Utilizando Métricas de *Ego-network* para Validação de Atributos dos Perfis de Usuários de Redes Sociais *On-line*

Hélder Seixas Lima, Humberto Torres Marques-Neto

Programa de Pós-Graduação em Informática
Pontifícia Universidade Católica de Minas Gerais (PUC Minas)
Belo Horizonte – MG – Brasil – 31980-110

helder.lima@sga.pucminas.br, humberto@pucminas.br

Abstract. *Online social network users identify themselves through their profiles, which are usually composed of attributes such as name, gender, age, city, among others. Since the profile attributes are self-declared, the possibility of malicious users creating accounts with false information arises. This work proposes a framework that determines a trustworthiness level for each attribute used in the profile of an online social network user. The proposed framework uses metrics in the ego-network context to verify common phenomena in social networks. The proposal was evaluated experimentally with two real samples and two synthetic samples of two social networks: Facebook and Google+. Synthetic samples simulate false users. The results showed that the trustworthiness levels determined by the framework are higher for most profile attributes of real samples when compared to those of synthetic samples.*

Resumo. *Os usuários de uma rede social on-line são identificados por meio dos seus perfis, os quais geralmente são compostos por atributos como nome, sexo, idade, cidade, entre outros. Como os atributos de perfil são autodeclarados, surge a possibilidade de que usuários mal-intencionados criem contas com informações falsas. Este trabalho propõe um framework que determina um nível de confiabilidade para cada atributo utilizado no perfil de um usuário de rede social on-line. O framework proposto utiliza métricas no contexto de ego-network para verificar fenômenos comuns nas redes sociais. A proposta foi avaliada experimentalmente com duas amostras reais e duas amostras artificiais de duas redes sociais: o Facebook e o Google+. As amostras artificiais simulam usuários falsos. Os resultados mostraram que os níveis de confiabilidade determinados pelo framework são mais elevados para a maioria dos atributos de perfil dos usuários das amostras reais quando comparados aos das amostras artificiais.*

1. Introdução

As redes sociais *on-line* fornecem um ambiente virtual para comunicação e interação entre seus usuários. Aproximadamente dois bilhões de pessoas participam de redes sociais *on-line* em todo o mundo[1], tornando este serviço um dos mais populares da *web*. Um usuário de uma rede social *on-line* é identificado por meio de um perfil autodeclarado que geralmente consiste em uma foto e informações pessoais como nome, idade, sexo, cidade,

[1]https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users

escola, profissão, empresa entre outros dados. O nível de detalhe dos perfis de usuário varia de acordo com a rede social; enquanto o Facebook, o Google+ e o LinkedIn permitem o registro de perfis detalhados, o Twitter e o Instagram permitem dados reduzidos para descrever um perfil.

A falsificação de perfil em redes sociais *on-line* é possível porque os perfis de usuários são autodeclarados e os *sites* de redes sociais não exigem a comprovação dos dados informados. Existem três tipos de perfis maliciosos: perfis falsos (pessoas que não existem no mundo real), perfis clonados (uma pessoa que finge ser outra pessoa que existe no mundo real) e perfis comprometidos (uma pessoa com intenção maliciosa que invade o perfil de outro usuário) [Soliman et al. 2016]. Algumas redes sociais implementam procedimentos para que sejam mais confiáveis, como a confirmação do número de telefone e/ou *e-mail*. Outro recurso aplicado em redes sociais *on-line* é permitir que seus usuários relatem contas suspeitas. No entanto, embora existam tais iniciativas, a validação da identidade de usuários em redes sociais *on-line* continua sendo um problema em aberto.

Este trabalho propõe um *framework* que determina níveis de confiabilidade para os atributos de perfil de um usuário de rede social *on-line*. Para isto, foi estabelecida a hipótese de que os atributos de perfil de um usuário podem ser validados com base em características presentes em sua *ego-network*, ou seja, na rede pessoal de um usuário [Leskovec and Mcauley 2012].

Experimentos foram realizados com amostras reais e artificiais de *ego-networks* de duas diferentes redes sociais *on-line* (Facebook e Google+) para avaliar o *framework* proposto. As amostras artificiais foram geradas a fim de simular usuários com atributos de perfil falsos. Os resultados obtidos mostraram que para a maioria dos usuários das amostras reais os níveis de confiabilidade retornados pelo *framework* foram mais significativos que para os usuários das amostras artificiais.

O trabalho está dividido da seguinte maneira: na Seção 2 são apresentados os pressupostos fundamentais para a compreensão da hipótese deste trabalho que é formulada na Seção 3. Na Seção 4 a hipótese estabelecida é validada. O *framework* considerado neste trabalho é apresentado na Seção 5. Os experimentos e resultados são descritos na Seção 6. Os trabalhos que já propuseram modelos para validação de usuários de redes sociais *on-line* são apresentados na Seção 7. Por fim, as conclusões e os trabalhos futuros são apresentados na Seção 8.

2. Premissas

As redes sociais do mundo real não são aleatórias, sendo que vários trabalhos já demonstraram que certos fenômenos interferem na formação deste tipo de rede [Newman 2003b]. A semelhança e a compatibilidade de características pessoais intensificam a formação de conexões entre usuários em uma rede social [Easley and Kleinberg 2010].

Homofilia representa a tendência de as pessoas estarem mais conectadas a outras pessoas com características semelhantes [McPherson et al. 2001]. Este conceito é um fenômeno frequente nas redes sociais e, portanto, objeto de vários estudos na sociologia. Esse fenômeno já foi verificado em características permanentes, como etnia, cor da pele, gênero e também em características mutáveis, como cidade de residência, profissão, empresa em que trabalha, escola frequentada, classe social, interesses, crenças e opiniões [McPherson et al. 2001, Currarini et al. 2009, Easley and Kleinberg 2010].

Vários estudos investigaram a homofilia nas redes sociais *on-line*. Por exemplo, Bhattacharyya, Garg e Wu [Bhattacharyya et al. 2011] propuseram um modelo de detecção de similaridade entre os usuários do Facebook pela análise semântica dos atributos de perfil. Os resultados mostraram que o nível de similaridade é maior em pares de amigos do que em pares de pessoas aleatórias. Mukta, Ali e Mahmud [Mukta et al. 2016] desenvolveram um modelo de identificação e validação de tipos de personalidade de usuários de redes sociais com base em traços de homofilia. Kwak et al. [Kwak et al. 2010] verificaram a ocorrência de homofilia no Twitter considerando a localização geográfica e popularidade dos usuários. Outros trabalhos verificaram o fenômeno da homofilia nas redes sociais *on-line* considerando a preferência política dos usuários [Colleoni et al. 2014, Himelboim et al. 2014, Halberstam and Knight 2014, Caetano et al. 2017].

Trabalhos de predição também consideraram o conceito de homofilia. Mislove et al. [Mislove et al. 2010] desenvolveram um método com taxa de precisão de 80% para inferir atributos ausentes em perfis de usuários usando indicadores de homofilia. Han et al. [Han et al. 2015] propuseram um método inspirado no princípio da homofilia para entender as preferências dos usuários sobre filmes, músicas, programas de televisão, entre outras. Eles concluíram que os níveis de homofilia em relação às preferências de um conjunto de usuários são maiores quando há semelhança nos atributos demográficos deles.

A formação de triângulos entre nós conectados é outra característica observada nas redes sociais. O termo fechamento triádico (*triadic closure*) define o processo no qual duas pessoas que possuem um amigo em comum também se tornam amigas em algum momento numa rede social [Easley and Kleinberg 2010]. Vários trabalhos descobriram que o início de novas amizades em uma rede social é mais provável de se concretizar quando ocorre um fechamento triádico [Rapoport 1953, Bianconi et al. 2014, Huang et al. 2015, Brandt and Leskovec 2014]. Esse fenômeno contribui para as redes sociais apresentarem formação de comunidades, em que um grupo de pessoas possui uma maior densidade de conexões entre elas e uma menor densidade de conexões com outros grupos [Newman 2003b].

Essas características de homofilia, fechamento triádico e formação de comunidades também foram observadas no contexto das *ego-networks* [Leskovec and Mcauley 2012, Wen and Yuan 2016], conceito que representa uma rede formada a partir de um nó central, chamado *ego*. Esta rede contém também os amigos do nó *ego*, chamado de *alters*. Além do nó *ego* e dos nós *alters*, as conexões entre todos os nós também integram esta rede. Em outras palavras, uma *ego-network* seria a rede de amizade de um usuário, considerando também as conexões entre os seus amigos [Leskovec and Mcauley 2012]. A Figura 1 ilustra uma *ego-network* com todos os seus componentes.

O escopo da *ego-network* é relevante para este trabalho, pois esse é o menor conjunto de dados de uma rede social *on-line* que permite caracterizar as propriedades de formação de homofilia e comunidade envolvendo um determinado usuário.

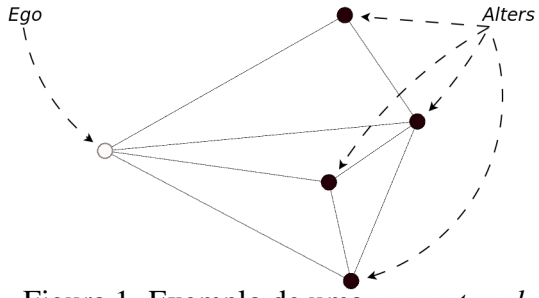


Figura 1. Exemplo de uma *ego-network*

3. Hipótese

Este trabalho considera que os atributos de perfil de um usuário de rede social *on-line* podem ser validados com base no nível de homofilia e no agrupamento dos nós em sua *ego-network*. Esta hipótese baseia-se nas premissas apresentadas na Seção 2, em que foi esclarecido que as redes sociais não são formadas aleatoriamente e são caracterizadas pela presença da homofilia e formação de comunidades.

A ideia considerada neste trabalho é que a *ego-network* de um usuário é coerente com seus atributos de perfil. Com base nisso, é esperado que, se um nó *ego* tiver um atributo específico, consequentemente ele irá se conectar a algum grupo de nós *alters* que também tenha esse mesmo atributo. Além disso, como as conexões sociais não são criadas aleatoriamente, espera-se que este grupo de nós *alters* represente uma rede caracterizada por níveis mais elevados de homofilia e agrupamento de usuários.

4. Validação da Hipótese

Esta seção apresenta a validação da hipótese descrita na Seção 3. Inicialmente, na Seção 4.1, se descrevem as métricas de *ego-network* que podem mostrar a coerência dos atributos de perfil com a *ego-network* de um usuário. Então, na Seção 4.2 se definem amostras de dados para verificar as métricas de *ego-network*. Por fim, na Seção 4.3, as métricas de *ego-network* apresentadas na Seção 4.1 são calculadas e analisadas considerando as amostras de dados apresentadas na Seção 4.2.

4.1. Métricas de *Ego-network*

Este trabalho define métricas que visam expressar a coerência de um atributo de perfil de um usuário de rede social *on-line* com sua *ego-network*. O primeiro passo para calcular essas métricas é encontrar o conjunto de nós *alters* que possui o mesmo atributo de perfil do nó *ego* a ser verificado. Neste trabalho, este conjunto é referenciado como S_i, onde i indica o atributo de perfil verificado. A seguir são listadas as métricas definidas:

- n_i: $|S_i|$, ou seja, o número de nós *alters* que possuem i;
- h_i: coeficiente de homofilia calculado para i considerando uma rede composta por todos nós *alters*;
- g_i: coeficiente de agrupamento médio para uma rede composta por nós de S_i.

Neste trabalho, h_i corresponde ao coeficiente de homofilia proposto por Newman [Newman 2003a], também chamado de coeficiente de assortatividade. Assim, h_i admite valores entre -1 e 1, em que 1 representa uma rede em que todas as conexões são formadas entre nós que possuem i e -1 representa uma rede em que todas as arestas são formadas entre nós onde apenas um deles possui i. Há homofilia quando h_i é maior que zero.

O coeficiente de agrupamento (*clustering coefficient*) é uma medida usada para verificar o grau de nós conectados em uma rede. O cálculo do coeficiente de agrupamento de um nó é equivalente ao número de arestas existentes entre seus vizinhos dividido pelo número máximo de arestas que poderia ser criado entre eles [Newman 2003b]. Esta medida está estritamente ligada ao conceito de fechamento triádico, considerando que quanto maior o número de triângulos conectados, maior será o resultado do coeficiente de agrupamento.

A métrica g_i determina o coeficiente de agrupamento médio de uma rede formada exclusivamente pelos nós que compõem S_i e suas arestas. A Equação 1 apresenta o cálculo para a métrica g_i, onde f representa a função que calcula o coeficiente de agrupamento de um determinado nó.

$$g_{\mathrm{i}} = \frac{1}{n_{\mathrm{i}}} \sum_{j=1}^{n_{\mathrm{i}}} f(S_{\mathrm{i}}[j]) \tag{1}$$

4.2. Amostras de Dados

Este trabalho utilizou duas amostras reais de redes sociais *on-line*, ambas coletadas por Leskovec e Mcauley [Leskovec and Mcauley 2012]. As amostras em questão são:

- *Amostra Real A*: consiste em uma amostra do Facebook com 4.039 perfis de usuários e 176.468 conexões entre os usuários. Os atributos de perfil considerados foram escola, sexo, cidade natal, cidade de residência, empresa e faixa etária. Os dados foram coletados a partir de 10 *ego-networks*. Esta amostra contém 37.257 atributos de perfil, sendo 62 atributos de perfil dos nós *egos*.
- *Amostra Real B*: consiste em uma amostra do Google+ com 107.614 perfis de usuários e 13.673.453 conexões entre os usuários, sendo que 2.865.008 são conexões com reciprocidade entre os usuários. Os atributos de perfil considerados foram sexo, cidade de residência, empresa e escola. Os dados foram coletados a partir de 132 *ego-networks*. Esta amostra contém 387.245 atributos de perfil, sendo 456 atributos de perfil dos nós *egos*.

As redes sociais Facebook e Google+ possibilitam que seus usuários cadastrem perfis com informações detalhadas. Desta forma, esse tipo de rede social é viável para a verificação da hipótese apresentada na Seção 3; e por isso, amostras de *ego-networks* dessas redes sociais foram consideradas neste trabalho. A rede social LinkedIn também possibilita que seus usuários cadastrem perfis com informações detalhadas. Entretanto, nenhuma amostra pública dessa rede social foi encontrada e também não é permitido coletar os dados necessários para análise da hipótese apresentada na Seção 3, conforme termos de uso dessa rede social.

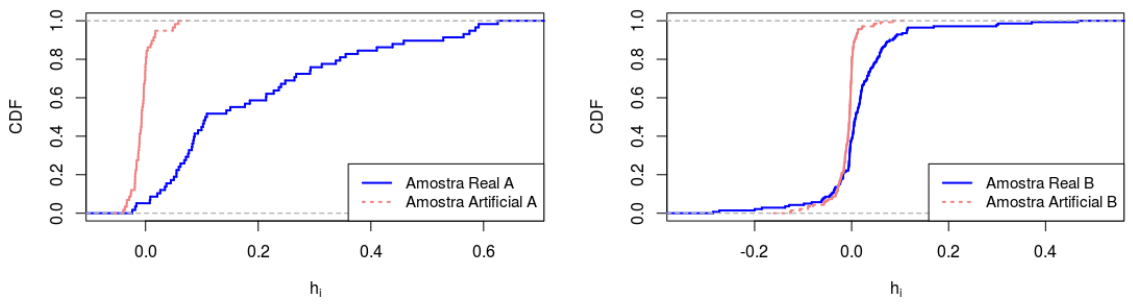
Grafos geralmente são utilizados para representar a estrutura de uma rede social, sendo que os usuários correspondem aos nós e as conexões entre os usuários correspondem às arestas. O grafo de uma rede social pode ser dirigido ou não dirigido. Por exemplo, o grafo do Facebook não é dirigido, porque esta rede social contém apenas conexões recíprocas entre os usuários. O grafo do Google+, por outro lado, consiste em uma rede que exibe conexões dirigidas entre os usuários. Neste trabalho apenas as conexões recíprocas entre os usuários são consideradas, pois esse tipo de relação indica um maior grau de proximidade entre os envolvidos [Colleoni et al. 2014].

Além das duas amostras reais, neste trabalho também foram geradas sinteticamente outras duas amostras chamadas de *Amostra Artificial A* e *Amostra Artificial B*. As amostras reais A e B foram a base para a construção das amostras artificiais A e B, respectivamente. Cada uma dessas amostras artificiais corresponde às amostras reais em relação ao número de *ego-networks*, usuários, conexões e estrutura da rede; a diferença é que, nas amostras artificiais, os atributos de perfil dos nós *alters* de cada *ego-network* foram distribuídos aleatoriamente. Essas amostras artificiais são relevantes para validar a hipótese porque permite confrontar as métricas das amostras reais com as métricas de *ego-networks* geradas aleatoriamente. A utilização de amostras sintéticas aleatórias é prática recorrente em trabalhos de validação de usuários de redes sociais *online* [Laleh et al. 2017, Soliman et al. 2016, Tran et al. 2011, Mulamba et al. 2016]

4.3. Análise das Métricas de *Ego-network*

Nesta seção são analisadas as métricas de *ego-network*. Inicialmente, foi definido que os atributos de perfil a serem avaliados seriam restritos àqueles com n_i maior ou igual a 3. Esse valor de n_i corresponde à menor quantidade de nós que permite a formação de um triângulo, sendo isso essencial para avaliar os valores de h_i e g_i. Dessa forma, 94% dos atributos de perfil das amostras do Facebook foram consideradas enquanto que apenas 31% dos atributos de perfil das amostras Google+ atenderam a esse requisito. O valor de n_i já demonstra que os usuários do Facebook, que foram considerados neste trabalho, apresentam *ego-networks* com maior similaridade entre nós *egos* e nós *alters* do que entre os usuários do Google+.

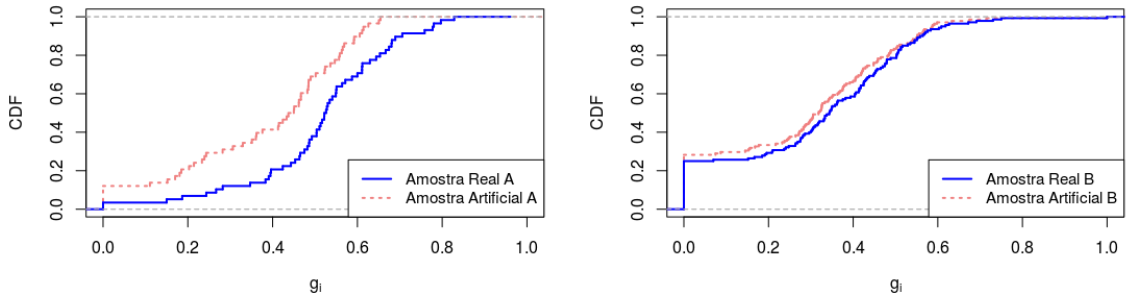
A Figura 2 apresenta a função de distribuição cumulativa (CDF) de h_i. Nota-se maior nível de homofilia nas amostras reais que nas amostras artificiais. No caso do Facebook (Figura 2a) é possível constatar grande diferença entre a amostra real e a amostra artificial. Por outro lado, no caso do Google+ (Figura 2b), os valores de h_i entre a amostra real e amostra artificial são mais aproximados, mas, ainda assim, a amostra real alcançou valores maiores, sendo que a média de h_i na *Amostra Real B* foi 0,02 e a média de h_i na *Amostra Artificial B* foi -0,01.



(a) Distribuição de h_i nas amostras do Facebook (b) Distribuição de h_i nas amostras do Google+

Figura 2. Distribuição de h_i nas amostras de dados

A Figura 3 apresenta CDF em relação à métrica g_i das amostras consideradas neste trabalho. A amostra real do Facebook alcança g_i maiores em relação à amostra artificial. Todavia, as amostras real e artificial do Google+ apresentam valores de g_i parecidos. O valor médio de g_i na *Amostra Real B* foi 0,31, pouco maior que o valor médio de g_i na *Amostra Artificial B*, que foi 0,28.



(a) Distribuição de g_i nas amostras do Facebook (b) Distribuição de g_i nas amostras do Google+

Figura 3. Distribuição de g_i nas amostras de dados

Cada rede social apresentou características específicas em relação aos níveis de homofilia e agrupamento considerando as amostras de dados analisadas neste trabalho. Para a amostra de dados do Facebook a hipótese de que as *ego-networks* não são aleatórias se mostrou verdadeira, pois foi apresentado que em geral os valores de h_i e g_i na amostra real são maiores que na amostra artificial.

Os valores de h_i e g_i para as amostras do Google+ mostraram que a diferença dos dados reais para os dados aleatórios é menor quando comparada com os resultados das amostras do Facebook. Características próprias de cada rede social *on-line* podem influenciar o nível de homofilia e agrupamento. Outro fator que pode ter influenciado nos valores das métricas é a densidade das amostras, sendo que a amostra real do Facebook possui uma média de 44 conexões por usuário, enquanto que a amostra real do Google+ possui apenas 27 conexões por usuário em média.

5. Framework para Validação de Atributos de Perfil

Esta seção apresenta um *framework* que define o nível de confiabilidade t_i para cada atributo de perfil i de um usuário de rede social *on-line*. Este *framework* é baseado nas métricas n_i, h_i e g_i, apresentadas na Seção 4.1. O cálculo de t_i está relacionado aos valores de desvio padrão e média dessas métricas, além dos pesos das métricas que são estabelecidos no Módulo 1 deste *framework* (Seção 5.1). O segundo e último módulo deste *framework* (Seção 5.2) é em que se obtêm os valores de t_i. Neste *framework*, t_i pode assumir valores de 0 até 1; a ideia é que quanto mais próximo de 1 maior é a confiabilidade de um atributo de perfil ser verdadeiro; e quanto mais próximo de 0 menor é a confiabilidade de um atributo de perfil ser verdadeiro.

Este *framework* foi proposto visando permitir flexibilização na sua aplicação, por isso, nesta seção alguns procedimentos são apresentados apenas conceitualmente. Opções de implementação são detalhadas na Seção 6, em que o *framework* foi experimentalmente aplicado.

5.1. Módulo 1: Preparação

O primeiro módulo consiste em calcular os valores de referência que serão utilizados no módulo seguinte. Para aplicar este *framework* é necessário selecionar uma amostra de *ego-networks* de usuários reais, na qual os atributos de perfil selecionados dos usuários

são reconhecidamente verdadeiros. Também é necessário selecionar uma amostra de *ego-networks* de usuários falsos. Essas duas amostras consistem nos parâmetros de entrada do Módulo 1, conforme ilustrado na Figura 4.

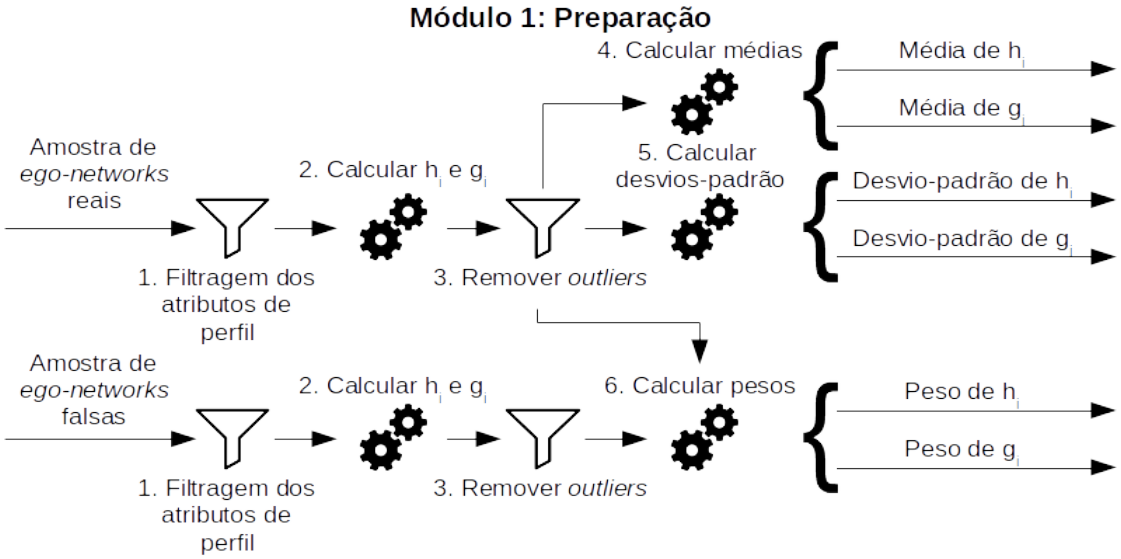


Figura 4. Primeiro módulo do *framework*

A primeira etapa deste módulo consiste em filtrar os atributos de perfil dos nós *egos*. Somente os atributos de perfil com n_i maiores ou iguais a 3 são considerados devido aos mesmos motivos mencionados na Seção 4.3. Em seguida, na etapa 2, as métricas h_i e g_i são calculadas para os atributos de perfil dos nós *egos*. Na terceira etapa, os valores *outliers* devem ser removidos para reduzir as distorções causadas por atributos de perfil que possuem métricas com comportamento extremamente fora dos padrões das amostras. Todas essas três etapas iniciais são executadas paralelamente na amostra de usuários reais e na amostra de usuários falsos.

Na quarta etapa deste módulo são calculadas as médias das métricas h_i e g_i considerando a amostra de usuários reais. Na penúltima etapa realizam-se os cálculos dos desvios-padrão para as métricas h_i e g_i também considerando apenas a amostra de usuários reais. Por fim, a última etapa deste módulo consiste em calcular os pesos atribuídos a cada uma das métricas. Esses pesos devem refletir a importância de cada métrica para determinar o valor de t_i para um atributo de perfil i. O cálculo desses pesos consideram tanto a amostra real como a amostra falsa de *ego-networks*. A soma dos pesos deve ser igual a 1 por motivo de normalização.

5.2. Módulo 2: Cálculo do Nível de Confiabilidade de Atributos de Perfil

O segundo módulo deste *framework* consiste no cálculo dos valores de t_i para cada atributo de perfil de um usuário. Conforme ilustrado na Figura 5, os parâmetros recebidos neste módulo são a *ego-network* de um usuário e os valores resultantes do Módulo 1.

Para cada atributo de perfil i de um nó *ego* é calculado o valor de t_i correspondente. Para isso, o primeiro passo consiste em calcular os valores de h_i e g_i para os atributos de perfil do nó *ego* em avaliação. Depois, uma nota entre 0 e 1 será definida para cada uma das métricas h_i e g_i. O Algoritmo 1 calcula uma nota para uma métrica, em que se recebe como parâmetro de entrada o valor da métrica em questão. Outros dois parâmetros de

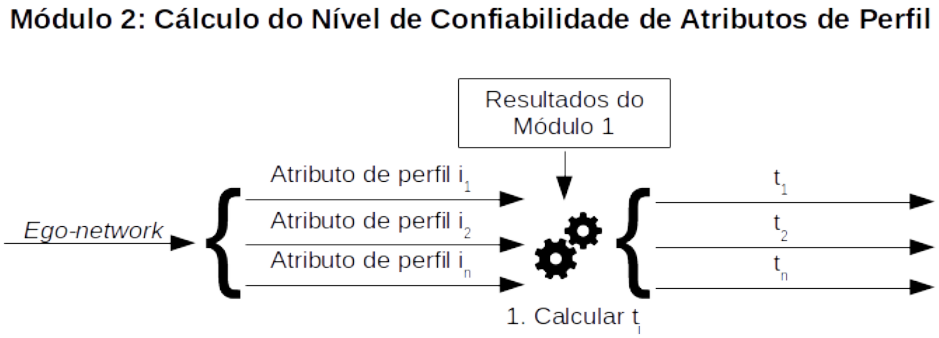


Figura 5. Segundo módulo do *framework*

entrada do Algoritmo 1 correspondem à média e ao desvio-padrão relacionados a essa métrica, ambos calculados no Módulo 1. A nota é definida proporcionalmente dentro de um intervalo de valores, em que os limites inferior e superior são definidos, respectivamente, nas linhas 1 e 2 do algoritmo.

Algoritmo 1 Cálculo da nota de uma métrica de *ego-network*

Entrada: valor v, média avg e desvio-padrão sd de uma métrica de *ego-network*
Saída: nota de uma métrica de *ego-network*

1: $bottom \leftarrow avg - sd$
2: $top \leftarrow avg + sd$
3: $amplitude \leftarrow top - bottom$
4: **se** $v < bottom$ **então retorna** 0
5: **fim se**
6: **se** $v > top$ **então retorna** 1
7: **fim se**
8: **retorna** $(v - bottom)/amplitude$

O passo final para calcular t_i de um atributo de perfil consiste em ponderar as notas das métricas h_i e g_i com seus respectivos pesos que foram definidos no Módulo 1 deste *framework*. O valor de t_i de um atributo de perfil é dado pela Equação 2, onde sh_i e sg_i são as notas das métricas h_i e g_i, respectivamente; e wh e wg correspondem aos pesos de h_i e g_i, respectivamente.

$$t_{\mathrm{i}} = (sh_{\mathrm{i}} * wh) + (sg_{\mathrm{i}} * wg) \tag{2}$$

6. Experimentos e Resultados

Nesta seção, apresenta-se os experimentos realizados no *framework* proposto. Foram realizados experimentos para as redes sociais Facebook e Google+, sendo que foram utilizadas as mesmas amostras de *ego-networks* apresentadas na Seção 4.2. Desta forma, a *Amostra Real A* e a *Amostra Real B* correspondem ao parâmetro de entrada denominado de *amostra de ego-networks reais* do Módulo 1 do *framework* para o Facebook e Google+, respectivamente. Ainda em relação aos parâmetros de entrada do Módulo 1 do *framework*, foi considerado que o parâmetro denominado de *amostra de ego-networks falsas* corresponde à *Amostra Artificial A* e *Amostra Artificial B* para o Facebook e Google+, respectivamente.

As etapas 1 e 2 do Módulo 1 do *framework* foram realziadas conforme especificado na Seção 5.1. Na etapa 3 foram utilizados diagramas *boxplot* para inspecionar e remover *outliers* em relação às métricas h_i e g_i das amostras. Os valores de quartil 1 *Q1*, quartil 3 *Q3* e o intervalo interquartil *IQR* foram usados para calcular os limites inferiores e superiores para remoção de *outliers*. Foi considerado o limite inferior igual a $Q1 - (1,5 * IQR)$ e o limite superior igual a $Q3 + (1,5 * IQR)$ [Silva et al. 2016].

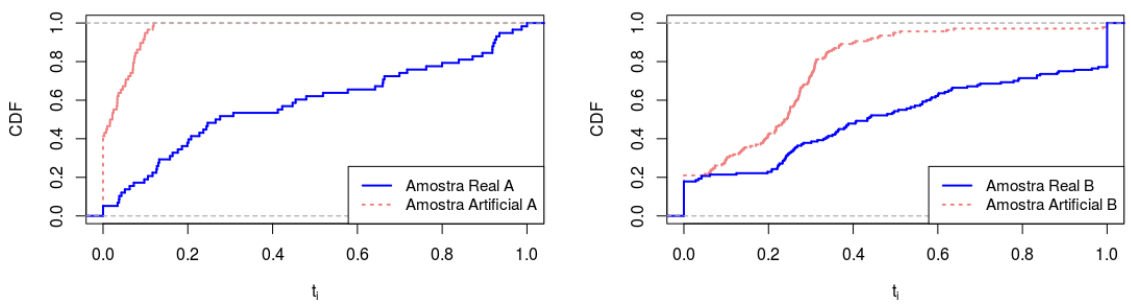
Após a remoção de *outliers* das amostras foram executadas as etapas 4, 5 e 6 do Módulo 1 do *framework*. Na etapa 6, que é a etapa em que se calculam os pesos das métricas, foi utilizado o algoritmo *RandomForests* [Breiman 2001] que tem como função identificar a importância das *features* (h_i e g_i no contexto deste trabalho) dentro de uma amostra com dados classificados.

A Tabela 1 apresenta os resultados do primeiro módulo do *framework* proposto. Para a amostra do Facebook a métrica h_i apresentou maior peso que a métrica g_i. No caso da amostra do Google+ foi constatado que o valor de g_i é indiferente para distinguir atributos de perfil reais de atributos de perfil falsos; desta forma, a métrica h_i assumiu peso 1.

Tabela 1. Resultados do Módulo 1

Métrica		Facebook	Google+
h_i	Média	0,209	0,013
	Desvio-padrão	0,180	0,033
	Peso	0,869	1,000
g_i	Média	0,548	0,296
	Desvio-padrão	0,130	0,209
	Peso	0,131	0,000

Os resultados apresentados na Tabela 1 consistem nos parâmetros de entrada do Módulo 2 do *framework* proposto. Após executar o Módulo 2 foi obtido valor de t_i para cada um dos atributos de perfil considerados neste trabalho. Isso consiste no nível de confiabilidade que o *framework* atribui a cada atributo de perfil analisado. A Figura 6 exibe gráficos da CDF de t_i nas amostras consideradas neste trabalho. Constata-se que, tanto no Facebook quanto no Google+, os níveis de confiabilidade são maiores nas amostras reais.



(a) Distribuição de t_i nas amostras do Facebook (b) Distribuição de t_i nas amostras do Google+

Figura 6. Distribuição de t_i nas amostras de dados

Os atributos de perfil analisados do Facebook apresentaram melhores resultados do que o Google+. Foi constatado que o maior t_i obtido por um atributo de perfil da amostra artificial do Facebook foi igual a 0,1192. Por outro lado, 78% dos atributos de perfil da amostra real do Facebook alcançaram valores maiores que 0,1192. Esse resultado indica que o *framework* proposto permite uma clara distinção entre atributos de perfil reais e falsos para a maioria dos casos nas amostras do Facebook.

A Figura 6b, mostra que no Google+, atributos de perfil falsos tiveram t_i tão elevados quanto os atributos de perfil da amostra real, sendo que houve atributo de perfil falso que obteve o valor máximo de t_i, que é igual a 1. Entretanto, numa análise minuciosa, verifica-se que 75% da amostra falsa atingiu valor t_i até 0,3052 e que 61% da amostra real atingiu valores superiores a 0,3052. Desta forma, os resultados indicam que o *framework* também permite distinguir atributos de perfil reais e falsos nas amostras do Google+ na maioria dos casos.

7. Trabalhos Relacionados

Vários trabalhos foram desenvolvidos visando combater o problema do crescente número de usuários mal intencionados nas redes sociais *on-line*. Destacam-se dois tipos de trabalho: (i) *detecção de contas falsas* e (ii) *validação da identidade de usuário*. Os trabalhos de *detecção de contas falsas*, em geral, procuram classificar os usuários como honestos ou falsos em uma larga escala de dados; enquanto os trabalhos de *validação da identidade de usuário* consideram uma avaliação individual do usuário para estimar o grau de confiabilidade das suas informações autodeclaradas [Bahri et al. 2016].

Nota-se na literatura que a maioria dos trabalhos sobre a detecção de conta falsa são focados no combate aos ataques *sybils*, um tipo de fraude em que vários perfis falsos são criados e controlados por uma única pessoa [Douceur 2002]. A tática considerada por esse tipo de ataque é tornar os perfis falsos mais parecidos com contas reais, a fim de potencializar práticas indesejáveis nas redes sociais *on-line* como *spams*, notícias falsas, manipulação de opinião e aplicação de golpes.

Os principais modelos propostos de detecção de ataques *sybils* são SybilGuard [Yu et al. 2008], SybilLimit [Yu et al. 2010], SybilInfer [Danezis and Mittal 2009] e SybilDefender [Wei et al. 2013]. Em geral, esses trabalhos assumem que os usuários honestos tendem a formar uma rede *fast-mixing*, ou seja, uma rede na qual os nós rapidamente formam um grafo densamente conectado. Esses trabalhos assumem que os perfis falsos utilizados neste tipo de ataque não possuem essa propriedade, pois, considera-se que é mais difícil contas falsas consolidarem amizades com usuários reais. Desta forma, isso pressupõe que os grafos das redes sociais formam regiões distintas entre usuários honestos e falsos.

Entre os trabalhos que abordam a *validação da identidade de usuário*, é comum a utilização de modelos de votação (*feedback* dos usuários da rede) e filtros coletivos para definir o nível de confiabilidade de um usuário [Cai et al. 2010, Sirivianos et al. 2014]. Também destacam-se trabalhos que utilizam técnicas de aprendizagem para identificar correlações de atributos de perfil nas redes sociais *on-line* e com isso determinar o nível de confiabilidade de um usuário [Bahri et al. 2014, Soliman et al. 2016, Bahri et al. 2016].

Diferentemente desses trabalhos, este artigo propõe um *framework* que determine o nível de confiabilidade de cada atributo de perfil autodeclarado por um usuário de

rede social *on-line*. A utilidade do *framework* proposto consiste em auxiliar os usuários a interagirem de forma mais segura com pessoas desconhecidas em um ambiente de rede social *on-line*. Outro diferencial deste trabalho consiste na utilização de métricas de *ego-network* para determinar os níveis de confiabilidade dos atributos de perfil de usuários de redes sociais *on-line*.

8. Conclusões

Neste trabalho, foi investigada a validação de atributos de perfil de usuário de rede social *on-line*. A hipótese de que os atributos de perfil de um usuário real são coerentes com sua *ego-network* foi estabelecida, assumindo que as redes sociais não são formadas aleatoriamente. Para isso, métricas relacionadas à homofilia e ao agrupamento dentro do contexto de uma *ego-network* foram definidas e calculadas para amostras reais e artificiais do Facebook e Google+. As amostras artificiais simularam usuários com atributos de perfil falsos. A análise dessas métricas indicaram que para a maioria dos casos os valores das métricas para amostras reais são maiores que nas amostras artificiais.

Com isso, um *framework* que indique o nível de confiabilidade dos atributos de perfil de usuários de redes sociais *on-line* foi proposto. Os experimentos realizados mostraram que as amostras reais obtiveram níveis de confiabilidade maiores que as amostras artificiais para a maioria dos atributos de perfil. Também foi constatado que a amostra de dados do Facebook proporcionou melhores resultados que a amostra de dados do Google+, revelando que a eficiência do *framework* proposto varia de acordo com as características de homofilia e agrupamento de usuários da rede social em análise.

Pretende-se estender este trabalho em algumas direções. Um trabalho futuro seria realizar novos experimentos em amostras maiores de redes sociais. Outro trabalho futuro consistiria em aperfeiçoar o *framework* proposto para que considere diferentes padrões de *ego-networks* e usuários e também considere outras métricas, como, por exemplo, informações do comportamento de publicação de conteúdos. Por fim, também se deseja especificar um protocolo de autorização baseado no OAuth 2.0 que incorpore o *framework* proposto. A ideia é que um sistema terceiro que utiliza dados autorizados por usuários de redes sociais *on-line* tenha informações da confiabilidade dos dados recebidos.

Agradecimentos

Este trabalho foi financiado pela CAPES, FAPEMIG e CNPq.

Referências

Bahri, L., Carminati, B., and Ferrari, E. (2014). Community-based identity validation on online social networks. In *Distributed Computing Systems (ICDCS), 2014 IEEE 34th International Conference on*, pages 21–30. IEEE.

Bahri, L., Carminati, B., and Ferrari, E. (2016). Coip—continuous, operable, impartial, and privacy-aware identity validity estimation for osn profiles. *ACM Transactions on the Web (TWEB)*, 10(4):23.

Bhattacharyya, P., Garg, A., and Wu, S. F. (2011). Analysis of user keyword similarity in online social networks. *Social network analysis and mining*, 1(3):143–158.

Bianconi, G., Darst, R. K., Iacovacci, J., and Fortunato, S. (2014). Triadic closure as a basic generating mechanism of communities in complex networks. *Physical Review E*, 90(4):042806.

Brandt, C. and Leskovec, J. (2014). Status and friendship: Mechanisms of social network evolution. In *Proceedings of the 23rd International Conference on World Wide Web*, WWW '14 Companion, pages 229–230, New York, NY, USA. ACM.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

Caetano, J., Lima, H., Santos, M., and Marques-Neto, H. (2017). Utilizando análise de sentimentos para definição da homofilia política dos usuários do twitter durante a eleição presidencial americana de 2016. In *Proceedings of the 6th Brazilian Workshop on Social Network Analysis and Mining*.

Cai, X., Bain, M., Krzywicki, A., Wobcke, W., Kim, Y. S., Compton, P., and Mahidadia, A. (2010). Collaborative filtering for people to people recommendation in social networks. In *Australasian Joint Conference on Artificial Intelligence*, pages 476–485. Springer.

Colleoni, E., Rozza, A., and Arvidsson, A. (2014). Echo chamber or public sphere? predicting political orientation and measuring political homophily in twitter using big data. *Journal of Communication*, 64(2):317–332.

Currarini, S., Jackson, M. O., and Pin, P. (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, 77(4):1003–1045.

Danezis, G. and Mittal, P. (2009). Sybilinfer: Detecting sybil nodes using social networks. In *NDSS*. San Diego, CA.

Douceur, J. R. (2002). The sybil attack. In *International Workshop on Peer-to-Peer Systems*, pages 251–260. Springer.

Easley, D. and Kleinberg, J. (2010). *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge University Press.

Halberstam, Y. and Knight, B. (2014). Homophily, group size, and the diffusion of political information in social networks: Evidence from twitter. Technical report, National Bureau of Economic Research.

Han, X., Wang, L., Crespi, N., Park, S., and Cuevas, Á. (2015). Alike people, alike interests? inferring interest similarity in online social networks. *Decision Support Systems*, 69:92–106.

Himelboim, I., Sweetser, K. D., Tinkham, S. F., Cameron, K., Danelo, M., and West, K. (2014). Valence-based homophily on twitter: network analysis of emotions and political talk in the 2012 presidential election. *new media & society*, page 1461444814555096.

Huang, H., Tang, J., Liu, L., Luo, J., and Fu, X. (2015). Triadic closure pattern analysis and prediction in social networks. *IEEE Transactions on Knowledge and Data Engineering*, 27(12):3374–3389.

Kwak, H., Lee, C., Park, H., and Moon, S. (2010). What is twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web*, pages 591–600. ACM.

Laleh, N., Carminati, B., and Ferrari, E. (2017). Risk assessment in social networks based on user anomalous behaviour. *IEEE Transactions on Dependable and Secure Computing*, PP(99):1–1.

Leskovec, J. and Mcauley, J. J. (2012). Learning to discover social circles in ego networks. In *Advances in neural information processing systems*, pages 539–547.

McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology*, pages 415–444.

Mislove, A., Viswanath, B., Gummadi, K. P., and Druschel, P. (2010). You are who you know: inferring user profiles in online social networks. In *Proceedings of the third ACM international conference on Web search and data mining*, pages 251–260. ACM.

Mukta, M. S. H., Ali, M. E., and Mahmud, J. (2016). Identifying and validating personality traits-based homophilies for an egocentric network. *Social Network Analysis and Mining*, 6(1):74.

Mulamba, D., Ray, I., and Ray, I. (2016). *SybilRadar: A Graph-Structure Based Framework for Sybil Detection in On-line Social Networks*, pages 179–193. Springer International Publishing, Cham.

Newman, M. E. (2003a). Mixing patterns in networks. *Physical Review E*, 67(2):026126.

Newman, M. E. (2003b). The structure and function of complex networks. *SIAM review*, 45(2):167–256.

Rapoport, A. (1953). Spread of information through a population with socio-structural bias: I. assumption of transitivity. *Bulletin of Mathematical Biology*, 15(4):523–533.

Silva, L. A. d., Peres, S. M., and Boscarioli, C. (2016). *Introdução à mineração de dados: com aplicações em R*. Elsevier.

Sirivianos, M., Kim, K., Gan, J. W., and Yang, X. (2014). Leveraging social feedback to verify online identity claims. *ACM Transactions on the Web (TWEB)*, 8(2):9.

Soliman, A., Bahri, L., Girdzijauskas, S., Carminati, B., and Ferrari, E. (2016). Cadiva: cooperative and adaptive decentralized identity validation model for social networks. *Social Network Analysis and Mining*, 6(1):1–22.

Tran, N., Li, J., Subramanian, L., and Chow, S. S. M. (2011). Optimal sybil-resilient node admission control. In *2011 Proceedings IEEE INFOCOM*, pages 3218–3226.

Wei, W., Xu, F., Tan, C. C., and Li, Q. (2013). Sybildefender: A defense mechanism for sybil attacks in large social networks. *IEEE Transactions on Parallel and Distributed Systems*, 24(12):2492–2502.

Wen, J. and Yuan, Q. (2016). Social circles discovery based on structural and attribute similarities. In *2016 IEEE Trustcom/BigDataSE/ISPA*, pages 1652–1659.

Yu, H., Gibbons, P. B., Kaminsky, M., and Xiao, F. (2010). Sybillimit: A near-optimal social network defense against sybil attacks. *IEEE/ACM Transactions on Networking*, 18(3):885–898.

Yu, H., Kaminsky, M., Gibbons, P. B., and Flaxman, A. D. (2008). Sybilguard: Defending against sybil attacks via social networks. *IEEE/ACM Transactions on Networking*, 16(3):576–589.