卒業論文　　2011年度（平成22年度）

メールヘッダ解析によるなりすましメールの検知手法の提案

慶應義塾大学 総合政策学部

氏名：関根　冬輝

担当教員

慶應義塾大学 環境情報学部

村井 純

徳田 英幸

楠本 博之

中村 修

高汐 一紀

Rodney D. Van Meter III

植原 啓介

三次 仁

中澤 仁

武田 圭史

平成24年2月13日

卒業論文要旨 - 2011 年度 (平成 22 年度)

メールヘッダ解析によるなりすましメールの検知手法の提案

近年電子メールを利用して, 特定の個人や集団を狙った攻撃が増加している．無差別に攻撃をしかけ，情報の価値を気にせずに入手，あるいは漏洩を目的とした従来の攻撃とは違い，攻撃対象を限定して，価値の高い情報のみを狙うといったものである，これらの攻撃が増加することで，大企業や政府機関，官僚にまで被害が及ぶケースも確認されている．このように特定の個人あるいは少数の団体を標的とした攻撃は,「標的型攻撃」と呼ばれる．標的型攻撃の特徴として，先程述べたような攻撃対象を限定することの他，攻撃手法を標的に特化させたり，既存のマルウェア対策では攻撃そのものを検知する事が難しい，などと言った点が上げられる．また電子メールを利用した標的型攻撃の代表例として，攻撃対象の知人や有名企業，政府機関に偽装して電子メールを送りつける「なりすましメール」が存在する．このような電子メールを利用した標的型攻撃に対する効果的な対策を早急に作り上げることが求められている．

本研究では攻撃利用される電子メールの特性に着目し，電子メールに記載されるヘッダ情報を分析，利用する対策手法を提案する．手順としては，送られてきた電子メールのヘッダ情報を取得し，必要な情報を抜き出す．次に抜き出した情報を元に, アドレス毎にそれぞれの精密な特徴を作り出す．特徴付けが済んでいるアドレスや人物から新規にメールが送られてきた場合，そのメールヘッダ情報を取得し, 作り上げた特徴と比較する．特徴から一定以上外れた情報を持っている場合攻撃と判断し，ユーザに報告する．これによってユーザの関係者になりすましたメールを検知することが可能になると考える．また検知したものは標的型攻撃のサンプルとして，今後の標的型攻撃の研究に利用する事が可能となる．

キーワード:
1. なりすましメール, 2. 標的型攻撃, 3. セキュリティ, 4. 侵入検知,

慶應義塾大学 総合政策学部

関根　冬輝

Abstract of Bachelor's Thesis - Academic Year 2011

A proposal of spoofed e-mail detection by mail header analysis

Recently,the number of attacks using E-mail which targetted specific individuals and groups increased.Different from the attacks which tried to attacked indiscriminately and getted the all information,the attacks using E-mail was limitted target and getted the only information has a high value.Increasing that type of attacks,various type of damages occur and the damages extends not only to the individuals but also corporations,officials and government agencies.

The attacks using E-mail is called 'targeted attack'.Targeted attack has a characteristic that limited target,change attack form of attack to specialize target,difficult to detect by existing measures like firewalls.And targeted attack has a representative examples,named'Spoofed e-mail'.Therefore required to develop effectivemeasures against such targeted attacks urgently.

The research focused on the characteristics of e-mail used in attacks and propose a method to use measures that analyzed the information in the e-mail header.Get the e-mail header information and extract the necessary information.Next,Produce a precise characteristics for each address.Compared the new get e-mail header infomation and these precise characteristics,if new e-mail header infomation outsided a certain level from precise characteristics,reported to the user.And the e-mail that reported by this way can be used in future studies of targeted attacks.

Keywords :
1. Spoofed e-mail, 2.targeted attack, 3.Internet security, 4.Instrusion Detection

Keio University, Faculty of Policy Management

Fuyuki Sekine

目 次

第 1 章　序論 **1**
1.1　背景 1
1.2　本研究の目的 1
1.3　本論文の構成 2

第 2 章　標的型攻撃となりすましメール **3**
2.1　標的型攻撃 3
2.2　なりすましメール 4
2.3　本論文の着眼点 6

第 3 章　関連研究 **8**
3.1　予防接種による対策 8
3.2　送信ドメイン認証を利用した攻撃対策 8
3.2.1　MTA から見た SPAM メールの特徴 10
3.2.2　送信ドメイン認証を利用したメールフィルタリング手法提案 11
3.3　添付ファイルに注目した対策 11
3.4　まとめ 11

第 4 章　メールヘッダ情報を利用した攻撃検知 **13**
4.1　概要 13
4.2　メールヘッダ 13
4.3　本研究で使用するヘッダ情報 13
4.4　提案手法 15
4.4.1　過去のメールデータの取得 15
4.4.2　点数表記による信頼度の提示 15
4.5　まとめ 16

第 5 章　実装 **19**
5.1　ヘッダ情報の収集と分析 19
5.2　データのプロファイリング 19
5.3　送信者の信用度評価 19
5.3.1　信頼度の点数表記 20
5.3.2　事前調査の結果 20

第 6 章　評価 **26**
6.1　評価の概要 26
6.2　評価の結果 27
6.2.1　IP アドレス 27
6.2.2　電子メールクライアント 29
6.2.3　日時 30
6.2.4　得点評価 31
6.2.5　異常検知の評価 32
6.3　考察 34

第 7 章　結論 **37**
7.1　まとめ 37
7.2　今後の課題と展望 38

謝辞 **39**

図 目 次

2.1 標的型攻撃の仕組み 4
2.2 なりすましメールの実例 (出展 [5]) 5
2.3 攻撃によって誘導されるサイト (出展 [5]) 6
2.4 東日本大震災に乗じたなりすましメール (出展:[7]) 7

3.1 送信元 IP による認証 9
3.2 電子署名による認証 10

4.1 メールヘッダに表示される情報 14
4.2 データ取得からプロファイリングまでの流れ 16
4.3 プロファイリング及び点数評価のフローチャート 17
4.4 点数計算表 18

5.1 抽出された情報 20
5.2 ユーザ 1 のプロファイリング結果：IP アドレス 21
5.3 ユーザ 1 のプロファイリング結果：クライアント 21
5.4 ユーザ 1 のプロファイリング結果：曜日 22
5.5 ユーザ 1 のプロファイリング結果：時間 22
5.6 ユーザ 2 のプロファイリング結果：IP アドレス 23
5.7 ユーザ 2 のプロファイリング結果：クライアント 23
5.8 ユーザ 2 のプロファイリング結果：曜日 24
5.9 ユーザ 2 のプロファイリング結果：時間 24
5.10 採点結果 25

6.1 IP アドレスの調査結果 27
6.2 IP アドレスの得点 28
6.3 電子メールクライアントの調査結果 29
6.4 電子メールクライアントの得点 30
6.5 曜日の調査結果 31
6.6 曜日の得点 32
6.7 時間の調査結果 33
6.8 時間の得点 34
6.9 得点評価 35
6.10 差異評価 36

表 目 次

第1章　序論

本章では，現在社会で起こっている電子メールを利用して特定の対象を狙う攻撃の脅威について説明し，その上で本論文の目的と構成を記述する．

1.1　背景

近年従来の不特定多数のコンピュータを狙った攻撃に代わり，攻撃対象を少数グループや個人に絞り情報を狙う形の攻撃による被害が増加している． 攻撃の手法として，攻撃対象にブラウザを用いて特定のサイトへ誘導して悪意のあるプログラムをダウンロードさせるもの，サーバやクライアントに直接アクセスして攻撃するもの，電子メールを利用して不正プログラムを送りつけたり本文に特定のサイトへのアクセスを促すものなど，様々な種類が存在する．これらの攻撃の事は標的を決めて攻撃を行うことから「標的型攻撃」と呼ばれる．

その中でも電子メールを利用して他人になりすますことで攻撃対象への危機感を薄れさせ，攻撃を行う「なりすましメール」による被害は近年増加しつつある．電子メールは本来それぞれのユーザが独自のアドレスを持ち，ネットワークを利用して特定の個人間，あるいは対象と連絡を取り合うことを目的としているため，その特性に注目した攻撃者が，特定の攻撃対象に接触するために利用するケースが増加しているのである．電子メールそのものに攻撃能力は無いが，攻撃手段を運ぶ事を目的に利用される．

1.2　本研究の目的

本研究の目的は，電子メールを利用した標的型攻撃に対して有効な対策を提示，実装する事で，これらの攻撃からの被害の増加を防ぐ事である．

電子メールを利用した標的型攻撃は，メールそのものに攻撃能力はないものの，メールが攻撃に使われているものか，本当にただの電子メールなのかを正確に区別することは難しい．特に送信者を友人や政府機関などになりすまし，内容をアンケートなどに偽装されると，文章を見ただけでは攻撃と判断する事は非常に難しい．判断の方法として電子メールに記載されているヘッダ情報が利用可能ではあるが，ヘッダ情報は文字と数字のみで構成されるため，前提知識の無いユーザが見ても内容を理解するのは難しい．そこで本研究では，メールヘッダ情報の中から特定の部分を抜き出し，各アドレス毎に特徴付けを行う．これを利用して，同じアドレスから送られてきた新規メールのヘッダ情報と特徴付け

されたデータを比較し，送信者の信憑性を点数で表記する．これによって電子メールを利用した標的型攻撃への有効な対策を作り上げる事が可能となる．

1.3 本論文の構成

本論文は，全 7 章で構成される．まず第 2 章において，標的型攻撃の説明と実例を挙げ，本研究で取り扱う標的型攻撃の形について述べる．次に第 3 章では，既存の標的型攻撃への対策に関しての論文と，本研究にて扱う標的型攻撃の実例の特徴に関係する技術を扱った論文を述べる．続いて第 4 章では本研究が提案する攻撃への対策手法と，評価基準について述べる．第 5 章ではその実装方法，第 6 章では評価と考察を述べ，第 7 章では結論として，本論文のまとめと今後の展望について述べる．

第2章　標的型攻撃となりすましメール

本章では，標的型攻撃についての説明と実例を記述すると共に，本研究で扱う攻撃の説明を記述する．

2.1　標的型攻撃

標的型攻撃とは，従来の無差別に攻撃を仕掛ける手法とは異なり，対象を限定して行う攻撃手法である．2007 年に JPCERT/CC より発表された「標的型攻撃について」[1] にある定義によると,「情報セキュリティ上の攻撃で，無差別に攻撃が行われるものではなく，特定の組織あるいはグループを標的としたもの，攻撃対象となる組織あるいはグループに特化した工夫が行われる事もある」とある．標的に対して様々な手法でアクセスをし，攻撃用のコードを含んだ PDF 形式や Word，Exel のような Microsoft Office のファイルを送りつけたり，あるいは形式を偽装したマルウェアそのものをてメールに添付したり，題名や本文内に特定のサイトへの誘導 URL を含んだメールを送ることで，ブラウザを通して相手に意図したサーバにアクセスさせるなど，多様な攻撃形式を持つ．ここに，標的型攻撃の実例を図示する．

それぞれの攻撃に利用されたマルウェア，プログラムは解凍されたりアクセスされた場合，解凍先やアクセス先のコンピュータ内部にキーロガーやバックドアを作り上げる事で，対象の PC 内に保存されているパスワードなどの個人情報や，攻撃者の狙っている情報を盗み出すといった働きを持っている．以下に標的型攻撃の実例を述べる．前章でも述べたように，標的型攻撃は攻撃対象に合わせて形を変えるため，様々なケースが存在する．

2011 年 11 月 9 日，総合機械メーカー「三菱重工」（東京）がサイバー攻撃を受け，コンピュタ 83 台がウイルスに感染，うちサーバ 2 台から情報が漏洩した形跡が発見された [2]．感染していないサーバの情報も，感染端末に勝手に移動されており，外部に送信された可能性も存在している．漏洩した情報は原子力発電プラントなどの設計情報や，防衛装備に関する情報など国家機密レベルの情報が含まれていた．

2011 年 11 月 25 日，経済産業省内の 20 人分のコンピュータに，情報漏洩を引き起こす効果を持つマルウェアの添付された電子メールが送付された [3]．幸いマルウェアによる被害はなく，情報は流出していないと発表されている．

2011 年 4 月 27 日，SONY と SCE(Sony Computer Entertainment America) から出された公表の中に， PSN(Play Station Network) から約 7700 万人分の個人情報とクレジットカード番号が流出した可能性がるという記述があった [4]．それから 7 日後の 5 月 2 日に SOE(Sony Online Entertainment) からも個人情報の流出が確認された．

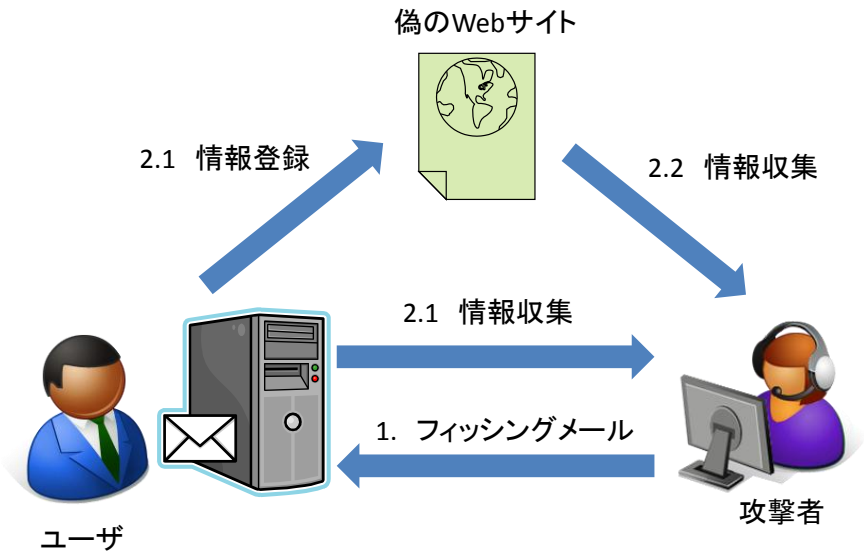

図 2.1: 標的型攻撃の仕組み

これらのように，大企業や政府機関などを中心に，情報価値の高い情報を目的としたケースが多い．

一般的に，標的型攻撃が行われる際に最初に攻撃対象へ接触する手段として利用されるのが「電子メール」である．これは電子メールに攻撃用のマルウェアを添付することで，一見無害な通信としてファイアーウォールなどのマルウェアの侵入を防ぐ壁を突破させる事が安易になるからである．更に電子メールを利用した標的型攻撃の中でも多く使われる手段に「なりすましメール」というものが存在する．

2.2 なりすましメール

特定の電子メールユーザに対して，送信者・題名・本文などを偽装してメールを送りつける．この際に送りつけられた電子メールを本研究では「なりすましメール」と定義する．標的となる受信者の知人になりすましたり，有名企業や政府機関を名乗ってアンケートを装って送りつけるなどのケースが多く見られる．以下になりすましメールの実例を図を用いて説明する．

図 2.2 のように有名な銀行を装って電子メールを送り，メール内に特定のサイト図 2.3 への誘導 URL を記載しておく．メールを受け取った標的が URL を踏み，パスワードな

図 2.2: なりすましメールの実例 (出展 [5])

どの個人情報を特定の場所に記入させることでそれを盗み出すといった手順である．

このように，メールの受信者は「ファイアーウォールなどが異常を検知しない」「見知った相手や会社名だから」などの理由から警戒心を解いてしまう．そうなれば内容を疑わずにメールを開き，添付されている攻撃用のマルウェアを解凍，あるいは記載されている URL をクリックしてしまうのである．これによって攻撃が開始される．

2011 年 11 月，経済産業省の職員宛に政府関係者になりすましたメールが送られてきた．このメールにはパソコン内の情報を抜き取るバックドア機能を持ったマルウェアが添付されていた．結果として情報漏洩は確認されなかったものの，省内の PC 約 20 台が感染したことが確認された．

2011 年 3 月 11 日に発生した東日本大震災に乗じた標的型攻撃も発見されている．IPA（独立行政法人情報処理推進機構）から出された「東日本大震災に乗じ た標的型攻撃メールによるサイバー攻撃の分析・調査報告書」[6]．によると，東日本大震災が発生してから，「発生した放射線量」などの震災に関係するメールを装い，エクセル形式やワード形式のファイルが添付された電子メールが届いたという．電子メールに添付された.doc ファイルを開こうとすると，Microsoft word がいったん指定したファイルを開き，その後そのファイルをすぐに閉じて改めてダミーの word ファイルが表示される．表示されたファイルをクリックすると，最終的にテキストファイルが開かれているだけに見えるが，開いたファ

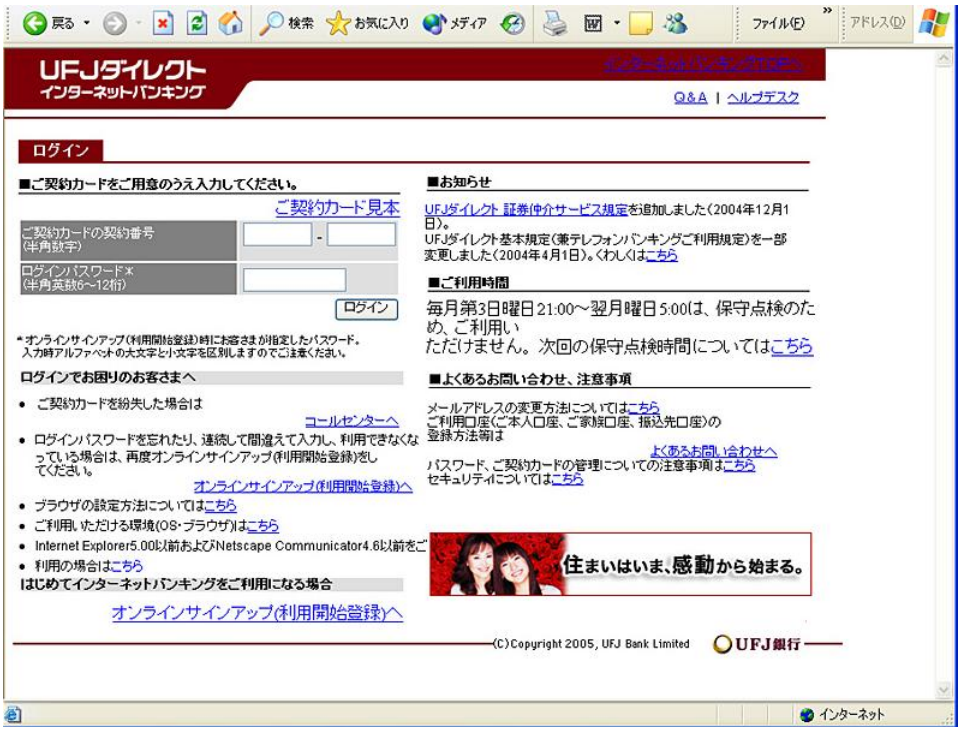

図 2.3: 攻撃によって誘導されるサイト (出展 [5])

イルと違う名前のダミーファイルが開かれていて，本命は C ドライブの中に．exe ファイル (バックドア) が作られているといったケースである．一見ただの悪戯に思えて，裏では情報を抜き出す準備が進められているのである．

このように，添付されたファイルを開くことで，バックドアの「入り口」がパソコン内につくられ，情報が流出してしまうといった手段である．

2.3 本論文の着眼点

標的型攻撃は攻撃対象を限定して行われるため，既存のマルウェア対策では様々な問題が生じる．まず攻撃対象を限定することで，第三者が攻撃を認知することが難しくなる．そのため標的となったユーザが攻撃されていることに気がつかず，個人情報などが流出して初めて気がつくといったケースが多い．さらには攻撃パターンが個人個人に特化されることで，攻撃の形態や攻撃に使われるコードが毎回パターンが変わってくる．これは従来のマルウェア対策に使われている，過去の攻撃データから攻撃パターンを検出してそれを検知した場合攻撃として検知するパターンマッチング方式では対応しきれない．また，攻撃が検知されにくいことから，攻撃報告も少なくなってしまうため，解析用のサンプルが少なく攻撃の特徴を見つけることが極めて難しい．

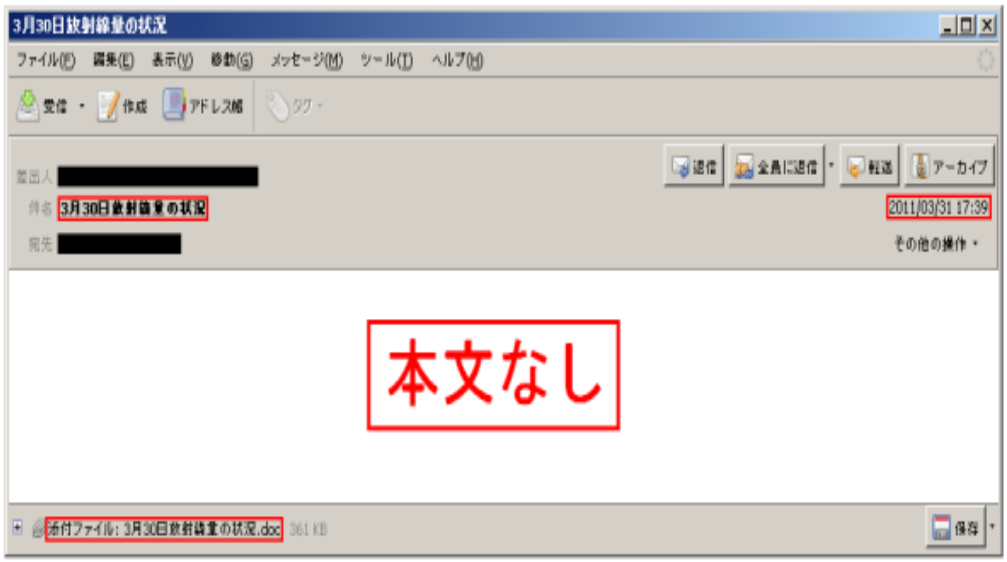

図 2.4: 東日本大震災に乗じたなりすましメール (出展:[7])

そこで本研究は，標的型攻撃が行われる際に電子メールを利用するケースが多いという点に着目した．電子メールを利用したなりすましなどの標的型攻撃を検知するために，電子メールの特性を利用する事で，有効な対応手段が作れると考えた．なりすましメールを有効的に検知する事で，攻撃として送られてくるマルウェアを確保して解析したり，誘導用の URL から攻撃に使うサイトを発見することが可能になる．それによって標的型攻撃の攻撃のサンプルが増え，今後の標的型攻撃への対策研究に大いに役立つのではないかと考える．

第3章　関連研究

本章では，標的型攻撃やなりすましメール，またメールヘッダに関する技術を取り扱う内容の既存研究を取り上げる．

3.1　予防接種による対策

標的型攻撃は行われる際に攻撃対象に特化される事が多い．これは，偽装する事で攻撃対象となるユーザの危険意識を薄れさせるためである．JPCERT コーディネーションセンターによって出された「標的型攻撃対策手法に関する調査報告書」[8] の中に,「予防接種によるセキュリティ意識向上」という方法がある．これは擬似標的型攻撃を各ユーザに体験してもらうことで，その攻撃に対する認知と警戒を促すと言ったものである．

最終的にメールやブラウザを扱うのはユーザである．そこでエンドユーザに対して実際の脅威に模したメール（ダミーメール）を受信させる．これによって実際に取り扱わせることでメールを用いた攻撃に対する危機意識を向上させ，適切な対処方法に関する理解を促す．ダミーメールには「無害な」添付ファイルをつけて送信し，メールを受信した対象者がこの添付ファイルを開いたか否かを記録すると同時に，開いてしまったユーザには事後同同様なメールに対して適切な取り扱いを行うよう促す。

この手法は，ユーザそのものに働きかけるので，技術的な対応だけでなく，使用するユーザの教育にも繋がってくる．結果として，攻撃に使われるメールの偽装内容に対して柔軟な対応を穫る事が可能となる．一方で，標的型攻撃は攻撃に利用する電子メールを「限りなく近づける」手法をとる為に，予防接種に使うダミーメールでは対応に限界がある．また，この対応は一般のユーザ全員に実施する場合，莫大な費用と時間を必要とするため，結果としてこの「予防接種」を受けられるのは，限られたユーザになってしまうのが問題である．

3.2　送信ドメイン認証を利用した攻撃対策

送信ドメイン認証とは，そのメールが，送信者と名乗っているアドレスに示されているドメイン（送信ドメイン）から本当に送られてきているかを確認するための仕組みである [9]．大きく分けて 2 種類ある．一つは送信元 IP アドレスを根拠に，正規のサーバから送られたかどうかを検証する技術．これには「SPF」「Sender IP」という技術がある．これは，エンベロープの送信者（SMTP プロトコルにおいて，MAIL FROM:の引数として与えられているアドレス）かメールヘッダ上の送信者（RFC で規定されている電子メー

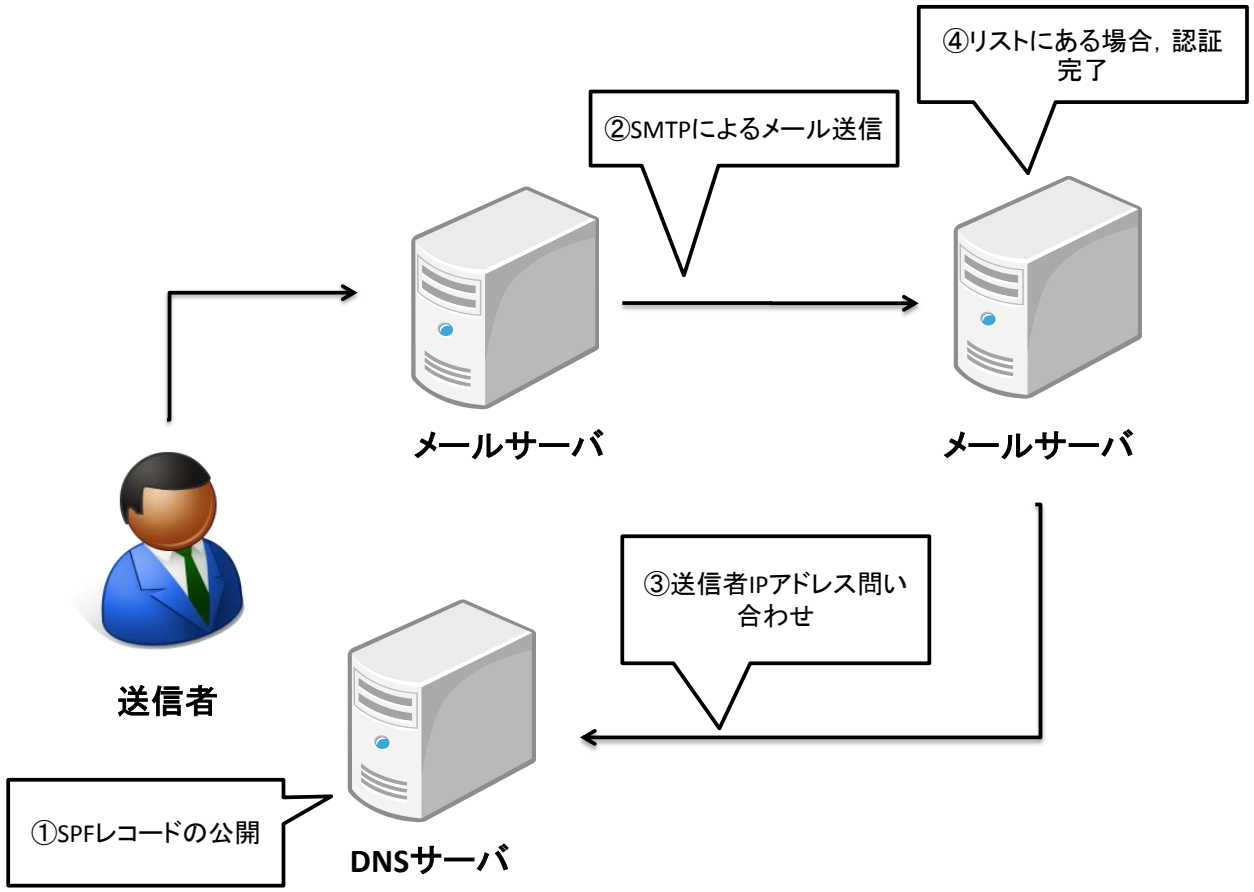

図 3.1: 送信元 IP による認証

ルヘッダに記録された送信者）のいずれかのアドレスを利用する．メールを送信する側では，送信者自身のドメインからメール送信を行う可能性のあるホストのリストを DNS に公開する．そしてメールを受信する側では，メールの受信中にメール送信者のドメイン部分を取り出し，DNS 側から公開されたレコードを読み出す事で，メール送信者の IP アドレスがそのリストに含まれているかどうかをチェックする．もし含まれていた場合，認証される．図 3.1

もう一つは送られたメールの中に電子署名を挿入し，その正当性を検証する技術であり，これには「DomainKey」「KDIM」といった技術がある．これらの技術は主にスパムメール対策に用いられる．両方ともあらかじめ電子署名の照合に利用する公開鍵を DNS サーバに設置し，電子メールにそのデータを元にした電子署名を付与して送信する．受け取った電子メールサーバはその電子署名の引数からドメインを取り出して，DNS に公開鍵を問い合わせることで，取得した公開鍵を使って電子署名を検証する．図 3.2

これらの送信ドメイン認証による対策は，なりすましを含む様々な標的型攻撃への対応

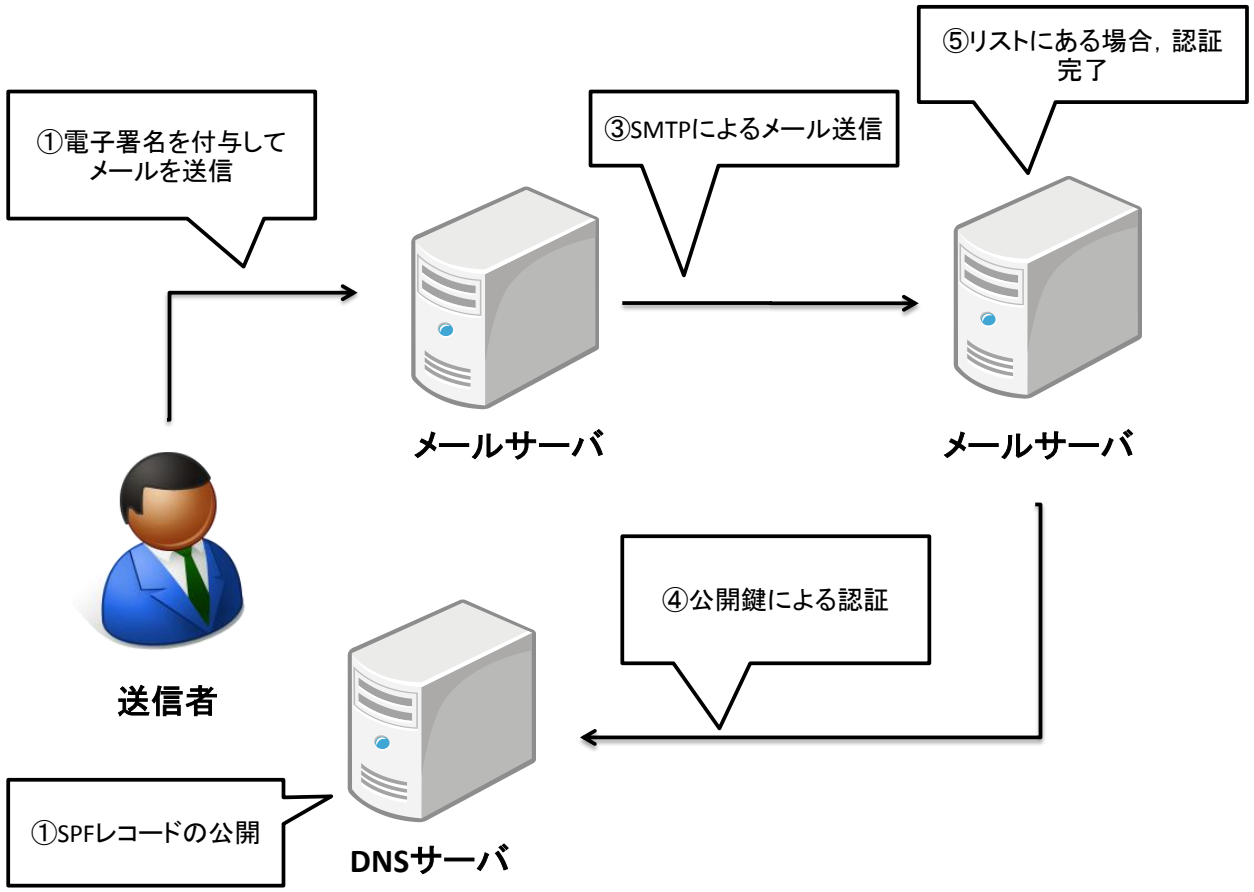

図 3.2: 電子署名による認証

が可能な対策手法であると考えられている．しかし，標的型攻撃はアドレスも偽装されているケースが存在するために，たとえ正規のサーバを利用して送られたものであっても，送信者本人かどうかの保証はない．そのため，ドメイン認証は完全な対策とは言えない．

3.2.1 MTA から見た SPAM メールの特徴

電子メールをインターネット内で配送するソフトウェアである「MTA」に注目し，SPAM メールに対して有効な対策である送信ドメイン認証，Greylisting，GreetPause，送信アドレス検証らの特徴をふまえた上で，メールヘッダ内の情報を取得して解析し，SPAM メールのヘッダ情報の特徴付けを行った研究である [10]．ヘッダ内に記載されている情報を可能な限り分析し，より強い SPAM メールの特徴を見つけ，その特徴を持つメールに対して MTA 側で設定を施す事で SPAM メールを検知，ブロックする流れを作り出している．

送信ドメイン認証は，Greylisting，GreetPause，送信者アドレス認証と比べて排除する

メールの条件を明確に定める事が可能であるが，SPAM 送信者の使い捨てドメインに対しても認証が通ってしまうという問題点を含んでいる．

3.2.2 送信ドメイン認証を利用したメールフィルタリング手法提案

送信者を偽装し，かつ指定したサイトへ誘導する URL を記載した悪意のある電子メールを送りつける「フィッシングメール」．これに対してメールヘッダを解析し，有効なフィルタリング手法を提案する研究である [11]．こちらの論文ではフィルタリング手法の中でもフィッシングメールに注目し，新たに定義された DNS 問い合わせベースの送信者信頼コストとコンテンツによるフィッシングメールを定量的に評価する手法を提案している．具体的には，メールヘッダ情報を解析し，送信者のメールアドレスのドメインが存在しない場合，送信者がメールを送信する際に経由した SMTP サーバののドメインが異なる場合などを評価して，それらを組み合わせたメール送信者の信用度を測定している．

送信ドメイン認証を利用した対策方法は非常に有効な手段であることがわかるが，この技術の導入には送信側サーバと受信側サーバの双方が対応しなければいけないという問題が存在する．これを解決しなければネットワーク全体で効果を期待するのは非常に厳しい．2005 年に始まった測定によると，2011 年 11 月辞典では，SPF[12] と DomainKeys[13] はそれぞれ 43.48 %，0.51 %と余り普及しているとは言えない．故に，長い時間を時間をかけてネットワーク全体に浸透させていかなければならず，素早い対応策として大きな効果は期待しにくい．

3.3 添付ファイルに注目した対策

攻撃用のメールに添付ファイルとして付けられるマルウェアに注目した対策である．攻撃に電子メールが使われても，メール自体には攻撃能力は無い．攻撃用のプログラムやマルウェアをファイル形式で添付し，ユーザの使用している PC 上で展開させる事で初めて効果を発揮するのである．そこで，添付されたファイルを切り離し，確認専用の形態端末や仮想空間上で展開し，中身を確認するといった対策である．[14]

この対策は，実際に攻撃のマルウェアを専用の端末で展開する事で，中身の確認が出来るとともに攻撃のサンプルとして確保する事が可能であるが，専門的な知識や手順が必要なことから，一般のユーザに浸透させる事は難しい．

3.4 まとめ

この章で，標的型攻撃やメールヘッダ情報に関連する様々な研究を挙げてきた．標的型攻撃に対しては，2.3 章で述べたように既存の攻撃対策手法では効果的な結果は期待しにくい．一方で，この章で述べたような標的型攻撃に対する新しい対策手法を生み出そうとする動きは，確実に増加してきている．しかしながら，予防接種にしろ送信ドメイン認証にしろ，コストや時間の関係から，一部のユーザには効果が期待出来ても，多くの一般

ユーザが活用出来る状態には至っていない．以上の事をふまえた上で，次章では本研究が提案する手法を述べる．

第4章　メールヘッダ情報を利用した攻撃検知

本章では、. 3 章で紹介した様々な既存の標的型攻撃攻撃，なりすましメールへの対策方法をふまえた上で，本研究が述べる対策手段の内容を記す.

4.1　概要

本研究では，攻撃が行われる際に利用される電子メールの特性に注目した．2 章で述べたように，電子メールを利用して攻撃対象に接触する場合，本文や題名・アドレスなどを攻撃対象に合わせて偽装する．これはユーザの攻撃に対する警戒心を薄めさせ，興味を引かせる事が目的となる．故に，題名や本文を見ただけでは本人からのメールかどうかを判断する事は難しい．そこで，攻撃時に利用される電子メールには，必ず「電子メールヘッダ情報」というものが存在していることを利用する，電子メーダヘッダ情報の中には様々な情報が含まれている．しかし，表記が英語 (専門的な用語) や記号，数字で構成されているため，前提知識の無いユーザが解読することは一般的には困難である．そこで本研究ではヘッダ情報を取得し，解析をしてアドレス毎に特徴付けを行い，その特徴とユーザから送られてきた新規メールを比較して信頼度を点数標記するといった形をとることで，前提知識の無いユーザにも理解する事が可能となる対策手法を提案する.

4.2　メールヘッダ

メールヘッダとは，電子メールが送られてくる際にそのメールの頭についている様々な情報をのせた部分である．ヘッダ情報には,「Return-Path」「X-Original-To」「Delivered-To」など送信者であるユーザの情報を特定する際に有益な情報となりうるものなどを含んでいる.

表記される形はクライアント側によって変わってくるが，内容は基本的に同じである．これらの情報を組み合わせて利用する事で，送信ユーザの識別が可能になってくる.

4.3　本研究で使用するヘッダ情報

本研究では,「Received」「X-Mailer」「Date」の領域にある情報を利用する.

Return-Path: ⟨*test@aaa.keio.ac.jp*⟩
Received: from example.aaa.wide.ad.jp (example.aaa.wide.ad.jp [aaa.bbb.xxx.yyy])
by ex.aaa.wide.ad.jp (Postfix) with ESMTPS id 45F3E2340AF;
Tue, 14 Dec 2010 10:40:34 +0900 (JST)
Received: from ex.aaa.keio.ac.jp ([aaa.bb.x.yyy])
by ex.aaa.wide.ad.jp with ESMTP; 21 Dec 2010 21:28:36 +0900
Received: from sample.aaa.keio.ac.jp ([aaa.bb.x.yyy])
by sample.sfc.keio.ac.jp with ESMTP; 14 Dec 2010 10:40:34 +0900
Message-ID: ⟨*4D109D71.3049604@aaa.keio.ac.jp*⟩
Date: Tue, 14 Dec 2010 10:40:34 +0900
From: test name ⟨*test@aaa.keio.ac.jp*⟩
User-Agent: Mozilla/5.0 (Windows; U; Windows NT 6.1; ja; rv:1.9.2.13)
Gecko/20101207 Thunderbird/3.1.7
MIME-Version: 1.0

図 4.1: メールヘッダに表示される情報

・送信元 IP アドレス

「Received」の領域内に，送信者元 IP アドレスが第 4 オクテットまで表記されている．本研究では，抜き出した送信元 IP アドレスの第 3 オクテットまでを利用する．これによってより厳密に送信元の地域を特定し，区別する事が可能になると考える．携帯電話やスマートフォンなどの携帯端末からのメールの場合，IP の大きな変化が予想されるが，デスクトップパソコンやノートパソコンの場合，利用可能なネットワーク環境が限られるため，送信元 IP が固定されるケースが多い．よって急なネットワーク環境の変化が少ない学生などにはプロファイリングをする際に強い効果を持つ要素になりうる．一方で，出張などで急にネットワーク環境が変化しやすい社会人などのプロファイリングにおいては，強い効果はあまり期待出来なくなる可能性が高い．

・メールクライアント情報

「X-Mailer」の領域内に，User-agent 情報が標記されている．これは送信者が利用したメールソフト，メールクライアントの情報を指し示す．各クライアントのバージョンも記載されるが，本研究ではバージョンは無視し，バージョンが違っても全て同一のクライアントと定義する．基本的に一つのアドレスを使用する際，固定されたクライアントを利用

する傾向にある．複数のアドレスを持っていも，同一のクライアント内で複数のアカウントを使い分けているケースが多い．その為，複数 PC アドレスを持たない人が電子メールを利用する場合，クライアントは固定されるケースが多い．また，複数使用するクライアントを持つユーザもいるが，職場と家のパソコンに入ってるメールクライアントが違う，あるいは家と学校で使用しているクライアントが違うと言った程度で，多くても 2 〜 3 つのケースがほぼ全てである．故に，送信者を判定する場合，このクライアント情報が違っていたら，攻撃を行うためのなりすましである可能性が非常に高いと言える．

・曜日と時間

「Date」の領域内に，メールの送信された日付，曜日，時間，タイムゾーンなどが表記されている．本研究では，曜日，時間（○○時までを）扱う．曜日は曜日別にメール受信数を調べ，時間は 1 時間単位でメール受信数を調べる．それによって得られたものをアドレス毎のプロファイリングに利用する．仕事以外では全くパソコンに触らないというユーザの場合，送信時間が仕事時間内に限定されるため，特徴が見られる可能性が高い．一方で常にパソコンを利用している学生などには，あまり強い特徴を得る事は難しくなると予想される．また曜日の場合，PC メールを仕事でメインに使う人などはメールが月曜日〜金曜日に集中し，週末はほとんど送受信が行われなくなると予想される．

4.4 提案手法

最終的な送信者の信頼度の評価を行うためには，アドレス毎の評価基準が必要となってくる．ここでアドレス毎に行うプロファイリング，点数評価の流れを記す．

4.4.1 過去のメールデータの取得

まず，ユーザが過去に取得した電子メールのデータからヘッダ情報を収集する．ユーザのローカル環境にある過去受信したメールのデータを取得し，アドレス毎にヘッダ情報を分割するスクリプトにかける．これによって取得したデータの中から，4.3 章で述べたデータをそれぞれ抜き出し，データベースに収集，蓄積する．これによって集められたデータを元に，アドレス毎のサンプリングを行い，できたデータを，送信者信頼度をはかる際に基準として利用する．

4.4.2 点数表記による信頼度の提示

次に新規にきたメールに対して，過去メールを取得した事のあるアドレスの場合，サンプリングした結果と比較し，送信者の信憑性を点数という形で提示する．これによってヘッダ情報の知識の無いユーザでも，結果を簡単に知る事が可能となる．

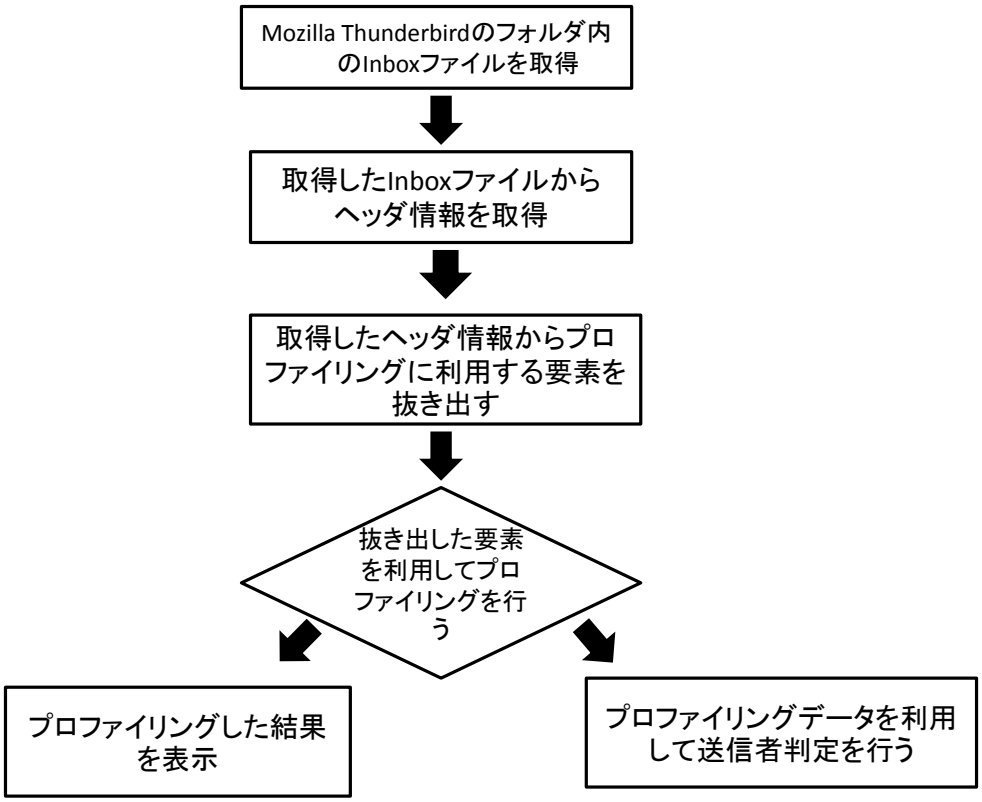

図 4.2: データ取得からプロファイリングまでの流れ

表記する点数の決め方だが，式を次のように定義する．まず点数の上限を 100 とし，利用するヘッダ情報にそれぞれ均等に点を振り分ける．4.3 章で記述したように，本研究で利用するヘッダ情報は 4 種類なので 1 つの項目毎に 25 点の配点となる．次に各要素の中で，どの要素がどれだけの割合を占めるかを計算する．以下に点数調節の手順を記した図を載せる．

曜日の情報を利用する場合を例に挙げると，過去 100 件分をサンプリングしたアドレスの曜日別送信数が，月曜日に 50 件送られてきていた場合，曜日の情報の中で「月曜日」は 50/100，つまり 1/2 を占める事になる．よって同じアドレスから新規にメールが送られてきた場合，それが月曜日に送られてきたものならば，25/2 の 12 点（端数は切り捨て）の信頼度を得る計算になる．この計算を 4 つの情報で行い，それぞれの合計点を総合的な信頼度としてユーザに報告する．この点数計算ならば，送信環境が違うアドレス毎に特徴を取る事が可能となり，柔軟な対応が可能になると考える．

4.5 まとめ

本章では，本研究が提案する対策手法と利用するデータの説明を行った．次章では，本章で述べた対策手法の実装環境を記す．

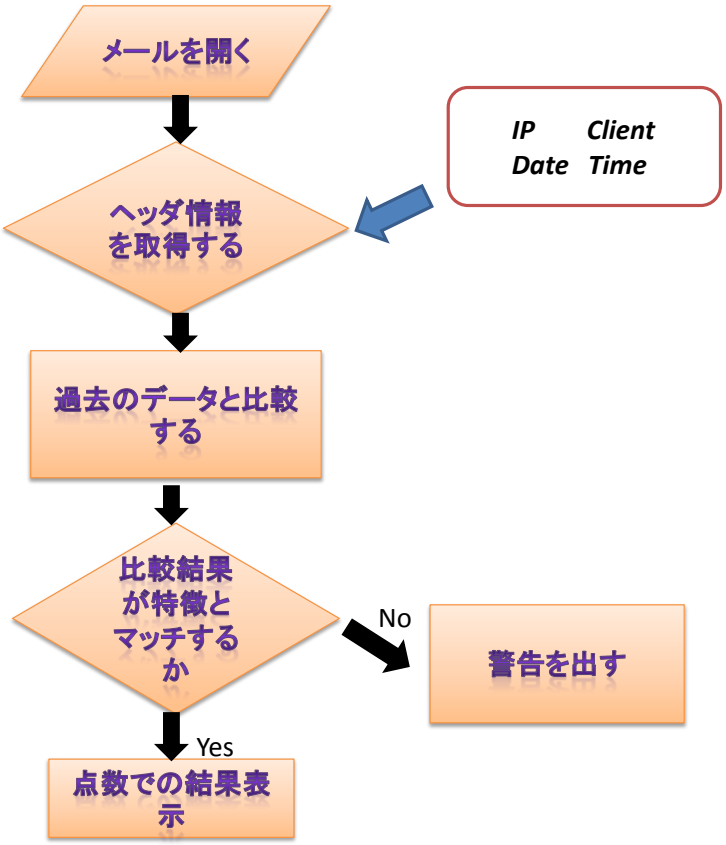

図 4.3: プロファイリング及び点数評価のフローチャート

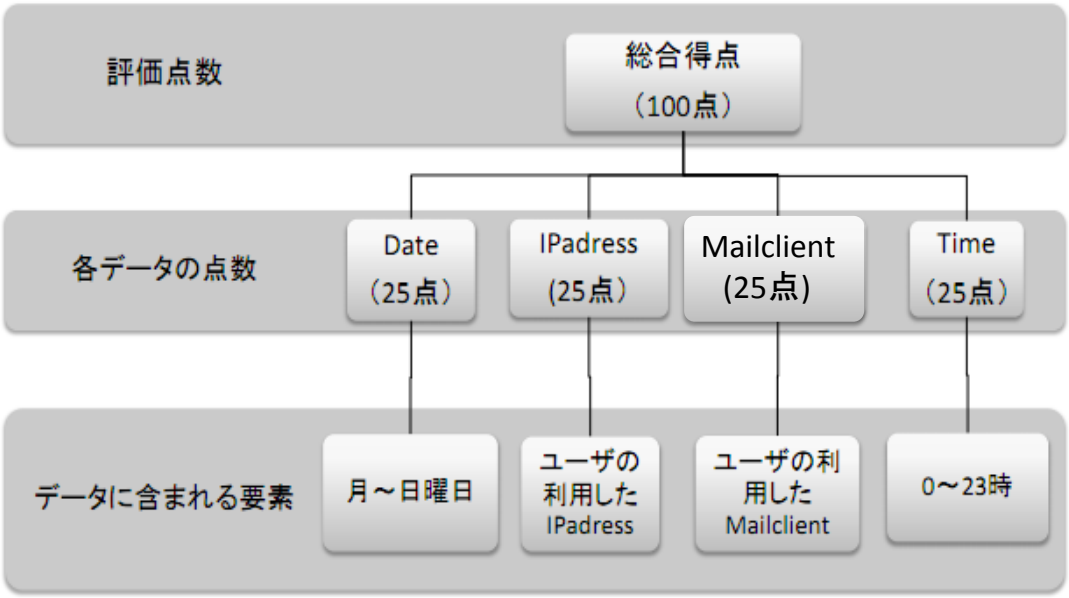
評価点数
総合得点
（100点）
各データの点数
Date
（25点）
IPadress
(25点）
Mailclient
(25点)
Time
（25点）
データに含まれる要素
月～日曜日
ユーザの
利用した
IPadress
ユーザの利
用した
Mailclient
0～23時

図 4.4: 点数計算表

第5章　実装

本研究の提案する手法の実装環境を記述する．以下に，順を追って詳しく説明する．

5.1　ヘッダ情報の収集と分析

電子メールのデータを取得する．ユーザが過去に取得した電子メールの情報を，ユーザのローカル環境から取得する．Mozilla Thunderbird[15] の場合，過去に取得した電子メールのデータは受信サーバに蓄積されているか，ユーザのローカル環境（特定の場所）に保存されている．それを取得し，取得した情報を解析スクリプトにかけて，アドレス毎にヘッダ情報を抽出する．これによって過去のメールからヘッダ情報のみを取得する．

次に，取得したヘッダ情報の中から，本研究で扱う情報のみを取得する．扱う情報に関しては，4.3 章で述べた通りである．更に，取得したデータをそれぞれの項目別に分け，ローカル環境に作っておいたデータベースにそれぞれ蓄積する．

5.2　データのプロファイリング

蓄積されたデータを元に，アドレス毎のプロファイリングを行う．以下にプロファイリングした結果をグラフに表す．

・ユーザ１には，主に IP アドレス (図 5.2)，クライアント (図 5.3) に強い特徴が見られた．また Date(図 5.4) にも若干の偏りが見られる．

・ユーザ２は，ユーザ１と対照的に使用するメールクライアント (図 5.7) が３種類でそれぞれ使用頻度が平均的である．また Date(図 5.8) と送信時間 (図 5.9) に偏りが見られる．

このように，同一のアドレスでも大きな特徴を持つものと，あまり特徴を持たない要素が存在する．またあるアドレスにとっては識別に非常に有効な情報でも，他のアドレスではあまり特徴付けに影響を及ぼさないものもある．これらを組み合わせる事で，各アドレス毎に有効なプロファイリングを施す事が可能になると考える．

5.3　送信者の信用度評価

プロファイリングの済んだアドレスから新規にメールが送られてきた場合，まずヘッダを取得する．次に取得したヘッダ情報から評価に必要な部分を抜き出してプロファイリングデータと比較する．比較した値を元に，ユーザの信用度を点数で表す．

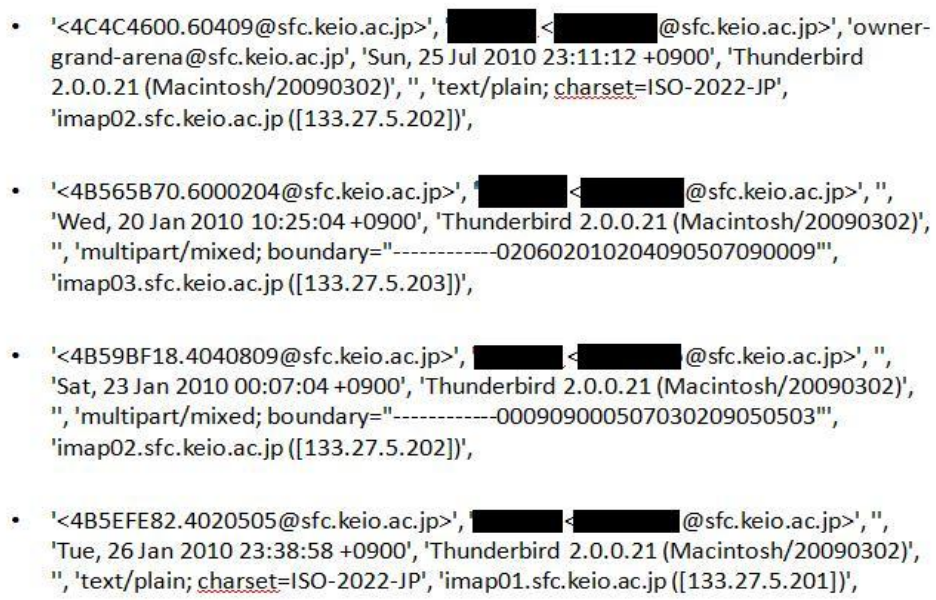

図 5.1: 抽出された情報

5.3.1 信頼度の点数表記

取得した情報をまとめた上で，4.4.2 章で提示した信用度の配点方法を基準に，事前調査として信用度の表記を行った．事前調査に利用したデータは，以下の通りである．

・過去自分が受信したメールデータ.

・プロファイリングに利用した過去データは 2010 年 1 月 1 日から 2010 年 12 月 31 日までの 1 年間.

・採点に利用したメールは 2011 年 1 月 1 日以降に受信したデータ.

・採点対象は SFC 生 5 名，社会人 1 名，企業 1 社.

以下に，結果の図 5.10 を記す.

このように点数表機にすることで，メールヘッダ情報に詳しくないユーザでもどれだけ信用がおけるかを理解することが可能となる.

5.3.2 事前調査の結果

5.2 章でグラフによる可視化を行った場合，様々な特徴がみられたが，今回事前調査として信頼度を数値によって表す事でも特徴を得る事が出来た.

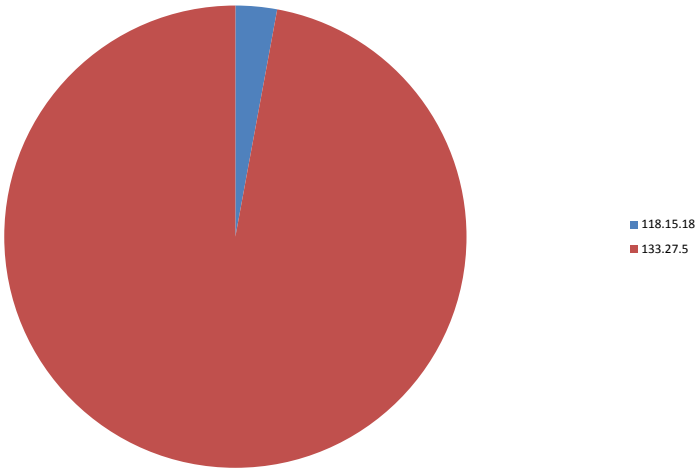

図 5.2: ユーザ 1 のプロファイリング結果：IP アドレス

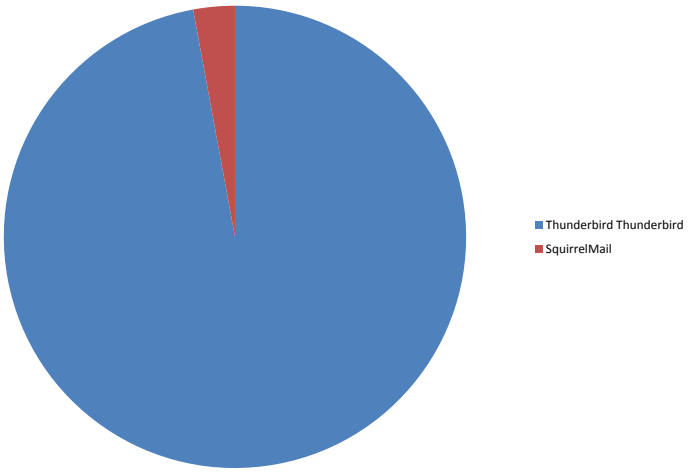

図 5.3: ユーザ 1 のプロファイリング結果：クライアント

収縮する平均点数の違い

点数は収縮する傾向にあるが，収縮する平均値に違いが見られた．SFC 生 1，2，3 と企業からのメールが 50 点台に収縮するのに対して，ユーザ 4，5 と社会人のデータは 30

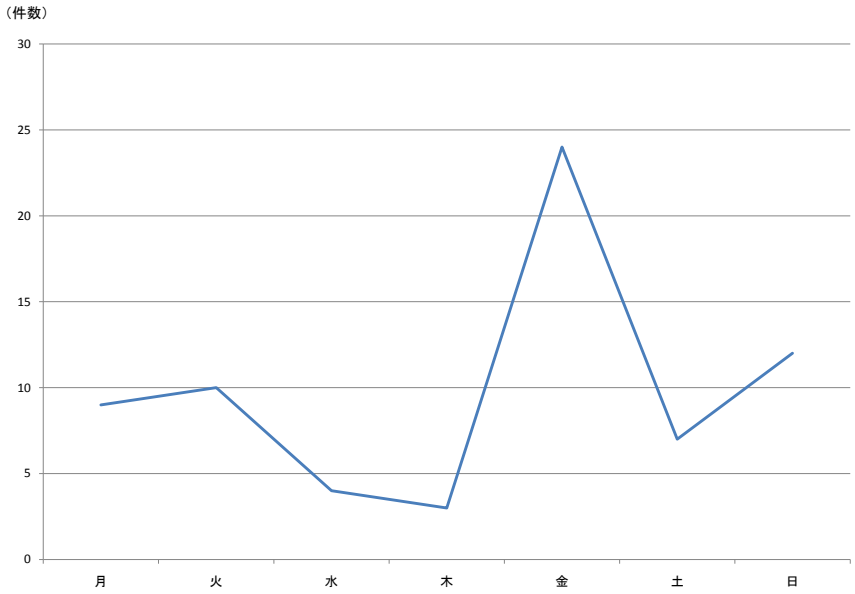

図 5.4: ユーザ１のプロファイリング結果：曜日

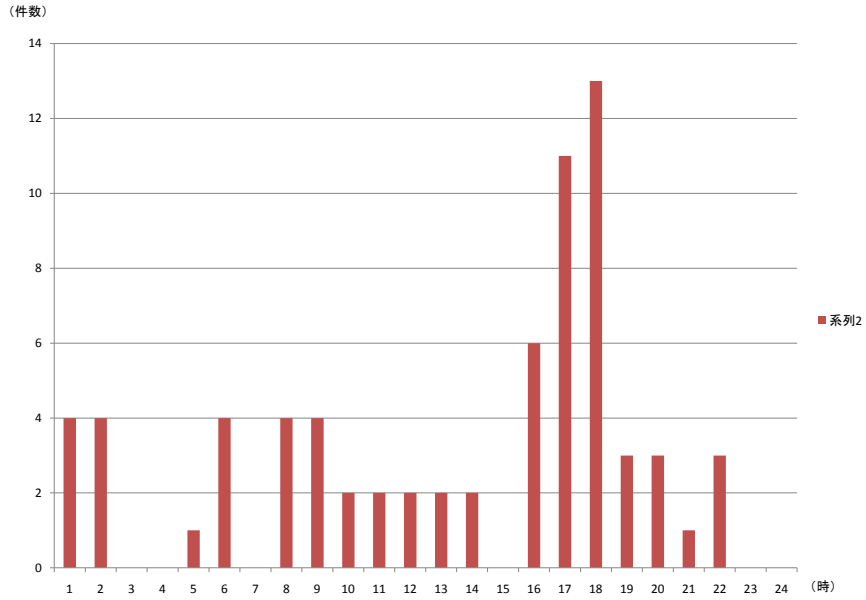

図 5.5: ユーザ１のプロファイリング結果：時間

点台に収縮した．調べてみたところ，違いとしてクライアントの利用頻度に特徴が見られた．平均点数の高いユーザは利用頻度の高いクライアントがほぼ１つに絞られるのに対し

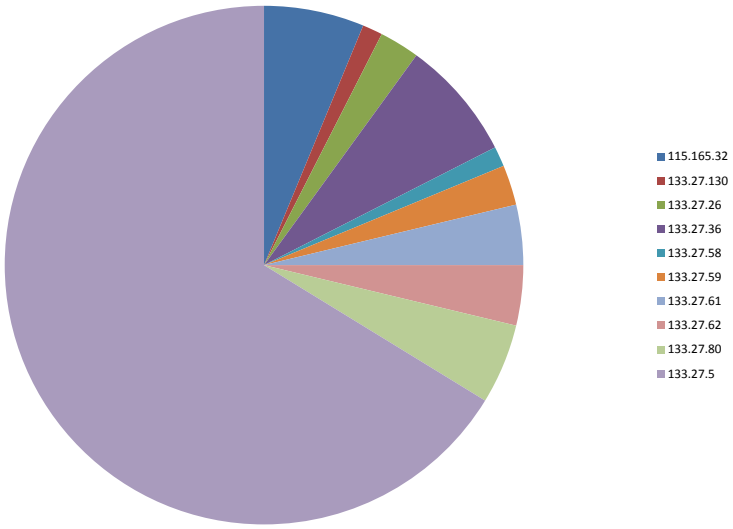

図 5.6: ユーザ 2 のプロファイリング結果：IP アドレス

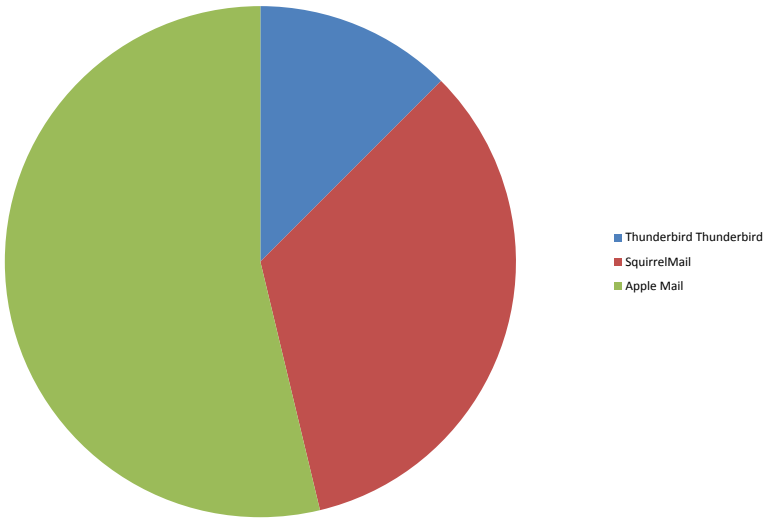

図 5.7: ユーザ 2 のプロファイリング結果：クライアント

て，平均点数の低いユーザは平均的に 2 つ以上のクライアントを使い分けている事が判明した．

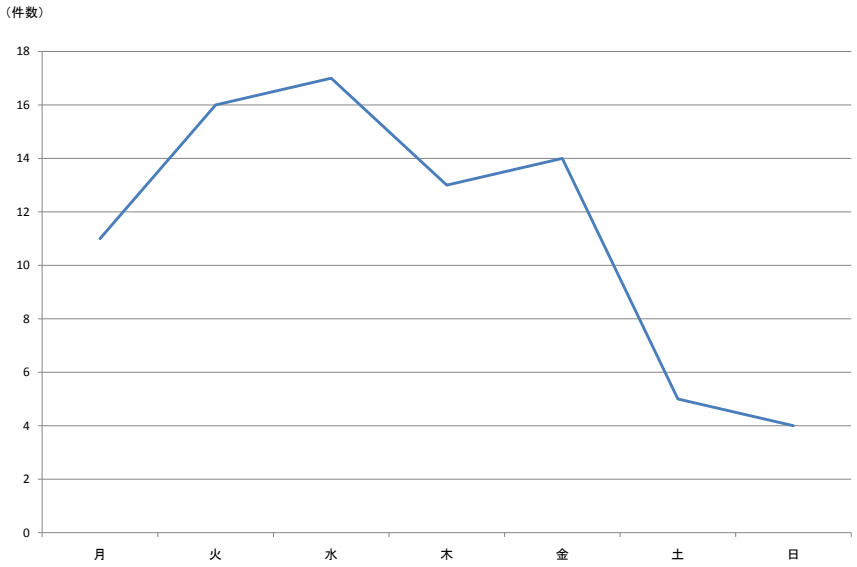

図 5.8: ユーザ 2 のプロファイリング結果：曜日

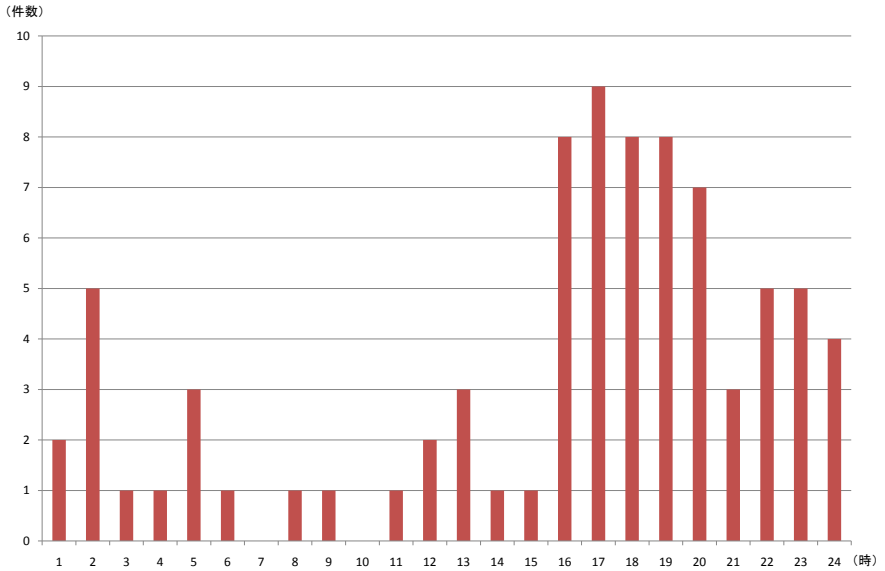

図 5.9: ユーザ 2 のプロファイリング結果：時間

振れ幅の広さ

SFC 生 4 には点数の振れ幅が広いという特徴が見られた．調べてみたところ，送信元 IP アドレスにばらつきが見られた．

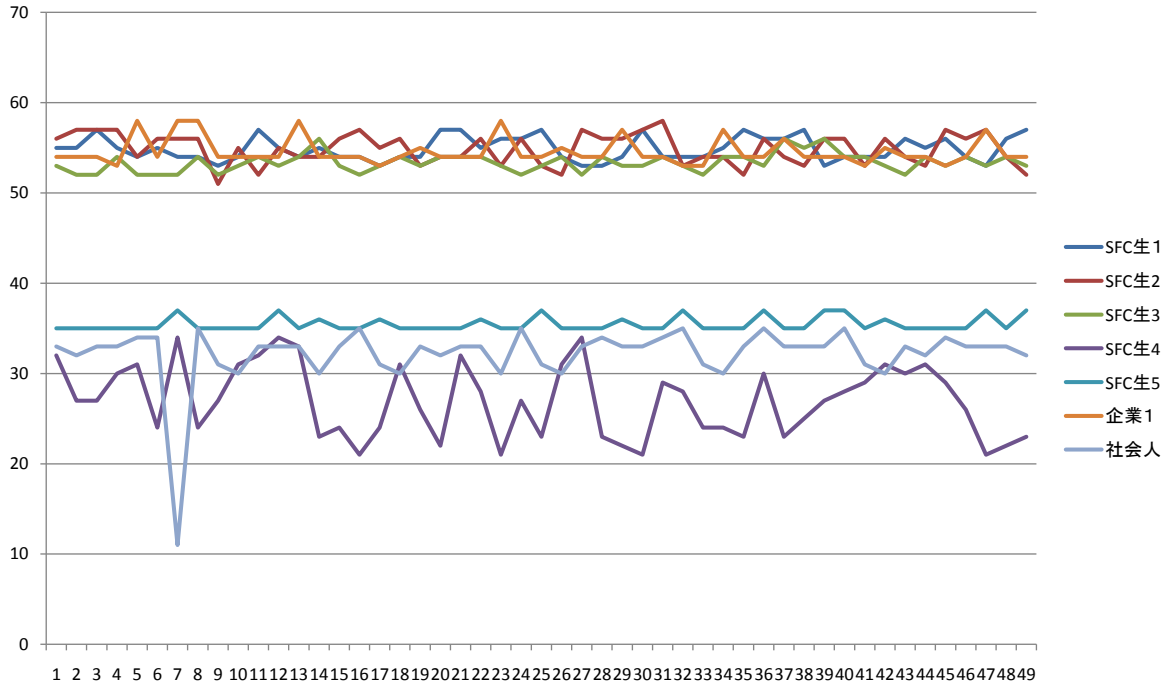

図 5.10: 採点結果

点数が大幅に低いメール

図 5.10 の中で，社会人からのメールの 1 つの得点が平均点数よりも大幅に下回ったため，そのメールデータを実際に調べてみたところ，送信環境が出張先だったために送信環境が大幅に変わっていた事が判明した．

これらの特徴をふまえた上で，第??章にて評価を行った．

第6章　評価

6.1　評価の概要

本研究においての評価は，5 章で実施した判別手法がどれだけ有効であるかによって評価する．

・評価内容

電子メールのヘッダ情報の中から 4.3 章で説明した情報を抜き出し，送信者毎に各情報を利用したプロファイリングを行う．これによって得られたデータを元に，新しく送られてきた電子メールの送信者がプロファイリングを行った送信者本人かどうかを確かめる．本研究の判別手法によってプロファイリングデータから外れた情報持つ電子メールが検知されたか否かを評価すると同時に，外れたデータを持つ電子メールが検知された場合その電子メールが本当になりすましなのかどうかを調べ，検知精度がどれだけのものかを評価する．

・評価データの取得先

筆者及び同研究室内の「Thunderbird」利用者複数名に届いたそれぞれの電子メールを利用．

・プロファイリングに利用する電子メール

それぞれのユーザが過去に受信した全ての電子メール

・評価対象となるメールの取得期間

ユーザ信頼度の点数評価に利用するデータ：2011 年 12 月 31 日以降に受信したメールデータ

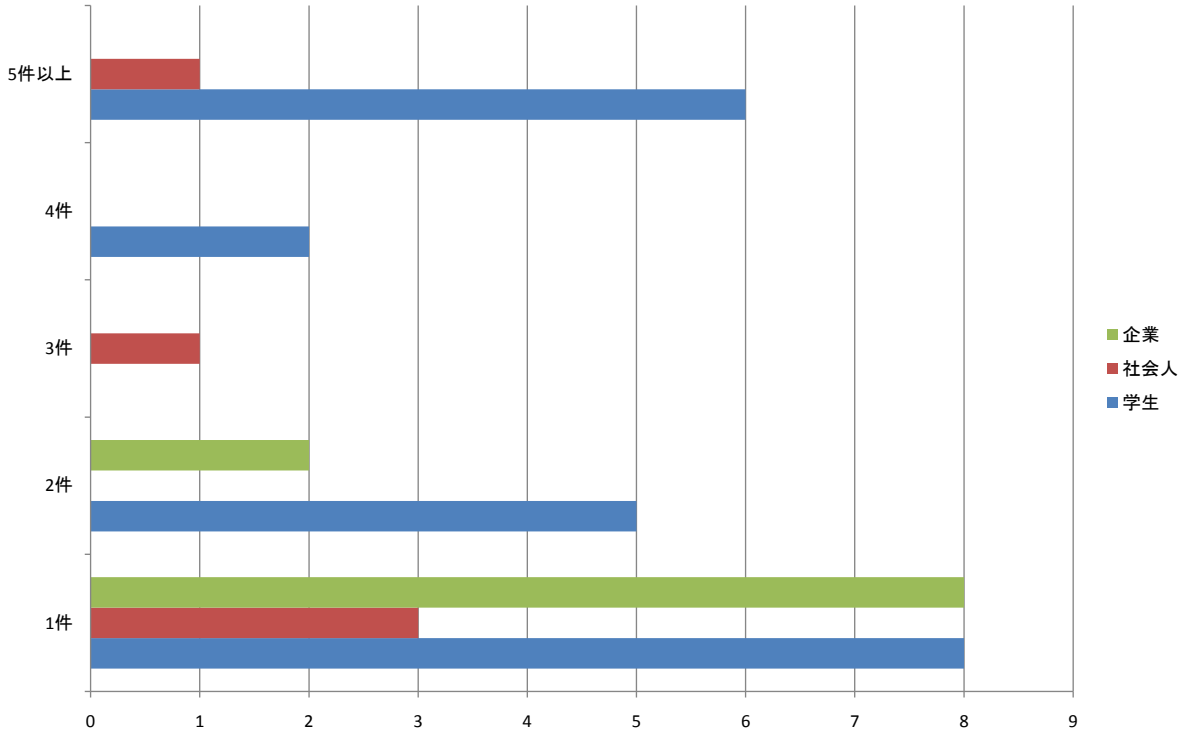

図 6.1: IP アドレスの調査結果

6.2 評価の結果

評価にはそれぞれ点数評価をしたメールアドレスを元に，学部生，企業，社会人の 3 種類に分類した．以下で本研究で利用したヘッダ情報（IP アドレス，メールクライアント，時間，曜日）をそれぞれ上記の 3 種類毎に見て，特徴を述べた上で結果を考察する．

6.2.1 IP アドレス

IP アドレスは、4.3 章にて説明したように，第 3 オクテットまでを判定に利用する．これは IP アドレス (v4) は，一般的に固定 IP アドレスを利用していない場合に第 4 オクテットが分散しやすいこと，一般ユーザは固定 IP アドレスを基本的には所持していないことを考慮したためである．以下に，点数評価とサンプリング結果の図 6.2.1 を記載する．

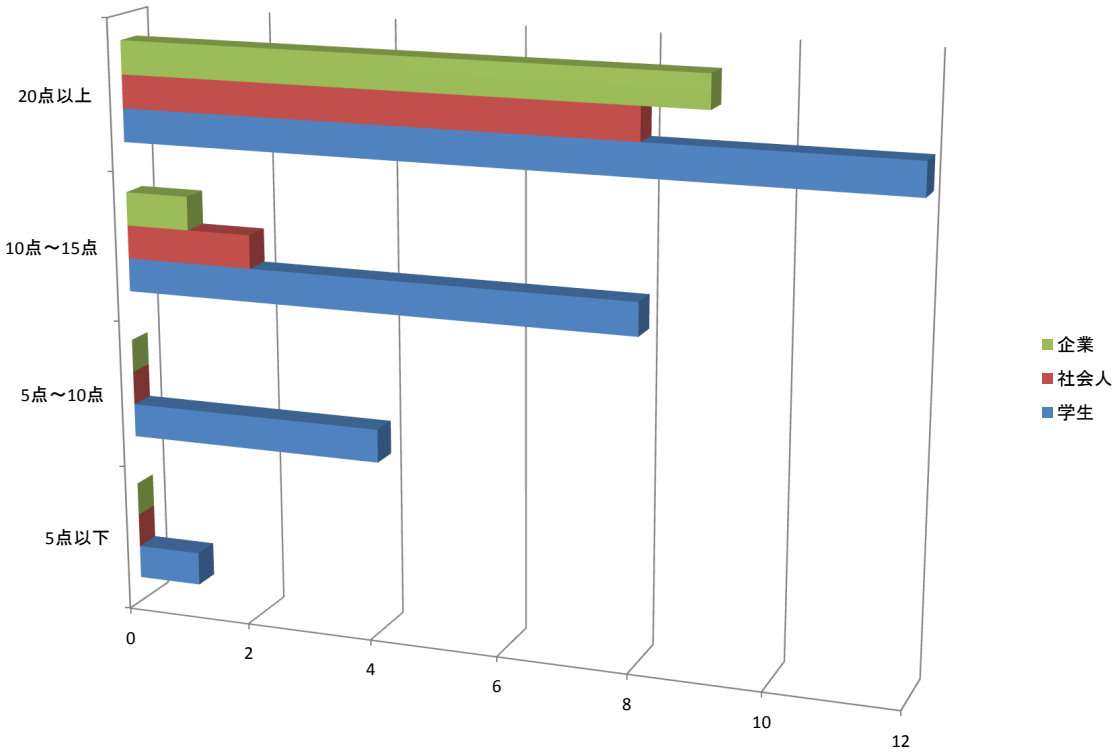

図 6.2: IP アドレスの得点

図からわかるように，企業と学生を比べた場合 IP アドレスの変化に差が生じていることがわかる. 学生の場合，家，学校，その他箇所での無線の使用が予想されるため，複数の IP アドレスがデータベースに蓄積される．一方で企業の場合，電子メールを送る場合会社の決まったコンピュータとなるため，決まったネットワークから有線を利用して送られる．その為，第 3 オクテットまでの IP アドレスに変化がほとんど生じないことが予想される．図 6.2 から，IP アドレスの得点の移り変わりにも特徴が見られた．複数の IP アドレスが同等の割合で過去のデータに存在する場合，本研究の配点基準だと点数がばらけるために得点が低くなる．また点数の推移にも特徴が見られた．これらのことから，なりすまし検知に非常に有効な性質を持つ事がわかる．

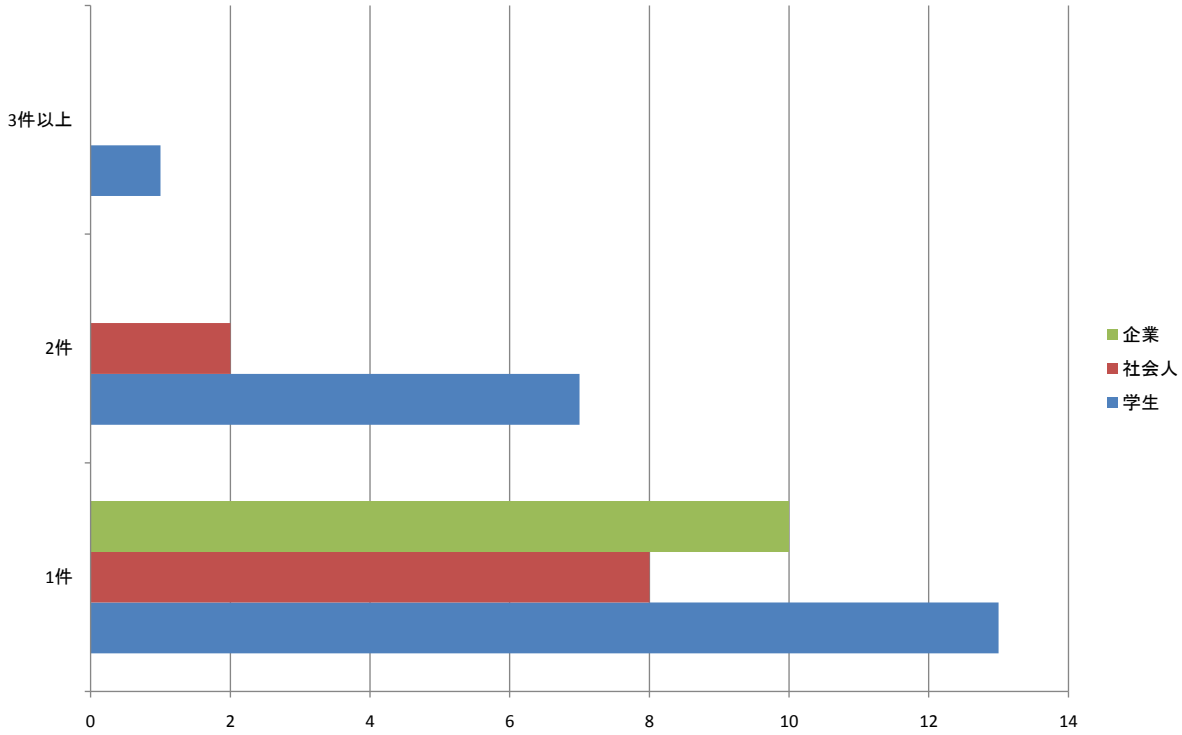

図 6.3: 電子メールクライアントの調査結果

6.2.2 電子メールクライアント

電子メールクライアントは，4.3 章にて述べたように，バージョンが違っても同一のクライアントとして扱う．電子メールクライアントは，１年の間に何度もバージョンアップが行われるケースが多いため，過去データの取得数を増やすことに比例してバージョンの種類も増えてしまうためである．以下に，結果の図を載せる．

図 6.3 からわかるように，使用されるクライアントは固定される傾向が強い．調べてみたところ，学生の場合主に Thunderbird が利用されていて，他には Becky![16]，SquirreMail[17]，AppleMail[18] などが確認はされたが，使用頻度としては Thunderbird が群を抜いて高かった．企業の場合 ClickM@iler[19] 及び Cuenote[20] から送られてきた者がほとんどであった．これらは，電子メールの配信を行うサービスであり，企業が電子メールを大量に一斉送信する際に利用しているからだと考えられる．

また図 6.4 から，クライアントの得点はほぼ同一なことがわかる．これらのことから，

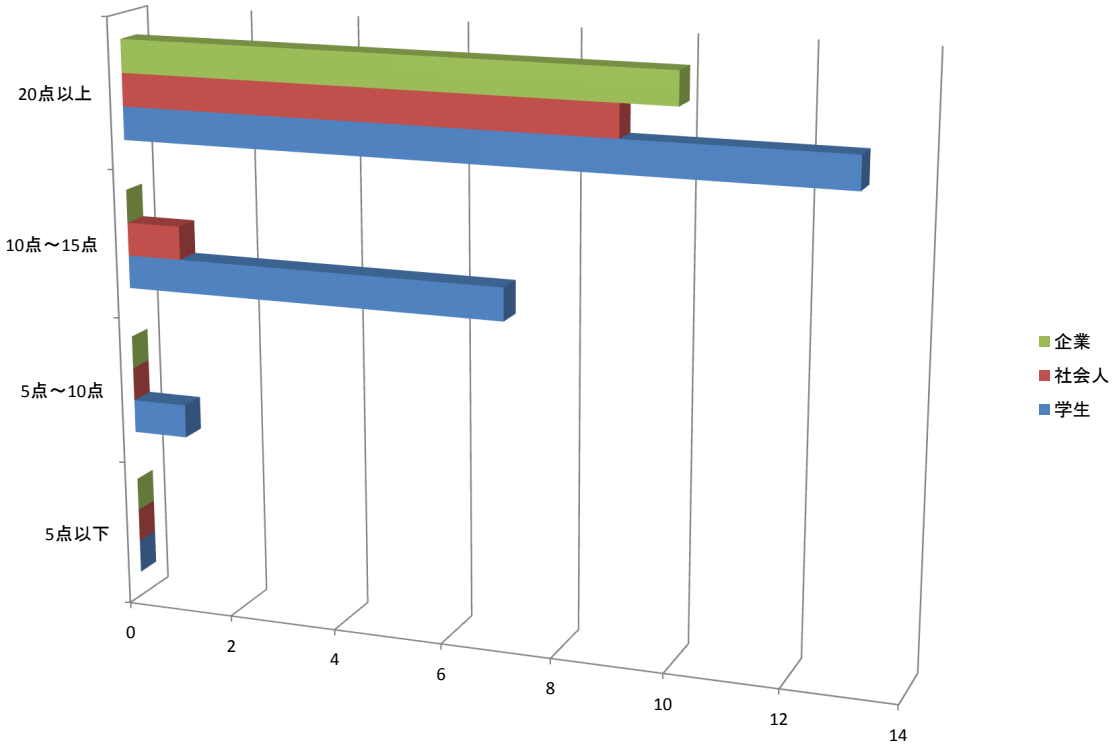

図 6.4: 電子メールクライアントの得点

電子メールクライアントはなりすましを判別する手法として非常に強い性質を持つ事がわかる．

6.2.3 日時

日時は，4.3 章で述べたように曜日．時間（○○時まで）を扱う曜日は 0 時 00 分から 23 時 59 分までを同一の曜日とし，時間は 00 分から 59 分までを同一の時間帯とする．以下に結果の図を載せる，

曜日に関しては，月曜日から日曜日のうち 7 日，6 日，5 日，4 日以下で分布を調べた．6.5 から，学生の場合，週間で 7 日以上の過去受信データがある場合が 95 %以上を占めたのに対して，社会人，企業のアドレスからの過去受信データは 7 日，6 日，5 日の分布が平均的になった．

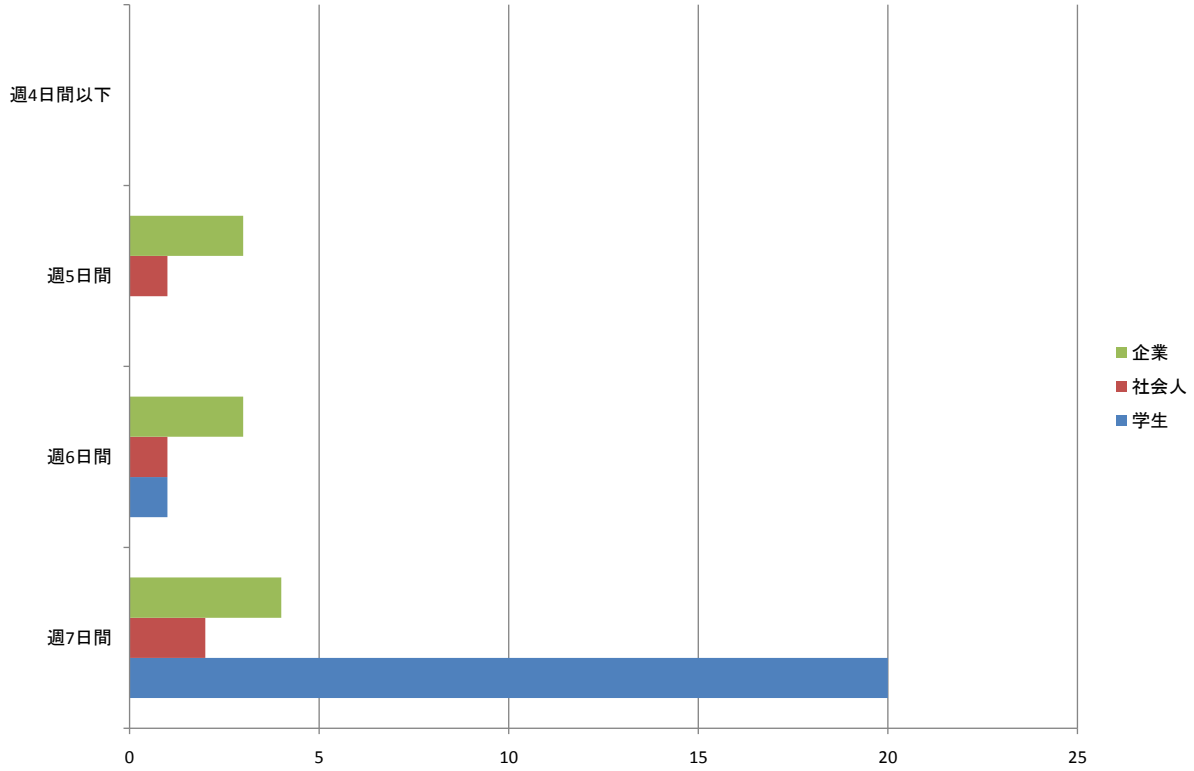

図 6.5: 曜日の調査結果

時間は 0 時から 23 時までのうち，15 時間以上，14 時間，13 時間，12 時間，11 時間，10 時間以下でデータの分布を調べた．6.7 からわかるように，相手が学生の場合，時間帯を問わずメールの送受信を行うため，15 時間以上が 95 %以上となる．対して，企業の場合，分布は平均的になった．ここから，企業や社会人が電子メールを送信する際，時間に左右されるケースが多い事が予想される．

6.2.4 得点評価

以下に，最終的な全メールの得点評価の図を載せる．

図 6.9 から，平均得点の分布にはそれぞれ特徴が見られた，企業の場合，決まった IP アドレスから，決まった電子メールクライアントを利用して送られてくる為に，平均の得点が高くなる．それに対して，IP アドレスや，クライアントに若干のばらつきが見られた学生や社会人の場合，平均の得点が企業に比べて低くなる事がわかる．更に調べた結果か

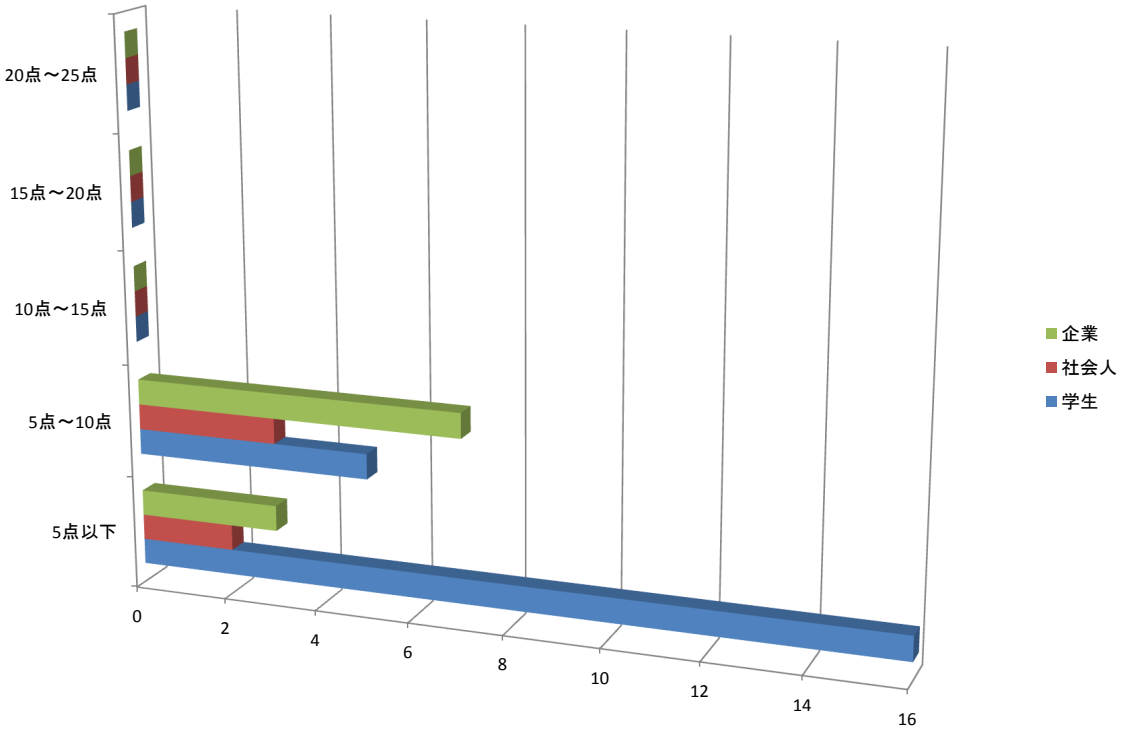

図 6.6: 曜日の得点

ら，アドレス毎に得点の幅があるものの，最大で平均値よりも+10 から-10 の範囲内におさまることがわかった．

また，今回実際に採点したメールを調べたが，なりすましメールと考えられるメールは存在しなかった．

6.2.5 異常検知の評価

実際に 4.3 章で述べた情報を利用して，本研究が提案する手法でアドレス毎のプロファイリングが可能である事は判明したが，プロファイリングデータから外れたものを異常として検知する事が可能かどうかは上記の実装では判断できなかった．本研究の提案する手法で異常検知が可能かについても評価するため，本研究が提案する手法でプロファイリングしたデータを利用して，本人からのメールと攻撃者と想定した別のアドレスからのメールを評価した．以下に詳しい条件を述べる．

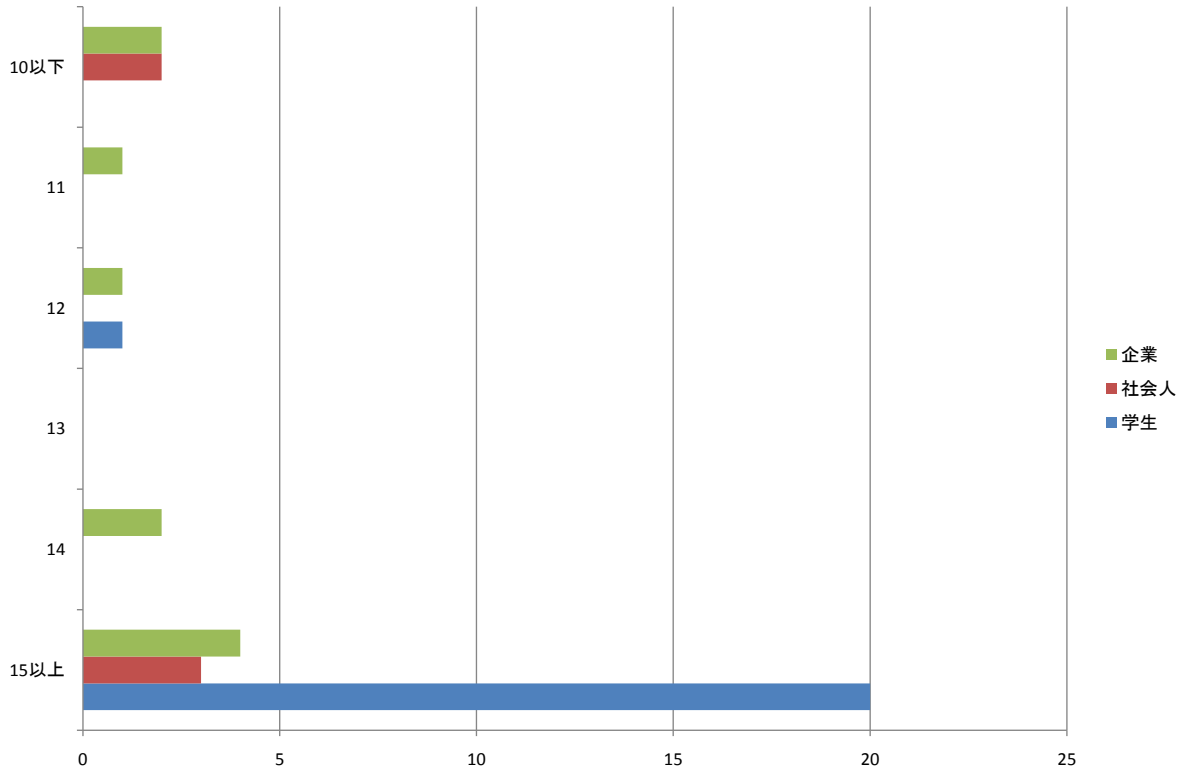

図 6.7: 時間の調査結果

・学生 A から受信したメールヘッダのデータを利用してプロファイリングを行う.
・学生 A，学生 B(別人)，企業，社会人それぞれから受信したメールヘッダ情報を解析.
・学生 A のプロファイリングデータを利用してそれぞれを点数評価する.
・評価するメールはそれぞれ新しく受信した物から 30 件.
以下に結果の図 6.10 を載せる

プロファイリングに使ったアドレス本人からのメールは，平均値の 57 点（小数点切り捨て）から大幅にブレる事がなかったが，学生 B，社会人，企業に関しては平均得点を大きく下回る点数を記録した．送信環境が近くなると予想される学生 B でも平均得点を 20 点以上も下回った.

このように，本研究が提案する手法で行ったプロファリングデータを利用する事で，本人からの電子メールと本人になりすました別人（攻撃者）からの電子メールの判別が可能である事がわかった．よって本研究の提案する手法は，なりすましメールを検知することに有効であると判断する事が出来る.

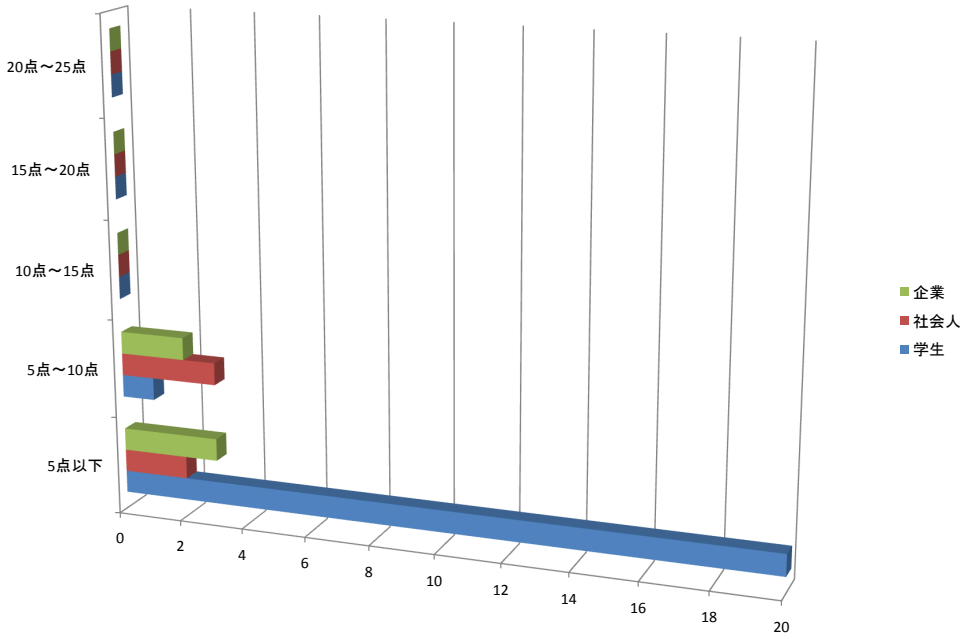

図 6.8: 時間の得点

6.3 考察

以上の調査から，メールヘッダに含まれる情報は，得点で表すことでアドレス毎にある値に収束することがわかった．そこから導きだされた値を利用して，それぞれアドレス毎に基準となる値を設定する事で，自分が意図した送信者からの電子メールとそれになりすました別人からの電子メールを有効的に検知する事が可能になると予想される．

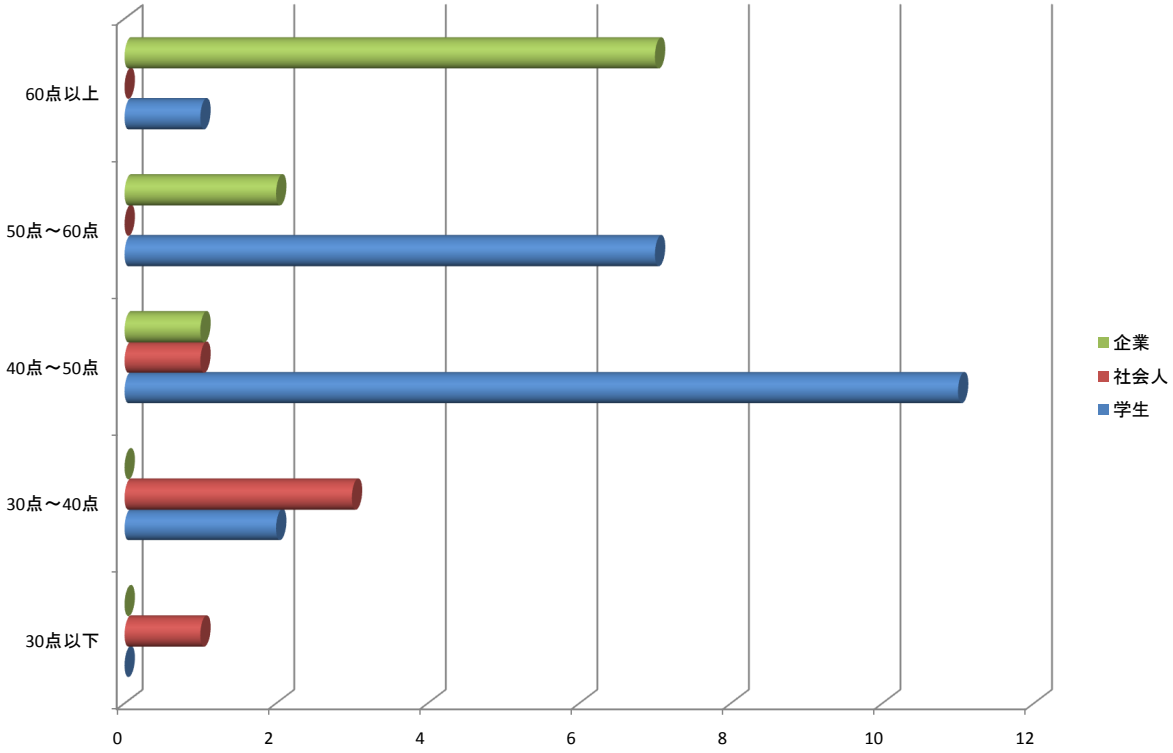

図 6.9: 得点評価

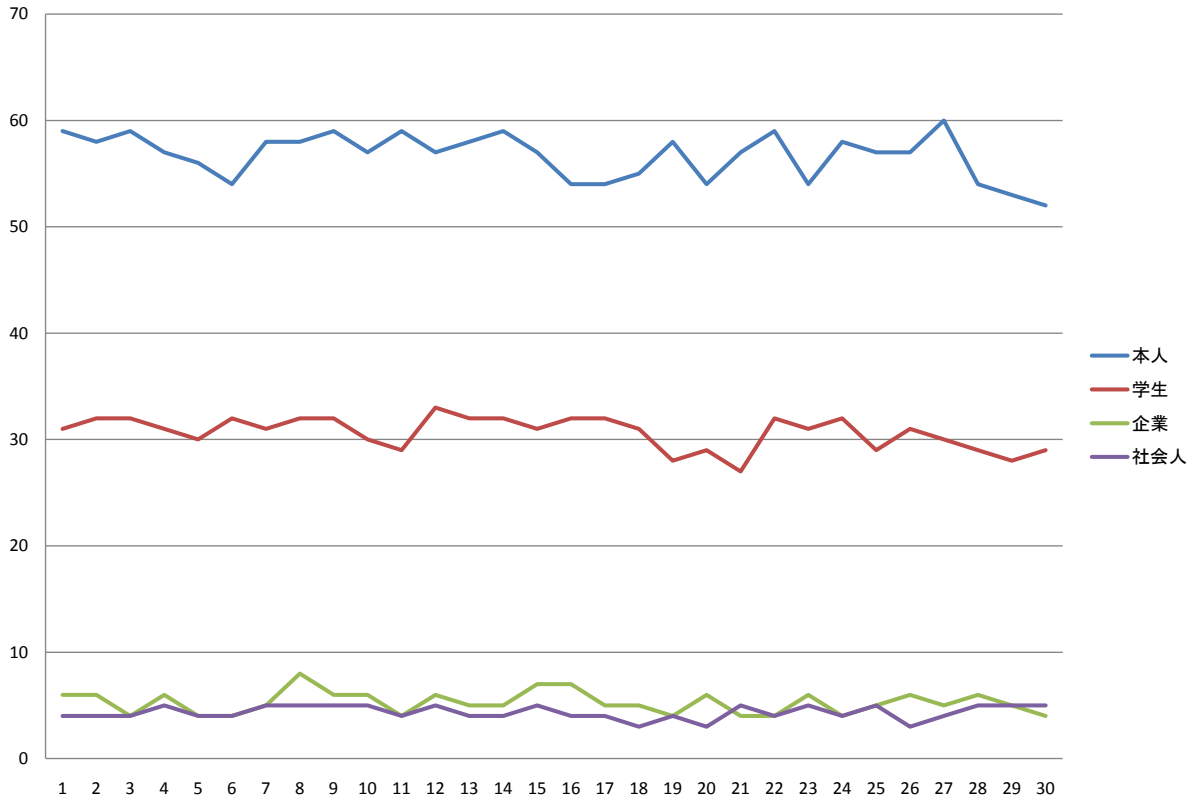
70
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
本人
学生
企業
社会人

図 6.10: 差異評価

第7章　結論

7.1　まとめ

本研究の目的は，電子メールを利用した標的型攻撃への有効な対策手段を作り上げるために，電子メールを利用した標的型攻撃の代表例であるなりすましメールへの有効な対策手段を作り上げる事を目標とした．そのために，電子メールの持つヘッダ情報を有効的に活用する解析用アルゴリズムを作成した．2章でも述べたように，標的型攻撃そのものは未だ一般ユーザに広く認知されているとは言えない．その一方で，いざ攻撃の被害にあった場合に被害が非常に大きく，更には攻撃による被害にすぐには気がつかないケースも確認されている．現状様々な会社から提供されているような既存のセキュリティ対策ソフトではなりすまし標的型攻撃への対策が十分ではなく，また攻撃の特性上対応が不可能な攻撃も多々存在する．同時に，既存のセキュリティ対策ソフトをコンピュータに導入しているからといって攻撃への対策意識が希薄になり，結果として攻撃を受けやすくなってしまう危険性もある．また，なりすましメールの場合，知人や公共機関になりすます為に，よりユーザの危険意識を希薄にしてしまう．

そこで、本研究はこのなりすましメールを，ユーザが受信した時点でなりすましかどうかを判定するアルゴリズムを作成した．これを利用する事で，ユーザが受信したメールを開く前に，攻撃かどうかを認知する事が可能となる．同時に，点数標記というわかりやすい形を利用する為に，電子メールの性質に関してあまり知識のないユーザにも電子メールの状態を理解する事が可能となる．過去受信した電子メールの情報を蓄積し，判定に使う事で，各アドレスの本来想定される送信環境から大きく外れたものを識別する事で，本来意図した送信者になりすました別人からのメールを判別する事が可能になる，更に，様々なアドレスに対応した特徴を数値として表す事で，容易に把握する事が可能となり，ユーザ毎に柔軟な対応が可能となる．

本研究では，攻撃に利用される電子メールの特性に注目し，電子メールに含まれる電子メールヘッダ情報を利用して判別を行う方法を提案した．まず電子メールヘッダに含まれる情報の中から，アドレス毎の送信環境を特徴づける有効な性質を持つ項目を調べ，判定に利用した．研究に先立って過去に自身が受信した電子メール情報を利用し，実際にアドレス毎に点数評価を行った結果，この方法で送信アドレス毎の特徴を掴む事が可能である事が判明した．本研究で利用した項目を順に述べる．まず送信者の送信元IPアドレス．これは送信者の年齢，職業によってそれぞれの特性が変わる性質を持つ事．また第4オクテットまでを判別してしまうと，同一の送信環境にいてもIPアドレスが変化してしまうため，第3オクテットまでに限定した．次に電子メールクライアント．これは送信者の職種に関わらず，ほとんどの場合が特定のクライアントを使用する事が判明した．また複数

のクライアントを利用する場合も，多くて２つから３つ，使用頻度も極端にばらける事はなかった．またクライアントのバージョンについてだが，バージョン毎に違うクライアントとカウントしてしまうと，電子メールクライアントは１年の間に何度もバージョンが変わるケースがあるので，本研究で判定の基準となる過去受信したメールデータが増える毎にバージョンも増えてしまう．その為，同一のクライアント名ならば，バージョンが違っても同一のクライアントと識別する事にした．最後に時間と曜日．これは主に企業などに強く傾向が見られた．特定の日にしか送ってこないケースや，ある時間を過ぎると一切メールが送られてこないなどの強い特徴が見られた．これらの情報を組み合わせる事で，アドレス毎に特徴を掴む事が可能となり，それを基準とする事で本来の送信環境から大きく外れたメールを識別する事が可能となる．更に，今後時間が進むにつれ，過去データが増えていく事でサンプルとなるデータが増え，より精度の高い検知を行う事が可能になると期待される．

本研究では，なりすましメールの特徴を利用した効果的な対策手法を提案，実装した．今後，なりすましメールのような電子メールを利用した標的型攻撃は更に増加されると予想される．そのため，本研究を元に，更に効果的な対策を作り上げ，企業，行政，少数の団体や個人への攻撃を未然に防げるよう研究を進めていく必要がある．

7.2 今後の課題と展望

本研究で提案した手法にはいくつかの問題点も確認された．まず大きく送信環境が外れた場合にそれを検知する事が目的だが，送信者が旅行や出張で海外にいる状態でメールを送信してきた場合，異常として検知してしまう可能性が非常に高い．また，判別に扱うデータは過去に受信したメールデータのため，新規に受信したメールや，受信頻度の低い電子メールに関しては，少し基準値から外れた場合でも異常として検知してしまう可能性もある．次に，点数評価において配点に偏りが存在している可能性がある．今回利用した要素の中で，曜日，時間の要素は月曜日から日曜日，0時から23時までとIPアドレス，クライアントに比べて構成要素が多い．そのため各構成要素の得点が軒並み低くなってしまうため，これに強い特徴を持つはずのアドレスの判定が，難しくなってしまうといった欠点がある．

そこで，１つのアドレスに対して，受信したユーザのデータだけでなく，同じアドレスから受信した事のある他のユーザの過去データからも情報を得る事で，本来受信したユーザが，そのアドレスからの受信頻度が低くても，より精度の高い識別が可能となってくる．更に，アドレス毎に識別するだけでなく，アドレス毎に学生，企業などのカテゴライズをし，カテゴリー毎に特徴を掴む事で，受信頻度が低いアドレスからのメールでも，そのアドレスがカテゴライズされているカテゴリー全体から特徴を掴む事で判定に利用する事が可能になる．

今後，本研究で提案した手法を利用して，更に高度な検知を行う場合に，より多くの過去データが必要となってくる．その為に，送受信サーバでのヘッダ情報の解析なども研究していく必要があると考える．

謝辞

本論分の作成にあたり，ご指導いただきました慶應義塾大学環境情報学部学長村井純博士，同学部教授徳田英幸博士，同学部教授中村修博士，同学部准教授楠本博之博士，同学部准教授高汐一紀博士，同学部准教授三次仁博士，同学部准教授上原啓介博士，同学部専任講師中澤仁博士，同学部准教授 Rodney D.VanMeter 博士，同学部教授武田圭史博士，同大学政策・メディア研究科特認講師斉藤賢爾博士，同学部政策メディア研究棟別研究講師佐藤雅明博士に感謝致します．

特に，武田圭史博士は，自分が本研究をするにあたり，セキュリティという分野に触れるきっかけを与えてくれた人であり，同時に研究を始めてから約 2 年間の間，様々ご指導を頂きました．更に研究に行き詰まったり，方向性を見失ったときに根気強く，何度も助言をして頂きました．本当に感謝しております．

また，本研究を薦めていく上で，様々な人に励まし，助言，叱咤を頂き，同時にお手伝いをして頂きました．村井研究室 OB 水谷正慶博士，田崎創博士，六田啓介氏，峯木厳氏，江村圭吾氏，黒宮祐介氏，梅田昂翔氏，Doan VietTung 氏，福岡英哲氏に感謝します．特に，梅田昂翔氏には，研究分野が近い事から，様々な，適切なアドバイスを頂きました．

慶應義塾大学政策・メディア研究科後期博士課程岡田耕司氏，空閑洋平氏，堀場勝広氏，工藤紀篤氏，久松剛氏，松園和久氏，松谷健史氏, 鈴木詩織氏，同修士課程永山翔太氏，波多野敏明氏，上原雄貴氏，山口修平氏，佐藤弘崇氏，三部剛義氏，米村茂氏，宮崎圭太氏，澤田暖氏，重松邦彦氏に感謝致します．特に，上原雄貴氏，重松邦彦氏の両氏とは所属する研究室が同じため，卒業論文を執筆するにあたり，様々な叱咤，助言を頂きました．両氏の協力なくして本研究を進める事は出来なかったと考えており，心から感謝致します．

慶應義塾大学徳田・村井合同研究室研究グループ ISC の研究生である相見眞男氏，碓井利宣氏， Vu Xuan Duong 氏，Pham Van Hung 氏，山本知典氏，吉原洋樹氏，有馬怜文氏，大矢崇央氏，鴻野弘明氏，小松真氏，中島明日香女史，三ツ木あかね女史，由井卓哉氏，吉原大道氏， Do Trung Kien 氏，露木航平氏，中安恒樹氏，深谷哲史氏，吉田裕氏，小澤麗女史，Nguyen Anh Tien 氏に感謝します．山本知典氏，碓井利宣氏，吉原洋樹氏，相見眞男氏は卒業論文を書くにあたって，苦楽を共にした仲間であり，様々な面での手助けをして頂きました．また吉原大道氏には，卒業論文の仮綴じ，添削などの細かな作業を手伝って頂きました．さらに，徳田・村井・楠本・中村・高汐・バンミーター・植原・三次・中澤・武田合同研究プロジェクトの皆様に感謝致します．

慶応義塾大学軟式野球サークル slayers に所属する方々に感謝します．大学生活の 4 年間を過ごす中で，彼らと過ごした時間は非常に有益な物でした．

最後に，大学および研究室で大きな怪我なく無事に 4 年間過ごせた事を，様々な面から

常鍄支え続けてくれた父吉昭，母幸子と家族に心から感謝致します.

参考文献

[1] JPCERT コーディネーションセンター. 「標的型攻撃について」. *JPCERT/CC ALL Rights Reserved.*, 6 2007.

[2] Inc. Nikkei Business Publications. 「三菱重工、国内 11 拠点でウィルス感染の事実を公表」, 9 2011.

[3] Ltd. Press Net Japan co. 「経産省にサイバー攻撃」, 10 2011.

[4] Inc. Nikkei Business Publications. 「sony 不正侵入による個人情報漏洩」, 5 2011.

[5] Akira Hisakawa All rights reserved. 「なりすましメールの実例」. `http://www.phishing.cc/ex/`.

[6] 独立行政法人情報処理推進機構. 「東日本大震災に乗じた標的型攻撃メールによるサイバー攻撃の分析・調査報告書」. *IPA,Japan. All rights reserved*, 12 2011.

[7] 独立行政法人情報処理推進機構. 「東日本大震災に乗じた標的型攻撃メールによるサイバー攻撃，実例」. `http://www.ipa.go.jp/about/press/pdf/110929_2press2.pdf`, 12 2011.

[8] PCERT コーディネーションセンター. 「標的型攻撃対策手法に関する調査報告書」. *JPCERT/CC ALL Rights Reserved.*, 8 2008.

[9] 末政延浩 and ITmedia. 「送信ドメイン認証の基礎知識」. `http://www.itmedia.co.jp/enterprise/articles/0603/24/news006.html`, 2006 3.

[10] 山本広志. Mta から見た spam メールの特徴. **山形大学紀要（工学）**, (32):27–36, 2 2010.

[11] 岡本圏 猪俣敦夫, ラーマン ミザヌール and 岡本栄司. フィッシングメール防御のためのメールフィルタリング手法の提案. 1 2005.

[12] 末政延浩 and ITmedi. 送信ドメイン認証の基礎. `http://www.itmedia.co.jp/enterprise/articles/0603/24/news006.html`, 3 2006.

[13] Incept Inc. 送信ドメイン認証の基礎. `http://e-words.jp/w/DomainKeys.html`, 4 2011.

[14] PCERT コーディネーションセンター. 「標的型攻撃対策手法に関する調査報告書」. *JPCERT/CC ALL Rights Reserved.*, page 9, 8 2008.

[15] Mozilla Japan. 「無料メールソフト thunderbird」. http://mozilla.jp/thunderbird, 2011 12.

[16] 株式会社リムアーツ. 「becky! internet mail」. http://www.rimarts.co.jp/becky-j.htm, 2 2009.

[17] The SquirrelMail Project Team. 「squirremail」. http://www.squirrelmail.org/, 9 2011.

[18] Weblio. 「apple mail」. http://www.weblio.jp/content/Apple+Mail, 9 2011.

[19] トランスコスモス株式会社. 「e メールマーケティングをトータルに支援するメールサービスー clickm@iler」. http://www.clickmailer.jp, 1 2011.

[20] YMIRINK.Inc. 「メール配信システムなら cuenote シリーズ」. http://cuenote.jp, 1 2011.