感情解析に基づく誘導型サイバー攻撃検知の検討

ファン タァンクァン[†1]　八槇 博史[†2]

概要：スマートフォンの普及により,個人をターゲットとする誘導型サイバー攻撃が近年増加傾向にある．これらの攻撃は LINE や Twitter などの SNS を通じてターゲットが興味を持ちそうな話題で接触を図り,巧みな言葉でユーザをマルウェアに感染させたり,アカウントを乗っ取るための PIN コードを発行するよう誘導したりしている．また,これらは迷惑メールなど一斉に同じ内容の文章を送信するバラマキ型の攻撃とは異なり,個人にカスタマイズされているため従来のベイズフィルターでは特徴を抽出するのが困難である．本稿では,フィルターに自然言語処理の感情解析を実装し,誘導サイバー攻撃の特徴を抽出できるか検討を行った．

キーワード：標的型サイバー攻撃，感情解析，SNS，詐欺，自然言語処理

Investigation of Induction Cyber Attack detection based on Emotional Analysis

PHAM THANH QUANG[†1]　YAMAKI HIROFUMI[†2]

Abstract: Due to the spread of smartphones, Induction Cyber Attacks targeting individuals have been increasing in recent years. Those are contacted through SNSs such as LINE and Twitter, to induce users to infect malware and to issue a PIN code to hijack accounts. Unlike broadcast type attacks that send sentences of the same content all together at the same time as spam mails, they are customized to individuals, so it is difficult to extract features with conventional Bayesian Filter. In this paper, we implemented Emotion Analysis on filter and investigated whether we can extract characteristics of Induction Cyber Attacks.

Keywords: Induction Cyber Attack, Emotion Analysis, SNS, Scam, Natural Language Processing

1. はじめに

近年は個人をターゲットにした標的型攻撃をはじめとしたユーザを誘導し，機密情報や情報端末のアクセス権限を奪取する攻撃が増加傾向にある．また，スマートフォンの普及により，これまでは企業を狙った攻撃が個人への攻撃にシフトしている．直近では任天堂ゲーム本体 Switch が品薄である事を利用し，本体をプレゼントすると騙し，前金として Amazon カードや iTunes カードの番号を送信させる手口が Twitter や LINE などの SNS 上で流行している．これらの攻撃の特徴としては，はじめにターゲットが興味を持ちそうな話題や，感情を持つ話題で接触し，巧みな言葉使いでユーザを攻撃者の望む行動をさせ，マルウェアに感染させたり，アカウントを乗っ取るための PIN コードを発行させたりしている．警視庁が発表した不正アクセスの件数[1]は，平成 23 年が 889 件だったのに対して，平成 27 年には 2,051 件とスマートフォンやメッセージアプリが普及するのに伴い増加傾向にある．

誘導型サイバー攻撃への対応策として広く知られているのが，ベイジアンフィルターである．事前に攻撃文章で利用される単語の出現頻度分布を計算し，頻出度が高いが文章中に含まれていた場合に攻撃だと判断するものである[2]．これは，同じ内容のメールを一斉に多くの人に送りつける従来の一対多のやり取りに効果的だが，標的型メールなど個人によって人間の心理的な隙を狙って[3]攻撃文章の内容が大きく変わる一対一のやり取りには効果を発揮しない．

今後，個人をターゲットとした攻撃(本稿での攻撃という言葉をサイバー攻撃の初期段階である「ターゲットに対してメールなどメッセージを送信し，遠隔操作ウィルスをダウンロードさせたり，アカウント乗っ取りを行うための情報を提供するよう仕向けたりして，より高度なサイバー攻撃を行うための事前準備」と定義する)がさらに増加するとみられる．そのため，文章の内容がターゲットによって変化しても，攻撃であるかを判断できる新しい尺度が必要とされる．本稿では，単語に感情ラベルを付与することで，文章を感情という新しい尺度で攻撃を判断できるシステムの検討及び実装を行った．

2. 巧妙化するサイバー攻撃

2.1 サイバー攻撃の変化

他人のコンピューターに侵入を試みるサイバー攻撃自体は 1980 年代から確認されており，1983 年に北米で公開された映画"War Game"によって，大衆にサイバー攻撃が広く知られるきっかけとなった．この頃の攻撃者のモチベーションはソフトウェアを改造することで，コンピューターの

†1 東京電機大学大学院情報環境学研究科 Tokyo Denki University, Graduate School of Information Environment
†2 東京電機大学システムデザイン工学部 Tokyo Denki University, School of System Design Engineering

限られたリソースを有効的に利用するのが目的である．あくまでも「自分の技術力を知ってほしい」「自分がこんなにも知識があることを認めて欲しい」という承認欲求が強かった．2003 年頃から日本ではオンラインバンキングを提供する銀行が登場し，インターネット上でお金のやり取りする機会が増加した．それをきっかけに，攻撃者は承認欲求から金銭目的に変わっていった． その後 2007 年の iPhone の発売をきっかけに，スマートフォンが一般に爆発的に普及した．小さくて，高性能，さらに拡張性の高いアプリケーションの登場で，人々はインターネットを通じて様々なことを行えるようになった．それに伴い，スマートフォン上に銀行口座やクレジットカードや PayPal の決済情報，Web サイトの ID やパスワードが多く保存され，価値の高い個人情報を持つ端末となった．その結果，個人をターゲットとしたサイバー攻撃が増加していった． また， 今後は価値の高い情報を持ち，さらに IT リテラシーや判断力が未熟な青少年が狙われる可能性が高い．内閣府の調べによると，青少年のスマートフォン普及率が平成 22 年で 1.5%だったが，平成 25 年では 34.8%まで上昇し，平成 28 年では 57.1%[4]と半数以上はスマートフォンを保有している．未成年の狙った攻撃の多くは Twitter や LINE など 10 代や 20 代の最も良く利用する SNS 上で起きている．

2.2 誘導型サイバー攻撃の手口

インターネット普及の草創期頃の攻撃の手口は単純なものが多かった．攻撃の代表例として迷惑メールが挙げられる．これは，マルウェア仕込んだファイルや URL に添付したメールを大量に一斉送信するものである．メールの内容は受信者全員に当てはまりそうなものである．電子メールは格安で国境に関係なく世界中に一斉送信できる特性があることから，外国人が日本人になりすましてメールを送信することが多かった．外国語から日本語へ翻訳する際，多くの攻撃者が Google 翻訳などを利用している．機械による翻訳の精度は低く，おかしな日本語になっている事が多い．そのため，迷惑メールの多くはひと目で怪しい内容だと気づくことができる．

ところが，近年では個人をターゲットとした攻撃は複雑で，巧妙になってきている．ユーザの端末やアカウントが持つ情報価値の上昇により，以前の大量バラマキ型ではなく，個人をターゲットにした攻撃を行っても利益を手に入れる事ができるからである．攻撃ではターゲット一人ひとりに合わせた文章と攻撃シナリオを用意し，それを元に攻撃者の言いなりになるよう誘導するケースが多い．例えば，LINE 詐欺ではまず初めにターゲットの知人や知人などよく知っている人間になりすまし，接触して来る場合が多い．攻撃方法は，ターゲットからお金をだまし取る手口とユーザカウントを乗っ取る手口がある．お金をだまし取る方法は，「もうすぐで給料日だからお金を貸して欲しい」，「急いでいるからとりあえず今はお金を立て替えて欲しい」など急用であることを理由に考える時間を与えずに iTunes カードなどのネットマネーを買わせ，券面に印字されている番号情報を送信させるものが多い．一方，ターゲットのアカウントを乗っ取る方法は，「LINE のアドレス帳が消えてしまった．新しく登録するから，電話番号と SMS 認証番号を教えてほしい」といかにも合理的な理由で，アカウント乗っ取りを行うための情報をターゲットから引き出している．乗っ取りに成功した攻撃者はアカウントのアドレス帳機能を利用し，被害者になりすまし，先程と同様な手口で被害をさらに拡大させている．誘導型サイバー攻撃の特徴は，従来のバラマキ型の攻撃とは異なり，ターゲットの身近な人間を装って攻撃している．その結果，身近な人間だからという安心感が生まれ，少し無理な内容でも誘導に従ってしまうのである．攻撃方法はオレオレ詐欺など実世界の詐欺に非常に似ている．

3. 誘導型サイバー攻撃への対策

3.1 誘導型サイバー攻撃への対応策

乗っ取りやインターネット上での詐欺行為に対して，各社様々な対応策を出している．例えば，LINE では平成 28 年にアカウント乗っ取り被害防止策として，PIN コードによる認証機能を廃止し，電話番号/SMS による 2 段階認証を導入した．また，平成 29 年 6 月から「LINE サイバー防災訓練」と呼ばれる体験型ムービーを通じて LINE 乗っ取りの疑似体験できるサービスの提供を開始した．乗っ取りの疑似体験を通じて，実際に攻撃のターゲットにされた際に正しい対処法を行えるようユーザを訓練することを目的としている．

だが，攻撃者は知人になりすまし，巧みな言葉遣いでターゲットの心のスキに付け込むことで情報を抜き出してしまうため，被害をゼロにはできていない．

3.2 ベイジアンフィルターの限界

従来のバラマキ型攻撃への対応策としてベイジアンフィルターが非常に有効的だった．ベイジアンフィルターは文章のカテゴリー化を統計のアプローチで行うものである．スパムメールの例として，まず事前に手動で単語をいくつかのカテゴリーに分類する．そしてスパムメールなどに含まれている文字集合と正規メールの文章に含まれる文字集合の確率を計算し，スパムメールか否かを判断する．ベイジアンフィルターのアルゴリズムは統計の性質上によりスパムメールのような同じ内容の文章が大量に送信される攻撃に対しては非常に有効的である．だが，誘導型サイバー攻撃のようにターゲットに合わせてカスタマイズされ，単語の出現頻度が一定とはなりえない文章の攻撃に対しては効果を発揮しない場面が多い．今後一般的になると予測される誘導型サイバー攻撃へ対処するためには，従来の単語の出現確率で文章を分類する方法に代わり，文章中の単語の出現頻度の変化に影響されずに，文章内容や出現単語が

ターゲットによって大きく変化しても文章を適切に評価できる新しい基準が今後必要である．

4. 感情解析による対策

4.1 文字から得られる人間の感情

人がソーシャルメディアなどで書き込みをする際,その内容がただの文字列の羅列で無機質に見えるが，実は書き込んだ人間の心理状態を内包している場合が多い．絵文字や顔文字などで感情を伝える方法もあるが，文字だけからでも人間は感情を読み解く事ができるのである．例えば，人に対して何か頼み事をする際は,「〜に困っている」,「〜を一人でやるのはつらい」など自分のネガティブな精神状態を，文字を通して相手に伝え，頼み事を聞いてくれる確率を高めようとする．逆に「頼み事を聞いたら，こんな素敵な経験を得られる事ができます」や「これができるのは貴方しかいない」など，文字を通じて相手のポジティブ精神に訴えかける場合がある．

文字情報のみから感情を読み取れる性質[5]を利用し，多くの企業ではソーシャルメディア上のユーザコメントを収集し，そこから感情を抽出する感情解析のソリューションを提供している． 例えば，jetrun 社が EC サイトの商品レビュー，掲示板に投稿された商品に対する意見を抽出し，感情解析を行うことでお客様の声を数値化し，商品開発やマーケティングのための指標を提供している．ここでは，ユーザの意見を，プラス感情を持つ"ポジティブ"，マイナス感情を持つ"ネガティブ", 感情を持たない"ニュートラル"の三段階に分類し数値化している．文章を解析し，感情を持つ単語の数を数え上げ，最終スコアがプラスであれば，ポジティブ意見として，マイナスであれば，ネガティブ意見として分類する．例えば,「この映画を見たが，内容がイマイチで全然内容が入ってこなかった．僕はあまり好きじゃない」というコメントがあれば，機械が文から"イマイチ"と"好きじゃない"という副詞と形容詞を抽出し，単語 DB にアクセスして感情を持つか調べる．この場合は，2 つともマイナス感情を持つことから，最終スコアが-2 となる．最終スコアがマイナスにより，この意見はネガティブ意見として分類され，商品開発の際の大切な指標として活用できるようになる．企業がリリースした製品の反応を随時確認したい時に非常に効果を発揮する．ユーザの感情をリアルタイム知ることで，適切な広告配信やユーザ視点に立った新商品開発に貢献できることや，面倒で高コストなアンケートを実施する必要がないことから街中や電話アンケートなど，従来の意見抽出の方法から取って代わられるようになった．

4.2 誘導型サイバー攻撃への応用

4．2 でのように，言葉には人間が意図しなかったにも関わらず，人間の感情を反映している場合が多い．これはソーシャルメディア上投稿の製品レビュー特有なものではない．小説など”話し言葉”以外の文章でも同様な性質が見られる．小説では読み手に共感してもらい，より深く物語に没頭できるよう，様々な感情表現を文章中に秘めている[6]. このように，文章にはその形態に関わらず，様々な感情を持っている．つまり誘導型サイバー攻撃の文章も同じ性質である可能性が高い．誘導型サイバー攻撃ではターゲットを自分の思い通りの行動をさせるため，相手を説得する必要がある．相手の理解を得るため，ターゲットが共感できる内容に文章がなっていなくてはならないのである．そのため，通常の文章以上に感情に関わる言葉を使う確率が高いと考えられる．この性質を利用し，攻撃文章に対して感情解析を行い，感情を数値化することによって目に見えない攻撃を可視化できる可能性がある．ターゲットの心理を揺さぶり，共感を得るための心理的な攻撃を検知できるかもしれない．

5. システム実装

5.1 システム構成

システム構成図を図 1 で示す．この構成で実際に攻撃に利用された文章の感情解析を行い,感情度の計算を行った．

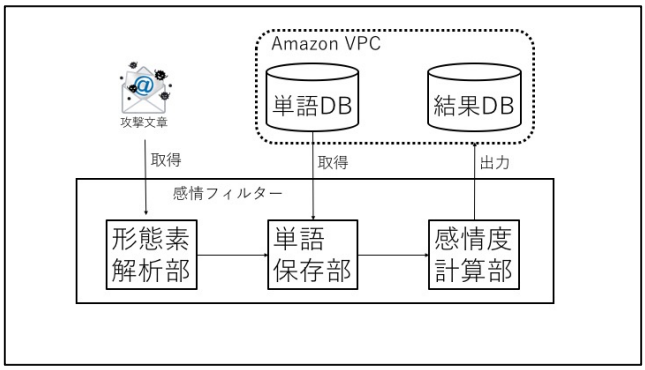

図 1 感情解析のためのシステム構成図
Figure 1 System Configuration for Emotional Analysis

システムは感情フィルター，単語 DB，結果 DB で構成されている．感情フィルターは形態素解析と単語保存部と感情度計算で構成されている．メンテナンス性を考慮し，単語 DB，結果 DB は外部に構築し，アクセスする方式を取り入れた．それぞれが完全に独立する事により，感情フィルターに変更を加えたとして他のシステムに影響が行かない設計となっている．

システム構築にあたり，プログラムの全ては Python3．5 で記述した．Python を利用した主な理由は XML 解析を行うためのライブラリーが充実していたことと，自然言語処理をするための形態素解析エンジンである MeCab の導入が簡単だったことがあげられる．

5.2 感情パラメータの決定

感情解析を行い際は最初に感情のパラメータパターンを数種類決めておく必要がある．一番単純なパラメータパターンは「肯定(ポジティブ)/否定(ネガティブ)」である．あとは利用目的によって基本パターンを拡張する方法が取られている．例えば，スマートメディカル株式会社のコールセンター向けの感情解析では基本パターンに「平常/喜び/怒り/悲しみ/元気度」の５つの感情パラメータでお客様の感情傾向を数値化する方式が取られている．また，メタデータ社の提供するテキスト用の感情解析 API では，「好ましい - 嫌い」「嬉しい - 悲しい」「怒り - 悲しみ」の 3 つの軸を用いて 7 段階で文章を評価できるシステムが提供されている．

今回, 攻撃用の文章中に感情の有無を調査するだけのため感情解析を利用する．上記の例のような複雑な感情パラメータは必要としない．今回のシステムでは「肯定(ポジティブ)/否定(ネガティブ)」の基本パターンだけを利用してシステムを構築していく．

5.3 単語データベースの構築

感情解析を行うためには, 事前に単語データベースに単語と感情カテゴリーを示すラベルを持った感情表現辞書と呼ばれるデータベースを構築する必要があった．感情 DB はメタデータ株式会社が感情解析 API を通じて提供されているが，1 日/100 回という制限があったことに加え，感情 DB を自分で拡張できないという理由から自分で構築した．

感情 DB 構築にあたり，利用した自作の感情 DB 構築プログラムのフローチャートを図 2 で示す．感情 DB 構築プログラムではまず初めにメールを読み込み, XML 解析を行うことでメール本文の抽出をする．次に文章を形態素解析し, 分解した品詞を単語保存用の配列に格納する. そして，配列の先頭から一つずつ単語を抜き出し，単語 DB に一致する単語が存在するか調べる．単語に一致する単語があった場合は処理を何も行わずに，次の単語の処理を行う．ここで，単語 DB に一致する単語がなかった場合は，単語を画面に出力し，手動で"POSITEVE"か"NEGATIVCE"のどちらかを入力し, 単語に感情パラメータを付与する. そして，単語と感情パラメータを Python の辞書形式で感情 DB へ保存する．配列から全ての単語を抽出したあとはプログラムの初めに戻り，新たなメールを解析することを繰り返す.

今回は感情 DB 構築プログラムで計 800 件の標的型メール，スパムメール，正規メール，LINE 詐欺の台本に対して形態素解析を行い，3000 語の品詞に対して感情パラメータを付与した．

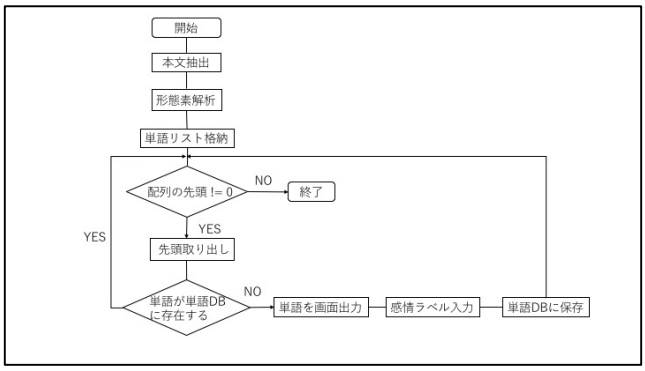

図 2 感情 DB 構築用プログラムのフローチャート
Figure 2 Flowchart of Emotion DB Construction Program

5.4 単語保存部による感情カウンター

単語保存部では形態素解析部から取得した単語リストの中から感情を持つ単語を調べる機能と感情含有している数を示す感情カウンター(変数名 POSITIVE, NEGATIVE)を持っている．処理内容として，取得した単語リストを一つずつ取り出し，その単語が単語 DB に含まれているかを調べる．単語 DB の単語と一致した場合にはそれがどちらの感情かを調べ, 結果に応じて適宜 POSITIVE か NEGATIVE の数を一つ増やしていく．単語リストの全ての単語の感情を調べたあとは, 感情カウンターを感情度計算部に引き渡す.

5.5 感情度計算部による感情度の計算

感情度計算部では攻撃文章中にどのくらい感情が含まれているのかという感情の度合いを計算する機能を持っている．感情度の計算を行うため，形態素解析部から感情度に影響を与えない助詞を除いた品詞の数，単語保存部からは感情カウンターのパラメータを取得する．そして，下記の式 1 にそれぞれを代入し，感情度を計算する．式では攻撃文章中における感情カウンターの割合を計算することによって，文章の感情度を求めている．また，感情度計算部では感情度を計算したあと，その結果を結果 DB に保存する設計になっている．

$$感情度 = \frac{感情カウンター}{文章全体におけ助詞を除いた品詞の数} \quad (1)$$

5.6 結果 DB による計算結果の保存

結果 DB では感情度計算部から受けった感情度のデータを保存している．また，今後データ解析を行うため，全ての結果を csv ファイルとして出力する機能を備えている．

6. 実験

6.1 実験用データの収集

実験を行うに辺り， 正規メール，標的型メール，LINE 詐欺のメッセージを 300 件収集した．正規メールは大学のメールボックスに送信されたメールを，標的型メールは IPA と TDU-CSIRT が提供している標的型メールの例を，

LINE 詐欺のメッセージは平成 26 年に流出した LINE 詐欺の台本をそれぞれ利用した．

今回は実験用のデータとして正規メール，標的型メール，LINE 詐欺メッセージをそれぞれ 10 サンプルずつ感情フィルターで解析を行った．また，今回は攻撃で使う会話のシナリオは「メッセージ内の URL までターゲットを誘導する攻撃」で統一した．

6.2 実験の流れ

実験は図 3 のフローチャートの流れに沿って実施した．まず，感情フィルターにメールを読み取らせ，形態素解析を行う．形態素解析を行った後，助詞を削除し，それ以外の単語を配列に格納する．次に，配列から単語を一つずつ取り出し，単語 DB に格納されている単語リストと比較する．取り出した単語と単語 DB の単語が一致した場合，ポジティブかネガティブかを調べ，感情パラメータのカウンターを一つ増やす．処理後はまた配列の先頭を取り出し，先程と同じ処理を行っていく．そして，配列の中身が 0 となった時点で感情パラメータのカンターの中身を感情度計算部に渡し，感情度の計算し，処理を終了させる．

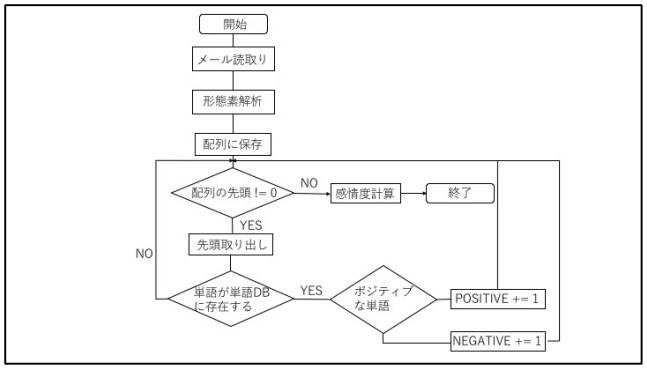

図 3 実験のフローチャート

Figure 3 Flowchart for Experiment

6.3 実験結果

結果 DB より感情度を取得し，グラフ化した．正規メール，標的型メール，LINE 詐欺を解析した結果をそれぞれ図 4，5，6 で示す．図中のx，y軸がそれぞれメッセージ内容，感情度を示している．グラフで青棒がネガティブな感情を，赤棒がポジティブな感情を表している．

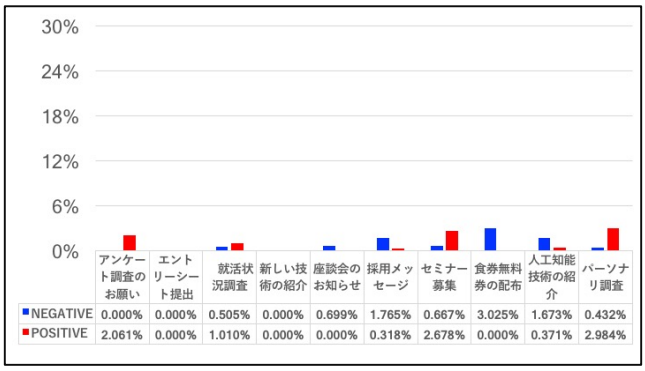

図 4 正規メールの感情度

Figure 4 Emotional degree of Regular Mail

正規メールでは，「エントリーシートの提出期限の通知」や「座談会に関するお知らせ」など，イベントの紹介や申込期限に関する文章に関しては感情が見られなかった．だが，「パーソナリティ調査」や「セミナーの参加案内」など，ターゲットに関係がありそうな内容については，攻撃者の感情が見られた．正規メール全体を通じて，ポジティブとネガティブの感情の両方が見られた．

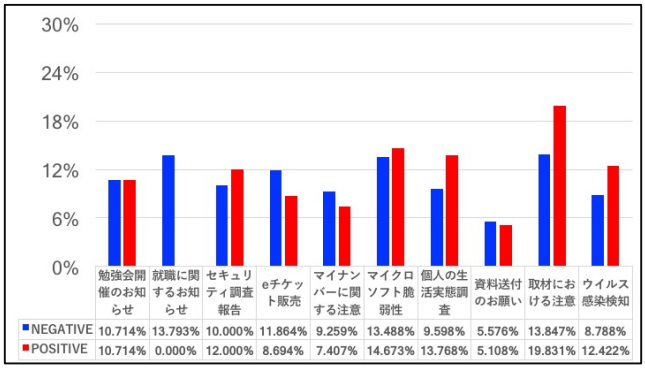

図 5 標的型メールの感情度

Figure 5 Emotional degree of Targeted Mail

標的型メールにおいては，全体を通じてポジティブとネガティブの感情度がバランス良く見られた．特に「マイクロソフト製品の脆弱性」と「セキュリティ調査に関する調査報告」などセキュリティに関する標的型メールでは，両方の感情度がほぼ同じくらい使われた．これは，まず初めにシステムには脆弱性などユーザの危機感を煽り，判断力を鈍らせたあとに問題への解決策を提示することでユーザを誘導させる攻撃手口だったと予測できる．このように，標的型メールではポジティブとネガティブの両方を利用し，ターゲットの心理状態を誘導する高度な攻撃が行われている可能性がある．

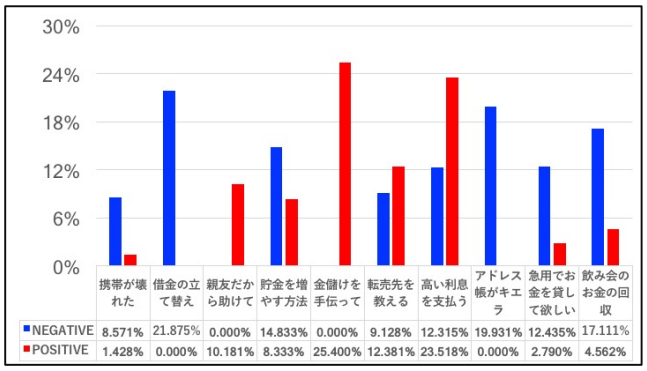

図 6 LINE 詐欺の感情度

Figure 6 Emotional degree of LINE Scam

LINE 詐欺においては，感情度が全体を通じて現れたが標的型メールとは違い，感情パラメータの出現の偏りが見られた．特に「借金の立て替え」や「急用でお金が必要になった」といった金銭問題に関する内容ではネガティブな感情が見られ，逆に「お金儲けを手伝ってほしい」や「高い利息を支払うからお金を貸して欲しい」など儲け話に関する内容のものにはポジティブな感情が多く見られた．一方，「良い転売先を教える」といった内容では感情パラメータがバランスよく出現するという標的型メールに似た攻撃も見られた．このことから， 今後は標的型メールのように感情をバランスよく利用して人間精神を誘導する攻撃が増えることが予測される．

7. 考察

今回の実験を通じて，正規メール，誘導型サイバー攻撃，LINE 詐欺間において感情度の傾向の違いが多く見られた．正規メールでは，全体に占める感情は低かった．これは，送信者がターゲットの誘導をする必要性があまりないからと考えられる．だが，送信者の意図に関わらず，感情が見られる文章もあったため，これが攻撃の誤検知の原因はとなるリスクはある．一方，標的型メールは感情のバランスが良かった．標的型メールは文字制限がないため，より巧妙で複雑な文章で誘導方法が行われている可能性がある．また，LINE 詐欺においてはどちらかの感情が極端に高かった．原因として LINE 詐欺の方が少ない文字数でターゲットを誘導する必要があるため，感情が出やすいからだと考えられる．また， LINE 上では短文のメッセージをお互いに送りつけ，会話を成立させていることから，日常の会話にとても似ている．LINE 詐欺のパターンを解析することによって，オレオレ詐欺や振り込め詐欺など従来の会話で相手を騙し，金銭を要求する詐欺も検知できる可能性がある．

8. 今後の課題

実験を通じていくつかの課題が見えきた．今回の「URL にアクセスさせる」シナリオでは感情度を検出できたが，他の攻撃シナリオでは感情度がどのようになるかはまだ分かっていない．今後は「ウィルスをダウンロードさせる」シナリオや「クレジットカード番号を送信させる」シナリオなど，他の攻撃のシナリオも分析していく．また，今後は感情を示唆する単語の利用を避けることでフィルターを回避するタイプの攻撃が出現する事が見込まれる．これに対しては，感情フィルターに新しく類語 DB を実装するなど，感情を検出できない単語を感情検出できる単語に置き換えるなどの対応策を練っていきたい．今後はもっとシナリオ数とサンプル数を増やすことで，シナリオと感情度の相関関係を分類問題で適用し，未知のシナリオが出てきた時に感情度を検知できるシステムを構築していきたい．

参考文献

[1] “平成 27 年における不正アクセス行為の発生状況等の公表について”. https://support.office.com/ja-JP/article/d38d6e47-f6fc-48eb-a607-1eb120dec563, (参照 2017-08-20).

[2] “ A PLAN FOR SPAM” , Paul Graham,2002(参照 2017-08-20) http://practical-scheme.net/trans/spam-j.html (参照 2017-08-20)

[3] 寺田剛陽，鳥居悟,安野智子,瀧澤弘和,新真知,“リスク認知に基づく標的型メール対策の検討”, 研究報告グループウェアとネットワークサービス（GN）, 2013-GN-88, No.9, pp. 1-8, 2013.

[4] “平成 28 年度 青少年のインターネット利用環境実態調査 調査結果（速報）”, http://www8.cao.go.jp/youth/youth-harm/chousa/h28/net-jittai/pdf/sokuhou.pdf(参照 2017-08-20)

[5] 小林のぞみ，乾健太郎，松本裕治，立石健二，福島俊一, “意見抽出のための評価表現の収集”, Journal of natural language processing 12(3), pp. 203-222, 2005.

[6] 三和義秀,小林久恵, “小説を対象にした感性語の分類の基礎研究:意味的類似性を基準として”, Journal of library and information science, vol. 17, pp.27-37, ,2003,