

インターネット障害の把握の在り方に係る調査研究報告

株式会社NTTデータ

# 目次

1. 調査概要	… P3
2. 調査1：インターネット・文献による調査	… P6
3. 調査2：ヒアリングおよびアンケート調査	… P11
4. 調査3：SNSを使用した実証調査	… P17
5. 総括	… P54

# 1. 調査概要

## 1.1. 調査の目的

本調査では、インターネット障害の現状と今後の把握の在り方について検証しました。

### 調査の背景

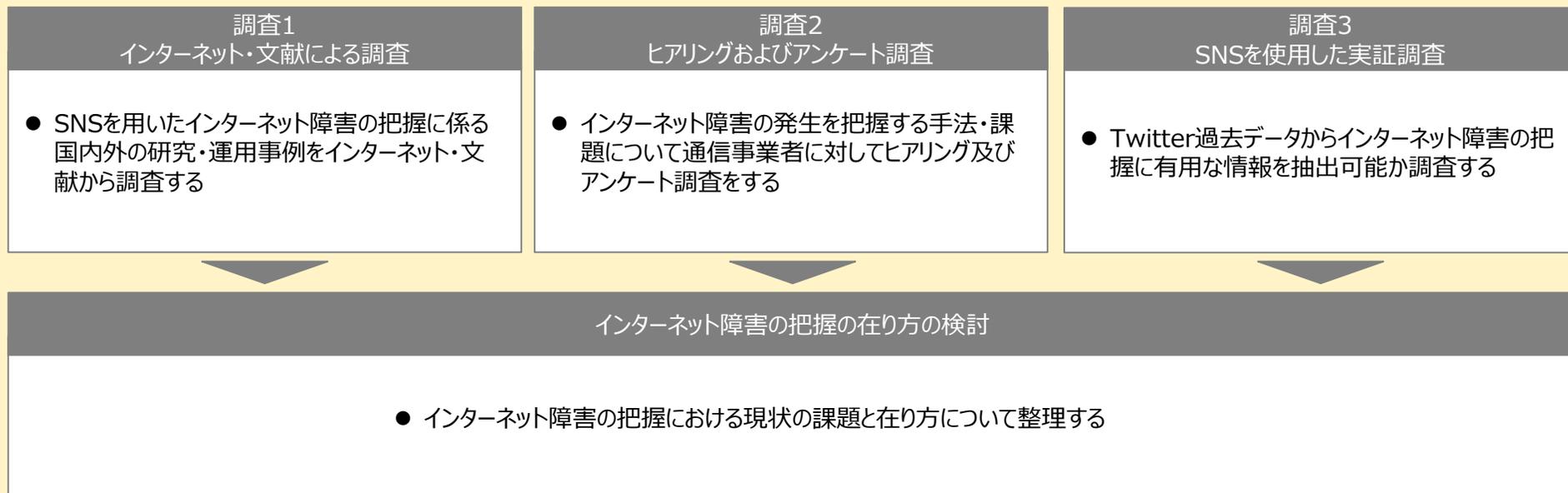
- ✓ 役務の停止に至らずとも、インターネットに接続しづらまたは特定のサービスやサイトにアクセスしづらいなど、通信速度やサーバ処理が著しく低下する事象が発生した場合、多くの利用者の混乱を招く事例が発生している
- ✓ このような事象を把握するには通信事業者に寄せられる問い合わせ内容から推測する手法が考えられるが、利用者の感覚的なとらえ方に依存する部分も多く、迅速かつ的確な事象の把握が難しい

### 調査の目的

- ✓ インターネット障害を把握するための現状の手法について整理する
- ✓ インターネット障害の発生状況、影響を迅速かつ的確に把握する手法として、SNSを用いたインターネット障害の把握の可能性について検討する

## 1.2. 調査実施方針

本調査では、インターネット・文献調査、通信事業者へのヒアリングおよびアンケート、SNSを使用した実証調査の3項目を実施してインターネット障害把握の在り方を検証しました。



## 2. 調査1：インターネット・文献による調査

## 2.1. 調査方法

本調査では、SNSを用いたインターネット障害の把握に係る国内外の研究・運用事例の概要と特徴をインターネット・文献から調査しました。

### 調査対象

- 2008年～2019年までに発表されたインターネット障害、通信障害に関するもの  
※Twitterの日本語版サービスが開始した2008年以降を対象とした

### 文献数

- 国内の研究等に関する文献：5件
- 海外の研究等に関する文献：4件（うち1件は参考文献としてインターネット障害以外に関する文献を調査）

## 2.2. 調査結果 -対象文献一覧

本調査で対象とした文献を以下に記載します。

文献No.	文献名	出典・著者
1	Twitter解析による通信品質低下傾向の早期検出手法の提案	FIT2012 第11回情報科学技術フォーラム 池田和史/服部元/小野智弘/麻生英樹
2	Syslog+SNS分析によるネットワーク故障検知・原因分析	NTT技術ジャーナル(2013.7) NTTネットワーク基盤技術研究所 木村達明/竹下恵/豊野剛/横田正裕/西松研/森達哉
3	国内文献 大規模災害時におけるSNSによる集合知に基づいたネットワークのQoE制御	マルチメディア, 分散協調とモバイルシンポジウム2016論文集 (2016) 丸千尋/榎美紀/中尾彰宏/山本周/山口実靖/小口正人
4	機械学習を用いたツイート解析と統計的異常検知による通信障害検出システム	研究報告コンシューマ・デバイス&システム (2019-CDS-25) 山田尚志/落合桂一/横井靖弘/神山剛/鳥居大祐
5	通信状況の可視化によるネットワーク障害の早期検出・分析	FIT2019 第18回情報科学技術フォーラム 桂康洋/内田真人
6	Measuring Online Service Availability Using Twitter	WOSN'10: Proceedings of the 3rd Wonference on Online social networks (2010) Marti Motoyama/Brendan Meeder/Kirill Levchenko/ Geoffrey M.Voelker/Stefan Savage
7	海外文献 Listen to Me if You can: Tracking User Experience of Mobile Network on Social Media	IMC'10 2010 Tongqing Qiu/Junlan Feng/Zihui Ge/Jia Wang/Jun (Jim) Xu/ Jennifer Yates
8	Uncovering Wireless Blackspots using Twitter Data	ELECTRONICS LETTERS Vol.00 No.00(2017) Weisi Guo, Jie Zhang
9 ※	SONAR: Automatic Detection of Cyber Security Events over theTwitter Stream	Ares'17(2017) Sceller, Quentin Karbab/Elmouatez/Debbabi/Mourad Iqbal/Farkhund

※No.9はインターネット障害に関する文献ではないが、SNSを活用してサイバー事件を検知する仕組みが通信障害把握手法の参考になるため参考文献として記載

## 2.3. インターネット・文献による調査のまとめ

インターネット障害の早期検知、サイレント障害の検知に係るSNSを活用した国内外の文献を調査し、障害検知に関して述べられている背景、目的、手法、検証結果、課題と今後の方向性についてまとめます。

### SNSを活用する背景

- SNS（特にTwitter）を用いる理由として、ユーザが体感したサービス状況や把握した問題を**リアルタイムに投稿**する傾向があること、特にモバイルサービスにおいてはその傾向が強いことが挙げられています。また、ネットワーク機器の情報だけでは把握できない、**障害地域、原因、ユーザ影響などの情報についてもツイートから得ることができる**ことも挙げられています。

### SNSを活用する目的

- 各文献におけるインターネット障害を検知する目的は、ネットワーク機器の故障予兆検知、5G到来を前に通信環境の悪いエリアの発見・改善、大規模災害時にネットワーク障害が発生した際の被害地域特定と対応優先度決めなど多岐にわたっていました。

### 障害検知手法

- 障害検知の手法についてはどの文献においても共通している部分が多く、また、2010年代前半の文献と2019年の最新の文献においても大枠は変わっておらず、**共通的な検知手法が存在する**ことが分かりました。（図1）

#### 共通的な部分

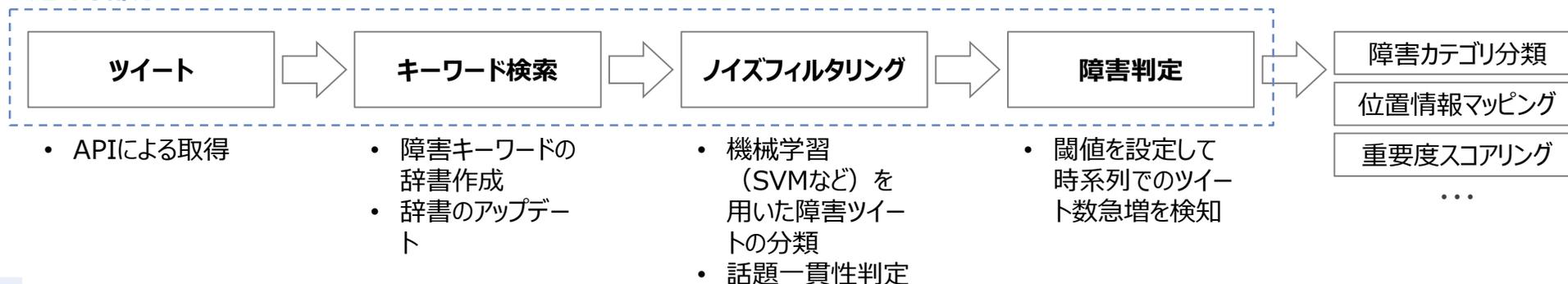


図1 Twitterを用いた標準的なインターネット障害検知手法

## 2.3. インターネット・文献による調査のまとめ

### 有用性の検証結果

- 障害検知の再現率、適合率は文献によって異なりますが、**利用者影響の大きい大規模な障害ほど早期に検知できる傾向**があることが分かりました。一方で、**中小規模の障害では検知まで時間がかかる（数十分～数時間）、または検知できないケース**があることも分かりました。
- 障害発生時に**カスタマーセンターへの問い合わせよりもTwitterの投稿が早い**という検証結果もあり、Twitterを使った早期検知の有用性が示されています。
- Twitterからの障害検知情報と**外部データ（企業データ、震度データなど）との組み合わせ**により、障害エリアの解像度を高めたり、障害の重要度を判定するなどTwitterデータの有用性を高められることも示されています。

### 現状の課題と今後の方向性

- 複数の文献において**障害の誤検知**が課題に挙げられていました。誤検知の要因として、ボットによる機械的な投稿、ユーザ実体験以外の投稿（ニュースへの反応など）、キーワード条件は一致するが無関係の投稿などが挙げられています。機械学習によるフィルタリング、自然言語処理の精度向上などが解決の鍵として挙げられています。
- 障害発生直後に**「通信障害発生に早期に気づくことができる」、「通信障害発生初期に通信障害状況を把握できる」ツイートが少量存在**することが確認されており、これらを検知することで中小規模の通信障害においても検知精度を上げることができると考えられます。

### 3. 調査2：ヒアリングおよびアンケート調査

## 3.1. アンケート – 調査方法

本調査では、まずインターネット障害の把握に係る事例、課題等を調査するために通信事業者へのアンケートを実施しました。

### アンケート対象

- 下記の電気通信4団体に加盟している通信事業者
  - 一般社団法人電気通信事業者協会
  - 一般社団法人テレコムサービス協会
  - 一般社団法人日本インターネットプロバイダー協会
  - 一般社団法人日本ケーブルテレビ連盟

### アンケート実施方法

- 電気通信4団体を通してエクセル形式のアンケート用紙を各通信事業者へ配布、回答記入後にメールにて回収

## 3.1. アンケート – 調査方法

アンケート調査では、インターネット障害に関する過去事例、課題、SNS利用状況について調査を実施しました。質問項目は下記の通りです。

大項目	小項目
1. インターネット障害の過去事例に関する調査	<ul style="list-style-type: none"><li>1. 障害内容</li><li>2. 発生日時</li><li>3. 把握日時</li><li>4. 把握手法</li><li>5. 影響（復旧までの時間、影響利用者数等）</li></ul>
2. インターネット障害の把握に関する課題調査	<ul style="list-style-type: none"><li>1. 把握までの時間に関する課題</li><li>2. 取得したいが出来ていない情報</li><li>3. その他の課題</li><li>4. 課題に対する取組み状況</li></ul>
3. インターネット障害把握のSNS利用に関する調査	<ul style="list-style-type: none"><li>1. 障害把握におけるSNSの利用状況</li><li>2. 利用しているSNSの種類と有用性</li><li>3. SNSを利用する上での課題</li><li>4. SNSを利用しない理由</li><li>5. SNSの今後の利用予定と種類</li></ul>

※「1. インターネット障害の過去事例に関する調査」では1事業者あたり最大3事例まで回答可とした

## 3.1. アンケート – 調査方法

アンケートに回答した事業者数と、回答があったインターネット障害の件数は以下の通りです。

アンケートに回答した事業者数	70 事業者
インターネット障害の件数 (うち影響利用者数が3万人以上の件数)	131 件 (9 件)

## 3.2. ヒアリング –調査方法

アンケート結果から生じた深掘りポイントを踏まえて、下記3つの観点を整理し、それぞれの観点ごとに選定した事業者に対してヒアリングを実施しました。

### 観点 1

SNSの利用を検討中または利用に踏み出せない理由がある事業者に、利用まで至っていないハードルは何かをヒアリングする

### 観点 2

SNSを利用している、かつ他社情報取得に関して課題を持っている事業者に、現状の取得方法と具体的な課題をヒアリングをする

### 観点 3

SNSで障害把握をしたと回答した事業者に具体的な把握内容をヒアリングする

### 3.3. ヒアリングおよびアンケート調査のまとめ

アンケート調査とヒアリング調査の結果を踏まえて、インターネット障害の発生・把握状況、事業者が抱える課題、障害把握におけるSNSの利用状況についてまとめます。

#### インターネット障害の発生・把握状況

今回の調査で、総務省の定める重大事故の条件に満たないインターネット障害は、今回回答のあった通信事業者の8割以上で発生しており、そのうち、**NW監視では把握できない所謂サイレント障害が全体の約3割を占めている**ことが分かりました。これらのサイレント障害の約半数が**発生から把握まで3時間以上**を要しています。

#### 事業者が抱える課題

通信事業者の6割強がインターネット障害の把握に関する課題感を持っています。

他社ネットワーク利用やクラウドサービスの普及等で外部との接続が増え、ネットワーク全体が複雑化している中、各事業者は自社の設備やネットワークを監視するだけでは不十分であり、接続先のネットワークやインターネット全体の状況を広く把握する必要があると認識しています。しかし、ネットワーク全体の状況把握は個社対応では限界があることや、**人員不足などの体制面の課題も抱えており、具体的な対応は進んでいない**状況です。

#### 障害把握におけるSNSの利用状況

約7割の事業者はSNSを障害把握に利用しておらず、その理由として**運用上の負荷が大きい、情報の信ぴょう性の低さ**を挙げています。また、これらの事業者の約9割が今後のSNS利用も未定と回答しています。

一方で、約3割の事業者は障害把握にSNSを利用しており、これらの事業者の大半はSNSの有用性を実感しています。

中でも障害把握に最も多く利用されている**Twitterでは、情報の量、即時性、検索性などを理由にインターネット障害の把握に有用と感じている事業者の割合が8割**を超えています。

## 4.調査3 : SNSを使用した実証調査

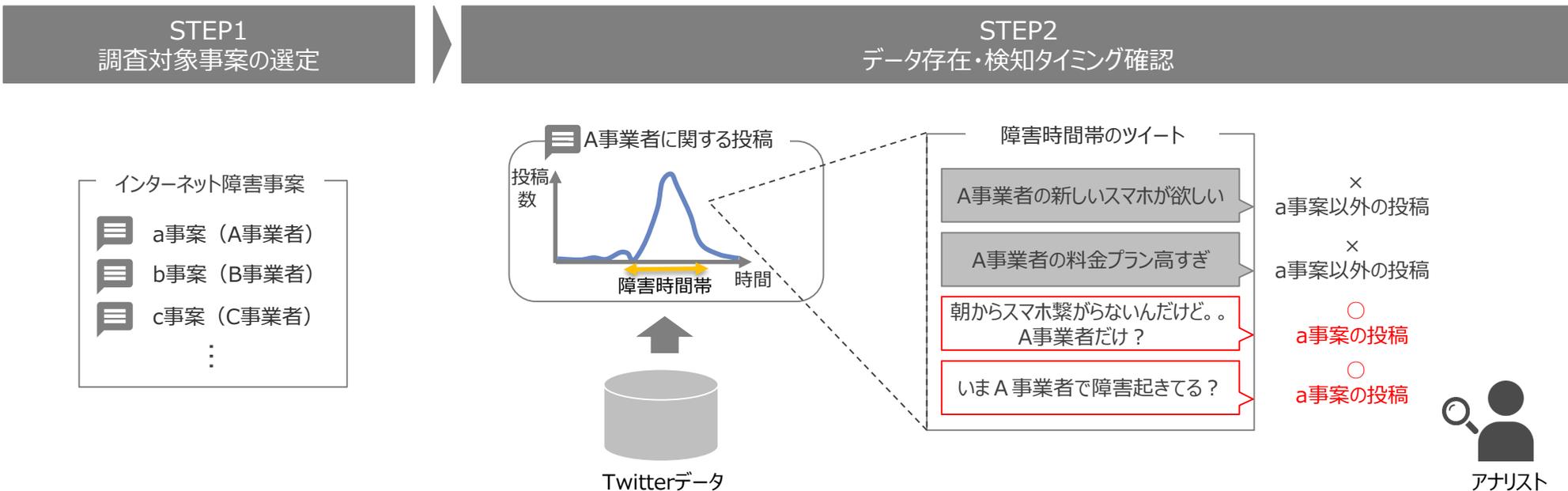
## 4.1. 調査方法 –調査フロー–

本調査では、SNS（Twitter）からインターネット障害の把握に関する有用な情報を抽出できるかを確認するために、Twitterの過去データを用いてインターネット障害を把握するための特徴的なキーワードを分析、事象抽出モデルを構築し、このモデルを用いて過去1年間（2019年）のインターネット障害がどの程度検知できるかの検証・評価を行いました。本調査は下記の4STEPで実施しました。



## 4.1. 調査方法 - 調査フロー

本調査の手順のSTEP1と2の概要を下記に記載します。過去に発生した障害から10事案を選定して、障害に関するツイートが存在するかを確認します。また、ツイートによる障害検知のタイミングが障害発生後どの時点であるかを確認します。



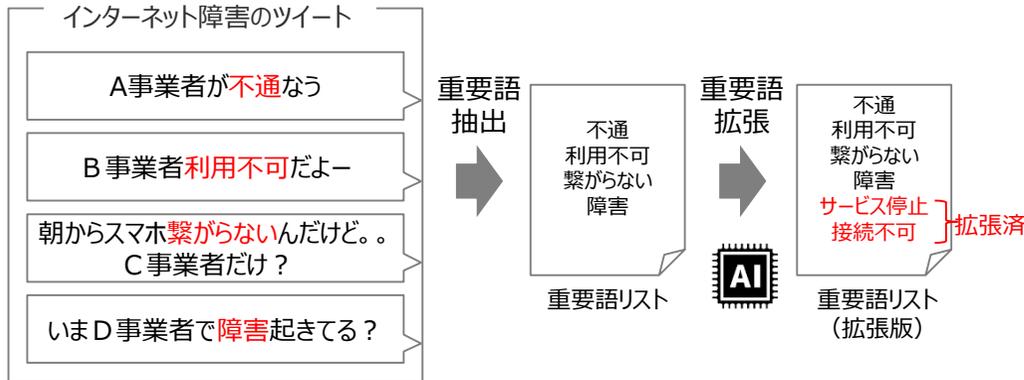
1. 過去に発生したインターネット障害のうち、Twitterを使った障害事象を抽出するモデルを作成するために使用する事案を10事案程度選定する

1. STEP1の選定事案について、障害発生前後の事業者名を含む投稿をTwitter全量データから抽出する  
2. 抽出したツイートが障害事案に関するデータであるかを目視確認する  
3. 公表されている障害発生時刻とツイート量増加のタイミングを比較して早期検知の可能性を検証する

## 4.1. 調査方法 -調査フロー-

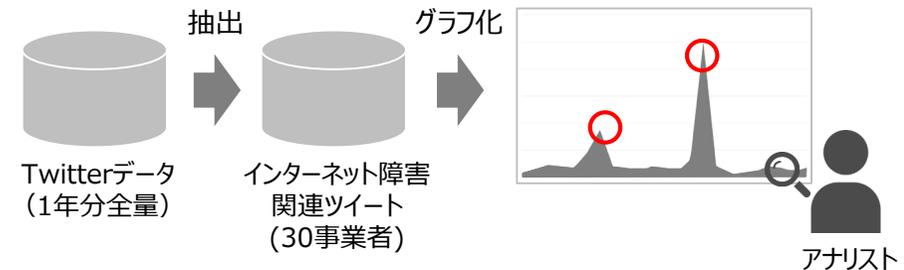
本調査の手順のSTEP3と4の概要を下記に記載します。STEP3ではSTEP2で確認した障害事案のツイートを分析することで、障害に特徴的な単語を抽出することで事象抽出モデルを作成します。STEP4ではこのモデルの評価を行います。

### STEP3 事象抽出モデルの構築



1. STEP2でデータ存在確認ができた事案について、インターネット障害のツイートに含まれる特徴的なキーワード（重要語）を分析して重要語リストを作成する
2. 機械学習の手法であるword2vec, Measure propagationを使って、重要語リストの単語を拡張する

### STEP4 事象抽出モデルの評価



1. 重要語リストを用いて、1年分のTwitter全量データから特定の30事業者のインターネット障害に関するツイートを抽出する
2. 事業者ごとにグラフ化を行い、グラフの波形（ピーク性）と投稿内容からインターネット障害を検知する
3. ニュース、事業者ホームページで公表されている障害の検知可否を調査して検知率を検証する

## 4.2. STEP1 調査対象事案の選定

## 4.2.1 調査対象事案の選定

STEP1  
調査対象事案の選定

STEP2  
データ存在・  
検出タイミング確認

STEP3  
事象抽出モデルの構築

STEP4  
事象抽出モデルの評価

STEP1では過去に発生した障害事案の中から、今回の調査対象として障害規模、サービス内容が異なり、かつ多くのツイート投稿数があると考えられる下記の10事案を選定しました。

No.	発生日	障害名	事業者名	障害が起きたサービス	影響利用者数
1	2019/9/9	台風15号による携帯電話通信障害	ドコモ、KDDI、ソフトバンク	音声通話 データ通信	不明
2	2019/8/23	AWS障害	アマゾン ウェブ サービス ジャパン	クラウド基盤	不明
3	2019/3/13	Gmail障害	Google	インターネットメール	不明
4	2018/12/6	ソフトバンク通信障害	ソフトバンク、LINEモバイル	音声通話 データ通信	約3070万
5	2017/8/25	海外事業者起因による インターネット通信障害	NTTコミュニケーションズ	インターネット回線	不明
6	2015/7/12	au携帯電話の 電子メールサービス障害	KDDI	携帯メール	最大約796万
7	2015/4/2	LINEサービス障害	LINE	無料音声通話 およびメッセージ	最大約5200万
8	③2014/11/6 ②2014/9/30 ①2014/9/27	yahoo!メール障害	ヤフー	インターネットメール	③約264万 ②約379万 ①約236万
9	2018/5/29	エネコムインターネット接続・ メール障害	エネルギア・コミュニケーションズ	インターネット回線 メール	約17万
10	2017/4/7	楽天、楽天コミュニケーションズ データ障害	楽天、 楽天コミュニケーションズ	データ通信	約22万

## 4.3 STEP2

### データ存在・検知タイミング確認

STEP2では、STEP1で選定した事案のツイートデータの存在確認と、ツイート件数が増加するタイミングを調査します。STEP2における調査方法を以下に記載します。

### データ抽出・集計方法

- 障害発生前後の計4日分のツイートデータを抽出対象とする
- 「企業名、サービス名」のみと「企業名、サービス名×障害に関するキーワード」について、ツイートを分単位で集計してそれぞれツイート件数グラフを作成する
- ツイートはオーガニック投稿のみを集計する（リツイートは含まない）

### データ確認方法

- 障害前後の時間帯のツイートを目視確認して障害に関するツイートが存在するかを調査する
- 障害発生からツイート増加開始までの時間とピーク時の件数を調査する

## 4.3.2. 調査結果 –AWSの障害事例

STEP1  
調査対象事案の選定

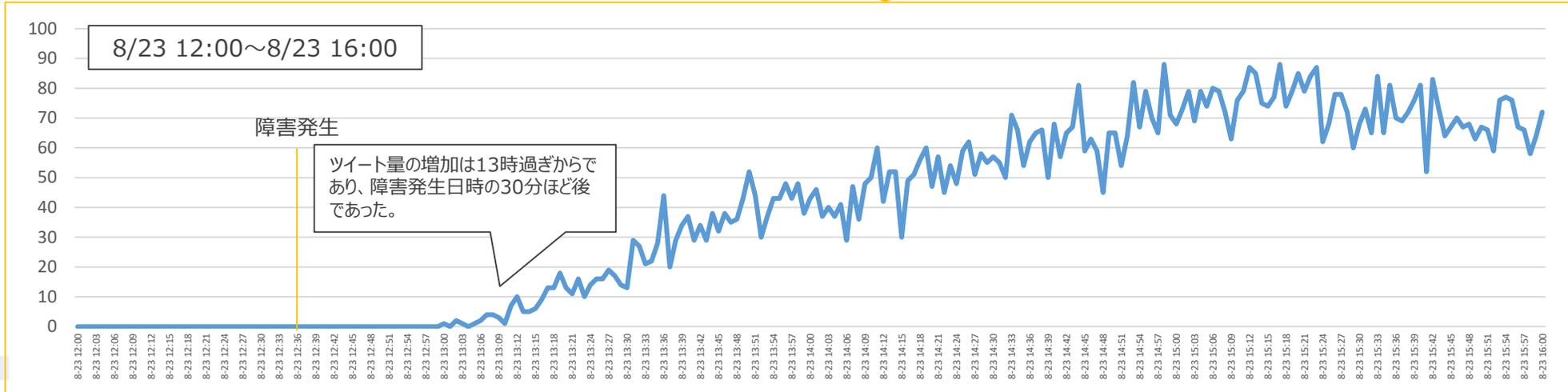
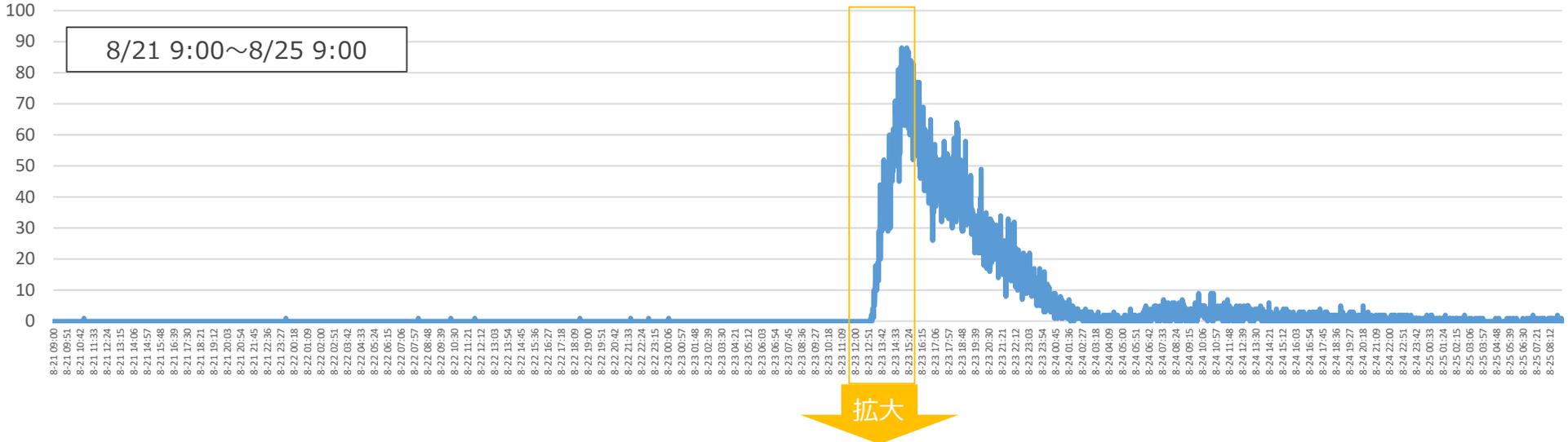
STEP2  
データ存在・  
検知タイミング確認

STEP3  
事象抽出モデルの構築

STEP4  
事象抽出モデルの評価

AWS障害（障害時間：2019/8/23 12:36～18:30頃）

「企業名、サービス名×障害に関するキーワード」で集計



## 4.3.2. 調査結果 –AWSの障害事例

STEP1  
調査対象事案の選定

STEP2  
データ存在・  
検知タイミング確認

STEP3  
事象抽出モデルの構築

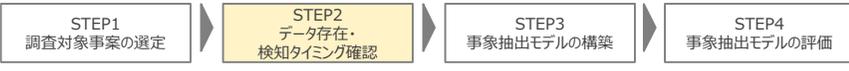
STEP4  
事象抽出モデルの評価

AWS障害（障害時間：2019/8/23 12:36～18:30頃）

ツイートサンプル

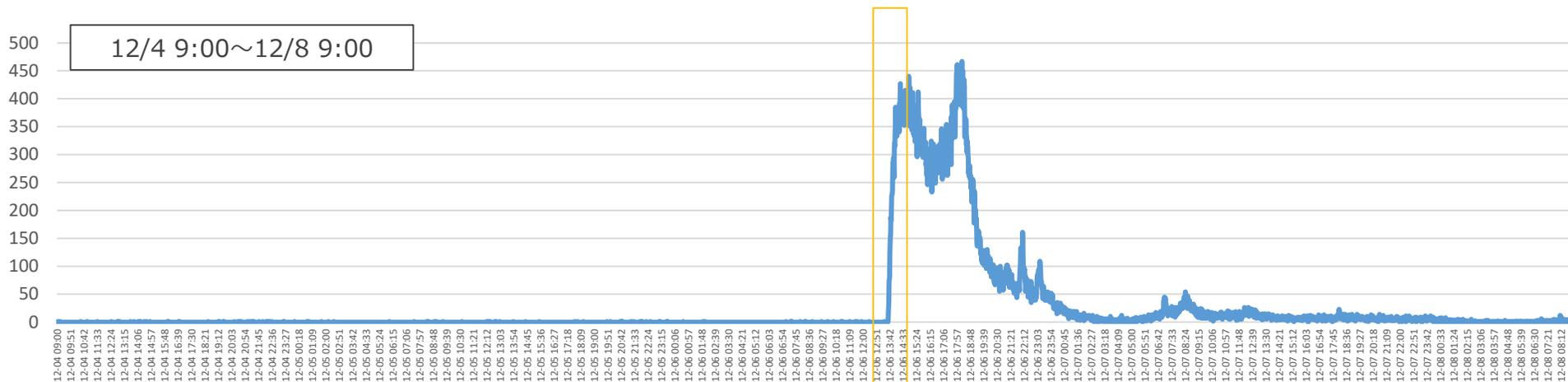
投稿時間	ツイート本文
2019-08-23 13:00:57	AWS ゾーン障害起きてる？一気に大量のFailover通知
2019-08-23 13:02:08	AWS障害？
2019-08-23 13:02:26	AWSのAZ範囲で障害？
2019-08-23 13:03:45	AWSでゾーンレベルの障害起きてるな
2019-08-23 13:05:11	AWS障害っぽ
2019-08-23 13:06:14	aws障害臭
2019-08-23 13:06:59	AWS 障害...
2019-08-23 13:07:21	AWSのap-northeast-1aにある多くのサーバやサービスで障害が発生してるっぽい。全てじゃないけど同時発生で台数も多い。
2019-08-23 13:07:33	AWSが落ちてとみどころさんがお祭り状態になってる
2019-08-23 13:07:43	awsが死んだので世界が終わった(障害出まくり)
2019-08-23 13:07:53	AWS障害.....??
2019-08-23 13:08:05	AWS死んでる
2019-08-23 13:08:17	夏のAWS障害祭り、嬉しくない
2019-08-23 13:08:21	AWS障害かねえ
2019-08-23 13:08:40	awsが死んでるらしい……
2019-08-23 13:09:25	AWS障害マ？
2019-08-23 13:09:51	AWSの障害かぁ
2019-08-23 13:09:58	aws死んでる
2019-08-23 13:10:44	AWS障害発生してるみたい？使ってるアプリはのきなみログインできないぞい:(;´ω`);
2019-08-23 13:11:00	AWSで大規模障害か

## 4.3.2. 調査結果 –ソフトバンク障害の事例



No.4 ソフトバンク通信障害（障害時間：2018/12/6 13:39～18:04）

「企業名、サービス名×障害に関するキーワード」で集計



拡大



## 4.3.2. 調査結果 –ソフトバンク障害の事例

STEP1  
調査対象事案の選定

STEP2  
データ存在・  
検知タイミング確認

STEP3  
事象抽出モデルの構築

STEP4  
事象抽出モデルの評価

### ソフトバンク通信障害（障害時間：2018/12/6 13:39～18:04） ツイートサンプル

投稿時間	障害関連ツイート
2018-12-06 13:34:59	なんかソフトバンクいきなり圏外になったんだけど
2018-12-06 13:35:02	なんかソフトバンク圏外なった <a href="https://t.co/zDbox2UL8A">https://t.co/zDbox2UL8A</a>
2018-12-06 13:35:06	どーしたのソフトバンクー！
2018-12-06 13:35:08	ソフトバンクの回線死んでる☺
2018-12-06 13:35:19	ソフトバンク圏外になった
2018-12-06 13:35:23	は？SoftBank圏外なんだけど
2018-12-06 13:35:36	ソフトバンクが圏外なってんやけど
2018-12-06 13:35:44	ソフトバンクいきなり圏外になったけど？？？ <a href="https://t.co/LMBS1jwhyS">https://t.co/LMBS1jwhyS</a>
2018-12-06 13:35:44	softbank 圏外 らしい
2018-12-06 13:35:47	自宅なのに圏外なんだけどソフトバンクどうした？
2018-12-06 13:35:50	SoftBankが圏外になった模様 <a href="https://t.co/PZXfSeSznk">https://t.co/PZXfSeSznk</a>
2018-12-06 13:35:53	ソフトバンク回線しんでね？？
2018-12-06 13:35:56	え、SoftBankの電波圏外になってblitz切断されたんやが
2018-12-06 13:35:57	なに？通信障害？？頼むぜソフトバンクさん
2018-12-06 13:35:58	ソフバン一瞬圏外になったか？
2018-12-06 13:35:58	ソフトバンク使ってるんですけど急に圏外になって
2018-12-06 13:36:00	ソフトバンク通信障害発生中
2018-12-06 13:36:00	ソフトバンクだけなん？圏外、
2018-12-06 13:36:00	ソフトバンク逝ったわ
2018-12-06 13:36:01	ソフトバンク圏外なんやが？おれだけじゃないみたいどうした

## 4.3.2. 調査結果

STEP1  
調査対象事案の選定

STEP2  
データ存在・  
検知タイミング確認

STEP3  
事象抽出モデルの構築

STEP4  
事象抽出モデルの評価

影響利用者数が多いほどツイート増加開始までの時間間隔は短く、ピーク時ツイート数は多くなっており、影響利用者数が少ないと時間間隔は長く、ピーク時ツイート数は少なくなる傾向が見られました。

ツイート数増加開始までの時間×ピーク時ツイート数×影響利用者数

		障害発生からツイート数増加開始までの時間		
		長い (30分以上)	中程度 (10分～30分未満)	短い (10分未満)
ピーク時の ツイート数 (1分間あたり)	多い (100件 以上)			<div style="background-color: #0056b3; color: white; padding: 2px;">No.4 ソフトバンク通信障害</div> <div style="background-color: #0056b3; color: white; padding: 2px;">No.7 LINEサービス障害</div>
	中程度 (10件～ 99件)	<div style="border: 1px solid black; padding: 2px;">No.2 AWS障害</div> <div style="background-color: #00a0e3; padding: 2px;">No.6 au携帯メール障害</div>	<div style="border: 1px solid black; padding: 2px;">No.3 Gmail障害</div> <div style="border: 1px solid black; padding: 2px;">No.5 海外事業者起因による障害</div> <div style="background-color: #00a0e3; padding: 2px;">No.8 Yahoo!メール障害①</div>	<div style="background-color: #00a0e3; padding: 2px;">No.8 Yahoo!メール障害③</div>
	少ない (10件未 満)	<div style="border: 1px solid black; padding: 2px;">No.1 台風15号による通信障害</div> <div style="background-color: #00a0e3; padding: 2px;">No.8 Yahoo!メール障害②</div> <div style="background-color: #add8e6; padding: 2px;">No.9 エネコムサービス障害</div> <div style="background-color: #add8e6; padding: 2px;">No.10 楽天通信障害</div>		

## 4.3.2. 調査結果

音声通話やチャットサービスでは、ツイート増加までの時間が短く、ピーク件数が多くなる傾向がみられました。これらはリアルタイムでコミュニケーションを取る必要性が高く、障害発生時に利用者がすぐに異変に気づきやすいためであると考えられます。

ツイート数増加開始までの時間×ピーク時ツイート数×影響サービス内容

		障害発生からツイート数増加開始までの時間		
		長い (30分以上)	中程度 (10分～30分未満)	短い (10分未満)
ピーク時の ツイート数 (1分間あたり)	多い (100件以上)			<div style="background-color: #0056b3; color: white; padding: 5px;">No.4 ソフトバンク通信障害</div> <div style="background-color: #00a651; color: white; padding: 5px;">No.7 LINEサービス障害</div>
	中程度 (10件～99件)	<div style="background-color: #333; color: white; padding: 5px;">No.2 AWS障害</div> <div style="background-color: #e67e22; color: white; padding: 5px;">No.6 au携帯メール障害</div>	<div style="background-color: #e67e22; color: white; padding: 5px;">No.3 Gmail障害</div> <div style="background-color: #663366; color: white; padding: 5px;">No.5 海外事業者起因による障害</div> <div style="background-color: #e67e22; color: white; padding: 5px;">No.8 Yahoo!メール障害①</div>	<div style="background-color: #e67e22; color: white; padding: 5px;">No.8 Yahoo!メール障害③</div>
	少ない (10件未満)	<div style="background-color: #0056b3; color: white; padding: 5px;">No.1 台風15号による通信障害</div> <div style="background-color: #e67e22; color: white; padding: 5px;">No.8 Yahoo!メール障害②</div> <div style="background-color: #663366; color: white; padding: 5px;">No.9 エネコムサービス障害</div> <div style="background-color: #0056b3; color: white; padding: 5px;">No.10 楽天通信障害</div>		

## 4.3.2. 調査結果



障害発生から最初の障害投稿までの時間は下記の通りです。平均7分程度で最初の投稿がされており、ツイート増加検知まで時間がかかる障害（No.1、No.10）であっても障害発生後すぐにツイートがされていることが分かりました。

No.	障害名	障害発生時刻から 最初の障害ツイート投稿までの時間 ① ※1	障害発生時刻から ツイート数増加開始までの時間 ②	差分②－①
1	台風15号による携帯電話通信障害	7分	389分	382分
2	AWS障害	24分	36分	12分
3	Gmail障害	不明 ※2	19分 ※3	－
4	ソフトバンク通信障害	-5分	-4分	1分
5	海外事業者起因によるインターネット通信障害	8分	10分	2分
6	au携帯電話の電子メールサービス障害	11分	40分	29分
7	LINEサービス障害	1分	3分	2分
	yahoo!メール障害①	15分	20分	5分
8	yahoo!メール障害②	1分	36分	35分
	yahoo!メール障害③	0分	4分	4分
9	エネコムインターネット接続・メール障害	13分	40分	27分
10	楽天、楽天コミュニケーションズデータ障害	3分	198分	195分
	平均	<b>7分</b>	<b>64分</b>	-

※1 公表されている障害発生時刻から障害に関する最初のツイートが見られるまでのおおよその時間

※2 障害発生時間が公表されていないため、不明としている

※3 障害発生時間が公表されていないため、障害に関する最初のツイート時刻と件数増加傾向が見られ始めるまでの差分時間としている

## 4.4. STEP3 事象抽出モデルの構築

STEP3では、Twitterを活用したインターネット障害事象を抽出するためのモデルを構築します。

### 事象抽出モデル

本調査では、インターネット障害のツイートを検知するための特徴的なキーワードを分析・抽出して事象抽出モデルとします。事象抽出モデルは、STEP2において抽出した影響利用者数やサービス内容の異なる10事案の障害時のツイートを分析することで、障害関連ツイートに特徴的なキーワードを抽出して作成します。

### 重要語の抽出

障害関連ツイートに特徴的なキーワードを抽出する方法として、障害時のツイートにおける出現頻度の高い単語を抽出します。ただし、障害時のツイートにおいて、障害が起きている企業名やサービス名を含むツイートであっても、障害とは無関係のツイートも存在します。障害関連ツイートに特徴的なキーワードを抽出するためには、分析時にノイズとなる障害とは無関係のツイートを除外して、障害関連のツイートのみを分析対象とすることが必要です。

そこで、障害関連ツイートには「つながらない、使えない、電波悪い」などの不満、怒り、苦情といったネガティブ表現を含むツイートが多いという特徴を利用して、STEP2で抽出した障害時のツイートのうち、ネガティブな表現を含むツイートを絞り込んで分析対象とします。今回、ネガティブな表現の抽出には日本語意味理解技術の「なずきSA」を使用します。これにより、ネガティブな表現を含むツイートから抽出した出現頻度の高いキーワードを障害関連ツイートの重要語としてリスト化します。

### 重要語の拡張

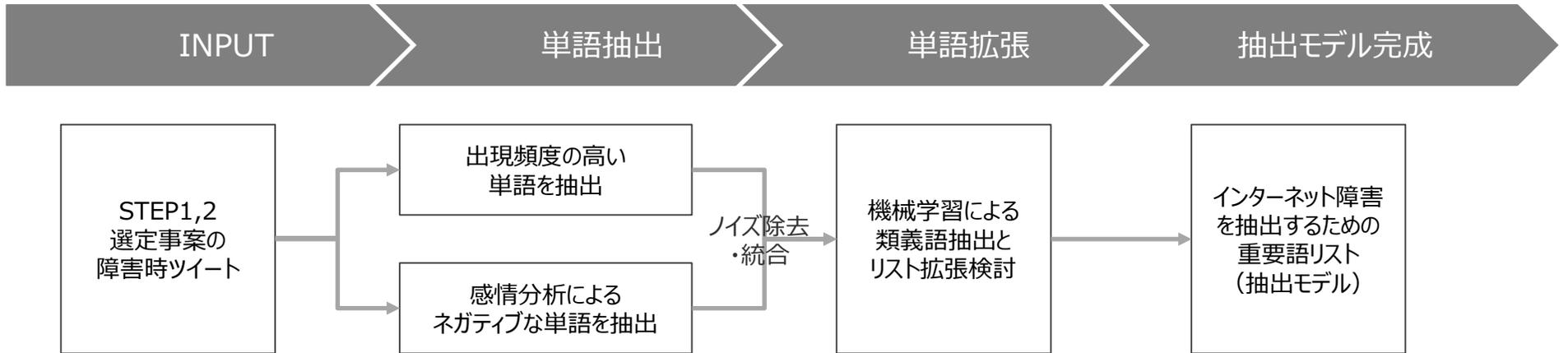
作成した重要語リストには下記2つの特性があります。

1. 抽出元のツイートが特定の10事案の障害ツイートに限られているため、抽出したキーワードに偏りがある可能性がある
2. 出現頻度の高い単語を抽出しているため、出現頻度は低い障害関連ツイートに特徴的なキーワードが存在する場合、重要語リストに含まれていない可能性がある

今回、この特性1,2に対処するため、2種類の機械学習手法を用いて重要語の拡張を行うことで、汎用的な重要語リストを作成します。今回、特性1に対してはWord2Vecを、特性2に対してはMeasure propagationを使用して拡張を行います。

## 4.4.1.事象抽出モデルの構築

障害関連ツイートから重要語の抽出、拡張を行い重要語リストを作成するまでの流れを下記に記載します。ここで作成した重要語リストを本調査におけるインターネット障害事象を抽出するためのモデルとします。



- STEP1,2で使用した対象事案の障害発生直後のツイート（企業名またはサービス名を含むツイート）をインプットとする
- 事案ごとに**ツイート中の出現頻度の高い単語を抽出**する
- 事案ごとにツイートの感情分析を行い、**苦情系などネガティブな単語を抽出**する（なずきSAを使用）
- それぞれ抽出した単語が**ノイズを多く含むものかを確認し、不要な単語を除外。上記を統合して単語リストを作成する**
- 単語リストを拡張する余地があるかを確認するため、機械学習（Word2Vec、Measure propagation）を使った**類義語抽出**を行う
- 抽出した単語をアナリストが確認して単語リストへの追加可否を判断する
- インターネット障害を抽出するための重要語リストが完成する

## 4.4.1. 事象抽出モデルの構築



本調査では、重要語リストを作成・拡張するにあたり日本語意味理解技術の「なずきSA」、機械学習アルゴリズムである「Word2Vec」と「Measure propagation」を使用しました。

### なずきSAによる感性分析

重要語リストを作成するにあたり、インターネット障害に関するツイート中に出現するネガティブな表現を抽出するために、ツイート本文を「なずきSA」を使って感性分析することで、「不満」、「怒り」、「苦情」などに関する表現を抽出しました。ここで抽出した表現を基にして重要語リストを作成しました。

### Word2Vecによる拡張

重要語リストはSTEP1,2で選定・確認した障害10事案を基にして、インターネット障害のツイートに特徴的な単語を抽出しています。10事案は異なる規模、サービス内容（音声通信、メール、インターネット、チャットなど）から構成されていますが、抽出した単語に偏りが生じている可能性も考えられます。このため、Word2Vecを用いて、特定の障害に関するツイートに限定することなく、あらゆるジャンルのツイートを学習させたモデルを作成して重要語リストの単語の類義語を抽出することで拡張を試みました。

### Measure propagationによる拡張

ツイート中の出現頻度自体は低いインターネット障害に関する特徴的な単語を抽出するために、Measure propagationを用いて、障害時に特徴的な単語を機械的にグルーピング・抽出を行い、重要語リストの拡張を試みました。

## 4.4.2.事象抽出モデルの構築結果

STEP1  
調査対象事象の選定

STEP2  
データ存在・  
検知タイミング確認

STEP3  
事象抽出モデルの構築

STEP4  
事象抽出モデルの評価

重要語を作成・拡張した結果、以下の134単語からなる重要語リストを作成しました。この重要語リストを本調査の事象抽出モデルとします。（単語抽出で126語、単語拡張で8単語を選定）

### 重要語リスト

利用できない	遅く	障害	使えな	繋がりにくい	フリーズ	なおってくれ	しっかりして
落ち	遅い	重たい	使えてない	繋がら	ぶっ壊れ	トラブル	こわれて
返信できない	送信不可	重すぎる	使えず	繋がってなくて	ぶっこわれ	トラブる	こわれた
復旧せん	送信失敗	重くね	混んでる	繋がってない	ひらけん	トラブっ	お亡くなり
復旧してほしい	送信できませんでした	重くて	故障	勘弁してくれ	開けない	どうしようもない	おかしく
復旧してくれ	送信できない	重い	固まっ	開けん	ひらけな	テザリングできな	おかしかった
復旧してください	送受信出来ない	受信できない	見れん	壊滅	パケット通信できな	テザリング出来な	おかしい
復旧してー	送れん	受け取れん	見れねえ	壊れとる	パケット通信出来な	つながん	応答せず
不通	送れなくて	事故ってる	見れなくなって	壊れて	バグっとる	つながりません	応答しな
不調	送れない	死亡	見れなくて	壊れた	ばぐっとる	つながりにくい	エラー
不具合	早く直せ	死んどる	見れない	回線悪くなって	バグって	つながら	動かな
不安定	早く何とかしろ	死んで	見えない	ログイン出来ない	ばぐって	つかえん	うごかな
届かない	全滅	死んだ	圏外	ログインできなくて	バグった	つかえない	受け取れな
読み込めない	接続できない	使えん	繋がん	ログインできない	ばぐった	ダメっぽい	イライラする
電波悪い	逝って	使えません	繋がりません	やらかし	なんか起きて	ダメだ	いい加減にしろ
調子悪い	逝った	使えへん	繋がりませーん	メンテ	なんかおきて	ダウン	
調子が悪い	逝く	使えね	繋がりにくくなって	みれない	なくなる	しっかりしろ	

※青字は単語拡張で選定した単語

## 4.5. STEP4 事象抽出モデルの評価

## 4.5.1. 調査方法 –データ抽出・集計・確認



STEP4では、STEP3で構築したインターネット障害抽出モデルの重要語リストを使って、過去に発生したインターネット障害がどのくらい検知できるかを評価します。本調査では、2019年1年間分の日本語全量ツイートを使用して、通信事業者30社の1年間に発生した障害を下記の方法で抽出・集計、確認します。

### データ抽出・集計方法

- 2019年1月1日から12月31日までの1年分のツイートデータを対象とする
- 「重要語リストのキーワード」と「企業名およびサービス名リストのキーワード」を両方含むツイートを日単位で集計して、事業者ごとにツイート件数グラフを作成する
- ツイートはオーガニック投稿のみを集計する（リツイートは含まない）

### データ確認方法

- ツイート件数グラフのピーク（山）が発生しているタイミング（日）のツイート内容を目視で確認して、障害に関するものか否かを判定する

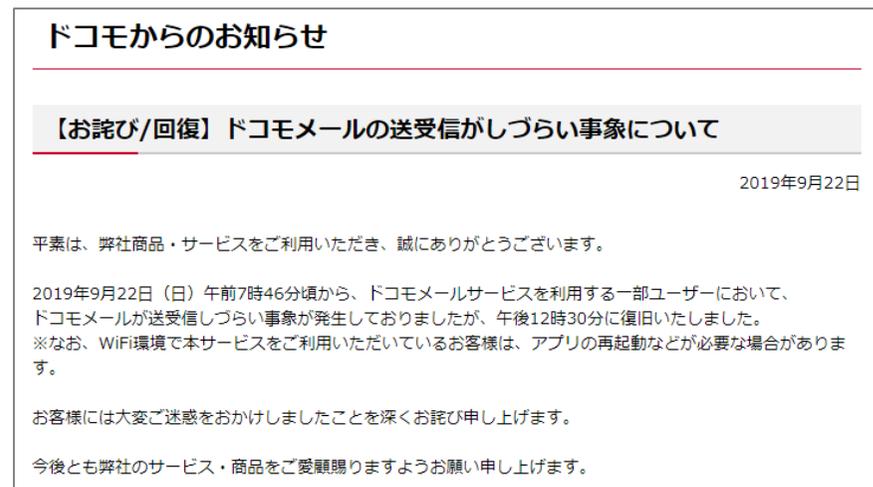
今回の調査ではツイートから障害を抽出する方法と比較するために、ニュースおよび事業者ホームページで公表されている障害についても調査しました。ニュースおよび事業者ホームページからの障害確認方法は以下の通りです。

### ニュースおよび事業者ホームページからの障害確認方法

- ニュース記事による障害事案の抽出は、「事業者名およびサービス名 AND 障害」等でGoogleニュース検索を行い抽出した
- 事業者ホームページからの障害事案の抽出は、各事業者公式サイト上の障害情報、ニュースリリース、お知らせ等の障害情報記載ページから抽出した（2020年3月時点で確認できるものが対象）



ニュースによる障害事案抽出の例



事業者ホームページからの障害事案抽出の例

## 4.5.1. 調査方法 - 評価

障害抽出モデルを評価するために、ツイートから確認できた障害、ニュースおよび事業者ホームページから確認できた障害、どちらか一方から確認できた障害、両方から確認できた障害の件数を調査します。

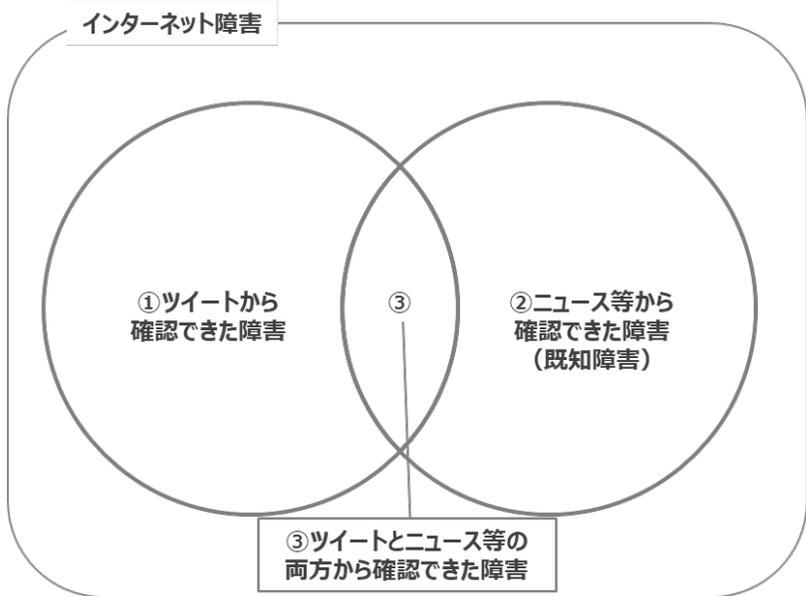


図1 障害検知手法のベン図

No.	調査対象項目		図1の該当箇所
(1)	ツイートまたはニュース等から確認できた障害	① ∪ ②	
(2)	ツイートから確認できた障害	①	
(3)	ニュース等から確認できた障害 (既知障害)	②	
(4)	ツイートとニュース等の両方から確認できた障害	③ (① ∩ ②)	
(5)	ツイートのみで確認できた障害 (未知障害)	① - ③	
(6)	ニュース等のみで確認できた障害	② - ③	

## 4.5.1. 調査方法 – 評価

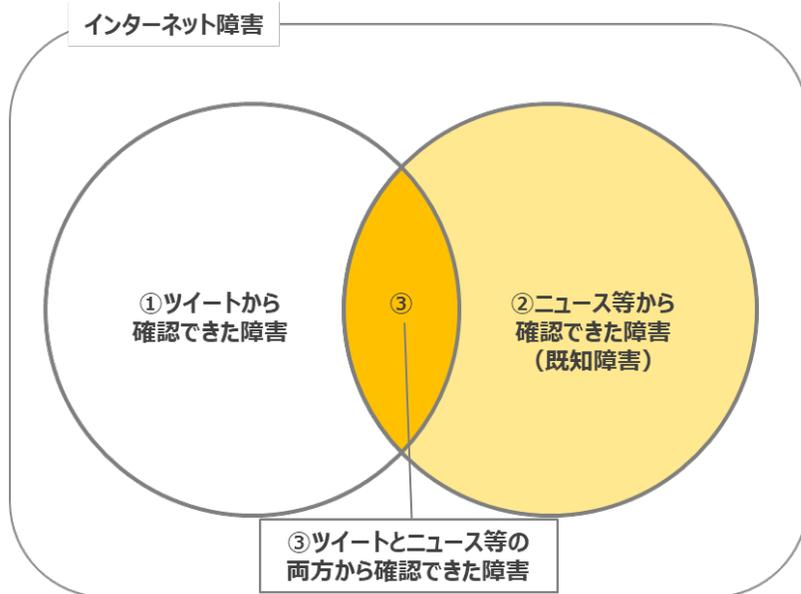
STEP1  
調査対象事案の選定

STEP2  
データ存在・  
検知タイミング確認

STEP3  
事象抽出モデルの構築

STEP4  
事象抽出モデルの評価

また、公表されている障害をツイートからどのくらい検知できるかを調べるために、ツイートによる既知障害の検知率を評価します。ツイートによる既知障害の検知率は、ニュースおよび事業者ホームページで公表されている障害件数のうち、ツイートから確認できた障害件数の割合から算出します。



No.	調査対象項目	
(7)	ツイートによる既知障害の検知率 (%)	③ ÷ ②

図2 ツイートによる既知障害の検知率

## 4.5.1. 調査方法

ツイートを抽出するための30事業者の企業名およびサービス名に関するキーワードは以下の通りです。

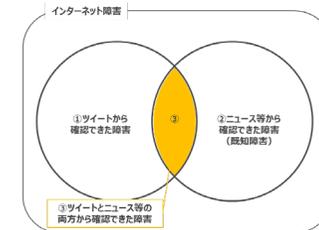
### 企業名およびサービス名リスト

No.	事業者名	キーワード（英語大小全半角ゆれ含む）				
1	NTT東	ntt東	フレッツ			
2	NTT西	ntt西	フレッツ			
3	NTTコミュニケーションズ	nttコミュニケーションズ	nttコム	nttcom	nコム	ncom    ocn
4	NTTドコモ	ドコモ	docomo			
5	KDDI	kddi	au			
6	ソフトバンク	ソフトバンク	softbank	ソフバン		
7	楽天モバイル	楽天	無料サポータープログラム			
8	ジュピターコム	ジュピターコム	ジェイコム	jcom	zaq	
9	ニフティ	ニフティ	nifty	nifmo		
10	ビッグロース	ビッグロース	biglobe			
11	So-net	so-net	sonet	ソネット	nuro光	nuroモバイル
12	UQコミュニケーションズ	uqコミュニケーションズ	uqモバイル	uq wimax	uq	
13	中部テレコミュニケーション	中部テレコミュニケーション	コムファ	中テレ		
14	オプテージ（ケイオプティコム）	オプテージ	ケイオプティコム	eo光	mineo	
15	エネルギア・コミュニケーションズ	エネルギア	エネコム	えねこむ	メガエッグ	メガ・エッグ    mega egg
16	QNet	qnet	bbiq	ビビック	qtモバイル	qt mobile
17	STNet	stnet	pikara	ピカラ	光ねっと	fiimo
18	インターネットイニシアティブ	インターネットイニシアティブ	インターネットイニシアチブ	ijj	ijjmio	
19	Google	gmail	googleクラウド	google cloud	gcp	
20	Amazon	aws				
21	Facebook	facebook	instagram			
22	Microsoft	office365	skype	windows	azure	
23	Twitter	twitter	ツイート			
24	LINE	line	ライン			
25	yahoo	yahoo	ヤフー			
26	宮崎ケーブルテレビ	宮崎ケーブル	mcnひかり	mcnモバイル		
27	豊橋ケーブルネットワーク	豊橋ケーブル				
28	CAC	cac				
29	秋田ケーブルテレビ	秋田ケーブル	cna光			
30	TOKAIケーブルネットワーク	TOKAIケーブル	ひかりdeネット			

## 4.5.2. 調査結果

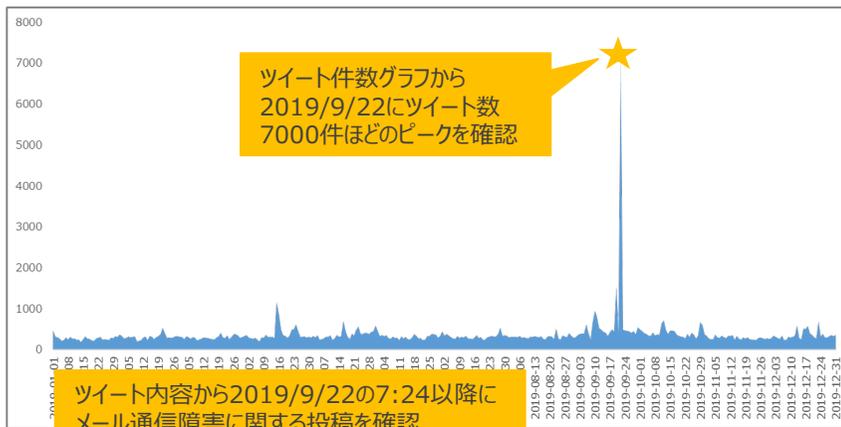


ツイートからとニュース・事業者ホームページからの両方から確認できた障害の一例を下記に記載します。



### ドコモ障害（2019/9/22発生）

#### ツイートからの障害確認



ツイート内容から2019/9/22の7:24以降にメール通信障害に関する投稿を確認

投稿時刻	ツイートサンプル
09-22 07:24	通信障害起きてるのか？ メールが受信出来ない…… #docomo #メール #通信障害 #長崎県 #大村
09-22 07:25	ドコモ 通信エラーでるな
09-22 07:35	ドコモのメール通信エラーで受信できない(&gt;&t;)
09-22 07:35	docomo通信エラーが☹️
09-22 07:36	あれ…？ドコモ通信障害？ メール問い合わせできないな。やっぱり台風のせいかしら？

#### ニュース・事業者ホームページからの障害確認



#### ドコモからのお知らせ

##### 【お詫び/回復】ドコモメールの送受信がしづらい事象について

2019年9月22日

平素は、弊社商品・サービスをご利用いただき、誠にありがとうございます。

2019年9月22日（日）午前7時46分頃から、ドコモメールサービスを利用する一部ユーザーにおいて、ドコモメールが送受信しづらい事象が発生しておりましたが、午後12時30分に復旧いたしました。  
※なお、WiFi環境で本サービスをご利用いただいているお客様は、アプリの再起動などが必要な場合があります。

お客様には大変ご迷惑をおかけしましたことを深くお詫び申し上げます。

今後とも弊社のサービス・商品をご愛顧賜りますようお願い申し上げます。

事業者ホームページから、2019/9/22 7:46頃にメールサービスの障害が発生していたことを確認

※事業者ホームページのURL  
[https://www.nttdocomo.co.jp/info/notice/page/190922\\_01\\_m.html](https://www.nttdocomo.co.jp/info/notice/page/190922_01_m.html)

## 4.5.2. 調査結果

STEP1  
調査対象事案の選定

STEP2  
データ存在・  
検知タイミング確認

STEP3  
事象抽出モデルの構築

STEP4  
事象抽出モデルの評価

調査した30社の障害について、ツイートから確認できた障害件数、ニュース・事業者ホームページから確認できた障害の件数、ツイートにおける障害検知率について下記に記載します。次頁に各事業者ごとの件数、検知率を記載します。

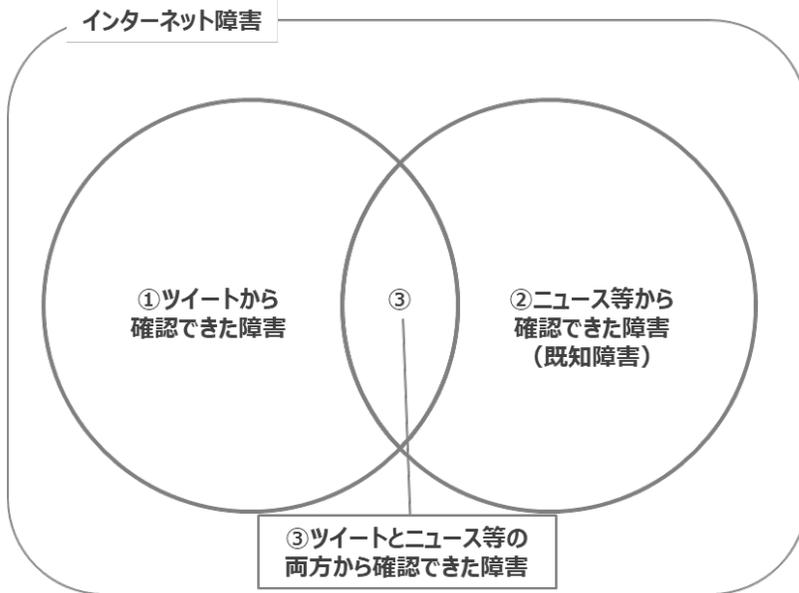


図1 障害検知手法のベン図

No.	障害検知の種類		図1の該当箇所	件数および検知率
(1)	ツイートまたはニュース等から確認できた障害	① ∪ ②		<b>165件</b>
(2)	ツイートから確認できた障害	①		<b>97件</b>
(3)	ニュース等から確認できた障害 (既知障害)	②		<b>119件</b>
(4)	ツイートとニュース等の両方から確認できた障害	③ (① ∩ ②)		<b>51件</b>
(5)	ツイートのみで確認できた障害 (未知障害)	① - ③		<b>46件</b>
(6)	ニュース等のみで確認できた障害	② - ③		<b>68件</b>
(7)	ツイートにおける既知障害の検知率 (%)	③ / ②		<b>43%</b>

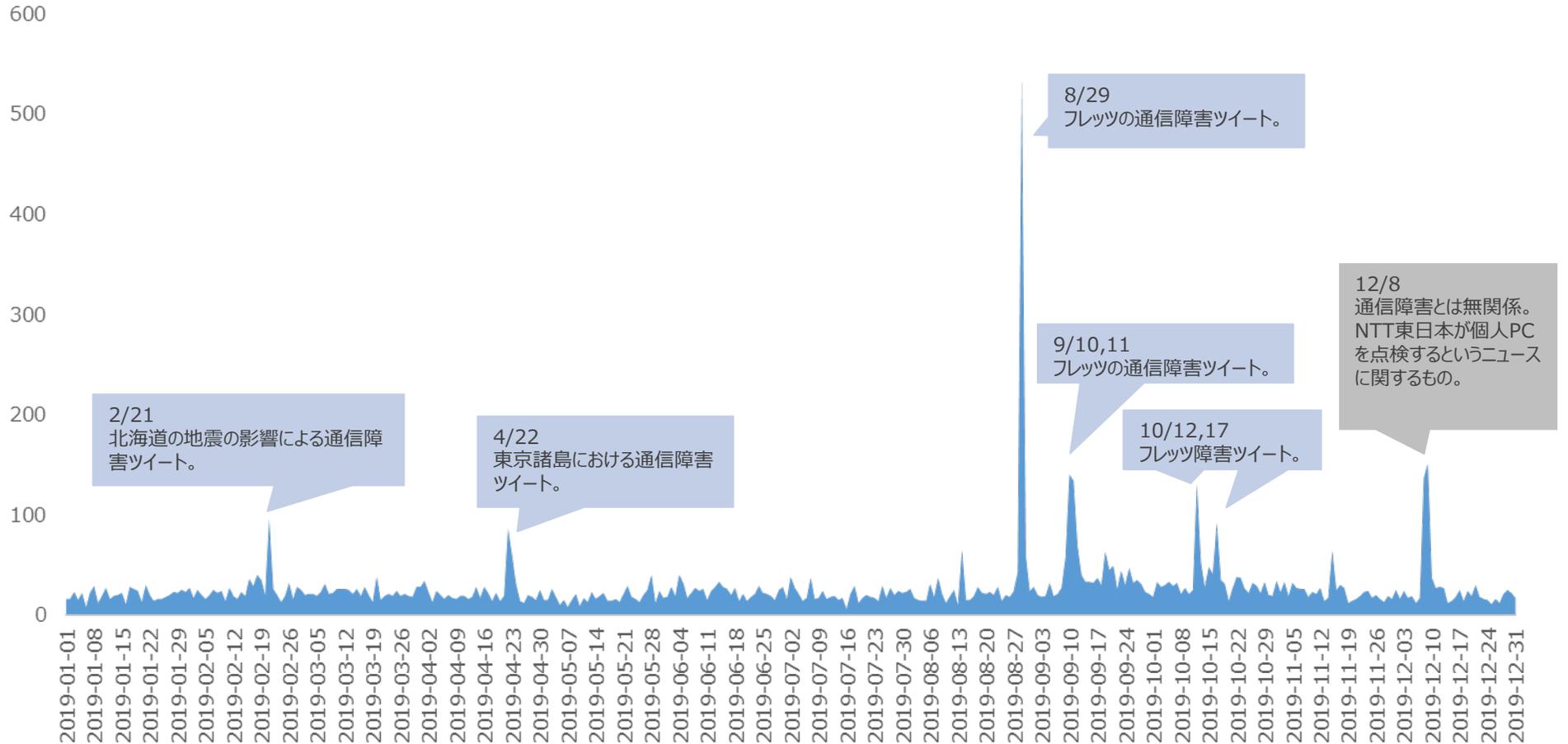
※調査対象30社の合計数

## 4.5.2. 調査結果

### 事業者ごとの障害件数および検知率

No.	事業者名	(1) ツイートまたはニュース等で確認できた障害件数	(2) ツイートから確認できた障害件数	(3) ニュース等から確認できた障害(既知障害)件数	(4) ツイートとニュース等の両方から確認できた障害件数	(5) ツイートのみで確認できた障害(未知障害)件数	(6) ニュース等のみで確認できた障害件数	(7) ツイートにおける既知障害の検知率
1	NTT東	9	6	7	4	2	3	57%
2	NTT西	5	4	3	2	2	1	67%
3	NTTコミュニケーションズ	5	5	2	2	3	0	100%
4	NTTドコモ	7	4	6	3	1	3	50%
5	KDDI	7	6	6	5	1	1	83%
6	ソフトバンク	6	6	2	2	4	0	100%
7	楽天モバイル	16	4	16	4	0	12	25%
8	ジュピターコム	3	3	1	1	2	0	100%
9	ニフティ	4	4	1	1	3	0	100%
10	ビッグロープ	1	1	0	0	1	0	-
11	So-net	16	8	11	3	5	8	27%
12	UQコミュニケーションズ	0	0	0	0	0	0	-
13	中部テレコミュニケーション	1	1	1	1	0	0	100%
14	オプテージ(ケイオプティコム)	4	2	3	1	1	2	33%
15	エネルギア・コミュニケーションズ	1	1	0	0	1	0	-
16	QNet	1	1	1	1	0	0	100%
17	STNet	0	0	0	0	0	0	-
18	インターネットイニシアティブ	3	1	2	0	1	2	0%
19	Google	4	4	3	3	1	0	100%
20	Amazon	4	3	3	2	1	1	67%
21	Facebook	9	8	6	5	3	1	83%
22	Microsoft	5	3	5	3	0	2	60%
23	Twitter	9	9	4	4	5	0	100%
24	LINE	13	5	9	1	4	8	11%
25	yahoo	3	3	1	1	2	0	100%
26	宮崎ケーブルテレビ	4	0	4	0	0	4	0%
27	豊橋ケーブルネットワーク	6	1	6	1	0	5	17%
28	CAC	0	0	0	0	0	0	-
29	秋田ケーブルテレビ	3	3	1	1	2	0	100%
30	TOKAIケーブルネットワーク	16	1	15	0	1	15	0%
	<b>合計</b>	<b>165</b>	<b>97</b>	<b>119</b>	<b>51</b>	<b>46</b>	<b>68</b>	<b>43%</b>

## 4.5.2. 調査結果 - NTT東日本



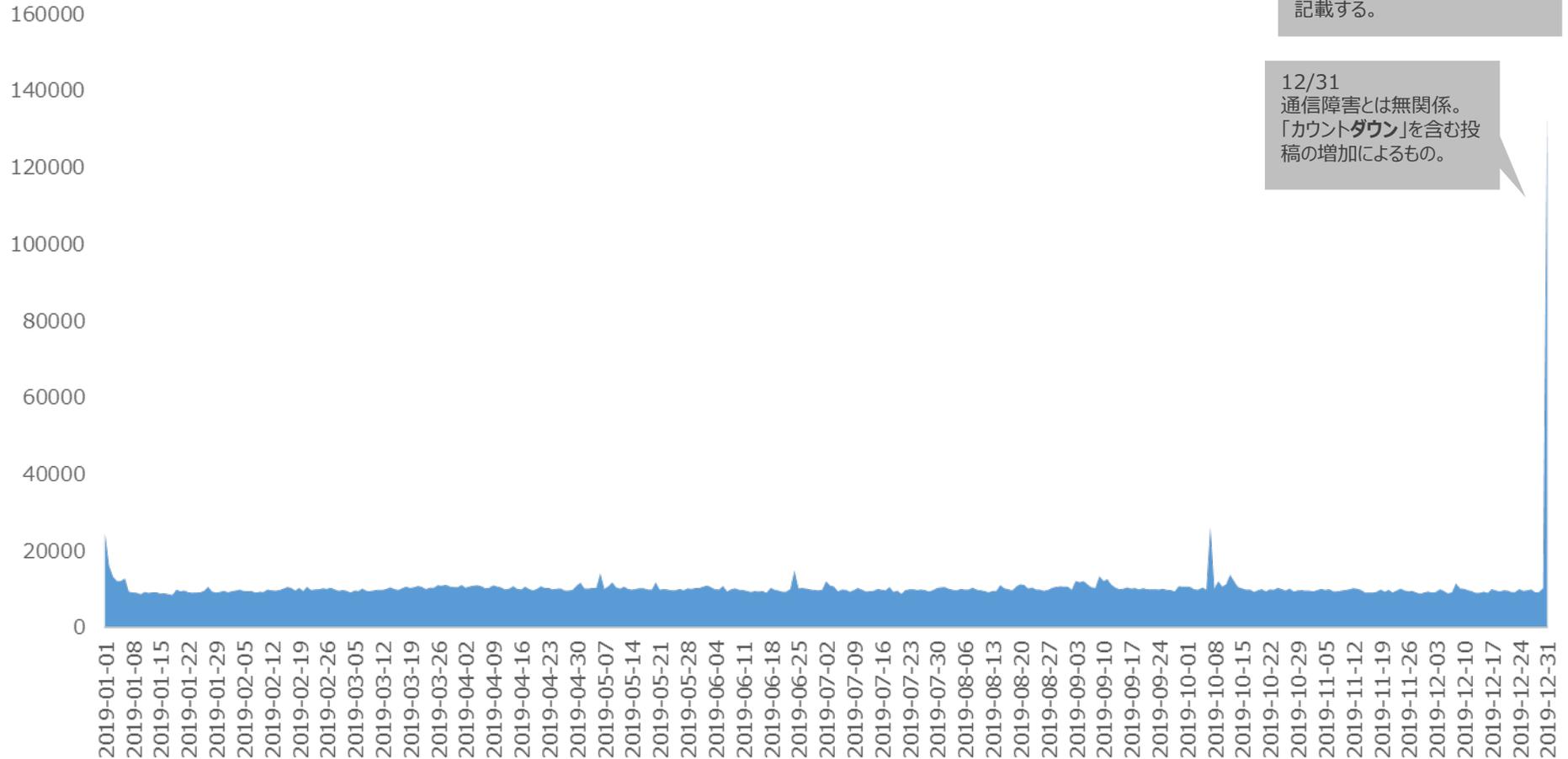
## 4.5.2. 調査結果 -LINE

STEP1  
調査対象事案の選定

STEP2  
データ存在・  
検知タイミング確認

STEP3  
事象抽出モデルの構築

STEP4  
事象抽出モデルの評価



## 4.5.2. 調査結果 -LINE

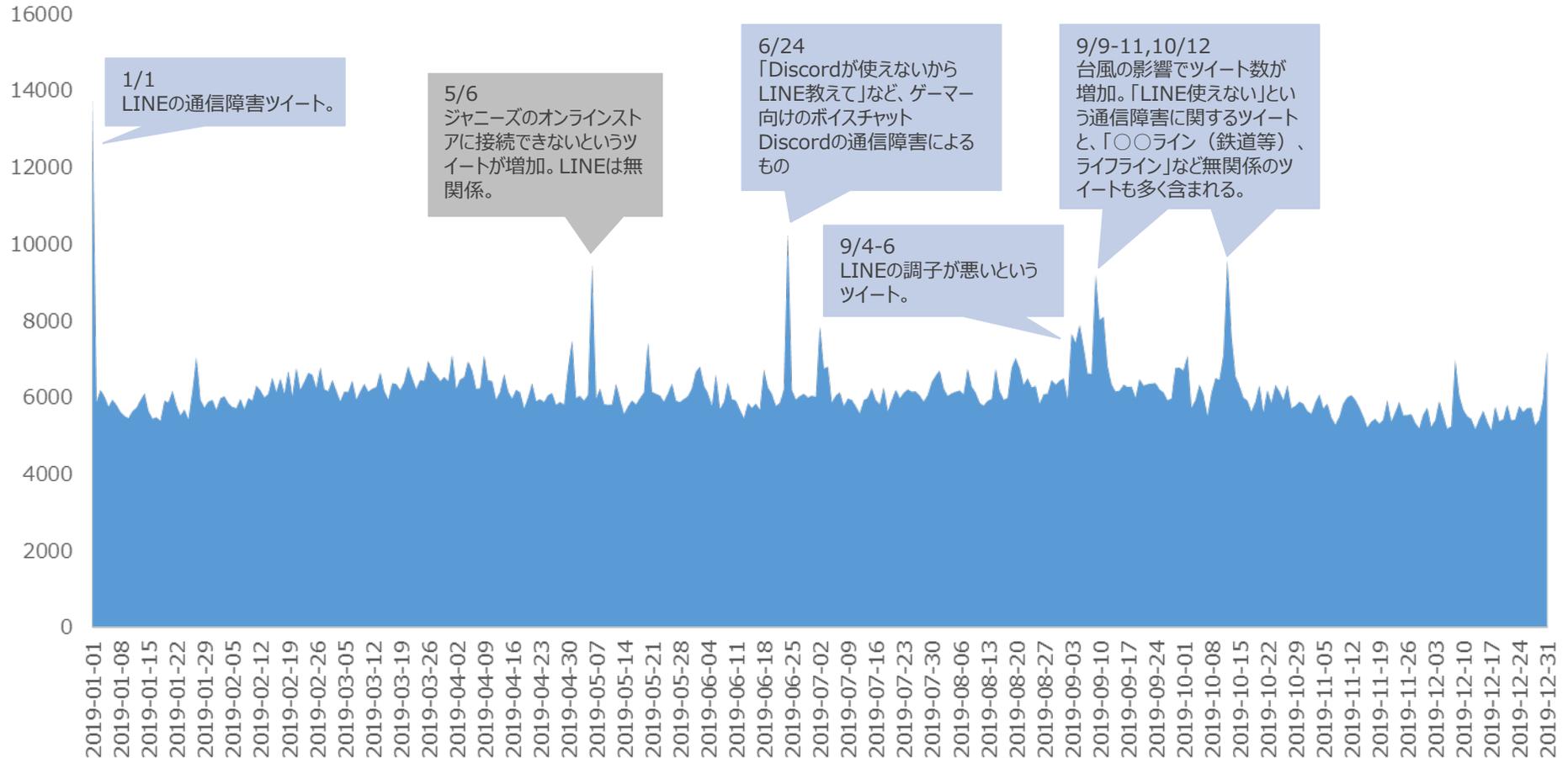
STEP1  
調査対象事案の選定

STEP2  
データ存在・  
検知タイミング確認

STEP3  
事象抽出モデルの構築

STEP4  
事象抽出モデルの評価

※BOT除去



## 4.6. SNSを使用した実証調査のまとめ

## 4.6. SNSを使用した実証調査のまとめ

Twitterを使った実証調査の結果を踏まえて、障害の重要度の把握、ツイートによる障害検知のタイミング、既知障害の検知率、未知障害の検知についてまとめます。

### 障害の重要度の把握

障害発生時のツイートは、「使えない、繋がらない、死んでる」などサービス全体に対して短文で不満を投稿するケースが大半ですが、一部のツイートでは、「メールの取得ができない、ログインできない、経路障害が起きているようだ」など、どの程度の障害かを推測する手掛かりになり得る、詳細な障害内容を投稿しているケースも確認されました。

例えば、Yahoo!メールの障害では、「yahooメールが使えへん。」というサービス利用不可の状況のみを投稿しているツイートがある一方で、「ヤフーメール今開けないか？モバイルはいけたけどPC版無理やねんけど」という、モバイル版は利用できるけれどもPC版は利用できない、といった状況まで詳細に投稿しているツイートも存在します。また、海外事業者の経路設定誤りが起因で発生したNTTコミュニケーションズの障害（P88参照）では、「あれ？OCN死んだ？」というサービス名と障害が起きていることのみを投稿しているツイートがある一方で、「OCNの経路障害が起きているようだ 何故かTwitterは繋がる謎」という、障害原因や一部サイトは繋がるなどの詳細状況を投稿しているツイートも存在します。

このため、**投稿内容を確認することで、障害の重要度をある程度把握することが可能なケースもある**と考えます。

また、AWSの障害では、AWSで障害が発生した数十分後に、AWSを利用していたPayPayにおいても障害が発生しており、インターネット上の広範囲に影響が及ぶ障害の場合、複数の事業者やサービスではほぼ同じタイミングで障害が発生する可能性があることが分かりました。このため、**複数サービスの状態を把握することで、広範囲に及ぶ障害の発生や影響範囲を把握できる可能性**があると考えます。

## 4.6. SNSを使用した実証調査のまとめ

### ツイートによる障害検知タイミング

STEP2の検証において、**障害発生からツイートがされ始めるまでの間隔は平均7分程度**であり、選定した10事案すべてにおいて発生後30分以内に障害に関するツイートが確認できました。障害発生からツイート数の増加が始まるまでの間隔は平均64分であり、事業者公表時刻よりも数分前にツイート数の増加が始まる障害から発生6時間後にツイート数の増加が始まる障害までばらつきが見られました。特に、ツイート数の増加が始まるまでに時間がかかる障害では、ピーク時のツイート数が1分間あたり数件程度であり、ツイート数自体が少ない障害である傾向が見られました。

障害による影響利用者数とツイートの関係では、影響利用者数が多いほどピーク時のツイート数は多く、ツイート数の増加開始までの間隔も短くなる傾向が見られました。また、障害が発生したサービス内容で比較すると、音声通話やチャットの障害などリアルタイムにコミュニケーションをとるサービスにおいて、ピーク時のツイート数は多く、ツイート数の増加開始までの間隔も短くなる傾向が見られました。

上記より、ツイート件数の増加傾向から障害検知を試みる場合、**影響利用者数が一定数以上いるサービスやリアルタイム性の高いサービスにおいては早期検知が可能**と考えます。一方で、ツイート件数の増加傾向のみからの検知方法では、影響利用者数の少ない障害を早期検知することが難しいケースも考えられます。このようなケースであっても、ツイートの中身を確認して**障害発生直後に少量存在する障害関連ツイートを見つけることによって、早期検知が可能**であると考えます。

## 4.6. SNSを使用した実証調査のまとめ

### ツイートによる既知障害の検知

通信事業者30社に対して、今回構築したTwitter障害検知モデルを使った**既知障害（ニュース等で確認できた障害）の検知率は43%**であり、また、各事業者ごとの検知率にはばらつきが見られる結果となりました。**検知率100%の事業者は（既知障害が存在する25社のうち）10社**あり、大手通信事業者から地方ケーブルテレビ局まで規模の異なる事業者においてもツイートによる障害検知が行えることが分かりました。一方で、一部の事業者においては検知率が低い結果となりました。この要因として、

1. サービスの利用者の全体数が少ない
2. サービス規模は大きい障害の影響利用者数は少ない（一部機能、一部地域の障害、短時間の障害など）
3. 企業名、サービス名を含まずに「ネット（メール、電話）繋がらない」とだけ投稿する
4. 平常時のツイート量が多く、障害時のツイートが目立たない（例えばLINEではネガティブな表現と組み合わせたツイートが一日に数万件あるが、その中でLINEモバイルの障害ツイートが数百あってもグラフ上埋もれてしまう）

などが挙げられます。

上記要因のNo.1、2のケースでは、ユーザのツイート自体が無いまたは少ない可能性もあり、この場合、Twitter以外の手法で障害を検知する必要があります。No.3のケースでは、通常1日に300件～400件のツイートが存在していることから、ユーザが自身の利用しているサービス名をあまり意識していないことが考えられ、ツイートによる障害検知が困難な事業者やサービスがあると考えられます。この場合、事業者側からは自社の障害かどうかの判別が難しいため、Twitter以外の情報と組み合わせるなどで障害を検知する必要があります。

No.4のケースでは、障害ツイート自体は存在するため、障害以外のツイートを除外する精度を高めることで検知率の改善が可能です。今回の実証調査では30社に一律で同じ重要語リストを適用して障害検知を試みましたが、今後、検知率を改善していく方法の一つとして、**各事業者のサービス内容に特化した重要語リストを作成することで、障害ツイートのみを抽出し易くなり、検知率を改善できる**と考えます。

## 4.6. SNSを使用した実証調査のまとめ

### 未知障害の検知

今回、構築したTwitter障害検知モデルを使って**未知障害（ツイートからのみ確認できて、ニュースや事業者ホームページからは確認できなかった障害）を46件確認**しました。また、未知障害におけるツイート内容では、既知障害のツイートと同様に「繋がらない、重い、電波が悪い」などの内容が主であり、ツイート内容において既知障害との差異は見られませんでした。

今回確認した未知障害の中に含まれるケースとして

- 事業者が障害を認識できていないサイレント障害
- 事業者は障害を認識しているが公表していない障害
- 利用者側の問題であり障害ではない事象
- 事業者による公表期間が終了して、かつ規模・内容等からニュースでは取り上げられていない障害

などが考えられます。

今回の調査では、確認できた未知障害が上記のどのケースに該当するかは分からないため、本当に障害が起きているかの判定などを含めて今後検討していくべき課題であると考えます。一方で、今回確認した未知障害の中にはサイレント障害が含まれている可能性もあり、ツイートから得られる情報と各事業者が持つ自社の情報とを関連付けていくことで、今まで早期検知が困難であった障害に対して有効な検知方法になり得ると考えます。

## 5. 総括

## 5.1. インターネット障害の把握の在り方についての整理

今回の調査から判明したインターネット障害の発生状況と通信事業者の取組み状況、Twitterの有用性と課題を踏まえて、インターネット障害の把握の在り方について整理します。

### 障害発生状況と事業者の取組み状況

今回調査した通信事業者の8割以上でインターネット障害が発生しており、このうち、NW監視では把握できない障害が約3割を占め、さらにこの約半数が障害発生から把握までに3時間以上の長時間を要しています。しかし、現状の障害把握手法はNW監視と問合せのみの状況であり、人員不足など**運用面の負担の大きさから、SNS等を活用したインターネット障害把握の取り組みは進んでいません。**

### Twitterの有用性と課題

SNSの中でもTwitterは、情報量・即時性・検索性などの観点からインターネット障害の把握に有用であると考えられており、今回**実施した3調査の全てにおいて、有用性を確認**できました。

アンケートではTwitterを障害把握に利用している事業者の8割以上が有用性が高いと回答しています。文献調査では、インターネット障害を検知する手法として定量的に障害判定を行う標準的な手法が存在しており、影響利用者数が多い障害ほど発生後早期に検知できることが確認できました。この点に関してはTwitterを使った本実証調査においても確認されています。

また、NW機器等では検知できないサイレント障害を把握できる可能性があることも本実証調査から示唆されました。

一方で障害把握におけるTwitterの有用性をさらに高めていくためには、下記の課題があることも分かりました。

- **ツイートの定量分析だけでは影響利用者数の少ない障害を早期検知できないケースがある**
- **ツイート内容だけでは障害の真偽や詳細内容を把握できない部分がある**

前者に対しては、ツイートの投稿内容から、障害に関連するツイートかどうかを目視確認するなど定性的に分析することで精度を高める必要があり、後者に対しては、NW監視情報や、利用者からの問い合わせ情報など、通信事業者が持っているSNS以外の情報と組み合わせることで精度を高めることが必要と考えます。

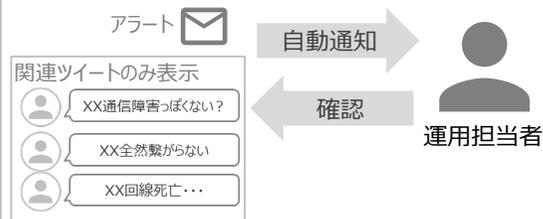
## 5.1. インターネット障害の把握の在り方についての整理

### 今後のインターネット障害把握

障害発生状況と事業者の取組み状況、Twitterの有用性と課題を踏まえると、今後のインターネット障害把握においては、下記に挙げる3要素を満たす仕組みと運用が必要であると考えます。

#### 1. 事業者が低負荷で運用可能な仕組み

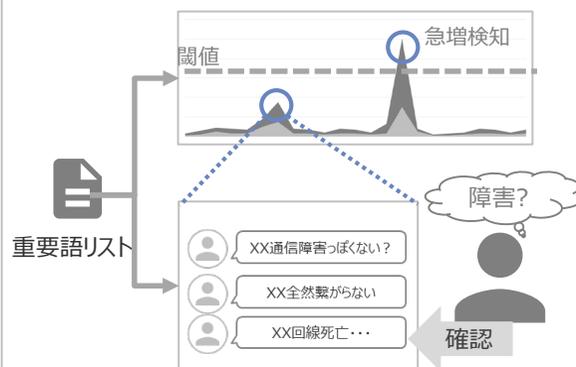
障害の可能性あるツイートを検知した場合のみにアラート通知して、障害関連ツイートのみを確認できるなど、運用負荷を極力少なくする仕組み



事業者側では定常的な運用は発生せずに事象が生じたときにのみ短時間で確認可能

#### 2. 定性・定量の両方からの分析

重要語リストを基にしたツイート量モニタリングによる定量分析とツイート内容の定性分析の両方から障害検知を行う仕組み



ツイート数が少ない障害であっても重要語リストが含まれる障害を目視確認することで障害把握できる

#### 3. SNSと外部データソースの掛け合せ

Twitterから障害が起きている可能性があることを把握しつつ、NW機器や問い合わせなどの情報と紐づけることで障害の真偽、詳細を確認する仕組み



複数の情報を掛け合せて確認することによって情報の真偽、正確性、詳細などを把握できる

## 5.2. インターネット障害の把握の将来像

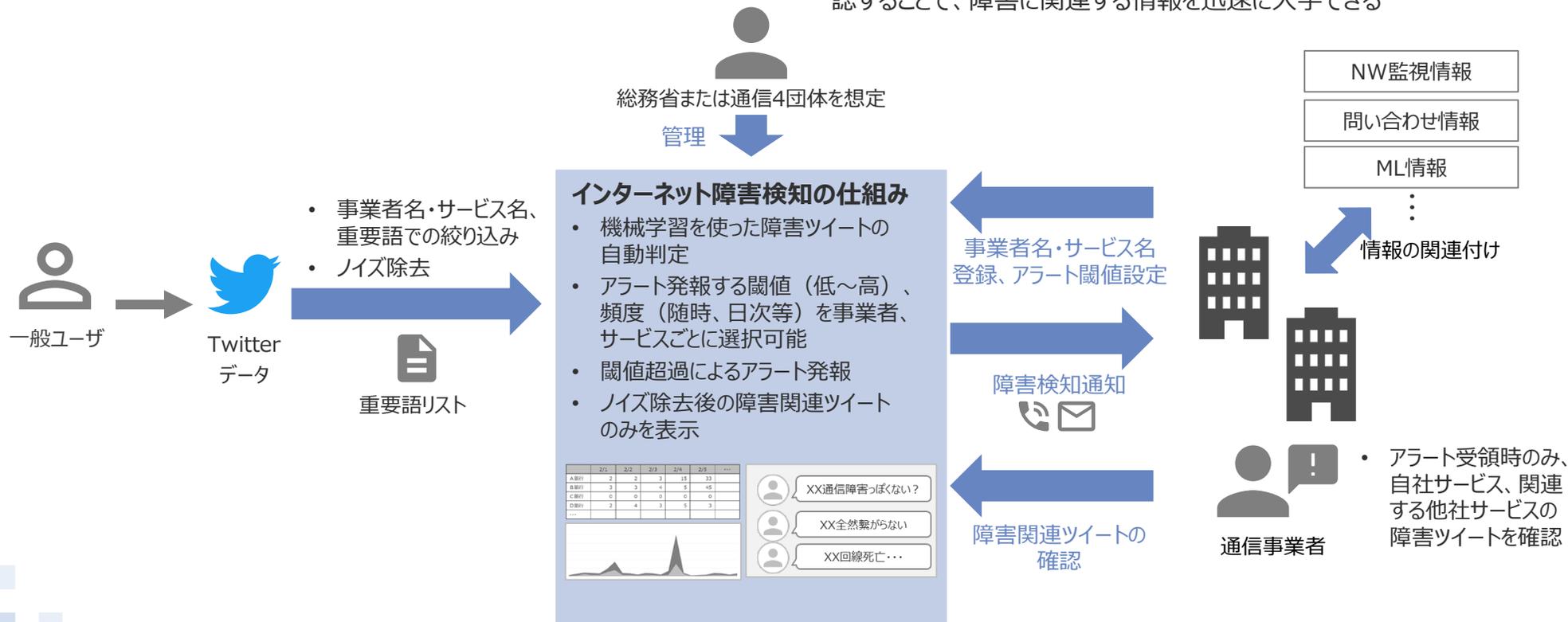
将来的には、複数の通信事業者が共同で利用できるインターネット障害検知の仕組みを構築して、インターネット障害把握に必要な3つの要素である、事業者の運用負荷が低く、SNSを定性・定量の両方から分析可能であり、外部データソースの掛け合わせることが容易な仕組みを構築することで、インターネット障害の迅速かつ的確な把握の実現に繋がっていくと考えます。

### ポイント① 通信事業者が共同で利用可能な仕組み

総務省または通信4団体が管理・運用して、各通信事業者が利用する共同型の仕組みとすることで、事業者のシステム導入や管理・運用に掛かる負担を低減することを可能にする

### ポイント② 通信事業者に合わせた設定が可能な仕組み

アラート発報する閾値や頻度を選択できるようにして、各事業者が許容できる運用負荷に合わせた設定を可能にする。各事業者はアラート通知を受け取ったときのみ、障害関連ツイートに絞り込まれたツイート詳細を確認することで、障害に関連する情報を迅速に入手できる





# NTT DATA

Trusted Global Innovator