Computer Security Symposium 2019
21 - 24 October 2019

SNSデータを用いたセキュリティインシデント早期検知に関する実現可能性検証

榊 剛史[1,3,a)]　大坪 雄平[2,4]　鳥海 不二夫[1]

概要：近年，SNS の普及に伴いユーザの様々な意見や行動の記録が SNS 上に蓄積されるようになった．それに伴い，SNS 上のデータから自然現象や社会現象をはじめとして様々な対象の観測や早期検知を試みている研究が増加している．他方，IT 技術の浸透に伴い，情報セキュリティの瑕疵に起因する情報漏洩や不正利用に関する事例は近年増加の一途を辿っている．そのような背景において，早期にセキュリティインシデントを検知することが重要となる．　本研究では，SNS のデータを分析することで，そのようなシステム・ソフトウェアに関するセキュリティインシデント早期検知の実現可能性について検証を試みる．具体的には，社会的に大きなインパクトがあったセキュリティインシデントの事例について，インシデント発生前後のデータを収集する．収集データについて，自然言語処理・社会ネットワーク分析・時系列解析などのデータマイニングの手法を適用し，セキュリティインシデントの検知に有効となり得る手掛かり情報や特徴量を明らかにすることを目指す．

キーワード：インシデント検知, SNS, ウェブマイニング, 計算社会科学

Feasibility study of security incident early detection using SNS data

Takeshi Sakaki[1,3,a)]　Yuhei Otsubo[2,4]　Fujio Toriumi[1]

Abstract: In recent years, along with the spread on SNS, various opinions and logs of actions have been accumulated by users on SNS. There are many researchers attempt to observe and perform early detection of various objects such as natural phenomena and social phenomena from SNS data. On the other hand, the number of security accidents related to information leakage and fraudulent is increasing in recent years. Under such circumstances, it is important not only to generate accidents, but to detect security incidents early. In this research, we try to perform a feasibility study of early detection of security incidents by analyzing the SNS data. We collect data before and after the occurrence of large security accidents. We apply data mining methods such as natural language processing, social network analysis to collected data to clarify clue information and features that can be effective for early security incident detection.

Keywords: incident detection, SNS, web mining, computational social science

1. はじめに

ソーシャルメディアによる情報拡散が実社会に大きな影響を与えるようになった現在，その正の側面・負の側面の双方が顕在化している．特に近年は，炎上やデマ，フェイクニュース等の負の側面が注目されている．他方，情報システムの普及に伴い，それらに内在するセキュリティインシデントが，世の中全体に与える影響が増加している，そ

1 東京大学
the University of Tokyo
2 警察庁
National Police Agency
3 株式会社ホットリンク
Hottolink, Inc.
4 情報セキュリティ大学院大学
Institute of Information Security
a) sakaki@ipr-ctr.t.u-tokyo.ac.jp

© 2019 Information Processing Society of Japan

のような背景のもと，セキュリティインシデントやセキュリティリスクに関する情報が SNS 上で大規模に拡散され，いわゆる「炎上」する事例は増加の一致を辿っている．ここ最近の炎上した事例は，Coincheck 事件[*1]や「ラブライブ！」公式サイト乗っ取り事件[*2]などが挙げられる．

このようにセキュリティインシデントに関する報告が SNS 上で発信・拡散されるようになったということは，SNS 上のデータを監視・分析することで，セキュリティインシデントの早期発見や影響規模の測定を実現できる可能性があると考えられる．

著者らは，既存研究において，ソーシャルメディア上での情報拡散を分析し，人間が直観的に理解可能な形で可視化する手法を提案している [1], [2]．これらの手法によりソーシャルメディア上での情報拡散を観測可能にすることで，情報拡散を俯瞰的に理解し，分析や対策を容易に行えるようにすることができるようになると考えられる．

そこで，本研究では大規模なセキュリティインシデントの SNS 上での拡散事例を対象とし，それらに既存研究の手法を適用することで，情報拡散を俯瞰的に捉え，さらには早期検知・影響規模推定の手掛かりになる情報が得られるかを検証する．本件では，電子決済の不正利用に関する SNS 上での拡散事例として，2019 年 7 月にサービスが開始され，8 月に終了が発表された「セブンペイ」を対象として分析を行った．また，比較対象として，2019 年 12 月に行われた「ペイペイ」における 100 億円キャッシュバックキャンペーンとそれに伴う不正利用に関する事例の分析も行った，

既存手法では，**図 1** のように Twitter 上の社会ネットワークを階層構造化することで，大規模かつ複雑なソーシャルメディア上の情報拡散を他の可視化手法と比較してより単純に可視化を実現している．本論文では，この Twitter 俯瞰図を元に大規模なセキュリティインシデント事例の情報拡散を可視化し，それらの情報が Twitter 上でどのように拡散していったかを明らかにする．その上で，どのような手掛かりを用いることで，セキュリティインシデントの早期検知が実現できるかを検討する．

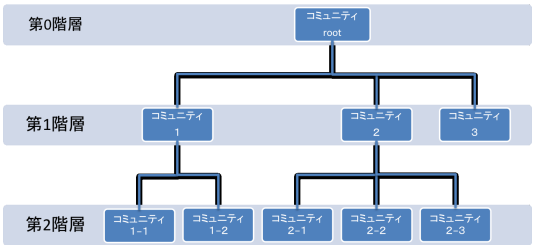


図 1　ソーシャルメディア上の社会ネットワークの階層構造化

*1　https://ja.wikipedia.org/wik#仮想通貨流出事件
*2　https://www.itmedia.co.jp/news/articles/1904/11/news024.html

2.　関連研究

SNS 上での情報拡散とは，実際には社会ネットワーク上での事象であるため，社会ネットワークの可視化手法が適用できると考えられる．社会ネットワークの代表的な可視化手法としては，クラスタリングを用いた可視化手法があげられる．柴田らの研究では，論文の引用ネットワークに Girvan-Newman 法 [3] を適用し，そこから得られるコミュニティ構造を用いて論文ネットワークを可視化している [4]．

しかし，Twitter 上での大規模な情報拡散においては次のような理由から，既存手法を利用することが難しい．

(1) 既存研究では，ノード数がたかだか数万のネットワークを扱っていたため，全ノードを点として表示しても，人間が理解可能な画像サイズに収めることができる．一方，SNS 上での情報拡散に関わるユーザ（ノード）の数は数十万〜数百万規模であり，全ノードを点として表示した場合，人間が理解可能な画像サイズに収めることが難しい．

(2) 論文ネットワークの可視化の場合，ジャンル毎に論文の部分集合を抽出した上で可視化を行うため，可視化を行う度に結果が異なってしまう．そのため，SNS 上での情報拡散に既存手法を適用した場合，各情報拡散の事例同士を比較することが難しい．したがって，比較可能な形で可視化することが求められる．

SNS 上の大規模拡散の可視化が，1. ノード数が大規模である，2. 個別の拡散事例を比較する必要がある，という特徴を持つことから，既存手法の適用が困難である，

著者らの既存研究では，これらの問題点を解消するための新たな可視化手法を提案した．本稿では，この手法を用いて，セキュリティインシデントに関する SNS 上の情報拡散事例を分析し，どのように情報が拡がっていったかを明らかにする．

3.　分析アプローチ

本研究では，Twitter 上の大規模な情報拡散を直観的に理解可能な形で可視化する著者らの既存研究で提案した手法を用いる．本節では，この手法について説明する．

前節において．SNS 上の大規模拡散の可視化には，1. ノード数の大規模性，2. 個別の拡散事例を比較する必要性，という 2 つの問題点があることを述べた．これらの問題点を解決する手段について述べる．

1. については，階層構造を取り入れるで解決することを目指す．複雑系科学においては，「複雑なシステムは一般に階層をもった構造を持つ」とされている [5]．本手法では，この発想を取り入れ，複雑な構造を階層構造により表現することで，直観的に理解可能な形で可視化する

2. については，Twitter 上の定常的な社会ネットワーク構

造を用いることで解決する．Twitter 全体の定常的なユーザ同士のインタラクション情報から社会ネットワークを構築した後，Twitter 全体を俯瞰できるような俯瞰図の作成を行う．俯瞰図上で，様々な情報拡散の事例を可視化することで，個別の情報拡散の事例を比較することが可能である．ただし，Twitter 全体を対象とすることが困難であるため，本研究では日本語 Twitter ユーザを対象とする．

実際に構築した日本語 Twitter ユーザ全体の俯瞰図は**図 2** である．以下では，図 2 を Twitter 俯瞰図と呼ぶこととする．

以下，本節では，Twitter 俯瞰図の見方および構築手順について説明する．

3.1 Twitter 俯瞰図の見方

Twitter 俯瞰図の見方について説明する．

Twitter 俯瞰図とは，図 1 のように階層構造化された Twitter ユーザネットワークのコミュニティ分布を可視化したものである．階層ごとに 1 枚の俯瞰図が作成される．最初のコミュニティを第 1 階層，それ以降の深さ n のクラスタリング結果を第 n 階層と呼ぶこととする．**図 2(a)** は第 1 階層の俯瞰図であり，**図 2(b)** は第 5 階層の俯瞰図である．用いるネットワークは個別の情報拡散事例のデータから構築されたものではなく，ある一定期間の全ユーザ（後述するが，本稿では日本語 Twitter ユーザ）の定常的な情報拡散データから構築されたものである．つまり，Twitter 俯瞰図は，その期間における定常的なユーザの情報拡散関係を表すデータと言える．

図 2(a) は，第 1 階層を表した俯瞰図である．図 2(a) 中の円が一つのコミュニティを表す．また円のサイズは，コミュニティに含まれるユーザ数により相対的に決定される．コミュニティに属するユーザ数が多いほどサイズが大きくなる．また丸ごとにラベルが付与されている．コミュニティ内のユーザプロフィール文から抽出した特徴語が用いられている（ただし，第 1 階層は得られた特徴語をの内容から，人手で再ラベリングを行った）．図 2 ではラベルの文字が読めないが，実際はブラウザで俯瞰図で表示し，マウスオーバーしたコミュニティのラベル名をポップアップで表示する仕様である．円の位置関係は適切に可視化できるように自動的にに決定されたものである．そのため，円同士の近さや隣接性とコミュニティ同士の関係性にはなにも相関性はない．つまり，Twitter 俯瞰図は，Twtiter ユーザの定常的な情報拡散に基づくコミュニティ分布のうち，規模と特徴ラベルを可視化したものである，1 つ下の層では，個別の円，つまり個別コミュニティの中のコミュニティ分布が再帰的に表現されている．

Twitter 俯瞰図に対し，個別の事例を用いて各円の塗色を決定する．各円の塗色は, 対応するコミュニティに属するユーザのうち，個別の拡散事例に関わったユーザの数で相対的に決定される，つまり，色が濃ければ濃いほど，関わったユーザが多いことがわかる．つまり，これが特定の円のみが濃い色で塗られていれば，ある情報拡散に関わったユーザ群はそのコミュニティに偏っていることが分かる．逆に，複数の円が薄い色で塗られていれば，そのユーザ群は複数のコミュニティに広く広がっていることとなる

Twitter 俯瞰図が階層化されているメリットは，全体的な俯瞰とより細部の俯瞰が両方実現できることである．全体を俯瞰したければ，第 1 階層の俯瞰図を見ればよい．より細部を見たければ第 2 階層，第 3 階層と下の層の俯瞰図を見ればよい．

次に具体的事例に基づいて Twitter 俯瞰図の見方を説明する．**図 3** では，著者本人（tksakaki）のフォロワーリストに基づいて，Twitter 俯瞰図に色を塗っている．

まず，第 1 階層について図 3 と図 2(a) を比較すると，「ミドルエイジ」のクラスタに多くのフォロワーがおり，逆にそれ以外のクラスタには殆どフォロワーがいないことがわかる．つまり，著者は日本語 Twitter ユーザ上では、非常に偏った領域のユーザにフォローされていることがわかる．次に第 5 階層について，図 3 と図 2(b) を比較すると，第 5 階層では，1 つのコミュニティの色が濃くなっている．これは「研究科の一覧,21 世紀 COE プログラム・プログラム一覧, 日本学術振興会賞」というラベルを持ったコミュニティである．つまり，著者は「ミドルエイジ」のクラスタの中でも，研究者関連の語をプロフィールに持つユーザに多くフォローされていることが分かる．これより，自分は Twitter の中でも研究者コミュニティに属していると判断することができる．

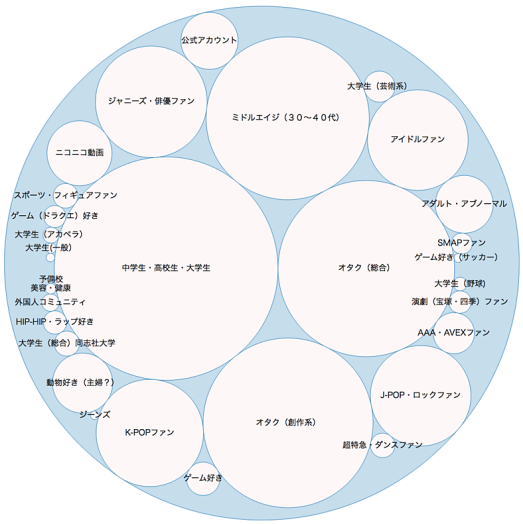
このように Twitter 俯瞰図を用いて個別の情報拡散事例に関わったユーザのコミュニティ分布を階層的に可視化することで，情報拡散の多様性と規模を直観的に理解できると考えられる．

3.2 俯瞰図構築の手順

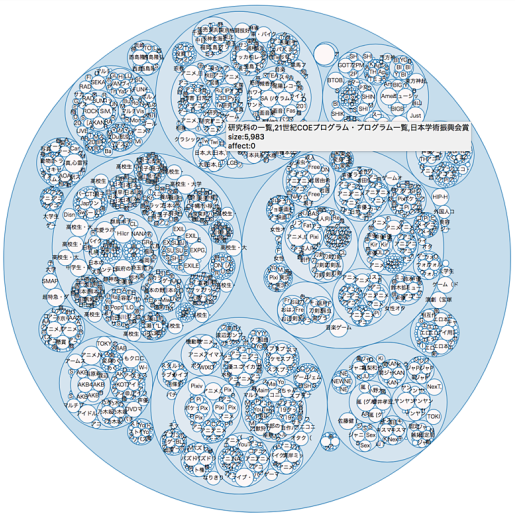
ここでは，Twitter 俯瞰図の構築手順について説明する．大まかな手順は以下の通りである．

(1) 一定期間の Twitter ユーザ同士のインタラクション情報から，Twitter ユーザ間のネットワークを構築する
(2) 構築したネットワークに対して，クラスタリング手法を適用し，ユーザをコミュニティに分割する
(3) 得られたコミュニティについてラベルを付与する
(4) コミュニティ規模が大きい場合，再帰的に (2)，(3) を適用する（再帰的なクラスタリングのことを以下、サブクラスタリングと呼ぶ）
(5) 得られた階層的なコミュニティ構造について, 階層構造に適した可視化手法を適用する

以下，各ステップについて詳細に述べる．



(a) 深さ 1 の Twitter 俯瞰図



(b) 深さ 5 の Twitter 俯瞰図

図 2　Twitter 俯瞰図

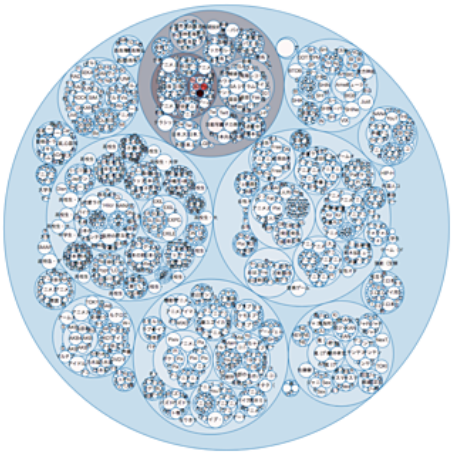


図 3　tksakaki のフォロワー群による俯瞰図

3.2.1　ネットワーク構築

Twitter ユーザの同士のインタラクションからネットワークを構築する．Twitter ユーザネットワークを構築するために手掛かりとして用いられるインタラクションはいくつかあるが，本稿では，リツイート情報を用いる．リツイートとは，1 クリックで，他のユーザの投稿をコピーし，元投稿のユーザを明示した上で，自身に向けてリンクを張っているユーザに対して再発信できる機能である．その手軽さ故に多くのユーザが使っている機能であり，またデータ収集も容易であることから，リツイート情報を用いることとした．

3.2.2　クラスタリング

本手法では，社会ネットワークのクラスタリングに用いられる Modularity-Optimization の手法を採用する．Modularity Q とはクラスタの結合度合いを表す指標であり，式 1 で表される．これが高ければ高いほどネットワーク全体のよくクラスタリングされていると考えられる．本手法では，既存研究で使われていた Louvain 法を用いる [6]．式 1 において，$e_{i,j}$ は「コミュニティ i, j に属するノード同士が繋がるリンク数の全リンク数に占める割合」，a_i^2 は $e_{i,i}$ の期待値を表す．

$$Q = \sum_i \left(e_{i,i} - a_i^2\right) \tag{1}$$

クラスタリングにより得られるクラスタをコミュニティとして抽出する．

3.2.3　コミュニティへのラベル付与

本稿では，各コミュニティに関連するドキュメント集合をコミュニティ文書と見なし，各コミュニティ文書に特徴的な語を抽出し，それによりクラスタをラベリングする．

最初に各コミュニティに含まれる全ユーザのプロフィール文を結合して，コミュニティ文書とした．

次に同じ階層にいる全コミュニティのコミュニティ文書を全文書集合として，term frequency - inversed document frequency(tf-idf) 法 [7] を適用することで，各コミュニティ文書に特徴的なキーワードを抽出し，コミュニティのラベルとして用いた．tf-idf 法とは，自然言語処理でよく使われる特徴語抽出手法であり，全文書集合に対して，個別の文書に特徴的な語を抽出する手法である．

ただし，第 1 階層・第 2 階層のコミュニティはサイズが大きいものが多く，説明力が高い特徴語を付与することが難し買ったため，得られた他特徴語から人手で再度ラベルを付与した．

3.2.4　サブクラスタリング

本節では，サブクラスタリングについて説明する．(2) において RT ネットワークからクラスタを抽出したが，可視化の観点からは，得られたコミュニティはサイズが大きく，これだけでは不適であった．

表 1　セブンペイを巡る一連の経緯

年月日	出来事
2018 年 6 月	スマホ決済サービス運営会社「セブン・ペイ」設立
2019 年 4 月 4 日	セブンペイを 7 月 1 日に開始と発表
7 月 1 日	セブンペイサービス開始
7 月 2 日	コールセンターに不正利用に関する問い合わせ
7 月 3 日	社内調査で不正利用を確認．入金を停止
7 月 4 日	被害者数・被害額の試算を公表
7 月 8 日	金融庁がセブン・ペイに報告要求
7 月 11 日	外部 ID からのログインを遮断
7 月 30 日	グループ共通 ID「7id」のパスワードを強制リセット
8 月 1 日	セブンペイのサービス終了を発表

そこで，1 回目のクラスタリングで得られたコミュニティのうち，サイズが閾値 th_{clsize} を超えるものについては、再度クラスタリングを行う．そうして 2 回目のクラスタリングで得られたクラスタのうち，サイズが閾値 th_{clsize} を超えるものについては再度クラスタリング行う，という手順で，再帰的にクラスタリングを適用する．

これにより一定規模以下のコミュニティが大量に得られることとなる．ただし，深さは最大 th_{depth} になるようにし, それ以上はクラスタリングを適用しないこととした．本稿では，$th_{depth} = 5$，$th_{clsize} = 1000$ とした．

3.2.5　可視化

本稿では，階層構造の可視化手法の筆頭である Circle Packing を用いることとした．具体的には，階層構造を可視化するためのパッケージである．ここでは D3.js に準備されている Circle Packing のモジュールを用いて可視化を行った[*3]．

4.　事例分析

本研究では，Twitter 俯瞰図を利用して，SNS 上でのセキュリティインシデントに関する情報拡散の分析を行う．

今回は，特に電子決済の不正利用に関する情報拡散を事例として分析をおこなった．具体的には，セブンペイおよびペイペイに関する情報拡散を対象とする．

4.1　利用データ

今回は，主たる分析対象をセブンペイのサービス開始から 4 週間弱の情報拡散，比較対象をペイペイのキャッシュバックキャンペーン前後の情報拡散として，Twitter の投稿を収集した．セブンペイを巡る一連の経緯を**表 1**，ペイペイのキャッシュバックキャンペーンを巡る一連の経緯を**表 2** に示す．

主として，サービスリリースから直近までのセブンペイに関する投稿を全て収集した．また，比較事例として 2018 年 12 月に発生したペイペイにおける不正利用に関連して，2018 年 12 月のペイペイに関するデータも収集した．ただし，データ取得方法の制約上，直近しか全投稿を取得できないため，ペイペイに関連した投稿は，10%サンプリング

[*3] https://beta.observablehq.com/@mbostock/d3-circle-packing

表 2　ペイペイ 100 億円還元キャンペーンを巡る一連の経緯

年月日	出来事
2018 年 10 月	PayPay アプリをリリース・サービス開始.
12 月 4 日	100 億円還元キャンペーン 第 1 弾開始.
12 月 13 日	100 億円還元キャンペーン 第 1 弾が終了.
12 月 17 日	Paypay 広報部がセキュリティリスクの存在を認める
12 月 18 日	セキュリティアップデートを行ったと発表
1 月 21 日	3D セキュアによる認証が導入

データからデータを抽出した．以下，簡略化のためにセブンペイに関するデータセットを「7pay」，ペイペイに関するデータセットを「paypay」と記述する．**表 3** にデータセットの概要を示す．

また，Twitter 俯瞰図の構築には，一定期間の Twitter ユーザのネットワーク情報が必要になる．本稿では下記のデータを用いて，Twitter ユーザネットワークの構築・コミュニティの抽出および Twitter 俯瞰図作成を行った．

投稿期間　2017 年 5 月 1 日〜7 月 31 日
利用データ　日本語 10%サンプリング
抽出投稿　公式 RT すべて

4.2　全体分析

まず，7pay・paypay の事例について，収集したデータ全体の傾向および不正関連した投稿に関する傾向を分析する．これにより，電子決済サービスとその不正利用に関心があるユーザ層の違いを明らかにしたい．

本稿では，収集したデータのうち「不正」を含む投稿を不正利用について言及したと推測される投稿（以下，不正関連投稿）として抽出した．以下，件数比較を行うために，7pay のデータの件数を 1/10 として比較を行う．

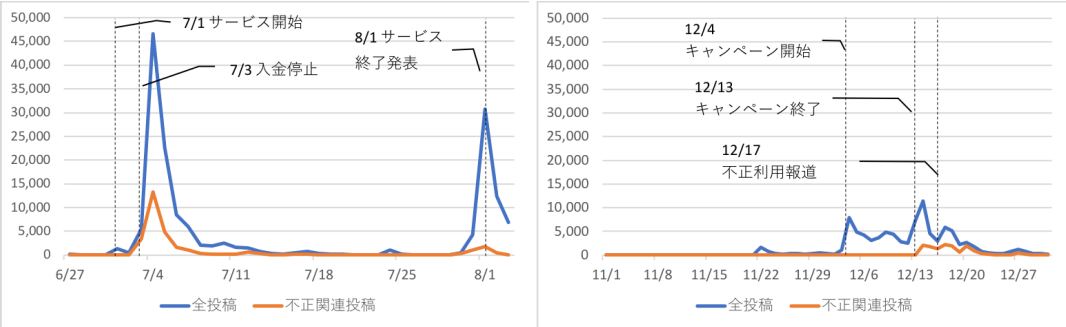
表 4 に 2 事例の全体件数および不正関連投稿の全件数を示す．**図 4** にそれらの時系列推移を示す．なお，本文投稿の収集が間に合わなかったが，時系列推移データのみ 2019 年 7 月 27 日〜2019 年 8 月 3 日までの 7pay データの件数を追加した．

表 4 の全投稿に着目すると，7pay と paypay でデータ収集期間に違いがあり，完全に条件が揃っていない面はあるが，この 2 事例について，事例全体の盛り上がりは同程度であったことが分かる．一方，不正関連投稿に着目すると，2 倍程度の開きがある．これより，paypay と比較して，7pay では不正利用により多くの注目が集まったと推測される．

図 4 より，2 事例の盛り上がりには異なる傾向があることが見てとれる．paypay については，キャンペーン開始日の 12 月 4 日から終了日の 12 月 13 日まで，日々一定量の投稿がある．一方，不正利用が報道された 12 月 17 日も含め，不正関連投稿についてはそれほど大きく盛り上がっているようには見えない．7pay については，サービス開始日当初はあまり盛り上がらず，不正利用が報道された 7 月 3 日とその翌日 7 月 4 日に，不正関連投稿の増加と共に大

表 3　データセットの概要

事例	検索キーワード	期間	取得方式	件数
7pay	7pay OR セブンペイ OR 7 ペイ OR セブン pay	2019.6.27-2019.7.26	全量	1,115,368
paypay	paypay OR ペイペイ	2018.11.1-2019.12.31	10%	93,090



(a) 7pay　　(b) paypay

図 4　投稿数の比較

表 4　データセットにおける全件数

事例	全投稿	不正関連投稿
7pay	111,537	26,473
paypay	93,090	14,379

きく盛り上がっているように見える．また，サービス終了が発表された 8 月 1 日も大きく盛り上がったが，不正利用が報道された翌日ほどではなかった．

つまり，paypay についてはキャンペーン開始当初からキャンペーンへの注目が高く，その背景のもと，不正利用にも一定の注目が集まっていたと言える．他方，7pay については，サービス開始当初にはそれほど注目が集まらず，不正利用により一気に注目度が高まったと言えるだろう．ただし，paypay と比較して，7pay はサービスが想定通りに運用されていた期間が短く，公正な比較が難しいため，これらはあくまで参考となる知見としてとどめておく．

次に**図 5** および **図 6** に 2 つの事例に関する Twitter 俯瞰図（全投稿と不正関連投稿）を示す．すると，2 事例ともに類似したコミュニティに属するユーザ群がデータセット全体の投稿に寄与していることがわかる．特に不正関連投稿については，色が濃い，多くのユーザが属するコミュニティは第 3 階層まで一致している．これより，電子決済に興味を持つユーザ，特にその不正利用に興味を持つユーザは Twitter 上で特定のコミュニティに属することが示唆された．

ただ，5, 6 だけでは，どのようなデータを入れても同じような Twitter 俯瞰図が得られる可能性を排除できない，そこで，7pay のデータセットと同じ収集期間に全く異なるキーワード「サッカー」を含む投稿を行ったユーザ群の分布を**図 7** に示す，図 7 よりデータセットによって，ユーザ分布が異なることが示唆された．

2 つの事例の全投稿において，拡散に寄与したユーザの多くが所所属しているコミュニティを**表 5** に示す．2 つの事例共に「ミドルエイジ」「オタク（創作系）」「オタク（総合）」というラベルを持つ第一階層のコミュニティに属す

表 5　投稿に寄与している上位コミュニティ（第 2 階層）

順位	7pay		paypay	
	第一階層	第二階層	第一階層	第二階層
1	ミドルエイジ	一般	ミドルエイジ	一般
2	オタク（創作系）	イラスト趣味	オタク（創作系）	イラスト趣味
3	ミドルエイジ	政治・社会 1	オタク（総合）	女性
4	オタク（総合）	女性	ミドルエイジ	政治・社会 1
5	オタク（創作系）	音楽ゲーム	ミドルエイジ	政治・社会 2

るユーザが拡散に寄与していることが分かる．

つまり，「コンビニの電子決済について興味を持っているユーザ」「コンビニの電子決済の不正利用について興味を持っているユーザ」というのは特定のコミュニティに偏っており，またそれらのユーザは似通っていると言えるだろう．

4.3　不正利用被害投稿に関する分析

次に，同じく 2 事例について，不正利用の被害に遭ったと思われるユーザが直接おこなった投稿（以下，不正利用被害投稿）を抽出し，その拡散状況について分析を行う．

本稿では，不正利用被害投稿を下記の様に定義する．

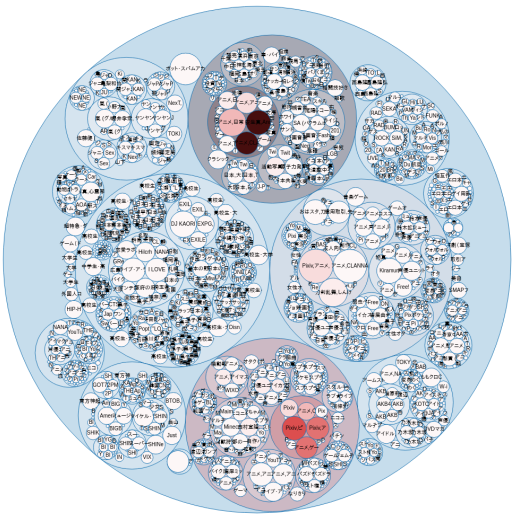
- 第三者から聞いた伝聞ではなく，投稿者本人が直接被害にあったと推測される内容である
- 具体的な被害金額が記載されている
- 不正利用が最初に報道されるよりも前に投稿されている

4.3.1　不正利用被害投稿の抽出

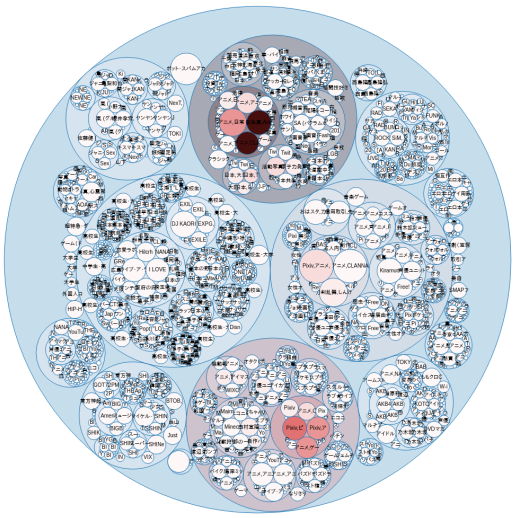
今回は，ナイーブな手法で不正利用被害者によると思われる投稿を抽出した．抽出方法は下記の通りである．

(1) 不正利用が最初に報道された日以前から「不正」「万円」を含む投稿を抽出
- paypay：2018 年 12 月 14 日以前（最初の報道は 2018 年 12 月 14 日）
- 7pay：2019 年 7 月 3 日以前（最初の報道は 2019 年 7 月 3 日）

(2) 抽出した投稿のうち，リツイート・メンションを除去

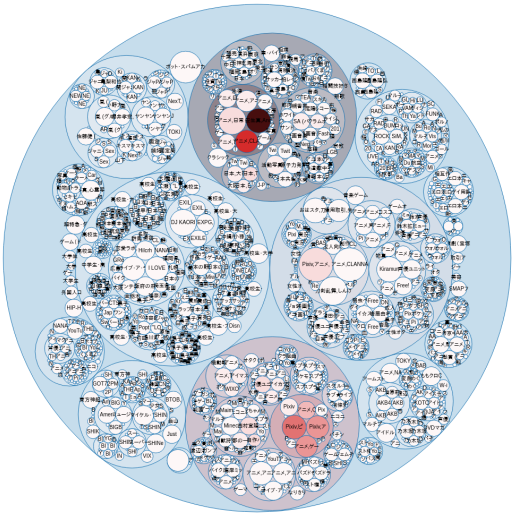


(a) 7pay 全投稿

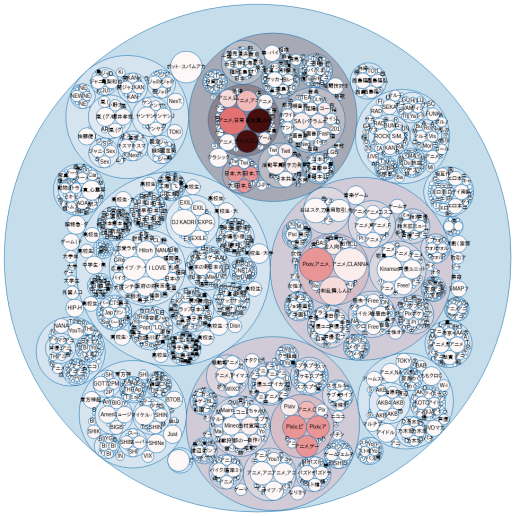


(b) 7pay 不正関連投稿

図 5　全投稿・不正関連投稿の分布: 7pay



(a) paypay 全投稿



(b) paypay 不正関連投稿

図 6　全投稿・不正関連投稿の分布: paypay

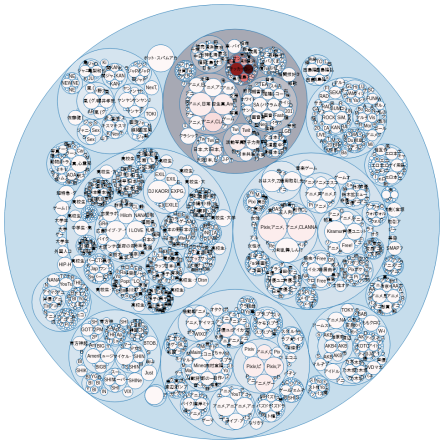


図 7　「サッカー」を含む投稿を行ったユーザの分布

(3) 人手で不正利用被害投稿であるかを判定

抽出された不正利用被害投稿の例を以下に示す．なお，個人情報，著作権等を考慮し，そのままの内容ではなく，投稿本文を一部改編し，投稿の大意を伝わる形で提示するにとどめる．

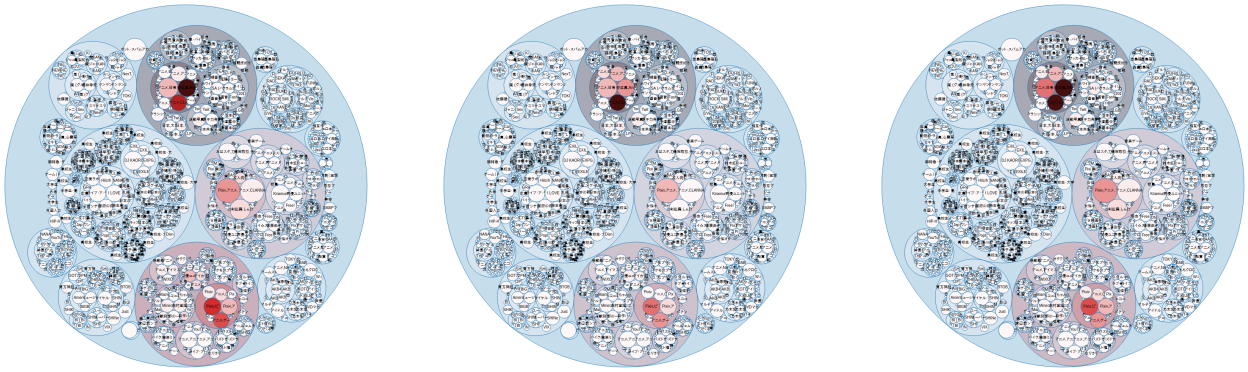
- 7pay

  その 1(1835RT)　7pay で不正利用あり。XX 万円の不正チャージ・利用。#7pay #サイバー犯罪

  その 2(773RT)　#7pay　に登録したら、XX 万円不正にチャージされてしまった

  その 3(5208RT)　7pay 不正利用問題。自分も同じ事態が発生して急遽パスワードを変更しましたが、XX 万円のチャージがされてしまいました。問合せをしたら、「不正利用の記録はありません。カード会社が原因では」という当事者意識の無い対応。危険過ぎる。再調査を依頼。

- paypay
  - PayPay 経由で自分のクレカが不正利用された模様



(a) 不正利用被害投稿 その 1(1835RT)　(b) 不正利用被害投稿 その 2(773RT)　(c) 不正利用被害投稿 その 3(5023RT)

図 8　不正利用被害投稿の拡散状況：7pay

– 僕のクレジットカード、ペイペイ経由で不正利用されてた.XX 万の被害.

これらの不正利用被害投稿について，それぞれ拡散状況の可視化を行う．しかし，paypay データセット内では，不正利用被害投稿について 100 件以上拡散された投稿は見当たらなかった．これは paypay データセットが 10%サンプリングデータから抽出された内容であるため，十分な量が得られなかったと思われる．今回用いる手法では，100 件未満の拡散では，正確な可視化が行えないため，paypay については可視化を省略する．7pay については，7pay データセット内で抽出された不正利用被害投稿のうち，十分に拡散されていた 3 件の投稿（先に提示した「その 1」「その 2」「その 3」の投稿）について，それぞれ 2019 年 7 月 13 日までの拡散状況の可視化を**図 8** 行った．

図 8 より，対象とした 3 つの投稿が拡散したユーザ群の分布は類似していることが明らかになった．特に 30 代〜40 代の会社員が多く含まれる「第一階層：ミドルエイジ，第二階層：一般」の中のさらに 1,2 個に属するユーザが偏って情報拡散に寄与してたことがわかる．

また 3 つの投稿全ての拡散に寄与したユーザを調べたところ，全部で 66 名のユーザが抽出された．これらのユーザは，「電子決済の不正利用に強い興味を持ったユーザ群」であることが推測される．このようなユーザ群がどのような特徴を持つかは今後明らかにしていきたい．

5. おわりに

本研究では，セキュリティインシデントに関する大規模な情報拡散について，事例分析を行った．今回は，特に電子決済の不正利用の 2 つの事例を対象として分析を行ったが，2 つの事例において，電子決済について投稿を行ったユーザ群，電子決済の不正に関連した投稿を行ったユーザ群は，ソーシャルグラフ上で特定のコミュニティに所属していることが明らかになった．また，3 つの不正利用の被害報告について拡散したユーザ群も同様のコミュニティに属していることが明らかになった．さらに 3 つの被害報告全ての拡散に関わっているユーザが一定数いることが明らかになった．これより，Twitter 上には，電子決済の不正利用に強い興味を持ったユーザ群がいることが推測される．このようなユーザ群の投稿を監視しておくことで，早期にセキュリティインシデントがで検知できる可能性がある．

謝辞　本論文は科研費 2019 年度 若手研究（課題番号 19K20413）の助成を受けた．

参考文献

[1] 鳥海不二夫，　榊剛史：バースト現象におけるトピック分析，情報処理学会論文誌，Vol. 58, No. 6, pp. 1287–1299 (2017).

[2] 榊　剛史，鳥海不二夫，大知正直：ソーシャルメディア上の大規模情報拡散に関する俯瞰的可視化手法の提案，人工知能学会全国大会論文集，Vol. JSAI2019, pp. 2D4OS1a01–2D4OS1a01 (2019).

[3] Girvan, M. and Newman, M. E. J.: Community structure in social and biological networks, *Proceedings of the National Academy of Sciences*, Vol. 99, No. 12, pp. 7821–7826 (2002).

[4] Shibata, N., Kajikawa, Y., Takeda, Y. and Matsushima, K.: Detecting emerging research fronts based on topological measures in citation networks of scientific publications, *Technovation*, Vol. 28, No. 11, pp. 758 – 775 (2008).

[5] ハーバート・A. サイモン：システムの科学〔第 3 版〕，パーソナルメディア (1999).

[6] Blondel, V. D., Guillaume, J.-L., Lambiotte, R. and Lefebvre, E.: Fast unfolding of communities in large networks, *Journal of Statistical Mechanics: Theory and Experiment*, Vol. 2008, No. 10 (2008).

[7] Manning, C. D. and Schütze, H.: *Foundations of Statistical Natural Language Processing*, MIT Press (1999).