

# 児童生徒を取巻く情報社会・インターネット関連犯罪の 分類の試み

阿 濱 志保里

(受付 2021 年 10 月 27 日)

## 1. はじめに

総務省の平成26年末の情報通信機器の普及率は世帯あたり、「携帯電話・PHS」は94.6%、「パソコン」は78.0%となっている。「携帯電話・PHS」の内数である「スマートフォン」については前年よりも1.6ポイント増加の64.2%と急速に普及が進んでいる [1]。これらのようにインターネットが児童生徒にとって身近なものになっている一方で、インターネットを利用することによる犯罪に巻き込まれる事例が増加傾向にある [2]。

安全にインターネットを利用するためには、情報安全教育の充実が求められるのはもとより、児童生徒を取巻くインターネット社会での出来事を定量的に把握し、その傾向や特徴を指導者が理解しておく必要があると考えられる [3]。また、多様化し、複雑化する犯罪やトラブルに対して、それぞれの特徴を正確に理解しておくことにより、情報安全の学習指導のカリキュラム開発や指導適時性の検討が進むことが期待できる。これまでに情報モラル教育分野を中心にインターネットに関連する犯罪や、犯罪に類するトラブルに関する研究は多岐にわたり行われてきている。情報モラル教育黎明期に高橋ら [4] は、情報社会の出来事について報道記事を基準にカテゴリ分析を試みている。その後、携帯電話などを用いた犯罪に巻き込まれないための教材開発や授業実践は事例が積み重ねられてきており、出会い系などコミュニケーションに起因するトラブルについても指導が行き渡り初めている [5]。しかし、情報通信機器や情報通信サービスが進化した後でも、それらに対応した犯罪の分類に関する研究は見当たらず、情報モラル教育を体系的に行っていくためには、情報基盤の種類に応じた分析を深化させる必要があると考えられる。そこで、本研究では平成28年度に発生した児童・生徒が関与するインターネットにおけるトラブルや犯罪を定量的に分析し、その傾向を把握する試みを行った。

## 2. 分析方法

### 2.1 概要

情報社会における諸々の出来事について、情報モラル教育の先行事例では、情報社会に関連

する法律やセキュリティ、コミュニケーションなどマナーやモラルに関する内容に分類されており、情報モラル指導モデルカリキュラムでは、法律やセキュリティ、倫理観や知識、情報社会に参画する意識など5つのカテゴリに分けて示されている [3]。それらの分類作業は、それぞれの経験知などに頼っているのが現状で、主観性を排除することは難しいと考えられる。

## 2.2 データの収集

本研究では、先行研究を参考にしつつ定量的に分析を行うために、情報社会の出来事について報道記事を基準にテキストマイニングによってカテゴリ分析を試みた。実際のデータは平成28年中にインターネットのニュースサイトで報じられた児童・生徒が関与した事件やトラブルを1000件収集した。情報収集期間は、2016年1月1日から12月31日とした。警察庁発表における平成28年度におけるサイバー犯罪の検挙件数は8,324件であり [6]、統計的に妥当性があると判断される1000件の記事を選択した。重複した記事については、削除した。情報の検索については、代表的な検索エンジンである Yahoo, Google, msn, goo を活用し、情報に関わる犯罪やトラブルについてのニュースを網羅的に検索した。ニュース記事が本研究の材料として適しているかの判断については、すべてのケースについて著者3名で妥当性を評定した。評定には、3名の評定値の相関は $r=.8$ であった。評定値の違いは3以上のケースについては、評定者による合意を行った。

## 2.3 分析の手続

ニュースサイトの本文で報じられた記事を対象に計量テキスト分析を行なった。計量テキスト分析では、先行研究 [7] を参考に形態素解析を用いて文章から単語を抽出した。分析には、日本語テキスト型データの分類に適したシステムとして KHCoder [8] を用いてニュース記事の記述について分析した。KHCoder は語彙の選択にあたり恣意的となり得る「手作業」を排し、多変量解析によってデータ全体を要約・提示することと、コーディング規則を公開するという手順を踏むことによって、操作化における自由と客観性の両立を可能にする。本研究においても操作の詳細を明示・公開した上で多変量解析によるデータの要約・提示を行なうことで、客観性を確保しつつ、対象データの特徴をとらえることを試みた。本研究では、ニュース記事で取り扱われている単語の表出傾向をもとに、児童生徒とインターネットに関連した犯罪やトラブルについて探索的な研究としての位置付けとした。

本研究では調査から得られた文章の記述データのテキスト化した後、計量テキスト分析により形態素解析を適用し、抽出されたキーワードから検討・考察を行なった。対象とした統計量は総抽出語数250,652語、異なり語数13,737語（文章：8,667文、段落：3,292段落）となった。

次に、調査から得られた文章の記述データによる共起ネットワーク分析を試みた。共起ネッ

トワークでは、出現パターンの似通った語、すなわち、共起程度の強い語を線で結んだネットワークを描くことができる語と語が線で結ばれているため、多次元尺度法のように単に語を布置する手法よりも、解釈が容易である。そのため、関係性の程度や文章の中で記述される回数によって、グループ分けし、表出された語彙の関係性については、出現数が多い語ほど大きく、共起の程度が強いほど太い線で描画することができる。今回は、語彙間の関係性を確認するため距離や配置のみに着目した。共起関係の強弱については、分析対象となった語のすべての組み合わせについて、集合間の類似性を表す Jaccard 係数を用いて計算した。Jaccard 係数の値を0.2以上として得られた出現パターンに類似性・共起関係があるものに対して線を用いてネットワークを描画した。描画には、表出した総語彙のうち、出現回数が200回を基準として抽出し40語を対象とした。

その後、カテゴリカルデータの解析方法の1つである対応分析を行なった。対応分析は、データ行列が十分に大きく簡単な統計分析ではデータの構造がわからないときや、変数が同質かつ行または列間の統計的距離を計算する意味があるとき、データ行列の行と列の幾何的図形による解釈が可能であり、分析を容易にし、関連性の探索に役立つ利点がある。

### 3. 結 果

#### 3.1 結 果

##### (1) 頻度の高い語彙

ニュース本文から得られたテキストに基づき、分析を行った。テキストマイニングでは、頻度の高い語彙の抽出を行った。結果を表1に示す。その結果、「攻撃」や「サイバー」「データ」などのデータの搾取を目標とした攻撃性に関わる語彙が上位に見られた。

##### (2) 語彙間の関係性

次に、抽出された語彙の関係性を明らかにするために、共起ネットワーク図の作成を行なった。図1に示す。

その結果、7つのカテゴリに分類されることが分かった。7つのカテゴリは、構成される語彙から見て、犯罪等に関わるカテゴリは「セキュリティに関わるサイバー犯罪」「ウイルス感染」「不正アクセスに伴う情報流出」「電話番号等の流出」であった。

##### (3) 語彙の分類

次に、質的データを分析する多変量解析法の1つである対応分析を行なった。クロス集計結果を用いて、行の要素と列の要素の相関関係が最大になるように数量化して、行の要素と列の要素を多次元空間に表現するもので、語の相関関係が高いほど近くにプロットされる特性がある。抽出された語彙を基に、対応分析を行なった。結果を図2に示す。



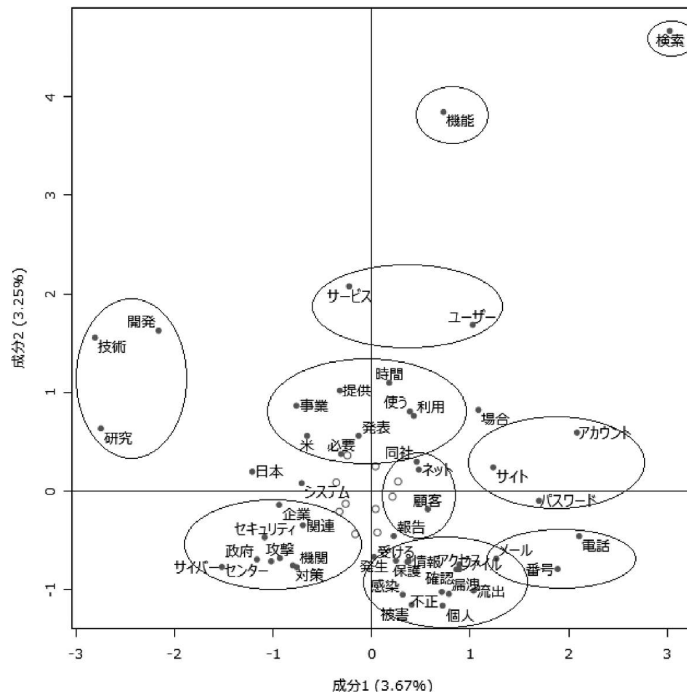


図2 対応分析の結果

その結果、特徴的な10のカテゴリが抽出された。1つ目は「開発」「技術」「研究」の表出が確認された。「開発」と「技術」を含む記述を抜粋すると、「膨大な情報を組みあわせて解析する技術の開発」などの記述が見られた。「研究」では、「情報処理技術の研究開発（略）」などの記述が見られた。これらのことから、「研究開発」に関連する情報の発展である。2つ目は「機能」の表出が見られた。「機能」を含む記述を抜粋すると、「(略) 同時通話の機能は (略)」などの記述が見られた。これらのことから“機能の多様化”であった。3つ目に「検索」の表出が見られた。「検索」を含む記述を抜粋すると、「忘れられる権利」をめぐって、検索エンジンの運用が変化する」などの記述が見られた。これらのことから、“検索”に関わるトラブルであった。4つ目は、「サービス」「ユーザー」の表出が見られた。「サービス」を含む記述を抜粋すると、「営利目的のコミュニケーションサービス」などの記述が見られた。また、「ユーザー」を含む記述では、「いたずらに機能を誤って使ったユーザーから多くの抗議が寄せられ (略)」の記述が見られた。これらのことから、“ユーザーサービス”に関わるトラブルであった。5つ目は、「提供」「発表」などの表出が見られた。「提供」を含む記述を抜粋すると、「サイバーセキュリティ保険の提供」が見られ、「発表」では、「サーバーが外部から不正アクセスを受けた疑いがあると発表した」の記述が見られた。これらのことから“トラブルの公表／発表”であった。6つ目は、「アカウント」「パスワード」などの表

出が見られた。「アカウント」を含む記述を抜粋すると、「12万5000件のアカウントを凍結したと発表」が見られ、「パスワード」では、「金融機関のネットバンキングサービスを使うとIDとパスワードが盗まれ」などの記述が見られた。これらのことから、「情報セキュリティ」に関連するトラブルであった。7つ目は、「ネット」「顧客」「報告」などの表出が見られた。「ネット」を含む記述を抜粋すると、「ネットを積極的に使う「アクティブシニア」が、ネット詐欺の被害に遭っている」などの記述が見られた。また、「顧客」「報告」における語彙からは、「不正アクセスを受け、登録顧客の個人情報が流出した」「2015年から法人ユーザーによる被害報告が急増しています」などの記述が見られた。これらのことから、「顧客の情報管理」に関するトラブルであった。8つ目は、「企業」「セキュリティ」「攻撃」「政府」「サイバー」などの表出が見られた。「企業」を含む記述を抜粋すると、「会社の顧客企業リストの一部を不正アクセスにより窃取した」などの記述が見られた。また、「セキュリティ」「攻撃」「サイバー」などの語彙を含む記述では、「データ越境におけるサイバーセキュリティの確保が必要」「アカウント情報などを騙し取るフィッシング攻撃が発生している」「外部からの大量データ送信によるサイバー攻撃とみられ」が見られた。9つ目は、「不正」「アクセス」「感染」「漏洩」「流出」などの表出が見られた。「不正」を含む記述を抜粋すると、「ネットバンキング不正送金ウイルスが添付されていた」などの記述が見られた。また、「アクセス」「感染」「漏洩」の語彙を含む記述は「インターネットを経由して遠隔の第三者から意図せずにアクセスされる事例」「ドライブバイダウンロード攻撃で感染を試みている」「悪意のあるコードの実行、情報漏洩、Webサイトの改ざんなどの影響を受ける」などが見られた。これらのことから、「情報漏えい」に関わるトラブルであった。10つ目は、「メール」「電話」「番号」などの表出が見られた。「メール」を含む記述を抜粋すると、「標的型攻撃メールを社員が開いてしまい、(略)」などの記述が見られた。また、「電話」「番号」などの語彙を含む記述は「(略)知らない男から電話があった」「近年、番号の流出被害が急増、社会的問題になっている」などが見られた。これらのことから、「個人情報」に関わるトラブルであった。抽出された語彙から、直接的に児童生徒に関わる語彙も見られた一方、社会的・経済的なトラブルと推察される語彙も見られた。しかし、実際に使用されている語彙の特徴をみると、児童生徒の情報や環境に関連する記事内容も見られ、犯罪・トラブルの幅広さの傾向が見られた。

#### 4. 考 察

これらの分析結果より、平成28年に報道されたインターネットに関連する犯罪は、サイバー攻撃などに関するものが多く、共起ネットワーク図（図1）による語彙間の関係性を検討すると、ウイルス感染や不正アクセスによる電話番号等の個人情報の流出が頻繁に報道さ



れていることが明らかになった。

対応分析（図2）の結果からは、これまでの先行研究と同様の情報セキュリティやサイバー攻撃などのカテゴリが表出するとともに、トラブルの公表／発表や顧客の情報管理など従来の研究ではあまり触れてこなかったカテゴリも表出しており、情報の管理などに関心が向けられているとともに、トラブルの公表や発表にも期待が寄せられていることが明らかになった。新たな情報公開の文化が醸成されたことから、新たに「トラブル公表／発表」や「顧客管理の立場に立った情報管理」などに着目され始めたとともに、「企業や政府機関を対象としたサイバー攻撃」などより具体的なサイバー攻撃についても取り上げられるようになったと考えられる。これは、児童生徒を取巻く環境や児童生徒が所持する情報そのものに対する犯罪やトラブルに巻き込まれることが、本研究を通じて知見として得られたことと示唆された。

## 5. ま と め

本研究では、平成28年に報じられたインターネットに関する事件についてテキストマイニングを用いて分析を試みた。その結果、情報通信サービスの変化に伴い新たな情報公開の文化の醸成がされたことから、トラブルの公表・発表、サービスを提供するものの情報管理の体制について着目されるとともに、企業や政府機関を対象としたサイバー攻撃などの具体的な問題について表出された。今後は、本研究で得られた新しいカテゴリを学習指導時に配慮・考慮するとともに、適した学習教材の開発が期待される。

## 参 考 文 献

- [1] 総務省（2015）情報通信白書平成27年度版
- [2] 警視庁（2016）平成28年上半年におけるサイバー空間をめぐる脅威の情勢等について
- [3] 文部科学省（2007）、情報モラル指導モデルカリキュラム、[http://www.mext.go.jp/a\\_menu/shotou/zyouhou/1296900.htm](http://www.mext.go.jp/a_menu/shotou/zyouhou/1296900.htm)（閲覧日：2017年9月1日）
- [4] 高橋参吉，阿濱茂樹，村田育也（2006）初等中等教育における情報モラル教育のための教材開発，電子情報通信学会技術研究報告106（249），17-22.
- [5] 渡辺真由子（2012）性的有害情報に関する実証的研究の系譜：～従来メディアからネットまで，情報通信学会誌 30（2），81-88.
- [6] 警察庁（2016）平成28年中におけるサイバー空間をめぐる脅威の情勢等について，[http://www.npa.go.jp/publications/statistics/cybersecurity/data/H28cyber\\_jousei.pdf](http://www.npa.go.jp/publications/statistics/cybersecurity/data/H28cyber_jousei.pdf)（アクセス日：2018年7月2日）
- [7] 樋口耕一：KHCoder，<http://khc.sourceforge.net/>（アクセス日：2017年9月10日）
- [8] 樋口耕一（2004）テキスト型データの計量的分析——2つのアプローチの峻別と統合——，理論と方法19（1）：101-115.