

Revista Eletrônica de Sistemas de Informação

ISSN 1677-3071

V. 14, n. 2

mai-ago 2015 - Edição Temática sobre Análise de Redes Sociais e Mineração

doi:10.21529/RESI.2015.1402

Sumário

Editorial

EDITORIAL

Jonice Oliveira

BrASNAM

EXTRAÇÃO, CARACTERIZAÇÃO E ANÁLISES DE DADOS DE CURRÍCULOS LATTES

Luciano Antonio Digiampietri, Jesús Pascual Mena-Chalco, José Jesús Pérez-Alcázar, Esteban Fernandez Tuesta, Karina Valdivia Delgado, Rogério Mugnaini, Gabriela Scardine Silva, Jamison José da Silva Lima

ANÁLISE DE SENTIMENTO DE TWEETS COM FOCO EM NOTÍCIAS

Paula Nascimento, Bruno Osiek, Geraldo Xexéo

MEDINDO SENTIMENTOS NO TWITTER POR MEIO DE UMA ESCALA PSICOMÉTRICA

Pollyanna Gonçalves, Wellington José das Dores, Fabricio Benevenuto

BOTS SOCIAIS: COMO ROBÔS PODEM SE TORNAR INFLUENTES NO TWITTER

Johnnatan Messias, Lucas Schmidt, Ricardo Oliveira, Fabricio Benevenuto



Este trabalho está licenciado sob uma [Licença Creative Commons Attribution 3.0](http://creativecommons.org/licenses/by/3.0/).

ISSN: 1677-3071

Esta revista é (e sempre foi) eletrônica para ajudar a proteger o meio ambiente, mas, caso deseje imprimir esse artigo, saiba que ele foi editorado com uma fonte mais ecológica, a *Eco Sans*, que gasta menos tinta.

This journal is (and has always been) electronic in order to be more environmentally friendly. Now, it is desktop edited in a single column to be easier to read on the screen. However, if you wish to print this paper, be aware that it uses Eco Sans, a printing font that reduces the amount of required ink.

MEDINDO SENTIMENTOS NO TWITTER POR MEIO DE UMA ESCALA PSICOMÉTRICA

MEASURING SENTIMENTS ON TWITTER BY MEANS OF A PSYCHOMETRIC SCALE

(artigo submetido em março de 2013)

Pollyanna Gonçalves

Mestre em Ciência da Computação pela
Universidade Federal de Minas Gerais
(UFMG)

pollyannaogoncalves@gmail.com

Wellington José das Dores

Mestrando em Ciência da Computação pela
Universidade Federal de Minas Gerais
(UFMG)

wellingtonjdores@gmail.com

Fabício Benevenuto

Professor do Departamento de Ciência da Computação
da Universidade Federal de Minas Gerais (UFMG)

fabricio@dcc.ufmg.br

ABSTRACT

Twitter has become an important social communication mean where users post messages about everything. Certain messages express information about their author's emotional state, which may be useful in developing applications that predict emotional trends of a population or simply to better understand the effects of global and local phenomena in the mood of people. In this study, we adapted a psychometric scale known as the PANAS-X, commonly applied as a questionnaire, to measure the feelings of Twitter users in many social, political and sporting events. Our results suggest that the PANAS-t, our adapted version of the PANAS-X, correctly captures feelings for the events analyzed.

Key-words: Twitter; emoticons; sentiment analysis; psychometric scale; public mood.

RESUMO

Twitter tem se tornado um meio importante de comunicação social, por meio do qual usuários postam mensagens sobre tudo. Algumas dessas mensagens expressam informações sobre o estado emocional do usuário, o que pode ser útil no desenvolvimento de aplicações que preveem tendências emotivas de uma população ou simplesmente para melhor entender os efeitos de fenômenos mundiais ou locais no humor das pessoas. Nesse trabalho, adaptamos uma escala psicométrica conhecida como PANAS-x, comumente aplicada em forma de questionário, para medir os sentimentos dos usuários do Twitter sobre uma série de eventos de tema social, político e esportivo. Nossos resultados sugerem que o PANAS-t, nossa versão adaptada do PANAS-x, captura corretamente sentimentos para os eventos analisados.

Palavras-chave: Twitter; emoticons; análise de sentimentos; escala psicométrica; humor do público.

1 INTRODUÇÃO

A rede social Twitter vem se tornando popular como uma plataforma que permite aos seus usuários a postagem de mensagens curtas sobre fatos do seu cotidiano, compartilhando informações em tempo real (MEEYOUNG *et al.*, 2012, COMARELA *et al.* 2012) sobre diversos eventos de escala global como tragédias e revoluções (MACASKILL, 2009; SAKAKI *et al.*, 2010; GOMIDE *et al.*, 2011), na previsão de ataques terroristas (TUMASJAN *et al.*, 2010; DIAKOPOULOS *et al.*, 2010, O'CONNOR *et al.*, 2010) e na compreensão de eleições políticas (CHEONG e LEE, 2011). Estimativas recentes sugerem que 200 milhões dos usuários ativos no Twitter postam cerca de 150 milhões de tweets diariamente (KIRKPATRICK, 2011; RAO, 2010), ou seja, se o Twitter fosse um país apenas esses usuários o classificariam como o quinto maior do mundo.

Como essas mensagens podem conter o humor de cada usuário, consideramos como hipótese que essas também são capazes de revelar o sentimento do autor em relação a um determinado evento. Dessa forma, detectar sentimentos das pessoas sobre eventos utilizando uma base de dados extraída do Twitter nos permitiria compreender como, quando e porque os sentimentos das pessoas oscilam em relação a algum evento. Além disso, essa medição permitiria uma melhor compreensão de características intrínsecas da natureza humana. Apesar de todas essas potenciais aplicações, existe um número limitado de esforços nessa direção, dos quais a grande maioria está focada na extração de opiniões e não de sentimentos (PANG e LEE, 2008), ou não foi devidamente projetada para esse contexto (DODDS e DANFORTH, 2010).

Nesse artigo propomos o PANAS-t, uma escala psicométrica de medição de sentimentos adaptada para o contexto do Twitter que se baseia na versão expandida da conhecida escala PANAS-x (*Positive Affect Negative Affect Scale*). O PANAS-t consiste em sentimentos positivos e negativos, sendo adequado para a medição de qualquer evento aleatório no Twitter. Para isso foram utilizados dados empíricos, a partir de um conjunto único de dados coletados contendo cerca de 1,8 bilhões de *tweets*. Utilizamos esses dados para computar valores de normalização para cada sentimento, de modo que o crescimento ou decréscimo de cada um pode ser medido na base de dados. Para validar nossa metodologia forneceremos uma análise sobre variações no humor dos usuários do Twitter em relação a eventos que abrangem tragédias, política, saúde, bem como eventos esportivos.

A nossa abordagem faz com que o PANAS-t seja bastante simples e prático de ser utilizado para grandes quantidades de dados e mesmo para análise em tempo real. Nossa avaliação quantitativa oferece evidências de que o PANAS-t captura corretamente os sentimentos esperados para os eventos analisados.

A seguir, fornecemos um breve histórico sobre metodologias existentes para medir sentimentos. E então detalhamos como o PANAS-t fun-

ciona. Além disso, fornecemos evidências experimentais que comprovam nossa abordagem. Finalmente, concluímos o artigo e oferecemos direções para trabalhos futuros.

2 ABORDAGENS PARA MEDIÇÃO DE SENTIMENTOS

Com o crescimento das redes sociais na Web, a análise de sentimentos e mineração de opiniões se tornaram assunto de estudos para muitas pesquisas. Uma visão geral das abordagens e técnicas de processamento de linguagem natural utilizadas por pesquisadores na extração de sentimento foi apresentada no *survey* (PANG e LEE, 2008). Algumas abordagens consistem no uso de técnicas para a aprendizagem de máquina. Embora aplicada em vários cenários, técnicas supervisionadas requerem uma intervenção manual para a pré-classificação dos dados de treinamento, o que pode se tornar inviável em alguns casos. A seguir, faremos um breve levantamento das diferentes técnicas já utilizadas para a medição de sentimentos e como elas foram aplicadas.

Um estudo na área (DODDS e DANFORTH, 2010) se define pela detecção de índices de felicidade de um texto. Os autores daquele estudo mostraram que o índice de felicidade em letras musicais possui uma tendência decrescente desde os anos de 1960 a meados de 1990. Porém, o contrário foi observado em *blogs*, onde o índice de felicidade aumentou entre 2005 e 2009. A desvantagem desse tipo de abordagem está na utilização de apenas uma escala para a medição do índice de felicidade, não permitindo nenhuma categorização para o sentimento, o que é o foco do nosso trabalho.

Em sua pesquisa, Balog *et al.* (2006) coletaram dados da rede social *LiveJournal*, que disponibiliza aos seus usuários um sistema de criação de *blogs* ou diário *online* para o compartilhamento de informações. O interessante desse sistema é que seus usuários também podem definir seu humor atual. Como resultado das análises, os autores concluíram que técnicas simples, baseadas em comparação de frequências, são eficazes na identificação de variações de humor de uma população em relação a eventos globais.

Três estudos estão mais estreitamente relacionados aos nossos objetivos. Kim e Gilbert (2009) propõem um método para detecção de sentimentos utilizando o *The Affective Norms for English Words* (ANEW), um conjunto de dados que contém uma classificação emocional normativa para 1034 palavras em inglês. Cada palavra da sua base de dados contendo uma amostra de *tweets* sobre Michael Jackson foi associada com uma escala de três dimensões emocionais: prazer, excitação e dominância. No entanto, como se pode perceber, essa escala é capaz de capturar apenas sentimentos positivos que um usuário possa expressar.

Já o segundo estudo fez uso da escala psicométrica *Profile of Mood States* (POMS) (BOLLEN *et al.*, 2009). O POMS consiste em um conjunto adjetivos que qualificam seis sentimentos específicos: tensão, depressão,

raiva, fadiga, vigor e confusão. Os autores aplicaram a escala para identificar sentimentos em uma amostra de *tweets* e avaliaram o humor dos usuários em relação a flutuações do mercado de ações e em eventos como as eleições nos Estados Unidos. No entanto, essa escala também deixa de fora sentimentos importantes que estão ligados à felicidade de um usuário, pois ela é capaz de detectar apenas sentimentos negativos.

Em sua pesquisa, Nascimento *et al.* (2013) utilizaram estratégias de aprendizado de máquina, com classificadores de linguagem natural para analisar reações positivas ou negativas dos usuários da rede social Twitter em relação a diversas notícias. Como resultado, os autores concluíram que o método proposto para classificação obteve acurácia satisfatória levando-se em conta um processo de avaliação manual associado à subjetividade de um texto.

Propomos aqui um método de detecção de sentimentos utilizando uma escala psicométrica que detecta até onze escalas de sentimentos, capturando não apenas sentimentos positivos como também sentimentos negativos. O trabalho aqui desenvolvido consiste de uma extensão de Gonçalves *et al.* (2012) e Gonçalves *et al.* (2013a). Adicionalmente às análises já efetuadas, mostramos que nosso método, o PANAS-t, funciona como esperado em uma base de dados que não foi utilizada para a construção do método. Ou seja, o PANAS-t pode ser utilizado em outras amostras de dados extraídas do Twitter associadas aos eventos desejados.

3 MEDINDO SENTIMENTOS COM O PANAS-T

Nossa abordagem de detecção e medição de sentimentos no Twitter consiste na adaptação de uma escala psicométrica bem conhecida, a *Positive Affect Negative Affect Scale* (PANAS). Começaremos descrevendo o PANAS-x, uma versão expandida do PANAS, adaptada para o contexto da Web.

3.1 A ESCALA PANAS-X

A escala PANAS foi desenvolvida por Watson e Clark (1985). Indivíduos recebem um questionário contendo adjetivos que descrevem cada sentimento. Cada palavra desse questionário é então classificada ao seu sentimento correspondente, em uma escala de 1 a 5, de acordo com o que o entrevistado está sentindo. Mais especificamente, o PANAS-x, versão expandida da escala original, mede onze afetos específicos: *medo, tristeza, culpa, hostilidade, timidez, fadiga, surpresa, jovialidade, autoconfiança, serenidade e atenção*.

A Tabela 1 sumariza os itens de composição da escala PANAS-x, que são utilizados no nosso trabalho para a análise de *tweets*. Antes de explicar como adaptamos o PANAS-x para o contexto do Twitter, precisamos descrever a nossa base de dados.

Tabela 1. Termos que compõem a escala PANAS-x

Estado afetivo	Palavras relacionadas (em inglês)
Medo	afraid, scared, frightened, nervous, jittery, shaky.
Hostilidade	angry, hostile, irritable, scornful, disgusted, loathing.
Culpa	guilty, ashamed, blameworthy, angry at self, disgusted with self, dissatisfied with self.
Tristeza	sad, blue, downhearted, alone, lonely.
Jovialidade	happy, joyful, delighted, cheerful, excited, enthusiastic, lively, energetic.
Autoconfiança	proud, strong, confident, bold, daring, fearless.
Atenção	alert, attentive, concentrating, determined.
Timidez	shy, bashful, sheepish, timid.
Fadiga	sleepy, tired, sluggish, drowsy.
Serenidade	calm, relaxed, at ease.
Surpresa	amazed, surprised, astonished.

Fonte: baseado em Watson e Clark (1985)

3.2 BASE DE DADOS DO TWITTER

Os dados utilizados neste trabalho incluem um extenso conjunto de dados proveniente de um estudo de medição anterior que incluía um histórico completo de *tweets* postados por todos os usuários desde sua criação, em 2006, até agosto de 2009 (CHA *et al.*, 2010). Nosso conjunto de dados contém 54.981.152 usuários com 1,9 bilhões de *links* entre eles e 1.755.925.520 de *tweets* postados. Cerca de 8% das contas desses usuários estavam definidas como privadas, o que implica que apenas seus amigos podiam visualizar suas ligações e *tweets*. Na nossa análise ignoramos esses usuários.

Esta base de dados é apropriada para fins do presente trabalho pelas seguintes razões. Primeiro, o conjunto de dados engloba todos os usuários cadastrados desde o surgimento da rede até agosto de 2009. Assim, ela não é baseada em técnicas de amostragem que podem introduzir tendências para algumas características dos usuários. Em segundo lugar, este conjunto de dados contém todos os *tweets* desses usuários, o que a torna essencial para a medição do aumento ou diminuição de certo sentimento relacionado a *tweets* de um evento específico.

Devido à maior porcentagem de usuários cadastrados até 2009 no Twitter estar localizada nos Estados Unidos (KULSHRESTHA *et al.*, 2012) focaremos nossas análises na medição de sentimentos apenas em *tweets* que estão na língua inglesa.

3.3 LIMPEZA E *PARSING* DOS DADOS

A fim de analisar apenas *tweets* que possivelmente expressam sentimentos individuais de um usuário, consideramos apenas *tweets* contendo as expressões em inglês "*I'm*", "*I am*", "*I*", "*am*", "*feeling*", "*me*", e "*myself*". No total, agrupamos 479.356.536 *tweets* que correspondiam a algum desses padrões, o que representa cerca de 27% da nossa base total de *tweets*.

Após a filtragem dos dados, uma limpeza e *parsing* dos dados foram feitos, seguindo as seguintes etapas:

- (i) Primeiramente aplicamos abordagens comuns de processamento de linguagem como radicalização, remoção de *stopwords*, vírgulas, traços e outros caracteres não alfanuméricos, *URLs* e formas verbais comuns.
- (ii) Em seguida, separamos termos individuais utilizando espaços em branco como delimitadores. Por exemplo, para o *tweet* "*I am so scared about swine flu*" teríamos o conjunto de termos normalizados para: [*i*, *am*, *scared*, *swine*, *flu*]. No restante desse artigo utilizamos essa base de dados filtrada contendo apenas *tweets* normalizados.

3.4 CALCULANDO O PANAS-T

Tweets que expressam certos sentimentos podem se tornar mais populares que outros. Assim, a fim de dizer se um sentimento cresceu ou decresceu em relação a um evento específico (ex. morte do Michael Jackson), precisaremos de uma base para comparação. Por meio da quantificação da fração de *tweets* que estão relacionados a cada sentimento, seremos capazes de dizer se um sentimento aumentou para um evento específico em relação à presença desse sentimento na base de dados inteira. Em seguida, apresentamos detalhadamente como calculamos essa base de comparação.

Cada *tweet* do conjunto de *tweets* normalizados foi mapeado a um único sentimento. Para isso, quando um adjetivo da Tabela 1 associado a um sentimento *s* aparece como um termo de um *tweet* normalizado nós consideraremos esse *tweet* como possuindo o sentimento *s*. Em casos de empate associaremos o *tweet* como possuindo o primeiro sentimento identificado. No entanto, casos de empate foram muito raros e, portanto, foram desprezíveis na quantificação. Com base nessa informação, definimos nossos níveis de base para cada sentimento, como segue.

Seja *T* o conjunto de *tweets* normalizados e *T_s* o subconjunto destes relacionados a um sentimento *s*. Assim, o valor base para cada sentimento, α , é definido como $\alpha = |T_s| / |T|$. Em outras palavras, nós dividimos a ocorrência de *tweets* de cada sentimento pelo número total de *tweets* normalizados na nossa base de dados. A Tabela 2 mostra os valores de base da ocorrência de cada sentimento considerado. Podemos notar que alguns sentimentos ocorrem em maior ordem de grandeza que

outros. Por exemplo, *tweets* que expressam sentimento de *fadiga* ocorrem cerca de 32 vezes mais do que os que expressam *timidez*, o que reforça que uma normalização é necessária.

Tabela 2. Fração de tweets para cada sentimento na base de dados total

Sentimento (s)	Baseline (α)
Medo	0,0063791
Hostilidade	0,0086279
Culpa	0,0021756
Tristeza	0,0018225
Jovialidade	0,0007608
Autoconfiança	0,0240757
Atenção	0,0084612
Timidez	0,0182421
Fadiga	0,0036012
Serenidade	0,0008997
Surpresa	0,0022914

Fonte: dados da pesquisa

Com base nesses valores, podemos calcular o aumento ou diminuição relativa de um sentimento em uma amostra particular de *tweets* como segue. Seja S um conjunto de *tweets* (por exemplo, uma amostra de *tweets* relacionados à morte do Michael Jackson) e S_s o subconjunto desses *tweets* relacionados ao sentimento s . Definimos β como $\beta = |S_s| / |S|$. E então, definimos a pontuação do PANAS-t como um vetor de dimensão igual a 11, onde a função de pontuação $P(s)$ é calculado na Equação 1.

$$\text{Se } \beta s \leq \alpha s, \text{ então } P(s) = -1(\alpha s - \beta s) / \alpha s \quad (1)$$

$$\text{Se } \beta s > \alpha s, \text{ então } P(s) = (\beta s - \alpha s) / \beta s$$

Baseados nessa estratégia, cada valor de $P(s)$ poderá variar de -1 a 1 para cada sentimento. Ou seja, um evento com $P(\text{medo}) = 0$ significará que o evento não resultou em um aumento ou diminuição do sentimento \$medo\$, em comparação ao conjunto de dados total de *tweets* publicados até 2009. Em oposição, um valor de P positivo igual a 0,3 significaria um aumento de 30% e assim por diante.

Nossa estratégia para calcular a pontuação do PANAS-t é simples e adequada para permitir a comparação do aumento ou diminuição de cada sentimento de uma base de dados não tendenciosa. Mais importante, a Tabela 2 fornece valores de base para comparação de qualquer tipo de amostra de *tweets*. Por exemplo, pode-se facilmente rastrear uma amostra de mensagens postadas por usuários a partir da API do Twitter e utilizar esses valores para normalizar e obter pontuações de sentimentos.

4 AVALIAÇÃO DO PANAS-T

Com o objetivo de avaliarmos se o PANAS-t mede com precisão sentimentos dos usuários do Twitter, precisaríamos de uma verdade absoluta para comparar os resultados com nossos métodos. No entanto, é difícil de obter tal verdade absoluta se tratando de medição de sentimentos, que são subjetivos por natureza. Consideraremos, portanto, uma estratégia para executar esta avaliação. Avaliaremos um conjunto de eventos populares para os quais acreditamos que os sentimentos associados sejam esperados ou fáceis de serem verificados. Em seguida, descrevemos essa análise.

4.1 ANÁLISE DE EVENTOS POPULARES

Escolhemos nove eventos que foram amplamente noticiados por terem sido populares no Twitter. Estes eventos, resumidos na Tabela 3, abrangem tópicos relacionados a tragédias, lançamentos de produtos e filmes, política, saúde e eventos esportivos. Para extrair apenas *tweets* relevantes de cada evento identificamos o conjunto de palavras-chave que o descreve em consultas a *websites*, *blogs*, na Wikipédia e informações individuais. Dada nossa lista de palavras-chave selecionadas, identificamos os eventos pesquisando por essas palavras na nossa base de dados de *tweets*. Limitamos a duração de cada evento no Twitter pois palavras-chave populares são tipicamente utilizadas por *spammers*, após certo tempo (BENEVENUTO *et al.*, 2010).

A Tabela 3 exibe as palavras-chave utilizadas e o número total de *tweets* filtrados para cada evento considerado. A primeira coluna contém o nome que identifica o evento a que iremos nos referir no restante desse artigo.

Com o intuito de demonstrar que o PANAS-t pode medir variações de sentimentos compararemos os nove eventos descritos na Tabela 3 utilizando a representação de Kiviat. Em cada gráfico de Kiviat, linhas radiais a partir de um ponto central igual a -1 representam um sentimento com valor máximo de 1 (JAIN, 1991). Na Figura 1 plotamos os onze sentimentos e cada subfigura dessa representa um evento específico.

Nosso primeiro caso de estudo é o dia em que a gripe suína (também denominada H1N1) foi declarada pela Organização Mundial da Saúde (OMS) como a primeira pandemia desde a gripe de Hong Kong, em 1968, causando grande preocupação em toda população mundial. Após o anúncio, a OMS lançou várias advertências e precauções que deveriam ser tomadas pelos governos e pelo público, levando a população mundial a um estado de alerta contra a doença. Como podemos visualizar na Figura 1(a) esse evento modificou o estado de humor dos usuários aumentando os sentimentos de atenção ($P(s) = 0,8774$) e medo ($P(s) = 0,6768$) nos dias posteriores ao anúncio. Na verdade, esses dois sentimentos correspondem aos sentimentos esperados mais prováveis para esse evento, já que as pessoas estavam atentas às precauções bem como com medo de uma epidemia mundial.

Tabela 3. Índice de eventos analisados por data e palavras-chave relacionadas

Evento	Data	Palavras-chaves	# Tweets
H1N1	09 – 16/06/2009	tamiflu, outbreak, antiviral, virus, influenza, pandemia, pandemic, flu, pig, mexico, h1n1, swine, World Health Organization.	35.559
AirFrance	01 – 06/06/2009	victims, passengers, A330, 447, crash, airplane, airfrance.	29.765
Barack Obama	18 – 22/01/2009	barack obama, politics, democrats, election, white house, vote, speech, presidential, republicans, candidate, yes we can.	29.765
Eleições-EUA	02 – 06/11/2008	clinton, biden, palin, voting, vote, elect, candidate, campaign, mccain, democrats, republicans, obama, bush.	43.015
Morte-Michael Jackson	25 – 30/06/2009	rip, mj, michael jackson, death, died, king of pop, overdose, drugs, heart attack, conrad murray, thriller.	185.477
Susan Boyle	11 – 16/04/2009	susan boyle, i dreamed a dream, britain's got talent, les miserables.	56.259
Olimpíadas-2008	06 – 26/08/2008	olympics, medals, china, beijing, sports, peking, sponsor.	7.142
Tsunami	28/09/28 – 04/10/2009	tsunami, samoa islands, tonga, earthquake.	12.815
Terremoto-Haiti	11 – 17/01/2010	haiti, earthquake, richter, port au prince, jacmel, leogane.	236.096

Fonte: elaborada pelos autores

Em seguida, analisamos a queda de um avião em 1º de julho de 2009, que causou uma grande comoção no Twitter. O voo comercial 447 da empresa aérea AirFrance estava programado para ir do Rio de Janeiro à Paris, mas caiu no oceano matando todos os 216 passageiros e tripulantes. Após o acidente, podemos esperar que as pessoas se sentissem tristes pelas vítimas e também com medo de que esse tipo de acidente pudesse acontecer com elas. A Figura 1(b) mostra a representação Kiviat para esse evento. Como esperado, os sentimentos de medo ($P(s) = 0,7291$) e tristeza ($P(s) = 0,6992$) foram os dois predominantes nos *tweets* associados ao evento.

Quando se trata de política, as eleições norte americanas costumam ser eventos com amplo envolvimento de usuários da Web. Com a eleição, muitos eleitores podem se sentir apreensivos e empolgados com o poder de escolha que lhes é dado. Nossos resultados mostram sentimentos nessa direção. Como podemos ver na Figura 1(c), os usuários tiveram os sentimentos de *autoconfiança* ($P(s) = 0,6741$), *jovialidade* ($P(s) = 0,4277$) e *medo* ($P(s) = 0,3072$) aumentados.

Após as eleições, a posse de Barack Obama como presidente dos Estados Unidos foi amplamente esperada e discutida no Twitter. Como relatado em HCD (2009), a maioria dos norte-americanos estava mais confiante na melhoria do país após o discurso do presidente Obama.

Nossa análise de estado de ânimo dos usuários do Twitter, executada no dia do discurso do novo presidente, mostra um grande aumento nos sentimentos de *autoconfiança* ($P(s) = 0,7980$), seguido por *surpresa* ($P(s) = 0,5802$) e *jovialidade* ($P(s) = 0,5227$). Mas, apesar de toda a manifestação positiva em relação ao novo governo, também pudemos perceber um pequeno aumento no sentimento de tristeza ($P(s) = 0,1789$), o que poderia naturalmente representar *tweets* de opositores ao Barack Obama. A Figura 1(d) mostra que os sentimentos medidos com o PANAS-t estão de acordo com os relatos da referência (HCD, 2009).

O gráfico de Kiviat seguinte mostra o impacto da morte do cantor Michael Jackson. De acordo com o jornal *DailyMail* (BATES, 2009), sete dos dez tópicos mais populares no Twitter no período foram dedicados à sua morte. Após a morte do cantor, muitas especulações surgiram sobre quem ou o que matou Michael Jackson. Fãs e críticos culpavam o estresse causado por *paparazzi* e pela mídia. Na Figura 1(e) podemos ver um aumento nos sentimentos de *tristeza* ($P(s) = 0,4055$), *medo* ($P(s) = 0,5676$), *timidez* ($P(s) = 0,4055$), *culpa* ($P(s) = 0,1616$) e *surpresa* ($P(s) = 0,0810$). É curioso perceber que, além dos sentimentos esperados, associados à morte súbita do ídolo, como tristeza e medo, podemos notar a presença do sentimento de culpa. De fato, após especulações sobre a causa da morte, é natural que alguns usuários do Twitter se sentissem com uma parte da culpa e assim expressaram em seus *tweets*.

Em seguida, analisamos o evento denotado por Susan Boyle. A aparição da Susan Boyle como uma concorrente de um programa de TV chamado *Britain's Got Talent* teve uma incrível repercussão na mídia. O interesse mundial foi desencadeado pelo contraste entre sua voz poderosa cantando "*I Dreamed a Dream*", da banda "*Les Misérables*", e sua aparência simples no palco. O contraste da primeira impressão dada aos telespectadores, seguida de uma ovação durante e após sua *performance*, levou a um imediato viral, que se espalhou pelas redes sociais e obteve uma atenção enorme dos meios de comunicação globais. A Figura 1(f) mostra que os primeiros sentimentos expressos no Twitter associados à Susan Boyle foram de *surpresa* ($P(s) = 0,9066$), seguidos por *autoconfiança* ($P(s) = 0,4751$) e *culpa* ($P(s) = 0,1367$). Surpresa é definitivamente a sensação que fez com que sua aparição na TV se tornasse um *hit* na Internet. Esperávamos também que algumas pessoas se sentissem autoconfiantes, ao ver uma mulher enfrentar com sucesso uma audiência que, primeiramente, a ironizou. Finalmente, o sentimento de culpa também era esperado, levando em consideração que o evento se baseou em um pré-julgamento de valores com base na aparência.

O último evento analisado foi relativo às Olimpíadas de 2008, que ocorreram em Pequim, China. Durante os jogos Olímpicos, vários sentimentos poderiam ter sido expressos, de acordo com diferentes jogos que estavam ocorrendo. Assim, ao invés de utilizarmos gráficos de Kiviat, fizemos a medição para cada sentimento diariamente, durante todo o período do evento.

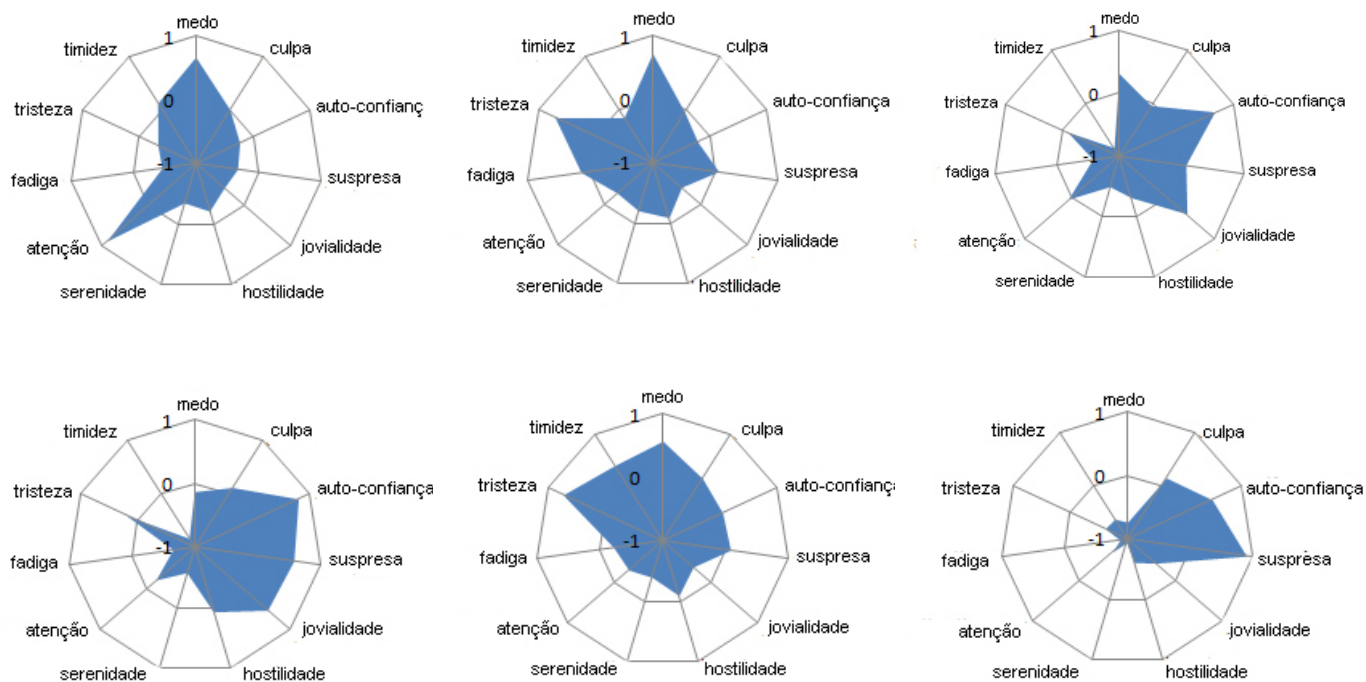


Figura 1. Sentimento, mostrado utilizando gráficos de Kiviat

Fonte: elaborada pelos autores

A Figura 2 apresenta um gráfico de *sparklines* que mostra as variações de cada um dos onze sentimentos definidos pelo PANAS-t durante os Jogos Olímpicos de Pequim. Como podemos perceber, no dia em que os jogos começaram, em 8 de agosto de 2008, tivemos predominância dos sentimentos de *surpresa* ($P(s) = 0,7024$), *atenção* ($P(s) = 0,4621$) e *jovialidade* ($P(s) = 0,3298$). No entanto, ao final do evento, no dia 24 do mesmo mês, podemos ver que esses sentimentos tiveram um decréscimo, enquanto o sentimento de *tristeza* aumentou de $P(s) = 0,1222$ para $P(s) = 0,5245$ no dia seguinte.

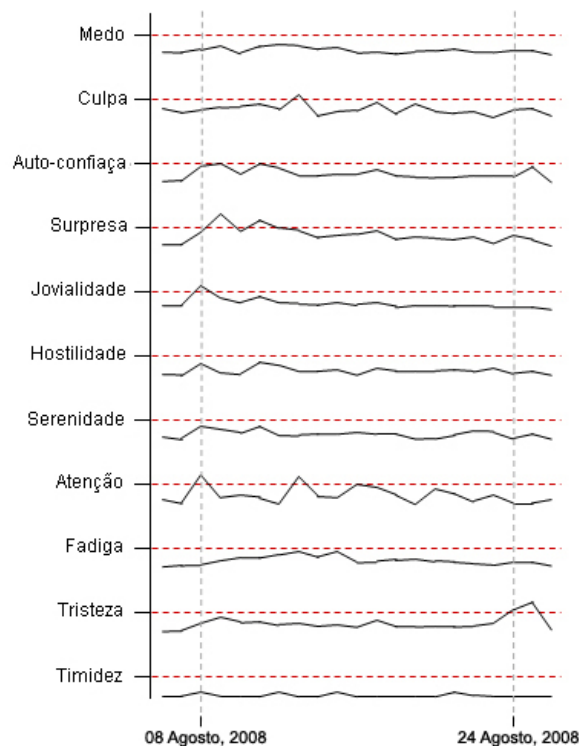


Figura 2. Exemplo gráfico de análise temporal do humor do público durante as Olimpíadas de Pequim, 2008

Fonte: elaborada pelos autores

4.2 GRIPE H1N1 VS. PANAS-T

Em 2009 o mundo foi surpreendido por uma epidemia do vírus H1N1 matando mais de meio milhão de pessoas. Também conhecida como "gripe suína", a doença foi declarada pela Organização Mundial da Saúde (OMS) como a primeira pandemia desde a Gripe de Hong Kong, em 1968. A partir de então, a OMS lançou várias medidas de precaução e preventivas que deveriam ser tomadas pelos governos e pela população em geral, gerando um estado de alerta contra a doença em todo o mundo.

Como explicado por Signorini *et al.* (2011), esse evento gerou uma elevada comoção entre os usuários do Twitter e geralmente as mensagens postadas por seus usuários estavam sincronizadas com os anúncios feitos pela OMS. Nessa seção mostraremos algumas análises feitas para comparar as variações de humor sofridas por usuários da rede em duas diferentes localidades. Mais especificamente, queremos mostrar como usuários norte-americanos e europeus se sentiram em relação a esse evento e se houve relação entre os sentimentos expressos por eles.

É importante ressaltar que nosso método é específico para textos em inglês, e, por essa razão, para usuários europeus filtramos apenas *tweets* de gente que vivia em países onde o inglês é o idioma principal. Em outras palavras, filtramos apenas mensagens publicadas por usuários que viviam nos Estados Unidos, Irlanda, Países Baixos, República Malta e Reino Unido. Para isso utilizamos uma base de dados coletadas por Kulshrestha *et al.*

(2012) e que contém uma lista de todos os *IDs* (que identificam cada usuário unicamente na rede social) de todos os usuários do Twitter separados por localização geográfica.

Os gráficos de *sparklines* apresentados na Figura 3 mostram as variações de humor dos usuários europeus e norte-americanos, respectivamente, para 4 dos 11 sentimentos definidos pelo PANAS-t. Como podemos perceber, os gráficos estão marcados com quatro datas que indicam o dia em que um importante anúncio foi feito pela OMS em 2009.

Em março de 2009, autoridades mexicanas começaram a anunciar casos do que a OMS definiu na época como "*influenza-like illness*", ou infecção respiratória aguda (IRA). Como apresentado no gráfico da Figura 3(a), usuários europeus tiveram um aumento no sentimento de *surpresa* ($P(s) = 0,8730$) nessa data, mas o mesmo não foi apresentado pelos usuários norte-americanos (Figura 3(b)).

No mês seguinte, em abril de 2009, o primeiro caso da gripe H1N1 nos Estados Unidos foi confirmado pela OMS, e os gráficos mostram um aumento similar no sentimento de *medo* para ambas às localidades com $P(s) = 0,7401$ para Europa e $P(s) = 0,6154$ para os EUA. Também podemos perceber um aumento no sentimento de *atenção*, mas apenas nos usuários europeus, de ($P(s) = 0,4423$).

A próxima data, em junho do mesmo ano, marca o dia em que a OMS declarou a gripe H1N1 como uma epidemia mundial, causando aumento nos sentimentos de *medo* ($P(s) = 0,5385$), mas também de *atenção* em usuários norte-americanos, ($P(s) = 0,3491$) e em usuários europeus ($P(s) = 0,3174$).

Em julho, quase 27 mil casos da doença haviam sido confirmados apenas na Europa. Nosso método mostra um aumento nos sentimentos de medo ($P(s) = 0,4887$), principalmente nos usuários que vivem naquele continente.

A última data marcada nos gráficos é de agosto de 2009, quando os dois países mais afetados pela doença estavam localizados na América e na Europa (WHO, 2009) Nesse período, usuários europeus tiveram um aumento no sentimento de *hostilidade* ($P(s) = 0,2542$), mas, diferentemente, os usuários norte-americanos tiveram aumento no sentimento de *medo* ($P(s) = 0,4112$).

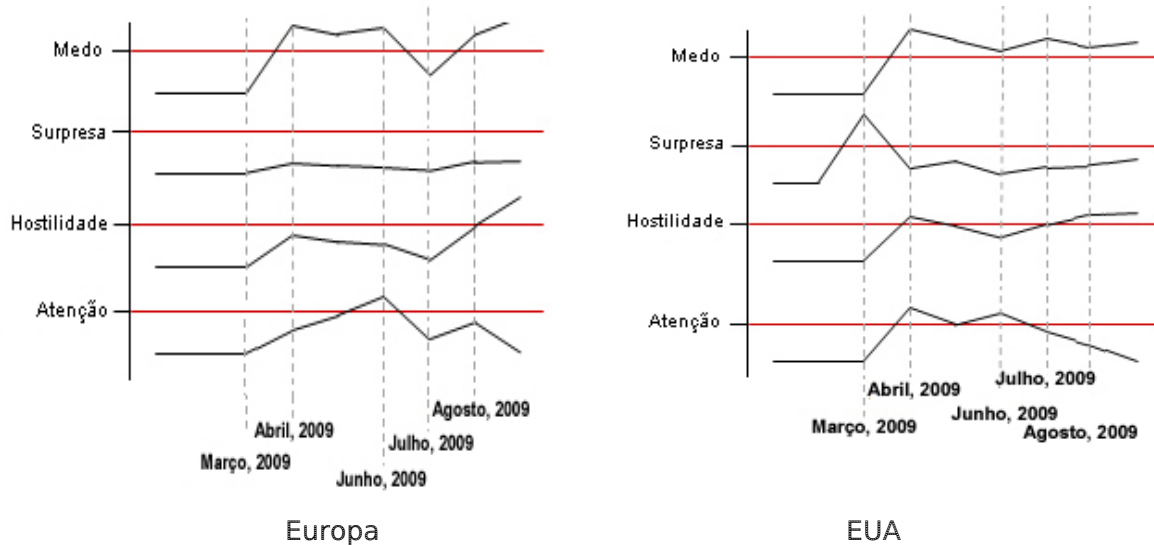


Figura 3. Humor dos usuários relacionado à H1N1 durante o ano de 2009 para Europa e Estados Unidos

Fonte: elaborada pelos autores

4.3 UTILIZANDO O PANAS-T EM OUTRA BASE DE DADOS

Os valores de base calculados para PANAS-t e apresentados na Tabela 2 também podem ser utilizados para detectar sentimentos em base de dados de diferentes períodos. Nessa seção, utilizamos uma base de dados diferente da base de dados original, que continha *tweets* produzidos entre 2006 e meados de 2009. Nossa nova base de dados, também proveniente da rede Twitter, contém *tweets* postados entre o final de 2009 e 2010, e foi coletada por Michael e Dredze (2011).

Para nossas análises utilizamos *tweets* da nova base que estivessem associados a dois eventos da Tabela 3. Em setembro de 2009 um terremoto submarino de magnitude 8.1, a maior de 2009, gerou um tsunami que atingiu as Ilhas Samoa e Tonga, matando pelo menos 189 pessoas. Esse evento causou uma grande comoção ao redor do mundo e gerou um estado de alerta em diversos países vizinhos aos países atingidos (MANAK, 2009). A Figura 4(a) apresenta o gráfico de Kiviat com os sentimentos que os usuários do Twitter expressaram no dia do acidente e um dia após a ocorrência do abalo. Como podemos perceber, o evento gerou um aumento nos sentimentos de *medo* ($P(s) = 0,9280$), *atenção* ($P(s) = 0,9932$), *hostilidade* ($P(s) = 0,8451$), *surpresa* ($P(s) = 0,6528$) e *tristeza* ($P(s) = 0,6483$).

Um evento similar aconteceu em janeiro de 2010, quando um terremoto atingiu o Haiti causando uma grande catástrofe natural no país. O evento causou maior dano na capital do país, Porto Príncipe e em Jacmel, matando pelo menos 250 mil pessoas. A Figura 4(b) mostra que sentimentos de *hostilidade* ($P(s) = 0,9280$), *atenção* ($P(s) = 0,3678$), *surpresa* ($P(s) = 0,4576$) e *tristeza* ($P(s) = 0,3975$) tiveram um aumento. Curiosamente também podemos perceber um aumento nos sentimentos de *timidez* e *culpa*. Alguns dias após o acontecimento o mundo se voltou

para o desastre. Pessoas de todo lugar uniram forças para ajudar o país que estava devastado (OPERATION COMPASSION, 2010). A hipótese é que essa situação tenha gerado o aumento nesses dois sentimentos, já que as pessoas só começaram a perceber a situação precária em que vivia o país, mesmo antes da tragédia, após os fatos.

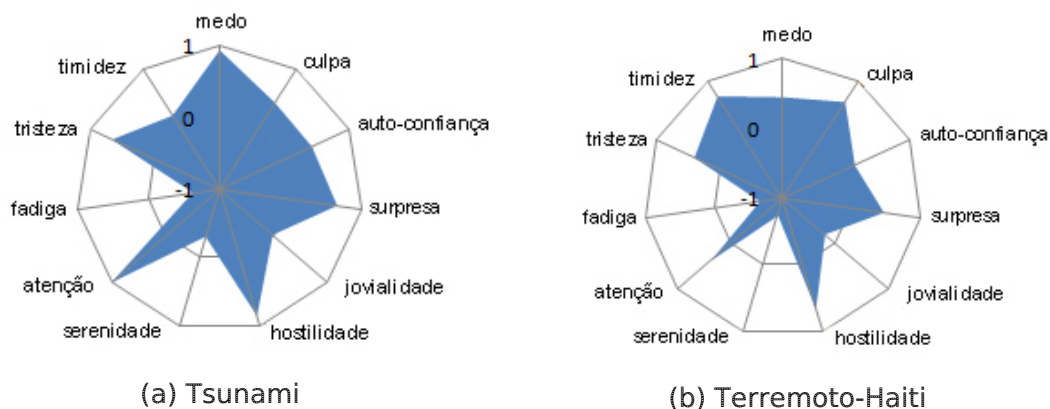


Figura 4. Eventos fora da base calculada e sentimentos associados

Fonte: elaborada pelos autores

5 CONCLUSÃO E TRABALHOS FUTUROS

Nesse trabalho, apresentamos o PANAS-t, uma escala psicométrica que mede onze sentimentos, adaptada ao contexto do Twitter, que se baseia na versão estendida da conhecida escala *Positive Affect Negative Affect Scale* (PANAS). Utilizando dados empíricos de um conjunto único de dados do Twitter contendo cerca de 1,8 bilhões de *tweets*, fomos capazes de calcular pontuações de normalização para cada sentimento.

Nossos resultados fornecem fortes evidências de que o PANAS-t é capaz de capturar com precisão sentimentos positivos e negativos sobre acontecimentos no Twitter. Mostramos também que o PANAS-t se sobressai quando comparado com métodos que utilizam *emojicons* para detectar sentimentos. Outra vantagem do nosso método é que ele pode ser utilizado em outras bases de dados com o objetivo de detectar sentimentos. Especulamos que esta escala de medição possa ser utilizada por quaisquer pesquisas com o objetivo de criar ferramentas que possam ser usadas por agências governamentais ou empresas interessadas em melhorar seus produtos utilizando redes sociais. De uma perspectiva da pesquisa, nosso método permite compreender como, quando e porque os indivíduos sentem e como seus sentimentos variam de acordo com eventos sociais, políticos e econômicos.

Existem várias frentes de pesquisa em que pretendemos utilizar o PANAS-t. Pretendemos combinar outras técnicas de aprendizagem de máquina para então, dinamicamente, incorporar sentimentos ao PANAS-t de

acordo com o contexto, assim como melhorar a detecção de sentimentos, acrescentando gírias, muito utilizadas principalmente pelos usuários do Twitter, já que há limitações na rede no tamanho das mensagens postadas (GONÇALVES *et al.*, 2016). Pretendemos ainda comparar o PANAS-t com outros métodos em tarefas como comparação de métodos (GONÇALVES *et al.*, 2013b; RIBEIRO *et al.* 2016). Além disso, pretendemos incorporar o PANAS-t como parte do sistema ifeel (ARAUJO *et al.*, 2014) e investigar o uso de ferramentas de tradução automática de texto para o uso do PANAS-t em múltiplas línguas (ARAUJO *et al.*, 2016).

6 AGRADECIMENTOS

Esse trabalho foi apoiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela Fundação de Amparo à Pesquisa do estado de Minas Gerais (Fapemig) e pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

REFERÊNCIAS

ARAUJO, Matheus; GONÇALVES, Pollyanna; CHA, Meeyoung; BENEVENUTO, Fabrício. iFeel: a system that compares and combines sentiment analysis methods. In: Proceedings of the companion publication of the 23rd international Conference on World Wide Web, 2014.

ARAUJO, Matheus; REIS, Julio; PEREIRA, Adriano; BENEVENUTO, Fabrício. An evaluation of machine translation for multilingual sentence-level sentiment analysis. In: ACM Symposium on Applied Computing (SAC), 2016.

BALOG, Krisztian; MISHNE, Gilad; Rijke, Maarten. Why are they excited? Identifying and explaining spikes in blog mood levels. 11th Meeting of the European Chapter of the Association for Computational Linguistics (EACL), 2006.

BATES, C. How Michael Jackson's death shut down Twitter, brought chaos to Google...and 'killed off' Jeff Goldblum. *Mail Online – Science Tech*. June, 2009. Disponível em: <http://bit.ly/16e6eM>. Acesso em 15/01/2012.

BENEVENUTO, Fabrício; MAGNO, Gabriel, RODRIGUES, Tiago; ALMEIDA, Virgílio. Detecting spammers on Twitter. 7th Annual Collaboration, Electronic messaging, Anti-Abuse and Spam Conference (CEAS), 2010.

BOLLEN, Johan; PEPE, Alberto; MAO, Huina. Modeling public mood and emotion: twitter sentiment and socio-economic phenomena. *Computer*, abs/0911.1:450–453, 2009.

CHA, Meeyoung; HADDADI, Hamed; BENEVENUTO, Fabricio; GUMMADI, Krishna. Measuring user influence in Twitter: the million follower fallacy. 5th International AAAI Conference on Weblogs and Social Media (ICWSM), May 2010.

CHEONG, Marc; LEE, Vincent. A micro blogging-based approach to terrorism informatics: exploration and chronicling civilian sentiment and response to

terrorism events via Twitter. *Information Systems Frontiers*, n. 13, p. 45–59, March 2011.

COMARELA, Giovanni; CROVELLA, Mark; ALMEIDA, Virgílio; BENEVENUTO, Fabrício. Understanding factors that affect response rates in Twitter. In: Proceedings of the 23rd ACM Conference on Hypertext and Social Media, p. 123-132, 2012.

DIAKOPOULOS, Nicholas; SHAMMA, David. Characterizing debate performance via aggregated Twitter sentiment. 28th International Conference on Human Factors in Computing Systems, CHI '10, p. 1195–1198, 2010.

DODDS, Peter; DANFORTH, Christopher. Measuring the happiness of large-scale written expression: songs, blogs, and presidents. *Journal of Happiness Studies*, v. 11, n. 4, p. 441–456, August 2010. <http://dx.doi.org/10.1007/s10902-009-9150-9>

GOMIDE, Janaina; VELOSO, Adriano; MEIRA, Vagner; ALMEIDA, Virgílio; BENEVENUTO, Fabrício; FERRAZ, Fernanda; TEIXEIRA, Mauro; Dengue surveillance based on a computational model of spatio-temporal locality of Twitter. In: ACM Web Science Conference (WebSci), 2011.

GONÇALVES, Pollyanna; ARAÚJO, Matheus; BENEVENUTO, Fabrício; CHA, Meeyoung. Comparing and combining sentiment analysis methods. In: ACM Conference on Online Social Networks (COSN), 2013b.

GONÇALVES, Pollyanna; BENEVENUTO, Fabrício; CHA, Meeyoung. Panas-t: A psychometric scale for measuring sentiments on Twitter. arXiv preprint arXiv:1308.1857, 2013a.

GONÇALVES, Pollyanna; DALIP, Daniel; COSTA, Helen; GONÇALVES, Marcos; BENEVENUTO, Fabrício. On the combination of “off-the-shelf” sentiment analysis methods. In: ACM Symposium on Applied Computing (SAC), 2016.

GONÇALVES, Pollyanna; DORES, Wellington; BENEVENUTO, Fabrício. PANAS-t: Uma escala psicométrica para medição de sentimentos no Twitter. 1st Brazilian Workshop on Social Network Analysis and Mining (BrasNAM), 2012.

HCD Research Inc. Confidence levels increase among democrats and independents, decrease among republicans after viewing Obama’s press conference. *Tell us Detroit*, 24 de março, 2012. Disponível em: <http://www.tellusdetroit.com/politics/hcd-research.html>. Acesso em 15/01/2012.

JAIN, Raj. *The art of computer systems performance analysis: techniques for experimental design, measurement, simulation, and modeling*, 1st edition. John Wiley and Sons, 1991.

KIM, Elsa; GILBERT, Sam. Detecting sadness in 140 characters: sentiment analysis and mourning Michael Jackson on Twitter. *Web Ecology*, n. 3, August, 2009.

KIRKPATRICK, Marshall. There are now 155m tweets posted per day, triple the number a year ago. *ReadWrite*, April 6, 2011. Disponível em: <http://rww.to/gv4VqA>. Acesso em 26/03/2013.

KULSHRESTHA, Juhi; KOOTI, Farshad; NIKRAVESH, Ashkan, GUMMADI, Krishna. Geographic dissection of the Twitter network. In: International AAAI Conference on Weblogs and Social Media, 1:1-8, 2012.

MACASKILL, Ewen. US confirms it asked Twitter to stay open to help Iran protesters. *The Guardian*, June 17, 2009. Disponível em: <http://www.theguardian.com/world/2009/jun/17/obama-iran-twitter>. Acesso em: 26/03/2013

MANAK, Mirian. Fear for new tsunami after earthquake hits Sumatra. *Digital Journal*, Sep. 2009. Disponível em: digitaljournal.com/article/279868. Acesso em: 26/03/2013.

MEEYOUNG, Cha; BENEVENUTO, Fabrício; HADDADI, Hamed; GUMMADI Krishna. The world of connections and information flow in Twitter. In: IEEE Transactions on Systems, Man and Cybernetics - Part A, 2012.

MICHAEL, Paul; DREDZE, Mark. You are what you tweet: analyzing Twitter for public health. In: 5th International AAAI Conference on Weblogs and Social Media (ICWSM), 2011.

NASCIMENTO, Paula; AGUAS, Rodrigo; LIMA, Débora; KONG, Xiao; OSIEK, Bruno; XEXÉO, Geraldo; SOUZA, Jano. Análise de sentimento de Tweets com foco em notícias. 1st Brazilian Workshop on Social Network Analysis and Mining (BrasNAM, 2012).

O'CONNOR, Brendan; BALASUBRAMANYAN, Ramnath; ROUTLEDGE, Bryan; SMITH, Noah. From tweets to polls: linking text sentiment to public opinion time series. In: International AAAI Conference on Weblogs and Social Media (ICWSM), 2010.

Operation Compassion. Operation compassion responds to Haitian earthquake victims. January 13, 2010. Disponível em: www.operationcompassion.org/2010/01/operation-compassion-responds-to-haitian-earthquake-victims. Acesso em 26/03/2013.

PANG, Bo; LEE, Lillian. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, v. 2, n. 1-2, p. 1-135, 2008.

RAO, Lenna. Twitter seeing 90 million tweets per day, 25 percent contain links. *Tech Crunch*, Sep 14, 2010. Disponível em: <http://techcrunch.com/2010/09/14/twitter-seeing-90-million-tweets-per-day/>. Acesso em 26/03/2013.

RIBEIRO, Filipe; ARAÚJO, Matheus; GONÇALVES, Pollyanna; BENEVENUTO, Fabrício; GONÇALVES, Marcos. A benchmark comparison of state-of-the-practice sentiment analysis methods. arXiv preprint. arXiv preprint arXiv:1512.01818, 2015.

SAKAKI, Takeshi; OKAZAKI, Makoto; MATSUO, Yutaka. Earthquake shakes Twitter users: real-time event detection by social sensors. In: 19th International Conference on World Wide Web, WWW '10, p. 851–860, 2010.

SIGNORINI, Alessio; SEGRE, Alberto, NIKRAVESH, Ashkan; POLGREEN, Philip. The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. PLoS ONE, n. 6, p. 1–10, 2011.

TUMASJAN, Andranik; SPRENGER, Timm; SANDNER, Philipp; WELPE, Isabell. Predicting elections with Twitter: what 140 Characters Reveal About Political Sentiment. Word Journal of the International Linguistic Association, p. 178–185, 2010.

WATSON, D.; CLARK L. Development and validation of brief measures of positive and negative affect: the Panas scales. *Journal of Personality and Social Psychology*, 54(1):1063–1070, 1985.

WHO - World Health Organization. Map of affected countries and deaths as of 23 august 2009. Disponível em: www.who.int/csr/don/2009_08_28/en/index.html. Acesso em 26/03/2013