Vol.2017-DPS-170 No.13
Vol.2017-CSEC-76 No.13
2017/3/2

標的型攻撃対策訓練メール自動生成のための受信メール分析手法の評価

岩田一希[†1]　中村嘉隆[†2]　稲村浩[†2]　高橋修[†2]

概要：近年，標的型メール攻撃の被害が増大している．現状の対策における課題として使用されるマルウェアは既存の対策ソフトでは検知できない場合が多いという点，またマルウェアへの防御システムは既知の攻撃にしか対応できない点があげられる．この課題に対して，「人間」に擬似的に攻撃を受けさせ，攻撃に対する訓練をすることで，標的型メール攻撃への耐性をつけるという手法が考えられている．しかし現在行われている訓練手法では個々の被訓練者に対して有効な訓練にはなっていない．そこで本研究では受信 BOX にある受信メールを送信者ごとに分類し，その送信者ごとのメール文章の特徴を分析し，普段被訓練者が受信するメールに類似した擬似攻撃メールを自動生成し，受信メールのように表示することで，効果の高い訓練を行うことが出来るシステムを提案し，メールの特徴を分析する部分についての実装を行った．さらに，評価として，我々の受信している実際の受信メールを用いて，特徴を抽出しメール文章生成が出来るかどうかを検証した．

キーワード：ネットワークセキュリティ，標的型攻撃，電子メール，文章分析．入口対策

Evaluation of incoming email analysis for automatic generation of training email against APT attack

KAZUKI IWATA[†1]　YOSHITAKA NAKAMURA[†2]
HIROSHI INAMURA[†2] OSAMU TAKAHASHI[†2]

1. はじめに

近年，組織の一個人を対象として，悪性サイトの URL やマルウェアを添付したメールを送信し，マルウェアをダウンロード・実行させて感染させ，組織内ネットワークにバックドアを仕掛けることで，情報の窃取を行う標的型攻撃の被害が増大している．標的型攻撃に用いられる攻撃メールには，被害者であるメール受信者にメールを送信する可能性の高い第三者になりすましたり，被攻撃者の業務内容に沿ったメール内容を装ったりしているという特徴がある．被害者はその組織の従業員もしくは社員であることが多く，そのような場で使われるメールはビジネスシーンで使われるメールの体裁であることが多い．したがって，攻撃メールもビジネスメールの体裁を装って送られてくるパターンが多くなっている．また，このような攻撃に使用されるマルウェアを実行後も，アンチウイルスソフトウェアでの検知が難しい上に，通常業務に影響が出ないように動作することが多いため，攻撃に気づきにくい．したがって，情報漏洩などの被害が生じてから初めて標的型攻撃を受けていたことに気づくケースも多く発生している．

標的型メール攻撃の攻撃手順には，図 1 のような 6 つの段階がある．

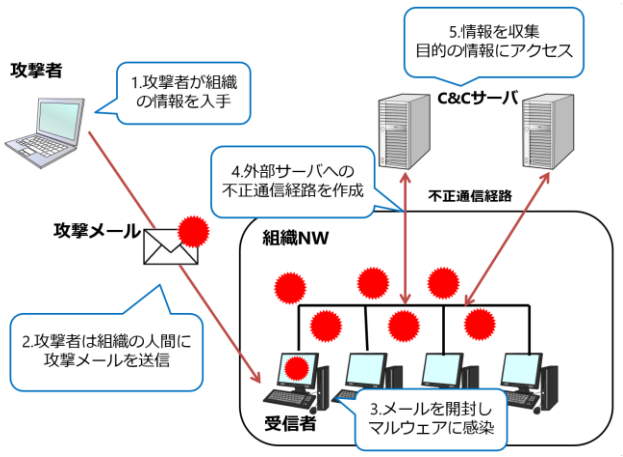

図 1: 標的型攻撃の攻撃手順

攻撃者は何らかの手段で受信者 A の情報を手に入れることが可能であるという前提のもと，この攻撃が実行される．受信者が攻撃メールに添付されているマルウェアを実行してしまうと，組織内の PC やサーバがマルウェアに感染してしまう．マルウェアは組織のネットワークに対して，不正な通信経路(バックドア)を作成する．攻撃者は C&C サー

†1 公立はこだて未来大学大学院 システム情報科学研究科
Graduate school of Systems Information Science, Future University Hakodate
†2 公立はこだて未来大学 システム情報科学部
School of Systems Information Science, Future University Hakodate

バ(コマンド&コントロールサーバ)から新たなマルウェアを実行させる.そのマルウェアを用いて,内部情報を収集,個人情報の窃取などを行う.

警察庁[1]によると，警察が報告を受けた 2014 年上半期から 2016 年上半期の間の標的型攻撃の件数の推移は図 2 のようになっている．この図を見ると，標的型攻撃の被害件数は年々上昇傾向にあることがわかる．

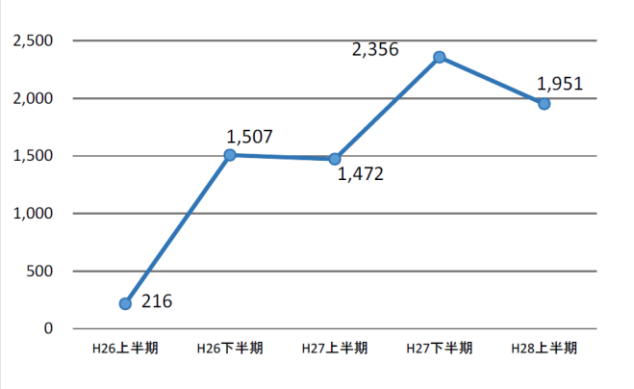

図 2: 標的型攻撃の件数の推移

また，2015 年 6 月には日本年金機構が標的型メール攻撃による大規模な情報漏洩を起こしている[2]．さらに，2016 年 6 月には JTB が標的型攻撃によって 793 万件の個人情報が流出した可能性があることを公表した[3]．このような状況にあるため，標的型攻撃への対策は社会的に急務となっている．

標的型攻撃において，マルウェアが侵入しないようにする対策を「入口対策」と呼び，外部へ不正な通信が行われないようにする対策を「出口対策」と呼ぶ.「入口対策」の課題として，この攻撃に使用されるマルウェアは既存のアンチウイルスソフトウェアでは検知できない場合が多いという点，またマルウェアへの防御システムは基本的に既知の攻撃にしか対応できない点がある．よって未知のマルウェアが添付された攻撃メールが送信されてきた場合に，マルウェアを侵入させないような手法が必要となる．

本稿では，2 章で入口対策の課題の解決方法とその関連研究・技術について述べる．3 章で提案システムについて説明する．4 章で提案手法の有効性を確認する評価実験について説明し，5 章では実験結果についての考察と本稿の結論を述べる．最後に今後の展望と本稿のまとめについて 6 章で述べる．

2. 関連研究・技術

本章では，入口対策の課題として挙げられた従来の防御システムやアンチウイルスソフトウェアでは既知の攻撃にしか対応できないという課題に対する解決手法について述べ，それに関連した研究と技術について紹介し，既存研究・技術の問題点について述べる．

2.1 入口対策の課題に対する解決手法

入口対策の課題の解決策として「人間」に擬似的に攻撃を受けさせる，つまり攻撃されることへの訓練を行うことで，標的型攻撃への耐性をつけるという手法がある．人間にはシステムと違い対応力が存在するため，訓練を経て耐性を持つことができれば，未知の標的型攻撃にも対応できると考えられる．訓練の方法の一つに，実際の攻撃メールに似せた訓練メールを用いるものがある．訓練メールは標的型攻撃メールに似せるためにビジネスシーンで使われるような体裁を使い，被訓練者が受信してもおかしくないようなメール文面を使用している．

2.2 訓練に関する関連研究・技術

(1) JPCERT/CC による訓練[4][5]

JPCERT/CC は 2008 年，2009 年に標的型攻撃対策訓練を行っている．訓練手順としては以下のとおりである．

1. 訓練を行うことを周知
2. 標的型攻撃についての教育
3. 訓練メール配信 1 回目
4. 訓練メール配信 2 回目
5. 被訓練者へのアンケート調査

事前に訓練を行うことの周知や，標的型攻撃についての教育を行っているが，訓練 1 回目よりも 2 回目の方が，添付ファイルの開封率が下がるという結果が確認できている．さらに，訓練対象の組織は，IT 企業から地方自治体まで幅広く，様々な業種の組織において標的型攻撃対策訓練の効果があるということが確認された．

(2) 内田[6]による訓練

内田[6]は，標的型攻撃対策訓練を以下の 1~4 の条件で実施し，表 1 のような結果を得ている．

1. 事前の情報提供をせず訓練を実施
2. 事前に情報提供を行って訓練を実施
3. 訓練実施 2 年後，事前の情報提供をせず実施
4. 訓練実施 2 年後事前に情報適用し実施

表 1: 訓練条件と添付ファイル開封割合

条件	添付ファイル開封割合
1	約 40%
2	約 10%
3	約 12.5%
4	約 6.3%

表 1 より，標的型攻撃対策として訓練が有効なこと，訓練を繰り返すことが重要であることが確認できた．

(3) 訓練企業による訓練

基本的には，標的型攻撃対策訓練を行う部署を持ってい

Vol.2017-DPS-170 No.13
Vol.2017-CSEC-76 No.13
2017/3/2

ない組織が訓練を実施する際には外部の訓練企業に委託することになる．訓練メールを一から訓練企業が作成するパターンもあるが，何点かのテンプレートの中から，被訓練組織の担当が訓練で使用したいメールのパターンを選び細部を変更することができるサービスも多く存在する．

2.3 関連研究・技術の課題

従来の訓練手法では，訓練メールは手作業で作成されている．よって毎回訓練を行う度に新たに訓練メールを作成する手間が生じる．テンプレート等を用いてこの手間を削減することも可能だが，結局細部の変更などの作業が必要になってしまう．一度作成した訓練メールを何度も使いまわした場合，訓練メールということを覚えられてしまい，訓練効果が薄れてしまう可能性があるため，細部の変更等の作業の必要性が生じることになる．このように，現状では訓練メール作成に大きな手間がかかることから継続した訓練は行われにくい状況にあり，訓練効果の維持という点で問題がある．さらに従来の手法で訓練メールを作成した場合，被訓練者全員に向けた同一内容の訓練メールが生成される．例えば，インフルエンザについての注意喚起を模した訓練メールや，全社で行われるイベント実施を伝えるメールを模した訓練メールなど，誰が受け取ってもある程度関連のあるメールが生成される．訓練の効果は被訓練者の特性 (普段メールをやりとりする相手，情報セキュリティ知識の有無など)に応じて変化するため，被訓練者ごとに合わせたメールを生成する必要があると考えられる．

3. 提案システム

本稿では，従来の攻撃対策訓練における，訓練継続性の問題，個人への特化への問題を解決するため，個々人の特性に応じた訓練攻撃メール自動生成を行う標的型攻撃対策訓練システムの提案を行う．

3.1 システム構成

提案システムは，図 3 のように，メールクライアント部を中心に構成する．提案システムのユーザ（ユーザ A）はこのメールクライアントを用いて日常的にメールの閲覧および送受信を行う．

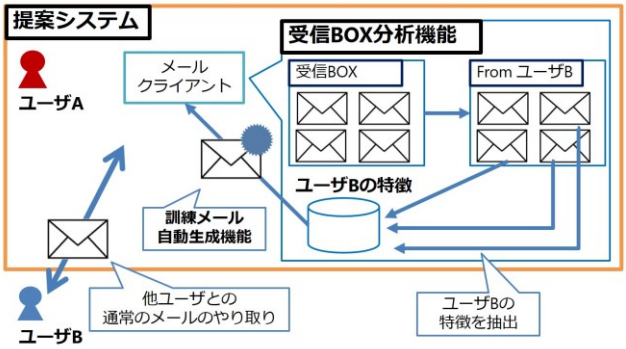

図 3: 提案システム構成

3.2 メールクライアントの機能

提案システムのメールクライアント部では，通常のメールの閲覧や送受信が可能である．この通常のメール閲覧・送受信機能に加えて，受信 BOX 内のメール内容を分析して，送信者ごとのメール特徴を抽出する受信 BOX 分析機能，抽出したメール特徴を用いて，各送信者を模倣した訓練メールを自動生成し，受信 BOX 内に通常のメールと同様に表示する訓練メール自動生成機能を備える．これらの機能が自動で動作することで「自動訓練」を実施することができる．

3.2.1 受信 BOX 分析機能

受信 BOX を分析することにより，ユーザが普段やり取りする相手の送ってくるメールの特徴を分析することができる．本稿で対象としているメールはビジネスシーンで使われるようなメールの体裁をもった文章で構成されている図 4 のようなメールである．

Subject	今週の会議の日程について
From	YY<YY@yyy.co.jp>
Message	XX 様 YY です. 今週の会議の日程を調整致しました 詳細は添付したファイルをご覧ください. よろしくお願い致します. 株式会社　○△□産業　営業部 YY
Attach	今週の会議の日程.txt

図 4: ビジネスメールの例

Vol.2017-DPS-170 No.13
Vol.2017-CSEC-76 No.13
2017/3/2

このときメール本文は以下のように構成されている．

1. 宛先
2. 挨拶
3. 内容
4. 添付ファイル確認促し
5. 〆の挨拶
6. 署名

この中でも，1,2,5,6 の部分に関しては，同一送信者なら同一文章になる場合が多いと考えられる．そこで，同一送信者からの各メールで同一文章が出現する部分をその送信者の「特徴」として保存する．この特徴を利用することでその送信者に似たメールを生成できると考える．

ここで．受信 BOX の中身を分析する手順について説明する．

(1) 受信 BOX 内のメールを送信者ごとに分割

まず，受信 BOX 内のメールを送信者ごとに分割する．分割の際にはメールヘッダの”Sender”を参照して送信者アドレスごとに分割する．新しいメールから得た特徴を使って訓練メールを作ることで，現在の被訓練者にとってより身近な訓練メールを生成できると考えられるので，一番新しいメールの送信者から順に分割していく．

(2) 送信者ごとに分割したメールから本文を取得

次に図 5 のように送信者ごとに分割されたメール群から 1 件ずつメール本文を取得する．この時も(1)と同様に最も新しいメールから取得する．

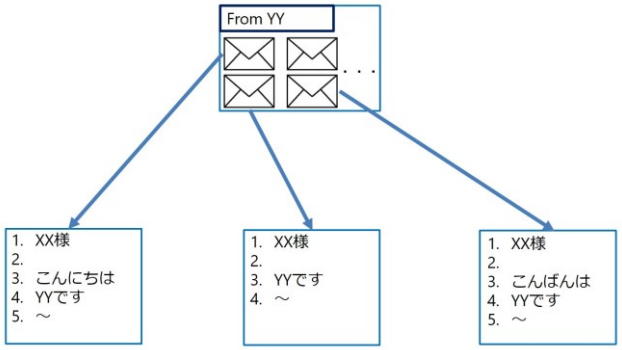

図 5: メール本文の取得

(3) メール本文を 1 行ごとに比較・加算

次にメール本文を 1 行ごとに比較する．比較は同一行に同一文があるかどうかを分析対象となっている全てのメールに対して行う．ここで「同一行」には昇順と降順で並べた場合の 2 パターン存在する．また，空白行は図 6 のように削除して詰めた状態で比較を行う．同様の処理を他の比較対象となるメールにも行う．

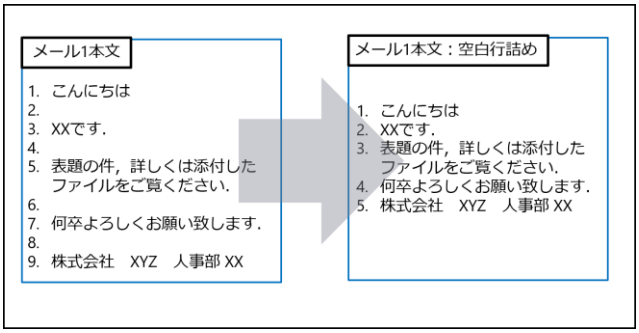

図 6: 空白行を削除して詰める工程

昇順で比較を行う場合，メールの先頭 1 行目から比較を行う．この時の 1 行目を「昇順 1 行目」と定義する．具体例を用いて説明すると，図 7 のメール 1 の 1 行目は「こんにちは」となっているため，「昇順 1 行目」の「こんにちは」の出現回数を「1」とする．

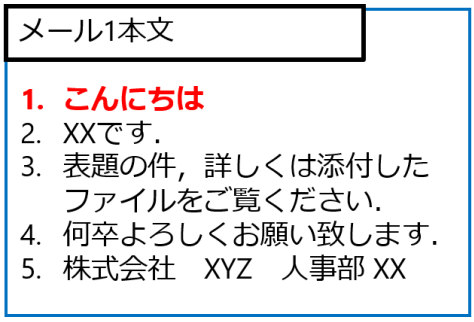

図 7: メール 1

次に図 8 のメール 2 の昇順 1 行目を確認すると「こんにちは」となっていて同一文が出現している．

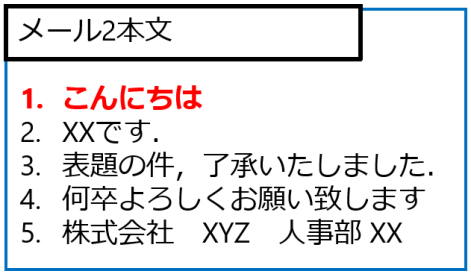

図 8: メール 2

よって，この送信者のメール群の昇順 1 行目の「こんにちは」の出現回数を加算する．次に図 9 のメール 3 の昇順 1 行目を確認すると「こんばんは」となっており，昇順 1 行目の「こんばんは」の出現回数を加算する．

Vol.2017-DPS-170 No.13
Vol.2017-CSEC-76 No.13
2017/3/2

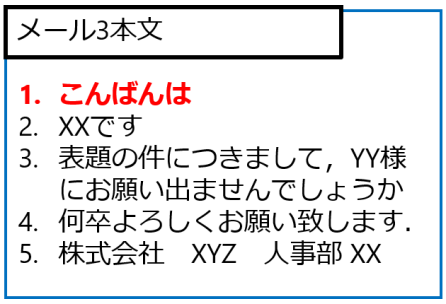

図 9: メール 3

これを全メールの昇順 1 行目に対して繰り返す．全メールを確認した後，昇順 2 行目以降の確認を行い，全行について繰り返す．これで昇順の場合の 1 行ごとの同一文の出現回数を記録することができる．この時，行ごとに回数が一番多い文をその行の特徴と捉えることできる．降順の場合も同様の手順で行う. このとき, 昇順の場合は特に挨拶や，名乗りなどの文頭部分，降順の場合は署名や〆の挨拶の特徴が顕著に現れるのではないかと考えられる．

3.2.2 訓練メール自動生成ツール

訓練メール自動生成ツールでは，3.3.1 節で述べた手法によって取得した送信者ごとの特徴を用いて，その送信者のよく作成するメールに似た訓練メールを自動生成する．送信者ごとの特徴を用いてメール文章を生成する際，挨拶，名乗り等の冒頭部分については降順特徴，〆の挨拶，署名など
の末尾部分については昇順特徴を用いて，それぞれの特徴が顕著な部分を貼り付けることで生成できる．しかし，メールを送信した目的によって，文章が大きく異なってしまう内容部分に関しては特徴がまちまちになるため，特徴を活用して生成することが難しい．そこで，添付ファイルの開封を促すようなテンプレート文を用意し，内容部分に差し込むことで送信者の特徴を損なわない訓練メールを生成できると考えられる．このようにして，訓練メールをある程度自動生成することが可能である．しかし，このまま訓練に用いると，この訓練メールは偽装した送信者の通常作成するメールの特徴を備えているため，訓練メールと通常メールとの見分けがつかなくなってしまい，訓練メールの添付ファイル開封確率が非常に高くなるなど，訓練として用いるためには支障を来す．ここで，実際の攻撃メールの事例[7]を見てみると，通常やりとりされるメールには無いような怪しい部分が存在することがわかる．このような部分を「不自然な点」と定義する．不自然な点が現れる理由として，攻撃者は，送信者になりすますための必要な情報がすべて入手できていない状態でメールを作成しているためであると考えられる．さらには日本人を対象とした攻撃メールを日本語のネイティブスピーカーでない攻撃者が作成している場合もあるため，言葉遣いなどに不自然な点が見られることも多い．提案システムでは，3.3.1 節で説明した通り，過去のメールを参照してメールを生成するため，送信者になりすますための情報は全て手に入れており，それをすべて活用して，訓練メールを生成しているため，外部の攻撃者では作成し得ないなりすましメール文章が生成できてしまう．そこで,「不自然な点」を訓練メールにも再現し，訓練メールを見分けるためのヒントとする．これによって，被訓練者は訓練を繰り返すにつれ，受信メールを注意深く観察するようになって「不自然な点」の発見が可能となり，標的型攻撃メールを見分けられるようになると考えられる．

4. 評価実験

4 章では受信 BOX 分析機能の有用性を示すために行った評価実験について述べる

4.1 評価実験の目的

本実験は受信 BOX の中身を分析し，送信者ごとの特徴を分析する機能について実装を行い，想定通り冒頭部分と末尾部分の特徴を取得できているかどうかを確認することが目的である．

4.2 実験環境

本実験は評価実験用のプロトタイプを用いて行う．評価用のプロトタイプとは，複数の文章(文章群)を与えた際に降順, 昇順で同 1 行の数え上げを行い, その文章群の降順, 昇順特徴を抽出するプログラムのことである．実験用のプロトタイプは Java を用いて実装を行った．また，使用する受信者メールアドレスは我々のものを使用した．

4.3 実験手順

本実験ではまず，メール群のパターンを 5 つ用意した．

1. 送信者 A からメーリングリスト a 宛に送られ，所属する個人に届いたメール群
2. 送信者 A からメーリングリスト b 宛に送られ，所属する個人に届いたメール群
3. 送信者 B から様々なメーリングリスト宛に送られ所属する個人に届いたメール群
4. 送信者 B からメーリングリスト c 宛に送られ，所属する個人に届いたメール群
5. 全く関連のないメール文章を寄せ集めた文章群

また一つの文章群を形成するメールは基本 20 件とし，1,2,3 に関してはその送信者の最新メールから 20 件，4 に関しては 17 件，5 に関しては送信者関係なく受信 BOX の最新メールから 20 件とした．

Vol.2017-DPS-170 No.13
Vol.2017-CSEC-76 No.13
2017/3/2

4.4 実験結果

メールの 1 行ごとの最多出現文の個数を最多フレーズ数として，昇順・降順それぞれについてこれをもとに評価を行う．

(1) **パターン 1**

(ア) 昇順

表 2: パターン 1・昇順

	最多フレーズ数	備考
1 行目	19	宛先を示す
2 行目	15	名乗りを示す
3 行目以降	2~4	

昇順では表 2 のように 1 行目の最多フレーズ数は 19，これは宛先を示す文であった．2 行目では 15 となった．これは名乗りを示す文であった．3 行目以降は 2~4 となった．

(イ) 降順

表 3: パターン 1・降順

	最多フレーズ数	備考
1~8 行目	20	署名部を示す
9 行目以降	2~3	

降順では表 3 のように 1~8 行目の最多フレーズ数が 20，これは署名部分を示す文であった．9 行目以降は 2~3 となった．

(2) **パターン 2**

(ア) 昇順

表 4: パターン 2・昇順

	最多フレーズ数	備考
1 行目	19	宛先を示す
2 行目以降	2~7	

昇順では表 4 のように 1 行目の最多フレーズ数は 19，これは宛先を示す文であった．2 行目以降は 2~7 となった．

(イ) 降順

表 5: パターン 2・降順

	最多フレーズ数	備考
1~8 行目	14	署名部を示す
9 行目	8	「＞＞」
10 行目以降	2~7	

降順では表 5 のように 1~8 行目の最多フレーズ数が 14，これは署名部分を示す文章であった．9 行目は 8 であった，これは返信メールのオリジナルメール引用部分につく「＞」が 2 つついた「＞＞」であった．以降は 2~7 となった．

(3) **パターン 3**

(ア) 昇順

表 6: パターン 3・昇順

	最多フレーズ数	備考
1 行目	4	宛先を示す
2 行目	17	名乗りを示す
3 行目以降	2~3	

昇順では 1 行目の最多フレーズ数は 4 と少なかった，最多となったのは宛先を示す文であった．2 行目は 17 となり，名乗りの文であった．3 行目以降は 2~3 であった

(イ) 降順

表 7: パターン 3・降順

	最多フレーズ数	備考
1~9 行目	14	署名部を示す
10~17 行目	4	下記参照
18 行目以降	2~3	

降順では 1~9 行目の最多フレーズ数が 14，これは主に署名部分を示す文章であった．10~17 行目は 4 であった．これは返信メールのオリジナルメールについていた署名部分であった．以降は 2~3 であった．

(4) **パターン 4**

(ア) 昇順

表 8: パターン 4・昇順

	最多フレーズ数	備考
1 行目	7	宛先を示す
2 行目	16	名乗りを示す
3 行目以降	2~7	

昇順では表 8 のように 1 行目の最多フレーズ数は 7 となった，最多となったのは宛先を示す文であった．2 行目は 16 となり，名乗りの文であった．3 行目以降は 2~7 であった．

(イ) 降順

表 9: パターン 4・降順

	最多フレーズ数	備考
1~9 行目	17	署名部を示す
10~17 行目	5	下記参照
18 行目以降	2~4	

降順では表 9 のように 1~9 行目の最多フレーズ数が 17，これは主に署名部分を示す文章であった．10~17 行目は 5 であった．これは返信メールのオリジナルメールについていた署名部分であった．以降は 2~4 であった．

(5) **パターン 5**

パターン 5 は昇順降順ともに最多フレーズ数は 2 以下となった．

5. 考察

パターン 1 は昇順降順ともに大きく特徴が現れたパターンとなっていて，昇順特徴からは宛先と名乗り，降順特徴からは署名部分の特徴を得ることができた．これは送信者 A が「就職支援担当」として「就職支援に関するメールを送信するため」のメーリングリスト a 経由で送信してくるメール内容はほとんど一貫していて，「就職支援」関連のものが多くなっている．よって，宛先にブレが生じにくかったのではないかと考えられる．例を挙げると，就職に関連のある学生は「B3」もしくは「M1」となっているため，図 10 のような書き出しで始まるメールが多かった．

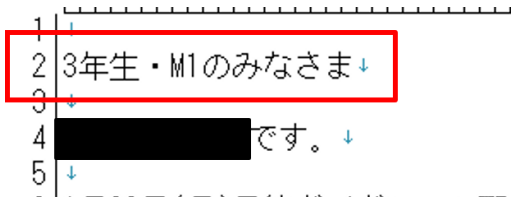

図 10: メールの出だし例

また，9 行目,11 行目の昇順特徴の最多フレーズ数が 4 となっていたが，この部分については「申し込み方法」を表す部分になっていて，同じような申し込み方法を促す文章が 4 件同行に書いてあったのではないかと推測される．他には，メール文章の全行数が偶然同一の値で署名部分がいくつか重複していた．降順特徴に関しては署名部分以外に特徴が大きく出たところはなく，偶然同一行に同一文が書いてあった部分が最多フレーズ数 2 や 3 となって現れたと考えられる．

パターン 2 については昇順では 1 行目，主には宛先の部分の特徴が大きく顕れ，降順では署名部分に特徴が顕れた．しかし署名部分に関しては，パターン 1 よりも最多フレーズ数が下がっている．これに関しては，メーリングリスト b は「全学生」を対象としたメールを送信するようなものとなっていて，そのメールの用途によっては送信者 A の肩書が変化するため，降順における署名部分に多少影響を与え，昇順における名乗り部分に関してはほとんど特徴に顕れなかったのではないかと考えられる．

パターン 3 に関しては送信者 B のメールをメーリングリストに関係なく最新のメールから分析した結果であるが，やはり，複数のメーリングリストにまたがって分析した場合では，送信対象が安定しないため，昇順特徴における宛先部分についてはほとんど特徴を取得することができなかった．

そこでパターン 4 では送信者 B のメールのうち，メーリングリスト c のみに絞って分析を行っている．昇順 1 行目の最多フレーズはパターン 3 と同じ宛先を示す部分であったが，パターン 3 が 4/20 であったのに対し，パターン 4 では 7/17 となりパターン 3 に比べ，大きな特徴が現れたものの，過半数には至らなかった．この理由は，メーリングリスト c では，ある程度個人宛のものも，メーリングリスト c への関連が高い場合はそのメーリングリストを介して，全体に伝えつつ個人宛に送るという使い方がされていたため，宛先に個人名が書いてあるものが何件か存在したためである．降順特徴についてはパターン 3，4 ともに，署名部分の特徴を取得することができた．さらに，返信メールに自動付加されているオリジナルメールの署名部分の特徴もある程度取得していた．

パターン 5 に関しては，送信者に関係なく最新のメールを 20 件分析したらどうなるのかを検証したが，昇順，降順ともに特徴は顕れなかった．唯一最多フレーズ数が 2 となった部分に関しては同一送信者からのものが偶然混ざっていたからである．

全体を通して，返信メールに自動付加されるオリジナルメール部分には「>」のような記号が付加されるが，今回は「>」のみがついている部分，つまり，オリジナルメールでは空行であった部分については詰めずに実験を行った．その結果として，オリジナルメールに空行が多く入っていた場合，「>」しかない行が多く存在し，元々は別の行であるはずの「>」同士が，同行のフレーズとして数えられてしまい，パターンによってはそれが 7, 8 などの数値を出していたことから，特徴として捉えられてしまう恐れがあることがわかった．よって，この現象については対策を行う必要があると考えられる．

結論として今回の実験ではパターン 5 以外の全てのパターンで昇順降順ともに特徴をある程度取得することが可能であることがわかった．さらに，パターン 1 と 2 を比較した場合，宛先範囲が狭まっているパターン 1 の方がより特徴が顕著に現れていたことや，パターン 3 と 4 を比較した場合，関連性の高いメールを多く使って分析を行ったパターン 4 の方が特徴が現れていたことから，話題が滝に渡るようなメール群よりも話題が絞られているメール群を用いた方が顕著な特徴が出る可能性が出ることがわかった．

6. まとめ

本稿では受信 BOX にある受信メールを送信者ごとに分類し，その送信者ごとのメール文章の特徴を分析し，普段被訓練者が受信するメールに類似した擬似攻撃メールを自動生成し，受信メールのように表示することで，効果の高い訓練を行うことが出来るシステムを提案し，メールの特徴を分析する部分についての実装を行った提案システムに対し，我々の受信している実際の受信メールを用いて，特

徴を取得しメール文章生成ができるかどうかを検証した．提案システムによって特徴の抽出は可能であるが，より顕著な特徴を抽出するためには話題が絞られているメール群を使ったほうが良いという結果が得られた．

今後はメール群内の話題を絞るために，メーリングリストごとにメールを分割できるような機能の追加や，自動生成したメールと実際のメールとの差を確認するための評価実験，および自動生成メールを使用した標的型攻撃対策訓練を実施し，その評価を行う必要があると考えられる．

参考文献

[1] “平成 28 年上半期におけるサイバー空間をめぐる脅威情勢等について”. http://www.npa.go.jp/kanbou/cybersecurity/H28_kami_jousei.pdf, (参照 2017-02-07).
[2] “年金機構の 125 万件情報流出 職員，ウイルスメール開封,” http://www.nikkei.com/article/DGXLASDG01HCD_R00C15A6000000/, 日本経済新聞, 2015/6/1, (参照 2017-02-07).
[3] “[詳報]JTB を襲った標的型攻撃,” http://itpro.nikkeibp.co.jp/atcl/column/14/346926/061500549/?rt=nocnt, ITpro by 日経コンピュータ, 2016/6/15, (参照 2017-02-07).
[4] JPCERT/CC, “2008 年度 IT セキュリティ予防接種調査報告書,” http://www.jpcert.or.jp/research/inoculation2008.html, 2009, (参照 2017-02-07).
[5] JPCERT/CC, “2009 年度 IT セキュリティ予防接種調査報告書,” http://www.jpcert.or.jp/research/inoculation2009.html, 2011, (参照 2017-02-07).
[6] 内田勝也, “大規模情報漏えいにおけるセキュリティマネジメントからの考察”, 情報処理学会第 78 回全国大会講演論文集, Vol.2016, pp.3_507-3_508 (2016).
[7] IPA, “IPA テクニカルウォッチ「標的型攻撃メールの例と見分け方」,” https://www.ipa.go.jp/files/000043331.pdf, (参照 2017-02-07).