

# 送受信者間の関係に基づく誤送信メール防止システムの試作

A System for Preventing Human Error in Sending Email based on Human Relationship

辻野友孝\*<sup>1</sup>

Tomotaka TSUJINO

清水堅\*<sup>1</sup>

Ken SHIMIZU

平田紀史\*<sup>1</sup>

Norihumi HIRATA

白松俊\*<sup>1</sup>

Shun SHIRAMATSU

大園忠親\*<sup>1</sup>

Tadachika OZONO

新谷虎松\*<sup>1</sup>

Toramatsu SHINTANI

\*<sup>1</sup>名古屋工業大学大学院工学研究科情報工学専攻

Dept. of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology

This paper presents a system for estimating human error in sending email. We implemented the system based on human relationship from email histories using machine learning. We particularly dealt with superior and subordinate categories as the human relationship. When the human relations of the automatically predicted are different from the human relations of the addressee, the system displays a warning to prevent sending error. This system was able to estimate a false transmission of a message email with precision of 93%.

## 1. はじめに

近年、誤操作による情報漏洩が年々増加し問題となっている。2009年に発表された2008年情報セキュリティインシデントに関する調査報告書によると情報漏洩の全体の35%以上が誤操作によるものと報告されている。中でも、メールによる情報漏洩の件数は多い。メールは簡単に相手に情報を伝達することができるため、日常のコミュニケーション手段として頻りに利用されている。しかしメールは、宛先を間違える、添付ファイルを間違えるなどの操作のミスで簡単に情報が漏洩してしまう。

本稿ではメールの送受信者間の関係を利用することで、メールの誤送信を防止するシステムを提案する。本稿では誤送信メールとは、メールの件名や内容が受信者と送信者の関係に適切ではないメールと定義する。本研究では、メールの送信者と受信者の人間関係に着目することにより誤送信メールの推定を行う。人間関係とは上司と部下の上下関係、社内と社外の所属関係、知り合いと友達の友人関係などの関係を指す。部下には連絡のメール、上司には報告のメールを送るなど、送受信者の関係が異なればメールの内容も異なると考えられる。そのため、複数の人間関係を用いる事で、送受信者間の関係とその場合における適切な宛先のメールが定義できると考えられる。本研究では人間関係の抽出には送信メール履歴の本文と件名を利用する。送信メールの特徴と送信メール履歴より得られた人間関係の特徴を利用しシステムは送受信者間の人間関係を推定する。システムによって推定された送受信者間の人間関係と送信メールの宛先との人間関係が異なった場合に誤送信であると判断しユーザに警告文を提示することでメールの誤送信を防ぐ。

本稿では第2節で本研究で用いた誤送信防止手法について述べる。第3節でシステムの流れを説明する。第4節で実験の結果をもとに本研究の手法を評価し、第5節で本稿をまとめ、第6節で今後の課題について述べる。

## 2. 誤送信防止手法

### 2.1 既存手法

現在、誤送信を防ぐための様々な方法が存在する。送信時に防ぐ方法として、送信時に計算問題を出し、計算に正解する事によりメールが送信される方法がある。この方法では送信時に正常な思考能力があるか確認する事ができる。他にも、送信時にメールの誤送信を防止する方法として、代表的な手法に、添付ファイルを添付した場合に警告文を提示する方法、キーワードマッチングにより重要な内容かどうか判断し警告する方法、宛先数が一定数以上超えると警告を表示する方法、異なるドメインにメールを送信する場合警告をする方法などがある。また送信後に対応できる方法としてメールを暗号化する方法、メール本文にはwebサイトのアドレスだけ記載し、本文はそのアドレスにアクセスしないと読めないようにする方法などがある。

しかし、既存の手法ではメールの内容を加味しておらず、精度良く誤送信メールを検出する事が困難である。そこで、多くの既存手法では再現率を向上させるため、キーワードマッチングを使いメールの重要度が高い場合に警告を発するようにしている。そのため実際には誤送信ではない多くのメールに対し警告を表示してしまい、ユーザは警告を軽視してしまうという欠点が生じている。それに対処して、信頼できる送り先をホワイトリストに追加してユーザへの警告の量を減らしたり、リスクを色分けして表示するシステムも存在する。

### 2.2 送受信者間の関係

本研究ではメールの内容に着目した誤送信メールの検出方法を提案する。しかし、各宛先ごとに対してのメールの適切さを利用する場合2つの問題が考えられる。1つ目はメールの本文は短かく、メールの件名や本文中に宛先を特定する特徴が現れない場合があることである。例えば「先ほどの件、了解しました」とメールの本文に書いてある場合、誰に送ったのか人間でも判断することができない。2つ目はメールでは同じ相手と様々な話題でメールをする点である。そのため、単純にキーワードで相手を特定することは困難である。

本研究では誤送信防止に人間関係を利用する。誤送信とは本質的には人間関係の違う相手にメールを送信してしまう事で

連絡先: 辻野友孝, 名古屋工業大学大学院 工学研究科 情報工学専攻, 〒466-8555 愛知県名古屋市昭和区御器所町, Tel:052-733-6550, Fax:052-735-5584, E-Mail:tomotaka.tsujino@toralab.ics.nitech.ac.jp

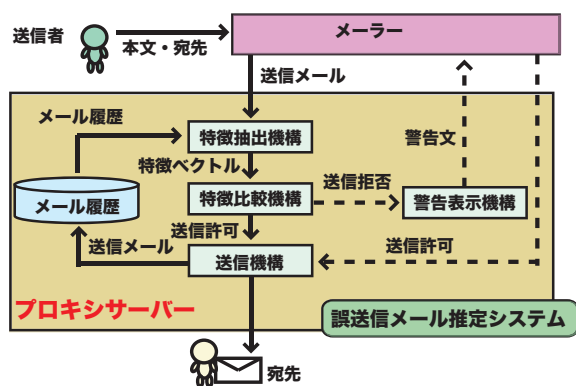


図 1: システム構成図

ある．送信者と一人一人の宛先との人間関係は異なっており，人間関係が異なる相手にメールを送信してしまうことが問題である．そこで，宛先への適切さではなく，人間関係クラスへの適切さを利用して誤送信を推定する．メールの本文と内容から宛先を推定するのは困難であるが，宛先がどの人間関係のクラスに属するか推定することは可能であると考えられる．

送受信者間には様々な関係がある．例えば社内 - 社外関係，友人関係，恋人関係，目上 - 目下関係などがある．相手により人間関係は異なるので複数の人間関係を利用することでユーザを識別することができると考えられる．また，所属関係や友人関係は比較的判別しやすい人間関係で，メールアドレスのドメイン名や文章の語尾から判別することができる．しかし目上 - 目下関係はメールの内容を考慮しなければ非常に推定することが困難であると考えられる．そこで本稿ではシステムが目上 - 目下関係を判別について述べる．

### 3. 誤送信メール防止システム

本システムでは送信メールの件名と本文と送信メール履歴を比較することで誤送信を判定する．本システムでは送信相手との人間関係は既知とする．本システムではメールから作成した特徴ベクトルと送信履歴の各人間関係のクラスから作成した特徴ベクトルを比較することで相手との人間関係を推定する．推定結果の人間関係と宛先との人間関係が異なる場合誤送信と判定し，ユーザに警告を出す．

#### 3.1 特徴ベクトル

本システムはメールを MeCab<sup>\*1</sup> を利用し形態素解析を行い特徴ベクトルを作成する．特徴ベクトルは  $tf-idf$  を用いて作成した． $tf-idf$  値は次の式で求められる．

$$tf(t) = \frac{t}{\sum t} \quad (1)$$

$$idf(t, m) = \log\left(1 + \frac{M}{m : m \subset t}\right) \quad (2)$$

$$tf-idf(t, m) = tf(t, m) \times idf(t) \quad (3)$$

これはメール  $m$  における語  $t$  についての  $tf-idf$  値を表している． $tf(t, m)$  はメール  $m$  での語  $t$  の出現頻度， $idf$  は  $\log(M/df(t) + 1)$  で， $M$  は全メール数， $df(t)$  は語  $t$  が含まれるメール数である．

\*1 <http://mecab.sourceforge.net>

#### 3.2 特徴ベクトル比較

特徴比較機構では SVM(Support Vector Machine) という学習を用いる手法で，本システムでは，データマイニングツール Weka<sup>\*2</sup> のライブラリを利用する事により特徴ベクトルの比較を行った [1]．Weka は，オープンソースのデータマイニングのフリーソフトである．本システムでは SVM に `weka.classifiers.functions.SMO` を用いた．特徴比較機構では，送信履歴の各人間関係の特徴ベクトルを SVM にトレーニングデータとして学習させ，送信メールの特徴ベクトルをテストデータとしてどの人間関係に属するか推定する．また，送信履歴にはシステムが警告を出さなかったメールとユーザが誤送信でないと判断したメールが保存されている．

#### 3.3 システム構成

本研究では学習を利用した誤送信メール推定システムの試作をした．本システムの構成を図 1 に示す．本システムはプロキシサーバ上に構築した．まず，ユーザは任意のメールエディタによりメールを作成し送信する．送信されたメールはプロキシサーバに届く．プロキシサーバ上にあるシステムはメールの件名と本文を抽出する．抽出した文を特徴抽出機構で形態素解析を行い，出現した形態素の  $tf-idf$  値から特徴ベクトルを作成する．送信メールから作成した特徴ベクトルと，メール履歴から作成した各人間関係の特徴ベクトルを特徴比較機構で比較を行う．システムは推定結果と実際の宛先の人間関係が等しい場合，メールの送信を許可し送信機構によりメールが送信する．人間関係が等しくなく誤送信と推定した場合はシステムは送信を拒否し，警告文表示機構がユーザに警告文を表示する．警告が間違いであった場合，システムはユーザにより送信許可が出されることで送信機構からメールが送信する．送信されたメールは送信メール履歴に保存され，次回からの推定に利用される．

### 4. 実験と評価

本研究では目上 - 目下関係に着目し目上クラスと目下クラスの識別をした．研究室でのメールを利用したため，メールの本文ではどちらの人間関係クラスのメールにも敬語を使用している．また，送信相手は同じ研究室に所属する人が大半であるため，人間関係クラスごとの話題の違いは少なく話題によるクラスの判別が困難であると考えられる．

実験データは大学院生 1 名の研究室 PC で送信したメール 200 通を対象とした．本実験では学生のメールを利用するので“目上クラス”は教員と先輩，“目下クラス”は同級生，後輩とした．メーリングリストのような複数にメールを送信する場合は宛先の中で最も地位の高い人が属する人間関係クラスをメールの人間関係クラスとした．以上のルールを適用すると，送信メール履歴 200 通のうち 127 通が目上クラス，73 通が目下クラスとなった．以下にそれぞれのクラスのメールの例を示す．

- 目上クラスメール
  - － 修正した論文を添付します．確認よろしくおねがいします．
- 目下クラスメール
  - － 明後日 18 日水曜日に部屋の配置換えをします．10 時から配置換えをするので必ず来てください．

\*2 <http://www.cs.waikato.ac.nz/ml/weka/>

表 1: 利用する品詞

|   | 名詞   |        |      | 助詞 | 他品詞 |
|---|------|--------|------|----|-----|
|   | 置換なし | 固有名詞置換 | 名詞置換 |    |     |
| A |      |        |      |    |     |
| B |      |        |      |    |     |
| C |      |        |      |    |     |
| D |      |        |      |    |     |
| E |      |        |      |    |     |
| F |      |        |      |    |     |
| G |      |        |      |    |     |
| H |      |        |      |    |     |

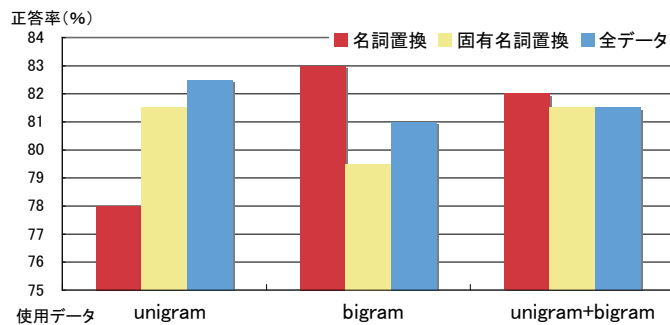


図 3: n-gram による違い

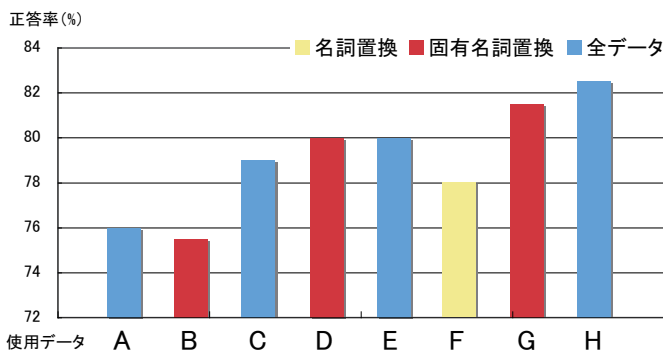


図 2: 品詞による違い

#### 4.1 品詞推定

特徴ベクトルを作成する際に利用する形態素の品詞により推定精度の違いがどのように現れるのか実験した。品詞は助動詞、名詞、名詞 + 助動詞、全品詞の 4 通り行った。今回は 8 パターンを比較した。今回試したパターンを表 4.1 に示す。A は助動詞のみ、B と C は名詞、D と E は名詞 + 助動詞、F と G と H は全品詞である。また、B と D と G は固有名詞を記号に置換、H は名詞を記号に置換したものである。名詞や固有名詞を記号に置き換えることにより文章の抽象度を高くすることにより精度がどのように変化するか調査した。

図 2 に unigram での品詞による人間関係推定精度を示す。表の縦軸は精度を示し、品詞ごとの精度の違いを示している。結果より、全品詞を用いた場合が最も精度が良いことが分かる。このことより、情報の多い方が正解率が高いと考えられる。また、品詞の中では名詞が最も正解率が高い。名詞 + 助動詞の固有名詞を置換精度の差が無いことから、今回利用したデータは人間関係の違いによる話題の違いは少なかったものと考えられる。

#### 4.2 n-gram 推定

利用する形態素の数によりどのように推定精度の違いが出るのか実験した。今回は、unigram, bigram, unigram + bigram の 3 通りを比較した。unigram は形態素 1 つ、bigram は形態素を 2 つを連結させたもの、unigram + bigram は unigram と bigram を利用したものである。

図 3 に n-gram の違いによる結果を示す。結果は、unigram では名詞置換、固有名詞置換、置換無し順で正解率が高くなっていることが結果よりわかる。このことより、unigram では情

表 2: 名詞 (tf-idf)

| 目下クラス ( $c_1$ ) |        | 目上クラス ( $c_2$ ) |        |
|-----------------|--------|-----------------|--------|
| 単語              | tf-idf | 単語              | tf-idf |
| 例               | 16.5   | 辻野              | 56.7   |
| 連絡              | 12.8   | お願い             | 42.7   |
| メール             | 12.4   | 写真              | 41.1   |
| 夏               | 12.0   | 3点              | 32.5   |
| 問題              | 11.8   | 人               | 32.4   |

報量が多いほど正解率が高くなると考えられる。また、bigram の場合では名詞を置換した場合が最も正解率が高く 83 % の正解率で分類でき、置換無し、固有名詞置換と正解率が下がっている。目上クラスを目下クラスと推定してしまう場合を誤送信と考えた場合 bigram で名詞を置換した場合は 93 % で誤送信を検知する事ができた。これは、名詞を置換することで人間関係の違いによる口調の違いをできたため、正解率が高くなったと考えられる。また、固有名詞を置換した場合の正解率が低いのは、今回利用したデータが人間関係によらず共通の話題だったためと考えられる。話題が同じであるため、人間関係によらず共通の名詞が多数出現したためと考えられる。unigram + bigram はどの場合も同じような正解率となった。利用するメールが増加するに従い、判別が困難なメールも増加するため正解率が変化しなかったと考えられる。

#### 4.3 各クラスの特徴語

助詞と助動詞の組み合わせは人間の仲の良さや上下関係を推定するのに有用である [2][3]。そこで今回は、助詞と助動詞に着目し、*tf-idf* を利用し特徴語を抽出した。その結果、目上、目下、共にほぼ同じような特徴語が得られた。このことから、今回利用するメールは内容を考慮しなければ人間関係を推定することができないデータである事が分かる。本実験ではピリオド、カンマなどの 30 語をストップワードとし特徴語にならないようにした。また、出現確率を利用した方法では同じ人間関係クラスに属するメールの 20 % 以上に出現しない語も特徴語にならないようにした。

本実験では各クラスの特徴語を *tf-idf* 値を利用して抽出する。

実験の結果を表 2, 3 に示す。表 2 は unigram の品詞が名詞であるもののみから *tf-idf* を利用し特徴語を抽出した結果である。各クラスの *tf-idf* 値が高い 5 つの語である。目下クラスは「例」「連絡」など説明や相談に用いられる語が抽出された。目上クラスにはフォーマルな形式でメールを書いている

表 3: bigram (tf-idf)

| 目下クラス ( $c_1$ ) |        | 目上クラス ( $c_2$ ) |        |
|-----------------|--------|-----------------|--------|
| 単語              | tf-idf | 単語              | tf-idf |
| てください           | 29.9   | 行目              | 82.1   |
| して              | 29.8   | します             | 68.0   |
| 行目              | 27.3   | ました             | 64.0   |
| ました             | 23.2   | して              | 61.7   |
| よろしく            | 22.9   | 辻野です            | 49.0   |

ため送信者名が抽出された。

また、表 3 は bigram から *tf-idf* を利用し特徴語を抽出した結果である目下クラスは「てください」、「をよろしく」など指示を与えている表現が抽出された。目上クラスからは「します」、「ました」など相手からの指示に受け答えしている表現が抽出された。また、両方のクラスに「行目」、「して」という語が特徴語として抽出されている。これらはプログラムの授業の話題の中で頻繁に利用し、先生と同級生とメールのやり取りを多くしていた。これは、*tf-idf* を用いた場合には良く利用する語は *tf* 値が高くなるため特徴語として検出されやすく、目上 - 目下両方のクラスで特徴語とされてしまったためだと考えられる。これらの結果より、名詞のみを抜き出すよりも bigram を用いることで指示や応答などの表現を特徴語として抽出することができた。そのため、目上 - 目下関係の抽出には unigram より bigram の方が優れていると考えられる。

#### 4.4 誤判定

本システムで誤判定をしたメールは、本文の短いメール、転送メール、幹事メール、等があった。「了解しました」等の短いメールは特徴を抽出することができず、誤判定をしてしまった。また、転送メールは転送される文の特徴を抽出してしまうため、正しく人間関係を推定することができなかった。また、幹事として研究室全体にメールを送信する場合、自分より目上の人にもメールが送るが、若干だけ表現を用いていたため目下クラスとシステムは誤判定してしまった。

## 5. おわりに

本稿では送受信者間の関係に基づく誤送信メール防止システムについて述べた。本システムは意図した宛先との人間関係と異なる人間関係の宛先を指定したメールを検知し警告を表示する。本研究ではメールの件名と本文を形態素解析し、特徴ベクトルを生成した。送信メールの履歴から各人間関係の特徴ベクトルと送信メールの特徴ベクトルを比較することにより宛先との人間関係を推定する。研究室で利用したメールから人間関係に着目して目上、目下の 2 クラス分類を行った。その結果、bigram で名詞を置換した場合が最も正解率が高く 83 % の正解率で分類できた。これは名詞を置換することで各人間関係へ言い回しを抽出できた為だと考えられる。目上の人に目下に送信するメールを送信してしまう誤送信は 93 % で誤送信を判定することができた。

## 6. 今後の課題

今後の課題として、人間関係の自動抽出が挙げられる。本稿では送信者と受信者間の人間関係はあらかじめ既知であるとしたが、送信相手の数が増えると全てを定義するのは非常に困難になる。そこで、web から人物に関するキーワードを抽

出する方法が利用できると考えられる [4]。人物に関するキーワードから人間関係を抽出することができれば、送信相手の増加による人間関係の登録の手間を減らすことができると考えられる。

また、複数人のメールの履歴を利用することが挙げられる。本研究では個人のメール履歴を利用して人間関係の推定を行っている。個人の送信履歴のみを利用することにより、ユーザに適した誤送信の推定ができる。しかし、個人の履歴だけでは学習に時間がかかり、また、未知の人間関係に対しては効果を発揮しない。そこで複数人のメール履歴を利用して人間関係の推定を行うことが考えられるが、大きく 2 つの問題が考えられる。1 つ目にプライバシーの問題が挙げられる。メールには他人に見られてはいけない物が記載されている場合がある。本研究では名詞や固有名詞を記号に置き換えたが、それだけでプライバシーを完全に守りきれているのかはわからない。他のアプローチや、他の品詞の置換なども検討する必要がある。2 つ目に人によって同じ人間関係でも、接し方が違うということが考えられる。目下の人に敬語を使う人もいれば、使わない人もいる。同じ目下クラスの相手にでも、人によって接し方が違うのでどのように個人による違いを吸収するのが問題である。そのために重み付けをすることでユーザ自身の送信メール履歴を重要視する必要があると考えられる。

## 参考文献

- [1] Lan Witten, Eibe Frank, "Data Mining: Practical machine learning tools and techniques with Java implementations", Morgan Kaufmann Series in Data Management Systems, June 2005.
- [2] 辻野友孝, 伊藤太樹, 柿元宏晃, 加藤健太, 白松俊, 大園忠親, 新谷虎松, "メール履歴を利用した学習に基づく誤送信メール推定システムの試作", 第 72 回情報処理学会全国大会講演論文集, (2010)
- [3] 西原陽子, 砂山渡, 谷内田正彦, "発話テキストからの人間の中の良さと上下関係の推定", 電子情報通信学会論文誌. Vol.J91D, No.1 pp. 77-88
- [4] 松尾豊, 石塚満, "Web からの人物に関するキーワード抽出", 人工知能学会論文誌. 17 巻 3 号 D (2002 年)