2K-02

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

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 　‡IBM Research - Tokyo 　§東京大学 　¶工学院大学

1. はじめに

ネットワークの状態は，一般に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る [1]．しかし，東日本大震災が発生した際には，被害状況の把握に必要な通信回線の混雑および故障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は，ネットワーク全体の状況を迅速に把握することが困難であった [2]．また，震災時には，電話やメールが使えないユーザが多い中，SNS は利用可能な場合が多い [3]．震災のような緊急時では，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，我々は SNS による集合知が，従来のネットワーク監視の相互補完的な情報取得手段として有効であると考える．本研究では，緊急時に自動的／自律的なネットワーク制御を実現するため，Twitter の集合知から通信障害に関する情報を市区町村単位で検出することを目指す．

2. 提案システムの概要

本研究では，ネットワーク障害検知システムを提案する．本稿では，通信障害の検知の対象を電話とする．提案システムの動作は以下の通りである．本稿では，(7) について詳しく説明する．その他は [4] 等で詳細に議論している．

(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) の市区町村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確率分布をもとに，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名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．

Information Retrieval using Collective Intelligence of SNS for QoE Control of Network in Large-Scale Disasters
† Chihiro Maru, ‡ Miki Enoki, § Akihiro Nakao, § Shu Yamamoto, ¶ Saneyasu Yamaguchi, † Masato Oguchi
Ochanomizu University (†), IBM Research - Tokyo (‡), University of Tokyo (§), Kogakuin University (¶)

この障害検知は通信障害を即座に検知できることが重要であるので，ネットワーク障害検知システムのリアルタイム処理を行う．リアルタイム化にあたり，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．

3. 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，ユーザが通信障害と思う程度を決定する．この結果は，緊急時にネットワークの経路制御を行う際に，地域ごとの復旧の優先順位を決定するために利用できる．

3.1 市区町村ごとの確率分布の作成

本研究では，「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．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 1 に仙台市における確率分布を示す．

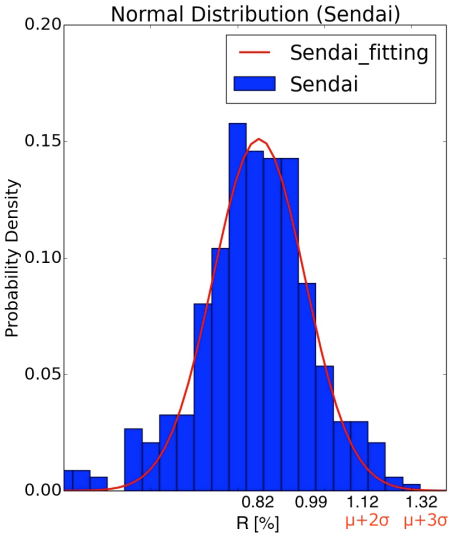


図 1: 仙台市における確率分布

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．

$$R = \frac{\text{ある市区町村に関する情報を含むツイートの総数}}{\text{地名情報を含むツイートの総数}} \quad (1)$$

図 1 の縦軸は，全ファイルにおける R の出現頻度である．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．図 1 の赤線は正規分布にフィッティングした結果である．正規分布の確率密度関数を以下に示す．

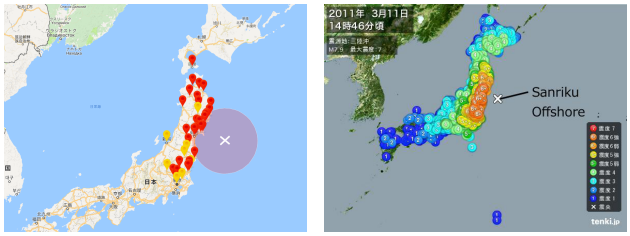
Copyright ©2017 Information Processing Society of Japan.
All Rights Reserved.

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

ここで，μ は平均，σ は分散を表す．次に，作成した正規分布から，μ+2σ と μ+3σ を求める．図 1 の仙台市の場合，μ+2σ は 1.12%，μ+3σ は 1.32%である．正規分布では，μ+2σ までで全体の事象の 95.45%，μ+3σ までで 99.73%をカバーしているため，R がこれらの値をほとんど超えることはない．よって，ある市区町村において，R が μ+2σ か μ+3σ 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，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 μ+2σ，μ+3σ 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．

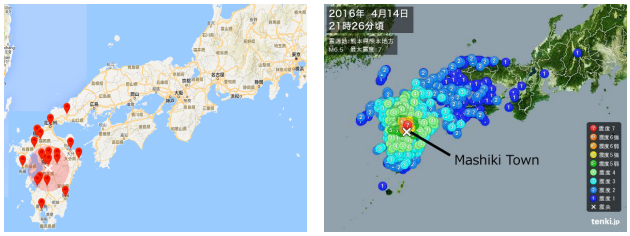
3.2 評価実験

東日本大震災時の 8,815,519 件のツイート，2016 年に発生した熊本地震時の 4,801,855 件のツイートを用いて評価実験を行った．本稿では，それぞれのコーパス中の地名分類処理によって抽出された，地震発生後 60 分間の通信障害に関するツイートを用いた．東日本大震災の結果を図 2(a)，震度図を図 2(b)，熊本地震の結果を 3(a)，震度図を 3(b) に示す．赤色のピンは $R_{Emergency}$ が μ+3σ を超えた市区町村，黄色のピンは $R_{Emergency}$ が μ+2σ を超えた市区町村をそれぞれ表している．



(a) 提案手法の結果　　(b) 震度図

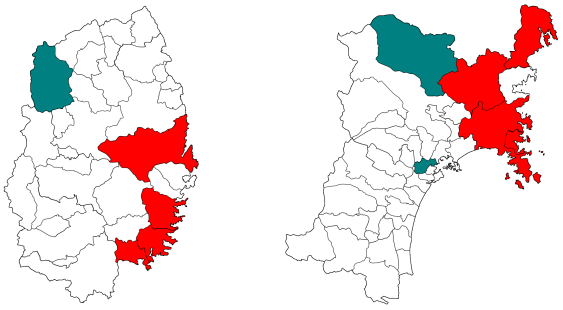
図 2: 東日本大震災時の通信障害の程度の結果



(a) 提案手法の結果　　(b) 震度図

図 3: 熊本地震時の通信障害の程度の結果

結果と震度図を比較した結果，どちらの地震の場合も，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．また，東日本大震災の結果について，更に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 4(a)，宮城県の市区町村を比較した結果を図 4(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．



(a) 岩手県の市区町村　　(b) 宮城県の市区町村

図 4: 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．

4. おわりに

本稿では，Twitter の集合知から通信障害に関する情報を市区町村単位で検出する，ネットワーク障害検知システムを構築した．東日本大震災，熊本地震時のツイートを用いて評価を行い，提案手法が，市区町村ごとに高い精度で通信障害を検出できることを確認した．

今度の課題として，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．また，現在，提案システムから検出された地域ごとの情報を利用して，通信経路を自動的に切り替えることが可能となっている．今後は，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．

謝辞

本研究は一部，総務省戦略的情報通信研究開発推進事業（SCOPE）先進的通信アプリケーション開発推進型研究開発によるものである．

参考文献

[1] ITU-T Forcus Group on Disaster Relief Systems, ”Monitoring Systems for Outside Plant Facilities,” ITU-T Recommendations, no.L.81, pp.1-10, 2009.

[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, 2011.

[3] 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, 2011.

[4] C. Maru et al., ”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.

Copyright ©2017 Information Processing Society of Japan.
All Rights Reserved.