# ネットワーク QoE 制御のための大規模災害時における SNS による集合知に基づいた情報抽出

 丸
 千尋†
 榎
 美紀††
 中尾
 彰宏†††
 山本
 周†††
 山口
 実靖††††

 小口
 正人†

† お茶の水女子大学 〒 112-8610 東京都文京区大塚 2-1-1 †† 日本アイ・ビー・エム (株) 東京基礎研究所 〒 103-8510 東京都中央区日本橋箱崎町 19-21 ††† 東京大学 〒 113-8654 東京都文京区本郷 7-3-1 ††† 工学院大学 〒 163-8677 東京都新宿区西新宿 1-24-2

 $E-mail: \ \dagger \{chihiro,oguchi\} @ogl.is.ocha.ac.jp, \ \dagger \dagger enomiki @jp.ibm.com, \ \dagger \dagger \dagger \{nakao,shu\} @iii.u-tokyo.ac.jp, \ \dagger \dagger \dagger sane @cc.kogakuin.ac.jp$ 

あらまし 東日本大震災に見られるように、大規模災害時に被害状況の把握に必要な故障および輻輳制御状態に関する情報が膨大になると、ネットワーク全体の状況を迅速に把握することが困難となる。また、災害時には、電話やメールが使えないユーザが多いものの、SNS は利用可能な場合が多い。災害のような緊急時には、ユーザがネットワークの状況に強い関心を寄せ、その情報を SNS を通して積極的に発信すると考えられる。そのため、ソーシャルデータを集合知として用いることが、従来のネットワーク監視の補完的な情報取得手段として有効であると期待できる。本研究では、緊急時の自動的/自律的なネットワーク制御を実現するため、SNS の集合知から通信障害に関する情報を検出するシステムを提案する。具体的には、障害が発生している地域を検出するため、ソーシャルデータとして Twitter におけるユーザのツイートを用い、そこに出現する市区町村名に着目し、確率分布に基づいて解析する手法を提案した。東日本大震災時のツイートを用いて評価を行い、市区町村単位で通信障害を検出する提案手法が、より一層障害検出の精度を高めることを示した。さらに、ネットワークテストベッド上で、検出した情報を利用した経路制御を実現できることを示した。

キーワード Twitter, SNS, 集合知, 障害検知, 通信障害, ネットワーク制御

# 1. はじめに

2011年3月に発生した東日本大震災は、広域的に大きな被 害をもたらした. 被害は通信ネットワークにも及び、電話が つながりにくくなったり、電子メールを送受信しにくくなると いう事態が発生した [1-3]. 固定電話に関して, NTT 東日本, KDDI、ソフトバンクテレコムの3社で約190万回線が不通と なり、携帯電話基地局に関しては、NTT ドコモ、KDDI、ソ フトバンクモバイル, イー・モバイル, ウィルコムの5社で最 大 29,000 局が停止した。震災時は、通信ケーブルの切断や基 地局の損壊、停電によるネットワーク機器の停止により、電話 が利用できない地域が多く存在した。また、情報連絡や安否確 認が集中したことにより、通常時の50~60倍ものトラフィック が発生するなど、トラフィックが大幅に増加したため、ネット ワークを構成するネットワーク機器に大きな負担がかかってい た。そのため、ネットワーク機器の処理能力を超え輻輳状態が 発生し、広範囲で通信障害が発生した. 震災時の情報行動に関 する調査 [4] によると、被災地でも被災地以外でも、地震発生 当日に知りたかった情報について「電話の状況」と回答する人 の割合が高かったが、ほとんどの人がその情報を得ることがで きなかった.

従来、ネットワークの状態は、運用管理システムが、ネットワーク機器を用いてネットワーク内部から情報を取得することで監視されていた [5]. しかし、大震災が発生した際には、被害状況の把握に必要な故障および輻輳状態に関する情報が膨大になっており、ネットワーク機器からの情報のみでネットワーク全体の状態を迅速に把握することが困難であった。本研究では、前述の課題を解決するために、ネットワーク機器により取得されるネットワーク内部の情報を補完しうる、ネットワーク外部からの情報を利用する方法を提案する.

また、大震災では、Social Networking Service (SNS) が、情報伝達のツールとして一定の役割を果たしていた [6]. 例えば Twitter (日本) では、地震発生から一時間以内に東京だけで毎分約 1,200 件以上のツイートが投稿されており、情報のやり取りが Twitter 上で活発に行われていたことがわかる。 SNS は多くのユーザから情報をリアルタイムに得ることができるという特徴がある。 震災のような緊急時では、ユーザが通信ネットワークの状態に強い関心を寄せ、その情報を積極的に SNS を通して発信すると考えられる。 従って、 SNS を用いることで、通信障害が発生している地域や原因、ユーザへの影響の度合いなど、

(注1):http://twitter.com/

ネットワーク機器から得られる情報だけでは把握することができない情報を得ることができる。そのため、SNSによる集合知が、従来のネットワーク機器による監視の補完的な情報取得手段として有用であると考えられる。

本研究では、最も広く使われている SNS の一つである,Twitter による集合知に基づいたネットワーク制御を自動的/自律的に行う,ネットワーク障害検知システムを開発することを目的とする。このシステムは,緊急時に通信障害を自動的に検知するという点で,ユーザやネットワークマネージャを補佐する。ここで,インターネットが使えない状況で Twitter が利用可能であるのかといった問題が生じる。しかし,例え無線 LANが利用できなくても、3G ネットワークや LTE など,他のサー

本論文の貢献は以下の2つである.

(1) SNS に基づくネットワーク障害検知システムの設計と 試作を行うことで、市区町村単位で通信障害に関する情報を検 出できることを示した.

ビスが使える可能性がある。また、障害が発生していない地域

にいる人が、通信障害に関する情報を発信すると考えられる。

(2) SNS に基づく障害検知をネットワーク制御に統合することで、SNS による集合知に基づき、ネットワーク制御を自動的/自律的に行えることを示した.

本論文の構成は以下の通りである。2. 章で関連研究について述べ、3. 章で提案システムの概要を紹介する。4. 章で本稿で利用する実験データについて紹介する。5. 章で初期キーワードの決定と候補データ抽出処理について、それぞれ概要を説明する。そして、6. 章で地名分類処理、7. 章でユーザが通信障害と思う程度の決定について説明する。8. 章で提案システムから取得した情報を用いたネットワーク制御システムの概要を説明する。最後に、9. 章で本稿をまとめる。

# 2. 関連研究

SNS でのデータを解析することで、地震や土砂災害や火災 など、実世界で発生する出来事を検知する手法が数多く存在す る [7-18]. 榊ら [7] は Twitter のユーザをソーシャルセンサと することで, 地震を早期に発見し, 発生場所を推測する手法を 提案している。国土交通省国土技術政策総合研究所 [8] は災害 が発生する可能性のある地域の住民等のツイートを元に、土砂 災害の予兆や発生の早期把握を行っている。これらの研究は出 来事の発生の有無に焦点を当てており、大きな出来事が発生し た際のより詳細な情報の抽出はなされていないため、大規模災 害が原因で起こる通信障害を検出することを目標としている本 研究とは異なる。また、Rudra ら [11] は災害時に収集したツ イートを複数のカテゴリに分類し、それぞれを要約する手法を 提案している. この研究は状況認識に貢献する, 地域ごとの状 況や救援活動に関するツイートを抽出している点で本研究と類 似している. Varga ら [13] は災害時にユーザが投稿した問題ツ イートとそれに対応する答えのツイートを見つける手法を提案 している この研究は対処する必要のあるユーザの問題を検出 している点で本研究と類似している。 榊ら [14] はドライバが投 稿したツイートの本文とツイートからの位置情報を用いて、ド

ライバ周辺の道路状況の収集を行っている。この研究は収集した道路交通情報をドライバに提示して利用している点で本研究と類似している。水野ら [17] は災害時に、ユーザによってなされた大量のツイートの本文と位置情報から、災害状況を検出するシステムを提案している。この研究は出来事の発生の有無だけではなく、その出来事によって引き起こされた二次的な被害についても検出し、管理者がその被害情報を利用している点で本研究と類似している。しかし、本研究は通信障害の検出に特化しており、検出された情報を用いたネットワーク制御の自動化を目的としているため、これらの研究とは異なる。

従来、ネットワーク制御はネットワーク機器から取得した情 報を用いて行われてきた [5]. ITU-T Focus Group on Disaster Relief Systems [5] は緊急時に、ワイヤレスセンサネットワー クを利用した監視システムを用いて、ネットワーク障害を検知 し、警告レベルを超えた際にネットワーク管理者にその旨を自 動的に通知する。一方、本研究では SNS による集合知を用い たネットワーク制御を行うという点で新規性がある. ユーザに とってネットワークがどのような状態であるかを従来の手法で 正確に掴む事は難しいが、本研究の提案手法でこれを補完する ことが可能である. Qiu [18] らはモバイルネットワーク機能に 問題があった場合、ユーザはサービスセンタに電話するよりも 早く Twitter にその問題を投稿することを示した。これは通信 障害を検知する際に、Twitter を利用することが有用であるこ とを示している。竹下ら [19] はネットワークの故障状況を把握 するために、Twitter 上に投稿されたネットワークの問題に言 及したツイートを利用している点で本研究と動機が似ている。 しかし、竹下らは大規模災害を対象としておらず、平常時にお ける障害検知を行う所までに留まっている点で、抽出した情報 を基にネットワーク制御を行うシステムを構築した本研究とは 異なる。更に、竹下らは通信障害以外にも様々な種類のネット ワーク障害を同時に検出しているため、ユーザへの影響の度合 いなど、通信障害に関するより詳細な情報を検出している本研 究とは異なる。

#### 3. 提案システムの概要

本研究では、以下のネットワーク障害検知システムを提案する. 提案システムの概要を図1に示す. 本稿では、障害検知の対象とするネットワークを電話網とする.

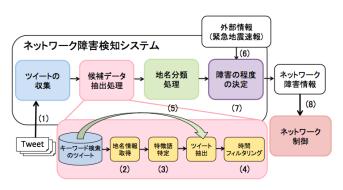


図 1 ネットワーク障害検知システム

図中(1)~(8)の動作は以下の通りである.

- (1) ブートストラップ法を用いた障害表現抽出によって,通信障害に関する初期キーワードを決定し,そのキーワードを 含むツイートを取得する.
- (2) (1) で取得したツイートを,同じ市区町村名が含まれるツイートごとにまとめる.
- (3) (2) の地名名詞ごとにまとめたツイートの中から特徴語を特定し、地名名詞を含まず特徴語を含むツイートを、(1)のツイートの中から抽出して加える。
- (4) 関係のないツイートを排除するために、ツイートされた時刻を考慮し、時間フィルタリングを行う。
- (5) 候補データ抽出処理によって抽出されたツイートに出現する地名を、「その地名で通信障害が発生しているのか」、「その他の場合であるのか」に分類し、前者のツイートを取得する。
- (6) 緊急地震速報からの情報を解析し、地震の発生時刻や 震源地やマグニチュードを取得する.
- (7) 平常時のツイート数の頻度の分布をもとに、緊急時のユーザが障害と思う程度を決定する.
- (8) 提案システムから取得した地域ごとの通信障害に関する情報をもとに、最適な経路を探索し、それを実際の通信経路に設定する.

Twitter から検出した通信障害に関する情報を出力することで、ネットワーク制御に直接連携するという目的を達成するために、ネットワーク障害検知の自動化システムを構築する。自動化にあたり、キーワード検索で取得したツイートのサンプリング間隔を1分とし、現在のツイートから60分遡った時刻までのツイートを障害検知対象のツイートとして扱う。このシステムにより得られた候補地ごとにまとめた障害情報が、出力データとして出力され、解析結果は1分以内に更新される。このシステムを用いることで、ユーザによるネットワークの状態に関する詳細な情報を獲得できる。本研究ではこの情報に基づき、8.章で述べる自動化されたネットワーク制御システムを構築する。

#### 4. 実験データの詳細

提案システムは大規模災害を対象としているため、2011 年に発生した東日本大震災、2016 年に発生した熊本地震前震、熊本地震本震、2014 年に北海道、茨城、長野で発生した地震の6つのツイートコーパスを利用する。本稿では、これらのコーパスをそれぞれ  $Dataset_{EastJapan}$ 、 $Dataset_{Kumamoto}$ 、 $Dataset_{Kumamoto}$ 、 $Dataset_{Nagano}$  と定義する。これらは、実験ごとに適宜利用される。コーパスの詳細を表1に、 $Dataset_{EastJapan}$  と $Dataset_{Kumamoto}$  に含まれるツイートの例を表2と3に示す。 $Dataset_{Hokkaido}$ 、 $Dataset_{Nagano}$  のツイート数は、5.2.1章で述べるキーワード検索後のツイート数を表してれるのコーパスは通信障害に関するツイート以外に、様々な種類のツイートを含んでいることがわかる。この中から、ネットワーク制御に有用なツイートのみを抽出することを目指す。

表 1 ツイートコーパスの詳細

データセット	発生日	発生時刻	ツイート数	最大震度	マグニチュード
$Dataset_{EastJapan}$	2011/03/11	14:46	8,815,519	7	9.0
$Dataset_{Kumamoto}$	2016/04/14	21:26	4,604,885	7	6.5
$Dataset_{Kumamoto2}$	2016/04/16	1:25	5,005,098	7	7.3
$Dataset_{Hokkaido}$	2014/07/08	18:05	184	5 強	5.8
$Dataset_{Ibaraki}$	2014/09/16	12:35	566	5 強	5.6
$Dataset_{Nagano}$	2014/11/22	22:08	808	6 強	6.8

表 2 東日本大震災時のツイートの例

- ・地震だ。結構揺れた。
- ・宮城県北部で震度7
- ・青森八戸津波注意してください!
- ・コンビニで食料調達しに行ったけど、ほとんど物がないです。
- ・ばあちゃんの安否が心配である。
- ・山手線は、地震の影響で、運転を見合わせています。
- ・盛岡市内は完全に停電しています。信号もついていません。
- ・岩手の母に電話がつながらない。

#### 表 3 熊本地震前震時のツイートの例

- ・熊本県で、震度七
- ·緊急地震速報! 熊本!
- ・熊本の叔父さん大丈夫だろうか…
- ・阿蘇山こわ
- ・熊本城の石垣崩れたとか……
- ・熊本駅近辺に人がけっこう溜まってるな。
- ・熊本市中心部、路面電車は止まってるけどクルマは普通に走ってる
- ・熊本市の親戚に電話したけど繋がらない。

#### 5. データ抽出の詳細

本章では、提案システムの一部である、 初期キーワードの決定、 候補データ抽出処理について概要を説明する.

#### 5.1 初期キーワードの決定

提案システムのキーワード検索で設定する初期キーワードを 決定するために、ブートストラップ法を用いた障害表現抽出を 行う [20]. ブートストラップ法を用いた障害表現抽出の流れを 図 2 に示す.

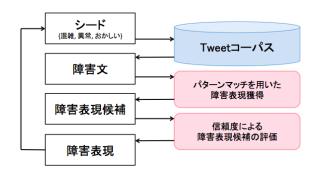


図 2 ブートストラップ法を用いた障害表現抽出

まず、Dataset<sub>EastJapan</sub> に対して初期シードを与え、初期シードを含むツイートを取得する。ツイートコーパスのツイートは、電話という単語を含む。そして取得したツイートを障害文と定義する。本稿では、通信障害に関するツイートを取得するため、混雑、異常、おかしいという単語を初期シードに設定

する. 次に、パターンマッチを用いて障害表現候補を獲得する. これは、障害文の中から、動詞の未然形に打ち消しの助動詞か 否定の助動詞が連続する組合せを、品詞のパターンマッチを用 いて獲得する. 組合せの例を以下に示す.

- (動詞の未然形) + ない
- (動詞の未然形) + ぬ
- (動詞の未然形) + ん
- (動詞の未然形) + ず

次に、ブートストラップ法では、一度適合率の低い結果が障害表現候補の出力の中にあると連鎖的に適合率が低下してしまうため、獲得した障害表現候補について信頼度のスコアリングを行う。障害表現候補 candidate の信頼度 score は、以下の式を用いて算出される。

$$score(candidate) = \sum_{s \in S} \frac{1}{distance(s)}$$
 (1)

S は障害表現候補 candidate が出現する障害文の集合, distance(s) は障害文での電話という単語と障害表現候補 candidate との距離である。この式は、障害文に多く出現し、電話という単語の近くに出現しやすい単語ほど、障害表現である可能性が高いという仮定に基づく。そして、上位 N%のものを障害表現とする。このようにして得られた障害表現を、次のステップの新しいシードとしてツイートコーパスに与え、同じ作業を行う。この作業を一定回数繰り返して得られた障害表現を、提案システムの初期キーワードに設定する。

ここで、初期キーワード決定の実験を行う。N を 1.0 から 10.0 まで 0.5 刻みで試した結果、今回は N を 2.5 に設定した。作業の繰り返し回数は 5 回とした。得られた全 18 個の障害表現のうち、人手で 10 単語を選択し、キーワード検索で設定する初期キーワードとした。初期キーワードの例を表 4 に示す。

繋がらない つながらない 通じない 繋がらん つながらん 使えない 繋がらず 繋がんない 出来ない つながんない

#### 5.2 候補データ抽出処理

本節では、候補データ抽出処理の概要について述べる.

#### 5.2.1 キーワード検索

Twitter 社の Search API<sup>(注2)</sup> を使って取得したツイートを, 5.1 節で決定した, 通信障害に関する初期キーワードを用いて キーワード検索を行う。キーワードは「電話 and 繋がらない」 などとし、リツイートは排除する.

#### 5.2.2 地名抽出

キーワード検索で取得したツイートの本文とユーザのプロファイルと GEO タグを  $MeCab^{(i\pm 3)}$  を使ってそれぞれ品詞分解し、地名名詞を抽出する。GEO タグの緯度・経度は、Yahoo! 社が提供する Yahoo! リバースジオコーダ Yahoo! を使って市

(注2):https://dev.twitter.com/rest/public/search

(注3):http://mecab.sourceforge.net/

(注4):  $\label{eq:http://developer.yahoo.co.jp/webapi/map/openlocalplatform/v1/reversegeocoder.html$ 

区町村に変換する。そして地名名詞ごとに、キーワード検索で 取得したツイートをまとめる。

#### 5.2.3 特徴語抽出

地名名詞ごとにまとめたツイートだけでは、同じ障害について言及したツイートであるが地名名詞を含んでいないツイートを全て捨ててしまうことになる。そのため、地名名詞ごとにまとめたツイートの中に出現する特徴的な単語を抽出し、抽出した特徴語を含み地名名詞を含まないツイートを、キーワード検索で取得したツイートの中から加える。

#### 5.2.4 時間フィルタリング

地名名詞ごとにまとめたツイートの中には、内容の異なるツイートが多く混じっている。そのため、ツイートされた時刻を考慮して関係のないツイートを排除することを考える。大規模災害が発生した際には、Twitter 上の複数のユーザが特定の時間に似たような内容のツイートをする [7]. 本研究では、その点に着目して、ツイートを排除する時間の閾値を決定する。閾値を決定するために、 $Dataset_{Hokkaido}$ 、 $Dataset_{Ibaraki}$ 、 $Dataset_{Nagano}$  を用いて、通信障害に言及するツイート数の時間変化を調べ、一般化する。

長野県で発生した地震について,通信障害に言及するツイート数の時間変化を図3の緑の棒グラフ,累積度数の値を赤線に示す.

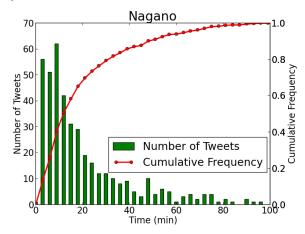


図 3 通信障害に言及するツイート数の時間変化

本研究ではサンプル数が少ないため、累積度数を考える。累積度数の時間変化を見ると、指数分布の累積分布関数に似た形になっている。この特徴は、北海道、茨城で発生した地震の場合にも当てはまる。そのため、累積度数の時間変化それぞれを、指数分布の累積分布関数にフィッティングする。指数分布の累積分布関数を以下に示す。

$$f(x) = 1 - e^{-\lambda x} \tag{2}$$

北海道, 茨城, 長野で発生した地震をフィッティングした結果を図4に示す. パラメータ  $\lambda$  はそれぞれ 0.07372231, 0.09200698, 0.0536034 とした.

図4より、どの場合も、累積度数の時間変化を指数分布の累積分布関数にフィッティングできることがわかる。指数分布であることから、閾値が60分で8割の事象を捉えることができ

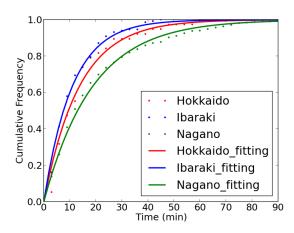


図 4 指数分布の累積分布関数にフィッティングした結果

るという結論が得られる。よって、障害検知のツイートを排除 する時間の閾値を 60 分に設定する。

#### 6. 地名分類処理

通信障害に関するツイートを抽出するために,5.1 節で説明した候補データ抽出処理によって得られたツイートの内容を見ていく。提案システムが電話が繋がらないという障害を取得した際には、抽出された地名に電話が繋がらないのか、それとも、その地名から他の地域に電話が繋がらないのか、ということを同時に判定する必要がある。候補データ抽出処理によって抽出されたツイートの例を図5に示す。

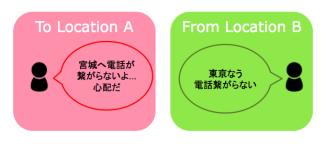


図 5 候補データ抽出処理のツイートの例

左のツイートは、抽出された地名 (宮城) に電話が繋がらないことを示している。一方、右のツイートは、抽出された地名 (東京) に電話が繋がらないのではなく、抽出された地名 (東京) から他の地域に電話が繋がらないことを示している。よって左のツイートと右のツイートは異なる種類のツイートであり、分類される必要がある。実際に障害が発生している場所は、Location A であると考えられるため、Location B ではなく Location A のツイートを取得する。これにより、通信障害が発生している場所に関する情報を含むツイートを抽出することができる。Location A のツイートを取得するために、機械学習を用いた手法を提案する [21]。

# 7. 市区町村ごとのユーザが通信障害と思う程度の決定

本章では, 市区町村ごとに, 電話の接続の可否だけではなく,

どのくらい困っている人がいるのか,ユーザが通信障害と思う 程度を決定する。この結果は、ネットワークの経路制御等で、 地域ごとに復旧の優先順位を決定する際に利用することがで きる

#### 7.1 市区町村ごとの確率分布の作成

本手法では、「平常時」と「非常時」の両方のソーシャルデータの解析を行う。平常時の確率分布を把握しておき、それをもとに非常時の障害の程度の判定を行う。ユーザが障害と思う程度を決定するために、平常時のツイートを利用して市区町村ごとに確率分布を作成する。図 6 に仙台市における確率分布を示す。ここで、 $\mu$  は平均、 $\sigma$  は標準偏差を表す。

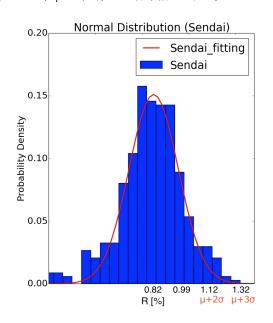


図 6 仙台市の確率分布

障害が何も発生しなかった 30 日分の全ツイートを、一時間単位で分割し、30\*24 個のファイルを作成する。そして、それぞれのファイルについて、市区町村ごとに、割合 R を求める。 R は以下のように求める。

図6の縦軸は、全ファイルにおける Rの出現頻度である。図6を見ると、仙台では、平常時、地名情報を含むツイートのうち、仙台に関するツイートが 0.82%の割合で出現することが多いことことがわかる。このように、平常時のツイートを用いて市区町村ごとに確率分布を作成する。そして、それぞれを正規分布にフィッティングする。正規分布の確率密度関数を以下に示す。

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx$$
 (4)

ここで、 $\mu$  は平均、 $\sigma$  は分散を表す。

次に、作成した正規分布から、 $\mu+2\sigma$  と  $\mu+3\sigma$  を求める。正規分布では、図 7 のように、 $\mu+2\sigma$  までで全体の事象の 95.45%、 $\mu+3\sigma$  までで 99.73%をカバーしているため、R がこれらの値

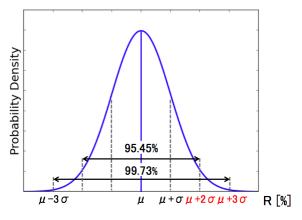


図7 正規分布

を超えることはほとんどない。よって,ある市区町村において,R が  $\mu+2\sigma$  か  $\mu+3\sigma$  を超えたら,この市区町村でツイートが普段よりも多くなされているため,その市区町村で障害が起きていると考えられる。地震発生後 60 分間の,6. 章の地名分類処理によって抽出された通信障害に関するツイートを用いて,市区町村ごとに  $R_{Emergency}$  を求める。そして,この $R_{Emergency}$  と平常時の正規分布から求めた  $\mu+2\sigma$ , $\mu+3\sigma$  と比較し, $R_{Emergency}$  がこれらの値のどちらかを超えたら,この市区町村は異常であると判定される。

図 6 の赤線は正規分布にフィッティングした線である。仙台では, $R_{Emergency}$  が  $\mu+3\sigma$  である 1.32%,もしくは  $\mu+2\sigma$  である 1.12%を超えたら,異常であると判定される。東日本大震災時では,仙台市の  $R_{Emergency}$  は 9.57%であり,この値はどちらの値よりも大きいため,大震災時,ユーザにとって仙台には非常に電話が繋がりにくい状況であったと判定される。

#### 7.2 評価実験

 $Dataset_{EastJapan}$ ,  $Dataset_{Kumamoto}$ ,  $Dataset_{Kumamoto2}$  の地震発生後 60 分間の通信障害に関するツイートを用いて、評価実験を行う。東日本大震災の結果を図 8(a)、震度図を図 8(b)、熊本地震前震の結果 9(a)、震度図を 9(b) に示す。赤色のピンは  $R_{Emergency}$  が  $\mu$ +3 $\sigma$  を超えた市区町村、黄色のピンは  $R_{Emergency}$  が  $\mu$ +2 $\sigma$  を超えた市区町村をそれぞれ表している。

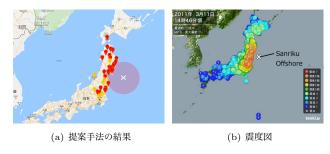


図8 東日本大震災時の結果

結果と震度図を比較した結果,東日本大震災,熊本地震前震の場合,被害が大きかった地域は通信障害の程度も大きいと判定されていることが分かる.しかし,熊本地震本震の場合は,市区町村を抽出することができなかった.これは,数日前に大きな前震があり,既に通信障害に関する対策が取られていたと



図 9 熊本地震前震時の結果

考えられることと, 地震発生時刻が他の地震と比べて遅いことが原因だと考えられる.

さらに、東日本大震災の結果について、より細かい市区町村単位で評価する.総務省が発行した東日本大震災における情報通信の状況 [2] に記載されている電話サービス不通地域と、先程の結果を比較する. 岩手県の市区町村を比較した結果を図10(a)、宮城県の市区町村を比較した結果を図10(b)に示す. 赤は提案手法もレポートも電話サービス不通地域と判定した市区町村、緑は提案手法は電話サービス不通地域と判定したが、レポートには記載されていない市区町村を表している.

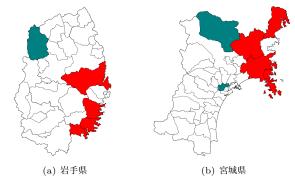


図 10 市区町村単位で評価した結果

結果より、提案手法が検出した11市区町村のうち、8市区町村が正確に検出されたことを確認できた。3市区町村はレポートには記載されていなかったが、提案手法は高い精度で検出することができていると評価できる。提案手法が検出することのできなかった市区町村に関しては、そもそもツイートにその市区町村名が出現していなかったことが原因であり、ツイートを利用する本手法の適用外の地域である。これは、Twitter 以外にネットワーク機器等、様々な種類の情報源を利用することで解決することができると考えられる。

# 8. SNS解析に基づくネットワーク制御システム

自動化された提案システムから抽出されたユーザによるネットワークの状態に関する情報をもとに、トラフィックの最適化を自動的/自律的に行うネットワーク制御システムを構築する。本研究では、経路制御を行うために、FLARE [22,23] と呼ばれるネットワークシステムを利用する。そして、FLAREによるネットワーク制御システムを、広域テストベッドネットワークである Japan Gigabit Network eXtreme (JGN-X) (注5) 上に

(注5):http://www.jgn.nict.go.jp/english/info/network.html

実装する.

#### 8.1 DPN/FLARE

近年、プログラミングによって自由にネットワークを制御する技術である Software Defined Network (SDN) が注目されている。OpenFlow [24] は SDN を実現する技術の 1 つである。ネットワーク上でデータを適切に送るためには、ネットワーク制御を行うコントロールプレーンと、コントロールプレーンの制御に従ってパケットを転送するデータプレーンが必要である。従来、これらはネットワーク機器ごとに組み込まれており、ユーザが機能を拡張することはできなかった。しかし OpenFlow では、これらが分離され、コントロールプレーンがプログラム可能になった。これにより、ユーザは、ハードウェアに直接アクセスすることなく、自らプログラムを組んで自由にネットワークトラフィックを制御することができる。

一方、コントロールプレーンのみならずデータプレーンもプログラム可能である、SDN を発展させた Deeply Programmable Network (DPN) が提唱されている。FLARE は DPN を実現した技術の 1 つである。OpenFlow がネットワーク層までのデータを扱うのに対し、FLARE はアプリケーション層までのデータを扱うことができる。これによりデータプレーンまでプログラムによる制御の対象となり、トラフィックからアプリケーションを識別すれば、アプリケーションの種類に基づく制御を行うことが可能となる。本研究では SNS 情報に基づいた高度で柔軟なネットワーク制御を目指しているため、アプリケーション層のデータまで制御に用いることができる FLARE は最適なプラットフォームであると言える。

#### 8.2 JGN-X

JGN-X とは、情報通信研究機構 (NICT) (注6) が運用している大規模な研究開発用テストベッドネットワークの名称である.このテストベッドを用いることで、新世代ネットワーク技術を実利用に近い環境で実装することができる。JGN-X 上には国内 25 か所と海外 5 か所にアクセスポイントが設置されており、これにより国内外の広域ネットワークが構築される。FLAREスイッチは、JGN-X 上の国内 8 か所に設置されている。

# 8.3 実験環境

本研究では、図 11 に示すネットワーク構成で実験を行う. FLARE Central は、FLARE 管理用のサーバである。このサーバ上に、経路制御を集中管理するコントローラを置く。コントローラで 4 台の FLARE スイッチを制御し、スイッチは指示に従ってパケットの転送を行う。我々は、ネットワーク障害検知システムを FLARE Central 上に実装する。そして、FLARE Central 上で、ソーシャルデータの獲得と解析を行い、解析結果に基づいてコントローラが FLARE スイッチそれぞれに制御指示を与える。

#### 8.4 SNS の情報に基づく経路制御

本節では、ネットワーク制御システムの実行例を示す.実験条件として、 $Dataset_{EastJapan}$ を再生し、経路制御を行うプログラムを JGN-X 上で検証する.図 11 の FLARE によるネッ

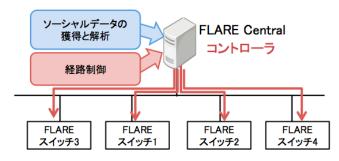


図 11 FLARE ネットワーク構成

トワーク制御システムを図 12 のように JGN-X 上に実装する. そして、ネットワーク障害検知システムから検出された地域ごとの情報をトリガとして、経路制御を行う制御プログラムを実行する. 具体的には、FLARE スイッチ間の経路のコスト値を、7. 章で算出された結果をもとに 1 分間隔で更新する. 最初、全ての経路のコスト値は共通のデフォルト値に設定されている. そして、ダイクストラ法を用いて最も小さいコスト値を持つ最適経路が検索され、その経路が実際の通信経路に設定される. 東日本大震災時のデータに基づく経路制御の実行例を図 12 に示す.



(a) 地震発生前



(b) 地震発生後

図 12 経路制御の結果

赤線は2点間の実際の通信経路を示している。地震発生後、ネットワーク障害検知システムは、電話が繋がりにくい地点を検出しており、図12(a)の赤線の通信経路は、この地点の近くを通っているため、輻輳していると考えられる。そのため、この経路のコスト値が大きくなり、候補となる経路の中で、図12(b)の赤線の経路のコスト値が最小と判断される。そして、図12(a)から図12(b)のように、経路が切り替わった。これに

より、提案システムを利用した自動的な経路制御を実現できたことがわかる。

# 9. おわりに

本稿では、Twitter による集合知を利用した自動的/自律的なネットワーク制御を行うため、緊急時に通信障害を検知するネットワーク障害検知システムの構築を行った。

まず、細かい地域の単位で通信障害に関する詳細な情報を抽出するために、ツイートに出現する市区町村名に注目し、ソーシャルデータを確率分布に基づいて解析することで、正確な情報を得ることを試みた。実際の地震が発生した際のデータを用いて評価実験を行い、結果と実際の震度図と照らし合わせると、地震の被害が大きかった地域は通信障害の程度も大きいと判定されていることがわかった。さらに、東日本大震災時の結果については、市区町村単位で評価を行うことで、提案手法は細かい単位でも高い精度で検出できていることがわかった。

次に、ネットワーク障害検知システムから検出された情報を用いた経路制御をテストベッド上で検証した。東日本大震災時の蓄積データを流し、経路間のコスト値を提案システムからの結果をもとに更新し、ダイクストラ法を用いて探索した最適経路を実際の通信経路に設定した。これにより、SNSの集合知に基づいた自動的な経路制御を実現できたことがわかる。

今後の課題として、電話以外の通信手段の通信障害を検知することが考えられる。通信手段によってツイートの内容が異なると考えられるため、それぞれに合う手法を検討する必要がある。ネットワーク制御では、通信経路を流れるトラフィックからアプリケーションを判別し、アプリケーション毎に経路を切り替えるといったきめ細やかな経路制御を実現することを検討している。

#### 文 献

- [1] 香川康介, 久野友也, 田村宏直, 高田 久, 古谷雅典, 南方伸哉, " 大規模災害時におけるオペレーションシステムの信頼性向上," NTT DOCOMO テクニカル・ジャーナル, vol. 20, no. 4, pp. 26-36, 2013.
- [2] 総務省, "東日本大震災における情報通信の状況," 平成 23 年版情報通信白書, http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h23/pdf/n0010000.pdf, 2011.
- [3] 総務省 総合通信基盤局 電気通信技術システム課,"東日本大震 災発生後の通信状況に関するアンケート,"http://www.soumu. go.jp/main\_content/000136157.pdf
- [4] 橋元良明,中村 功,関谷直也,小笠原盛浩,山本太郎,千葉直子, 関 良明,高橋克巳,"被災地住民の震災時情報行動と通信不安," 情報学研究調査研究編, no. 28. 2012.
- [5] ITU-T Focus Group on Disaster Relief Systems, "Monitoring systems for outside plant facilities," ITU-T Recommendations, no. L.81, pp. 1-10, 2009.
- [6] 吉次由美, "東日本大震災見る大災害時のソーシャルメディアの役割," 放送研究と調査 2011 年 6 月号, https://www.nhk.or.jp/bunken/summary/research/ report/2011\_07/20110702.pdf, 2011.
- [7] T. Sakaki, M. Okazaki, and Y. Matsuo. "Earthquake shakes Twitter users: real-time event detection by social sensors," Proceedings of the 19th International Conference on World Wide Web, pp. 851-860, 2010.
- [8] 国土交通省 国土技術政策総合研究所,"ソーシャルメディア情報 分析による土砂災害の予兆・発生情報検知に向けた取り組み," 2014.

- [9] 斎藤翔太, 伊川洋平, 鈴木秀幸. "Twitter を用いた災害情報の早期発見 (言語理解とコミュニケーション)," 電子情報通信学会技術研究報告 信学技報, vol. 114, no. 81, pp. 7-12, 2014.
- [10] S. Verma, S. Vieweg, W. Corvey, L. Palen, J. Martin, M. Palmer, A. Schram, and K. Anderson, "Natural Language Processing to the Rescue?: Extracting "Situational Awareness" Tweets During Mass Emergency," ICWSM, 2011.
- [11] K. Rudra, S. Ghosh, N. Ganguly, P. Goyal, and S. Ghosh, "Extracting Situational Information from Microblogs during Disaster Events: a Classification-Summarization Approach," Proceedings of the 24th ACM International Conference on Information and Knowledge Management, pp. 583-592, 2015.
- [12] S. Panem, M. Gupta, and V. Varma, "Structured Information Extraction from Natural Disaster Events on Twitter," Proceedings of the 5th ACM International Workshop on Web-scale Knowledge Representation Retrieval&Reasoning, pp. 1-8, 2014.
- [13] I. Varga, M. Sano, K. Torisawa, C. Hashimoto, K. Ohtake, T. Kawai, J. Oh, and S. Saeger, "Aid is Out There: Looking for Help from Tweets during a Large Scale Disaster," ACL, pp. 1619-1629, 2013.
- [14] 榊 剛史,柳原 正,那和一成,松尾 豊,"Twitter を用いた道 路交通情報の抽出,"電子情報通信学会論文誌 D, vol. J98-D, no. 6, pp. 1019-1032, 2015.
- [15] M. Cameron, R. Power, B. Robinson, and J. Yin, "Emergency Situation Awareness from Twitter for Crisis Management," Proceedings of the 21st International Conference Companion on World Wide Web, pp. 695-698, 2012.
- [16] Y. Qu, C. Huang, P. Zhang, and J. Zhang, "Microblogging after a Major Disaster in China: a Case Study of the 2010 Yushu Earthquake," Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work, pp. 25-34, 2011.
- [17] 水野淳太,後藤 淳,大竹清敬,川田拓也,鳥澤健太郎,クロエツェージュリアン,田仲正弘,橋本 力,奥村明俊,"対災害情報分析システム DISAANA 及びその質問応答モードの性能評価,"情報処理学会 コンシューマ・デバイス&システム, vol. 2015-CDS-14, no. 14, pp. 1-13, 2015.
- [18] T. Qiu, J. Feng, Z. Ge, J. Wang, J. Xu, and J. Yates, "Listen to me if you can: tracking user experience of mobile network on social media," Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement, pp. 288-293, 2010.
- [19] K. Takeshita, M. Yokota, and K. Nishimatsu, "Early Network Failure Detection System by Analyzing Twitter Data," IFIP/IEEE International Symposium on, pp. 279-286, 2015.
- [20] 栗原光平, 嶋田和孝, "ブートストラップ法を用いた Twitter から の不具合文抽出,"言語処理学会第 21 回年次大会, pp. 341-344, 2015.
- [21] C. Maru, M. Enoki, A. Nakao, S. Yamamoto, S. Yamaguchi, and M. Oguchi, "Development of Failure Detection System for Network Control using Collective Intelligence of Social Networking Service in Large-Scale Disaster," In Proc. the 27th ACM Conference on Hypertext and Social Media (HT2016), pp. 267-272, 2016.
- [22] A. Nakao, "Software-Defined Data Plane Enhancing SDN and NFV," Special Section on Quality of Diversifying Communication Networks and Services, IEICE Transactions on Communications, vol. E98-B, no. 1, pp. 12-19, 2015.
- [23] A. Nakao, "FLARE: Open Deeply Programmable Network Node Architecture," Stanford Univ. Networking Seminar, http://netseminar.stanford.edu/10\_18\_12.html, 2012.
- [24] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Lexford, S. Shenker, and J. Turner, "Open-Flow: enabling innovation in campus networks." ACM SIG-COMM Computer Communication Review, vol.38, no.2, pp.69-74, 2008.