédio da utilização da ferramenta NodeXL, aprofundando os resultados mais periféricos e menos concentrado no núcleo principal, resultando no aparecimento de nós fortes com fluxo informacional e identificação de direção.

O NodeXL sai dos grandes laboratórios e ganha a contribuição coletiva, se diferenciando frente às franquias comercializadas encontradas no mercado e utilizado por agências digitais e de publicidade.



## **BIBLIOGRAFIA**

BARABÁSI, A.-L.; **Bursts: the hidden pattern behind everything we do**, Dutton, 2010.

\_\_\_\_\_ ; **Linked: a nova ciência dos network**, Leopardo, 2009.

EASLEY, D.; KLEINBERG J. **Networks, Crowds, and Markets: Reasoning About a Highly Connected World**, Cambridge University Press, 2010.

GALLOWAY, A. **Protocol: how control exist after decentralization**. MIT Press, 2004.

HANSEN, DEREK L. **Analyzing Social Media Network with NodeXL: insights from a connected world**, Morgan Kaufman, 2010.

KNOKE, D.; YANG S. **Social network analysis – Quantitative applications in the social sciense**, SAGE, 2008.

NEWMAN, M.; **Networks, an Introduction**, Oxford University Press, 2010.

WASSERMAN, S.; FAUTS K. **Social Network Analys – Methods and Applications**. Cambridge University Press, 1994.

WU, T. **The Master Switch – the rise and fall of information empires**, Knopf Doubleday Publishing Group, 2010

Documentos eletrônicos online:

LIMA JR, **Mídia social conectada: produção colaborativa de informação de relevância social em ambiente tecnológico digital**. **Líbero**, São Paulo, Ano XII, n.24, p.95-106, 2009.

MANOVICH, L. **Software takes command**. 2008. Disponível em: [http://softwarestudies.com/softbook/manovich\\_softbook\\_11\\_20\\_2008.pdf](http://softwarestudies.com/softbook/manovich_softbook_11_20_2008.pdf). Acesso em: 22/09/2010.

NODEXL. **Networking Oveview, Discovery and Exploration for Excel**. Disponível em: <http://nodexl.codeplex.com>. Acesso em 10/09/2011.