L-025

メール送信サーバ情報送信量特性

E-mail server information transmission amount characteristic

山口翔生 † 中平 勝子 † 北島 宗雄 †
Kakeru Yamaguchi Katsuko T. Nakahira Muneo Kitajima

1 はじめに

本稿では, メール受信サーバにおいてスパムと判定されたメールのメールヘッダを元に収集したメール送信サーバの情報を観測し, その収集データを処理することで, メール送信サーバの行動や変化を明らかにする. 観測により得られたサーバの情報送信量特性の変化により, スパム送信サーバの実態を把握し, スパム防御技術の発展に貢献することを本研究の目標とする. ここでいうスパム送信サーバの情報送信量特性とは, スパム送信の時間密度パターン (定期的な送信, 短期的な送信など) の遷移の様子であり, それはスパム送信時刻などの他に, 国やドメインなどのサーバの置かれた環境などによって推定する. 本稿では, 情報送信量特性を得るために有効なパラメータをスパム送信の時間密度パターンから探索する.

2 スパムメールに関する先行研究

この章では既存の研究と本研究の差異について説明する. 既存のスパム送信についての研究の多くは, メールフィルタリングを扱っている. メールフィルタリングはスパムへの対処の一手法であり, ベイジアンフィルタ, ブラックリスト方式, ホワイトリスト方式などのいくつかの種類がある.[2] 従来のスパムに関する研究では, これらの既存のメールフィルタをどの様に改良するかが中心だった. しかしメールフィルタリングは受動的なものであるため, スパムと健全なメールの振り分けは可能だが, 根本的な問題であるスパム送信自体を止めることはできない. 従来の研究より, スパムを送信するサーバには様々な特徴が現れることが報告されている.[3][4] 例えばドメイン名に特徴が現れる, スパム送信元の国に偏りが現れるなどである. 本研究では, スパム送信の根本的な問題点に言及するための一手法として, サーバの情報送信量特性を捉え, スパム送信サーバの特徴を見出そうと考えている.

本稿における大きな仮定は, スパム送信の時間密度パターンにも特徴が存在することである. スパム送信の時間密度パターンは, 能動的に任意のサーバを観測し続けることで得られる. そしてそこから前章で述べたように, 情報送信特性を得て, スパム送信サーバの行動の特徴を観察する.

3 情報送信量特性の定義

情報送信量特性とはサーバのメール送信パターンや, サーバの置かれた環境などのサーバの特徴をパラメータ化したものである. 本稿ではサーバの状態を最も的確に表すことができ, なおかつ状態遷移を予想することができるパラメータの探索が目標である. この章ではスパム送信の時間密度パターンにおける特性のパラメータ化について記す.

スパム送信の時間密度のパターンは, 観測した任意のサーバからのスパムより導き出される. u 番目のスパム送信サーバから送信された k 個目のスパムを S_{uk}^{send} (ただし, $0 \leq k \leq N_u$) とする, N_u は任意のスパム送信サーバ S_u^{send} からのスパム送信数である. 観測者が受信に利用するサーバが受信した u 番目のスパム送信サーバからの k 個目のスパムを S_{uk}^{rec} とする. また個々のスパム S_{uk}^{rec} からはサーバが受信した時刻 t サーバにおけるシステム時間で得らる. t は, 固定時間 δt の整数倍 ν の時刻に生起する連続量として観測されると考える. すなわち, 任意の測定開始時刻 t_0 から起算して i ($i \in \mathbb{N}$) 番目のイベント生起時刻 t_i は次の様に表される.

$$t_i = t_0 + \nu_i \times \delta t \tag{1}$$

ただし, ν_i は 0 以上の整数である. この時, イベント観測のウィンドウ $\Theta(t)$ は, $[t_0+\alpha\delta t-(n-\frac{1}{2})\delta t, t_0+\alpha\delta t+(n+\frac{1}{2})\delta t]$ の間で定義し,

$$\begin{cases} (t_0+\alpha\times\delta t-(n-\frac{1}{2})\delta t,\ \ t_0+\alpha\delta t+(n+\frac{1}{2})\delta t) \text{ では } 1 \\ (t_0,\ \ t_0+\alpha\times\delta t-(n-\frac{1}{2})\delta t) \text{ および} \\ (t_0+\alpha\delta t+(n+\frac{1}{2})\delta t,\ \ \infty) \text{ では } 0 \end{cases} \tag{2}$$

α, n は自然数

となる. これを, ヘビサイド関数 $\theta(x)$ を用いて次の様に定義する. ここで, $\theta(x)$ は次の性質を持つ関数を採用する.

$$\begin{cases} \theta(0) = \dfrac{1}{2}, \\ \theta(x) = 1, \ \ \text{ただし } x > 0, \\ \theta(x) = 0, \ \ \text{ただし } x < 0 \end{cases} \tag{3}$$

すると,

$$\Theta(t_0+\alpha\times\delta t) = \theta(t_0+(n-\frac{1}{2})\delta t)\times\left(1-\theta((n+\frac{1}{2})\delta t)\right) \tag{4}$$

となり, ここで $t_0+\alpha\times\delta t$ を t と記述し直すと

$$\Theta(t) = \theta\left(t-\frac{\delta t}{2}\right)\times\left(1-\theta\left(t+\frac{\delta t}{2}\right)\right) \tag{5}$$

となる. ある時刻 t に u 番目のスパム送信サーバから受信したスパムの数を $\boldsymbol{S}_u(t)$ と定義する.

$$\boldsymbol{S}_u(t) = (S_{u1}^{send},\ S_{u2}^{send}, \cdots,\ S_{uN-1}^{send},\ S_{uN}^{send}) \tag{6}$$

ただし

$$0 \leq t \leq T$$

† 長岡技術科学大学

表 1　情報送信量特性分類手法

手法	隣接分布	Evolution Diagram	SOND
注目個所	スパムの連続送信区間	スパム送信頻度	スパム送信密度
情報送信特性	時間密度パターン (送信幅特性, パターン数)	スパム送信状態 (健全, 悪性) の遷移	スパム送信期間
必要変数	$\boldsymbol{a}$,$\boldsymbol{e}$,L, K	$\boldsymbol{a}$, $\boldsymbol{i}$, L	$\boldsymbol{a}$,$\boldsymbol{e}$,L, K, T
表記法	$B(n\delta t) = \frac{1}{L-K}\sum_{i=1}^{L-K}\prod_{j=1}^{K} C(a_{i+j},\ e_j)$	$I(n\delta t) = \frac{1}{L}\int_1^L A'(\nu_i\delta t,\ t)di$	$H = -\int_1^L B(\nu_i\delta t)\log(B(\nu_i\delta t))di$

T は受信サーバがスパムを受信したことを観測する期間である，この $\boldsymbol{S}_u(t)$ を受信した時間に沿って並べることで, スパム送信の時間密度パターンを表す集合が得られる．しかしスパムを大量に送信するサーバでも 1 ヵ月に 100 通程度であるため, この集合は非常に疎な値になる．よって任意の区間幅 $n\delta t$ でのスパム送信の時間密度パターンを得る式を定義する．

$$\boldsymbol{\nu} = (\nu_1,\ \nu_2, \cdots,\ \nu_n) \tag{7}$$

$$A(\boldsymbol{\nu},\ t) = \frac{\boldsymbol{S}_u(t)}{|\boldsymbol{S}_u(t)|}\Theta(t) \tag{8}$$

$A(\boldsymbol{\nu},\ t)$ はスパムの時間密度パターンに注目するため, 値を規格化しておく．$\boldsymbol{S}_u(t) = 0$ の場合は, $A(\boldsymbol{\nu},\ t) = 0$ とする．この $A(\boldsymbol{\nu},\ t)$ は, 任意の区間 $t_0 \pm n\delta t$ のスパム状態を真偽値で返す関数である．時刻 t に沿って並べることで, 任意の区間幅 $n\delta t$ で観測したスパム送信の時間密度パターンの集合が得られる．実際の解析では離散的な値を利用するため, 式 2 を変更する．

$$A'(\nu_i,\ t) = \frac{\boldsymbol{S}_u(t_i)}{|\boldsymbol{S}_u(t_i)|}\Theta(t_i) \tag{9}$$

式 (9) より時間密度パターンの集合 $\boldsymbol{a}$ が得られる．

$$\boldsymbol{a} = (a_1,\ a_2, \cdots,\ a_i, \cdots,\ a_L) \tag{10}$$

ただし,

$$a_i = A'(\nu_i, t) \tag{11}$$

$$1 \leqq i \leqq L$$

$$L = \lfloor \frac{T}{n} \rfloor \tag{12}$$

である．本稿ではこの $\boldsymbol{a}$ を軸に研究を進めていく．

4　情報送信量特性のパラメータ

この章では前章で得た集合 $\boldsymbol{a}$ より隣接分布, Evolution Diagram(以下, ED 図), second-order neighborhood distribution(以下, SOND) を作成することで, 情報送信量特性を分類する．表 1 は本章で解説する 3 つの手法について簡単にまとめたものである．注目箇所は, スパム送信サーバのどの行動に注目して解析したかである．情報送信特性は, その手法で解析の結果得られるそのサーバの状態であり, これにより情報送信特性を分類する．必要変数は, その手法を使う際に必要な変数である．表記法は, その手法で使われる主な式である．

4.1　隣接分布

この章では隣接分布について解説する．表 1 で示したスパムの連続送信について調べるために, いくつかの段階を踏む．まずは集合 $\boldsymbol{a}$ から任意のスパム発生パターンの出現頻度を得る．任意のスパム発生パターンの集合を

$$\boldsymbol{e} = (e_1,\ e_2, \cdots,\ e_m, \cdots,\ e_K) \text{ただし, } e_i = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

とする．$\boldsymbol{a}$ と $\boldsymbol{e}$ のマッチングは, $C(a,\ b)$ を
a と b の排他的論理和の否定の結果, $\begin{pmatrix} 0, 1 \end{pmatrix}$ を返す関数と定義すれば次の式で求められる．

$$B(n\delta t) = \frac{1}{L-K}\sum_{i=1}^{L-K}\prod_{j=1}^{K} C(a_{i+j},\ e_j) \tag{13}$$

ただし, $n = \nu_i$ として

$$b_i = B(\nu_i\delta t) \tag{14}$$

とする．

$B(n\delta t)$ はを区間幅 $n\delta t$ で区切られた集合 $\boldsymbol{a}$ に含まれるパターン $\boldsymbol{e}$ の割合を示している．この処理は, 集合 $\boldsymbol{a}$ の一部の要素である a_j から a_{j+K} と, 任意のパターン $\boldsymbol{e}$ の要素 e_1 から e_K までを, それぞれ対応する要素どうしの排他的論理和の否定をとることで一致しているかを調べている．全ての要素が真の値を返したならば a_j から a_{j+K} までは任意のパターン $\boldsymbol{e}$ と一致していると判断する．この処理を全ての集合 $\boldsymbol{a}$ の要素に行うことで, 集合 $\boldsymbol{a}$ には任意のパターン $\boldsymbol{e}$ がどの程度の頻度で含まれているのかを示す．そしてこの $B(n\delta t)$ を区間幅 $n\delta t$ に沿って並べることで, パターン $\boldsymbol{e}$ の発生頻度を区間幅ごとに表した集合 $\boldsymbol{b}$ が得られる．

$$\boldsymbol{b} = (b_1,\ b_2, \cdots,\ b_m, \cdots,\ b_{L-K}) \tag{15}$$

今回は, 任意のパターンを $\boldsymbol{e} = (1, 1)$, K=2 とした．これはスパムが連続した区間で送信されていることを表す最小のスパム送信の時間密度パターンである．得られた集合 $\boldsymbol{b}$ は, ある区間幅 $n\delta t$ でスパム送信を観測したとき, スパム発生区間がどの程度連鎖しているかを表す．この $\boldsymbol{b}$ をプロットしたものを隣接分布図とする．本稿で例 (1(a)(c)) としてあげる隣接分布図は筆者の大学から観測されたスパム送信サーバを対象として作られている．この際, 観測期間は $T = 1$ ヵ月 $(2592000s)$ とした．この隣接分布図より得られるスパム送信サーバの状態について考察する．図 1(a)(c) の縦軸はパターン発生頻度 $\boldsymbol{b}$ を表し, 横軸は区間幅 $n\delta t$ を表す．縦軸横軸ともに対数表示である．図 1(b)(d) は図 1(a) と同じスパム送信サーバから得られた t_i をスペクトルとしてプロットしたものである．図 1(b)(d) の横軸は時間を表す．まずは隣接分布図の特性について考える．隣接分布図からは, 各 b_i の値が $b_i = 0$ か $b_i > 0$ のどちらかの状態だという

情報が得られる．隣接分布図の特性として，値が $b_i = 0$ の場合とは，スパムが発生した区間が 1 つも隣合わなかった場合である．この値が 0 になる場合は，2 つの状況が考えられる．1 つは区間幅 $n\delta t$ が小さいため，スパムが発生した区間が飛び飛びに発生し連続した区間として現れない場合．もう 1 つは，区間幅 $n\delta t$ が大きいため，1 つの区間にすべてのスパムの発生時刻が含まれた場合である．区間幅 $n\delta t$ は順にに大きくなっていくため，上記の状態は順に発生すると考えられる．そして，$b_i > 0$ の状態が連続している期間をスパム連鎖期間 S（以下，期間 S）とする．そしてこの期間 S の始端となる $n\delta t$ をスパム送信間隔特異点 P(以下，点 P) とし，末端となる $n\delta t$ をスパム送信期間特異点 Q(以下，点 Q) とする．まず期間 S について考察する．期間 S は対象のサーバにおけるスパム送信の時間密度パターンを表している．そして期間 S は図 1(c) の様に隣接分布図に複数現れる場合がある．この際の期間 S の個数は，対象のサーバが持つスパム送信の時間密度パターンの数だと考えられる．図 1(c) のサーバは図 1(d) から見えるように，短期的なスパム送信 (期間 $S1, S2$) が何度も現れる．そしてそれらが繰り返されることで，長期のスパム送信 ($S3$) として表される．このいくつかの時間密度パターンが図 1(c) に複数の期間 S として現れている．この期間 S の個数とそれぞれの点 P,Q はサーバに内在する時間密度パターンを調査するうえで重要な指針になると考えられる．ここから点 P と点 Q について考察する．まず期間 S の始端である点 P は，スパム送信が発生した区間が初めて隣接した区間であるから，そのサーバのスパム送信の時間密度パターンの最短の送信間隔を表していると考えられる．ただ，点 P は最短の送信間隔を表すだけでなく，その時間密度パターンのスパム送信間隔の特性を表すのではないかと考えられる．例えば図 1(a) の期間 S2 の点 P のとる $n\delta t$ の値と，図 1(b) のスパム送信幅はおよそ 30000 秒と一致している．図 1(b) の様に，ある程度規則的にスパムを送信するサーバの場合は，点 P は平均的なスパム送信幅に近い値を示すまた図 1(c)，図 1(d) の様に短期的な送信と，長期的な送信の様な 2 つ以上の時間密度パターンが現れた場合，それぞれの点 P は対応するパターンの平均的な送信間隔の値をとっている．この点 P の表す値の厳密な定義は今後の調査の課題である．次に点 Q について考える．図 1(a) の点 Q は，全てのスパムが 1 つの区間に収まる直前の $n\delta t$ である．これはスパム送信期間のおおよその幅を表していると考えられる．図 1(a) の点 Q の $n\delta t = 100000$ を示し，図 1(b) のスペクトルからも，およそ 100000 秒程度の区間でスパムを送信していることを示していることが確認できる．ただ点 Q は実際のスパム送信の時間密度パターンの送信期間より常に小さくなることには注意が必要である．ここまでの考察で隣接分布図から 3 つの特性を得ることができた．スパム連続期間 S の個数はスパム送信サーバに内在する，スパム送信の時間密度パターンの数を表してくれる．またこの期間 S から得られるスパム送信幅特異点 P は各々の時間密度パターンのスパム送信間隔の特性を表し，スパム送信期間特異点 Q はスパム送信期間などを表していることが示唆された．

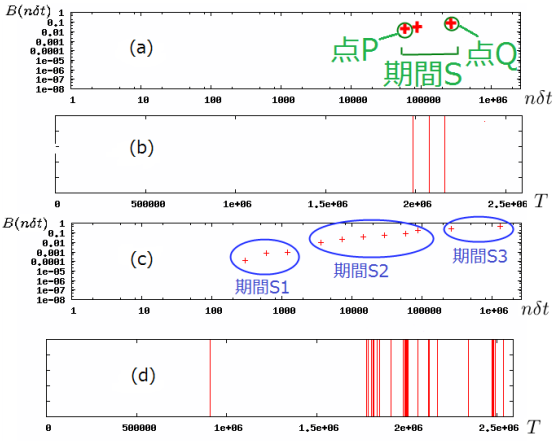

図 1　隣接分布図

4.2 Evolution Diagram

この章では Evolution Diagram(以下，ED 図) について考える．ED 図は集合 $\boldsymbol{a}$ より得られるスパム発生区間の頻度より求められる．まず $\boldsymbol{a}$ は区間幅 $n\delta t$ を変化させることで，分割区間数 $L(\boldsymbol{a}$ の要素数) とスパム発生区間数が変化する．そして $\boldsymbol{a}$ における，分割区間数 L とスパム発生区間数の比率をスパム発生頻度 $I(n\delta t)$ は，以下の式で得られる．

$$L = \frac{T}{n\delta t} \tag{16}$$

$$I(n\delta t) = \frac{1}{L}\int_1^L A'(\nu_i \delta t,\ t)di \tag{17}$$

ただし，$n = \nu_i$ として

$$i_i = I(\nu_i \delta t) \tag{18}$$

とする．

この $I(n\delta t)$ を区間幅 $n\delta t$ に沿って並べることで，スパム送信頻度を表す集合 $\boldsymbol{i}$ が得られる．

$$\boldsymbol{i} = (i_1,\ i_2, \cdots,\ i_i, \cdots,\ i_L) \tag{19}$$

$I(n\delta t)$ は累積密度関数であり，そこから得られる 集合 $\boldsymbol{i}$ の特徴は，区間幅 $n\delta t$ が増加するにつれて i_i も増加し，その比率が 1 に近似していく点である．区間幅 $n\delta t$ が増加することで，区間数 L とスパム発生区間は減少する．スパム発生区間は区間幅 $n\delta t$ が増加することで，同じ区間中に発生するスパムが増えるために，減少することはあれ増加することはない．区間幅 $n\delta t$ が広くなるにつれ，L とスパム発生区間数の個数は近づいていき，その比率である集合 $\boldsymbol{i}$ は増加し，ある値 i_iで 1 に に収束する (全ての区間にスパムが発生)．この集合 $\boldsymbol{i}$ をグラフとしてプロットしたものが ED 図である．この ED 図を実際に観測したサーバに適用しプロットしたところ，グラフの種類は大きく分けて 3 つのタイプに分類できる．それを表したのが図 2 である．縦軸は $I(\nu_i \delta t)$ であり，横軸は区間幅 $n\delta t$ である．横軸は対数表示となっている．図のパターン 1 が示すグラフは，比較的小さい区間幅 $n\delta t$ で 1 に収束しているタイプである．このタイプは区間分割数 L が大きいにも関わらず，多くの区間にスパムが発生している．観測の結果，長期的に継続してスパムを送信する

サーバが, このタイプであった. パターン 2 は区間幅 $n\delta t$ が十分大きくなった後, 急激に 1 に近づくタイプである. このタイプは分割区間数 L が十分小さくなってから初めて値が上昇することから, スパム発生区間は狭い区間にしか存在しないと考えられる. 観測の結果では, やはり観測期間中, ある一時期のみスパムを送信したサーバがこのタイプに分類されていた. パターン 3 はパターン 1,2 の中間のであり, 小さい区間幅 $n\delta t$ から緩やかに上昇し続けるタイプである. 上昇し続けつつも, パターン 1 の様にすぐに収束しないということは全域にスパム発生区間が散らばっているわけではないということを表し, パターン 2 と違い常に値がグラフにわかる形で上昇していることから, ある程度のスパム発生区間は存在していると考えられる. 観測した結果, パターン 2 は全区間の内, ある程度の大きさを持った一部の区間 (例えば前半, 後半の区間) のみにスパム発生区間が存在しているサーバが多かった. このタイプは健全なサーバがスパム送信サーバになった, もしくはその逆などのサーバの状態の遷移を観測するのに有効だと考えられる. この様に ED 図を観測することで, おおよそのスパム送信サーバの状態とその遷移を調べることができた. 今回はパターン 1,2,3 のタイプ分けを手動で行ったが, 将来的にはパターンマッチングなどを行う予定である. どの様なパターンマッチングを行えば, 適切にサーバを分類できるのかは今後の課題である.

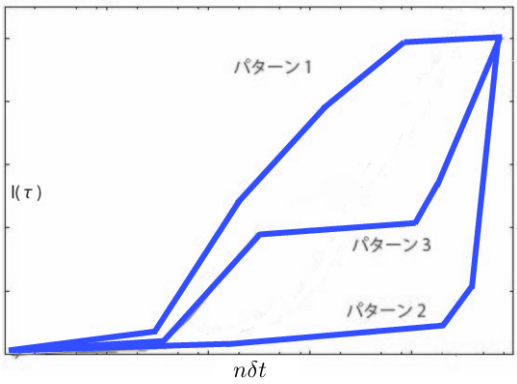

図 2　ED 図

4.3　Second-orderNeighborhoodDistribution

最後に Second-orderNeighborhoodDistribution(以下, SOND) について説明する. SOND は隣接分布図から得られ, 以下の様に定義される.

$$H = -\int_1^L B(\nu_i \delta t)\log(B(\nu_i \delta t))di \qquad (20)$$

この H はスパム送信が観測期間全体を通して, どの程度の密度で送信されているかを表している. 隣接分布図は小さい $n\delta t$ でスパム発生区間が連鎖するより, 値の大きい $n\delta t$ でスパム発生区間が連鎖するほうが全体の区間数が少ないため $B(n\delta t)$ の値が大きくなる. よって H が大きいということは, 値の大きい $n\delta t$ で連鎖しているということであり, つまり広い期間でスパムが送信され続けているということを示す. 図 3 はスパム発生区間数と H の関係について示したものである. 2012 年 6 月に筆者の学校にスパムを送信してきたサーバ 25000 個を観測し, そのスパム発生区間数と H の平均を計算した. 区間分割数 L=720($n\delta t$=1 時間) として, スパムの発生した区間数ごとにサーバを分類し, その H の平均をとったのが図 3 である. 図 3 の縦軸は H であり, 横軸はスパム発生区間数である. 図からわかるようにスパム発生区間数が増えるにつれて, H が増加している. 一般的にスパム発生区間が多いほど, スパムを長期的に送信していると考えられるため, この結果は H が短期的なスパム送信か, 長期的なスパム送信かを定量的に測る 1 つの基準になることを示唆している. またスパム発生区間数が少ない場合でも, 長期的なスパム送信である場合が多く確認できた. そのような場合でも H を見ることにより, 単にスパム区間数数で判断するより正確に, 時間密度パターンの判断できる. つまり集合 $\boldsymbol{a}$ や ED 図と組み合わせることで, より効果的に情報送信特性を分類できるのではないかと考える.

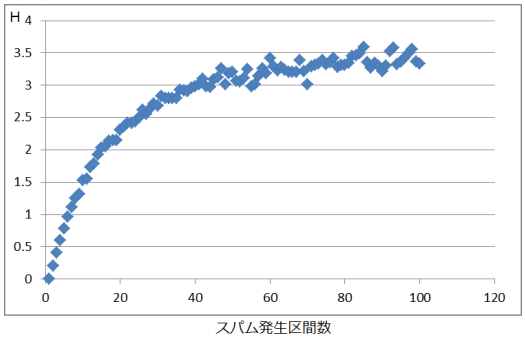

図 3　SOND

5　まとめ

本稿では, スパム送信サーバに着目し, その情報送信量特性を隣接分布図, Evolution Diagram として表すことで, スパム送信サーバの情報送信量特性を得るのに有効だと思われるパラメータが得られた. 今後の課題は, より有効なパラメータの探索, 各パラメータにより情報送信量特性がどの様に分類されるかの調査, そして情報送信量特性によるスパム送信サーバの行動の予測である.

参考文献

[1] Joshua Goodman, Gordon V. Cormack, David Heckerman , Spam and the Ongoing Battle for the Inbox, COMMUNICATIONS OF THE ACM, Vol. 50, No. 2, 25, 2007.

[2] Zhengchuan Xu, Qing Hu, Chenghong Zhang,Why computer talents become computer hackers, COMMUNICATIONS OF THE ACM, vol. 56, no. 4, 64, 2013.

[3] 竹下峰弘, 中平勝子, 三上喜貴, スパムメール発信源分析によるサーバ・ドメイン管理実態の推定, 全国大会講演論文集, 2011(1), 499-501, 2011.

[4] 澤谷雪子, メッセージ本文受信前でのスパムメール探知方式の制度向上に関する一検討, 信学技報 IEICE Technical Report, ICSS2009-57, 19, 2009.