UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"
Faculdade de Filosofia e Ciências - Campus de Marília
Departamento de Pós-Graduação em Ciência da Informação

JOSÉ ANTONIO MAURILIO MILAGRE DE OLIVEIRA

INTELIGÊNCIA CIBERNÉTICA E USO DE RECURSOS SEMÂNTICOS NA DETECÇÃO DE PERFIS FALSOS NO CONTEXTO DO BIG DATA

Marília
2016

JOSÉ ANTONIO MAURILIO MILAGRE DE OLIVEIRA

INTELIGÊNCIA CIBERNÉTICA E USO DE RECURSOS SEMÂNTICOS NA DETECÇÃO DE PERFIS FALSOS NO CONTEXTO DO BIG DATA

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, Campus de Marília, como exigência parcial para a obtenção do título de Mestre em Ciência da Informação.

Orientador: Dr. José Eduardo Santarem Segundo

Marília
2016

Oliveira, José Antonio Maurilio Milagre de.
O48m Inteligência cibernética e uso de recursos semânticos na detecção de perfis falsos no contexto do Big Data / José Antonio Maurilio Milagre de Oliveira. – Marília, 2016.
130 f. ; 30 cm.

Orientador: José Eduardo Santarem Segundo.
Dissertação (Mestrado em Ciência da Informação) – Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, 2016.
Bibliografia: f. 122-130

1. Crime por computador. 2. Cibernética. 3. Big data. 4. Web semântica. I. Título.

CDD 004.67

JOSÉ ANTONIO MAURILIO MILAGRE DE OLIVEIRA

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, Campus de Marília, como exigência parcial para a obtenção do título de Mestre em Ciência da Informação.

Área de concentração: Informação, Tecnologia e Conhecimento.

Linha de Pesquisa: Informação e Tecnologia.

Data de aprovação: 29/04/2016

Comissão examinadora:

_____
Dr. José Eduardo Santarem Segundo (Orientador)
Universidade de São Paulo – USP/ Faculdade de Filosofia e Ciências – UNESP.

_____
Dr. Ricardo César Gonçalves Sant'Ana
Universidade Estadual Paulista, Faculdade de Ciências e Engenharia, Tupã.

_____
Dr. Mário Furlaneto Neto
Centro Universitário Eurípides de Marília – UNIVEM.

Marília
2016

Dedicatória

A DEUS, em primeiro lugar. À minha mãe Maria Elisa e à minha filha Stephanie. Dedico também à amada Jaline Gilioti. À minha família por todo o incentivo e suporte.

Agradecimentos

Ao Professor Doutor José Eduardo Santarém Segundo, por acreditar em nosso projeto, por nos acompanhar a cada minuto de nossa pesquisa e por todo o apoio e motivação, sem os quais jamais chegaríamos até aqui.

Agradecimentos especiais aos professores Doutor Ricardo César Gonçalves Santana e Doutor Mário Furlaneto Neto pelas importantes orientações fornecidas no exame de qualificação e por integrarem a banca examinadora desta pesquisa.

Resumo

O desenvolvimento da Internet transformou o mundo virtual em um repositório infindável de informações. Diariamente, na sociedade da informação, pessoas interagem, capturam e despejam dados nas mais diversas ferramentas de redes sociais e ambientes da *Web*. Estamos diante do Big Data, uma quantidade inacabável de dados com valor inestimável, porém de difícil tratamento. Não se tem dimensão da quantidade de informação capaz de ser extraída destes grandes repositórios de dados na *Web*. Um dos grandes desafios atuais na Internet do "Big Data" é lidar com falsidades e perfis falsos em ferramentas sociais, que causam alardes, comoções e danos financeiros significativos em todo o mundo. A inteligência cibernética e computação forense objetivam investigar eventos e constatar informações extraindo dados da rede. Por sua vez, a Ciência da Informação, preocupada com as questões envolvendo a recuperação, tratamento, interpretação e apresentação da informação, dispõe de elementos que quando aplicados neste contexto podem aprimorar processos de coleta e tratamento de grandes volumes de dados, na detecção de perfis falsos. Assim, por meio da presente pesquisa de revisão de literatura, documental e exploratória, buscou-se revisar os estudos internacionais envolvendo a detecção de perfis falsos em redes sociais, investigando técnicas e tecnologias aplicadas e principalmente, suas limitações. Igualmente, apresenta-se no presente trabalho contribuições de áreas da Ciência da Informação e critérios para a construção de ferramentas que se destinem à identificação de perfis falsos, por meio da apresentação de uma proposta de modelo conceitual. Identificou-se, na pesquisa, que a Ciência da Informação pode contribuir com a construção de aplicações e *frameworks* para que usuários possam identificar e discernir perfis reais de perfis questionáveis, diariamente despejados na *Web*.

Palavras-chave: Computação Forense. Inteligência Cibernética. Big Data. *Web* Semântica. Perfis falsos.

Abstract

The development of the Internet changed the virtual world in an endless repository of information. Every single day, in an information-society, people change, catch and turn out files in different tools of social network and *Web* surrounding. We are in front of “The Big Data”, an endless amount of data with invaluable, but hard treating. It doesn’t have some dimension of measure information to be able of extracting from these big *Web* data repositories. One of the most challenges nowadays on the Internet from the “Big Data” is to identify feelings, anticipating sceneries dealing with falsehood and false profiles social tools, which cause fanfare, upheavals and significant financial losses worldwide in front of our true scene. The cyber intelligence has by objective to look for events and finding information, subtracting dates from the *Web*. On the other hand, the Information Science, worried with the questions involving recovery, processing, interpretation and presentation of information that has important areas of study capable of being applied in this context hone the collection and treatment processes of large volumes of information (datas). Thus, through this research literature review, documentary and exploratory, the researcher aimed to review the International studies implicating the analysis of large volumes of data on social networking tools in falsehoods detection, investigating applied techniques and technologies and especially their limitations. Based on the identified concepts and areas of Information Science, also (equaly), it’s the scope of this research to show a suggestion of “framework” that is able to detect or indicate falsehoods, specifically, has the ability to detect profiles, false identities and misinformation, rumors, or unreliable information, processing large volumes of data on social networking tools. It was identified with this research, the Information Science can contribute to building skilled frameworks to provide criteria so that users can identify and discern actual content of questionable content, dumped daily on the *Web*.

Keywords: Computer Forensics. Intelligence Cybernetics. Big Data. Semantic Web. Fake profiles.

Lista de Figuras

Lista de quadros

Lista de siglas

AP	Análise do publicador
API	*Application Programming Interface*
CI	Ciência da Informação
CN	Credibilidade da Notícia
DAN2	*Dynamic Architecture for Artificial Neural Networks*
EIS	*Eletronic Information System*
Exifs	*Excheangeable image file format*
FCN	Frequência de comentários negativos
FPA	*Fake Profile Attack*
HTML	*Hypertext Markup Language* (Linguagem de Marcação de Hipertexto)
HTTP	*Hypertext Transfer Protocol*
ICA	*Identify Cloning Attacks*
ICC	*International Color Consortium*
IDF	*Inverse Document Frequency*
IPTC	*International Press Telecommunications Council*
JSON	*JavaScript Object Notation*
OSINT	*Open Source Intelligence*
OWL	*Web Ontology Language*
RDF	*Resource Description Framework*
RDF-S	*Resource Description Framework Schema*
SMS	*Short Message Service*
SPARQL	*Protocol and RDF Query Language*
SVN	*Support Vector Machine*
TF	*Term-Frequency*
URI	*Uniform Resource Indicator*
URL	*Uniform Resource Locator*
VC	Veículos Confiáveis
W3C	*Word Wide Web Consortium*
XML	*eXtensible Markup* Language (Linguagem de Marcação Extensível)

Sumário

1 Introdução

A internet desponta como fonte de informação para pessoas físicas e jurídicas, sendo que as ferramentas de redes sociais contribuem para a publicação, transferência e difusão de informações.

É inegável a importância da Internet como fonte de acesso às preciosas e relevantes informações a toda sociedade. No entanto, nos últimos anos, o crescimento exponencial do volume de informações nos meios digitais tem acarretado a demanda por instrumentos e métodos capazes de suprir as novas necessidades informacionais. (RAMALHO, 2006).

De fato, tratar informação digital não é tão simples como parece. Nem todos os conceitos, recursos e informações estão localizáveis e podem ser indicadas ou indexadas por meio de *Uniform Resource Locators* (URLs) e *Uniform Resource Indicators* (URIs). Páginas dinâmicas dificultam a coleta e tratamento de dados por outras aplicações.

Já advertia Santarém Segundo (2004, p. 135)

> Durante este trabalho verificamos que o tratamento da informação digital está sofrendo mudanças, acarretadas principalmente pelo crescimento da quantidade de informação disponível na Internet e que ferramentas e conceitos estão sendo desenvolvidos, procurando facilitar o trabalho de armazenamento, descrição, indexação e recuperação de informações na *Web*. Observamos que as ferramentas de busca são serviços importantes, mas ainda muito limitados, pois descrevem e indexam uma quantidade de informação muito pequena em relação ao tamanho da *Web*; como produzem resultados, na maioria das vezes, indesejados ou insignificantes, grande parte deste problema é gerado pela forma de armazenamento das informações que se apresentam sem um padrão ou formato que permita a estes serviços um melhor resultado para o usuário.

De fato, recuperar informações na *web* demonstra-se complexo, sobretudo pela quantidade de informações e dados não relevantes que podem ser coletados, sem a existência de recursos e critérios.

Lixo virtual é uma consequência do “Big Data”, ou seja, dados e recursos sem relevância para uma específica pesquisa, ruídos. Assim, grande parte dos autores entende que as informações *web* necessitam de padronização para que possa existir uma recuperação mais objetiva. (SANTAREM SEGUNDO, 2004).

Organizar a informação é essencial não só para a recuperação, mas para atribuição de significado e aquisição de conhecimento e neste trabalho, será demonstrado que é importante para detectar perfis e páginas falsas na rede.

Uma das propostas para organizar o conteúdo disposto na *Web* escora-se no conceito da *Web* Semântica, que segundo Berners-Lee, Hendler e Lassila (2001, p. 2, tradução nossa) pode ser "[...] uma extensão da *Web* atual, onde a informação possui um significado claro e bem definido, possibilitando uma melhor interação entre computadores e pessoas".

Assim, emergem-se duas grandes áreas desafiadoras para a Ciência da Informação no tratamento de grandes volumes de dados: a) A coleta adequada destes dados; b) Técnicas e ferramentas para análise dos dados, com extração de significados e detecção de falsidades.

Partindo deste pressuposto, cumpre investigar a literatura sobre o tema, se as principais ferramentas de redes sociais dispõem de conectividades que permitam as atividades citadas, e principalmente, se a padronização ou descrição da informação despejada nestas ferramentas auxiliará na extração de seu significado e detecção de um problema muito comum nas redes sociais: A falsidade.

1.1 Problema

Sabe-se que a *Web* Sintática está relacionada à ideia de informações apresentadas em páginas, onde a interpretação é papel dos usuários ou sujeitos informacionais. (BREITMAN, 2005). Com o advento das redes sociais, tornou-se difícil sequer ler, quanto mais interpretar a infinidade de dados apresentados em tais ambientes, que caracterizam o denominado "Big Data". É preciso ir além de meras representações sintáticas, mas extrair significado deste precioso contexto informacional como condição para a detecção de falsidades na rede.

Ramalho (2006, p. 41) já advertia sobre a problemática envolvendo a falta de padronização de manipulação da informação ao asseverar que:

> Nesta perspectiva, para a concretização do projeto *Web* Semântica torna-se necessário padronizar a maneira pela qual os diversos tipos de softwares, utilizados no ambiente *Web*, manipulam as informações, assim como possibilitar meios que possam ser utilizados para descrever os aspectos semânticos de cada recurso, de modo que possa haver intercâmbio de informações de maneira

> padronizada, e que os recursos sejam descritos a partir de bases tecnológicas compatíveis.

Este problema aparentemente ainda não fora solucionado, embora certamente passe pela criação de “agentes computacionais”, que possibilitem coletar as informações de fontes diversas, relacioná-las automaticamente e retorná-las de forma organizada ao consumidor da informação. Quando se trata de Big Data, a problemática está ainda mais longe de ser solucionada.

Assim, tem-se uma enorme dificuldade em se recuperar e tratar o Big Data gerado nas ferramentas de redes sociais. Sem controle, inúmeros perfis falsos se proliferam, lesando usuários e empresas. Estes “*bots*” ou aplicações que simulam perfis são criadas para gerarem audiência a determinada página ou mesmo para atrair usuários para armadilhas cibernéticas. O uso de *bots* desvaloriza o sistema, pois por se tratar de *spam*, perde-se a confiança na rede. (MÍDIAS SOCIAIS, 2014).

Ademais, não se tem métricas para se detectar conteúdos fraudados e perfis falsos nas redes. Hoje já se fala em “*Web* Pragmática”, onde dados são gerados a partir dos sujeitos informacionais e onde se permite adicionar contexto às informações, com a criação de ontologias dinâmicas. (VECHIATTO, 2013). Considerar as características dos leitores e consumidores da informação no processo informacional demonstra uma forma de tratamento e organização mais adequada da informação nas redes sociais. Mais que isso, nesta pesquisa se demonstrará que poderá ser elemento para detecção de falsidades. A *Web* Pragmática deve considerar os sujeitos nas relações com os signos. Nesta abordagem o controle sobre a representação deve mudar do produtor para o consumidor da informação. (ANDRADE, 2012). Usa-se na Web Pragmática o contexto pragmático para manipular recursos semânticos, oferecendo aos usuários serviços mais personalizados. No entanto, não foram identificadas aplicações práticas que considerem “em tempo real” o sujeito que busca uma informação ou mesmo, o sujeito que posta uma “informação” nas redes sociais. Não se identificou estudos que considerem as expressões ou manifestações do sujeito cognoscente ao ter contato com a informação, seja ela verdadeira ou não.

Por outro lado, a criação de *“profiles”* (perfis relativos ao buscador da informação), e uma metodologia de busca de informações no mar do “Big Data” poderia favorecer um ambiente de extração de conhecimento para determinada área do saber em específico. O que se verifica, no entanto, é o crescimento das redes

sociais como plataforma de desinformação e falsidades. Milhares de notícias falsas são postadas diariamente por perfis falsos ou “robôs”, aplicativos que simulam um perfil real nas redes sociais. O intuito é evitar que na apuração da autoria, uma pessoa real possa vir a ser responsabilizada.

Nesta perspectiva, apresenta-se como problema de pesquisa a investigação de meios e métodos para extrair significado de *data sets* de ferramentas de redes sociais, possibilitado a construção de conhecimento, avaliação de cenários e detecção de informações e perfis falsos.

1.2 Objetivos

A presente pesquisa tem por escopo realizar a revisão de literatura relativa a modelos, propostas e técnicas para detecção de falsidades envolvendo grandes repositórios de informações, especificamente, em ferramentas de redes sociais e relativas a perfis falsos, e neste contexto, avaliar sua aplicabilidade e limitações. Igualmente, valendo-se da contribuição de conceitos e técnicas da Ciência da Informação, a ser aliada à Inteligência Cibernética e Computação Forense, é objetivo deste trabalho apresentar uma proposta de um **modelo conceitual para desenvolvimento de aplicações computacionais** que, sendo aplicado em ferramentas de redes sociais ou grandes repositórios de informações, possa avaliar possibilidades de perfis e identidades falsas (*fakes*), com isso, contribuindo para construção de um ambiente informacional mais seguro e de credibilidade.

1.2.1 Objetivos específicos

Apresenta-se como objetivos específicos da presente pesquisa:

a) Apresentar os desafios envolvendo a organização e recuperação da informação, trazidos pelo Big Data das redes sociais, no contexto da detecção de perfis falsos;
b) Apurar como as principais áreas estudadas no âmbito da Ciência da Informação e *Web* Semântica podem contribuir no processo de organização, mineração e significação de dados na *Web* Social, na detecção de falsidades na rede;

c) Investigar os trabalhos e pesquisas já desenvolvidos abrangendo metodologias para investigação de perfis falsos (*fakes*) em ferramentas de redes sociais, apresentando seus pontos positivos e limitações;
d) Apresentar contribuições e uma proposta de modelo conceitual para rotular informações, apontar e detectar a existência de perfis falsos em ferramentas de redes sociais, contando com aportes de áreas estudadas na Ciência da Informação, de aplicação em investigação e inteligência cibernética ou mesmo na computação forense, contribuindo para uma *web* com informações mais precisas e minimizando a desinformação, reduzindo os danos causados a pessoas no mundo todo por perfis falsos atuantes nas redes sociais.

1.3 Justificativa

A busca pelo melhor resultado na recuperação da informação é tema que já vem sendo abordado na Ciência da Informação há algum tempo. Como estabelece Santarém Segundo (2010, p. 25)

> Desde a publicação do “Manual de Documentação”, de Paul Otlet em 1937 (LÓPES YEPES, 1989) e do MEMEX de Vannevar Bush em 1945 (BARRETO, 2008), que diversos estudos vêm apresentando métodos e técnicas para evoluir o processo de recuperação da informação.

Embora tenhamos uma gama infindável de informações com o Big Data, não sabemos se utilizamos todas as potencialidades desta informação. Com as redes sociais, todos são publicadores, muitos postam informações verídicas, porém, outros usam a rede para espalhar desinformações e fraudar, incluindo a prática de crimes informáticos. Mooers (1951) criou o termo “recuperação da informação” e nesta época já apontava os problemas para esta nova área do saber.

No Brasil, o uso de perfis falsos nas redes sociais pode ser considerado crime de falsa identidade, previsto no Código Penal Brasileiro, em seu art. 307[1], com pena de detenção de três meses a um ano ou multa. (BRASIL, 1940).

[1] Art. 307 - Atribuir-se ou atribuir a terceiro falsa identidade para obter vantagem, em proveito próprio ou alheio, ou para causar dano a outrem: Pena - detenção, de três meses a um ano, ou multa, se o fato não constitui elemento de crime mais grave.

De acordo com levantamento da Solve Media, o mercado publicitário mundial na Internet perde U$$ 1,5 bilhão ao ano com os robôs ou “*bots*”, ferramentas de *web* automáticas que geram comportamentos repetitivos e em série, falsos, como se fossem o de um usuário real. (SOLVE MIDIA, 2012).

Embora existam pesquisas que tratam da recuperação da informação no ambiente cibernético, percebe-se uma temática esquecida nas pesquisas: a extração de significados na mineração de grande volume de dados ou de dados inseridos neste contexto e a criação de um ambiente virtual seguro, livre de fraudes, diga-se, um ambiente com neutralidade de conteúdo.

E os problemas persistem, pois, a quantidade de dados despejados de forma desestruturada na rede é imensa. Assim, a presente pesquisa se justifica, pois, um ambiente seguro na Internet passa necessariamente pelo estudo de meios que permitam auferir a autenticidade de perfis e pessoas nas redes sociais. Deste modo, esta, além de investigar as principais análises sobre o tema, vem averiguar as melhores técnicas e práticas estudadas na Ciência da Informação e como podem ser empregadas na detecção de perfis falsos.

Deste modo, evidencia-se que são necessárias soluções inovadoras que efetivamente insiram a Ciência da Informação como base na arquitetura de um ambiente que pretenda extrair significados e diminuir a falsidade no ambiente das redes sociais, ou que pelo menos ofereçam indícios aos usuários antes que estes possam acessar, dar crédito ou compartilhar uma “notícia” da rede social oriunda de perfis de identidade duvidosa.

Neste contexto, esta pesquisa, ao revisar o estado da arte das pesquisas sobre o tema, no âmbito global, alerta sobre a preciosidade e volatilidade das informações disponíveis nas redes sociais, ao mesmo tempo em que aponta caminhos e uma proposta de um modelo para construção de mecanismos que poderão apurar perfis falsos nas redes sociais.

Como inovação, se verifica que o modelo descentralizado de descrição de objetos gerados nas redes sociais, apresentado neste trabalho científico, poderá aprimorar a organização, estruturação e recuperação da informação e principalmente, ser utilizado como aliado de usuários da Internet na recuperação de informações idôneas e investigação de falsidades.

1.4 Percurso metodológico

Na presente pesquisa, estuda-se a problemática envolvendo a detecção de perfis falsos nas redes sociais, investigam-se os trabalhos e pesquisas sobre o tema e aplicando conceitos e aportes teóricos da Ciência da Informação, bem como da revisão internacional da literatura, apresenta-se uma proposta de modelo de sistema para detecção de perfis falsos em tais ambientes, aplicável às áreas de inteligência cibernética, mas que pode ser adaptável a demais campos do saber e áreas do conhecimento, de notório interesse para empresas da publicidade, autoridades de aplicação de lei, investigadores, peritos, aplicativos, empresas de ciber inteligência, ferramentas de redes sociais, dentre outras.

Trata-se o presente trabalho de natureza aplicada, sob o prisma de seus objetivos, de uma pesquisa teórica e exploratória, descritiva no que diz respeito aos fins e de investigação bibliográfica, no que concerte aos meios ou delineamento da pesquisa. (MORESI, 2003). Quanto à abordagem trata-se de uma pesquisa qualitativa. (GIL, 2002).

Já na doutrina de Silva e Menezes (2005), a presente pesquisa, quanto à forma, pode ser representada como uma pesquisa aplicada, no escopo e produzir conhecimentos para aplicação prática, com foco à solução de problemas específicos.

Exploraram-se técnicas e sistemas disponíveis que podem ser usados no tratamento de grandes volumes de dados, e na detecção de falsidades. Posteriormente, investigaram-se por meio de levantamento bibliográfico em teses, livros, artigos, em bases de dados internacionais, dentre outros, os avanços da pesquisa mundial envolvendo detecção de perfis falsos em redes sociais.

Assim, de modo a cumprir os objetivos do presente trabalho, foi investigado no corpus teórico da Ciência da Informação, conceitos que podem ser aplicados e utilizados na construção de métodos que possam ser usados para apurar a credibilidade de perfis e informações nas redes sociais.

No desenvolvimento da pesquisa, foram considerados documentos pertinentes e atuais em português e inglês. Não se realizou limitação cronológica ou geográfica. No levantamento bibliográfico, foram realizados fichamentos e registro dos artigos científicos, com apontamentos e para destaques, referências e uso posterior.

Após o primeiro momento, quando se formou o marco teórico e conceitual da pesquisa, passou-se à segunda etapa do trabalho, quando foram refletidas as temáticas envolvidas no escopo da produção de novos conhecimentos para elaboração de uma proposta de modelo conceitual para detecção de perfis falsos nas redes sociais.

1.5 Estrutura do trabalho

A **Introdução** apresenta a problemática que motiva a pesquisa envolvendo o grande volume de dados hoje disponível nas redes sociais e os perfis falsos em tais ambientes. **O Capítulo 2**, as áreas de estudo da Ciência da Informação e as contribuições identificadas para o melhor tratamento de informações na detecção de perfis falsos. **O Capítulo 3** investiga desafios na fase de coleta de dado na detecção de perfis falsos em ferramentas de redes sociais, problemas e limitações.

O **Capítulo 4** investiga a produção de literatura nacional e internacional a respeito da temática envolvendo detecção de perfis falsos, avaliando suas vantagens e limitações e pontos relevantes à presente pesquisa. Também se avaliam pontos que foram considerados na construção desta proposta objeto do presente trabalho. Já no **Capítulo 5**, considerando a contribuição da Ciência da Informação, apresenta-se uma proposta de modelo conceitual para inteligência cibernética em grandes volumes de dados dispostos nas redes sociais, especificamente, em relação à detecção de perfis falsos em tais ambientes, apontando-se ainda quesitos de observação e pontos importantes a aplicações e parâmetros que tratem deste contexto, demonstrando a relevância do profissional de Ciência da Informação no desenho destas aplicações. Em **considerações finais, Capítulo 6**, consolida-se as conclusões acerca da pesquisa sobre a importância do uso de conceitos e técnicas da Ciência da Informação que podem auxiliar na construção de aplicações que se destinem a apurar perfis falsos em redes sociais, checa-se se os objetivos foram cumpridos e apresenta-se o diferencial do modelo proposto. **Em trabalhos futuros, Capítulo 7,** apresenta-se os trabalhos futuros, dentre eles é abordado o intento de se investigar o comportamento dos metadados em ferramentas de redes sociais, bem como técnicas para descrição independente de objetos destas redes.

2 Áreas da Ciência da Informação e suas contribuições na detecção de perfis falsos em ferramentas de redes sociais

É sabido que inúmeras ferramentas de redes sociais permitem a criação e perfis falsos, também chamados de “*fakes*”. Conquanto os termos de uso e políticas destas ferramentas exijam que pessoas reais usem as ferramentas, não se pode garantir que usuários mal-intencionados não criem personagens ou usem imagens de outras pessoas na criação de perfis falsos, que podem ser utilizados para aplicação de golpes e fraudes.

2.1 A problemática dos perfis *Fakes*

É preciso distinguir duas formas de *fakes*. A primeira, onde atribui-se a um perfil fotos de desconhecidos ou de bancos de imagens, onde embora esteja violando-se termos de uso, não existe crime, a princípio. A segunda, onde o *fake* é criado a partir de uma pessoa real, onde é possível caracterizar o crime de falsidade ideológica. Além disso, está a se tratar de violação ao art. 5º, inciso X da Constituição Federal, que protege o direito de imagem (ATHENIENSE, [2010?]).

Assim dispõe o Código Penal Brasileiro sobre o crime de falsidade ideológica:

> Art. 299 - Omitir, em documento público ou particular, declaração que dele devia constar, ou nele inserir ou fazer inserir declaração falsa ou diversa da que devia ser escrita, com o fim de prejudicar direito, criar obrigação ou alterar a verdade sobre fato juridicamente relevante.
> Pena - reclusão, de um a cinco anos, e multa, se o documento é público, e reclusão de um a três anos, e multa, se o documento é particular. (BRASIL, 1940).

Esta é a conclusão trazida por Atheniense ([2010?]), ao inferir que

> Mas se o fake é criado a partir de uma pessoa real, viva ou morta, o responsável poderá cometer o crime de falsidade ideológica, desde que cause dano à vítima. O ato de incorporar a personalidade de outras pessoas e manifestar em nome de outrem, inserindo declaração falsa ou diversa da que devia ser escrita, com o fim de prejudicar direito, criar obrigação ou alterar a verdade sobre fato juridicamente relevante é crime de falsidade ideológica.

Criar um perfil *fake* é conduta extremamente simples no meio eletrônico, sem que haja por parte do provedor deste serviço que o suporta qualquer tipo de autenticidade e identidade. Conquanto muitos provedores já tenham sido condenados no Brasil por permitirem perfis *fakes* (BLUM et al., 2009) é fato que hoje, nos termos do Marco Civil da Internet, o provedor de aplicação não será responsável por conteúdos gerados por terceiros. (JESUS; MILAGRE, 2014).

Os perfis *fakes* hoje são utilizados para inúmeras atividades, muitas delas ilícitas, dentre as quais as identificadas neste trabalho:

- Envio de mensagens não solicitadas (Spam);
- Postagens e comentários com links e códigos maliciosos;
- Envio de falsas notícias;
- Crimes contra a honra;
- Propaganda irregular;
- Remoção de conteúdos da Internet;
- Direcionamento de conteúdos a internautas;
- Aumentar resultado e popularidade de páginas e perfis;
- Extorsão;

Quanto ao spam, mensagem não solicitada enviada em massa, é importante destacar que no Brasil não é crime e embora existissem no Congresso Nacional projetos de Lei que tratavam da matéria, todos foram arquivados. Contudo, o spam com a finalidade de furto de dados financeiros ou pessoais pode ser caracterizado como crime de estelionato ou furto mediante fraude, na modalidade consumada ou tentada de acordo com o Código Penal Brasileiro.

Neste sentido:

> Art. 171 - Obter, para si ou para outrem, vantagem ilícita, em prejuízo alheio, induzindo ou mantendo alguém em erro, mediante artifício, ardil, ou qualquer outro meio fraudulento.
> Pena - reclusão, de um a cinco anos, e multa, de quinhentos mil réis a dez contos de réis.
> Art. 155 - Subtrair, para si ou para outrem, coisa alheia móvel:
> § 4º - A pena é de reclusão de dois a oito anos, e multa, se o crime é cometido:
> II - com abuso de confiança, ou mediante fraude, escalada ou destreza; [...]. (BRASIL, 1940).

Outro problema em relação à falsidade na rede são os *bots*, que nada mais são do que perfis criados nas redes sociais e que não correspondem a pessoas

reais, mas em verdade, são aplicativos que simulam interações humanas, também utilizadas para criminalidade, compartilhamento de ofensas e para aumentar a "importância" online de pessoas. O Facebook informa que descobre de 67 a 137 milhões de perfis falsos por ano. É preciso esclarecer que o Bot é um aplicativo que simula um perfil na Internet, podendo ser ou não um fake (falso). Temos, por outro lado, fakes administrados por pessoas. Assim, o Bot seria uma das formas pelas quais fakes podem ser criados e administrados. O Twitter afirma que cinco por cento, ou 24 milhões, de suas contas são *bots*. O Instagram está lotado de milhões deles, que copiam os perfis das pessoas, compartilham suas fotos e deixam até comentários nas imagens. (BILTON, 2014).

Esta audiência *fake* traz inúmeros danos para a sociedade, além de alterar o fluxo normal da informação, o que se nominou de "floodagem da opinião" (GOUVEIA, MALINI, CIARELLI, 2014), serve para a prática de golpes e crimes. Não bastasse, expõe claramente que os métodos de detecção e repressão a perfis falsos utilizados pelas ferramentas de redes sociais são obsoletos, pois não incomum verificar-se que páginas falsas permanecem no ar, mesmo diante de inúmeros indícios de serem robôs. Em sentido contrário, páginas reais são removidas pela ação de falsos denunciantes, através de perfis *fakes*.

Algumas ferramentas, como TwitterAudit[2], permite que usuários detectem quantos perfis *fakes* estão linkados a seus perfis. Elas consideram o número de *tweets*, data do último *tweet* e relacionamento dos perfis com amigos. Mas a problemática persiste, pois na era dos grandes volumes de dados (Big Data), poucos são os critérios para coletar e classificar dados, fornecendo elementos para apuração de perfis falsos.

De fato, encontra-se diante de um conjunto de dados não estruturados, extremamente grandes e precisam de ferramentas especialmente preparadas para lidar com grandes volumes, denominado Big Data, de forma que informações neste ambiente possam ser encontradas, analisadas e aproveitadas em tempo hábil. (NEVES, 2014).

Diante desta problemática, foram revisadas áreas afetas à Ciência da Informação, identificando como podem contribuir em mecanismos para detecção destes perfis forjados.

---

[2] https://www.twitteraudit.com/periciadigital

2.2 Áreas da Ciência da Informação e suas contribuições em projetos de detecção de perfis falsos

A Ciência da Informação tem como escopo investigar o comportamento informacional e as forças que governam este fluxo. Preocupa-se igualmente teorizar a informação que propiciará melhorias de várias instituições procedimentos dedicados ao acúmulo e transmissão do conhecimento. (BORKO, 1968).

Um ambiente virtual seguro pressupõe “autenticidade” de seus atores e a apropriação da identidade visual de outrem na rede é um problema atual. Neste contexto, é inegável que o estudo das técnicas de detecção, perfis falsos e a concepção de modelos, neste sentido, contribuem para produção de conhecimento, evitando-se crimes e golpes.

Esta informação, produzida a partir dos estudos ora refletidos neste trabalho, permitirá a construção e arquitetura de ferramentas que, sejam capazes de descrever objetos e minimizar falsidades, especificamente perfis falsos, podendo inclusive detectar o uso indevido de imagens, textos, ou mesmo reconhecer padrões que sugiram um perfil falso.

Para tanto, é preciso interpretar adequadamente elementos informacionais dispostos nas ferramentas de redes sociais de modo a investigar se replicados indevidamente ou relativos a perfis falsos, elementos estes que podem se perder facilmente, considerando o Big Data produzido nestes ambientes.

À medida que os sistemas de informações se tornam mais globais e interconectados, as informações implícitas são muitas vezes perdidas, o que desafia a Ciência da Informação a ser mais receptiva aos impactos sociais e culturais dos processos interpretativos, e também, às diferenças qualitativas entre diferentes contextos e mídias. Esta mudança pressupõe a inserção de processos interpretativos com condição indispensável nos processos de informação. (CAPURRRO; HJORLAND, 2007).

Estes sistemas, na era do Big Data, podem conter elementos não perceptíveis em um primeiro momento aos usuários, mas que se tratados, poderiam revelar a existência de perfis falsos. É papel da Ciência da Informação possibilitar a representação de informações não aparentes. Segundo BUCKLAND (1991, p. 9)

> determinar o que pode ser informativo é uma tarefa difícil. Árvores, por exemplo, provêm madeira, assim como lenha para construção e carvão para calefação. Alguém naturalmente pensa em árvores como informação, mas árvores são informativas no mínimo de duas maneiras. Obviamente, assim como representativas as árvores são sobre elas mesmas. Não tão óbvio assim, as diferenças na espessura das árvores são causadas por variações no tempo, e portanto, são suas evidências. Padrões refletem um específico ciclo de anos constituindo informações valiosas por arqueólogos que precisam datar antigas vigas.

A mineração de grandes volumes de dados nas redes sociais pode nos revelar novas informações, não tão óbvias, mas que podem ser preciosas na detecção de fraudes.

Assim, verificamos que a Ciência da Informação poderá ser útil ao oferecer elementos para detecção de evidências e padrões na construção de propostas para detecção de perfis falsos, que consigam extrair outros elementos informativos de perfis nas ferramentas de redes sociais.

Neste contexto, cumpre investigar processos e campos de estudo da Ciência da Informação que podem ser considerados na construção de modelos que se disponham a detectar robôs e *fakes*.

Apresentamos a seguir o resultado de nossa análise de áreas de estudo da Ciência da Informação que podem ser apurados e refletidos no escopo do presente trabalho.

2.2.1 *Web* Semântica

Conforme explicado por Breitman (2005), vivencia-se hoje uma *Web* Sintática, onde a informática apenas apresenta a informação, cabendo aos seres humanos interpretar, avaliar, classificar e selecionar informações.

Como estabelece Santarém Segundo (2004, p. 106),

> A facilidade de interpretação que o ser humano tem em distinguir uma palavra de um determinado contexto não é encontrada nos computadores e nos robôs de busca, não permitindo, assim, que os mesmos consigam entender o conteúdo significativo de uma página *Web* antes de descrevê-la e informá-la como resultado a um usuário.

Já com *Web* Semântica, espera-se recuperar o significado de páginas permitindo aos computadores entenderem o significado dos conteúdos. Para isso é

indispensável embutir semântica nos recursos *Web*, como às postagens lançadas em ferramentas de redes sociais.

Neste contexto, teremos a facilitação da recuperação da informação, considerando que as máquinas poderão realizar associações e deduções para inferir o conteúdo de determinada postagem.

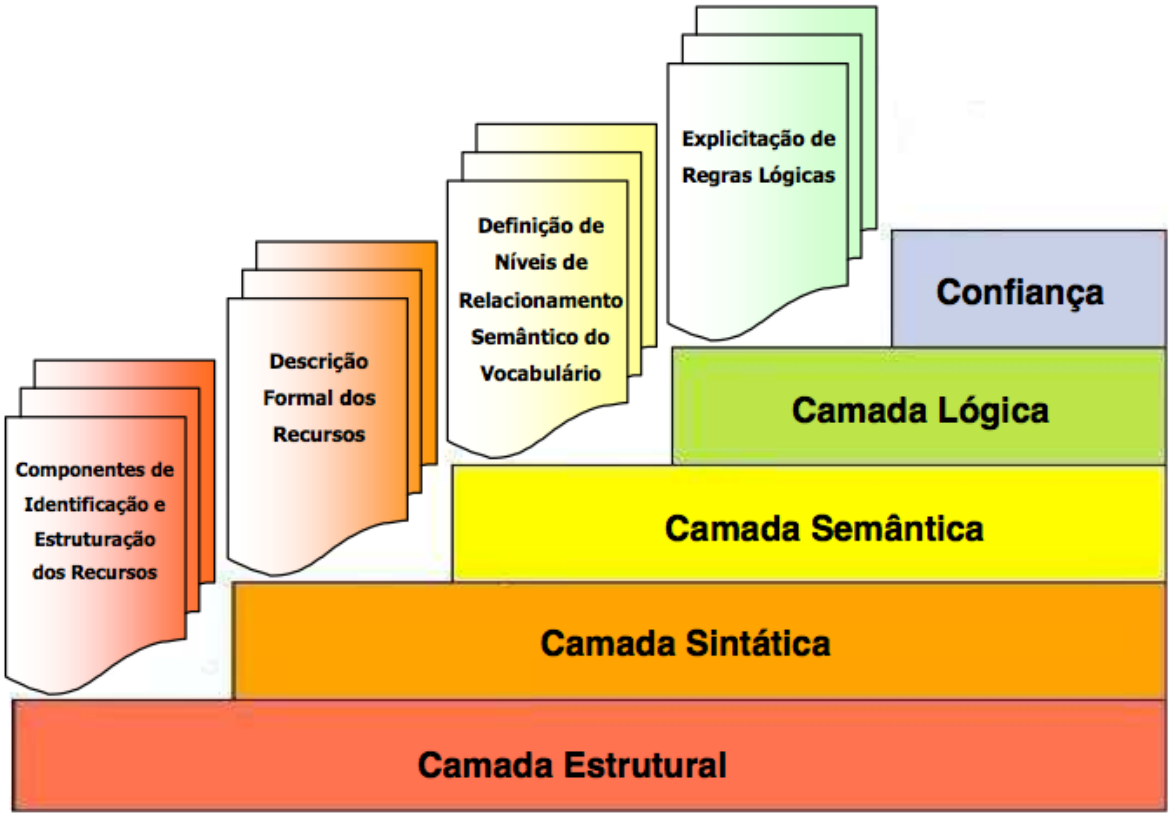
A *Web* atual e sua vasta quantidade de informações foram desenvolvidas para o consumo humano, com suporte limitado para o processamento da máquina. Mesmo em se tratando de dados bem estruturados, a variação das terminologias dificultam as máquinas de automaticamente “entenderem” os dados. Isto reflete na qualidade dos dados retornados nos buscadores. Uma mesma palavra pode representar diversos significados e estar em contextos diversos, o que não será entendido pela máquina, reduzindo a precisão dos buscadores. (BERNERS-LEE, 1998).

Para Ramalho (2006, p. 98),

> [...] a partir da análise apresentada, pode-se afirmar que o objetivo principal do projeto *Web* Semântica é possibilitar a classificação de recursos informacionais disponíveis no ambiente *Web* a partir de categorias que possam ser “interpretadas” automaticamente por computadores, possibilitando a realização de tarefas mais sofisticadas de forma automatizada.

Tem-se assim o espectro funcional da *Web* Semântica, vejamos a figura 1.

Figura 1 – Espectro funcional da *web* semântica



Fonte: RAMALHO (2006, p.101)

Assim, a *web* semântica nasce com a proposta de não ser uma nova *Web*, mas a extensão da *web* atual, onde a informação é bem definida em seu significado (BERNES-LEE, et al. 2001). Em projetos de mineração de dados em redes sociais, tal proposta auxilia na coleta dos dados e até mesmo em uma possível análise em tempo de coleta de elementos que sugiram perfis falsos.

> Metadados são atributos que representam uma entidade (objeto do mundo real) em um sistema de informação. Em outras palavras, são elementos descritivos ou atributos referenciais codificados, que representam características próprias ou atribuídas às entidades; são ainda dados que descrevem outros dados em um sistema de informação, com o intuito de identificar de forma única uma entidade (recurso informacional) para posterior recuperação. Os padrões de metadados são estruturas de descrição constituídas por um conjunto predeterminado de metadados (atributos codificados ou identificadores de uma entidade) metodologicamente construídos e padronizados. (ALVES, 2010, p. 47).

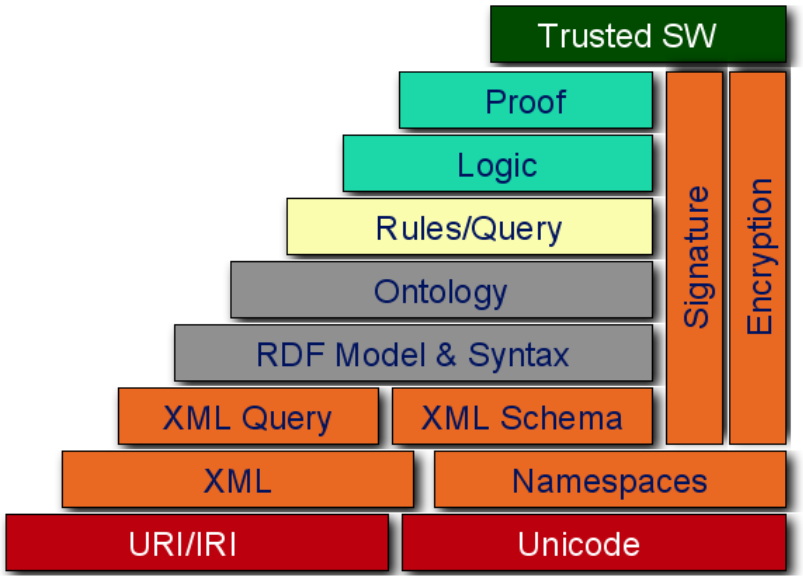
Ao se criar um perfil na ferramenta de rede social, um exemplo de metadados para este recurso seria, data de criação, tamanho e formato da imagem de perfil, origem da conexão (número IP), etc.

De maneira que, informação sobre dados é um dos caminhos para a obtenção de uma *web* com mais significado, especialmente na era do Big Data. Recursos como “metadados” (dados sobre dados) auxiliam no processo de descrição de um “dado”. Evidencia-se que na *Web*, o Big Data (imenso conteúdo disponível) e a heterogeneidade dos recursos evidenciam cada vez mais a necessidade de adoção de padrões para metadados.

O uso da *Web* Semântica proporciona a combinação de dados e metadados permitindo inferências de conclusão, explicitando um conhecimento até então oculto, como por exemplo, o uso de imagens de um perfil original por um perfil falso. Tal recurso demonstra-se indispensável na proposta deste projeto de pesquisa.

A estrutura da *Web* Semântica é descrita abaixo.

Figura 2 – Arquitetura em Camadas da *Web* Semântica



Fonte: Disponível em: http://www.w3c.it/talks/eva2004Jerusalem/all.html

Na camada base da Figura 2 identificamos o URI (*Uniform Resource Identifiers*), padrão para descrição de identificadores universais de recursos e códigos internacionais de dados.

Se rotular a informação é essencial, é necessário que os objetos sejam identificados para depois serem comparados. Para isso, também existe o padrão URI, que pode ser descrito como um localizador ou nome de um objeto. O padrão URI é definido na norma RFC3986[3]. A Norma em comento estabelece protocolos padrões de rastreabilidade para a Internet, definindo o URI como uma sequência compacta de caracteres, uniforme, que identifica em abstrato um recurso na rede.

Uma estrutura que permita coletar grandes volumes de dados em ferramentas de redes sociais precisa definir uma URI única para os diversos objetos capturados, que permita que o objeto seja único em toda a rede. Algumas ferramentas de redes sociais e comunicadores não identificam seus objetos, como o WhatsApp, o que dificulta ou impede a rotulação adequada da informação, grupos, postagens e usuários.

O WhatsApp é um aplicativo para troca de mensagens disponível para diversas plataformas. Permite a troca de mensagens por equipamento celular sem o pagamento de mensagens SMS (*short message service*). A aplicação tem hoje 600 milhões de usuários ativos com 45 milhões de usuários no Brasil[4].

Ao contrário de outras ferramentas sociais, o WhatsApp não descreve um recurso com um índice ou numeração única. Por exemplo, no Facebook um grupo, usuário, vídeo, foto ou página tem uma URL própria ou mesmo um código único que o identifica. Já no WhatsApp grupos são identificados apenas pelo nome e usuários se valem dos seus números de telefone para se identificarem na ferramenta.

Na segunda camada (XML + NS + xmlschema) verificamos a importância da linguagem XML e os schemas na criação e interoperabilidade dos metadados. São recursos considerados sustentação da *Web* Semântica. Temos também os namespaces que são URIs que garantem segurança entre vocabulários de metadados.

A terceira, quarta, quinta e sexta camada estão contidas em um grupo chamado assinatura digital (*Digital Signature*) onde se validam a integralidade dos dados usados para as tarefas dos agentes da *Web* Semântica. Na terceira camada apresenta-se o RDF, juntamente com o RDF *Schema*, arquitetura concebida para descrever metadados sobre recursos em termos de suas propriedades e valores.

---

[3] http://tools.ietf.org/html/rfc3986

[4] http://www.gazetadopovo.com.br/economia/whatsapp-atinge-600-milhoes-de-usuarios-ativos-diz-cofundador-eckqsierkjxfcufa8ty5cqfym

A quarta camada “Ontologia” ou também denominada como “Vocabulário Estruturado”, pode ser entendida como:

> A camada de ontologia [que] representa a veia semântica central de metadados na *Web*, onde simples descrições para complexos esquemas classificatórios devem ser criados e registrados, de modo que os agentes inteligentes possam interpretar dados, fazer inferências e executar tarefas. (GRENNBERG apud SANTAREM SEGUNDO, 2004, p. 112).

A quinta camada apresenta a lógica, onde se pode afirmar que agentes computacionais podem tomar decisões lógicas baseados nas descrições encontradas nas estruturas inferiores.

Por fim, a sexta e sétima camadas (*Proof* e *Trust*) têm por escopo testar a confiabilidade da informação a ser recuperada, com base nos metadados existentes nas camadas inferiores.

2.2.1.1 Uso de metadados e linguagens de representação XML

Verificou-se que a *Web* semântica pode oferecer importante contribuição na detecção de perfis falsos, seja pela construção de ontologias, seja pela aplicação ou coleta descentralizada de metadados em objetos que trafegam pelas redes sociais, oportunizando análises de contextos informacionais. A aplicação de metadados é feita pela descrição que pode ser feita por uma linguagem de marcação.

Como constatado, dados que definem dados recebem o nome de “metadados” (SANTAREM SEGUNDO, 2004, p. 109). Ao se criar um perfil na ferramenta de rede social, um exemplo de metadados para este recurso seria, data de criação, tamanho e formato da imagem de perfil, origem da conexão (número IP), etc. Quando referidos a imagens, os metadados recebem o nome de EXIFs.

Exifs (*Exchangeable Image File Format*) são padrões de metadados embutidos e que descrevem informações dentro de arquivos de imagens. Exifs são dados gerados automaticamente por dispositivos de imagens, câmeras, e embutidos em arquivos. (GUERREIRO, 2009).

Do mesmo modo, outro metadado aplicável a imagens que pode ser coletado é o relacionado ao “ICC Profile”, que define a forma que os pixels de cores são codificados, no padrão de metadado definido pela ICC (*International Color Consortium*). Estes metadados podem revelar inúmeras informações, pois são

gerados pela própria ferramenta de rede social em determinados casos. Eles definem, por exemplo, a sombra, a luminosidade, a capacidade de cores capazes de serem exibidas e outros dados. Cada ferramenta de rede social deverá ser estudada para se identificar critérios que podem ser “índices” para objetos visuais. Outro padrão existente em imagens é o IPTC (*International Press Telecommunications Council)*, desenvolvido na década de 1970 pela *International Press Telecommunications Council*, e que graças ao mesmo, hoje, aplicações conseguem cadastrar e editar metadados em arquivos de imagens, sobretudo para proteção de direitos autorais. (CHASTAIN, 2014).

Assim, os metadados podem ser extraídos e organizados por meio de linguagens de marcação, passível de serem pesquisados por agentes computacionais, como o XML.

XML (*Entensible Markup Language*) é uma linguagem de marcação de texto que permite codificar documentos arbitrários em forma taggeada (marcada) e consistente podendo ser lida por máquinas e humanos. (DOSSIS, 2012).

A definição de dados no padrão XML auxilia a leitura dos mesmos por diversos tipos de programas. Embora seja possível definir vocabulários e relações entre os elementos, é sabido que a estrutura é limitada e não permite extração de significados pelos sistemas.

A coleta de dados em ferramentas de redes sociais como Facebook ou Twitter imprescinde de uma padronização ou estruturação mínima de descrição dos recursos, para que após a coleta, outras ferramentas sejam utilizadas para analisar os dados.

Como se sabe, após a criação da Linguagem XML os engenheiros perceberam que não bastava descrever recursos informacionais sintaticamente, mas principalmente, deveriam ser criadas tecnologias que descrevessem o significado das informações. (SANTAREM, 2004).

Quanto se exibe uma página de posts na Internet, utiliza-se a linguagem HTML (*Hypertext Markup Language*), que tem por escopo exibir conteúdo de um documento. Este conteúdo pode sumir em questão de minutos. Por isso, deve-se considerar no projeto, a descrição de recursos por meio de metadados, podendo ser utilizada a linguagem XML, criada para estruturar informações sobre informações, sejam estas páginas, comentários, perfis ou postagens.

Assim, uma proposta para análise dinâmica de perfis em ambiente de grandes volumes de dados deve ser capaz de descrever recursos por meio de metadados, podendo-se ainda construir arquivos XML (ou outro padrão) em tempo de postagem, com atributos do usuário, e já armazenando adequadamente para tratamentos, como no exemplo, onde verificamos uma descrição gerada com o uso de metadados:

Quadro 1 – Exemplo de descrição de metadados de uma postagem na rede social

<estudante>
<postagem id =“919919191”>Vamos quebrar esta pessoa</postagem>
<estado>Fraude XX</estado> → resultado da aplicação do algoritimo
</estudante>

Fonte: Elaborado pelo autor

No exemplo do Quadro 1, a fraude pode estar relacionada a uma outra estrutura de perfis falsos, ou seja, pelo conteúdo, pode-se inferir emanado de perfis falsos ou não.

Deste modo, o recurso XML seria usado como formato de saída dos metadados, resultado dos analisadores de perfis falsos, que considerariam as ontologias, e ficariam atrelados à informação, como uma “sugestão” ou “carimbo” de análise para os usuários que detivessem permissão de acesso a tais análises.

Construir dinamicamente um atributo de alerta de acordo com o conteúdo da postagem. Com base nas características do perfil e da postagem, a ferramenta já direciona para uma área e insere *tags* dinamicamente, com alertas refletido pela postagem ou perfil, permitindo maior velocidade na análise de grandes volumes de dados e detecção de *fakes* e *bots*.

2.2.1.2 Uso de linguagem de representação RDF

O modelo RDF *(Resource Description Framework)* foi proposto em fevereiro de 1999 pelo W3C (*Word Wide Web Consortium)*, no escopo de possibilitar maior interoperabilidade no ambiente *Web*, oferecendo um padrão aberto para descrição de recursos. Permite a declaração de recursos sem interferir no mesmo, sendo a descrição uma entidade separada. (SANTAREM, 2004).

Trata-se de um padrão para embutir metadados em objetos. Igualmente, uma das linguagens utilizadas para especificação de ontologias. O conceito para utilização de metadados para definir recursos pode auxiliar o processamento e pesquisa automatizada, possibilitando um retorno de maior relevância. RDF Schema (RDF-S) é uma linguagem para construção de vocabulários RDF e então prover um *framework* para o tratamento de classes e propriedade de um domínio específico. (DOSSIS, 2012).

Neste contexto, a cada postagem na rede social, um modelo para detecção poderia aplicar o RDF *Schema* adequado, identificando, sujeito, predicado (propriedade) e identificando o objeto e valor, descrevendo de forma organizada, dados e metadados de publicações. Assim, a aplicação que implementa tal modelo poderá, por exemplo, analisar a rede e informar àquele que publica foto ou texto pertencente e outra pessoa ou protegida por direitos, evitando falsidades. Assim, as ontologias para descrição e publicação de imagens, aliadas à disponibilização de recursos em formato RDF podem ser importantes recursos para a marcação de objetos em ferramentas de redes sociais.

Deste modo, como verificado, o instituto será utilizado para rotulagem e descrição de objetos que são submetidos às ferramentas de rede sociais.

2.2.1.3 Ontologias e linguagem OWL

A crescente complexidade dos objetos armazenados na Internet e o grande volume de dados demandam processos de recuperação mais sofisticados e neste sentido as ontologias são técnicas de organização que vêm recebendo especial atenção no que diz respeito à representação formal do conhecimento. (MORAIS; AMBROSIO, 2007).

Passa-se a um melhor detalhamento sobre as ontologias Santarém Segundo e Coneglian (2015, p. 227) indicam que

> para o uso como tecnologia da *Web* Semântica, entende-se as ontologias como: artefatos computacionais que descrevem um domínio do conhecimento de forma estruturada, através de: classes, propriedades, relações, restrições, axiomas e instâncias.

Uma definição de ontologia é tratada por Ramalho (2006, p. 97), ao asseverar se tratar de

> Um artefato tecnológico que descreve um modelo conceitual de um determinado domínio em uma linguagem lógica e formal, a partir da descrição dos aspectos semânticos de conteúdos informacionais, possibilitando a realização de inferências automáticas por programas computacionais. Destacando assim o seu propósito e as novas possibilidades oferecidas no contexto da recuperação de informações.

A definição de ontologia esclarece ser esta uma explícita e formal especificação de uma conceitualização (GRUBER, 1993). As ontologias podem ser usadas na percepção do conhecimento de um domínio específico e para representar este conhecimento em entidades e relacionamentos.

Para Jacob (2003 apud SANTAREM SEGUNDO, 2010, p. 102)

> Ontologias são categorias de coisas que existem ou podem existir em um determinado domínio particular, produzindo um catálogo onde existem as relações entre os tipos e até os subtipos do domínio, provendo um entendimento comum e compartilhado do conhecimento de um domínio que pode ser comunicado entre pessoas e programas de aplicação.

Rotular recursos é algo essencial para uma recuperação adequada da *web*. Igualmente, é desejável que se defina tipos diferentes de relacionamentos entre estes recursos. Sobre a contribuição das ontologias na representação de informações na *Web*, Ramalho (2006, p. 59) estabelece que

> Desta maneira, verifica-se que a utilização de ontologias no âmbito da *Web* Semântica favorece o compartilhamento da mesma estrutura de informações entre pessoas e softwares, pois possibilita a descrição formal das relações existentes entre os objetos em um formato que as máquinas possam identificar, permitindo inclusive o reuso de conhecimentos dentro de um determinado domínio, pois torna possível, por exemplo, associar uma ontologia a uma página *Web*, definindo o significado de cada uma das informações existentes e possibilitando a integração e reutilização de ontologias entre diversos domínios. A partir desta perspectiva, uma página poderia ser relacionada automaticamente com outras, utilizando regras de inferência, e possibilitando inferir novas informações.

Assim, um dos principais objetivos das ontologias é a construção de conhecimentos interoperáveis e melhor estruturados. (MOREIRA; ALVARENGA; PAIVA, 2004). Tratamos de termos associados a textos (e objetos *web*) que descrevem o que os mesmos significam.

Segundo Breitman (2005, p. 7):

> ontologias são especificações formais e explícitas de conceitualizações compartilhadas. Ontologias são modelos conceituais que capturam e explicitam o vocabulário utilizado nas aplicações semânticas. Servem como base para garantir uma comunicação livre de ambiguidades. Ontologias será a língua franca da *web* semântica.

Sabe-se que o ponto principal da *Web* Semântica está na separação da apresentação e da estrutura do conteúdo disponível. Com as ontologias, pode-se organizar conteúdos de determinados domínios, como, por exemplo, da “fraude”, sendo tais recursos descritos e assim, proporcionando uma recuperação mais adequada da informação, facilitando associações e revelando informações até então ocultas.

As ontologias podem ser expressas usando modelos de representação como *Resource Description Framework* (RDF), linguagem para representar informações na *web* permitindo a interoperabilidade de dados e a recuperação das informações.

Enquanto um Schema-RDF permite relacionamentos básicos entre as propriedades e hierarquias, uma ontologia suporta muito mais expressividade e maiores relacionamentos como restrição de valores, restrições em propriedades de classes, dentre outras. A linguagem OWL *(Web Ontology Language)* é o padrão definido pela W3C. (DOSSIS, 2012). A Linguagem OWL também pode ser utilizada na especificação das ontologias. A OWL é uma linguagem para ser utilizada quando informações contidas em documentos *web* precisam ser passadas por aplicações em situações em que seu conteúdo precisa ser mais do que apresentado apenas para humanos, podendo ser usada para representar explicitamente o significado de termos em vocabulários e os relacionamentos entre os termos. (MORAIS; AMBRÓSIO, 2007). No contexto de ontologias, a linguagem OWL (*Web Ontology Language*) é considerada importante para a construção da *Web* Semântica, sendo avaliada como uma evolução das linguagens de descrição de conteúdos, fornecendo um vocabulário formal. Contendo elementos da linguagem RDF, avança, permitindo a definição de vocabulários para a descrição de classes e propriedades.

O modelo RDF é composto por objetos básicos denominados recursos, propriedades e declarações, sendo que os recursos são informações identificadas por um URI (*Uniform Resource Identificator*), as propriedades, as características do

recurso e a declaração a informação completa, envolvendo recurso, as propriedades e valores das propriedades.

Deste modo, identificou-se nesta pesquisa que as ontologias, estudadas no âmbito da Ciência da Informação, são indispensáveis para organização de uma classificação de conteúdos, no escopo de fornecer elementos para a detecção de falsidades. Do mesmo modo, a descrição dos recursos e metadados, por meio de linguagens como RDF, são fundamentais para organização das informações das ferramentas de redes sociais, possibilitando associações e interoperabilidade de aplicações para consulta dos dados.

Algumas propriedades do RDF como *"type", "subClassOf", "subPropertyOf" e "comment"* serão úteis para apresentar postagens e objetos *web* e suas relações com outros domínios envolvendo fraudes e falsidades.

Um exemplo seria uma postagem que fosse considerada tipo *(type*) de uma classe envolvendo fraude que por sua vez é uma subclasse (*subClassOf*) de outra classe relativa a um perfil falso, o que seria capaz de gerar a indução ou inferência de uma fraude ou de provável perfil falso, gerando esta informação para o usuário.

Logo, as ontologias, institutos estudados no âmbito da Ciência da Informação, desempenham papel fundamental nesta pesquisa. Conforme esclarece Uschold (1998, p. 12)

> Uma ontologia pode assumir vários formatos, mas necessariamente deve incluir um vocabulário de termos e alguma especificação de seu significado. Esta deve abranger definições e uma indicação de como os conceitos estão inter-relacionados, o que resulta na estruturação do domínio e restringe possíveis interpretações de seus termos.

É exatamente o que se pretende apresentar na presente pesquisa, classificando significados para objetos de ferramentas de redes sociais, postagens e links, com a possibilidade de detecção de associações com perfis falsos. Vocabulário, como se pode verificar, é essencial e básico, sendo primeira alerta, na detecção de expressões e postagens em ferramentas de redes sociais.

No modelo conceitual proposto na presente pesquisa, a concepção de tecnologias como ontologias e linguagem de descrição de recursos nas ferramentas de redes sociais proporcionará associação de dados e dedução de informações, momento em que tratamos de inferência e revelação de conhecimento sobre falsidades.

Neste contexto, identificou-se ainda que as ontologias são mais adequadas a este objetivo que os tesauros, considerando que estes, apesar de serem definidos por alguns autores como vocabulários controlados, não pretendem realizar na *Web* o que os tesauros fizeram na organização tradicional do conhecimento, por serem mais flexíveis e proporcionarem novos relacionamentos entre os termos (NICOLINO, 2014).

Aplicando-se, por fim, tais conceitos à proposta deste trabalho, é possível categorizar espécies de fraudes e postagens nas redes (ódio, financeiras, difamação), e pela inferência, avaliar perfis similares que possam ser considerados falsos. Pode-se igualmente categorizar objetos que possam revelar o uso indevido por autores de perfis falsos.

Como exemplo, podemos citar um objeto descrito em uma ontologia de fraude bancária que está relacionada ou associada a perfis falsos, sendo que outros objetos idênticos ou similares na rede poderiam gerar alerta de potencial perfil falso na rede. Outro exemplo seria o cadastro de uma foto ou vídeo categorizada como utilizada por um “robô” (aplicativo que simula uma aplicação nas ferramentas de redes sociais). Ferramentas poderiam considerar esta ontologia para identificar outros perfis usando a foto sinalizada, aclarando um indício que o perfil comparado também é um robô.

A partir das ontologias aplicadas aos conteúdos das ferramentas de redes sociais, neste contexto, serão utilizadas linguagens, agentes computacionais e protocolos, como SPARQL para pesquisa nos conteúdos gerados e publicados pelo sistema, com apoio das tecnologias da *Web* Semântica. O SPARQL é um conjunto de especificações que fornece linguagens e protocolos para consultar e manipular conteúdo publicado em RDF (SANTAREM SEGUNDO; CONEGLIAN, 2015). Igualmente para representar as relações entre os objetos de forma gráfica, pode-se utilizar ferramentas como GEPHI [5] ou plug-ins para construção de grafos disponíveis em ferramentas como Protégé[6], conhecido editor e construtor de ontologias, desenvolvido pela Universidade de Stanford.

[5] http://gephi.github.io/
[6] http://protege.stanford.edu/

2.2.2 Folksonomia

A *folksonomia* é entendida como uma das principais funcionalidades na construção coletiva de inteligência informacional. (SANTAREM SEGUNDO, 2010). A funcionalidade consiste em taggear ou etiquetar recursos ou recursos na *web*, basicamente visando sua recuperação. Está relacionada à *Web* 2.0.

Por outro lado, na proposta deste trabalho, a *folksonomia* seria utilizada na construção de um ambiente mais seguro, permitindo que organicamente, usuários da rede taggeassem perfis suspeitos. O apontamento de conteúdos por usuários permitiria o uso de ferramentas para recuperação, por exemplo, em uma pesquisa, de potencial conteúdo ilegal ou falsidades na rede social.

Etiquetar conteúdos ilícitos ou falsos auxiliaria no processo de identificação de contrafações, perfis falsos e outros. Outro ponto a se considerar são os comentários sobre determinado recurso. A análise dos comentários de postagens de um perfil, embora não possa ser considerada absoluta, demonstra-se importante na apuração da autenticidade do perfil. Na visão proposta neste trabalho e que norteia toda esta pesquisa, muito há de “conhecimento” em comentários espalhados pelas ferramentas de redes sociais, muitas vezes desconsiderados, que se denomina aqui de uma “*folksonomia* inconsciente”, pois os usuários de ferramentas de redes sociais contribuem para a classificação de objetos, sem perceberem que estão agindo nesta finalidade.

3 Big Data e coleta de dados em ambientes *Web*

Não se nega que a grande dificuldade na recuperação da informação é atender às expectativas dos usuários. Um sistema de recuperação de informação deve observar um modelo, que por sua vez influencia diretamente no modo de operação do sistema. (FERNEDA, 2003).

Os sistemas de armazenamento de dados estão crescendo de forma exponencial. Hinshaw (2004) já salientava que taxa de dados dobrava a cada nove meses, diga-se, taxa muito maior do que a da Lei de Moore. Em quatro anos, estima-se que tráfego IP Global anual será de 1,3 (zetabyte), o que equivale a 1 trilhão de gigabytes. Este tráfego de Internet deve crescer 29% até 2016. Neste ambiente, cresce a preocupação para com a preservação da informação, que é cada vez mais volátil na era da Internet e tecnologia da informação.

Ao contrário da corrente doutrinária que trata dos riscos envolvendo a exposição da privacidade com o advento da Internet e a persistência das informações na rede, Solove (2007), Mayer-Schroenberger (2011) e Ambrose (2013), apresentam importantes pesquisas relativas ao problema da não persistência da informação, ou seja, da rápida capacidade da informação “sumir”.

A pesquisa em “persistência da informação” debruça-se no estudo e é dedicada a mensurar por quanto tempo a informação permanece acessível e íntegra (isenta de modificações), sobretudo na Internet.

Kahle (1998) ao apresentar o projeto “Internet Archieve”, afirmou que a média de vida de uma página *web* é de aproximadamente 100 dias. O mesmo autor afirma que o futuro da *web* pode ser o mesmo da biblioteca perdida de Alexandria, se medidas como o Internet Archieve, não forem propostas.

Setenta e sete por cento (77%) do conteúdo *web* permanecem vivos após um dia, conforme pesquisa que realizou o download de 720.000 páginas de *web* servers populares, diariamente, por quatro meses para estudar como os documentos haviam mudado. Mais de 20% das páginas mudavam a cada download, que era diário. Estima-se que 50% do conteúdo *web* podem sumir após 100 dias. Outra pesquisa ainda aponta que 65% do conteúdo *web* continuavam vivos após uma semana. Em um ano, 10% do conteúdo *web* continuam vivos, conforme pesquisa analisada pela autora. (AMBROSE, 2013).

Gomes e Silva (2006) estudaram a persistência do conteúdo *web* de 2006 a 2007 e descobriram que 55% do conteúdo permanecem vivos após um dia, 41% após uma semana, 13% após 100 dias e 15% após um ano.

A cada minuto milhares de páginas *web* são atualizadas e abandonadas. Deste modo, a *web* não é permanente. Ao contrário daqueles que pensam que a *Web* eterniza informações, temos na verdade que a *Web* é efêmera e não pode se autopreservar. E o grande volume de dados (Big Data) é um fator que prejudica a recuperação de conteúdos *web* e de redes sociais antes que “evaporem”.

Perfis falsos podem ser criados para espalharem desinformações ou códigos maliciosos específicos e sumirem no dia seguinte. O processo de detecção de perfis falsos enfrenta necessariamente o desafio em lidar com uma grande quantidade de informações e que pode desaparecer rapidamente.

3.1 O desafio envolvendo o Big Data e a coleta de dados

Não se tem um conceito único de Big Data, Breternitz e Silva (2013), esclarecem que se pode- usar o termo para designar um conjunto de tendências que favorecem o tratamento e entendimento de grande volume de dados, segundo os autores, para fins de tomada de decisões. Big Data também é relacionado à ideia de grande volume de informações.

Assim, associa-se ao termo Big Data à disponibilidade, ao crescimento e ao uso exponencial de informações, sejam elas estruturadas ou não estruturadas.

Laney (2001) é considerado o primeiro a tratar das dimensões do Big Data, apresentando o termo “3vs do Big Data”, o qual designou como “volume”, “velocidade” e “variedade”.

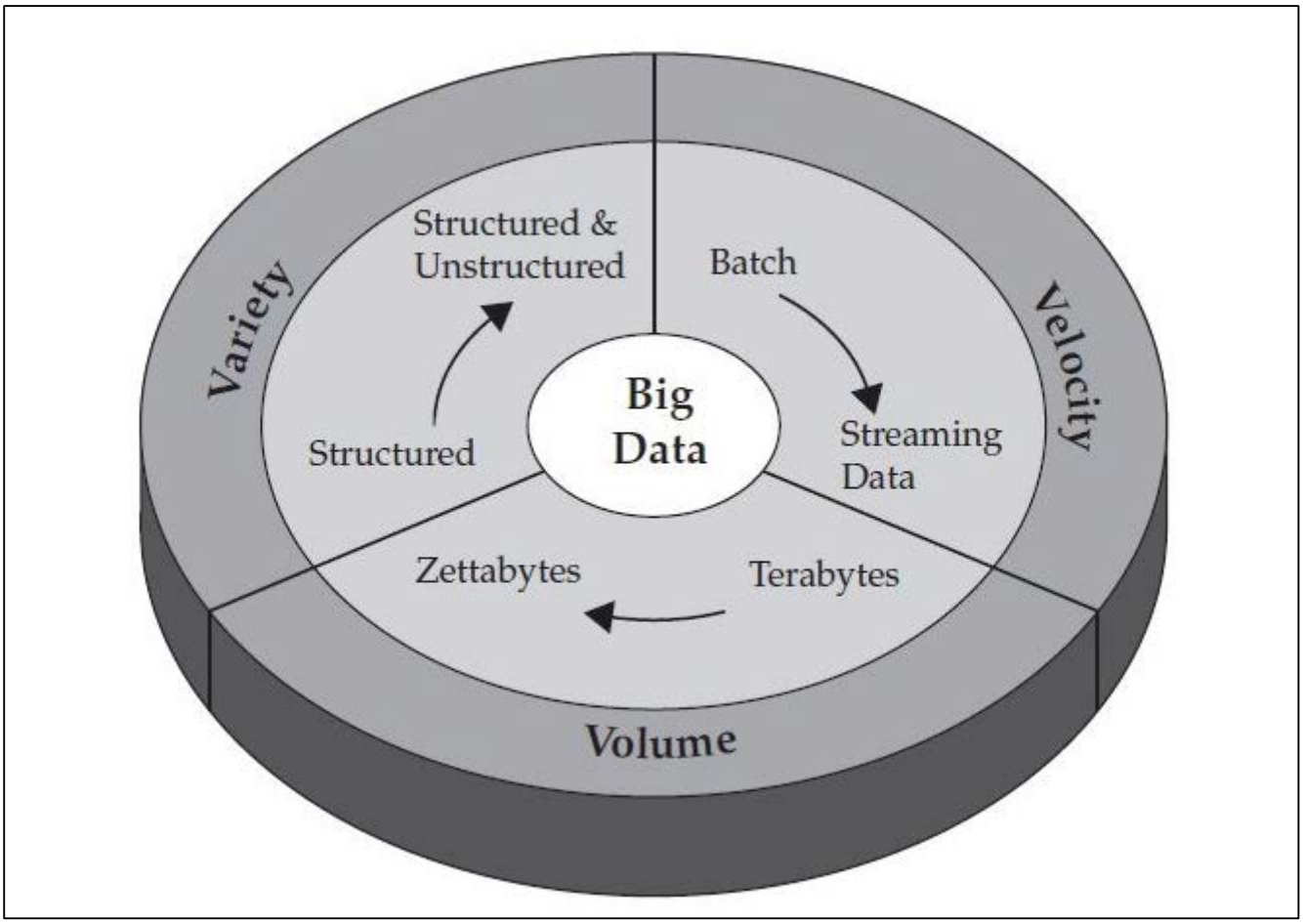
Ao tratarmos de “volume”, apresenta-se a enorme quantidade de informações geradas com as novas tecnologias. Muitos fatores são considerados os responsáveis deste volume, como a custódia de transações, dados e textos em ferramentas de redes sociais, aumento da quantidade de sensores de dados, dentre outros. A grande preocupação é como armazenar ou determinar a relevância dos dados neste cenário de extremo volume. (LANEY, 2001).

O termo “velocidade”, uma dimensão de Big Data, refere-se ao estudo da velocidade em que os dados são produzidos e quão rápido deve ser o seu

tratamento, não só para atendimento de demandas, mas para que não se “percam”. Reagir rápido é preciso. (LANEY, 2001).

Já a dimensão “variedade” implica em um dos principais desafios impostos com o advento do Big Data: lidar com formatos heterogêneos de dados. Imagens, sons, textos, bases de dados, e-mails, dados de sensores, compõem apenas alguns formatos que demonstram a variedade do Big Data. (LANEY, 2001).

Figura 3 – Os 3 “vs” do Big Data



Fonte: Disponível em:
https://a.disquscdn.com/uploads/mediaembed/images/656/7545/original.jpg

Ao tratar de Big Data, Zikopoulos et al. (2012) além das dimensões já estudadas, apresenta mais um “v”, que nomina de “veracidade”. A problemática da “veracidade” envolve, especificamente, lidar com dados imprecisos, sendo diariamente gerados em diversas fontes, problema da presente pesquisa.

Além das dimensões estudadas, podem-se identificar como outras dimensões do Big Data[7] a “variabilidade”, em que fluxos de informações podem variar em picos periódicos, como eventos sazonais e a “complexidade”, dimensão que estuda a problemática de os dados que compõe o Big Data se originarem de diversas fontes, o que implica em um grande esforço para se determinar.

Importantes contribuições são encontradas nos estudos de McAfee e Brynjolfsson (2012) a respeito do tema, inclusive a constatação de que empresas que utilizam o Big Data são 5% mais produtivas e 6% mais lucrativas. Os mesmos autores apontam que cerca de 2,5 *exabytes* de dados são criados a cada dia e que o número iria dobrar em aproximadamente 40 (quarenta) meses.

A exploração adequada do Big Data permite uma revolução no processo decisório. Mas, explorar adequadamente um cenário com grande variedade de formatos é realmente um desafio na atividade de investigação, detecção de fraudes, computação forense e inteligência cibernética. Como esclarece Lohr (2012), além do modelo convencional de tomada de decisão, onde a organização filtra dados de vários sistemas, tem-se também em alguns casos a necessidade do emprego da Inteligência artificial, onde se destacam os estudos envolvendo reconhecimento de padrões, processamento de linguagem natural e aprendizado de máquina, capazes de extrair significado de grandes volumes de dados para a tomada de decisões.

Os serviços de inteligência têm por missão a obtenção de informações estratégicas. Tais informações podem significar uma demanda social ou mesmo uma ameaça à segurança ou à integridade de pessoas e serviços. Identificar a origem de crimes informáticos é também desafio dos serviços de inteligência, como por exemplo identificar o responsável por um perfil difusor de códigos maliciosos na Internet.

A inteligência voltada para os meios eletrônicos pode ser conceituada de “inteligência cibernética” (WENDT, 2010). Montanaro (2014) explica que “inteligência” dirige operações, tratando-se de um mantra conhecido no militarismo, relacionado ao planejamento sobre o alvo, ecossistemas e afins. Já a “inteligência cibernética” monitora e analisa ameaças que surgem dentro do espaço cibernético e que podem causar danos, sendo responsável por antever fraudes e vulnerabilidades.

---

[7] http://www.sas.com/pt_br/insights/big-data/what-is-big-data.html

Como as atividades afetam a inteligência cibernética, podem-se antecipar cenários, aumentar a segurança de pessoas, e evitar fraudes digitais. Sabe-se que os índices de criminalidade cibernética vêm aumentando no mundo, o que demanda apuração ágil dos responsáveis. Neste sentido, a computação forense é ciência dedicada à coleta, à preservação e à análise de evidências que tenham relevância a uma investigação. Tem como escopo transformar a informação em evidência da prática de ilícitos, golpes e fraudes (FARMER; VENEMA, 2007). Tal disciplina também se vale da inteligência cibernética para identificar ações ou mesmo reconstruir o que ocorreu com sistemas comprometidos. As áreas de inteligência cibernética e computação forense hoje se deparam com a problemática de processar grandes volumes de dados, na identificação de fraudes, *fakes*, falsidades e autores de crimes cibernéticos. Os peritos trabalham com quesitos apresentados por autoridades, devendo respondê-los para esclarecimento de casos. Norteado pelos quesitos, os peritos irão desenvolver o trabalho de análises forenses. Ocorre que é muito difícil analisar milhares de registros em pouco tempo, para resposta aos quesitos.

Assim, tem-se hoje, como principal desafio à computação forense e inteligência cibernética, a análise de “Big Data”, grande volume de dados dispostos em dispositivos informáticos e na Internet. (BEEBE; CLARK, 2005). A grande maioria dos procedimentos, hoje, ignora o “Big Data” como preciosa fonte de informação estratégica na segurança da informação e na tomada de decisões em geral. A falta de procedimentos específicos tem contribuído para a realização de análises superficiais, pois soluções que tentam analisar “Big Data” são simplistas e não aplicam significado aos dados brutos coletados, dependendo da atuação do sujeito cognoscente, ser humano absolutamente limitado. E o desafio não está só na análise, mas na coleta dos dados.

Como dito, quando tratamos de Big Data, existem desafios em duas grandes áreas para que se possa identificar um perfil falso, sendo elas: a) A coleta; b) A análise de dados.

Na coleta, o grande volume de informações e a velocidade com que são geradas demandam das aplicações e *frameworks,* técnicas de coleta em tempo real, evitando o perecimento da informação;

Já no desafio envolvendo a ***análise***, lida-se justamente, com a variedade de formatos dos dados, sendo que para cada formato, deve-se propor uma forma

especial de tratamento. Mais que isso, no contexto de detecção de falsidades, lida-se com a problemática em se “atribuir significado” a um *data-set* de informações. E sem descrição não se pode cogitar extração de “significado”.

As atuais ferramentas de computação forense ou inteligência cibernética não estão preparadas para manipular *terabytes* de informações de uma maneira eficiente, mesmo com um *data-set* moderado de 200 *gigabytes*, por exemplo, as ferramentas de extração se demonstram fracas e ineficientes e o tempo de processamento para pesquisas por *keywords* pode levar dias. (BEEBE; CLARK, 2005).

Evidencia-se assim, a necessidade de estudos de novas técnicas e métodos para a recuperação da informação em ambientes de grandes volumes de dados, como a *Web*. Santarém Segundo (2010), ao descrever os modelos de recuperação da informação, estabelece que dentro da dimensão da Internet está evidente o esgotamento de alternativas com relação aos modelos conhecidos. De fato, a *web* é uma imensa base de dados onipresente e desestruturada. (BAEZA-YATES; RIBEIRO NETO, 1999).

Do mesmo modo, apurar a falsidade em ferramentas de redes sociais não é um desafio simples. Neste contexto, é evidente a carência de pesquisas relacionadas ao tratamento de grandes repositórios de dados, com a finalidade de extração de significados que possam detectar falsidades em identidades. As soluções que pretendam investigar esta temática deverão considerar desafios impostos na era dos *terabytes*, sobretudo nas frentes envolvendo “coleta” e “análise” dos dados.

A questão da falsidade em redes sociais oferece riscos imensuráveis à sociedade da informação. Difamação, estelionato e extorsão são apenas alguns exemplos dos riscos decorrentes da criação destes perfis. Algumas ferramentas como *Identify Badge* e *MysafeFriend*[8] ajudam a validar solicitações de perfis falsos nas redes sociais. Porém não são suficientes.

Até hoje, todas as análises para detecção de perfis falsos valem-se de critérios subjetivos e análises humanas. Agregar conhecimento a dados nos parece, sempre foi atividade desenvolvida por pessoas com pouco auxílio da tecnologia. Os

[8] https://vimeo.com/28954841

reportes de um perfil *fake*, a exemplo, comumente são manuais e baseados em achismos e presunções[9].

Ravenscraft (2014) estabelece as formas disponíveis para se detectar um perfil falso nas redes sociais sendo elas: a) sistemas automatizados; b) reporte de usuários; c) investigadores do Facebook (ferramenta de rede social). O autor, no entanto, deixa claro que não se sabe quais os critérios usados pelas redes sociais para tais detecções. O resultado é inúmeros falsos positivos e negativos e pessoas reais tendo problemas com seus perfis nas redes sociais.

Ademais, as formas citadas muitas vezes não estão disponíveis ao usuário convencional em busca de critérios para se verificar a autenticidade de um perfil.

De maneira que, não se tem dúvidas de que o crescimento das redes sociais também trouxe em seu bojo o crescimento das atividades criminosas, fraudes, golpes, falsidades. Basicamente, todas as ferramentas de redes sociais permitem a criação de *profiles* (ou perfis) que servem para conectar usuários. Estes perfis podem ser usados por criminosos para rastrear pessoas e para espionagem.

Os estudos de Athanasopoulos et al. (2008) apresentam propriedades intrínsecas que fazem das ferramentas de redes sociais um local ideal para exploração por adversários, sendo elas a alta distribuição de usuários, grupos ou clusters de usuários com os mesmos interesses e principalmente, plataforma aberta para aplicações que podem ser usadas para atividades fraudulentas. Identificado nas pesquisas que usuários não checam as aplicações que anexam a seus perfis nas redes sociais.

No mesmo sentido, é evidente a lacuna envolvendo métodos padronizados para investigações em ferramentas de redes sociais. Técnicas de investigação, não se valendo dos aportes da Ciência da Informação, se tornam obsoletas rapidamente, com a evolução da tecnologia. É neste contexto, que identificamos a importância da Ciência da Informação na construção de futuros métodos e mesmo para a criação de *frameworks* que se destinem à finalidade de extrair conhecimento de grandes volumes de dados, incluindo, mas não se limitando, à detecção de falsidades.

Resta evidente que pesquisas que pretendam enfrentar a problemática do Big Data, considerando apenas institutos da Computação, como Data Mining, não resolverão a problemática considerando que esta tecnologia se aplica apenas a

[9] http://www.technologyreview.com/news/535901/*fake*-persuaders/

dados estruturados, o que é bem distinto dos dados dispostos em ferramentas de redes sociais, sem um padrão de organização de dados. Algumas pesquisas sobre o tema já concluem neste sentid.

Beebe e Clark (2005) apresentam a importância do *data mining* para se lidar com grandes volumes de dados em investigações. A pesquisa é mais focada em *"data-sets"* de dados já coletados. Mas a coleta é um desafio que precisa ser enfrentado. De fato, ao se buscar por falsidades e perfis falsos na rede, o primeiro ponto a se pensar diz respeito à coleta destas informações, que estão a trafegar na rede em volume infindável.

Para os autores em comento o

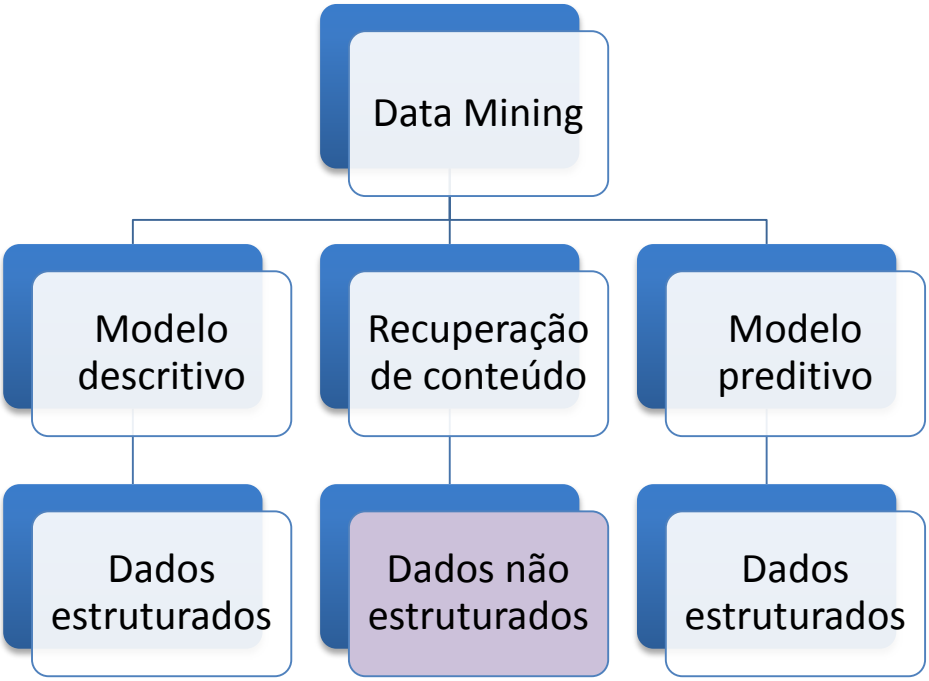
> *Data mining* incorpora uma abordagem multidisciplinar para encontrar e recuperar informação e está assentada em disciplinas referenciais que gozam de ricas correntes de pesquisa incluindo a matemática, estatística, Ciência da Computação e Ciência da Informação. (BEEBE; CLARK, 2005, p. 4).

Dentre as técnicas que podem ser utilizadas na análise de grandes volumes de dados, podemos citar a inteligência artificial, aprendizagem de máquina, reconhecimento de padrões, dentre outras.

O trabalho dos autores citados classifica o *data mining* em descritivo e preditivo, sendo que no primeiro modelo, haveria a sumarização de dados e no segundo, teríamos a identificação de características que podem auxiliar na predição do futuro. (BEEBE; CLARK, 2005).

Todas as técnicas acima citadas dependeriam de dados estruturados (bancos de dados, documentos descritos XML, documentos e gráficos). Porém sabemos que no Big Data muitas das informações estão em formato desestruturado. Neste sentido e para fazer frente a esta problemática, Beebe e Clark (2005), apresentam o conceito de *content retrieval* (recuperação de conteúdo) que são técnicas usadas em dados complexos e comumente desestruturadas, como dados *Word Wide Web*, imagens, links, vídeos, dentre outros.

Figura 4 – Data-mining em grandes volumes de dados



Fonte: Elaborado pelo autor com base em Beebe e Clark (2005)

O modelo descritivo de data-mining estaria baseado na generalização e conceitualização para gerar descrições e facilitar a caracterização (sumarização) e comparação (descriminação dos dados). O objetivo é transformar longos *data-sets* em menores e mais significativos. As técnicas de comparação seriam utilizadas na produção de regras para comparar características gerais de objetos encontrados em duas compilações de dados distintos, como por exemplo, comparar as características de quem compra mais pela internet com as características gerais de quem compra menos pela Internet. (BEEBE; CLARK, 2005).

O modelo preditivo se subdividiria em três subclasses, sendo elas: associação baseada em regras, regressão e classificação. Na associação baseada em regras busca-se identificar relacionamentos ou “associações” entre itens. De maneira que dado um número de itens esta regra é capaz de predizer a ocorrência de outro item, com base na ocorrência de outros itens, calculando-se a probabilidade de coocorrerem. A classe de regressão consiste na observação dos dados que, modelados, podem gerar inputs de predição com base em funções matemáticas. Já a última classe, que chamamos de classificação, pode ser utilizada tanto na investigação preditiva como na descritiva, e se vale de métodos de classificação de

dados como *decision tree induction, bayesian classification/belief networks, neural networks, nearest neighbor classification, genetic algorithms, case-based reasoning, rought sets e fuzzy logic.* (BEEBE; CLARK, 2005).

Mas é no modelo Content Retrieval de *Data mining* que encontramos os maiores desafios, considerando que em análises de redes sociais, lida-se, comumente com dados não estruturados ou semiestruturados. E neste momento os autores chama a atenção para a importância da Ciência da Informação. Para Beebe e Clark (2005, p. 7)

> Considerando que as técnicas de mineração de dados descritiva e preditiva em largamente utilizam disciplinas de referência matemática e estatística, content retrieval depende fortemente da pesquisa em Ciência da Informação e Ciência da Computação, particularmente nas áreas de recuperação de informação, inteligência artificial, aprendizado de máquina e processamento de linguagem natural.

Este é o modelo de mineração de dados que mais se enfrenta ao analisar ferramentas de redes sociais, considerando os limites impostos por estas redes no acesso a dados estruturados, por vezes exigindo coletas de informações sem qualquer estrutura ou limitando dados estruturados a poucas informações, considerando todas as implicações de privacidade.

Os autores em estudo estabelecem a distinção entre conteúdos textuais e conteúdos multimídia, onde apresentam a relevância das buscas semânticas e dos metadados em ambas as abordagens.

Importa dizer que toda a técnica de classificação apresentada por Beebe e Clark (2005) impõe a fase de aprendizagem, e é possível definir classes para os dados de forma supervisionada e manual e até mesmo não supervisionada, como no caso da *cluster analysis*. Na detecção do crime ou fraude, os autores concluem que ambos os modelos, descritivos e preditivos, podem ser utilizados.

Como se verifica na pesquisa analisada, o data mining não é totalmente adequado para lidar com dados desestruturados, sendo a Ciência da Informação elemento indispensável na construção de soluções com este objetivo. Após a estruturação das informações, papel da Ciência da Informação, então sim, pode-se cogitar da aplicação de data mining e outras tecnologias computacionais.

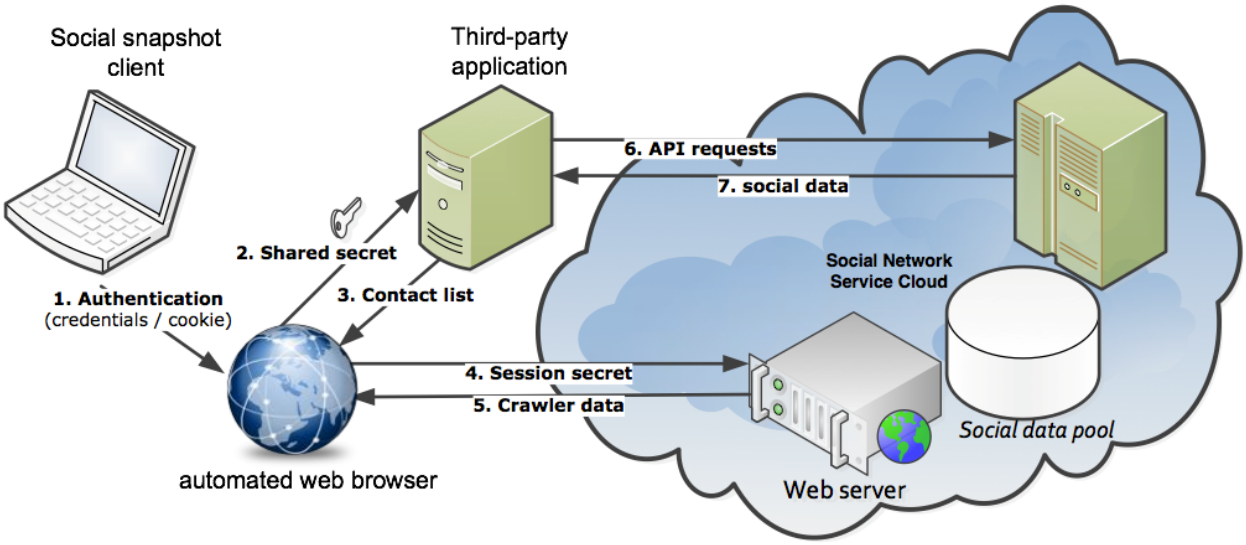
Zainudin, Merabti e Llewellyn-Jones (2010) apresentam um processo para a condução de perícias e investigações em ferramentas de redes sociais. O processo

basicamente replica outros procedimentos já estabelecidos e não descreve o “como realizar”. No entanto, os autores deixam clara a lacuna de padrões e *frameworks* para a área, fato que registramos neste trabalho (ZAINUDIN; MERABTI; LLEWELLYN-JONES, 2010).

Evidências em ferramentas de redes sociais podem ser usadas como prova de um crime, fraude ou delito. Enquanto os métodos tradicionais de computação forense e investigação são baseados na análise de sistemas de arquivos e tráfego de rede capturado, novas abordagens para extração de dados de redes sociais são necessárias. Muitos dos métodos atuais de coleta de dados em redes sociais ainda são baseados em “*webcrawling*” (técnica para se clonar páginas *web*) o que se demonstra inviável para a efetiva coleta, considerando a quantidade de tráfego, características distintas, bem como os dados ocultos que não são coletados pelos *webcrawlers*, além da estrutura dos *web*sites que mudam constantemente. (HUBER et al., 2011).

No que tange à coleta de dados, Huber et al. (2011) desenvolveram um método para investigação forense em redes sociais denominado “Social Snapshot”, que representa a atividade na rede social de um usuário específico, troca de mensagens, fotos, etc. A proposta é apresentada no seguinte diagrama

Figura 5 – Fluxo do projeto Social Snapshot



Fonte: HUBER et al. (2011, p. 3)

A proposta apresentada, porém, não é focada na detecção de perfis falsos e precisa de *"token"* de autenticação de um usuário que seria monitorado ou investigado. Os autores deixam claro que o consentimento seria uma forma de acesso ao *token* (credencial de acesso) da pessoa cujas atividades serão analisadas. Outra forma seria o denominado "*Hijack social networking sessions"* onde seria possível descobrir *tokens* trafegando, por exemplo, em redes Wi-Fi e lans desprotegidas e a partir disso, instalar o aplicativo. Outra forma ainda seria o acesso ao computador do suspeito mediante extração física de cookies no disco rígido. (HUBER, et al., 2011). A aplicação instalada em um perfil voluntário varreria todos os contatos deste voluntário, coletando informações não acessíveis convencionalmente via *crawling* ou download direito por meio da função disponibilizada pelo Facebook. Os dados coletados são apresentados a seguir, o que demonstra ganhos importantes na recuperação da informação, em comparação com formas convencionais de acesso a dados, como o "*Download*".

Figura 6 – Dados coletados pelo Social Snapshot em comparação com a funcionalidade de download do Facebook e *WebCrawling*

Element	Download	social snapshot
Contact details	–	✓Crawler
News feed	–	✓Graph API
Checkins	–	✓Graph API
Photo Tags	–	✓Graph API
Video Tags	–	✓Graph API
Friends	name only[a]	✓Graph API
Likes	name only[a]	✓Graph API
Movies	name only[a]	✓Graph API
Music	name only[a]	✓Graph API
Books	name only[a]	✓Graph API
Groups	name only[a]	✓Graph API
Profile feed (Wall)	limited[b]	✓Graph API
Photo Albums	limited[b]	✓Graph API
Video Uploads	limited[b]	✓Graph API
Messages	limited[b]	✓Graph API

[a] No additional information available.
[b] Missing meta-information such as UIDs.

Fonte: HUBER et al. (2011, p. 5)

Como se pode constatar, o acesso a dados de redes sociais por meio de aplicativos constitui-se mais efetivo do que o acesso via *"webcrawlers".*

Os pesquisadores esclarecem, no entanto, que "a interface padrão *web* não fornece informação se a conta de um usuário é verificada [...]". (HUBER, et al., 2011, p. 7, tradução nossa). Assim, o projeto "Social *Snapshot*" constitui-se uma aplicação a ser rodada a partir de um perfil *"loggado",* permitindo coletar informações dos contatos deste perfil, mas também servir para que os próprios usuários tenham a dimensão da quantidade de informações publicadas na rede social. Caso não tenha acesso à sessão do usuário, tais sessões poderão ser tomadas, por meio de técnicas e aplicações, onde inclusive já existem provas de conceito a respeito, sendo elas nos códigos *Firesheep*[10] e *FaceNiff*[11].

O *Firesheep* é uma aplicação para computadores que tem a capacidade de sequestrar sessões de usuários do Facebook, que utilizem, por exemplo, redes wireless desprotegidas. Já o *FaceNiff* é a versão similar ao *Firesheep* para Android, tendo a capacidade de capturar sessões de usuários do Facebook que estejam conectados na mesma rede wireless aberta e que não encriptem seus dados. (HUBER, et al., 2011).

Os autores concluem a pesquisa informando que, das abordagens pesquisadas, desconhecem alguma que foi desenhada para comparar imagens na rede social a partir de um disco suspeito.

Leung et al. (2009) apresentam o projeto "*iCloner*" de um sistema que coletaria dados de um dado usuário em ferramentas de redes sociais. A proposta, inicialmente com base em uma conta específica, coleta estes dados e verifica se os mesmos são encontrados em outras redes sociais. Caso não seja encontrado o *framework* cria um perfil nestas redes. Por fim, ocorre o disparo de solicitações de amizades a amigos identificados na rede social base.

O sistema, no entanto, parte do pressuposto de um perfil *fake* conectado na rede social e que tenha estabelecido contato primário com a vítima. Para a coleta de dados da vítima, os autores usaram a linguagem Python, tendo em vista a existência de uma biblioteca HTTP, pronta para interagir com Facebook, não só coletando informações, mas já classificando, no conceito denominado "*parsing".* Com isso, conseguiram coletar as requisições de amizade de um usuário.

---

[10] https://codebutler.github.io/firesheep/
[11] http://faceniff.ponury.net/

Os dados são então remetidos a um banco de dados.

A proposta utiliza os denominados “*crawlers*”. Os *crawlers* têm em sua essência a capacidade de navegar por meio de páginas HTTP e coletar hyperlinks em uma página. O crawler então acessa os links que continuam repetindo a tarefa enquanto encontrar links nas páginas. O grande diferencial da pesquisa é que adapta o conceito de *crawling*, para que sejam coletados perfis que o usuário investigado estaria interessado. (LEUNG, et al., 2009).

Dessa forma, é preciso especificar um nome como parâmetro do *crawling*,sendo que a partir de então, o sistema procura pelo nome salvo em uma lista. O sistema irá identificar todos os amigos deste usuário e salvar em uma lista. Outros parâmetros interessantes a serem identificados no sistema podem ser os interesses do usuário em grupos e *fan pages*. A partir do momento que se tem as listas começam as requisições de amizade do perfil “*fake*” gerado, com intervalo de 10 a 20 segundos de espaço entre uma requisição e outra para o Facebook não detectar o robô e bloquear a aplicação.

Destaca-se que “*fake*” é um vocábulo do idioma inglês que significa falso, falsificado ou falsificação. Também é empregado para designar perfis falsos criados na Internet para ocultar a identidade real de um usuário.

Após este pedido de amizade, o *crawling* continua agindo e coletando dados dos amigos do perfil, que são inseridos em um banco de dados, momento em que, no trabalho dos autores, são apresentadas informações coletadas como nome, relacionamento, aniversário, etc.

Na pesquisa de Leung et al. (2009), fora criado um perfil falso de uma jovem, com foto atrativa e solteira, e realizadas 200 solicitações de amizade, sendo que 130 retornaram positivas, ou seja 65% das solicitações.

Figura 7 – Perfil falso criado para o projeto

Figure 4: Profile Picture and Information

Fonte: LEUNG et al. (2009, p. 3)

Com isso verifica-se o risco que é a aceitação de amizade com pessoas desconhecidas nas redes sociais, eis que para coletar dados de um perfil alvo para a criação do perfil falso, basta ser amigo deste. Para um futuro, os autores advertem para o risco da criação de um *"pool of fake profiles"* que poderia processar o crawler em várias contas nas redes sociais e que poderia criar mais links e relacionamentos entre usuários alvos. (LEUNG, et al., 2009). Os desafios futuros do trabalho estão em contornar restrições do Facebook, como o recurso de segurança captcha, capaz de detectar ataques automatizados.

Pelo projeto em estudo, foi possível processar scripts para adicionar amigos e amigos de amigos e consequentemente, ingressar no "ciclo de amigos mútuos", sendo possível coletar informações estatísticas como idade e relacionamentos destes usuários.

Os pesquisadores não apresentam o protótipo para análises. No entanto, embora se trate de uma pesquisa que demonstra o êxito na clonagem de um perfil, a ideia pode ser utilizada, por exemplo, na investigação e detecção de falsidades, para descobrir um ciclo de amigos de um *fake* e aplicando cálculos probabilísticos, apurar o responsável pelo perfil falso ou pessoa real muito ligada ao *Fake*.

De outra ordem, a proposta pressupõe a criação de um perfil falso que irá processar um perfil alvo e partir dele, levantando o número de amigos, disparar

solicitações de amizade. A proposta esclarece a alta taxa de aceitação de amizade diante de amizades em comum, o que pode ser utilizado em inteligência cibernética.

3.2 Reflexões sobre as questões legais e privacidade

Tanto as técnicas expostas por Beebe e Clark (2005), tanto o projeto Social Snapshot e iCloner destinados à fase inicial do processo de detecção de perfis falsos, que é a fase de coleta, precisam de reflexões sob o prisma da legalidade.

É sabido que apenas autoridades legais possuem acesso a algumas informações sobre usuários das ferramentas das redes sociais (por meio de acessos especiais)[12]. Mais que isso, o processo de computação forense, em grandes volumes de dados na detecção de perfis falsos, deve observar a proteção dos dados pessoais, estampada no Marco Civil da Internet.

Neste sentido, dispõe a Lei 12.965/2014, o Marco Civil da Internet Brasileira que em seu Art. 7º

> O acesso à internet é essencial ao exercício da cidadania, e ao usuário são assegurados os seguintes direitos:
> II - inviolabilidade e sigilo do fluxo de suas comunicações pela internet, salvo por ordem judicial, na forma da lei;
> III - inviolabilidade e sigilo de suas comunicações privadas armazenadas, salvo por ordem judicial; [...]. (BRASIL, 2014).

Deve-se destacar igualmente que o Marco Civil da Internet protege os dados pessoais, assegurando o direito dos usuários de não terem seus dados revelados a terceiros, inclusive registros de conexão ou acesso à aplicação. Ressalvadas as hipóteses previstas em lei ou ordens judiciais.

Do mesmo modo, a interceptação não autorizada de comunicações é conduta incriminada no Brasil, nos termos da Lei 9.296/1996. (BRASIL, 1996). Ferramentas que analisam dados e falsidade em redes sociais devem considerar este ambiente regulatório. Não se pode, no Brasil, sequestrar a sessão de um usuário (perfil) para lá adicionar aplicativo que coletaria seus dados ou de seus amigos.

Como se verifica, desafio que antecede a análise de grandes volumes de dados nas redes sociais é coletar um set de informações e contornar as restrições das próprias redes sociais, além de manter-se em legalidade. No Brasil, ainda, o

[12] A exemplo, o serviço do Facebook acessível em: https://www.facebook.com/records/x/login/

acesso indevido a perfis suspeitos pode encontrar restrições na Lei 12.737/2012 (BRASIL, 2012), a Lei Carolina Dieckmann, que pune a “invasão de dispositivo informático”, com pena de detenção de três meses a um ano e multa:

> Art. 154-A. Invadir dispositivo informático alheio, conectado ou não à rede de computadores, mediante violação indevida de mecanismo de segurança e com o fim de obter, adulterar ou destruir dados ou informações sem autorização expressa ou tácita do titular do dispositivo ou instalar vulnerabilidades para obter vantagem ilícita:
> Pena - detenção, de 3 (três) meses a 1 (um) ano, e multa. (BRASIL, 2012).

Portanto, propostas que visem à coleta de dados de ferramentas de redes sociais para avaliação ou apuração se perfis reais ou falsos, deverão considerar a coleta de dados públicos e sem restrições de privacidade, ou mesmo, em determinados casos, o consentimento do investigado com a permissão para instalação de aplicações.

3.3 Ferramentas e recursos análise de dados na detecção de perfis falsos

É sabido que o combustível do Big Data são as redes sociais (CONVERGÊNCIA DIGITAL, 2012), foco, aliás, desta pesquisa. Normalmente, as ferramentas de redes sociais são ambientes *web* fechados, com controle de *logins* e *sessions*, de maneira que só se coleta informação quem tem uma conta devidamente criada na rede.

Ainda assim, os dados são limitados a sessões, rolagens e recursos disponíveis. Coletar dados em ferramentas de redes sociais de forma nativa ou via “*web crawling*” é complexo, pois os recursos *web* hoje envolvem tecnologias como Java, Flash, Ajax, PDF-A e outros formatos que prejudicam a recuperação da informação com base em pesquisas textuais por expressões regulares.

Como visto anteriormente, não se pode invadir um perfil suspeito em busca de informações sobre o mesmo.

Neste ponto do trabalho, apresenta-se um levantamento feito de principais formas e ferramentas para acesso a grandes volumes de dados dispostos nas ferramentas de redes sociais, igualmente, são apresentados alguns protótipos e projetos para detecção de perfis falsos e *bots* nestas redes.

3.3.1 APIS

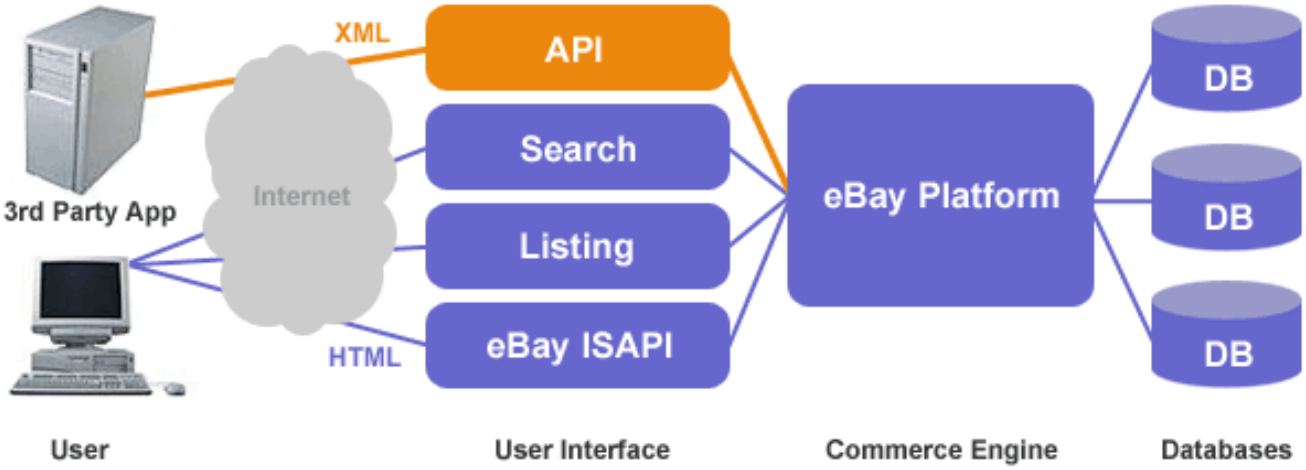
Application Programming Interface (API). Pode-se classificar APIs como a interface que existe entre uma “rede de dados” e as aplicações que a acessarão. Também podem ser chamadas de “*Web* APIS”. Por intermédio APIs, muito utilizadas em “*web services*” (solução para integração de sistemas e na comunicação de aplicativos diferentes), normalmente ocorrem requisições HTTP (*Hypertext transfer protocol*) que devolverão mensagens usualmente em XML (*Extensible Markup Language*) ou JSON (*JavaScript Object Notation*).

A finalidade prática das APIs é proporcionar o compartilhamento de conteúdo. Assim, uma API pode ser entendida como um conjunto padrão de instruções computacionais para acesso a uma aplicação *web-based*, ou ferramenta online. (ROSS, 2007).

Um exemplo claro de como a API funciona pode ser visualizado na figura 8. Usuários convencionais utilizam as ferramentas de busca e pesquisas dispostas pela aplicação *web* para minerar informações em suas bases de dados. Logicamente que estas pesquisas são limitadas ao que o programador da aplicação definiu, especialmente quando tratamos de ferramentas de rede sociais e seu grande volume de dados.

Porém, por meio de uma API, pode-se construir uma aplicação que acesse diretamente a base de dados da aplicação, de acordo com a documentação, existindo a possibilidade de um retorno rico de informação, comumente não retornável nas buscas convencionais (interfaces). Vejamos no exemplo da figura 8, em que um usuário convencional busca a informação no sistema “E-BAY” e outra aplicação faz o mesmo, por meio de uma API.

Figura 8 – Esquema de funcionamento de uma API



Fonte: Disponível em: http://developer.ebay.com/images/api-flow.gif

Pela figura 8, visualizamos que um usuário convencional utiliza as interfaces disponibilizadas pelo proprietário da plataforma ou serviço para desenvolver suas pesquisas na base de dados. Já uma aplicação de terceiro pode "contornar" as limitações destas interfaces, acessando diretamente a plataforma e os bancos de dados, logo, conseguindo acesso mais rápido a um volume maior e mais detalhado de informações.

Assim, as APIS são importantíssimas interfaces na recuperação de grandes volumes de dados nas redes sociais, eis que contornam limites oferecidos a usuários convencionais no acesso às informações. Com este acesso, é possível, por exemplo, copiar "*data-sets*" para posterior tratamento e análise.

As principais redes sociais utilizadas no Brasil possuem APIs (*webservices*), sendo elas:

a) Facebook[13]
b) Twitter[14]

Sabe-se, porém, que as regras de privacidade dos países têm proporcionado mudanças nas APIs das grandes ferramentas de redes sociais, com restrições a acessos não consentidos ou dados considerados privados, como preferência religiosa e sexual.

[13] https://developers.facebook.com/docs/graph-api
[14] https://dev.twitter.com/rest/public

No Brasil, o Projeto de Lei Proteção de Dados Pessoais, PL 330/2013 (BRASIL, 2013), recentemente aprovado na Comissão de Constituição e Justiça, elenca dentre direitos básicos dos usuários:

> Art. 7º São direitos básicos do titular de dados:
> V – o consentimento prévio como requisito à coleta e ao tratamento de dados pessoais sensíveis, bem como à interconexão internacional de dados realizada por banco de dados privado (art. 10);

Por fim, conclui-se que a criação de APIs, para que aplicações diversas se comuniquem com a plataforma de um serviço de dados, não é obrigação do titular da plataforma, ou seja, o proprietário da ferramenta de redes sociais não é obrigado a implementá-la. Logo, muitos repositórios de grandes volumes de informações não disponibilizam tais interfaces, essenciais para ferramentas que busquem coletar grandes volumes de dados para análises e investigações.

3.3.2 *Crawling*

Como visto anteriormente, as APIS são importantíssimas interfaces na recuperação de grandes volumes de dados nas redes sociais, eis que contornam limites oferecidos a usuários convencionais no acesso às informações. As APIs também permitem recuperar grandes volumes de dados, eis que os dados comumente são acessados em formato simples e nativo, sem padrões de estilo e outros formatos da interface do usuário, que tornam a recuperação lenta.

Muitas vezes os grandes repositórios de informações restringem suas APIs ou não possuem uma interface de acesso aos dados. Em situações extremas, a única forma disponível para a coleta de dados é a utilização de *crawlers* ou *spiders* que podem realizar a cópia de conteúdo *web* para posterior análise.

Alguns *crawlers* já varrem as páginas por links e já os inserem em sua lista para que sejam copiados também. As grandes limitações dos *crawlers* são que muitas redes sociais possuem mecanismos anti-robô, o que inviabilizam o processamento dos mesmos. Além disso, os dados coletados via "*web crawler*" são desestruturados, variam de página para página e demandam especial habilidade para remoção de *tags* e para a estruturação dos dados:

Algumas ferramentas comumente utilizadas são:

a) HTTrack[15]: Onde é possível copiar um site todo, incluindo sua estrutura de diretórios e arquivos desprotegidos para navegação e tratamento off-line;
b) Wget[16]: Software livre que permite a cópia do conteúdo de um site e que pode atuar de forma recursiva.

No entanto, sabe-se que existem diferenças entre o *crawling* de páginas simples e de posts dinâmicos em ferramentas de redes sociais. Para isso algumas ferramentas disponíveis podem ser listadas:

a) *Social Media Crawler*[17]: Um script que tem a capacidade de coletar dados em ferramentas de redes sociais;
b) *Social Snapshot*[18]: Script automatizado capaz de coletar dados sobre usuários nas redes sociais;

Deste modo, na proposta desta pesquisa, os conceitos de *crawling* têm sua especial valia nas hipóteses em que a coleta massiva de dados não poderá ser feita via aplicativo (por falta de consentimento do titular do perfil) ou API.

3.3.3 Google Hacking (Search by image)

O buscador Google pode ser utilizado para minerar dados em redes sociais, eis que alguns dados são também indexáveis. Inúmeros operadores do Google estão disponíveis a usuários para serem manipulados. Estes operadores, se combinados, podem realizar e apresentar informações úteis para se extrair significados e detectar falsidades em redes sociais.

Usuários podem utilizar a pesquisa avançada[19] ou mesmo os operadores de pesquisa disponível, que se combinados, podem servir para o desenvolvimento da atividade de inteligência cibernética.

Alguns operadores de pesquisas disponíveis no buscador Google podem ser visualizados na figura 9.

---

[15] https://www.httrack.com/
[16] http://www.delorie.com/gnu/docs/wget/wget_5.html
[17] https://github.com/aalexandrov/social-media-crawler
[18] https://github.com/search?o=desc&q=social+snapshot&s=&type=Repositories
[19] http://www.google.com/advanced_search

Figura 9 – *Tags* e operadores para utilização do Google em pesquisa avançada

Operador	Como usar
`site:`	Consiga resultados a partir de determinados sites ou domínios. Exemplos: `olimpíadas site:nbc.com` e `olimpíadas site:.gov`
`link:`	Encontre páginas vinculadas a uma página específica. Exemplo: `link:youtube.com`
`related:`	Encontre sites semelhantes a um endereço da Web que você já conhece. Exemplo: `related:time.com`
`OU`	Encontre páginas que podem usar uma das várias palavras. Exemplo: `maratona OU corrida`
`info:`	Receba informações sobre um endereço da Web, incluindo a versão em cache da página, páginas semelhantes e páginas vinculadas ao site. Exemplo: `info:google.com.br`
`cache:`	Veja como estava a página na última vez que o Google visitou o site. Exemplo: `cache:washington.edu`

Fonte: Disponível em: https://support.google.com/*web*search/answer/2466433

Algumas simulações que podem ser utilizadas para inteligência cibernética são exemplificadas abaixo:

“site:legaltech.com.br ext:pdf”

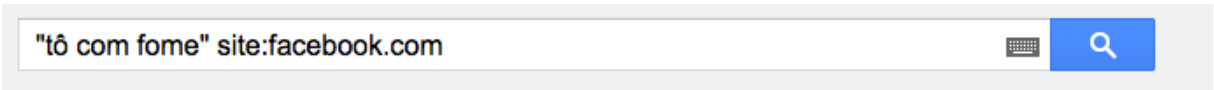
Retorna todos os arquivos disponíveis no padrão PDF em um dado site;

“allintext:lixo site:legaltech.com.br”

Retorna todas as expressões “lixo” dentro de um determinado site. Outro exemplo poderia ser: *allintext:internet* site: Facebook.com/professormilagre. Este exemplo retornaria todas as expressões “internet” publicadas em uma página específica “*fan page*”, na rede social Facebook.

Em nossa pesquisa exemplificativa realizamos a busca:

Figura 10 – Exemplo de pesquisa de expressão que sugira sentimento



Fonte: Elaborado pelo autor utilizando a ferramenta de busca Google

E o resultado identificado é apresentado na figura 11.

Figura 11 – Resultado da pesquisa de sentimento

Fonte: Elaborado pelo autor utilizando a rede social Facebook

O buscador Google, no entanto, ranqueia páginas nos primeiros resultados e apenas ao final as postagens públicas de perfis. Ponto interessante a ser observado na pesquisa é que é possível usar o CACHE do buscador para identificar comentários removidos em páginas, o que pode ser útil na atividade de inteligência cibernética para detecção de conteúdos já excluídos.

Provedores de aplicação podem ser instados, por ordem judicial, a apresentarem o cache (versões anteriores armazenadas) relativo a uma URL ou mesmo a guardarem tais registros por mais tempo.

Assim é possível igualmente buscar por comentários feitos por um usuário em páginas das redes sociais, o que normalmente não seria coletado via APi ou aplicações. Comentários dizem muito sobre uma pessoa. A Universidade Cambridge disponibiliza uma API denominada *"Apply Magic Sauce",* apelidada de *predictionAPI* e que é capaz de dizer muito sobre uma pessoa somente com base, até mesmo, nos seus "*likes*"[20].

Ademais, a partir de um bot suspeito identificado, é possível realizar buscas por imagens associadas ou mesmo conteúdos de texto, o que poderá revelar postagens em massa e outras evidências.

O Google "*Search by Image*" também pode ser utilizado para verificar imagens correspondentes e semelhantes na Internet, inclusive nas redes sociais (postagens públicas), podendo-se identificar por onde anda uma imagem e eventualmente suas alterações, detectando ainda se uma imagem foi utilizada em

[20] É possível se registrar na aplicação e enviar os likes para a aplicação que traçará a personalidade da pessoa. http://applymagicsauce.com/test.html

um perfil falso. A combinação de ferramentas de investigação com estes recursos pode resultar de aplicações eficientes.

3.3.4 *Tineye*

O Projeto *Tineye*[21] também apresenta uma ferramenta que varre a Internet a partir de imagens. Porém aqui não se busca apenas imagens iguais ou semelhantes, mas a pesquisa é feita por palavras-chave, metadados e marcas d'água existentes na imagem. O sistema já tem mais 11 bilhões de imagens crawleadas (capturadas e registradas).

A ferramenta então permite detectar montagens e derivações de uma imagem, bem como recuperar tais imagens (originais). É possível identificar de onde uma imagem veio e todos os locais em que está sendo utilizada, principalmente, onde foi postada pela primeira vez.

3.3.5 *Truthy Project*

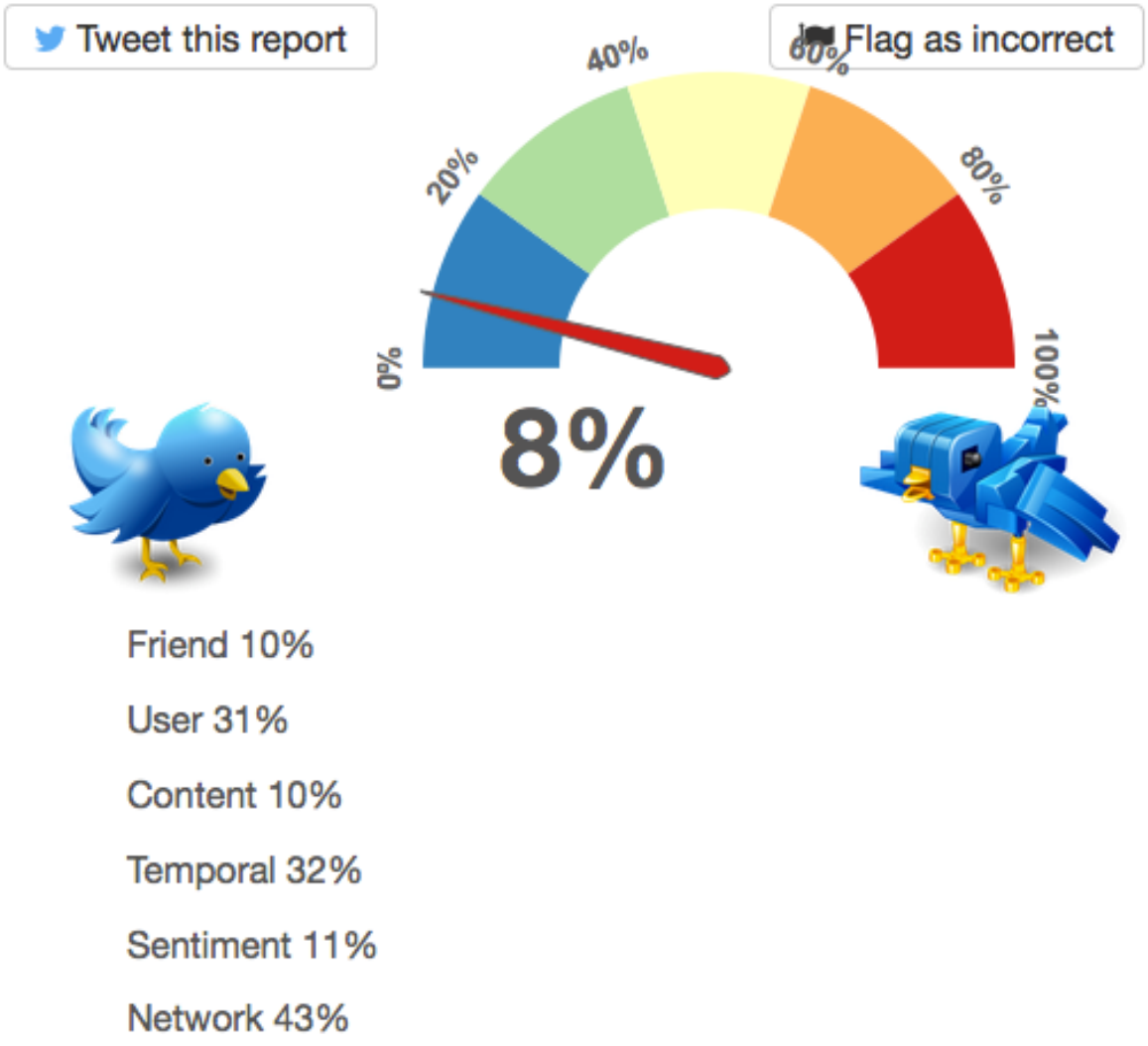
O Projeto TRUTHY[22] é um projeto de pesquisa da *Indiana University* que vem estudando viralidades e como os memes se espalham nas redes sociais. O projeto compõe-se de várias ferramentas sendo que em uma delas, em desenvolvimento, será possível acompanhar a difusão de memes que interessem um usuário.

Uma das ferramentas do Projeto é a "BOT OR NOT". Em uma das pesquisas deste trabalho simulou-se o uso do BOT OR NOT com um perfil no Twitter periciadigital[23]. O resultado apresentado em 05/06/2015 foi que o perfil não era um BOT.

---

[21] http://www.tineye.com/
[22] http://www.truthy.indiana.edu/
[23] http://truthy.indiana.edu/botornot/?sn=periciadigital

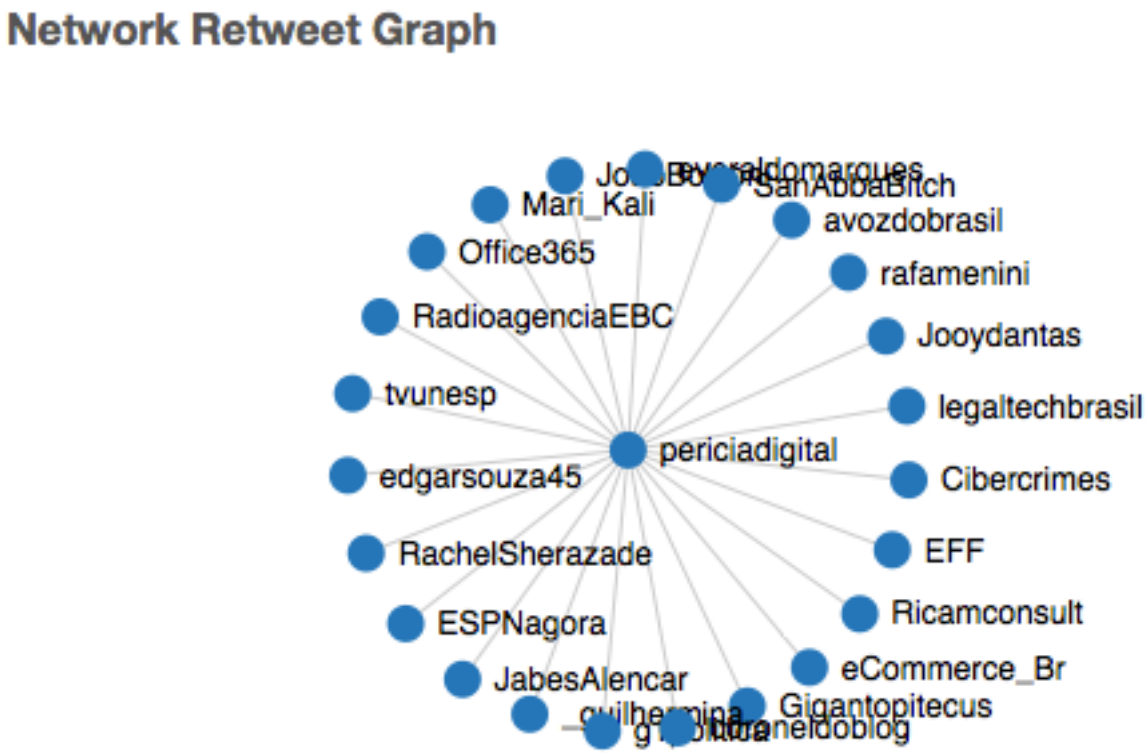
Figura 12 – Resultados no uso da ferramenta *Truthy Project*



Fonte: Elaborado pelo autor utilizando a ferramenta TruthyBotOrNot

Igualmente, com a *network graph* é possível identificar “preferências” ou temas ou pessoas pelas quais o perfil se interessa, vejamos:

Figura 13 – Gráfico de preferências ou retuítes do perfil @periciadigital



Fonte: Elaborado pelo autor utilizando a ferramenta TruthyBotOrNot

Assim, extraindo tais informações, em uma rápida investida visual na figura 13, é possível identificar as principais influências e interesses de um perfil. O que poderia ser categorizado e podendo revelar as “preferências” de um usuário, o que pode ainda ser utilizado para detectar falsidades e seus sentimentos (expressões que contrariam seu “agir” no mundo virtual). A ferramenta, no entanto, está disponível apenas para o Twitter. Além disso, a ferramenta não detecta perfis falsos, mas BOT, ou seja, perfis criados por robôs (aplicações) para postagens automáticas nas redes sociais.

3.3.6 Photo DNA

O PhotoDNA é uma tecnologia desenvolvida pela empresa Microsoft juntamente com a Universidade de Darthmouth e vem sendo utilizada por grandes ferramentas de redes sociais para impedir imagens de abuso de crianças. Trata-se de um sistema que realiza um resumo criptográfico de imagens suspeitas,

taggeando as mesmas que são inseridas em um banco de dados. Twitter, Facebook e Microsoft são algumas empresas que utilizam a tecnologia. O Google utiliza tecnologia semelhante.

Deste modo, quando uma imagem é postada, o PhotoDNA produz uma soma matemática sobre o conteúdo (*hash*) e coleta algumas informações biométricas da figura, sendo que o método continua funcionando mesmo com a imagem alterada em seu formato e tamanho.

O grande diferencial da ferramenta é que, além da geração do resumo criptográfico (*hash*), ela converte uma imagem que ingressa na rede social em escala de cinza, divide a imagem em regiões, calculando a densidade derivativa para cada pixel nas regiões, criando um histogram e adicionando o valor absoluto para cada região. Assim, a imagem pode perder seus metadados, mas ainda pode ser rastreada graças ao algoritmo matemático aplicado. A grande diferença do *hashing* do PhotoDNA é que este é baseado no contexto da imagem e não no arquivo. (MICROSOFT PHOTODNA, 2009).

A ferramenta não está disponível a pesquisadoras e não é aplicada em casos de proteção autoral, direito de imagem ou perfis falsos.

4 Revisão de pesquisas que tratam de perfis falsos em redes sociais

No Capítulo 3 identificamos alguns recursos e ferramentas disponíveis que podem ser utilizadas para a construção de ferramentas de análise de dados em grandes volumes e na detecção de sentimentos, falsidades, perfis falsos e sentimentos em redes sociais.

Muitos são os desafios e limitações das ferramentas e recursos avaliados. Assim, neste capítulo, avançou-se nas pesquisas, com a revisão das principais técnicas e metodologias propostas por pesquisadores, em âmbito internacional, para a construção de modelos e protótipos para detecção de falsidades em redes sociais.

Revisaram-se pesquisas internacionais que tratam da detecção de *fakes*, avaliando os métodos propostos nas pesquisas, suas limitações e principalmente, eixos que podem ser úteis nesta pesquisa, apresentadas ao final do trabalho.

4.1 Método para classificar e avaliar a credibilidade de usuários no Twitter

Como enfatizado, tem-se como principal desafio nas redes sociais decidir quais usuários seguir ou informativos assinar, considerando sua relevância, qualidade e credibilidade da informação recebida. É tarefa difícil avaliar a credibilidade de usuários na rede. No escopo de contornar este problema, Canini, Suh e Pirolli (2011) apresentam um método para automaticamente identificar e ranquear usuários de redes sociais de acordo com a relevância e expertise para um dado tópico ou área de conhecimento.

O crescimento das redes sociais proporcionou a facilidade em compartilhar conteúdos como nunca antes constatado. Por outro lado, redes como Twitter, por exemplo, contendo aproximadamente 200 milhões de usuários registrados, são responsáveis por uma infinita produção de dados, mas para recuperação da informação, no entanto, oferecem apenas um simples mecanismo de busca que retorna informações em ordem cronológica. Como analisar a credibilidade de informações que são publicadas em uma velocidade incrível?

Pela pesquisa proposta, dado um tópico de interesse, seria possível identificar quem provê informações de credibilidade neste tema, considerando para fins da pesquisa, credibilidade, como a junção de experiência e confiança. (CANINI; SUH;

PIROLLI, 2011). Tal percepção é humanamente impossível a um usuário de redes sociais, sem suporte de técnica que auxiliem no tratamento destas informações.

Neste contexto, importante igualmente o trabalho anterior de Bernstein et al. (2010), que concebeu um sistema para organização e apresentação de tweets em tópicos. Weng et al. (2010) combinou a modelação em tópico para medir a influência de usuários do Twitter.

Canini, Suh e Pirolli (2011) se valem dos estudos de Birnbaum e Stegner (1979), onde em seus experimentos os participantes foram convidados a julgar o valor de mercado de carros usados antes e depois de conferirem a avaliação por terceiros, que variava de acordo com o conhecimento de domínio de carros usados.

Os autores selecionaram 98 participantes ativos de redes sociais e criaram cinco diferentes domínios do conhecimento*, carros, investimentos, vinhos, futebol e namoro.* Para cada domínio, selecionaram 10 contas no Twitter com alta relevância e experiência, coletadas a partir do serviço WeFollow[24] para cada domínio, que cria listas de usuários influentes no Twitter. Também foram selecionadas 10 contas de pessoas sem experiência em nenhum domínio particular. Foram coletadas as postagens dos 60 (sessenta) perfis. (CANINI; SUH; PIROLLI, 2011). Assim os participantes faziam, dada uma ficha de um veículo um "pré-julgamento" do valor do carro. Após, tinham contato com a avaliação de um perfil de terceiro, e então, realizava um "pós-julgamento", mantendo o preço, subindo ou baixando.

Foi possível calcular um valor e peso para a avaliação em determinado assunto. Foi detectado que o fator experiência no domínio tem forte influência no julgamento de credibilidade e o "*social status*" (número de seguidores, quem segue e tweets) tem menor influência. Percebeu-se que o domínio do conhecimento também influencia na classificação e avaliação de credibilidade do usuário. Identificado que nuvens de *tags*, sozinhas não se referem à classificação elevada de credibilidade para um usuário do Twitter, mas está mais relacionada a tendências. (CANINI; SUH; PIROLLI, 2011).

Os resultados da pesquisa esclareceram que a credibilidade de uma conta no Twitter dependerá em grande parte da força de associação entre o conteúdo textual e o domínio em questão e menos do "status social do domínio".

[24] http://wefollow.com/

Os autores em estudo, no mesmo trabalho, também apresentam um algoritmo para ranquear usuários relevantes por tópico, que combina a pesquisa básica na rede social Twitter (sobre um domínio específico), com a aplicação do algoritmo de ranqueamento e tópico. A pesquisa condiz-se da seguinte forma, sem síntese:

a) Identifica-se um set de perfis que são potencialmente relevantes para um tópico de interesse (estes são chamados de eleitores ou *voters*);

b) A aplicação então observa a opinião dos eleitores, como base em quem eles seguem. Os autores da pesquisa partem da premissa que se um usuário segue outro no Twitter, significa que ele valoriza a informação de quem é seguido;

c) A partir desta premissa, o sistema construiu um conjunto de usuários (chamados de candidatos ou candidates), incluindo qualquer um que é seguido por pelo menos um dos eleitores;

d) Para cada usuário no set de candidatos, é extraído o número de votantes que seguem o usuário candidato e o número total de usuários de Twitter que seguem o usuário, assim é possível computar o “social score” de cada membro na tabela de candidatos;

e) A partir do momento em que se identificaram candidatos e recuperados seus números relevantes, passa-se ao cálculo de relevância de cada usuário candidato em relação ao “tópico” pesquisado;

f) Este processo pode ser feito de várias formas. Os autores apresentam 3 (três) algoritmos. A medida *“Numvotes”* mensurará e contará quantas vezes os seguidores do usuário candidato recentemente tweetaram sobre o determinado tópico.

g) O processo também pode ser feito pela medida de relevância que considera-se a proporção *“DivF”* de números de seguidores associado ao tópico pesquisado, quanto maior a proporção de seguidores de um usuário que está associado com um tópico, mais confiável o “Candidato” é. Embora os autores reconheçam limitações neste algoritmo chegando a propor um algoritmo derivado de nome “DivLogf”. Os autores, porém, consignam que a medida de

relevância preferida é a “Betabin (α , β)”, calculada a partir de probabilidade bayesiana. Os autores em estudo consideram o Betabin mais efetivo que o “DivF” e “NumVotes”. (CANINI; SUH; PIROLLI, 2011, p. 5).

Figura 14 – Recuperação de votantes e candidatos para medição do “social score” de um usuário na rede social

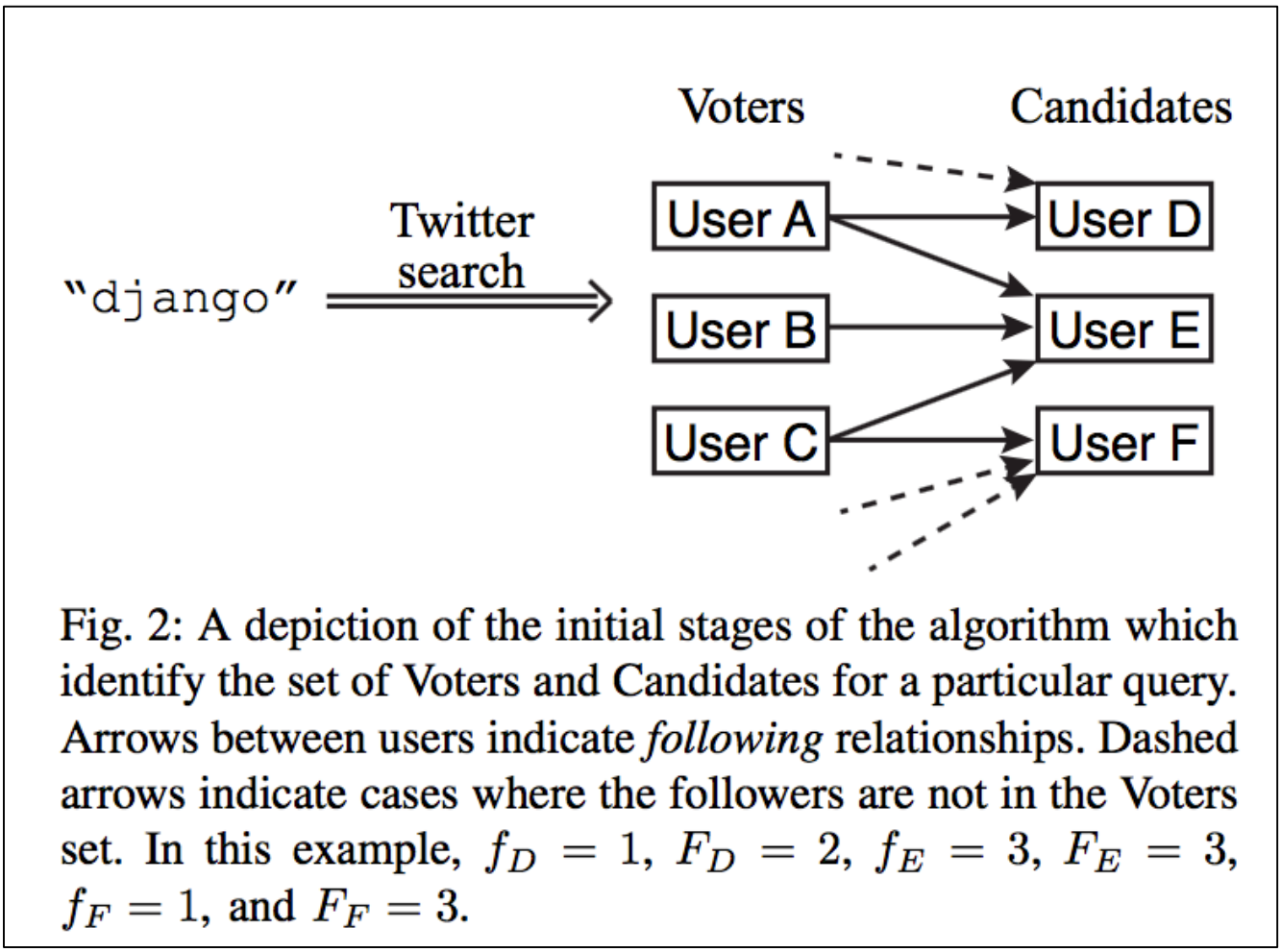


Fig. 2: A depiction of the initial stages of the algorithm which identify the set of Voters and Candidates for a particular query. Arrows between users indicate *following* relationships. Dashed arrows indicate cases where the followers are not in the Voters set. In this example, $f_D = 1$, $F_D = 2$, $f_E = 3$, $F_E = 3$, $f_F = 1$, and $F_F = 3$.

Fonte: CANINI; SUH; PIROLLI (2011, p. 6)

A inferência bayesiana afeta a estatística e é fundamentada no Teorema de Bayes, utilizado na teoria da probabilidade. O teorema é capaz de apresentar a relação entre uma probabilidade condicional e a sua forma inversa. Pelo teorema é possível extrair a probabilidade de um acontecimento dada a observação de uma evidência, bem como a probabilidade da evidência a partir do acontecimento.

Assim, apresenta-se a alteração da probabilidade anterior, a partir de uma evidência, gerando-se uma probabilidade posterior. A fórmula da probabilidade ou inferência bayesiana é assim definida:

Figura 15 – Fórmula da Inferência Bayesiana

$$P(H \mid E) = \frac{P(E \mid H) \cdot P(H)}{P(E)}$$

Fonte: CANINI; SUH; PIROLLI (2011, p. 6)

“H” significa qualquer hipótese cuja probabilidade possa ser alterada por dados, provas ou evidências. Ex: A probabilidade de uma mãe dar a luz menino;
“E” é a evidência. Ex: Ultrassonografia;

“P(H)” é a probabilidade da hipótese antes da observação da evidência. Ex: Existe a possibilidade de nascer um menino ou uma menina;

“P(H|E)” é a probabilidade da hipótese com observação da evidência. Ex: Como a evidência é capaz de predizer que será um menino;

“P(E|H)” é a probabilidade de se observar a evidência dada a hipótese. Ex: É possível que uma mulher tenha um filho cão? Não. Neste caso, mesmo apresentando uma imagem de um cão, não há como afirmar compatibilidade da evidência com a hipótese dada;
“P(E)” é a probabilidade marginal ou evidência modelo, um fator que é o mesmo para todas as hipóteses possíveis.

Para validar os algoritmos de medição de credibilidade acima narrados, Canini, Suh e Pirolli (2011) se valeram da seguinte prova de conceito:

a) Pesquisaram pela query “django” e levantaram 1.500 tweets de 980 usuários únicos (formaram os eleitores);
b) Após expandiram os seguidos destes usuários, gerando 234.166 registros (formaram os candidatos);
c) Os candidatos foram ranqueados de acordo com as medidas estudadas;
d) Para as comparações, os autores coletaram os top 200 usuários de acordo com a mesma query “django” no site *WeFollow;*
e) Na sequência, foi realizada a medição da precisão de cada algoritmo estudado. Foi- então extraído o top 20 para cada medida de relevância e

misturada com a lista top 20 do *WeFollow*, o que produziu 97 candidatos. Foram recrutados dois dos usuários do Twitter com experiência no tema (django) e que classificaram cada um dos usuários como relevantes ou irrelevantes, realizando tal tarefa para os 97 registros de candidatos.

f) Posteriormente foi checado o "*recall*" de cada algoritmo. Isto é: Dada uma lista de conhecidos especialistas, quantos deles são identificados nos algoritmos. A lista foi de 25 especialistas. Foi calculado quantos deles apareciam na lista top 100 de cada algoritmo e na lista do *WeFollow.*

O resultado é interessante, pois apresenta nas colunas *"Precision"* da figura 16, o número de usuários em cada lista de "top 20" (geradas pelos algoritmos e pelo *WeFollow*) que foram julgados relevantes para dois categorizadores humanos que entendem do assunto. A coluna *"recall"* por sua vez, mostra quantos usuários da lista de 25 conhecidos especialistas foram identificados por cada algoritmo. (CANINI; SUH; PIROLLI, 2011). O Melhor algoritmo evidencia-se:

Figura 16 – Precisão dos algoritmos utilizados na pesquisa

Measure	Formula	Precision 1	Precision 2	Recall
NumVotes	f	7	6	6
DivF	f/F	0	2	0
DivLogF	$f/\log F$	13	12	8
BetaBin$(1, 10^2)$	$(f+1)/(F+10^2+1)$	15	11	11
BetaBin$(1, 10^3)$	$(f+1)/(F+10^3+1)$	**19**	**17**	**13**
BetaBin$(1, 10^4)$	$(f+1)/(F+10^4+1)$	17	15	11
WeFollow	N/A	**19**	14	10

Fonte: CANINI; SUH; PIROLLI (2011, p. 7)

O algoritmo BetaBin demonstrou-se com grande potencial para auxiliar pessoas a identificar usuários interessantes e realmente relevantes sobre um dado tema no Twitter, sendo que teve precisão até maior que o serviço "*WeFollow*" já disponível na Internet. (CANINI; SUH; PIROLLI, 2011).

Embora os autores em estudo acreditem que o método descrito possa ser apresentado em outras redes sociais, o trabalho limitou-se ao experimento no Twitter. O sistema, usado em forma reversa, poderia ser derivado para detectar

“falsidades” ou inverdades nas redes sociais. Muitos jornais utilizam “perfis” e páginas para divulgarem notícias estampadas em sites estáticos. A proposta é interessante, pois medir a falsidade sobre determinada informação passa necessariamente por conhecer a credibilidade da fonte. Assim, o sistema de credibilidade aqui proposto, na detecção de perfis falsos pode, por exemplo, distinguir dentre vários perfis, dado um tema de especialidade, qual efetivamente é o real, considerando que os *fakes* não seriam capazes, em tese, de usurpar o conhecimento do clonado. Não se teve relatos de avanço nas pesquisas.

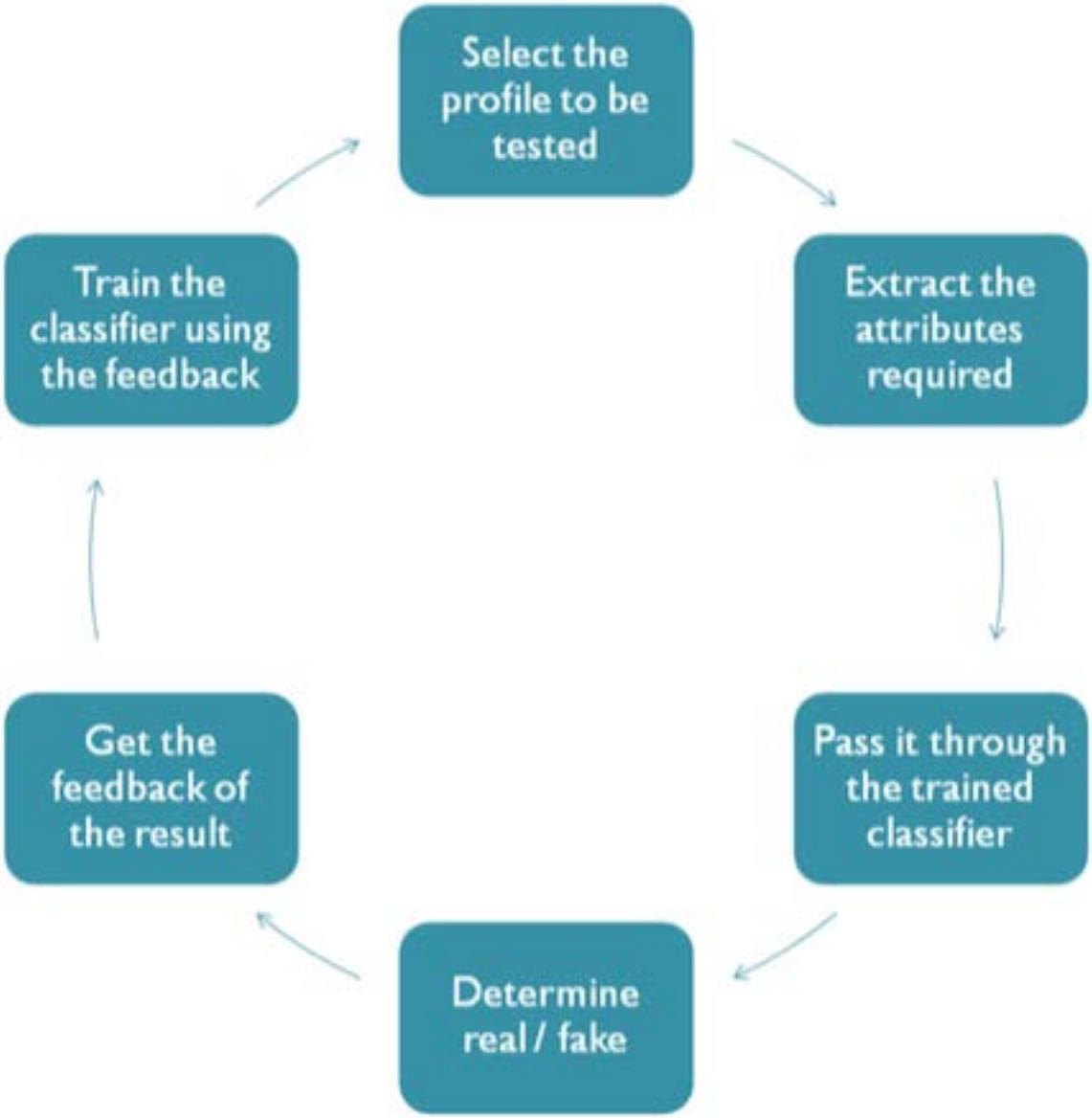
4.2 Detecção automatizada de perfis falsos e a vantagem do algoritmo “*Support Vector Machine*”

Cada perfil na rede social contém muitas informações como gênero, número de amigos, número de comentários, educação, etc. É sabido que algumas informações são privadas e outra públicas.

Reddy e Kumar (2012) apresentam um método automatizado para detecção de perfis falsos nas redes sociais, baseados apenas em informações públicas (o que em tese não violaria a legislação brasileira e internacional), a seguir descrita.

Ocorre a seleção do perfil a ser testado e a extração de atributos que são passados ao classificador, que é treinado regularmente. Este classificador determinaria se o perfil é *fake* ou real e nestes casos considerando que não se tem precisão de 100% (cem por cento), um aviso seria remitido ao profile para que prove sua identificação e sendo a identificação fornecida o classificador seria ajustado. Logo, uma aplicação de uso anexado às ferramentas de redes sociais. A proposta é apresentada através do fluxo apresentado na figura 17.

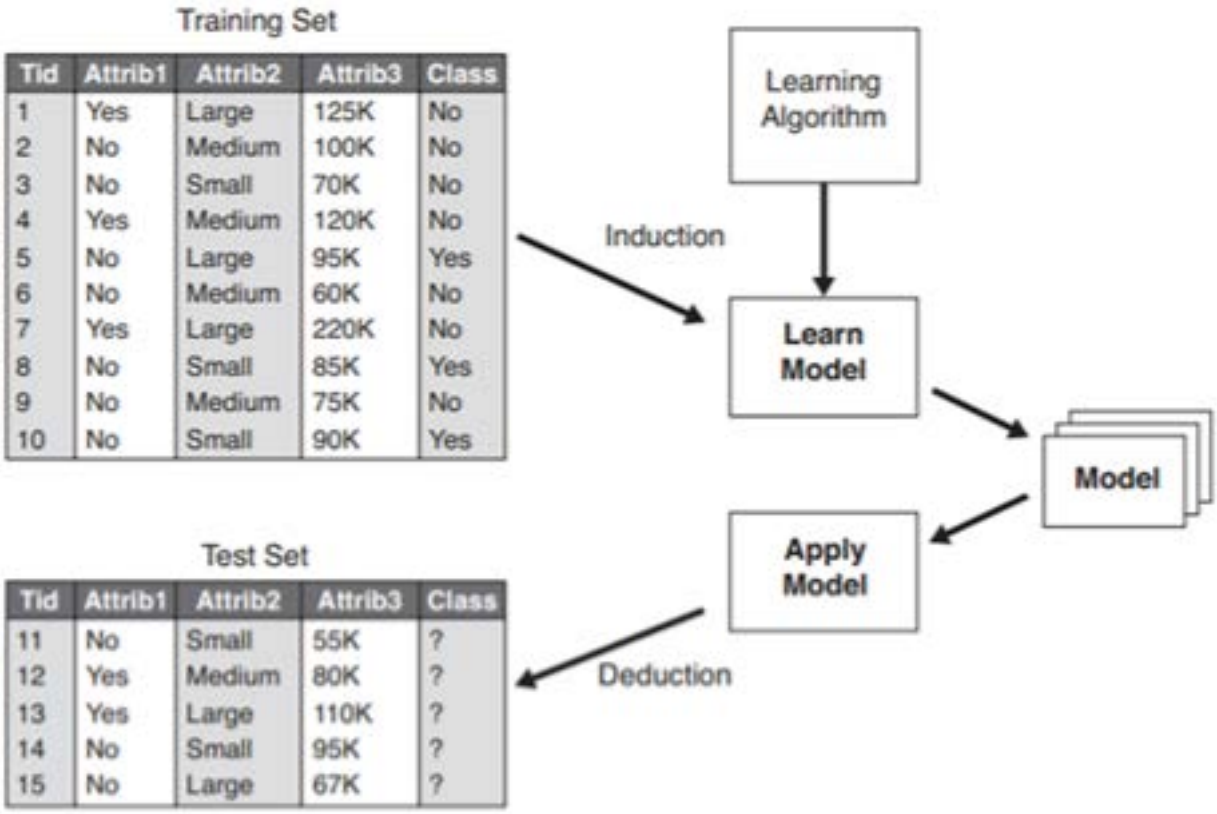
Figura 17 – Fluxo do sistema para detecção de perfil falso



Fonte: REDDY; KUMAR (2012, p. 15)

Os autores do trabalho em estudo apresentam uma proposta para detecção de perfis falsos que considera um “algoritmo de aprendizado”, que primeiramente é treinado com o *Data-set* de dados com “classes definidas” (chamado de *set de treinamento*) e então se extrai- modelos com base na indução. Após, é submetido o *data-set* real a ser testado (chamado de *set de teste*) e são aplicados os modelos, onde pela “dedução” podem ser sinalizados possíveis perfis falsos, vejamos:

Figura 18 - Abordagem geral para construção de um modelo de classificação de perfis falsos com base em indução e dedução



Fonte: REDDY; KUMAR (2012, p. 16)

Indicam os autores no presente trabalho que o “classificador” na pesquisa fora implementado utilizando os algoritmos *“Navie Bayes”, “Decision Tree” e “Support Vector Machine”*, algoritmos amplamente utilizados em problemas envolvendo a detecção de e-mails não solicitados (*SPAM)* e outras funções. (REDDY; KUMAR, 2012).

As dificuldades encontradas na pesquisa estão relacionadas à coleta dos *data-sets* considerando as configurações de privacidade de ferramentas de redes sociais. Neste sentido, fora criado um perfil e adicionados amigos, em seguida foi realizado o procedimento denominado *“scrapping data”*, onde scripts em *python* foram utilizados para coletar os dados dos amigos.

Os dados coletados foram:

a) Número de amigos;
b) Educação e trabalho;
c) Gênero;

d) Número de colunas preenchidas sobre o usuário;
e) Número de fotos da pessoa taggeada;
f) Número de posts;
g) Número de upload de fotos pela pessoa;

A pesquisa considerou dados coletados de 15 de maio de 2011 a 15 de setembro de 2011. Os atributos acima usados foram os únicos que os autores conseguiram coletar das ferramentas de redes sociais. (REDDY; KUMAR, 2012).

Os parâmetros de avaliação foram definidos com base em “Eficiência” (número de predições corretas dividido pelo número total de predições); “Taxa de falso positivo” (Número de perfis reais detectados como falsos, dividido pelo número total de perfis falsos a serem detectados); “Taxa de falso negativo” (Número de perfis falsos detectados como reais, dividido pelo número total de perfis reais).

Figura 19 – Métricas para apuração dos perfis falsos

$$\text{Efficiency} = \frac{\text{No. of correct predictions}}{\text{Total No. of Predictions}}$$

$$\text{False Positive rate} = \frac{\text{No. of real profiles detected as fake}}{\text{Total No. of fake profiles to be detected}}$$

$$\text{False Negative rate} = \frac{\text{No. of fake profiles detected as real}}{\text{Total No. of real profiles}}$$

Fonte: REDDY; KUMAR (2012, p. 25)

Como resultado, foi observado que a eficiência do algoritmo “*Support Vector Machine*” é alta quando os dados são bem treinados. Observou-se também que conforme o número de atributos cresce no *data-set* de treinamento, cresce também a eficiência dos algoritmos.

A pesquisa, no entanto, não apresenta o protótipo, tampouco apresenta como realizou a coleta de dados automatizada nas ferramentas de redes sociais, outro desafio envolvendo grande volume de dados.

As contribuições da pesquisa em estudo são claras. Ferramentas para detecção de perfis falsos podem se basear na extração de uma “impressão digital” formada pelos atributos de um perfil real (condensando os que variam menos) e neste contexto, ser a base para comparação com outros perfis. De outro lado, pode-se ainda extrair atributos de perfis reconhecidamente *fakes* e com base neles, formar um classificador automático com novos perfis a serem testados.

4.3 Detecção de imagens *fakes* e a viralidade da desinformação

Gupta et al. (2013) esclarecem que extrair boa informação e de qualidade é um dos grandes desafios na utilização da informação disponível nas redes sociais.

Os autores em estudo trabalharam com o episódio do Furacão Sandy que causou a destruição em massa por todos Os Estados Unidos, de 22 a 31 de outubro de 2012, causando danos de 50 bilhões de dólares. Foi evidenciado que as ferramentas de redes sociais foram utilizadas para espalhar rumores e falsas imagens, que se tornaram virais e causaram caos e pânico.

Ao revisitarem os trabalhos existentes sobre o tema, Gupta et al. (2013), identificaram a contribuição de Corvey et al. ([2011?]) que chega a concluir que durante situações de emergência, usuários usam um vocabulário específico em redes sociais. Já Mendoza, Poblete e Castillo (2010) concluem que a propagação de tweets relacionados a rumores versus notícias reais diferem e podem ser utilizadas para o desenvolvimento de soluções para identificação da informação correta.

No trabalho em análise, utilizou-se a “Streaming API” do Twitter para coletar dados envolvendo “Sandy” e “*hurricane*”, o que gerou aproximadamente 1.8 milhões de tweets de 1.2 milhões de usuários únicos. Os autores então submeteram os Tweets com URLs às comparações de imagens “*Fake*” do furação publicadas em revistas. Posteriormente, analisaram os retuítes destas postagens. Após esta primeira triagem (coletando os dados dos links), a pesquisa selecionou os tweets com imagens falsas e aplicou a classificação que foi sumarizada em dados do usuário que postou e dados da mensagem postada, a seguir classificados na figura 20.

Figura 20 – Dados levantados dos usuários e sobre as mensagens envolvendo o furacão Sandy

User Features [F1]
Number of Friends
Number of Followers
Follower-Friend Ratio
Number of times listed
User has a URL
User is a verified user
Age of user account
Tweet Features [F2]
Length of Tweet
Number of Words
Contains Question Mark?
Contains Exclamation Mark?
Number of Question Marks
Number of Exclamation Marks
Contains Happy Emoticon
Contains Sad Emoticon
Contains First Order Pronoun
Contains Second Order Pronoun
Contains Third Order Pronoun
Number of uppercase characters
Number of negative sentiment words
Number of positive sentiment words
Number of mentions
Number of hashtags
Number of URLs
Retweet count

Fonte: GUPTA et al. (2013, p. 5)

Com isso, puderam perceber que a grande maioria dos tweets eram retuítes e que 0,3% dos usuários eram responsáveis por 90% dos retuítes das imagens falsas. Igualmente, analisando dados e comportamento dos usuários, foi identificado que em tempos de crises, usuários retuítam informações de outros independentemente de segui-los ou não. (GUPTA et al., 2013).

Uma das deficiências deste processo é que necessita de imagens “base” que seriam consideradas como “verdadeiras” para a comparação/caracterização em relação ao *data-set* coletado. Em seguida, passariam a fazer a classificação para se estudar a “viralidade”, a característica dos usuários e suas influências na

transmissão da desinformação. Os autores finalizam sua pesquisa informando que pretendem avançar em um protótipo automatizado. Não se identificou relatos do lançamento deste protótipo.

Na detecção de perfis falsos a proposta em estudo pode ser adaptada para identificar a origem de uma desinformação, bem como na detecção de robôs programados para compartilhar determinado conteúdo.

4.4 Detecção de perfis falsos com base em "*profile similarities*"

Jin, Takabi e Joshi (2011) propõem um *framework* para a detecção de perfis clones em redes sociais com base em comparações entre os perfis. O *framework* considera dois atributos como fundamentais, sendo eles: a) atributos de similaridade e b) similaridades em redes de amigos.

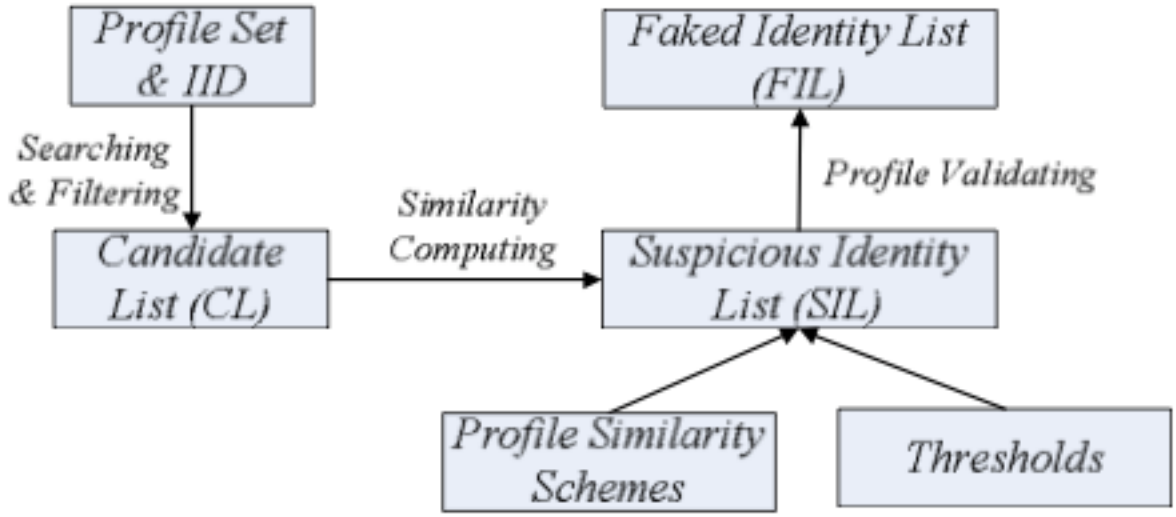
O escopo seria contribuir para evitar os ataques de clone de identidade, que tem por escopo o roubo de dados pessoais de usuários de redes sociais.

Quando um indivíduo solicita amizade com outro em uma rede social, ocorre um "link", sendo que chamamos de "*nodes*" a representação de usuários e outras entidades, como grupos, em um contexto social.

Os autores utilizam a *"cosine similarity"* como medida de similaridade e analisam critérios como, tendo por base a vítima, amigos do Id *Fake*, amigos recomendados da vítima (gerados pela própria rede social comumente) e do *fake*, amigos mútuos da vítima e do *fake* e amigos excluídos da vítima e do *fake*.

A pesquisa chega a se preocupar com o atacante que pode querer falsificar também os amigos do perfil falso, por meio da técnica *"Multiple faked identities* in *friend networks*" (JIN; TAKABI; JOSHI, 2011, p. 33). O processo descrito pelos pesquisadores é apresentado a seguir:

Figura 21 – Modelo de detecção de perfis falsos com base em similaridades de atributos



Fonte: JIN; TAKABI; JOSHI (2011, p. 34)

Pelo processo descrito na figura 21, um perfil base é selecionado e pesquisado por técnicas de filtragem, os chamados perfis candidatos (nomes similares) e de posse destes perfis, são realizadas análises de similaridade, gerando assim os chamados perfis suspeitos. Estes por sua vez são submetidos a outros processos de similaridade e análise, gerando a chamada lista de identidades *fakes*.

Os pesquisadores não abordam como conseguiram acesso ao *data-set* de dados nas ferramentas de redes sociais. Igualmente, as limitações apresentadas por esta técnica evidenciam, eis que nem sempre pessoas usam ferramentas de redes sociais, logo, não teriam perfil real para as comparações. Em muitos casos, constata-se a criação de robôs e não propriamente *fakes*.

O método também não cobre a questão da criação de perfis com personagens que não existem no mundo real. Os testes não foram realizados em sistemas reais e os pesquisadores planejaram implementar o sistema por meio de uma aplicação para Facebook, no futuro. (JIN; TAKABI; JOSHI, 2011).

4.5 Detecção de perfis falsos com base na evolução do tempo

As redes sociais também atraem atividades criminosas interessadas na criação de perfis falsos para causar danos a pessoas ou obter informações pessoais por meio da interação com amigos da vítima. Já se tratada na literatura especializada sobre os a chamados "*Identify Cloning Attacks"* (ICA) capazes de criar clones a partir da rede de relacionamentos da vítima. Outro ataque comum é o *"Fake Profile Attack" (FPA)*, onde informações reais de uma pessoa são usadas em outras redes sociais, comumente onde a vítima não está presente. (CONTI; POOVENDRAN; SECCHIERO, 2012).

Robustos mecanismos para detectar perfis falsos, são debatidos na comunidade científica. Conti, Poovendran, Secchiero (2012), desenvolvem os primeiros esforços no escopo de detecção de perfis falsos no Facebook. A abordagem utiliza análises empíricas e a estrutura das interações nas redes sociais e suas estatísticas (*Graph*s). Em suas pesquisas, os autores realizaram o uso de uma aplicação no Facebook para a coleta de *data-sets* de usuários reais e analisaram os dados, chegando à conclusão que a evolução do tempo *"time evolution"* nas novas amizades pode ser considerada um alerta quando ocorrem desvios do comportamento padrão, na detecção de falsidades. (CONTI; POOVENDRAN; SECCHIERO, 2012).

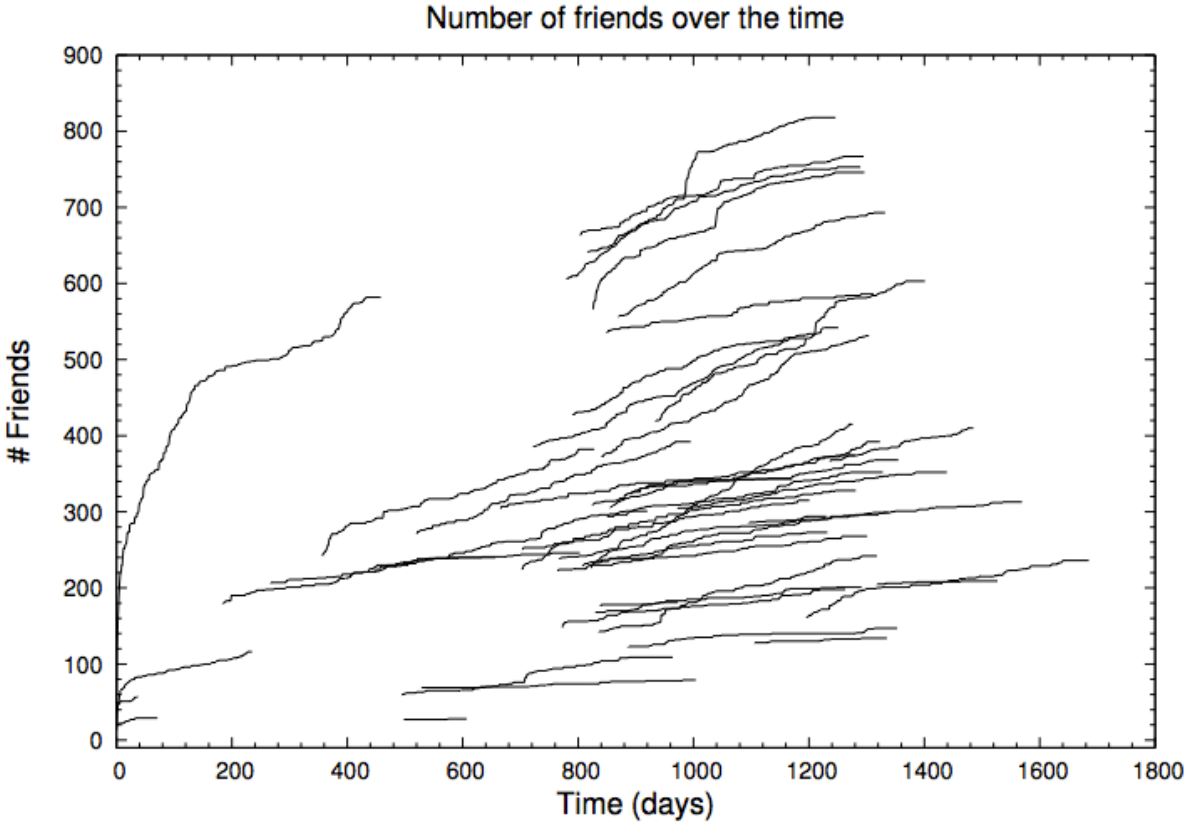
Pelo trabalho fica demonstrado que a comparação de perfis com base em apenas similaridades pode não ser efetiva, considerando que nem sempre se tem o perfil real (original) base para a comparação.

Deste modo, os pontos de análise em ferramentas que pretendam tratar da detecção de perfis falsos devem considerar: a) a evolução no tempo do número de amigos, b) as reais interações do usuário nas ferramentas de rede sociais, c) a evolução no tempo da formação da estrutura e estatísticas do "*graph*" da ferramenta da rede social. (CONTI; POOVENDRAN; SECCHIERO, 2012).

Na pesquisa realizada e aqui em análise, os pesquisadores desenvolveram uma aplicação que coletava informações de um perfil (como o id dos amigos, gênero, data de nascimento), convidando usuários que conheciam a instalarem e processarem tal aplicação (pessoas que eles conheciam). Foram coletados dados de 80 perfis. Os dados foram processados sob diversas óticas, e uma das avaliações realizadas calculou o número de amigos ao decorrer do tempo, quando

se percebeu, excetuado o período inicial, que a taxa em que usuários adicionam novos amigos é quase constante no tempo (foram calculadas a taxa em três intervalos distintos).

Figura 22 – Taxa de adição de novos amigos em dias



Fonte: CONTI; POOVENDRAN; SECCHIERO (2012, p. 1074)

Para os pesquisadores, para evitar ser detectado nesta abordagem o adversário (*fake*) precisaria criar um perfil falso e continuar adicionando amigos a uma taxa satisfatória às condições padrões. Se ainda assim conseguir, parte-se para a segunda parte da pesquisa, que é a análise das interações nas redes sociais.

Com base no perfil real, poderia ser analisada a intensidade dos contatos com os amigos e comparada com o perfil suspeito, o que iria expô-lo. A frequência, por exemplo, com que membros da família da vítima interagem com o perfil real seria mínima ou quase nula em relação ao perfil falso.

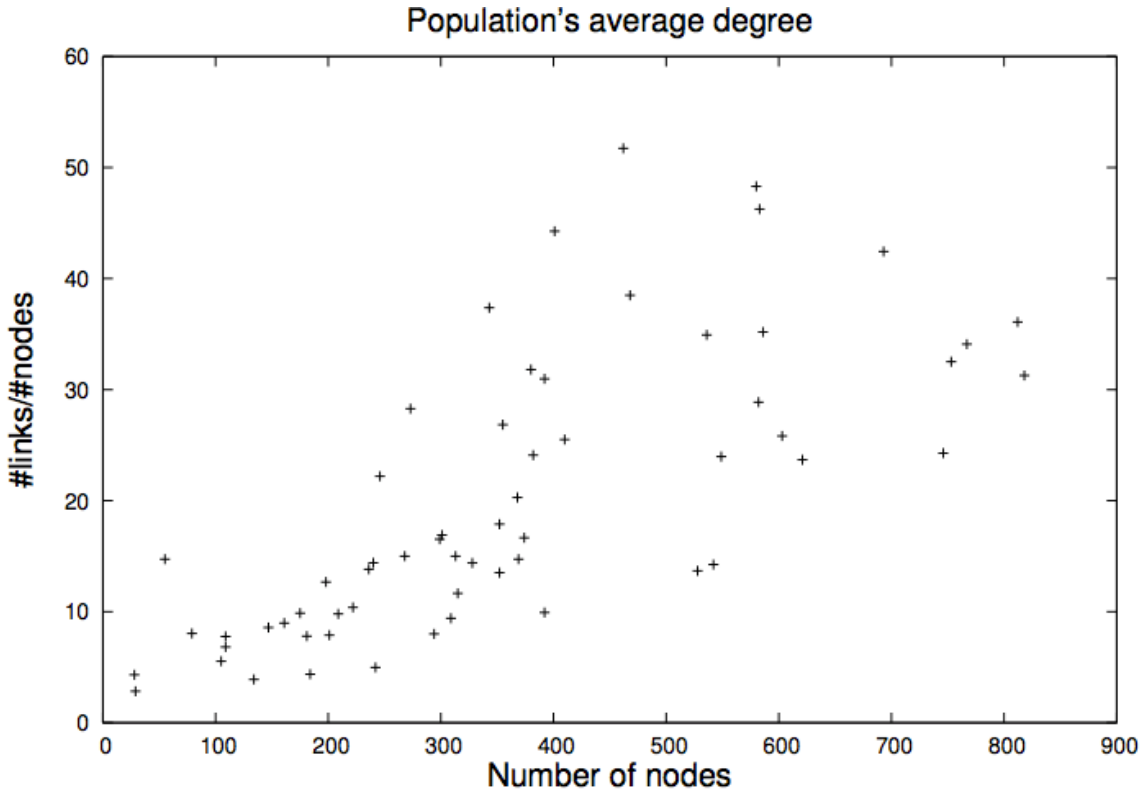
Por fim, como último bloco para a possível criação de um mecanismo de detecção de perfil falso, está a análise estrutural da rede social, onde o escopo seria uma comparação da estrutura de perfil falso em relação a estrutura geral da

população coletada nos *data-sets* reais. Podem ser detectados: número de cliques, número de componentes conectados, dentre outros. Neste ponto duas observações são importantes, sendo elas a análise do grau médio de nós e número de amigos únicos no gráfico da Rede Social.

De fato, o perfil falso tentará evitar contato e requisições de amizades e pessoas que sejam amigas do perfil real. Foram indagados aos participantes da pesquisa que fornecessem a lista dos amigos que possuem contato na vida real. Foram observados os links em comum, compartilhamentos e os nós e também amigos em comum. Pessoas que se veem frequentemente tendem a ter mais estes elementos. Logo, o perfil *fake* teria um esforço para manter amigos em comum e compartilhamentos de links. Se duas pessoas se reúnem comumente na vida real, elas também terão muitos amigos em comum na rede e um alto grau de links em comum. (CONTI; POOVENDRAN; SECCHIERO, 2012).

Nos "*nodes*" seriam considerados não só os amigos que o perfil real vê com mais frequência, mas aqueles que compartilham links em comum com este perfil (conexões em comum). O gráfico proposto é claro:

Figura 23 – Gráfico de *nodes*

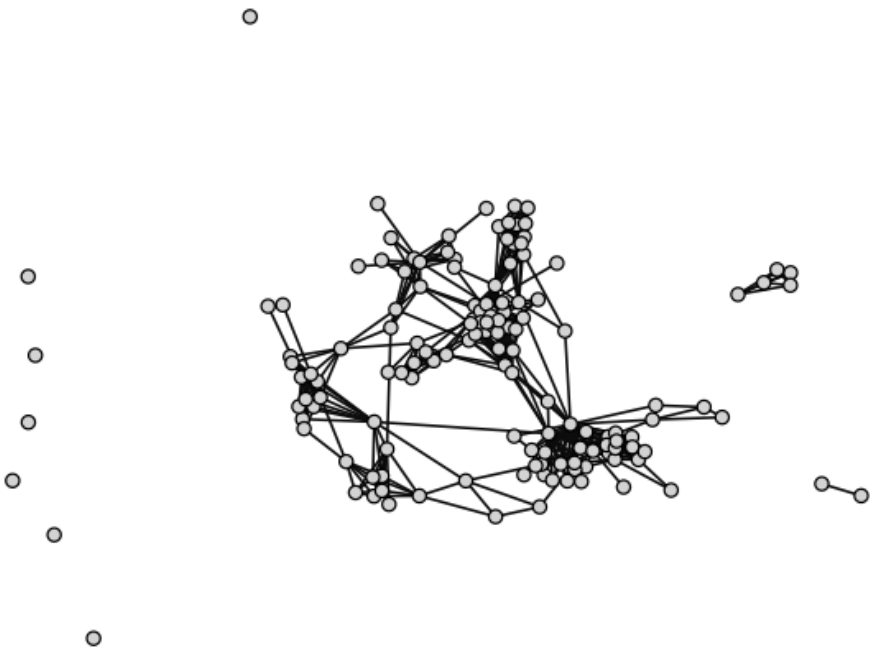


Fonte: CONTI; POOVENDRAN; SECCHIERO (2012. p. 1076)

Deste modo, um adversário ou perfil falso para não ser detectado na rede teria que trabalhar com um número limitado de “*nodes*” (nós/conexões) e neste momento sendo exposto na análise acima proposta.

Outra análise a ser considerada é a de componentes conectados no perfil, onde os pesquisadores avaliaram, por exemplo, a análise de requisições de amizade e nódulos isolados e sem relacionamentos, que também poderiam ser considerados como critério para a observação de perfis falsos. Pessoas não ligadas a qualquer outra pessoa podem ser consideradas um indício (mas é apenas probabilidade). As análises foram realizadas com *data-sets* de perfis reais, após 1 dia e 30 dias da criação dos perfis, e demonstraram um “padrão” de componentes conectados:

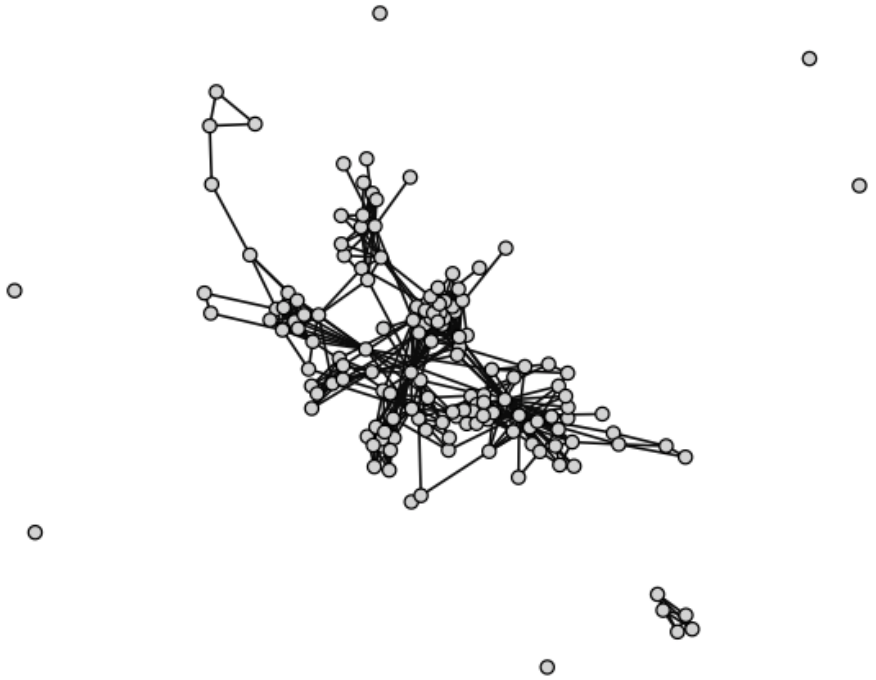
Figura 24 – Padrão de componentes conectados no dia 1 após a criação do perfil



(a) Day 1: Graph

Fonte: CONTI; POOVENDRAN; SECCHIERO (2012. p. 1078)

Figura 25 – Padrão de componentes conectados no dia 30 após a criação do perfil



(c) Day 30: Graph

Fonte: CONTI; POOVENDRAN; SECCHIERO (2012. p. 1078)

Percebe-se, porém que a proposta pode ser aprimorada, pois primeiramente tem sempre como base um perfil real e não se aplica a perfis imaginários, ou seja, onde não há furto de identidade, mas a criação de uma “identidade inexistente”. Por outro lado, a contribuição da proposta é justamente oferecer um caminho para a detecção de perfis falsos que não só considere apenas similaridades, mas a interatividade do perfil falso.

4.6 Método automatizado para detectar perfis falsos e *botnets* em redes sociais

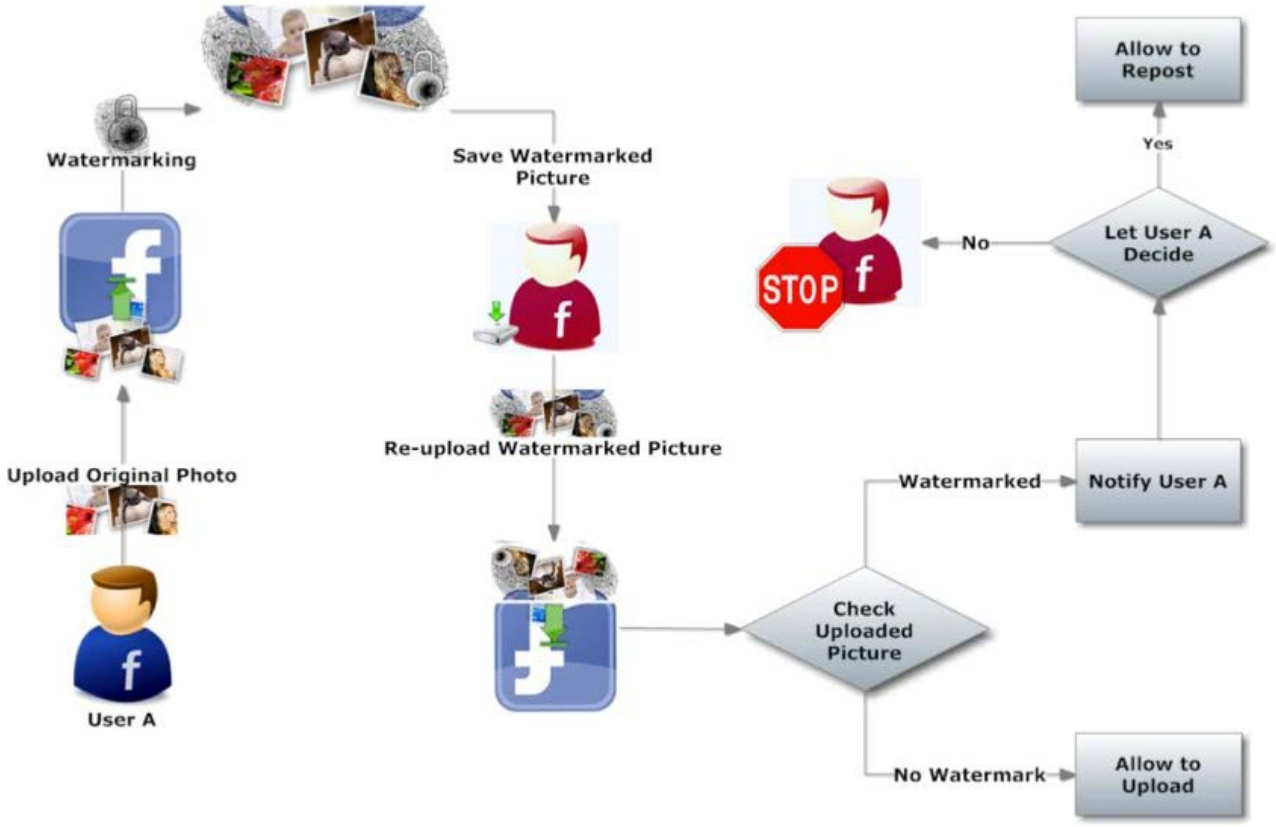
Ahmadizadeh et al. (2015) desenvolve importante contribuição nos estudos de modelos para detecção de perfis falsos em redes sociais. Segundo os

pesquisadores, dos 955 (novecentos e cinquenta e cinco) milhões de usuários ativos do Facebook, 8,7% (oito vírgula sete por cento) são perfis falsos.

Como se sabe, as atividades dos perfis falsos envolvem não só difamação e violação da privacidade, mas principalmente, podem espalhar spam, códigos maliciosos, etc. Enquanto redes sociais agem reativamente após a denúncia de perfis, a proposta é desenvolver um sistema automatizado e preventivo.

Considera-se que a foto de perfil é um ponto fundamental na criação de um perfil falso. Portanto, na abordagem proposta, poderia utilizar técnicas de esteganografia para, por exemplo, quando um usuário subisse (*upload*) uma foto, fossem gravados metadados como *username*, e-mail e data na imagem, o que seria uma espécie de rotulagem prévia dos objetos das redes sociais. Assim, se alguém salvar a foto e tentar subir novamente, o sistema automaticamente detectaria a fraude e impediria o upload. (AHMADIZADEH et al., 2015). A proposta, segundo os autores, pode ser ilustrada na figura 26.

Figura 26 – Detecção de perfis falsos com base em esteganografia e descrição prévia de objetos



Fonte: AHMADIZADEH et al. (2015, p. 68)

Porém, entende-se que este método seria de fácil adulteração com técnicas antiforenses para remoção de Exifs *(Excheangeable image file format).* Aliás, hoje as próprias redes sociais fazem o trabalho de sanitizar as informações de conteúdos que sobem no perfil. Para contornarem as limitações, os autores propõem uma checagem em tempo de “upload” ao Google Imagens, que remeteria o disparo de uma informação ao real proprietário da foto.

Assim, a proposta apresentada possui evidentes limitações, pois em tese depende da concordância do mantenedor da ferramenta de rede social em implantar o procedimento. Um descritor independente seria muito útil ou mesmo uma forma de preservar metadados em imagens que integram ferramentas de redes sociais.

Uma aplicação mediando a informação que coletasse tais dados, e quando alguém salvasse, registrasse o usuário que está salvando a informação seria um ponto interessante a se considerar e faz parte da proposta apresentada neste trabalho. Parte-se do princípio de que é necessário manter os metadados afastados do dado. Nesta pesquisa, apresenta-se um método para geração de metadados sobre conteúdos das ferramentas das redes sociais e a possibilidade de controle destes artefatos com base na análise dos metadados, de fundamental importância para detecção de falsidades.

4.7 Identificação de perfis falsos no Linkedin

Ao tratar especificamente da ferramenta de rede social Linkedin, Addikari e Dutta (2014) propõem uma sistemática para análise de perfis nesta que é considerada uma das principais redes profissionais do Brasil, onde elenca o número mínimo de dados necessários para identificar “*fake profiles*” e mais, apresentam um método de *data mining* para esta tarefa. A identificação de perfis falsos seria realizada com 84% (oitenta e quatro por cento) de precisão e 2,44% de falsos positivos.

A grande quantidade de dados despejada nas redes sociais torna difícil a detecção de atividades e comportamentos suspeitos na rede, bem como separá-los de comportamentos legítimos. Porém, comportamento do perfil *fake* é diferente de usuários legítimos. (KROMBHOLZ; MERKL; WEIPPL, 2012).

Uma das análises propostas consiste em selecionar os componentes comumente verificados nos perfis e comparar a variabilidade da existência dos

mesmos em relação aos dados existentes, com isso, identificando componentes que variam muito em relação a perfis falsos e autênticos.

Em outra análise, apresentam os pesquisadores Addikari e Dutta (2014) algoritmos de redes neurais aplicados aos *data-set*s. Ainda, apresentam uma análise com base em *Support Vector Machine (SVM)*, denominada *“polynomial kernel”,* técnica que, segundo os pesquisadores, apresenta a maior precisão e a menor taxa de falsos negativos.

Assim, os autores comparam técnicas de *data mining* para determinar a melhor abordagem para diferenciar perfis legítimos de ilegítimos no Linkedin.

A pesquisa não apresenta detalhes dos *Data-set*s, o que torna difícil sua reprodução em outro ambiente, bem como sua avaliação sobre importância de uso de seus conceitos nesta proposta de pesquisa. Igualmente, trata-se de uma pesquisa com foco apenas na rede social Linkedin.

4.8 Detecção de perfis falsos baseada na força do relacionamento

Rizi, Khayyambashi e Kharaji (2014) apresentam em sua pesquisa o que consideram uma nova abordagem para detecção de perfis clones nas redes sociais.

Nas ferramentas de redes sociais, milhões de pessoas estão compartilhando informações pessoais. De modo a evitar os chamados ataques de clonagem de perfis, é proposto um método com base na similaridade de perfis e na força dos relacionamentos.

Nas redes sociais o conceito de amizade nada mais representa do que um link baseado na confirmação de duas pessoas que enviam uma solicitação de amizade uma para com a outra e é diferente da amizade que ocorre no mundo real. Recentemente fora descoberto o ataque *Identify clone attack (ICA)*, que falsifica perfis de usuários nas redes sociais, com o objeto de obter informações pessoais dos amigos deste perfil ou prejudicar sua reputação. (RIZI; KHAYYAMBASHI; KHARAJI, 2014).

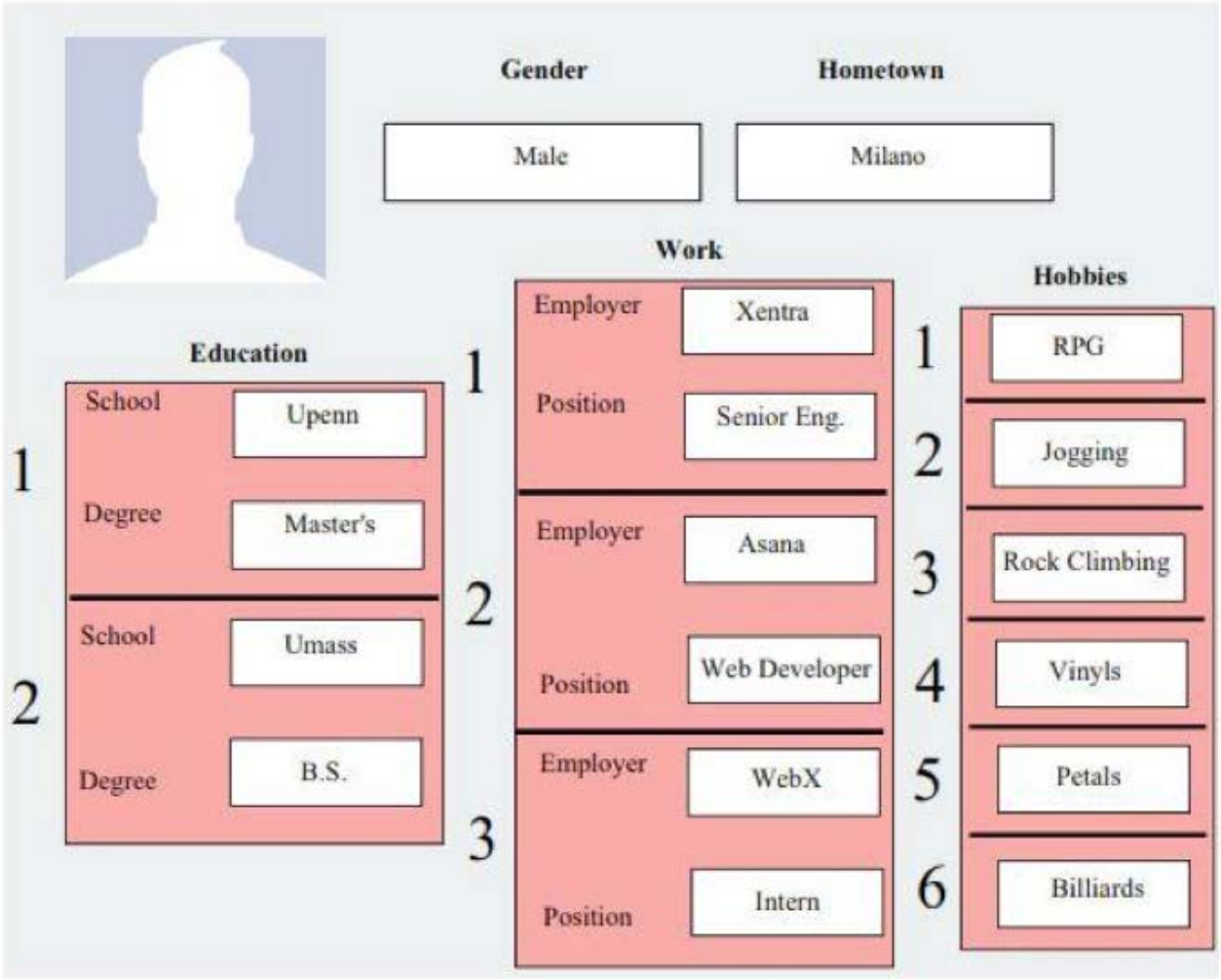
A proposta visa determinar qual de dois perfis é clone e qual perfil é efetivamente genuíno, garantindo-se assim uma rede social mais segura e motivada, que será capaz de promover um nível adequado de privacidade.

As pesquisas de Bhumiratana (2011) apresentam um modelo automático para criar perfis falsos de usuários reais, aplicando técnicas de ataque nas ferramentas

de redes sociais, apresentando ainda uma técnica para ativar perfis falsos nas comunicações reais de usuários. Por sua vez as pesquisas de Jin, Takabi e Joshi (2011) também apontam um *framework* para descobrir perfis falsos onde resta claro que não só um perfil pode ser falsificado, mas também uma lista de amigos.

Apesar das ferramentas de redes sociais apresentarem muitas diferenças, é comum o conceito de que cada usuário apresenta uma única entidade que o exibe, podendo apresentar sua imagem, bem como outros dados. Muitos perfis incluem duas partes, sendo elas campos de atributos e lista de amigos. Alguns campos de atributos podem ser visualizados na figura 27:

Figura 27 – Atributos comuns de um perfil



Fonte: RIZI; KHAYYAMBASHI; KHARAJI (2014, p. 28)

Deste modo, os ataques de clonagem de perfis podem ocorrer de duas formas:

- Na primeira forma, um perfil já existente na rede social é copiado e muitas solicitações de amizade são enviadas aos contatos da vítima (perfil real). Neste caso o atacante copia as informações do perfil real e começa sua trajetória para criar os mesmos relacionamentos da vítima. Estudos comprovam que usuários típicos aceitaram uma requisição de amizade de um usuário que já é confirmado em sua lista de amigos, o que revela falta de atenção. (RIZI; KHAYYAMBASHI; KHARAJI, 2014).
- Na segunda forma, usuários são registrados em algumas redes sociais, mas não em outras. O atacante então cria o perfil e reconstrói a rede de relacionamentos da vítima. Este ataque demonstra-se absolutamente efetivo. (RIZI; KHAYYAMBASHI; KHARAJI, 2014).

A proposta para detecção de perfis falsos de Rizi, KhayyambashI e Kharaji (2014) envolve as seguintes fases:

A) Coleta de perfis suspeitos: Conduzem-se buscas no serviço de pesquisas das redes sociais a partir dos dados dos perfis reais;
B) Avaliação dos perfis: Com base nos perfis coletados na fase anterior, os mesmos são avaliados com base na similaridade de atributos e na medida da força dos relacionamentos, em relação ao perfil real. É medido então se a força de relacionamentos de um perfil suspeito é menor que outros perfis.

Assim, a avalição dos perfis é medida, em síntese, com base na similaridade de atributos e na força dos relacionamentos dos perfis em análise, sempre em comparação entre o perfil real e o perfil supostamente falso. Como já tratado no presente trabalho, importante pesquisa, mas que também exige um perfil real para as comparações. Partirmos para os estudos em detalhes do modelo proposto pelos autores.

4.8.1 Avaliação da similaridade

A avaliação da similaridade é medida com base nos valores dos campos ou atributos de um perfil. Se ambos os campos possuem o mesmo valor, é retornado

1(um), já se os campos não têm o mesmo valor, é retornado 0 (zero). Porém estes campos em redes sociais podem ser valores únicos ou simples, ou valores múltiplos.

O cálculo da similaridade ocorre da seguinte maneira:

Figura 28 – Cálculo de similaridade entre dois perfis

$$Sim_{sv}(i_v, i_c) = \frac{1}{/Sb/} \times \sum_{n \in Sb} \begin{cases} 1 & if\ i_v^{(n)} = i_c^{(n)} \\ 0 & if\ i_v^{(n)} \neq i_c^{(n)} \end{cases} \quad (1)$$

Fonte: RIZI; KHAYYAMBASHI; KHARAJI (2014, p. 29)

Temos por “i” um valor único de um perfil (e que contém subcampos, denominados de “Sb”), assumindo que “iv” e “ic” constituem o valor único de “i” para os perfis “v” e “c”, respectivamente. O valor do subcampo “n” para o campo “i” é comparado para ambos os perfis “v” e “c”. Logo se iguais, retornará 1 e se diferente, retornará 0.

Logicamente, alguns itens do perfil podem ter mais que um valor, e então esta similaridade deve ser somada. Nestes casos onde existem mais itens para cada valor, a fórmula da similaridade aplicada é:

Figura 29 – Cálculo da similaridade de itens que podem ter mais de um valor

$$Sim_{iv}\ (v, c) = \frac{1}{|\text{values } (i_v)|} \times \sum_{h \in values\ (i_v)} max\ (SimItem(h))$$

Fonte: RIZI; KHAYYAMBASHI; KHARAJI (2014, p. 30)

Por fim a soma das similaridades deve ser realizada mediante a seguinte aplicação, onde βi e Simiv (v, c) são a importância e a semelhança dos itens “v” e “c” e “I” é item definido no perfil do usuário. Pela fórmula dos autores em estudo, o coeficiente βi (de importância) pode ser definido pelo usuário:

Figura 30 – Fórmula para soma das similaridades

For $v, c \in G$:

$$Sim_{prf}(v,c) = \frac{1}{/I/} \sum_{i \in I} \beta_i \times Sim_{iv}(v,c)$$

Fonte: RIZI; KHAYYAMBASHI; KHARAJI (2014, p. 30)

4.8.2 Medida de força dos relacionamentos

Após realizar a medição da similaridade, passa-se para a análise da força dos relacionamentos que considera que os links entre os perfis podem ser avaliados como suas relações de amizade.

a) A primeira medida realizada diz respeito a amigos ativos, que realiza a medição da frequência de interação de um usuário com sua rede de amigos, e pode ser representada por:

Figura 31 – Avaliação de amigos ativos em comum

$$F_i^a = F_i \cap I_i \qquad (4)$$

$$F_{ij}^a = F_i^a \cap F_i^a \qquad (5)$$

Fonte: RIZI; KHAYYAMBASHI; KHARAJI (2014, p. 31)

Temos para um usuário qualquer o valor “Fi”, que significa o “set de amigos” e “Fi*a*” que representa a interação computada entre os amigos “Fi” e os amigos do usuário que foram contatados por ele ou interagiram com ele de alguma forma, com postagens em mural, comentários, *tags*, dentre outros, que denominamos de “Ii”. Assim, para um usuário o valor de “Fia” (amigos ativos) representa justamente este dado.

Tendo por base os dados “amigos ativos” de um usuário, os amigos ativos em comum também podem ser calculados “Fija”, e será feito, como na Figura 31, a partir

da intersecção entre os amigos ativos de um usuário “Fia”, em relação aos amigos ativos de outro usuário “Fja”. No caso, a comparação é feita entre o perfil original e o suposto perfil falso.

b) A segunda medida para se apurar a força da rede de relacionamentos de um perfil, no escopo da detecção do perfil falso, são as preferências de páginas, que como na fórmula apresentada, se dá com o cálculo das “páginas comuns curtidas” entre o perfil oficial e o suposto *fake*, a seguir representada:

Figura 32 – Cálculo de páginas em comum curtidas

$$P_{ij} = P_i \cap P_j \qquad (6)$$

Fonte: RIZI; KHAYYAMBASHI; KHARAJI (2014, p. 31)

c) Do mesmo modo, avalia-se também o padrão de compartilhamento de URLs entre os usuários investigados, calculando-se a intersecção entre as URLs compartilhadas pelos usuários em análise. Este atributo de URL é calculado como uma fração de links usados por ambos, e pode ser representada pela seguinte fórmula:

Figura 33 – Cálculo de similaridade com base nas URLs compartilhadas

$$U_{ij} = \frac{U_i \cap U_j}{U_i \cup U_j} \qquad (7)$$

Fonte: RIZI; KHAYYAMBASHI; KHARAJI (2014, p. 31)

Apurados então os 3 (três) itens para análise de força dos relacionamentos é hora de atribuir um peso *(w(eij)* que é calculado pelos autores estudados com base na soma do valor de cada um dos itens apurados, mediante a seguinte fórmula:

Figura 34 – Fórmula para cálculo do peso da força dos relacionamentos

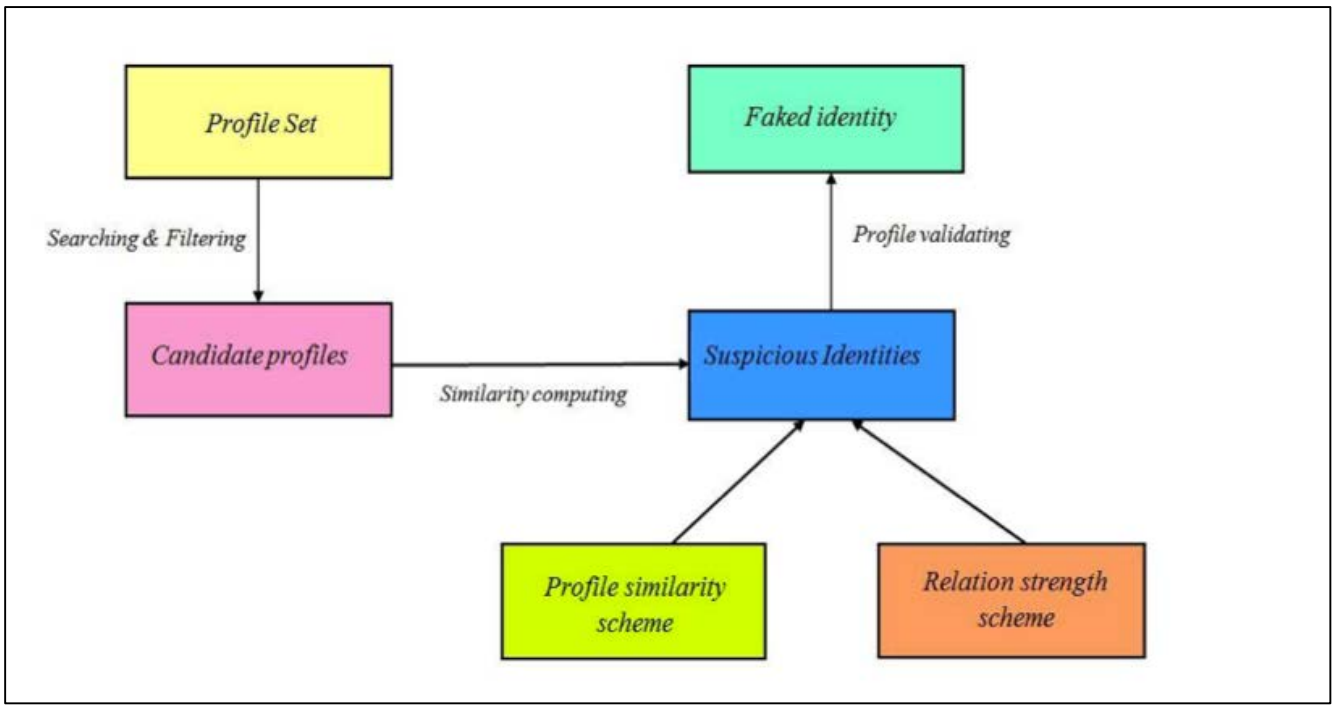
$$w(e_{ij}) = |F_{ij}^{a}| + |P_{ij}| + |U_{ij}| \qquad (8)$$

Fonte: RIZI; KHAYYAMBASHI; KHARAJI (2014, p. 31)

Assim, para cada perfil suspeito (edge) é atribuído um peso que mede a força do relacionamento. Deste modo, é comum a um atacante tentar passar despercebido tentando se relacionar com amigos da vítima. Embora possa passar despercebido por pessoas, por meio da técnica proposta pelos autores em estudo, fica claro identificar as similaridades e dissimilaridades que podem identificar a existência de um perfil falso.

Portanto, quando a similaridade de um perfil do candidato é maior do que um ponto definido e a força das relações são menores do que outros, este perfil é considerado suspeito e cada identidade suspeita deve ser validada. Em síntese, a proposta ora detalhada é exemplificada por meio do diagrama na figura 35.

Figura 35 – Diagrama da detecção de perfis falsos com base em similaridades e força de relacionamentos



Fonte: Fonte: RIZI; KHAYYAMBASHI; KHARAJI (2014, p. 34)

A grande deficiência da pesquisa é que realiza comparações necessariamente entre um perfil real e um perfil suspeito. Outra deficiência é que se faz necessário o acesso aos *data-sets* de dados que hoje são restritos aos titulares e que não são acessados via APIs, pois envolvem privacidade.

Ademais, basear a força dos relacionamentos apenas em links compartilhados em comum pode apresentar falsos positivos e negativos.

5 Uma proposta de modelo conceitual para o desenvolvimento de aplicações computacionais ou *frameworks* para inteligência cibernética e detecção de falsidades nas redes sociais

Baseado nos conceitos, técnicas e áreas anteriormente pesquisadas neste trabalho, este capítulo tem por escopo apresentar uma proposta de modelo conceitual para investigação digital e exercício da atividade de inteligência cibernética em redes sociais, de modo a compreender o Big Data gerado nestes ambientes.

Da mesma forma, deve o modelo conceitual contribuir para aplicações que visem implementar ferramentas de redes sociais mais verdadeiras e seguras, com maior credibilidade, permitindo aos seus usuários investigar perfis falsos, considerando a ausência de pesquisas neste tema e a infinidade de informações a que os usuários são submetidos, diariamente, nas ferramentas de redes sociais.

A proposta vale-se das pesquisas já realizadas e neste trabalho revisadas em detalhes, sobre o tema, aproveitando conceitos e técnicas, tecnologias e do aporte teórico da Ciência da Informação, grande diferencial na construção um modelo de interface que organizando e interpretando o significado dos dados, seja capaz de predizer ou informar a probabilidade de um perfil falso.

Seguindo o entendimento de Santarém Segundo (2010), a recuperação do significado da informação não pode se limitar a comparações sintáticas de caracteres e objetos textuais. Se assim agíssemos, estaríamos a desconsiderar polissemias, homógrafos e palavras diferentes com mesmo significado.

Neste sentido, a proposta de um modelo de detecção de falsidades em redes sociais pode ser formulada como um modelo conceitual teórico para construção de aplicações e *frameworks*. Um *framework* provê as funcionalidades comuns de várias aplicações pertencentes a um mesmo domínio de problema, fornecendo uma solução para uma família de problemas semelhantes. Trata-se de uma abstração que reúne código comum de vários projetos de aplicações, concebendo uma funcionalidade genérica. (MULLER, 2008).

5.1 Especificações para aplicações e critérios para detecção de perfis falsos

Apresentam-se neste ponto, antes da exposição do modelo conceitual proposto, critérios e especificações importantes a aplicações que sejam construídas com a finalidade de detecção de perfis falsos. Inicialmente deve-se conceber que uma ferramenta para detecção de perfis falsos poderá se dar através de duas modalidades de recuperação de informação:

- **Interface *web* disponível para consulta:** Onde o usuário pode indicar a URI ou URL de determinado objeto e consultar a falsidade ou não; Haverá um usuário responsável por desenvolver a interface entre o pesquisador ou sistema e a interface de dados das ferramentas de redes sociais;
- **Processamento nas ferramentas de redes sociais:** Onde uma aplicação varre um ou mais objetos em ferramentas de redes sociais em busca de perfis falsos ou na rotulação de objetos.

Em ambas as interfaces, a aplicação deverá considerar as fases de: **a) preparo; b) coleta; c) processamento; d) análise; d) informe e armazenamento.**

Em cada fase, alguns itens devem ser observados, sendo eles:

a) **Preparo:** Na fase de preparo o desenvolvedor deverá atentar para as seguintes questões:

- *Aspectos legais:* Deverá considerar e revisar a legislação no que diz respeito à coleta de dados, revisando inclusive os termos de uso da ferramenta de rede social, evitando-se violações à privacidade;
- *Conectores:* Deverá avaliar quais mecanismos a ferramenta de rede social disponibiliza para a interação com dados. Identificar quais pontos existentes permitem o envio e recebimento de dados por meio de aplicativos, identificando ainda quais dados são liberados e quais são protegidos (dependendo da aceitação do usuário);
- *Ontologias:* Deverá ser descrito pelo profissional de Ciência da Informação a ontologia do domínio falsidade/fraudes e dos objetos de redes sociais, que

conterá classificação de itens e sua associação a outros itens que informem a probabilidade de perfil falso ou outras fraudes;

- *Descritor independente:* A aplicação deve contar com módulo preventivo, onde o usuário pode submeter seus ativos, perfis e objetos pessoais à descrição independente, que indexará os elementos do seu perfil, que ficarão armazenados na aplicação e servirão de base para detecção de fraudes, ou usos indevidos;
- *Descritor de objetos:* A aplicação deverá gerar arquivo RDF descrevendo cada objeto coletado do Big Data das redes sociais, possibilitando consultas e inferências por meio de linguagens especializadas.

b) **Coleta:** Na fase de coleta de dados para análise de perfis falsos deverão ser observados os seguintes requerimentos:

- *Formas de coleta:* A coleta poderá se dar por adição do perfil falso como “amigo” e criação de um node, poderá se dar pelo uso de APIs da rede social ou ainda poderá se dar pela captura (*crawling*) de dados exibíveis ao público. Para cada ferramenta de rede social, será necessário um crawler específico, que será parametrizado pelo profissional de informação. Outra forma de coleta ainda é pelo “consentimento”, convidando ou solicitando ao investigado que instale aplicativo destinado a esta finalidade. Para cada modalidade é preciso avaliar as implicações legais e obter as autorizações necessárias; na ausência de APIs e possibilidade de crawler o “sequestro de sessão” de um usuário suspeito só pode se dar ordem judicial. Pesquisadores denominam esta técnica de *network traffic analysis.* (CANALI; COLAJANNI; LANCELOTTI, 2011).
    - Aplicativo de extração: Deverão ser observadas as políticas das ferramentas de redes sociais, de forma que a aplicação não seja desativada por atividade suspeita ou hostil; O módulo de extração deve coletar conteúdos textuais, imagem e vídeo, links, organizando-os por data, autor e por tipo de conteúdo.

c) **Processamento:** Na fase de processamento de dados para análise de perfis falsos deverão ser observados os seguintes requerimentos:

- *Base de conhecimento:* Já na fase de processamento podem ser fornecidos objetos de base para comparação (uma foto, um padrão textual). Neste momento, uma ontologia pode descrever itens associados a um domínio

específico de falsidade, podendo auxiliar para classificação em tempo de coleta; A base de conhecimento conterá os dados coletados e organizados de outras atividades na ferramenta de rede social;

- *Metadados:* De acordo com a ferramenta de rede social, deverão ser realizados testes para se identificar quais metadados gravados em objetos permanecem intactos no *upload* à ferramenta, quais metadados são sobrescritos, se o resumo criptográfico dos arquivos é alterado com o upload, se existem *thumbnails* (miniaturas de objetos de imagem) acessíveis diretamente, mesmo após o conteúdo excluído, dentre outros itens;
- *Avaliação dos metadados:* Se metadados permanecerem após upload, o profissional de informação deverá arrolar quais são eles, tamanho em bytes, localização, descrevendo ainda se podem ser utilizados como indexador ou não; A aplicação deverá facilmente retornar, via pesquisa, dois conteúdos com mesmo metadado, o que pode inferir uma possível fraude;
- *Hashing:* Na fase de processamento é necessário também gerar a soma criptográfica dos objetos do perfil base e cadastrar no banco de dados. Neste ponto, a Ciência da Informação pode contribuir avaliando se o *hash* do arquivo é suficiente ou, considerando que o mesmo pode ser alterado, direcionar o processamento do *hash* para um resumo matemático diferenciado, que considere aspectos visuais da imagem;
- *Marcas d´água*: Neste momento, pode-se avaliar a possibilidade de inserção de marca d'água que possa ser reconhecida pela ferramenta de processamento;
- *Contorno de zeração de metadados:* Uma das alterativas identificadas em ferramentas de redes sociais que apagam os metadados, é realizar o upload do objeto a ser monitorado e após capturá-lo com a aplicação, onde o profissional de Ciência da Informação deverá avaliar quais metadados gravados pela própria rede podem servir de índice para o arquivo.
- *Coleta de padrões textuais:* A ferramenta não deve processar apenas imagens, mas deve coletar padrões textuais, expressões, postagens, links, comentários e outros textos que serão submetidos às ontologias e às sugestões folksonômicas sobre um objeto ou postagem. Por exemplo, um link suspeito pode estar cadastrado na ontologia de fraude, associada a um perfil que esteja também referenciado na base de "perfil falso"; links podem ser comparados em bases de

mensagens maliciosas e *blacklists*. (AGGARWAL; RAJADESINGAN; KUMARAGURU, 2013).

- *Processamento:* A aplicação deve ter a capacidade de gerar “loops” para processar sets de perfis submetidos para análise, processando grandes volumes de dados ou postagens em tempo de publicação.

d) **Análise:** Na fase de análise de perfis falsos deverão ser observados os seguintes requerimentos:

- *Ontologias:* Deverá se descrever as ontologias por tipos de “falsidades” comuns em ferramentas de redes sociais, como exemplo “desinformação”; “pornografia infantil”, “código malicioso”, “links maliciosos” e dentro destes domínios especificará conteúdos, identificados e associados ao tema (esta especificação poderá se dar com *folksonomia* sugerida que considere comentários de postagens); em comparação com ontologias de perfis falsos ou robôs, poder-se-á, pelo conteúdo cadastrado em uma ontologia de falsidade, identificar um perfil *fake* associado;
- *Critérios para análise:* Dentre os critérios para análise de perfis falsos que se identificou na presente pesquisa, sugere-se às aplicações que adotem:

a) Análise em bancos de imagens e buscadores de conteúdos similares (incluindo cache para conteúdos apagados);
b) Análise da progressão em *nodes* e amigos;
c) Análise do tempo de postagens;
d) Análise de conteúdo de comentários;
e) Análise de conteúdos das postagens em se relacionando com as ontologias;
f) Análise da repetição de postagens;
g) Análise de dados similares na mesma ou em outras ferramentas de redes sociais;
h) Extração de metadados e EXIFS e comparação com dados pré-rotulados;
i) Técnicas de *data mining* podem ser usadas para extrair características mais comuns de perfis *fakes* e *bots,* contribuindo para aprimoramento das regras de análise;
j) Comparação de atributos entre perfil oficial e suspeito, com geração de um modelo de aprendizado que pode servir de base a comparações futuras;

k) Adoção de técnica envolvendo coleta de postagens de um perfil base e com a análise do perfil, se traçar o *profile* comportamental, formando um modelo estatístico que considere, por exemplo, data das postagens, aplicação usada para postar, links na mensagem, navegador utilizado, linguagem, etc. Com isso, detecta-se anormalidades que possam indicar o uso indevido de um perfil ou perfis falsos. (EGELE et al., 2013).

a) **Informe e armazenamento:** Na fase de informe e armazenamento, na análise de perfis falsos, deverão ser observados os seguintes requerimentos:

- *Critérios:* A ferramenta deve descrever os critérios e pesos utilizados para apurar a probabilidade de um perfil falso ou *fake*;
- *Opções:* Devem ser oferecidas ao titular do conteúdo, identidade ou perfil usurpado, opções para denunciar, ou caso implementada pela ferramenta de rede social, esta deve oferecer a possibilidade de prevenir a inserção do conteúdo. Deve ser disponibilizada interface para que o usuário possa comprovar sua identidade ou propriedade sobre o conteúdo;
- *Relatório:* Com base nos registros, deve ser possível especificar, a localização geográfica ou mapa de *fakes* usando conteúdos de terceiro, identificando ainda a origem e com análises de disseminação e viralidade, alcance de conteúdos e dados pessoais copiados indevidamente. Deve ser possível indicar, dado um objeto (Ex. uma foto), quantas versões copiadas existem nas redes sociais;
- *Volatilidade:* Perfis falsos não permanecem no ar por muito tempo, neste sentido a aplicação deverá gerar um *screenshot* do perfil e armazená-lo, permitindo a reprodutibilidade ou mesmo gerando provas de que existiu, registrando seus identificadores como user id, dentre outros;
- *Pesquisa semântica:* Deve ser disponibilizado um repositório dos dados em padrão RDF, online, para que aplicativos, usuários e redes sociais possam consultar objetos, perfis falsos e controlar o uso de seus dados.

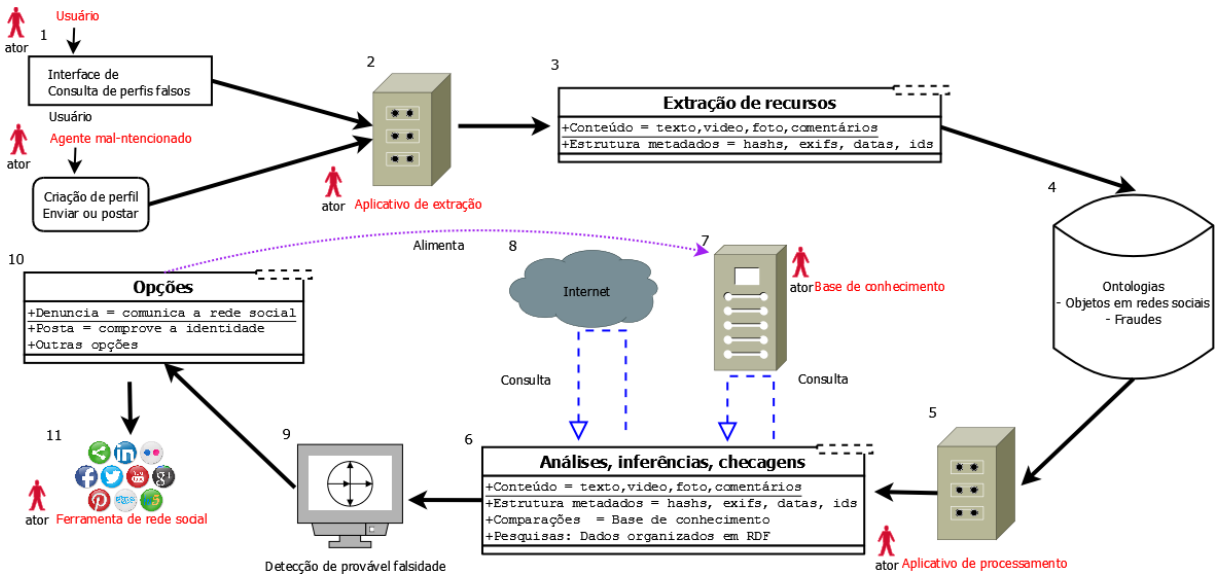
Como se verifica, a proposta de modelo conceitual para referência em aplicações, dividida em fases, vem no escopo de organizar a informação e possibilitar meios para que usuários detectem *fakes* e controlem parte de seus recursos nas ferramentas de redes sociais.

O profissional de Ciência da Informação demonstra-se indispensável em todas as fases, trabalhando na arquitetura da aplicação, na construção das interfaces, avaliando os metadados possíveis de serem coletados, mensurando a questão dos dados públicos e pessoais, oferecendo parâmetros para o profissional de computação, definindo e relacionando as ontologias e realizando ajustes nos projetos, com vistas à melhor coleta, análise e apresentação das ocorrências de falsidades.

5.2 O modelo conceitual

Na presente pesquisa, avança-se em relação aos trabalhos de detecção de perfis falsos nas redes sociais, aqui amplamente revisados, sobretudo por envolver tecnologias da Ciência da Informação na organização de dados destes ambientes. O modelo proposto pode ser representado pela figura 36.

Figura 36 – Modelo de inteligência cibernética e uso de recursos semânticos na detecção de perfis falsos



Fonte: Elaborado pelo autor

Podem-se definir os atores que fazem parte do modelo conceitual da Figura 36 na seguinte classificação:

Ferramenta de rede social: Aplicação que permite a interação entre usuário por meio de plataforma na Internet;

Usuário: Pessoal responsável por consultar um perfil “suspeito” na interface de consulta ou por criar um perfil ou postar um conteúdo na rede social;

Modelo de inteligência cibernética: Conjunto ordenado de práticas e atividades para proporcionar ao usuário a detecção de perfis falsos;

Agente mal-intencionado: Usuário que se vale das ferramentas de redes sociais para criação de perfis ou postagens de conteúdos falsos;

Aplicação (aplicativo de extração e aplicativo de processamento): Sistema ou agente computacional dividido em módulos de “App de extração” e “App de processamento”, como o escopo de coletar dados e metadados de objetos nas redes socais, organizar as informações e realizar as análises de inferência para detecção de possíveis perfis falsos;

Base de conhecimento: Banco de dados armazenando conteúdos e metadados de conteúdos, descritos em linguagem padronizada e disponíveis a agentes computacionais.

Assim, detalhando o fluxo da Figura 36 e os atores que interagem com o fluxo, temos a seguinte dinâmica, associada à figura exposta pela numeração:

a) Inicia-se o processo de duas formas. Um usuário ou usuário mal-intencionado envia um conteúdo para a ferramenta de rede social, ou cadastra um novo perfil. A segunda opção disponível para começar o processo é o módulo “Consulta perfis falsos” em que um usuário pode checar informações sobre um perfil suspeito (1);
b) Neste momento o aplicativo de extração (2) realiza a coleta dos metadados disponíveis para o recurso (*hash*s, códigos, marca d’água, datas), além do próprio conteúdo do recurso (texto, foto, vídeo, comentários) (3);
c) A descrição das informações será organizada de acordo com ontologias pré-estabelecidas, com foco na classificação de perfis falsos e fraudes (4).
d) Depois de descritos e identificados como conteúdo e estrutura informacional, o aplicativo de processamento é acionado (5), onde atuará realizando testes de duas maneiras (6), cadastrando os recursos na “Base de conhecimento”(7), sendo:

- **No conteúdo, baseado em sintaxe:** buscará a existência de conteúdos similares em outras bases de dados, pesquisas em buscadores e na própria rede social; pesquisará por dados do perfil similares, dados pessoais, etc.,

poderá utilizar comentários e palavras nas postagens para sugestões folksonômicas do tipo de conteúdo existente;

- **Na estrutura e no conteúdo com apoio de recursos semânticos:** realizará inferências baseadas em consultas às ontologias e por conteúdo associado a fraudes, golpes, perfis falsos, coletará data dos recursos, identificará imagens e conteúdos com os mesmos metadados de imagens (EXIFS), como por exemplo, mesma máquina fotográfica, mesmo software editor de imagens, etc.

e) O aplicativo poderá realizar consultas na Internet para detecção de similaridades (8);

f) Com base nos resultados das análises (9), a aplicação então poderá oferecer opções (10) ao usuário (que pode ser um usuário mal-intencionado) que posta um conteúdo/cria um perfil ou consulta um perfil, sendo elas:

- **Denúncia:** O usuário poderá denunciar o perfil falso pesquisado ou a própria rede social emitirá alerta de provável *perfil fake*;
- **Postagem:** Neste momento ele deverá provar sua identidade, se provada, o sistema irá cadastrar todos os objetos a um perfil real e ficará de base para futuras comparações com perfis suspeitos;
- **Opções:** Outras opções poderão ser adicionadas ao sistema.

g) Feita a opção e de acordo com a validação, o conteúdo é liberado para ser postado ou publicado na ferramenta de rede social (11).

Assim, além da descrição do fluxo para recuperação e rotulagem da informação nas redes sociais, o modelo conceitual aqui proposto deve considerar os seguintes pontos, seguir os métodos e considerar as seguintes análises:

a) Inicialmente, a proposta apresentada contorna as restrições das redes sociais em acessar metadados sobre informações postadas (considerando que as redes sociais em tese limpam os metadados de arquivos, sobrescrevendo-os para outros metadados), eis que todo o trabalho pode ser feito por “agente intermediário” ou “descritor independente”, podendo se constituir em uma aplicação instalada. Logicamente, um usuário mal-intencionado, não irá instalar o agente intermediário na construção de um *fake*, mas a vítima pode

ter usado (descrevendo os dados e armazenando-os na base de inteligência da aplicação), não só descrevendo e protegendo seus dados de uso indevido, mas, principalmente, contribuindo para rotulagem das informações nas redes sociais e detecção de falsidades. Um aplicativo ou ferramenta de rede social pode ser programado para que o download de conteúdos de um perfil permita apenas downloads de objetos rotulados; além disso, diante da suspeita de um perfil falso, qualquer pessoa poderia indicar a URL (*Uniform Resource Locator*) na interface de consulta “Consulta perfis falsos” que poderá identificar se os objetos daquele perfil constituem objetos já rotulados, fornecendo elementos para detecção de falsidade;

b) A interface de entrada de uma informação (dado que se quer proteger, seja uma foto, um vídeo, um post, etc.) pode ser a aplicação intermediária. O usuário também pode consultar em um portal que disponibilizará na Internet a probabilidade de perfis falsos ou mesmo a anterioridade de recursos em redes sociais (fotos, vídeos, etc.);

c) Submetido o conteúdo à aplicação, extrai os metadados dos documentos ou objetos, inclusive calculando o *Hash* dos arquivos (resumo criptográfico), textos, mídias, dentre outros. Posteriormente os recursos são classificados de acordo com as ontologias e na sequência os agentes computacionais (App de processamento) realizarão as inferências e análises sobre o conteúdo. O conteúdo, dados e inferências integrarão a “base de conhecimento” para confrontação com outros conteúdos e para novas análises de inferências;

Os dados extraídos são os dados passíveis de serem capturados mediante interação com as interfaces das redes sociais (o que irá variar de rede para rede), seja a foto, seja vídeo ou imagens. Deste modo a aplicação construída deve seguir este modelo e considerar a coleta de:

- **Conteúdo:** O conteúdo pode ser o texto de uma postagem, uma foto, vídeos ou perfis. A coleta de conteúdo se dará pela aplicação interagindo com os conectores das ferramentas de redes sociais e em casos de limitações de redes sociais, poderá ocorrer o *crawling* (captura direta).
- **Os metadados (estrutura informacional):** A URL da postagem, nome do arquivo, data da postagem, tipo da postagem, usuário responsável pela postagem. Neste contexto poderão ser utilizadas ferramentas que atuem

como o software Findmyfbid[25] que, dado uma página ou perfil, forneça um índice único (ID) para este recurso na rede social. Deve ser realizada pesquisa por conteúdos similares nos principais buscadores, sendo as ocorrências listadas e cadastradas na ferramenta. Igualmente, neste momento, serão extraídos os EXIFS do conteúdo ou postagens envolvendo imagens.

- **O resumo criptográfico**: Em conteúdos envolvendo texto, imagem ou vídeo, é preciso gerar também um resumo criptográfico (ou *hash code)* e armazená-los como metadados relativos aos objetos(URIs). Eles serão igualmente membros da ontologia e descrito por linguagem RDF. O resumo funciona como uma espécie de assinatura do objeto, sendo que em tese seria possível identificar outro objeto com o mesmo *hash* circulando pela rede social. O *hash* também servirá para declaração de conteúdos originais, sendo que conteúdos adulterados ou pirateados nas redes sociais não terão o mesmo resumo criptográfico. Neste sentido, esclarece Pereira (2009, não paginado) que

> Assim, antes de alguém começar a baixar aquele conteúdo, o programa utilizado pede a mesma sequência de letras e números para a máquina que está servindo o arquivo: se a sequência estiver diferente, significa que o arquivo foi alterado e que não deve ser baixado. Caso a conferência esteja correta, a transferência é iniciada. Essa sequência é conhecida como "*Hash* info" e pode ser facilmente verificada em indexadores como o The Pirate Bay, exibido na imagem acima. Porém, o uso dessa técnica vai muito além dos torrents e é conhecida no mundo da informática como *hash* ou *hash*sum.

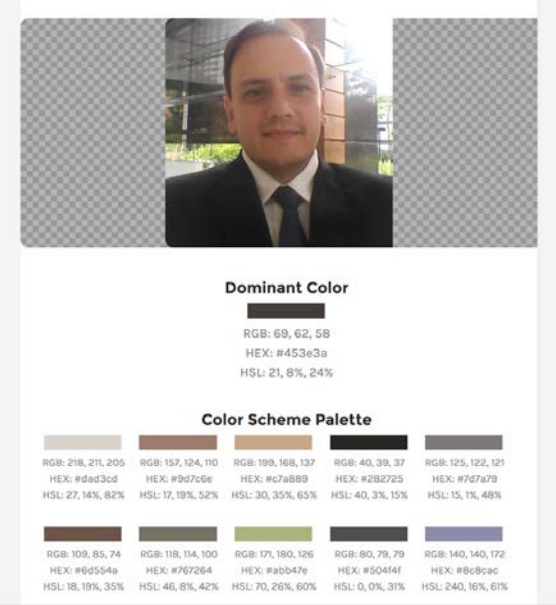
Logicamente, o resumo criptográfico pode ser diferente a cada re-upload (nova inserção) da mesma imagem, o que demanda a comparação com outros elementos do conteúdo ou que seja gerado a partir de outros elementos do conteúdo não modificáveis, porém, é um elemento identificador que deve ser considerado em ferramentas com as quais se pretende analisar perfis falsos. Para geração do resumo, podem ser utilizadas ferramentas como OnlineMD5[26].

- **O código das cores:** Para os conteúdos envolvendo imagens, deve haver a extração do código de cores das imagens (RGB, HEX, HSL), em ferramentas

[25] http://findmyfbid.com/
[26] http://onlinemd5.com/

como Color Code Picker[27], permitindo-se mapear a codificação dominante de uma imagem que serviria de índice ou identificação para imagens semelhantes ou idênticas:

Figura 37 – Extração da codificação de cores da imagem



Fonte: Elaborado pelo autor por meio da ferramenta Color Code Picker

Como se pode verificar, é muito difícil duas imagens distintas terem a mesma codificação de cores. Assim, este código pode ser comparado a um identificador da imagem. Se alguém salvar uma imagem cujo código de cores (metadado) foi registrado e tentar subi-la novamente na rede social, a aplicação poderia detectar a atividade. Tal atividade pode ser utilizada para identificar imagens de vítimas usadas para criação de perfis falsos.

d) Com base na extração de conteúdo e estrutura, cadastrados de acordo com a classificação definida nas ontologias, a aplicação rotularia as postagens e objetos do perfil, auxiliando na recuperação de informação e comparação entre estados ou detecção de padrões *fakes*, *bots* ou crimes;

As ontologias de suporte a aplicações com as quais se pretendam detectar perfis falsos, podem se estruturar da seguinte forma: a ontologia de objetos de redes sociais, que coletará e identificará: textos, vídeos, imagens e associará ao perfil responsável. Igualmente, outra base de fraudes classificará conteúdos fraudulentos

[27] http://www.colorcodepicker.com/

na rede. Em comparações e inferências, podem-se identificar conteúdos relacionados a perfis falsos ou vice-versa.

Definir-se-á uma ontologia para classificação de objetos de redes sociais, de modo que seja possível identificar com as informações disponíveis que tipo de conteúdo está sendo tratado. Neste contexto é importante identificar e organizar informações como: perfil associado (uid, url); data de criação; atributos (será avaliado de acordo com a rede social, por exemplo, número de amigos); data de postagem; tipo de objeto (*link,* vídeo, texto, imagem); resumo criptográfico; metadado 1 (autoral); metadado 2 (outros critérios a serem extraídos); conteúdo (texto, nome do arquivo, etc.); código de cores; identificadores em buscadores (sim ou não).

Outra ontologia a ser definida para uso neste modelo trata da organização, identificação e classificação de fraudes em redes sociais, sendo elementos importantes em sua construção: tipo de fraude (autoral, financeiro, pornografia, *fake*); conteúdo; palavras associadas (rol de palavras associadas à fraude); identificador do objeto (*link*, postagem, imagem, vídeo, etc);

Trata-se de sugestões de ontologia, que poderão ser derivadas e ajustadas à realidade de cada ferramenta de rede social.

Neste contexto, por inferência, podem-se descrever objetos de redes sociais, e na sequência compará-los com ontologias de fraudes associadas, o que poderá revelar um conteúdo (foto, texto, vídeo) associado a um perfil falso ou reutilização de objetos para construção destes perfis. No mínimo, ter-se-á maior controle dos objetos que se divulga e publica em redes sociais. Detalhe interessante é que mesmo excluído o conteúdo após ser postado em uma ferramenta de rede social, o registro RDF continuaria ativo e pesquisável por ferramentas, contribuindo para análises estatísticas sobre fraudes e perfis falsos ou mesmo para comprovar que uma fraude existiu.

A aplicação poderá, por exemplo, detectar dois perfis usando um mesmo objeto (ex. foto) e neste sentido, avaliar outros metadados como “data de criação” para verificar qual perfil pode ser considerado *fake*. Igualmente, poderá listar rapidamente conteúdos imagéticos que foram encontrados em buscadores (atividade comum na construção de um perfil *fake*).

e) Assim, a base se vale de ontologia (e descrição de vocabulários) específica e submete os dados para processamento, onde será checado o match de *hash*es, outras imagens semelhantes, outros dados idênticos, outros textos, metadados, dentre outros. O conteúdo deverá ser descrito por meio de linguagem RDF e deve ficar disponível em repositório adequado para que a própria aplicação pesquise novos conteúdos que serão inseridos ou verificados na ferramenta de rede social; Todo o conteúdo poderá ser pesquisado usando-se a linguagem Sparql, em ambientes definidos com bancos de dados de triplas, incluindo ambientes com estrutura do tipo Sparql-endpoint bastando-se indicar a URI dos RDFs gerados pela aplicação. Os RDFs serão atualizados diariamente;
f) De acordo com o identificador, o sistema deverá emitir um aviso informando perfis ou o uso indevido ou duplicado do conteúdo, possibilitando ao usuário titular visualizar o material ou os materiais similares e realizar uma denúncia. O usuário que está subindo o conteúdo, igualmente, receberá um aviso para comprovar sua identidade, caso tenha a aplicação instalada. Um terceiro recurso disponível é que pessoas possam consultar perfis e objetos na ferramenta identificando conteúdos rotulados ou semelhantes, momento em que o sistema apresentará as relações que identificou;
g) As aplicações baseadas no modelo aqui proposto poderiam, igualmente, automaticamente, denunciar à ferramenta de rede social a existência de um perfil supostamente falso, para as providências.

Assim, como não se podem obrigar pessoas (ou usuários mal-intencionados) a utilizarem uma aplicação intermediária para descrição de recursos, as pessoas seriam livres para usar somente os conteúdos que desejam proteger ou acompanhar replicações indevidas, o que pode ser utilizado na detecção de perfis falsos, mas também na proteção intelectual. Outra opção seria as próprias ferramentas de redes sociais instalarem o recurso.

A proposta poderia ser implementada por ferramentas de redes sociais que pretendam assegurar maior privacidade, veracidade, e evitar desinformações oriundas de robôs e perfis falsos. Difere-se das pesquisas revisadas neste trabalho, pois permite identificar, por meio das ontologias, metadados e descrição de objetos descentralizada, robôs e *fakes*, não somente com base nos seus atributos (dados),

mas em relação ao seu conteúdo compartilhado, resumo criptográfico e demais metadados, por meio de análises de inferência.

Mesmo sabendo-se que algumas ferramentas de redes sociais excluem os metadados de uma imagem que é inserida (excluindo, por exemplo, metadados relativos ao programa de edição utilizado na imagem), eles seriam coletados pelo software desenvolvido com base no modelo aqui proposto. O sistema poderia realizar buscas de imagens similares e semelhantes e em sendo localizadas em redes sociais e perfis, imediatamente realizaria a comparação (o que pode ser feita mediante consulta aos RDFs gerados).

Parte-se do princípio que o titular de um recurso ou objeto é o primeiro a publicá-lo, embora possam ocorrer casos em que o usuário mal-intencionado utiliza recursos da vítima para criação de contas em outras ferramentas de redes sociais que a vítima, ou seja, o usuário não tenha conta. Assim, diante de um perfil suspeito, basta inseri-lo na ferramenta proposta, e o sistema irá buscar referências primárias para os objetos utilizados, apontando então o “real” (ou mais antigo) titular dos objetos que estiverem na base de dados. Mais que isso, os softwares que pretendam implementar este modelo conceitual devem considerar o conteúdo do perfil suspeito, submetendo-os às ontologias que possam organizar os dados para que aplicações possam sugerir grau de integridade do perfil avaliado.

Mesmo que um usuário mal-intencionado não use a aplicação, se ele se valer de objetos que já foram rotulados pela aplicação, em algumas redes sociais (que permitem a inserção de metadados em objetos), seria possível detectar e alertar o titular dos dados sobre possível upload suspeito, tão logo o conteúdo suspeito estivesse sendo inserido.

Para redes que não permitem metadados em objetos, a vítima ou um terceiro poderia localizar o uso indevido dos mesmos, consultando perfis suspeitos na aplicação (que segue o modelo conceitual aqui especificado). A proposta vale também para objetos já nos servidores do Facebook, que poderiam ser baixados, classificados e indexados e reinseridos, desde que haja permissão do usuário.

Dentre as limitações do modelo encontra-se o da ferramenta não varrer automaticamente perfis na internet em busca de objetos copiados, sendo necessária indicação de um perfil supostamente falso na interface do modelo conceitual. Isso se deve às ferramentas de redes sociais que bloqueiam este tipo de atividades.

Igualmente, deve-se atentar, na implementação, para as diretivas legais e de privacidade vigentes no país em que se pretenda usar o modelo conceitual.

Como proposta futura, pode-se oferecer a possibilidade de avaliar também a falsidade com base em outros critérios a serem implementados em módulos e até mesmo a existência de desinformações nas redes sociais (coletando-se e organizando, por exemplo, comentários relativos às informações que sugiram uma classificação de notícia falsa).

Portanto, a proposta apresentada, valendo-se de diversas áreas da Ciência da Informação e tecnologias da *Web* Semântica, vem no escopo de oferecer um rotulador nas ferramentas de redes sociais, bem como um classificador de atividades suspeitas inerentes a *fakes* e *bots*, por meio de ontologias, oferecendo maior significado aos dados disponíveis nestes ambientes, auxiliando na rastreabilidade de objetos publicados, detecção de perfis falsos e cooperando para um ambiente digital mais rico e seguro.

5.3 Análise de persistência de metadados na ferramenta de rede social Facebook

Nos experimentos realizados nesta pesquisa, uma imagem já submetida no perfil Facebook.com/josemilagre (perfil) foi baixada para o computador e identificado seu *hyperlink*[28].

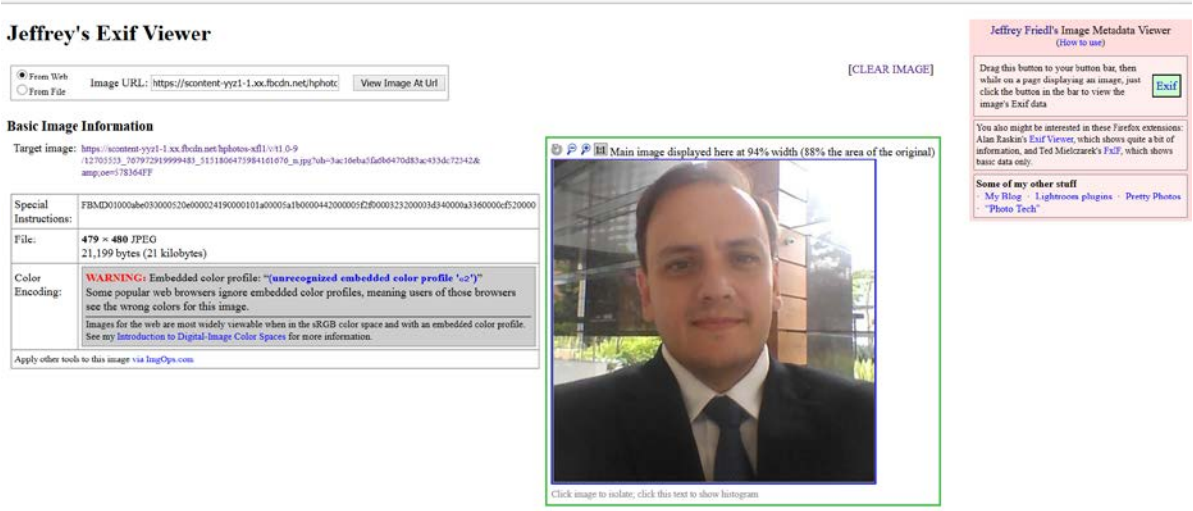
[28] https://scontent-yyz1-1.xx.fbcdn.net/hphotos-xfl1/v/t1.0-9/12705553_767972919999483_5151806475984161676_n.jpg?oh=3ac16eba5fa6b6470d83ac433dc72342&oe=578364FF

Figura 38 – Download de imagem de perfil na rede social Facebook

Fonte: Elaborado pelo autor utilizando a rede social Facebook

Posteriormente o Hiperlink foi submetido a um verificador de EXIFs online de nome Jeffrey's Exif Viewer[29] que extraiu os metadados da imagem ainda no servidor da ferramenta da rede social:

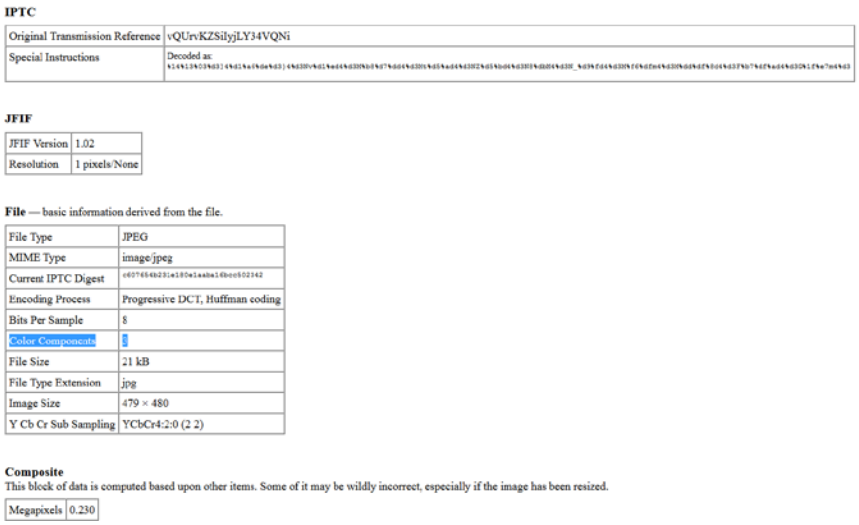
Figura 39 – Verificação dos metadados da imagem baixada da rede social



Fonte: Elaborado pelo autor utilizando o verificador Jeffrey's Exif Viewer

[29] http://regex.info/exif.cgi

Figura 40 – Metadados extraídos da imagem baixada da rede social



IPTC

Original Transmission Reference	vQUrvKZSiIyjLY34VQNi
Special Instructions	Decoded as:

JFIF

JFIF Version	1.02
Resolution	1 pixels/None

File — basic information derived from the file.

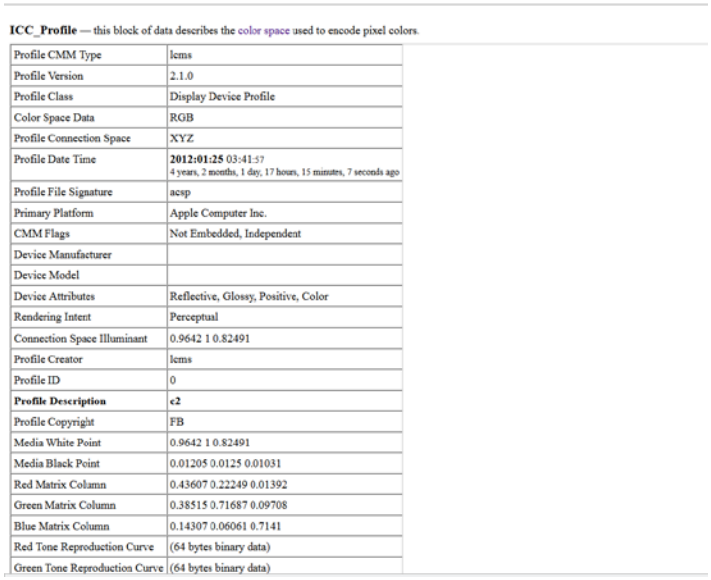
File Type	JPEG
MIME Type	image/jpeg
Current IPTC Digest	
Encoding Process	Progressive DCT, Huffman coding
Bits Per Sample	8
Color Components	3
File Size	21 kB
File Type Extension	jpg
Image Size	479 × 480
Y Cb Cr Sub Sampling	YCbCr4:2:0 (2 2)

Composite
This block of data is computed based upon other items. Some of it may be wildly incorrect, especially if the image has been resized.

Megapixels	0.230

Fonte: Elaborado pelo autor utilizando o verificador Jeffrey's Exif Viewer

Figura 41 – Metadado ICC Profile extraído da imagem baixada da rede social



ICC_Profile — this block of data describes the color space used to encode pixel colors.

Profile CMM Type	lcms
Profile Version	2.1.0
Profile Class	Display Device Profile
Color Space Data	RGB
Profile Connection Space	XYZ
Profile Date Time	**2012:01:25** 03:41:57 4 years, 2 months, 1 day, 17 hours, 15 minutes, 7 seconds ago
Profile File Signature	acsp
Primary Platform	Apple Computer Inc.
CMM Flags	Not Embedded, Independent
Device Manufacturer	
Device Model	
Device Attributes	Reflective, Glossy, Positive, Color
Rendering Intent	Perceptual
Connection Space Illuminant	0.9642 1 0.82491
Profile Creator	lcms
Profile ID	0
Profile Description	**c2**
Profile Copyright	FB
Media White Point	0.9642 1 0.82491
Media Black Point	0.01205 0.0125 0.01031
Red Matrix Column	0.43607 0.22249 0.01392
Green Matrix Column	0.38515 0.71687 0.09708
Blue Matrix Column	0.14307 0.06061 0.7141
Red Tone Reproduction Curve	(64 bytes binary data)
Green Tone Reproduction Curve	(64 bytes binary data)

Fonte: Elaborado pelo autor utilizando o verificador Jeffrey's Exif Viewer

Constatou-se que a rede social renomeia arquivos que são reinseridos, altera seu resumo criptográfico e não apresenta dado indexador de conteúdos idênticos postados novamente, um grande dificultador da detecção de uso indevido de objetos.

Neste contexto, salvou-se a imagem de perfil aqui estudada e inseriu-se diretamente um metadado de autor na imagem (contendo o nome e o telefone do proprietário da imagem). Na sequência inseriu-se a imagem em outra página, simulando-se o procedimento de um criminoso criando um *fake*:

Figura 42 – Inserção de uma imagem com metadados de autor na rede social Facebook

Fonte: Elaborado pelo autor utilizando a rede social Facebook

Coletou-se então a URL[30] do objeto de imagem cadastrada novamente e se submeteu ao analisador de EXIFs novamente. O resultado foi positivo e demonstrou que a ferramenta de rede social manteve o metadado, exibindo os dados do autor da imagem e seu telefone, como mostrado na figura 43.

[30] https://scontent-yyz1-1.xx.fbcdn.net/hphotos-xpf1/v/t1.0-9/10330405_762813110485558_5646164794334499417_n.jpg?oh=3a29785d36f79ea8ee4506075bba71c7&oe=577F1F1E

Figura 43 – Verificação da persistência de metadados após upload

Fonte: Elaborado pelo autor utilizando o verificador Jeffrey's Exif Viewer

Figura 44 – Metadado inserido permaneceu na imagem após upload à rede social

Fonte: Elaborado pelo autor utilizando o verificador Jeffrey's Exif Viewer

Assim, como se verificou, a ferramenta de rede social Facebook aceita a inserção de alguns metadados e não os remove quando o conteúdo ingressa na rede ou é reinserido. Estes critérios podem ser classificados com apoio de tecnologias como ontologias e descritos em formato RDF, possibilitando a aplicações a busca por conteúdos de determinado usuário ou usados para perfis falsos com base em metadados específicos. O desafio identificado é preservar estes metadados para que não seja zerado, o que pode ser contornado com mecanismos

descentralizados de descrição dos objetos. Outro desafio é a investigação do comportamento dos metadados em diversas e distintas ferramentas de redes sociais.

6 Considerações finais

A construção dessa pesquisa se deu diante da necessidade de se extrair significados de grandes volumes de informações e da necessidade de se criar um ambiente mais seguro na Internet, sobretudo, combatendo o problema da falta de crédito que as ferramentas de redes sociais enfrentam na atualidade. Não é de hoje que as redes são comentadas como fontes de *"hoaxes"* (boatos) e do crescente número de perfis falsos.

Apresentou-se, com o presente trabalho, como o acesso e tratamento das informações das redes sociais podem ser potencializados com o uso da Ciência da Informação. O escopo do trabalho foi avaliar o estado da técnica e o atual estágio das pesquisas (apresentando suas limitações) na detecção de perfis falsos nas redes sociais

Os objetivos de revisitar a literatura e melhores práticas para tratamento de dados e detecção de perfis falsos para amparo à inteligência cibernética e à computação forense foram contemplados. Igualmente, realizou-se uma análise exploratória de recursos e ferramentas disponíveis para análise de dados e detecção de perfis falsos onde se evidenciaram suas limitações e pontos de aproveitamento.

Constatou-se que embora haja uma infinidade de dados e informações circulando na Rede Mundial de Computadores, as ferramentas disponíveis não são capazes de atribuírem significado em tempo real, no processamento de tais informações, revelando conhecimento, o que contribui para a criação de *fakes* e envio de mensagens não solicitadas.

Para isso, apresentou-se uma proposta de modelo conceitual para detecção de perfis falsos, contornando as limitações destas ferramentas de redes sociais e com relevância para a sociedade, que poderá avaliar perfis, controlar alguns de seus recursos (objetos) que transitam pelas redes e também identificar prováveis golpes crimes e *bots* (com ontologias específicas).

O projeto se diferencia dos demais, pois se vale de questões ligadas à *web* semântica, ontologias, inteligência artificial, *folksonomia*s e outras áreas relacionadas à Ciência da Informação. Organizando e descrevendo dados relativos a perfis falsos facilitará o trabalho da Computação, na aplicação de probabilidades e algoritmos computacionais para análise de dados.

É notável, na inteligência cibernética e computação forense, a falta de projetos que consideram campos de aplicação e conceitos tratados no âmbito da Ciência da Informação, o que resulta de aplicações e *frameworks* engessados, que poderiam extrair muito mais significados de grandes volumes de dados despejados nos locais da *Web*, cooperando para organização e representação desta informação aparentemente oculta.

Identificou-se que a Ciência da Informação tem muito a contribuir na temática e no desenvolvimento de uma *web* mais “verdadeira” e com menos “falsidades” e de maior crédito, considerando que todo o desenvolvimento de técnicas e modelos sobre o tema da pesquisa em algum momento se relacionam com institutos pesquisados no âmbito da disciplina ou foram criados a partir de aportes teóricos de especialistas e trabalhos desenvolvidos na área.

Identificamos que a despeito das limitações envolvendo a utilização de dados públicos gerados em redes sociais, bem como limitações legais e de privacidade, técnicas podem ser utilizadas para o tratamento destas informações, como a rotulação e descrição independente, dentre outras.

Identificou-se que o uso de conceitos e ferramentas para recuperação e tratamento de informações debatidos na Ciência da Informação pode ser aliado no aprimoramento das técnicas de inteligência cibernética em grandes volumes de dados, especificamente, na detecção de *fakes* e robôs (bots) nas redes sociais, igualmente podendo associar conteúdos postados a potenciais atividades criminosas ou que sejam relacionadas a potenciais perfis falsos.

Conclui-se que o uso de recursos da Ciência da Informação na análise de grandes volumes de dados é um processo que elevará a rotulagem e significado dados disponíveis nas ferramentas de redes sociais, bem como alternativa a ser refletida pelos pesquisadores para preservar ou retardar a volatilidade dos dados na era do Big Data.

Dentro deste contexto, o presente projeto cria uma rotulagem paralela, descentralizada, mas conectada por meio das URIs de cada objeto do Big Data das ferramentas de redes sociais, o que contribui e avança nas pesquisas na busca por significados e extração de inteligência e conhecimento a partir destas redes, aprimorando a autenticidade e fornecendo recursos para que usuários possam pesquisar por perfis potencialmente falsos.

Com isso, um repositório paralelo de pesquisa, descrito de forma semântica, permitirá a usuários pesquisarem por perfis, antes de avalizarem ou compartilharem conteúdos, de modo a apurar se possuem credibilidade ou não, inclusive, sabendo se o objeto ou recurso já fora ou não pesquisado por alguém e por quanto tempo está na rede. Resta, em fase futura de implementação, especificar detalhes das ontologias, desenvolver a programação da proposta e aplicar em "*data-set*" extraído de ferramentas de redes sociais para experimentos práticos.

O presente modelo conceitual não é restrito à área de investigação, computação forense e inteligência cibernética, mas poderá ser derivado e adaptado, parametrizado com "*profiles*" para diversas áreas de negócio, mercado ou de conhecimento, e então detectar não só perfis, mas conteúdos contendo falsidades, estados, objetos protegidos por direitos autorais, boatos, dentre outros, gerando significados inerentes e importantes às precitadas áreas do conhecimento, com base nas informações das redes sociais. Logo, aplicável a outros tipos de ambientes, podendo-se resultar em ferramentas que irão organizar a informação existente em ferramentas de redes sociais, permitindo ágil consulta e inferências, revelando conhecimento. Como apresentado, espera-se, com o presente trabalho, minimamente, ter contribuído para os avanços na construção de modelos para representação e significado da informação em grandes volumes de dados, com a possibilidade de detecção de perfis falsos nas redes sociais, sobretudo, demonstrando a indispensável contribuição da Ciência da Informação na construção de modelos e propostas que se destinem a esta finalidade.

7 Trabalhos futuros

Será objeto de trabalhos futuros a avaliação de *webcrawlers* e APIs de ferramentas de redes sociais como Facebook e Twitter, para se identificar exatamente o que permitem e o que não permitem em termos de coleta e extração de dados, para que se possa dimensionar como vão interagir com aplicações construídas com base no modelo conceitual aqui proposto e que precisará ser adaptado.

Serão realizadas pesquisas de avaliação do comportamento dos metadados após o "*upload*" de recursos nas redes sociais, para se avaliar quais metadados são eliminados e quais são mantidos e neste caso, quais outros índices possíveis para identificar esta informação no Big Data, detectando falsidades e violações autorais. Este processo será feito com imagens e vídeos.

Serão realizadas simulações com a ferramenta *Firesheep* para identificar se é possível coletar dados de usuários de ferramentas de redes sociais que estejam no mesmo segmento de rede, possibilitando massa de dados para aplicações em casos de não amizade do operador do *framework* com o suspeito, avaliando-se as implicações legais deste contexto.

Continuar-se-á na pesquisa envolvendo as características dos ataques de clone de identidade (ICA)

Pretende-se ainda simular cálculos de similaridade entre perfis, com base nas pesquisas estudadas, o que poderá ser usado na construção da ferramenta de detecção de perfis falsos.

Avaliar-se-á a possibilidade de rotular objetos já inseridos nas redes sociais, coletando-os e mantendo-os em uma base de dados paralela, com especificação em ontologia própria. Será pesquisado ainda acerca da coleta e descrição independente dos Hyperlinks de imagens em redes socais, que podem preservar o conteúdo mesmo após o delete da postagem, cooperando para recuperação de ofensas, falsidades e usos temporários de perfis falsos nas redes sociais.

Verificar-se-á também como as aplicações construídas com base no modelo conceitual poderão ser disponibilizadas e acopladas a grandes fontes de dados (se agente intermediário ou embutida nas ferramentas de redes sociais). Realizar-se-á a simulação de funcionamento do modelo conceitual, ainda sem aplicativo ou

ferramenta, mas com base em coleta manual de dados, observando-se os resultados e gerando-se sugestões de q*ueries* para pesquisa.

Na questão legal, em conclusão, serão avaliadas as implicações legais e de privacidade decorrentes da implementação de projetos com base no modelo conceitual proposto no presente trabalho.

Referências

ADDIKARI, S.; DUTTA, K. Identifying *fake profiles* in Linkedin. In: Pacific Asia Conference on Information Systems(PACIS), 19., 2014. [S. l.]. **Proceedings...**, 2014. p. 1-16. Disponível em: <aisel.aisnet.org/cgi/viewcontent.cgi?article=1110&context=pacis2014> Acesso em: 20 jul. 2015.

AGGARWAL, A.; RAJADESINGAN, A.; KUMARAGURU, P. **PhishAri**: automatic realtime phishing detection on Twitter. 29 jan. 2013. Disponível em: <http://arxiv.org/abs/1301.6899>. Acesso em: 12 fev. 2016.

AHMADIZADEH, E. et al. An Automated Model to Detect *Fake Profiles* and botnets in Online Social Networks Using Steganog*raph*y Technique. IOSR **Journal of Computer Engineering**, v. 17, n. 1, p. 65-71, 2015. Disponível em: <http://www.iosrjournals.org/iosr-jce/papers/Vol17-issue1/Version-4/L017146571.pdf> Acesso em: 20 jul. 2015.

ALVES, R. C. V. **Metadados como elementos do processo de catalogação**. 2010. 132 f. Tese (Doutorado em Ciência da Informação)–Faculdade de Filosofia Ciências, Universidade Estadual Paulista, Marília, 2010.

AMBROSE, M. L. It's about time: Privacy, Information life cycles, and the right to be forgotten**. Stanford Technology Law Review**, v. 16, n. 2, 2013. Disponível em: < https://journals.law.stanford.edu/sites/default/files/stanford-technology-law-review/online/itsabouttime.pdf>. Acesso em: 20 jul. 2015.

ANDRADE, I. A. de. **As dimensões semântica e pragmática da Web e dos mecanismos de busca no ciberespaço**. Londrina, 2012. 121 f. Dissertação (Mestrado Profissional em Gestão da Informação)-Centro de Educação, Comunicação e Artes, Universidade Estadual de Londrina. 2012. Disponível em: <http://www.bibliotecadigital.uel.br/document/?code=vtls000181112> Acesso em: 05 maio 2016.

ATHANASOPOULOS, E. et al. **Antisocial networks**: turning a social network into a botnet. 2008. p. 146-160. Disponível em: <http://link.springer.com/chapter/10.1007%2F978-3-540-85886-7_10>. Acesso em: 20 jul. 2015.

ATHENIENSE, A. Ter um perfil falso na internet é crime? **O Direito e as Novas Tecnologias**, [2010?]. Disponível em: <http://www.dnt.adv.br/salas-do-conhecimento/ter-um-perfil-falso-na-internet-e-crime/#>. Acesso em: 10 mar. 2016.

BAEZA-YATES, R.; RIBEIRO-NETO, B. **Modern information retrieval.** New York: Addison-Wesley, 1999.

BEEBE, N. CLARK, J. Dealing with terabyte data sets in digital investigations. In: IFIP International Conference on Digital Forensics, National Center for Forensic Science. **Advances in digital forensics**. Florida: Springer, 2005. p. 3-16. Disponível

em: <http://link.springer.com/chapter/10.1007%2F0-387-31163-7_1>. Acesso em: 20 jul. 2015.

BERNERS-LEE, T. **Semantic *Web* Road map**. [S.l.: s.n.], 1998. Disponível em: <http://www.w3.org/DesignIssues/Semantic.html>. Acesso em: 20 maio. 2015.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. **The semantic *Web***: a new form of *Web* content that is meaningful to computers will unleash a revolution of new possibilities. Scientific American: London, 2001.

BERNSTEIN, M. et al. Interactive topic-based browsing of social status streams. In: 23rd ACM Symposium on User Interface Software and Technology (UIST), 23., 2010, october 3-6. New York, NY. NY: ACM. **Proceedings...**, 2010. Disponível em: <http://hci.stanford.edu/publications/2010/eddi/eddi-uist2010.pdf>. Acesso em: 25 jul. 2015.

BHUMIRATANA, B. A Model for Automating Persistent Identity Clone in Online Social Network. In: 10th International Conference on Trust, Security and Privacy in Computing and Communications (IEEE TrustCom-11), 10., 2011. Changsha, China. **Proceedings...**, 2011. p. 681-686. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6120880&tag=1>. Acesso em: 23 jun. 2015.

BILTON, N. Bots de redes sociais: amigos falsos, lucros reais. **Gazeta do Povo**, 29 dez. 2014. Disponível em: <http://www.gazetadopovo.com.br/mundo/new-york-times/bots-de-redes-sociais-amigos-falsos-lucros-reais-ehtknb5jv956w8fpyevo8h07i>. Acesso em: 5 mar. 2016.

BIRNBAUM, M. H.; STEGNER, S. E. Source credibility in social judgment: Bias, expertise, and the judge's point of view. **J. Personality and Social Psychology**, v. 37, n. 1, p. 48-74, 1979.

BLUM, O. et al. O perigo dos perfis falsos em redes sociais. **Opice Blum**, 2009. Disponível em: <http://www.opiceblum.com.br/o-perigo-dos-perfis-falsos-em-redes-sociais/>. Acesso em: 11 mar. 2016.

BORKO, H. Information Science: What is it? **American Documentation**, v.19, n.1, p. 3-5, Jan. 1968. Disponível em: <http://www.scribd.com/doc/533107/Borko-H-v-19-n-1-p-35-1968>. Acesso em: 25 mar. 2015.

BRASIL. **Código penal**. Decreto-lei n° 2.848, de 7 de dezembro de 1940. Rio de Janeiro, 7 de dezembro de 1940. Disponível em: <http://www.planalto.gov.br/ccivil_03/decreto-lei/Del2848compilado.htm>. Acesso em: 16 mar. 2016.

BRASIL. **Lei Nº 9.296, de 24 de julho de 1996**. Regulamenta o inciso XII, parte final, do art. 5° da Constituição Federal, Brasília, 24 de julho de 1996. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/L9296.htm>. Acesso em: 13 mar. 2016.

BRASIL. **Lei nº 12.737, de 30 de novembro de 2012**. Dispõe sobre a tipificação criminal de delitos informáticos; altera o Decreto-Lei no 2.848, de 7 de dezembro de 1940 - Código Penal; e dá outras providências, Brasília, 30 de novembro de 2012. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/l12737.htm>. Acesso em: 13 mar. 2016.

BRASIL. Senado Federal. **Projeto de Lei do Senado nº 330, de 2013**. Dispõe sobre a proteção, o tratamento e o uso dos dados pessoais, e dá outras providências, 2013. Disponível em: <http://www25.senado.leg.br/web/atividade/materias/-/materia/113947>. Acesso em: 13 mar. 2016.

BRASIL. **Marco Civil da Internet**. Lei 12.965, de 23 de abril de 2014. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm>. Acesso em: 11 jun. 2015.

BREITMAN, K. ***Web* Semântica**: a Internet do futuro. Rio de Janeiro: LTC, 2005.

BRETERNITZ, V. J.; SILVA, L. A. BIG DATA: bringing new opportunities and challenges. In: 10th International Conference on Information Systems and Technology Management - CONTECSI, 10., 2013, São Paulo. **Proceedings...**, 2013b. p. 2906-2914. Disponível em: <http://www.contecsi.fea.usp.br/envio/index.php/contecsi/10contecsi/paper/viewFile/3564/2086>. Acesso em: 20 jul. 2015.

BUCKLAND, M. K. Information as thing. **Journal of the American Society for Information Science**, v. 45, n. 5, p. 351-360, 1991.

CANALI, C.; COLAJANNI, M.; LANCELLOTTI, R. **Data acquisition in social networks: issues and proposals**. 2011. Disponível em: <https://weblab.ing.unimore.it/papers/sos11.pdf>. Acesso em: 10 mar. 2016.

CANINI, K. R.; SUH, B. ; PIROLLI, P. L. Finding credible information sources in social networks based on content and social structure. In: International Conference on Privacy, Security, Risk and Trust (PASSAT) and IEEE Third International Conference on Social Computing (Cicalo) IEEE, 3., 2011, october 9-11. Boston, MA**. Proceedings...**, 2011. p. 1-8. Disponível em <http://www.peterpirolli.com/ewExternalFiles/Canini-2011-SocialCom.pdf>. Acesso em: 22 jul. 2015.

CAPURRO, R.; HJØRLAND, B. O conceito de Informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 12, n. 1, p.148-207, jan./abr. 2007. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/54/47>. Acesso em: 15 abr. 2015.

CHASTAIN, S. **What is Metadata?** 16 dez. 2014. Disponível em: <http://graphicssoft.about.com/od/glossary/f/metadata.htm>. Acesso em: 10 mar. 2016.

CONTI, M.; POOVENDRAN, R.; SECCHIERO. *Fake*Book: detecting *fake profiles* in on-line. In: International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2012 IEEE/ACM, 2012, aug. 26-29. Istanbul. **Proceedings...**, 2012. p. 1071-1078. Disponível em: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6425616&tag=1>. Acesso em: 25 jul. 2015.

CONVERGÊNCIA DIGITAL. **Rede social é 'combustível' do Big Data**. 2012. Disponível em: <http://convergenciadigital.uol.com.br/cgi/cgilua.exe/sys/start.htm?infoid=32130&sid=97#.VhFlFflViko >. Acesso em: 16 jun. 2015.

CORVEY, W. et al. **Foundations of a multilayer annotation *framework* for *Twitter* communications during crisis events**. [S. l.: s. n.], [2011?]. Disponível em:<http://epic.cs.colorado.edu/wp-content/uploads/lrec_2012_final_120523.pdf>. Acesso em 25 jul. 2015.

DOSSIS, S. **Semantically-enabled digital investigations**: a method for semantic integration and correlation of digital evidence using a hypothesis-based approach. 2012. 118 f. Degree project at the master level (Computer and Systems Sciences)-Stockholm University, Suécia, 2012. Disponível em: <http://www.isaca.org/chapters4/Sweden/events/Documents/Dossis_ThesisDSV%20spyridon.pdf>. Acesso em: 10 jun. 2015.

EGELE, M. et al. COMPA: detecting compromised accounts on social networks. In: 20th Annual Network & Distributed System Security Symposium, 20., 2013. San Diego, CA United States. **Proceedings...**, 2013. p. 1-17. Disponível em: <http://www.internetsociety.org/doc/compa-detecting-compromised-accounts-social-networks>. Acesso em: 13 mar. 2016.

FARMER, D.; VENEMA, W. **Pericia Forense Computacional**: teoria e prática. São Paulo: Person, 2007.

FERNEDA, E. **Recuperação da informação**: análise sobre a contribuição da Ciência da Computação para a Ciência da Informação. 2003. 147 f. Tese (Doutorado em Ciências da Comunicação)–Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2003. Disponível em: <http://www.teses.usp.br/teses/disponiveis/27/27143/tde-15032004-130230/público/Tese.pdf>. Acesso em: 20 jul. 2015.

GIL, A. C. **Como elaborar projetos de pesquisa.** 4. ed. São Paulo: Atlas, 2002.

GOMES, D.; SILVA, M. J. **The Viuva Negra crawler**. Lisboa: [s.n.], 2006. Disponível em: <http://visibilidade.net/daniel/docs/papers/vnRT.pdf>. Acesso em: 20 jul. 2015.
GOVEIA, F.; MALINI, F.; CIARELLI, P. **Bots contra a Sociedade**. 30 set. 2014. Disponível em: <http://www.labic.net/blog/internet-2/bots-contra-a-sociedade/>. Acesso em: 25 fev. 2016.

GREENBERG, J. The Semantic Web: more than a vision. Bulletin for the **American Society for Information Science and Technology**, v. 29, n. 4, p.6-7, apr./may, 2003

GRUBER, T. R. Toward principles for the design of ontologies used for knowledge sharing. In: INTERNATIONAL WORKSHOP ON FORMAL ONTOLOGY, 1993, Padova, Italy. **Technical Report** *KSL 93-04*. Stanford: Knowledge Systems Laboratory, Stanford University, 1993. Disponível em: <http://tomgruber.org/writing/onto-design.pdf>. Acesso em: 05 ago. 2015.

GUERREIRO, D. Exif, o que é? **Fotografia-DG**, 1 maio 2009. Disponível em: <http://www.fotografia-dg.com/exif-o-que-e/>. Acesso em: 26 mar. 2016.

GUPTA, A. et al. Faking Sandy: characterizing and identifying *fake* images on *Twitter* during hurricane Sandy. In: 22nd International Conference on World Wide *Web*, 22., 2013, Rio de Janeiro, Brazil. **Proceedings...**, 2013. p. 729-736. Disponível em: <http://www2013.w3c.br/companion/p729.pdf> Acesso em 20 jul. 2015.

HINSHAW, F. D. Data Warehouse Appliances: driving the business intelligence revolution. **DM Review**, v. 14, n. 9, 2004. Disponível em: <http://connection.ebscohost.com/c/articles/14532711/data-warehouse-appliances-driving-business-intelligence-revolution>. Acesso em: 25 jul. 2015.

HUBER, M. et al. Social snapshots: digital forensics for online social networks. In: 27th Annual Computer Security Applications Conference, 27., 2011. Orlando, Florida, USA. **Proceedings…**, 2011. Disponível em: <http://publik.tuwien.ac.at/files/PubDat_202726.pdf>. Acesso em: 20 jul. 2015.

JACOB, E. K. Ontologies and the semantic *web***. Bulletin for the American Society for Information Science and Technology**, v. 29, n. 4, p. 19-22, abr./maio 2003.

JESUS, D. de; MILAGRE, J. A. **Marco Civil da Internet**: comentários à Lei n. 12.965/14. São Paulo: Saraiva, 2014.

JIN, L.; TAKABI, H.; JOSHI, J. B. D. Towards active detection of identity clone attacks on online social networks. In: ACM Conference on Data and Application Security and Privacy (CODASPY), 1., 2011, february 21–23. San Antonio, Texas, USA. **Proceedings...**, 2011. p. 27-38. Disponível em: <http://dl.acm.org/citation.cfm?id=1943520>. Acesso em: 15 jun. 2015.

KAHLE, B. Preserving the Internet. **Scientific American**, 1998. Disponível em: <http://*web*.archive.org/*web*/19980627072808/http:/www.sciam.com/0397issue/0397kahle.html>. Acesso em: 13 jul. 2015.

KROMBHOLZ, K.; MERKL, D.; WEIPPL, E. *Fake* identities in social media: a case study on the sustainability of the Facebook business model. **Journal of Service Science Research**, v. 2012, n. 4, p.175-212. Disponível em: <https://www.sba-research.org/wp-content/uploads/publications/krombholzetal2012.pdf> Acesso em 25 jul. 2015.

LANEY, D. **3D Data Management**: controlling data volume, velocity, and variety. 2001. Disponível em: <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>. Acesso em: 18 jun. 2015.

LEUNG, A. et al. **Implementation of a Focused Social Networking Crawler.** [S. l.: s. n.], 2009. Disponível em: <http://courses.ece.ubc.ca/412/term_project/reports/2009/focused_social_net_crawler.pdf>. Acesso em: 20 jul. 2015.

LOHR, S. The age of Big Data. **The New York Times**, 2012. Disponível em: <http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html >. Acesso em: 02 jan. 2015.

MAYER-SCHÖNBERGER, V. **Delete**: the virtue of forgetting in the digital age. [S. l.: s. n.], 2011.

MCAFEE, A; BRYNJOLFSSON, E. Big Data: The Management Revolution. **Harvard Business Review**, v. 90, n. 10, p. 60-68, 2012.

MENDOZA, M.; POBLETE, B.; CASTILLO, C. *Twitter* under crisis: can we trust what we rt? In: Workshop on Social Media Analytics, SOMA, 10., 2010, New York, NY, USA. **Proceedings…**, New York: ACM, 2010. p. 1-9. Disponível em: <http://snap.stanford.edu/soma2010/papers/soma2010_11.pdf>. Acesso em: 13 jul. 2015.

MICROSOFT PHOTODNA. **Frequently Asked Questions**. 2009. Disponível em: <https://www.microsoft.com/en-us/PhotoDNA/FAQ>. Acesso em: 14 mar. 2016.

MÍDIAS SOCIAIS. **Bots**: os robôs das redes sociais. 21 maio 2014. Disponível em: <http://www.magicwebdesign.com.br/blog/redes-sociais/bots-os-robos-das-redes-sociais/>. Acesso em: 4 mar. 2016.

MONTANARO, D. **Inteligência Cibernética**. 30 dez. 2014. Disponível em: <http://corporate.canaltech.com.br/noticia/seguranca/Inteligencia-Cibernetica/>. Acesso em: 12 mar. 2016.

MOOERS, C. Zatocoding applied to mechanical organization of knowledge. **American Documentation**, Washington, v. 2, n. 1, p.20-32, 1951.

MORAES, E. A. M.; AMBRÓSIO, A. P. L. **Ontologias**: conceitos, usos, tipos, metodologias, ferramentas e linguagens. Goiás: Universidade Federal de Goiás, 2007. Disponível em: <http://www.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_001-07.pdf>. Acesso em: 26 fev. 2016.

MOREIRA, A.; ALVARENGA, L.; OLIVEIRA, A. de P. O nível do conhecimento e os instrumentos de representação: tesauros e ontologias. **DataGramaZero**, v. 5, n. 6, 2004. Disponível em: <http://www.brapci.ufpr.br/documento.php?dd0=0000007546&dd1=c1a8e>. Acesso em: 10 fev. 2016.

MORESI, E. (Org.). **Manual de metodologia da pesquisa.** Brasília: Universidade Católica de Brasília, 2003. Disponível em: <http://www.inf.ufes.br/~falbo/files/MetodologiaPesquisa- Moresi2003.pdf>. Acesso em: 23 jun. 2015.

MÜLLER, N. ***Framework*, o que é e para que serve?** [S. l.: s. n.], 2008. Disponível em:<https://www.oficinadanet.com.br/artigo/1294/*framework*_o_que_e_e_para_que_serve>. Acesso em: 19 jun. 2015.

NEVES, C. Big Data: uma fonte de poder? **Administradores.com**, 29 jul. 2014. Disponível em: <http://www.administradores.com.br/mobile/artigos/marketing/big-data-uma-fonte-de-poder/79336/>. Acesso em: 26 fev. 2016.

NICOLINO, M. E. V. P. **Diretrizes para a utilização de ontologias na indexação automática**. 2014. 101 f. Dissertação (mestrado)-Faculdade de Filosofia e Ciências, Universidade Estadual Paulista Júlio de Mesquita Filho, Marília, 2014. Disponível em: <http://repositorio.unesp.br/handle/11449/121979>. Acesso em: 15 fev. 2016.

PEREIRA, A. P. **O que é hash?** 4 mar. 2009. Disponível em: <http://www.tecmundo.com.br/o-que-e/1663-o-que-e-hash-.htm>. Acesso em: 19 fev. 2016.

RAMALHO, R. A. S. ***Web* Semântica**: aspectos interdisciplinares da gestão de recursos informacionais no âmbito da Ciência da Informação. 2006. 120 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2006. Disponível em:< http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/ramalho_ras_me_mar.pdf>. Acesso em: 25 jul. 2015.

RAVENSCRAFT, E. **How to use a *fake* name on Facebook without getting flagged**. [S.l.: s.n.], 2014. Disponível em: <http://lifehacker.com/how-to-use-a-*fake*-name-on-Facebook-without-getting-flag-1637644101>. Acesso em: 10 jul. 2015.

REDDY, N. R.; KUMAR, N. **Automatic detection of *fake profiles* in online social networks**. 2012. 31 f. Thesis (Bachelor of Technology Degree in Computer Science and Engineering)-National Institute of Technology Rourkela, Rourkela-769 008, Orissa, India, 2012. Disponível em:<http://ethesis.nitrkl.ac.in/3578/1/thesis.pdf> Acesso em 20 jul. 2015.

RIZI, F. S.; KHAYYAMBASHI, M. R.; KHARAJI, M. Y. A new approach for findings cloned *profiles* in online social networks. **Int. J. of Network Security**, v. 6, p. 25-37, 2014.Disponível em: <http://arxiv.org/ftp/arxiv/papers/1406/1406.7377.pdf>. Acesso em: 25 jul. 2015.

ROSS, D. **How to Leverage an API for Conferencing**. [S. l.: s. n.], 2007. Disponível em: <http://money.howstuffworks.com/business-communications/how-to-leverage-an-api-for-conferencing1.htm>. Acesso em: 20 jul. 2015.

SANTAREM SEGUNDO, J. E. **Recursos tecno-metodológicos para descrição e recuperação de informações na *Web***. Marília, 2004. 157 f. Dissertação (Mestrado em Ciência da Informação)-Faculdade de Filosofia e Ciências, Universidade Estadual Paulista. 2004. Disponível em: <http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/ santaremsegundo_je_me_mar.pdf>. Acesso em: 05 maio 2015.

SANTAREM SEGUNDO, J. E. **Representação Iterativa**: um modelo para Repositórios Digitais. Marília, 2010. 244 f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista. 2010. Disponível em: <http://repositorio.unesp.br/bitstream/handle/11449/103346/santaremsegundo_je_dr_mar.pdf?sequence=1&isAllowed=y >. Acesso em: 5 maio 2015.

SANTAREM SEGUNDO, J. E. ; CONEGLIAN, C. S. Tecnologias da web semântica aplicadas à organização do conhecimento: padrão SKOS para construção e uso de vocabulários controlados descentralizados. In: GUIMARÃES, J. A. C.; DODEBEI, V. (Org.). **Organização do conhecimento e diversidade cultural**. Marília: ISKO-Brasil; FUNDEPE, 2015. p. 224-233. Disponível em: <http://isko-brasil.org.br/wp-content/uploads/2015/09/Organiza%C3%A7%C3%A3o-do-Conhecimento-e-Diversidade-Cultural-ISKO-BRASIL-2015.pdf>. Acesso em: 26 fev. 2016.

SILVA, E. L.; MENEZES, E. M. **Metodologia da pesquisa e elaboração de dissertação.** 4 ed. Florianópolis: UFSC, 2005. Disponível em: <https://projetos.inf.ufsc.br/arquivos/Metodologia_de_pesquisa_e_elaboracao_de_teses_e_dissertacoes_4ed.pdf>. Acesso em: 18 jul. 2015.

SOLOVE, D. J. **The future of reputation**: grossip, rumor, and privacy on the internet. New Haven: Yale University Press, 2007. Disponível em: <http://docs.law.gwu.edu/fac*web*/dsolove/Future-of-Reputation/text/futureofreputation-ch1.pdf>. Acesso em: 18 jul. 2015.

SOLVE MIDIA. **Solve Media Uncovers $1.5 Billion in Wasted Ad Spend**. 28 set. 2012. Disponível em: <http://news.solvemedia.com/post/32457007008/solve-media-uncovers-15-billion-in-wasted-ad>. Acesso em: 05 maio 2016.

USCHOLD, M. Knowledge level modelling: concepts and terminology. **The Knowledge Engineering Review**, v. 13, n. 1, p. 5-29, 1998. Disponível em: <http://journals.cambridge.org/download.php?file=%2FKER%2FKER13_01%2FS0269888998001040a.pdf&code=2441dfe1531c7262cc1b3917051f9ad5>. Acesso em: 16 jul. 2015.

VECHIATO, F. L. **Encontrabilidade da informação**: contributo para uma conceituação no campo da Ciência da Informação. 2013. 206 f. Tese (Doutorado em Ciência da Informação)–Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2013. Disponível em: <http://repositorio.unesp.br/bitstream/handle/11449/103365/000735214.pdf?sequence=1&isAllowed=y>. Acesso em: 18 jun. 2015.

WENDT, E. **Inteligência Cibernética**: introdução ao assunto. Porto Alegre: Inteligência Policial, 2010. Disponível em: <http://www.inteligenciapolicial.com.br/2010/03/inteligencia-cibernetica-introducao-ao.html>. Acesso em: 18 jul. 2015.

WENG, J. et al. TwitterRank: finding topic-sensitive influential Twitterers. In: ACM international conference on *Web* search and data mining, 3., 2010, february 04-06. New York, New York, USA. **Proceedings...**, 2010. Disponível em: <http://ink.library.smu.edu.sg/cgi/viewcontent.cgi?article=1503&context=sis_research>. Acesso em: 13 jul. 2015.

ZAINUDIN, N. M.; MERABTI, M.; LLEWELLYN-JONES, D. A Digital forensic investigation model for online social networking. In: 11th Annual Conference on the Convergence of Telecommunications, Networking & Broadcasting (PGNet), 11., 2010, june. Liverpool, UK. **Proceedings...**, 2010. p. 21-22.

ZIKOPOULOS, P. et al. **Harness the power of Big Data**: the IBM Big Data platform. New York: McGraw-Hill, 2012. Disponível em: <http://www.dotgroup.co.uk/wp-content/uploads/2014/11/Harness-the-Power-of-Big-Data-The-IBM-Big-Data-Platform.pdf>. Acesso em: 15 jun. 2015.