

震災時の情報流通を支える Twitter アカウントの発見

Twitter accounts to support the information distribution of earthquake disaster

石原裕規^{*1} 諏訪博彦^{*1} 鳥海不二夫^{*2} 太田敏澄^{*1}
 Hiroki Ishihara Hirohiko Suwa Fujio Toriumi Toshizumi Ohta

^{*1} 電気通信大学
 The university of Electro-Communications

^{*2} 東京大学
 The University of Tokyo

The aim of our study is to identify active account of the online communication for the Great East Japan and to understand the type of their communication on Twitter. There was a lot of information about disaster on Twitter. At that time, the account which diffuses or mediates information is important. For discovering those accounts, we construct a communication network per 1day and calculate two centrality measures (degree and betweenness) on Twitter. In addition, we compare the type of their communication before and after the disaster. As a result, we found that 1)communication is increased after the earthquake, 2)important account is newspapers, news sites and bot before the earthquake, 3)important account is public broadcasting and celebrities after the earthquake, 4)most important accounts have been retweeted after the earthquake.

1. はじめに

2011年3月11日に発生した東日本大震災においては、Twitter に多くの情報が流通していた。この情報流通においては、情報拡散の起点となるアカウントや、情報を仲介するアカウントが重要となる。情報拡散の起点となるアカウントは、数多くのアカウントとコミュニケーションを行うアカウントである。情報を仲介するアカウントは、様々なカテゴリのアカウントとコミュニケーションを行うアカウントである。これらのアカウントは、ネットワーク分析における次数中心性と媒介中心性で表現できると考える。

本研究の目的は、震災前後において重要なアカウントを発見することである。更に、それらのアカウントがどのように他アカウントと繋がっていたのかを明らかにすることである。

2. 関連研究

災害時における Twitter に関する研究として、Vieweg[1]らは、2009年に発生したオクラホマの山火事やレッドリバーでの洪水における Twitter 利用について分析している。投稿されたツイートの内容を地域情報や災害の最新情報の有無で分類し、その広がり方を調査している。Twitter での分析から、緊急事態へのシチュエーション・アウェアネス(事態への気づき)の情報取得について、フレームワークの提案を行っている。

Cheong ら[2]は、2010年から2011年に発生したオーストラリアでの洪水における Twitter 利用について分析している。ユーザがツイートに返答する関係をネットワークとして生成し、ネットワーク分析における様々な中心性の指標を算出している。その結果、地方自治体、国会議員、社会メディアボランティア、公共のメディアレポーター、非営利組織が重要なアカウントであったことを確認している。

東日本大震災における Twitter に関しても多くの研究が行われている[3][4]。鳥海ら[3]は、2011年3月7日から23日の Twitter の日本語ツイートデータを用い、東日本大震災前後における関連ネットワークの変化を分析している。震災後は震災前と比較し、ユーザが密に繋がり、情報がより広範囲に伝播するようになったことを示している。

東日本大震災における重要アカウントに関する研究では、鳥海ら[3]が次数中心性に基づく分析を行っている。しかしながら

次数中心性だけでは、情報を仲介するアカウントを抽出できないと考える。そこで我々は、次数中心性と媒介中心性に基づいて重要アカウントの発見を試みる。

Cheong ら[2]は、オーストラリアの洪水を対象に次数中心性と媒介中心性に着目した分析を実施しており、本研究もその手法を参考とする。Cheong らは、災害後の単一ネットワークを分析対象としているのに対し、我々は、震災前後の複数ネットワークを対象とする。

本研究では、次数中心性と媒介中心性の高いアカウントの変化を震災前後で分析する。また、その際に各中心性が高いアカウントが、どのようなコミュニケーション形態をとっているのか分析する。

3. ネットワーク分析による重要アカウントの特定

震災前後の情報流通において重要なアカウントを抽出するために、ネットワーク分析を行う。情報拡散の起点となるアカウントは、ネットワーク分析の代表的な指標である次数中心性が高くなると考える。また、情報を仲介するアカウントは、媒介中心性が高くなると考える。本章では、震災前後の重要アカウントを2つの指標を用いて特定し、その差異を示す。

3.1 分析概要

分析の手順は以下の3つのステップで構成される。

1. ネットワーク生成
2. 次数中心性・媒介中心性の算出
3. 震災前後における重要アカウントの比較検討

まず、東日本大震災前後のツイートデータを基に、1日毎にコミュニケーションネットワークを生成する。次に、ネットワーク毎に次数中心性と媒介中心性を算出する。各中心性が高いアカウントを重要アカウントとみなし、震災前後の変化を比較する。

本研究の分析対象は、以下の通りである。2011年3月5日から24日に日本語で Twitter に投稿されたツイートの収集を行った。取得した総ツイート数は3億3235万8199件である。また取得したツイートの日ごとのコミュニケーション(リプライ、リツイート)に基づいて無向ネットワークを作成する。さらに、次数中心性、媒介中心性を日ごとに算出する。

情報拡散の起点となるアカウントを特定するために、次数中心性を算出する。次数中心性(Degree Centrality)とは、ノード同士をつなぐ関係(リンク)の多寡によりノードの重要性を評価する

指標である。ノードが持つ、リンク数が大きければ大きいほど、他ノードと関係をもっているため、情報拡散の起点となる重要なノードと考える。次数中心性は下記の式で表される。

$$DC(v) = \frac{\text{Degree}(v)}{N - 1}$$

DC(v) : 各ノード v の次数中心性

Degree(v) : 各ノード v のリンク数

N : ノード数

情報を仲介するアカウントを特定するために、媒介中心性を算出する。媒介中心性(Betweenness Centrality)とは、ノード間の連結関係上の重要性を評価する指標である。情報が最短経路上で伝わるという前提の元、最短経路上に現れるノードほど、情報を媒介する重要なノードと考える。全対全の最短経路を求め、各ノードが最短経路上にどれだけ現れたかにより算出することができる。媒介中心性は下記の式で表される。

$$BC(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

BC(v) : 各ノード v の媒介中心性

: ノード s とノード t 間の最短経路の数

: s から t の最短経路上においてノード v を通るものの数

Freeman[5]が提案した媒介中心性は、全対全の最短距離を求める事が必要になり、ノード数が増加すると計算時間がかかるため、様々な高速化の研究がなされている本研究では、計算時間の短縮のために、近似解を用いることとする[6]。

3.2 分析結果

図 1 は、震災前後におけるノード数とリンク数の推移を表している。震災前では、ノード数の最大は 129 万であり、リンク数は最大 342 万である。震災発生日である 11 日のノード数は 158 万であり、リンク数は 746 万である。震災後である 12 日から 24 日においては、ノード数が徐々に減少していることが確認できる。このことから、震災によりコミュニケーションが急激に活発化し、徐々に収束していったことがわかる。

そこで、具体的に震災前後でどのような変化があったかを明らかにするために、以後の分析では 3 月 6 日から 10 日を震災前、11 日から 15 日を震災後として分析を行う。

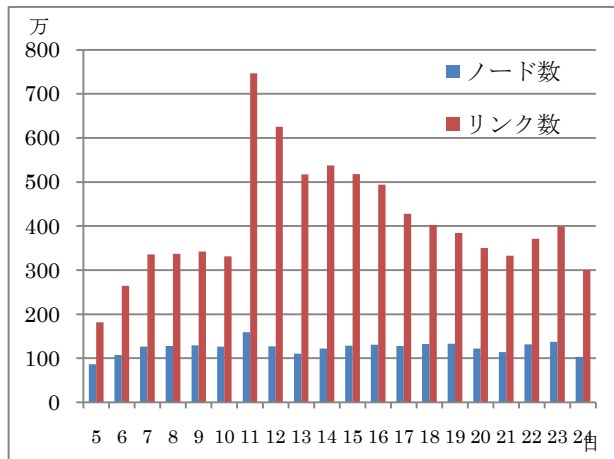


図 1 震災前後におけるリンク数とノード数の推移

表 2 は、震災前である 6 日から 10 日の次数中心性の上位 10 アカウントである。wwwwww_bot, shuumai などは Bot が数多く確認されている。Bot 以外のアカウントとしては、youtube, foursquare などの web サービスのアカウントや、Yomiuri_Online, 47news などのニュースアカウントが確認されている。

表 3 は、震災前である 6 日から 10 日の媒介中心性の上位 10 アカウントである。表 2 と同じように、shuumai, youtube などのアカウントが確認できる。一方で、表 2 では上位に現れなかった AddThis(ソーシャルブックマーク), justinbieber(ジャスティンビーバー:外国人歌手)などのアカウントが確認できる。また、上位 10 位以内ではないが、沢山 RT されているツイートを RT する viratter や Retweeterjp などの Bot が上位となっている。

表 4 は、震災後である 11 日から 15 日の次数中心性の上位 10 アカウントである。NHK_PR, nhk_seikatsu, nhk_HORIJUN など NHK のアカウントが数多く確認できる。また、震災に関連して、FDMA_JAPAN(総務省消防庁), radio_rfc_japan(ラジオ福島), Kantei_Saigai(首相官邸災害)などが確認できる。それ以外では、放射能についてツイートした hayano(早野龍五)や tsuda(津田大介), ayu_19980408(浜崎あゆみ)などが確認できる。

表 2 震災前の次数中心性の順位 1 位~10 位のアカウント名

順位	6 日	7 日	8 日	9 日	10 日
1	youtube	shuumai	shuumai	shuumai	youtube
2	shuumai	youtube	youtube	youtube	shuumai
3	wwwwww_bot	exitelecaster	wwwwww_bot	earthquake_jp	Setsulla
4	kopipedoujou	wwwwww_bot	0101xiahtic	wwwwww_bot	natalie_mu
5	sazae_f	kopipedoujou	kaorugundam	lgm_	wwwwww_bot
6	Nijinoasi	sazae_f	SoalCINTA	michan39	foursquare
7	0101xiahtic	gizmodoJapan	foursquare	Twinavi	Swedenhills
8	rinningo2525	natalie_mu	karashichan	natalie_mu	Yomiuri_Online
9	foursquare	MentionKe	mainichijpnews	Foursquare	SoalCINTA
10	Komatchy	foursquare	MentionKe	sazae_f	47news

表 3 震災前の媒介中心性の順位 1 位~10 位のアカウント名

順位	6 日	7 日	8 日	9 日	10 日
1	youtube	youtube	youtube	youtube	youtube
2	foursquare	shuumai	shuumai	shuumai	foursquare
3	shuumai	foursquare	foursquare	earthquake_jp	Shuumai
4	sazae_f	AddThis	AddThis	foursquare	AddThis
5	kopipedoujou	exitelecaster	wwwwww_bot	AddThis	Swedenhills
6	wwwwww_bot	MentionKe	kaorugundam	wwwwww_bot	wwwwww_bot
7	rinningo2525	wwwwww_bot	0101xiahtic	sazae_f	natalie_mu
8	AddThis	justinbieber	SoalCINTA	natalie_mu	Setsulla
9	tetra_d	sazae_f	justinbieber	lgm_	Justinbieber
10	h_ototake	kopipedoujou	MentionKe	NHK_PR	sazae_f

表 4 震災後の次数中心性の順位 1 位~10 位のアカウント名

順位	11 日	12 日	13 日	14 日	15 日
1	NHK_PR	NHK_PR	NHK_PR	nhk_seikatsu	NHK_PR
2	kazhime1977	nhk_seikatsu	masason	NHK_PR	takapon_jp
3	FDMA_JAPAN	nhk_HORIJUN	nhk_seikatsu	Asahi_Shakai	tsuda
4	earthquake_jp	twj	Asahi_Shakai	technon_yousuke	kopipedoujou
5	oohamazaki	earthquake_jp	pref_iwate	nhk_kabun	hayano
6	ayu_19980408	touhokujishin	h_ototake	Kantei_Saigai	tarourakami
7	tsuda	takapon_jp	radio_rfc_japan	think_justice	itsumonoTL
8	itkz	fukanju	touhokujishin	takapon_jp	twinavi
9	shi_tya	haiirochuuhi	HandsNet	itsumonoTL	nhk_kabun
10	nhk_news	TMR15	takapon_jp	nhk_news	routao

表 5 震災後の媒介中心性の順位 1 位～10 位のアカウント名

順位	11 日	12 日	13 日	14 日	15 日
1	NHK_PR	NHK_PR	NHK_PR	nhk_seikatsu	NHK_PR
2	kazhime1977	nhk_seikatsu	masason	NHK_PR	takapon_jp
3	FDMA_JAPAN	nhk_HORIJUN	nhk_seikatsu	Asahi_Shakai	kopipedoujou
4	earthquake_jp	twj	Asahi_Shakai	technon_yousuke	tsuda
5	oohamazaki	earthquake_jp	hatakezo	nhk_news	tarourakami
6	ayu_19980408	touhokujishin	h_ototake	youtube	hayano
7	tsuda	fukanju	radio_rfc_japan	takapon_jp	youtube
8	itkz	takapon_jp	touhokujishin	Kantei_Saigai	nhk_seikatsu
9	shi_tya	ayu_19980408	Addthis	think_justice	Twinavi
10	youtube	masason	takapon_jp	nhk_kabun	itsumonoTL

表 5 は、震災後である 11 日から 15 日の媒介中心性の高い上位 10 アカウントである。表 4 と同じように、NHK_PR, nhk_seikatsu, nhk_HORIJUN や radio_rfc_japan, Kantei_Saigai などが確認できる。

図 2 の左図は、各アカウントの次数中心性の順位と媒介中心性の順位の散布図である。次数中心性順位と媒介中心性順位の相関が高いことが確認できる。また、数は少ないが、次数中心性の順位は高いが媒介中心性の順位が低いアカウントや、次数中心性の順位は低いが媒介中心性の順位が高いアカウントの存在が確認できる。

各アカウントを特徴付けるために、2 つの指標を用いてアカウントを分類する(図 2, 右図)。次数中心性と媒介中心性が両方とも高い(100 位以上)アカウントをアクティブアカウントとする。また、どちらかの指標が高い(100 位以上)にもかかわらず、もう一方の指標が低い(300 位以下)のアカウントを、それぞれ局所ハブアカウント・仲介アカウントとする。それ以外は非アクティブアカウントとする。

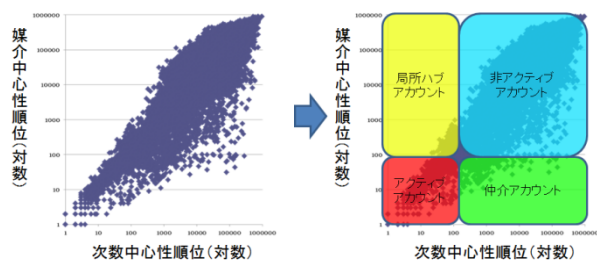


図 2 次数中心性と媒介中心性の順位相関と分類結果

表 6 に震災前後のアクティブアカウントの一例を示す。震災前後において、NHK_PR, 47news, Yomiuri_Online らのニュースアカウントは、アクティブアカウントであった。また、震災前においてアクティブアカウントでない nhk_seikatsu, nhk_news, nhk_kabun, nhk_HORIJUN, Asahi_Shakai らは、震災後にアクティブアカウントとなっている。

個人アカウントとして、masason, ayu_19980408, らは、震災前後においてアクティブアカウントである。震災後にアクティブアカウントとなった個人アカウントとして、hayano(早野龍五)などがある。hayano は、放射能に関するツイートを数多く行っている。

その他に分類した shuumai, kopipedoujou, sazae_f らは、Bot である。震災前は、これらのアカウント以外にも数多くの Bot が確認できていた。一方で、震災後は、kopipedoujou, sazae_f 以外の Bot は確認出来なかった。また震災後は、震災に関する情報を提供していた OfficialTEPCO(東京電力), Kantei_Saigai(首相官邸災害), team_nakagawa(東大病院放射線治療チーム),

radio_rfc_japan(ラジオ福島)らがアクティブアカウントとなったことが確認できる。

表 7 に震災前後の局所ハブアカウントを示す。局所ハブアカウントは、震災前において FlandreBot, ReiuiziUtsuhoBot などのいくつかの Bot のみである。また、震災後では、局所ハブアカウントは確認されていない。

表 8 に震災前後の仲介アカウントを示す。個人アカウントとして、震災前では、ladygaga(レディガガ), 23kin(日本語, インドネシア語, 韓国語などの言語でツイートするアカウント), Charlesheen(チャリー・シーン)などがあげられる。震災後では、ladygaga(レディガガ), detikcom(インドネシアのサービス), justinbieber, TheDenials_Risa などが確認できている。

その他のアカウントとして、震災前では viratter, Retweeterjrp(RT されてるツイートを RT)らの Bot が確認できているが、震災後では確認されていない。

表 6.震災前後のアクティブアカウント

	震災前アカウント	震災後アカウント
ニュースアカウント	NHK_PR(NHK 広報局) 47news(共同通信ニュース) Yomiuri_Online(読売新聞ニュースサイト)	NHK_PR(NHK 広報局) 47news(共同通信ニュース) Yomiuri_Online(読売新聞ニュースサイト) nhk_seikatsu(NHK 生活情報部) nhk_news(NHK ニュース) nhk_kabun(NHK 科学文化部) nhk_HORIJUN(NHK 堀潤) Asahi_Shakai(朝日新聞社会部)
個人アカウント	masason(孫正義:実業家) ayu_19980408(浜崎あゆみ:アーティスト) sasakitoshinao(佐々木俊尚:ジャーナリスト)	masason(孫正義:実業家) ayu_19980408(浜崎あゆみ:アーティスト) sasakitoshinao(佐々木俊尚:ジャーナリスト) hayano(早野龍五:大学教授)
その他	Kopipedoujou(コピペ道場 bot) sazae_f(サザエさん bot) shuumai(シュウマイ君 bot) wwwwww_bot(笑ったら RT してね bot)	Kopipedoujou(コピペ道場 bot) sazae_f(サザエさん bot) touhokujishin(東日本大震災の情報センター) OfficialTEPCO(東京電力) Kantei_Saigai(首相官邸災害) team_nakagawa(東大放射線治療チーム) radio_rfc_japan(ラジオ福島)

表 7.震災前後の局所ハブアカウント

	震災前アカウント	震災後アカウント
その他	FlandreBot(キャラクターbot) ReiuiziUtsuhoBot(キャラクターbot)	

表 8.震災前後の仲介アカウント

	震災前アカウント	震災後アカウント
個人アカウント	ladygaga(レディガガ) 23kin(インドネシア語, 韓国語でもツイート) Charlesheen(チャリー・シーン:女優)	ladygaga(レディガガ) detikcom(インドネシアのサービス) justinbieber(ジャスティンビーバー) TheDenials_Risa
その他	viratter(RT されてるツイートを RT) retweeterjrp(RT されてるツイートを RT)	

図 3 は仲介アカウントの一例である。TheDenials_Risa というアカウントから 2-hop までを描画したコミュニケーションネットワーク図である。青点で表したものはアカウント、青線はリンクである。また、黄点のアカウントは日本とインドネシアを仲介するアカウントである。黄点のリンクは赤線で示す。また、図 3 左上のクラスターは、SoalCINTA を介して繋がっているインドネシアクラスターである。それ以外のクラスターは、nhk_seikatsu, nhk_kabun, touhokujishin, kopipedoujou など日本のクラスターである。黄色で示したアカウントは、媒介中心性順位と次数中心性順位を比較した場合、媒介中心性順位の方が高い傾向にある。

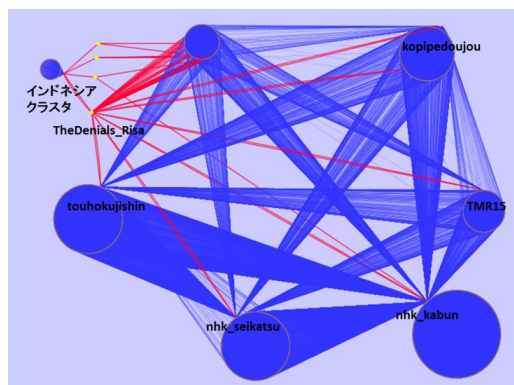


図 3. 仲介アカウントの一例

4. コミュニケーション形態に基づく分析

本章では、抽出された重要アカウントが、どのようなコミュニケーションをとっているか明らかにする。また、震災前後でどのような変化が起きているか明らかにする。

4.1 分析概要

重要アカウントのコミュニケーション形態を分類するために、震災前、震災後の 5 日間毎の投稿形態(ツイート)を基に分析を行う。アカウント毎に、総コミュニケーション数に占めるコミュニケーションの違いにより、コミュニケーション形態を分類する。分類するための指標として以下の 4 つの指標を用いる。

Out_Reply 率 = 自分発の総 Reply 数 / 総コミュニケーション数

Out_RT 率 = 自分発の総 RT 数 / 総コミュニケーション数

In_Reply 率 = 相手発の総 Reply 数 / 総コミュニケーション数

In_RT 率 = 相手発の総 RT 数 / 総コミュニケーション数

3.2 節で抽出を行った震災前後の重要アカウントのコミュニケーション形態の推移をみる為、K-means 法を用い分類を行う。分類期間は、震災前は 2011 年 3 月 6 日から 3 月 10 日までとし、震災後は 2011 年 3 月 11 日から 3 月 15 日までと区分する。震災前において次数中心性または媒介中心性が一日以上 100 位以上に該当するアカウント数は 315 である。震災後において、次数中心性または媒介中心性が一日以上 100 位以上に該当するアカウント数は 396 である。また、震災前後において重複しているアカウント数は 57 である。

4.2 分析結果

表 9 にコミュニケーション形態の分類結果を示す。Out_Reply 型は、Out_Reply 率が高く、自アカウントから他アカウントへ Reply をする割合が高いアカウントである。震災前では 29 アカウント(9%)が該当し、FlandreBot, reiuiziUtsuhoBot などである。震災後では、アカウントは確認されていない。震災前後において、9%から 0%と減少している。

Out_RT 型は、Out_RT 率が高く、他アカウントのツイートを RT する割合が高いアカウントである。震災前では 16 アカウント(5%)が該当し、habu_bot, nannsuba_bot などである。habu_bot は「はぶ」というキーワードが含まれているツイートを RT する bot である。震災後では 10 アカウント(3%)が該当し、Litha_Sacra, cuppekezuke など、日本語だけでなく日本語以外の言語をツイートするアカウントが含まれている。震災前後において、5%から 3%と減少している。

In_Reply 型は、In_Reply 率が高く、他アカウントから自アカウントへ Reply される割合が高いアカウントである。震災前では、67 アカウント(21%)が該当し、TMR15, ayu_19980408 などがある。震災後では 10 アカウント(3%)が該当し、ladygaga, justinbieber, mijeje などである。震災前後において、21%から 3%へ減少している。

In_RT 型は、In_RT 率が高く、他アカウントから RT される割合が高いアカウントである。震災前では、203 アカウント(64%)が該当し、sazae_f, asahi, NHK_PR などである。震災後では、殆どの 376 アカウント(96%)が該当し、Kantei_Saigai, hayano, asahi, NHK_PR などがある。震災前後において、63%から 96%に増加している。

表 9. 中心性が高いアカウントのコミュニケーション形態分類

	クラスタ(K=4, 母数 711 前 315 後 396)			
	Out_Reply 型	Out_RT 型	In_Reply 型	In_RT 型
該当 アカウント数 (%)	29 (4%)	26 (4%)	77 (11%)	579 (81%)
前 (%)	29 (9%)	16 (5%)	67 (21%)	203 (64%)
後 (%)	0 (0%)	10 (3%)	10 (3%)	376 (96%)
前アカウント	FlandreBot ReiuiziUtsuhoBot	habu_bot nannsuba_bot	TMR15 ayu_19980408	sazae_f asahi NHK_PR
後アカウント	---	Litha_Sacra cuppekezuke	ladygaga justinbieber mijeje	Kantei_Saigai hayano asahi NHK_PR

5. 結論

本研究では、次数中心性、媒介中心性を各アカウントに対して求めることで、重要なアカウントを発見することが出来た。その結果、震災前では、新聞社やニュースサイト、bot のアカウントが重要であったが、震災後では、公式アカウントや有名人、震災に関する情報を発信しているアカウントが重要であったことが特定できた。震災前では、重要アカウントとして、bot が多く抽出されたことから、楽しむためのツールとして利用されている側面が強いと考える。震災後では、重要アカウントとして、震災に関する情報発信を行っているアカウントが数多く抽出されたことから、情報収集のために利用されていると考える。

また、コミュニケーション形態の差異により、震災後では、震災前より、殆どの上位アカウントが RT されることにより、他アカウントと繋がっていることが明らかとなった。このことから、震災後では、リツイートが重要な情報流通手段となっていると考える。

参考文献

- Vieweg, S., L.Hughes, A., Starbird, K., and Palen, L., Microblogging during two natural hazards events: what twitter may contribute to situational awareness. In Proceedings of the 28th international Conference on Human Factors in Computing Systems, 1079-1088, 2010.
- Cheong, F., and Cheong, C., Social Media Data Mining: A Social Network Analysis of Tweets During The 2010-2011 Australian Floods. PACIS 2011 proceedings, 7-11, 2011.
- 鳥海不二夫, 篠田孝祐, 栗原聡, 榎剛史, 風間一洋, 野田五十樹, 震災がもたらしたソーシャルメディアの変化. JWEIN'11, 41-46, 2011.
- 風間一洋, 鳥海不二夫, 篠田孝祐, 榎剛史, "名詞出現頻度の時間的変化に着目した東日本大震災時の Twitter のトピック分析," WebDB forum, 2011.
- Freeman, L.C., A Set of Measures of Centrality Based on Betweenness. Sociometry, 35-41, 1977.
- Bader, D., Madduri, K., Guilbert, Kepner, J., Meuse, T., and Krishnamurthy, A., Scalable synthetic compact applications for benchmarking high productivity computing systems. CTWatch Quarterly, 41-51, 2006.