情報処理学会論文誌 Vol. 51 No. 3 1083–1093 (Mar. 2010)

推薦論文

社会ネットワーク分析を用いたスパム対策：固有ベクトル中心性に基づくメールフィルタリング

白石善明[†1] 福田洋治[†2]
溝渕昭二[†3] 鈴木貴史[†1]

営利目的で一方的に送りつけられる電子メールは一般にスパムメールとして知られており，過去数年間，その影響は世界中に広がり続けている．本論文では，社会ネットワークの特性を表す指標の 1 つである固有ベクトル中心性に基づいたスパムメールフィルタリングの手法を提案する．提案手法は，受信したメールに含まれるメールアドレスの固有ベクトル中心性の得点を評価し，スパム以外のメールを区別するためのメールアドレスのホワイトリストを構成する．固有ベクトル中心性の得点とメールネットワークの形態との関係や，関連する手法との違いを考察し，著者のメールボックスのデータを用いた実験により提案手法の有効性を示す．

Antispam Method Using Social Network Analysis: Eigenvector Centrality-based Email Filtering

Yoshiaki Shiraishi,[†1] Youji Fukuta,[†2]
Shoji Mizobuchi[†3] and Takashi Suzuki[†1]

Unsolicited commercial email is generally known as spam and its negative effects continue to spread around the world in the past few years. In this paper, we propose a method for spam email filtering based on eigenvector centrality which is one of indexes expressing property of social network. The proposed method evaluates score of the centrality for each mail address included in received emails and constructs a whitelist of the reliable mail address to distinguish non-spam. This paper shows the effectiveness of the proposed method by some experiments using author's mailbox. Then, we give some considerations on the relation between score and form of mail network in order to show differences from relative methods.

1. はじめに

営利目的で一方的に送りつけられる電子メールは一般にスパムメールとして知られており，近年では，世界でやりとりされる電子メールの全体の 75%以上がスパムメールであるという報告がなされている[1]．メールサービスの利用者は，大量のスパムメールを受信することによって，スパムメールの除去に時間を費やし，サービスを快適に利用できなくなっている．またメールシステムの管理者および提供者は，大量のスパムメールによって，メールシステムやネットワークシステムの資源が浪費され，サービスを維持するための付加的なコストが発生し，経済的な損失を受けている．日本国内では，総務省による特定電子メールの送信の適正化等に関する法律や経済産業省による特定商取引に関する法律において，スパムメールに関する条文が整備されつつあるが，スパムメールの減少には至っていない．

技術的なスパムメール対策としては，メールの送信者が偽造メールアドレスを使えないようにする DKIM（Domain Keys Identified Mail）[2] や，ISP の加入者が ISP のメールサーバを介さずに外部ネットワークにメールを送信することを禁止する OP25B（Outbound Port 25 Blocking）[3]，メールサービスの利用者に送られるスパムメールを抑止，選別する Graylisting [4),5)]，Throttling [6]，Filtering などがあげられる．Filtering は，メールのヘッダ部にあるメールアドレスやサブジェクト，メールのボディ部にあるテキストなどの情報を用いて，スパムメールとその他のメールを区別するものであり，代表的な手法として，ルールベースの手法[7] やコンテンツベースの手法[8)–10)]，ホワイトリストによる手法[11)–13)] がある．

ホワイトリストによる手法は，メールの受信者が持つホワイトリストにあらかじめ登録されている人からのメールだけを受信するという手法であり，単純ではあるが効果的な手法として知られている．ホワイトリストに登録されている識別情報が搾取された場合，それを用いたフィルタリング回避に対処できないため，この手法では，一般に他のルールベースの

†1 名古屋工業大学
Nagoya Institute of Technology
†2 愛知教育大学
Aichi University of Education
†3 近畿大学
Kinki University
本論文の内容は 2007 年 7 月のマルチメディア，分散，協調とモバイル（DICOMO2007）シンポジウムにて報告され，CSEC 研究会主査により情報処理学会論文誌ジャーナルへの掲載が推薦された論文である．

手法やコンテンツベースの手法との併用が前提となっている．近年，メールのヘッダ部のFrom:や To:，Cc:のメールアドレスを抽出し，メールのやりとりを表現したネットワークを構成して，その社会ネットワークの特性を利用してメールアドレスのホワイトリストとブラックリストを自動的に作成する手法[14] が提案されている．この手法では，メールの社会ネットワークの特性が推移性指標であるクラスタリング係数によって評価されており，ヘッダ部が詐称されたスパムメールの存在を仮定した場合でも，誤判定が少ないという意味で精度の高いホワイトリストとブラックリストを作成できることが示されている．

本論文では，社会ネットワーク分析を用いてメールアドレスのホワイトリストを自動的に作成するスパムメールフィルタリングに注目し，中心性指標の 1 つである固有ベクトル中心性に基づいた手法を提案する．固有ベクトル中心性の得点を用いたメールアドレスの信頼度の評価方法を示し，そこで計算される得点とメールネットワークの関係や，既存手法との違いを考察し，著者のメールボックスのデータを使用して提案手法を評価し，その有効性を示す．

本論文の構成は以下のとおりである．2 章では，関連研究として，推移性指標であるクラスタリング係数に基づくメールフィルタリング手法（既存手法）について述べる．3 章では，中心性指標である固有ベクトル中心性に基づくメールフィルタリング手法を提案し，計算される固有ベクトル中心性の得点についての定理を与える．4 章では，著者が過去に受信したメールを用いて，固有ベクトル中心性の得点の分布を確認する実験や，提案手法と既存手法を単独で用いた場合と，提案手法とコンテンツベースの手法を併用した場合について，フィルタリングの性能を評価する実験を行う．5 章では，判定可能なメールタイプやホワイトリストとブラックリストの精度を低下させる行為に関して既存手法との違いを考察する．

2. 関 連 研 究

社会ネットワーク分析を用いたメールフィルタリングの関連研究として，推移性指標であるクラスタリング係数に基づくメールフィルタリング（Clustering Coefficient-based Method; CCM）[14] について述べる．以降，この手法を既存手法と呼ぶことにする．

まず，メールサービスの 1 人の利用者は 1 つのメールシステムだけに属しているものと仮定して，そのメールボックスのデータからメールのやりとりを表すネットワークを構成する．このネットワークは，個々のメールのヘッダ部から From:，To:，Cc:のメールアドレスを抽出して，**図 1** のような無向グラフで表現する．From:のメールアドレスから To:，Cc:のメールアドレスへのメール送信の様子を，ノード（メールアドレス）どうしを無向リンクで結ぶことで表している．メールボックスの所有者のメールアドレスが From:，To:，Cc:のいずれにも含まれないメールは，メールボックスの所有者のメールアドレスが Bcc:に含められて送信されたものと考える．

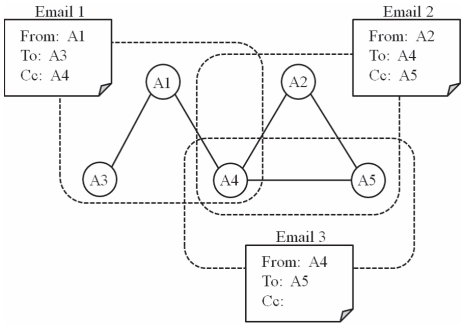

図 1　メールのやりとりを表現した無向グラフ
Fig. 1　Undirected graph expressing email exchanges.

次に，この無向グラフからメールボックスの所有者のメールアドレスのノードを除き，そのときできる個々の部分グラフ（コンポーネント）について，社会ネットワーク分析における推移性の指標であるクラスタリング係数を計算する．コンポーネントの中の 1 つのノード i に注目したとき，ノード i のクラスタリング係数 C_i は，次の式で定義される．

$$C_i = \frac{E_i}{{}_{k_i}C_2} = \frac{2E_i}{k_i(k_i-1)}. \tag{1}$$

ただし，ノード i の隣接ノードの集合を S_i としたとき，集合 S_i に含まれるノードの個数（ノード i の次数）を $k_i = |S_i|$，集合 S_i の中の 2 つのノード間に存在する無向リンクの数を E_i とする．クラスタリング係数は，ある 1 つのノードを定めたときそれに隣接する他のノードの間に関係がどの程度あるかを表す指標であり，隣接するノードの間に関係があるほど値が大きくなる．**図 2** は，各ノードのクラスタリング係数の関係の一例を示したものである．

コンポーネントのクラスタリング係数は，コンポーネントの中のノードのクラスタリング係数の平均をとったものであり，次の式で定義される．

$$C = \frac{1}{N_2}\sum_i C_i = \frac{1}{N_2}\sum_i \frac{2E_i}{k_i(k_i-1)}. \tag{2}$$

ただし，コンポーネントのクラスタリング係数を計算するときには，次数が 2 以上のノード

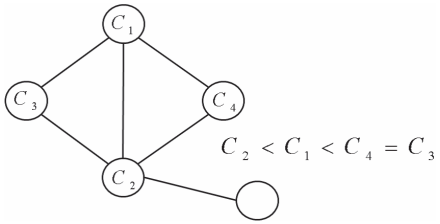

図 2　ノードのクラスタリング係数の関係
Fig. 2　Relation of clustering coefficient for each node.

を対象としており，N_2 はグラフの中の次数が 2 以上のノードの個数とする．

次に，個々のコンポーネントのクラスタリング係数から，スパムメールの送受信関係あるいはスパム以外のメールの送受信関係を判定して，コンポーネントに含まれるメールアドレスを前者の場合はホワイトリストへ，後者の場合はブラックリストへ登録する．

メールボックスの所有者の知人は互いに知り合いであり，その間でメールがやりとりされる可能性が高く，コンポーネントの推移性すなわちクラスタリング係数は大きくなる．また，スパムメールの送信者同士，被害者同士が互いに知り合いであることはほとんどなく，スパムメールに対応するコンポーネントのクラスタリング係数は小さくなる．このような考え方に基づいて，推移性の指標であるクラスタリング係数により，メールボックス所有者の知人のメールネットワークと，スパムの送信者と被害者のメールネットワークをコンポーネント単位で識別している[14]．

メールフィルタリングの一連の手続きは，次の 3 つのステップにまとめることができる．

(1) メールのフィルタリング ··· 受信したメールのヘッダ部の From:からメールアドレスを取り出して，ホワイトリスト，ブラックリストのメールアドレスと比較する．ホワイトリストのメールアドレスと一致する場合は通常のメール（ハム）と判定し，ブラックリストのメールアドレスと一致する場合はスパムと判定し，それ以外の場合は判定不能とする．

(2) メールのやりとりの無向グラフの更新 ··· 受信したメールのヘッダ部の From:，To:，Cc:のメールアドレスから，そのメールがどのメールアドレスの間でやりとりされたのかという関係を，メールの送受信関係の無向グラフに追加する．ただし，メールボックスの所有者のメールアドレスとの関係は含めない．

(3) ホワイトリストとブラックリストの更新 ··· 更新したメールの送受信関係の無向グラフにおいて，各コンポーネントのクラスタリング係数を求める．クラスタリング係数からコンポーネントをハムとスパム，その他に分類し，ハムコンポーネントに属するメールアドレスをホワイトリストへ，スパムコンポーネントに属するメールアドレスをブラックリストへ追加する．

評価対象のコンポーネントは，コンポーネントの中のノードの個数すなわちコンポーネントサイズが S_{min} 以上であり，かつ（コンポーネント中のノードの最大次数+1）/（コンポーネントのサイズ）が K_{frac} 以下のものとする．コンポーネントに属するメールアドレスは，2 つの閾値 C_{max}，C_{min} を用いて，クラスタリング係数が C_{max} より大きい場合はホワイトリストへ，C_{min} より小さい場合はブラックリストへ分類される．パラメータは，受信するメールの傾向に合わせて設定する必要があるが，$S_{min} = 10 \sim 20$，$K_{frac} = 0.6 \sim 0.8$，$C_{min} = 0.01$，$C_{max} = 0.1$ が適当とされている．

上記の手法に関連して，ホワイトリストやブラックリストの判定精度を高めること，ホワイトリストのメールアドレスをヘッダ部の From:に含めるようなヘッダ部を詐称したスパムメールに対処することなどを目的とした，改良手法[15),16)]や統合手法[17)]が提案されている．文献 15) の手法は，複数のメールボックスのデータを用いて社会ネットワークを共有することで，従来の単一のメールボックスのデータに基づいたものよりも判定の成功率を高めている．文献 16) の手法は，コンポーネントのクラスタリング係数を求める過程で，メールを受信したある時間の範囲でコンポーネントを構成することで，従来の時間の概念を含めていないものよりも判定の成功率を高めている．文献 17) の手法は，社会ネットワーク分析を用いたメールフィルタリングとベイジアンフィルタリングを連携させたものであり，フィルタリングの精度を高めるための統合手法（A）と，ホワイトリスト，ブラックリストを作成する際に必要となるメールの送受信関係のグラフの精度を高めるための統合手法（B）が示されている．

次章では，ホワイトリストによる手法として，中心性指標の 1 つである固有ベクトル中心性に基づくメールフィルタリングを提案する．なお，文献 14)–16) の手法は，メールのやりとりの関係を無向グラフで表し，クラスタリング係数に基づいてコンポーネント単位でメールアドレスの信頼度を評価し，ホワイトリストとブラックリストを作成するものである．一方で，提案手法は，メールのやりとりの関係を有向グラフで表して，固有ベクトル中心性に基づいて個々のメールアドレスの信頼度を評価し，ホワイトリストを作成する．以上のような違いはあるが，提案手法は既存の手法と同様に，文献 17) の手法のような他のフィルタリング手法との併用を前提としている．

3. 固有ベクトル中心性に基づくメールフィルタリングの提案

社会ネットワーク分析における中心性指標の 1 つである固有ベクトル中心性に基づくメールフィルタリング（Eigenvector Centrality-based Method; ECM）を提案し，固有ベクトル中心性の得点についての定理を与える．以降，ここで述べる手法を提案手法と呼ぶことにする．

まず，メールサービスの 1 人の利用者は 1 つのメールシステムだけに属しているものと仮定して，そのメールボックス中のメールのヘッダ部の情報からメールのやりとりを表すネットワークを構成する．このネットワークは，メールのヘッダ部から From:，To:，Cc: のメールアドレスを抽出して，**図 3** のような有向グラフで表現する．既存手法とは異なり，From:のメールアドレスから To:，Cc:のメールアドレスへのメールの送信の様子を，ノード（メールアドレス）どうしを有向リンクで結ぶことで表現する．メールボックスの所有者のメールアドレスが From:，To:，Cc:のいずれにも含まれないメールは，メールボックスの所有者のメールアドレスが Bcc:に含められて送信されたものと考える．

次に，有向グラフ上の個々のノードについて，社会ネットワーク分析の中心性指標である固有ベクトル中心性の得点を計算する．ノード i の固有ベクトル中心性の得点 x_i は，ノード i への有向リンクを持つノードの固有ベクトル中心性の得点の線形結合として，次の式で定義する．

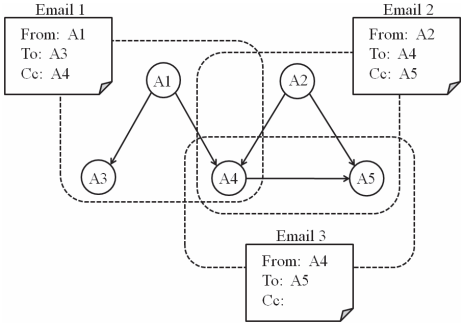

図 3　メールのやりとりを表現した有向グラフ
Fig. 3　Directed graph expressing email exchanges.

$$x_i = \frac{1}{\lambda}\sum_{j=1}^{M} a_{j,i}x_j. \tag{3}$$

ただし，有向グラフのノードの個数を M とし，ノード j からノード i へのリンクの状態を $a_{j,i}$ とする．x_i，$a_{j,i}$，λ は正の実数とする．ノード j からノード i への有向リンクが存在する場合は $a_{j,i} = (1-\epsilon)/l_j + \epsilon/M$，有向リンクが存在しない場合は $a_{j,i} = \epsilon/M$ とする．ノード j からの出力リンクの個数を l_j とし，ある 1 つのノードの得点をリンク先以外のノードに伝搬させる割合を ϵ，$0 < \epsilon \ll 1$ とする．固有ベクトル中心性の得点は，集団においてどの程度中心的な役割を果たしているかを表す指標であり，入力リンクの多いノードおよび入力リンクの多いノードからのリンクを持つノードほど値が大きくなる．**図 4** は，各ノードの固有ベクトル中心性の得点の関係の一例を示したものである．

得点 x_i，$i = 1, 2, \ldots, M$ を求める問題は，次の式を満足するような，行列 $\mathbf{A}^T$ の固有値，固有ベクトルを求める問題として扱うことができる[18]．

$$\mathbf{A}^T\mathbf{x} = \lambda\mathbf{x}. \tag{4}$$

ただし，j 行 i 列の要素を $a_{j,i}$ とする M 次正方行列を $\mathbf{A}$ とし，得点の列ベクトルを $\mathbf{x} = [x_i]_{i=1}^{M}$ とする．行列 $\mathbf{A}$ は要素がすべて正の行列であり，ペロン・フロベニウスの定理から，絶対値最大の固有値は正の実数であり，対応する固有ベクトルは要素がすべて正でノルムが 1（$\mathbf{x} > 0$，$\sum_{i=0}^{M} x_i = 1$）となる．また，行列 $\mathbf{A}$ は，$0 < a_{j,i} < 1$，$\sum_{i=1}^{M} a_{j,i} = 1$ を満たすことから，マルコフ連鎖の推移確率行列と見なすことができる．この行列の絶対値最大の固有値は 1 であることが知られている．したがって，図 4 の固有ベクトル中心性の得点の関係は，行列 $\mathbf{A}$ の最大固有値（$\max\lambda = 1$）に対する固有ベクトル（優固有ベクトル）において成立するものであり，個々のノードの得点はべき乗法などの近似法により容易に計算できる．

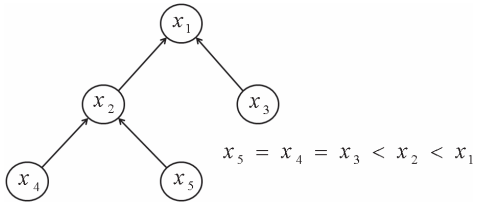

図 4　ノードの固有ベクトル中心性の得点の関係
Fig. 4　Relation of score assigned to eigenvector centrality for each node.

ノード j からノード i へのリンクが存在する場合を $a_{j,i}=1$，存在しない場合を $a_{j,i}=0$ のように単純に設定すると，個々のノードの得点が一意に定まらない場合や，意図したノードの得点の関係が得られない場合がある．固有ベクトル中心性の得点を計算する際は，一般に，先に示したような方法により，ノード間のリンクの状態を，$0<a_{j,i}<1$，$\sum_{i=1}^{M} a_{j,i}=1$ を満たすように設定し，これに対処する[18]．

本手法の有向グラフから行列 **A** を構成して固有ベクトル中心性の得点を求める過程は，Web ページのスコアリングの手法として知られる PageRank 法[19)–21)] と同じものになっている．PageRank 法は Web ページのリンク関係を固有ベクトル中心性の指標で評価するものであり，固有ベクトル中心性の得点を計算する過程は，本手法とは適用対象が異なるが本質的に変わらない．しかし，本手法では後述するように，信頼できるノードの得点が高くなるように有向グラフを補正しているところが異なる．そして，本手法では陽に与えられる有向グラフ中の得点の最小値をもとにホワイトリストを自動生成する．

ϵ は，個々のノードの得点差に関係するものであり，これを PageRank 法では dampening factor と呼び，$\epsilon=0.1\sim0.2$ の範囲で，固有ベクトル中心性の得点を求めている[20]．ϵ の値を小さくとると，ノード間に有向リンクが存在するとき得点が大きく伝播して，ノードのリンク状況の違いにより，個々のノードの得点差を計算機上で識別しやすくなる．ただし，ϵ の値を小さくとりすぎると，全体のノードの個数が増加した際に，$(1-\epsilon)/l_j$ と ϵ/M の差が大きくなり，計算機上の計算誤差が大きくなるため適切な調整が必要である．

多数の出力リンクを持つノードと，少数の出力リンクを持つノードにおいて，出力リンクの重みを一定にした場合，出力リンクを多く持つノードほど，得点の影響を他のノードに与えることになり，個々のノードの影響力が一定にならない．有向リンクが存在する場合を，$a_{j,i}=(1-\epsilon)/l_j+\epsilon/M$ とすることで，多数の出力リンクを持つノードと，少数の出力リンクを持つノードの，他のノードに与える得点の影響を公平にしている．

本手法では，メールボックスの所有者のメールアドレスは，信頼度が最も高いメールアドレスであることから，このノードに対して有向グラフ上のすべてのノードから有向リンクを張るという補正を行い，得点が最大値をとるようにする．**図 5** はこの補正の様子を示している．また，メールボックスの所有者の知人のものだと確定しているメールアドレスが存在する場合は，そのノードとメールボックスの所有者のメールアドレスのノードとの間に双方向の有向リンクを張るという補正も可能である．これらの補正は，信頼度の高さが明らかなメールアドレスが存在する場合に，そのノードから直接あるいは間接的に有向リンクを受けているノードに得点の影響を伝搬しやすくするための付加的な操作である．メールボックスの所有者やその知人のメールアドレスのノードから有向リンクを受けていないノードであっても，隣接するノード間に図 4 であげたような送受信関係があれば得点は大きくなる．

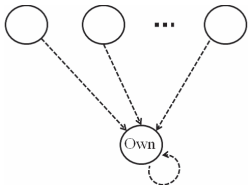

図 5　所有者のノードの受信スコアを最大にするための補正
Fig. 5　Revision for owner node's having maximum score.

次に，個々のノードの固有ベクトル中心性の得点を信頼度と見なして，スパム以外のノードを抽出し，そのノードに対応するメールアドレスをホワイトリストに登録する．

メールボックスの所有者の知人は，互いに知り合いである可能性が高く，メールのやりとりがあると考えられ，そこには中心的な役割を果たすノードが存在し，そのノードの固有ベクトル中心性の得点は大きくなる．スパムメールの送信者は，互いにメールのやりとりを行う可能性が少ないと考えられるため，そのノードの固有ベクトル中心性の得点は小さくなる．またスパムメールの被害者は，スパムメールを多数受信することで，そのノードの固有ベクトル中心性の得点は大きくなると考えられ，スパムメールの送信者と区別できる．

信頼度すなわち得点の大きなノードから直接的または間接的にメールを受けているノードは，固有ベクトル中心性の得点の定義から，同様に大きな得点を持つことになり，信頼度の大きい一連のノード群として容易に抽出できる．本手法では，メールボックスのデータのみを使用することを前提としているが，メールボックスの所有者が送信したメールのデータ（送信済みメールデータ）を利用するよう拡張すれば，知人のメールアドレスが容易に得られるので，先にあげた補正を実施してハムのノードの抽出精度を高めることも可能となる．

メールフィルタリングの一連の手続きは，次の 3 つのステップにまとめることができる．

(1)　メールのフィルタリング ··· 受信したメールのヘッダ部の From:からメールアドレスを取り出して，ホワイトリストのメールアドレスと比較する．メールアドレスが一致する場合は通常のメール（ハム）と判定し，それ以外の場合は判定不能とする．

(2)　メールのやりとりの有向グラフの更新 ··· 受信したメールのヘッダ部の From:，To:，Cc:のメールアドレスから，そのメールがどのメールアドレスからどのメールアドレスへ送られたのかという関係を，メールの送受信関係の有向グラフに追加する．メール

ボックスの所有者のメールアドレスのノードに対して有向グラフ上のすべてのノードからリンクを張るように補正を行う．

(3) ホワイトリストの更新 ··· 有向グラフから構成した行列 $\mathbf{A}$ から，べき乗法などの近似法により，各ノードの固有ベクトル中心性の得点を求める．閾値 S を定めて，ノードの固有ベクトル中心性の得点がその閾値以上の場合，そのメールアドレスをホワイトリストへ追加する．

本手法では，他のノードから有向リンクを受けていない，出力リンクのみのノードの固有ベクトル中心性の得点について，次の定理が成り立つ．

定理 1　メールのやりとりの関係を表した有向グラフにおいて，出力リンクのみのノードの固有ベクトル中心性の得点は ϵ/M となる．

■

（証明）入力リンクのみを持つノード，入力リンクと出力リンクを持つノードの集合を S_1，出力リンクのみを持つノードの集合を S_2 とおく．行列 $\mathbf{A}$ は M 次正方行列であり，集合 S_1 のノードの個数を k，集合 S_2 のノードの個数を $M-k$ とおく．集合 S_1 のノード n_i，集合 S_2 のノード n_j に，$i<j$ となるように重複なく番号を付け，行列 $\mathbf{A}$ の行と列の番号に対応させる．このとき，行列 $\mathbf{A}$ は，次のような 4 つの部分行列 $\mathbf{V}_{S_1,S_1}$，$\mathbf{V}_{S_1,S_2}$，$\mathbf{V}_{S_2,S_1}$，$\mathbf{V}_{S_2,S_2}$ に分けることができる．

$$\mathbf{A}=\begin{bmatrix}\mathbf{V}_{S_1,S_1} & \mathbf{V}_{S_1,S_2}\\ \mathbf{V}_{S_2,S_1} & \mathbf{V}_{S_2,S_2}\end{bmatrix}, \tag{5}$$

$$\mathbf{V}_{S_1,S_1}=\begin{bmatrix}(1-\epsilon)\frac{v_{1,1}}{l_1}+\frac{\epsilon}{M} & \cdots & (1-\epsilon)\frac{v_{1,k}}{l_1}+\frac{\epsilon}{M}\\ \vdots & \ddots & \vdots\\ (1-\epsilon)\frac{v_{k,1}}{l_k}+\frac{\epsilon}{M} & \cdots & (1-\epsilon)\frac{v_{k,k}}{l_k}+\frac{\epsilon}{M}\end{bmatrix}, \tag{6}$$

$$\mathbf{V}_{S_1,S_2}=\begin{bmatrix}\frac{\epsilon}{M} & \cdots & \frac{\epsilon}{M}\\ \vdots & \ddots & \vdots\\ \frac{\epsilon}{M} & \cdots & \frac{\epsilon}{M}\end{bmatrix}, \tag{7}$$

$$\mathbf{V}_{S_2,S_1}=\begin{bmatrix}(1-\epsilon)\frac{v_{k+1,1}}{l_{k+1}}+\frac{\epsilon}{M} & \cdots & (1-\epsilon)\frac{v_{k+1,k}}{l_{k+1}}+\frac{\epsilon}{M}\\ \vdots & \ddots & \vdots\\ (1-\epsilon)\frac{v_{M,1}}{l_M}+\frac{\epsilon}{M} & \cdots & (1-\epsilon)\frac{v_{M,k}}{l_M}+\frac{\epsilon}{M}\end{bmatrix}, \tag{8}$$

$$\mathbf{V}_{S_2,S_2}=\begin{bmatrix}\frac{\epsilon}{M} & \cdots & \frac{\epsilon}{M}\\ \vdots & \ddots & \vdots\\ \frac{\epsilon}{M} & \cdots & \frac{\epsilon}{M}\end{bmatrix}. \tag{9}$$

$\mathbf{V}_{S_1,S_1}$ は集合 S_1 のノードから集合 S_1 のノードへの有向リンクを表した k 行 k 列の行列，$\mathbf{V}_{S_1,S_2}$ は集合 S_1 のノードから集合 S_2 のノードへの有向リンクを表した k 行 $M-k$ 列の行列を表す．$\mathbf{V}_{S_2,S_1}$，$\mathbf{V}_{S_2,S_2}$ についても同様に考える．集合 S_2 のノードは出力リンクを持たないことから，行列 $\mathbf{V}_{S_1,S_2}$，$\mathbf{V}_{S_2,S_2}$ の要素は，有向リンクなしを表す ϵ/M となる．ここで，行列 $\mathbf{A}$ を転置した行列を $\mathbf{A}^T$，単位行列を $\mathbf{E}$ とおくと，各ノードの固有ベクトル中心性の得点は，次の式を満足するような列ベクトル $\mathbf{x}$ を求める問題となる．

$$\begin{aligned}&(\mathbf{A}^T-\mathbf{E})\mathbf{x}\\ &=\begin{bmatrix}(1-\epsilon)\frac{v_{1,1}}{l_1}+\frac{\epsilon}{M}-1 & \cdots & (1-\epsilon)\frac{v_{k,1}}{l_k}+\frac{\epsilon}{M}\\ \vdots & \ddots & \vdots\\ (1-\epsilon)\frac{v_{1,k}}{l_l}+\frac{\epsilon}{M} & \cdots & (1-\epsilon)\frac{v_{k,k}}{l_k}+\frac{\epsilon}{M}-1\\ \frac{\epsilon}{M} & \cdots & \frac{\epsilon}{M}\\ \vdots & \ddots & \vdots\\ \frac{\epsilon}{M} & \cdots & \frac{\epsilon}{M}\end{bmatrix}\\ &\quad\begin{bmatrix}(1-\epsilon)\frac{v_{k+1,1}}{l_{k+1}}+\frac{\epsilon}{M} & \cdots & (1-\epsilon)\frac{v_{M,1}}{l_M}+\frac{\epsilon}{M}\\ \vdots & \ddots & \vdots\\ (1-\epsilon)\frac{v_{k+1,k}}{l_{k+1}}+\frac{\epsilon}{M} & \cdots & (1-\epsilon)\frac{v_{M,k}}{l_M}+\frac{\epsilon}{M}\\ \frac{\epsilon}{M}-1 & \cdots & \frac{\epsilon}{M}\\ \vdots & \ddots & \vdots\\ \frac{\epsilon}{M} & \cdots & \frac{\epsilon}{M}-1\end{bmatrix}\begin{bmatrix}x_1\\ \vdots\\ x_k\\ x_{k+1}\\ \vdots\\ x_M\end{bmatrix}=\begin{bmatrix}0\\ \vdots\\ 0\\ 0\\ \vdots\\ 0\end{bmatrix}.\end{aligned} \tag{10}$$

ここで，列ベクトル $\mathbf{x}$ の要素 $x_{k+1},\cdots,x_M$ に注目すると，次の式が明らかに成立している．

$$\frac{\epsilon}{M}(x_1+\cdots+x_M)=x_{k+1}=\cdots=x_M. \tag{11}$$

$x_{k+1},\cdots,x_M$ は，出力リンクのみを持つノードの得点であり，$\sum_{i=1}^{M}x_i=1$ であることから，次の式が書ける．

$$x_{k+1} = x_{k+2} = \cdots = x_M = \frac{\epsilon}{M}. \tag{12}$$

（証明終）

出力リンクのみのノードの固有ベクトル中心性の得点は，上記の定理から一定の値となり，また固有ベクトル中心性の定義からすべてのノードの得点の中で最小の値をとる．本手法では，出力リンクのみのノードの得点を x_{min} とおいたとき，実数 k（≥ 1）を定めて，閾値 S を次のように設定することにする．

$$S = k \cdot x_{min}. \tag{13}$$

ここで，メールボックスの所有者のノードの得点を x_{max}，このノードからの出力リンクの数を l とおいたとき，次のような不等式が書ける．

$$S < \frac{x_{max}}{l}. \tag{14}$$

メールボックスの所有者のノードは最も信頼できるノードであり，固有ベクトル中心性の定義から，その得点 x_{max} は最大の値をとる．メールボックスの所有者のノードから直接有向リンクを受けるノードは，メールボックスの所有者のノードに次いで信頼できると考えられる．このノードの得点は固有ベクトル中心性の定義より x_{max}/l 以上となることから，閾値 S の選び方に関して上記のような不等式が得られる．

4. 実　　験

著者が過去に受信したメールを用いて，スパムノードとハムノードで固有ベクトル中心性の得点の分布に違いが現れることを確認する実験と，提案手法と既存手法を単独で用いた場合と，提案手法とコンテンツベースの手法を併用した場合について，フィルタリングの性能を評価する実験を行う．

評価用メールセットとして，著者らが 2007 年 4 月から 7 月にかけて受信した，日本語と英語の混在したメールセットを用意した．このメールセットは，5,585 通（ハムが 3,013 通，スパムが 2,572 通）のメールを含んでおり，メールのやりとりの有向グラフのノードは 2,390 個，有向リンクは 4,027 本であった．実際のメールサービスにおいて収集したメールセットであることから，スパムメールの中にはヘッダ部の From:，To:，Cc:のメールアドレスを詐称したメールが含まれていると考えられる．

まず，評価用メールセットを用いて，ノードの固有ベクトル中心性の得点を計算したところ，**図 6** のような得点の分布が得られた．この実験では，ハムメールとスパムメールを混在させ，メールのヘッダ部の From:，To:，Cc:の情報から，メールの送受信関係の有向グラフを構成し，各ノードの固有ベクトル中心性の得点を計算している．有向グラフから構成した行列 $\mathbf{A}$ の要素に関するパラメータ ϵ については，$\epsilon = 0.1$ としている．固有ベクトル中心性の得点の度数は，得点を区間 $\Delta = 10^{-5}$ で各級に分けた後，カウントしている．図 6 では，ノードのメールアドレスがハムメールのヘッダ部に含まれるか，スパムメールのヘッダ部に含まれるかによって，ハムとスパムに分類している．ハムメールとスパムメールの両方に含まれる場合は，メールボックスの所有者またはその知人のメールアドレスを詐称したものと見なし，ハムに分類している．

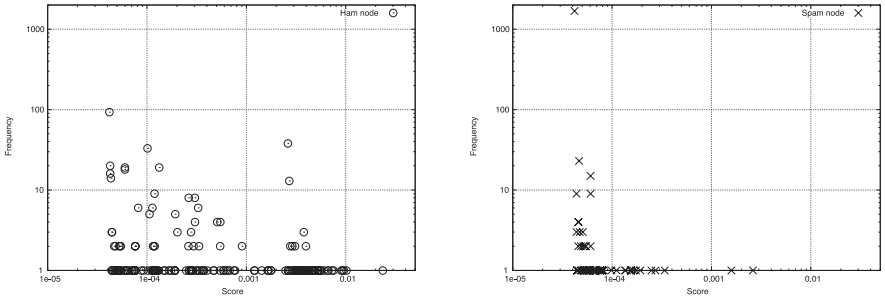

図 6　固有ベクトル中心性の得点分布
Fig. 6　Score distribution for each node.

図 6 では，スパムのノードは得点 $x_{min} = \epsilon/M = 0.1/2{,}390 = 4.1841 \times 10^{-5}$ のところに全体の約 92%が存在している．また，ハムのノードは得点 $x_{max}/l = 2.3573 \times 10^{-2}/127 = 1.8562 \times 10^{-4}$ 以上のところに全体の約 60%が存在しており，スパムのノードはこの得点以上のところに全体の約 1%が存在しているだけであった．このことから，ハムのノードとスパムのノードの得点分布の違いから，閾値 S を x_{min} から x_{max}/l の間に設定することで，提案手法を単独で用いた場合でも，良好な精度でハムのノードを判定できることを確認した．

次に，評価用メールセットに対して提案手法を適用して，メールフィルタリングの性能を確認したところ，**表 1** のような結果が得られた．この実験では，メールを 1 通ずつ受信した順に取り出し，ハム，スパム，判定不能のいずれかに判定し，さらにそのメールのヘッダ部の情報からメールのやりとりのグラフ，ホワイトリストを逐次更新している．パラメータ ϵ については $\epsilon = 0.1$ とし，ハムのノード抽出に関するパラメータ S については $S = 2\ x_{min}$ とした．

表 1　提案手法のメールフィルタリングの結果
Table 1　Filtering result of the proposed method.

	Success rate	False rate	Unsure rate
Ham	0.7866 (2,370/3,013)	—	0.2134 (643/3,013)
Spam	—	0.0292 (75/2,572)	0.9708 (2,497/2,572)

表 2　既存手法のメールフィルタリングの結果
Table 2　Filtering result of the existing method.

	Success rate	False rate	Unsure rate
Ham	0.8038 (2,422/3,013)	0.0003 (1/3,013)	0.1958 (590/3,013)
Spam	0.0000 (0/2,572)	0.0257 (66/2,572)	0.9743 (2,506/2,572)

表 1 において，提案手法はホワイトリストのみを作成することから，ハムのメールセットにおける失敗率（FPR）の項目，スパムのメールセットにおける成功率の項目は存在しない．ハムのメールセットにおいて正しく判定できたものは全体の約 79%，スパムのメールセットにおいて誤って判定したものは全体の約 3%であり，提案手法の良好な性能を確認した．

次に，提案手法のときと同様に，評価メールセットに対して文献 14) の既存手法を適用して，メールフィルタリングの性能を確認したところ，**表 2** のような結果が得られた．評価対象のコンポーネントを決定する際使用するパラメータ S_{min}，K_{frac} については $S_{min} = 10$，$K_{frac} = 0.6$ とし，ハムおよびスパムを抽出する際使用するパラメータ C_{max}，C_{min} については $C_{min} = 0.01$，$C_{max} = 0.1$ とした．

表 2 において，既存手法のスパムのメールセットの判定不能率が大きい理由は，評価用メールセットから構成した無向グラフのコンポーネントサイズが小さいためと考えられる．提案手法と既存手法はともにパラメータの設定値によって判定精度が変化することから，厳密な性能の比較を行うことは困難であるが，表 1 と表 2 から提案手法は既存手法と同等の精度でハムのメールを判定する能力を有していることを確認した．

表 1 と表 2 のハムのメールセットにおいて，ハムと正しく判定できたもののうち，提案手法のみで正しく判定できたものが約 7%，既存手法のみで正しく判定できたものが約 8%，既存手法と提案手法で共通して正しく判定できたものが約 72%であった．また，スパムのメールセットにおいて，ハムと誤って判定したもののうち，提案手法のみで誤って判定したものが約 1%，既存手法のみで誤って判定したものが約 1%，既存手法と提案手法で共通して誤って判定したものが約 2%であった．このことから，既存手法と提案手法を併用することで，特にハムのノードの判定精度の向上が期待できる．

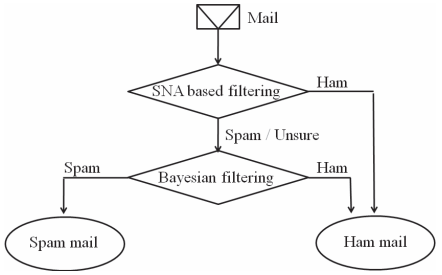

図 7　統合手法（A）
Fig. 7　The integrated method (A).

表 3　統合手法（A）のメールフィルタリングの結果
Table 3　Filtering results of the integrated method (A).

	Success rate	False	Unsure rate
ECM	0.4244 (2,370/5,585)	0.0134 (75/5,585)	0.5622 (3,140/5,585)
bsfilter	0.8417 (4,701/5,585)	0.1583 (884/5,585)	—
ECM+bsfilter	0.9386 (5,242/5,585)	0.0614 (343/5,585)	—

次に，文献 17) の統合手法（A）に基づいた，提案手法とコンテンツベースの手法（ベイジアンフィルタリング）[8)–10)] との連携の効果について確認する．統合手法（A）は，**図 7** のように，社会ネットワーク分析を用いたフィルタリングでスパム，判定不能と分類されたものを再度ベイジアンフィルタに通すことでハムメールの誤遮断率（FPR）を低減しようとするものである．ベイジアンフィルタリングは，メールのボディ部に現れる単語の出現確率を解析して高い精度でハム/スパムを判定できるコンテンツベースの手法である．

ここで，評価用メールセットに対して統合手法（A）を適用して，メールフィルタリングの性能を確認したところ，**表 3** のような結果が得られた．この実験では，社会ネットワーク分析に基づくフィルタリングに提案手法（ECM）を，ベイジアンフィルタリングに bsfilter [22)] を利用して，評価用メールセットからメールを 1 通ずつ受信した順に取り出して，統合手法（A）を適用している．パラメータ ϵ については $\epsilon = 0.1$ とし，またハムのメールアドレスを抽出する際使用するパラメータ S については $S = 2\ x_{min}$ とした．bsfilter では，ベイジアンフィルタの方式として Gary Robinson-Fisher 方式を，スパムメールを判定するときの閾値として 0.95 を指定している．フィルタ学習用の clean および spam の token には，評

価用メールセットとは別に，著者が 2006 年 11 月に受信したハム 480 通とスパム 338 通を使用した．ignore-header オプションを指定し，auto-update オプションは指定しない．

表 3 は，提案手法を単独で適用した場合，bsfilter を単独で適用した場合，両者を統合手法（A）に従って併用した場合のそれぞれについて成功率，失敗率，判定不能率をまとめたものである．提案手法を単独で適用した場合に比べて，bsfilter を併用した場合の方が成功率が約 51%向上しており，これは提案手法により判定不能とされたメールが，bsfilter によるメールのボディ部の情報を用いたフィルタリングにより救われたと考えられる．

5. 考　　察

判定可能なメールタイプやホワイトリストとブラックリストの精度を低下させる行為について，既存手法との違いを考察する．

提案手法と既存手法では，利用する社会ネットワーク分析の指標が異なることから，判定可能/不可能なメールタイプがそれぞれある．1 つ目は，既存手法では，メールのやりとりの様子を無向グラフで表現して，それを推移性指標であるクラスタリング係数を用いて評価することから，3 つのノードに閉路が存在しない場合，それをハムと判定できないというものである．これに対して，提案手法では，メールのやりとりの様子を有向グラフで表現して，各ノードを固有ベクトル中心性の得点により評価するので，ノード間の閉路の存在の有無にかかわらずノードの判定が可能である．2 つ目は，提案手法では，メールアドレスのノード 1 つに対して固有ベクトル中心性の得点を求め，その得点が閾値より大きいかどうかで個々のノードを判定しており，得点が最小となる送信のみのノードは，それがハムのノードであっても抽出できないというものである．これに対して，既存手法では，メールボックスの所有者のメールアドレスのノードを除いてできるコンポーネントを評価単位として，ハムのノードの集合をまとめて判定することから，送信のみのノードであってもハムのコンポーネントに含まれればそれを正しく抽出できる．3 つ目は，既存手法では，コンポーネントを評価単位とすることから，スパムと判定されたコンポーネントにハムのノードが含まれるような場合（スパムメールの From:，To:，Cc:のどれかにハムのメールアドレスが含まれる場合）があり，ハムメールを誤遮断してしまうというものである．これに対して，提案手法は，ノード単位で評価することから，このような場合であっても，ハムまたは判定不能のどちらかにノードが判定され，スパムメールの誤通過率（FNR）が増加することはあっても，ハムメールの誤遮断率（FPR）は増加しない．

次に，提案手法と既存手法では，メールのヘッダ部の From:，To:，Cc:のメールアドレスが詐称されるとホワイトリストやブラックリストの精度が低下する可能性があるので，その詐称する 3 つの方法について考える．1 つ目は，スパムメールの To:，Cc:に複数のメールアドレスを設定しないことで，コンポーネントのサイズおよびノードの次数を小さくして，評価対象から意図的に除外するという方法である．これは，提案手法では影響がないが，既存手法ではメールの判定不能率を増加させ，ブラックリストの精度を低下させる．2 つ目は，スパマがスパイウェアなどを利用して，事前にメールボックスの所有者の知り合いのメールアドレスを調べておき，スパムメールの To:，Cc:にそれを含めて，ハムと判定されたノードとの間に関係を持たせるという方法である．この方法はホワイトリストのメールアドレスを調べて，スパムメールの From:のメールアドレスになりすます行為とコストが同等であり，ホワイトリストによる手法の潜在的な脅威と同等と考えられる．これに対しても提案手法では影響がないが，既存手法ではホワイトリストの精度を低下させ，FNR を増加させる．3 つ目は，スパムメールの From:，To:，Cc:に含めるメールアドレスに送受信関係を持たせて，ハムと判定されるメールネットワークを意図的に作り出すという方法である．これは，既存手法と提案手法のどちらも，ホワイトリストの精度を低下させ，FNR を増加させる．しかし，ホワイトリストとブラックリストの精度を低下させる行為は，メール交換のグラフを構成する際に，メールボックスの所有者の送信済みメールや統合手法（B）を用いたり，統合手法（A）に従い他のルールベースの手法やコンテンツベースの手法と併用したりすることで対処可能[17] と考えられる．

6. おわりに

本論文では，社会ネットワーク分析における中心性指標の 1 つである固有ベクトル中心性に基づくメールフィルタリング手法を提案した．提案手法は，メールボックスに格納されたメールのヘッダ部から From:，To:，Cc:のメールアドレスを取り出し，メールのやりとりを表す有向グラフを構成する．有向グラフ上の各ノードについて固有ベクトル中心性の得点を求め，これを信頼度と見なし，メールボックスの所有者および知人のメールネットワークを抽出して，メールアドレスのホワイトリストを作成する．そして受信したメールのヘッダ部から From:のメールアドレスを抽出して，そのメールアドレスがホワイトリストに含まれるかどうかで，ハムメールを判定する．

提案手法は，推移性指標であるクラスタリング係数に基づく手法（既存手法）ではメールのやりとりの関係を無向グラフで評価していたところを，メールの送信と受信を区別して扱うことができる有向グラフで評価している．また，既存手法ではメールボックスの所有者の

メールアドレスのノードを除いて構成される部分グラフに対してクラスタリング係数を用いた評価が行われていたところを，所有者のメールアドレスのノードを含めて各ノードを固有ベクトル中心性の得点により評価している．提案手法では，評価値として固有ベクトル中心性の得点を採用することでノード間の得点すなわち信頼度の連鎖に基づく評価を実現しており，送信のみのノードの得点を基準にして信頼度を一意に定義，評価できるため，既存手法に比べてノードの分類を直感的に行える．

著者の用意した評価用メールセットを用いて，固有ベクトル中心性の得点の分布を確認する実験，提案手法や既存手法のメールフィルタリングの実験を行い，提案手法が既存手法と同様に高い判定性能を有することを確認できた．また，既存手法と提案手法は，利用している社会ネットワーク分析の指標の違いから，抽出可能なメールネットワークが異なり，提案手法と既存手法を併用することで，ハムのメールの判定において精度向上が期待できる．さらに，提案手法とベイジアンフィルタリングを連携させる実験を行い，提案手法で判定不能となっていたメールが，ベイジアンフィルタリングにより正しく判定され，判定の成功率が改善されることを確認した．

ノードの固有ベクトル中心性の得点は，ノードどうしの接続関係によって様々な値をとるため，今後の課題として，さらなる詳細な分析手法の検討があげられる．

参　考　文　献

1) Symantec Corporation: The State of Spam (2008). http://www.symantec.com/business/theme.jsp?themeid=state_of_spam
2) Allman, E., Callas, J., Delany, M., Libbey, M., Fenton, J. and Thomas, M.: Domain Keys Identified Mail (DKIM) Signatures, IETF, RFC4871 (2007).
3) Japan Email Anti-Abuse Group: JEAG Recommendation – Outbound Port 25 Blocking について (2006). http://jeag.jp/news/pdf/op25b20060223.pdf
4) 前野年紀：MTA でできる spam 撃退術，情報処理学会第 45 回プログラミング・シンポジウム報告書，pp.135–145 (2004).
5) 吉田和幸：greylisting による spam メールの抑制について，情報処理学会研究報告，Vol.2004, No.96, pp.19–24 (2004).
6) 吉田和幸：throttling による spam メール抑制の効果について，情報処理学会研究報告，Vol.2005, No.39, pp.69–73 (2005).
7) SpamAssassin Development Team: SpamAssassin. http://spamassassin.org/
8) Graham, P.: A Plan for Spam (2002). http://paulgraham.com/spam.html
9) Graham, P.: Better Bayesian Filtering (2003). http://www.paulgraham.com/better.html
10) Robinson, G.: A statistical approach to the spam problem, *Linux Journal*, Vol.107 (2003).
11) Hall, R.J.: Channels: Avoiding Unwanted Electronic Mail, *1996 DIMACS Symposium on Network Threats*, pp.85–103, American Mathematical Society (1997).
12) Gabber, E., Jakobsson, M., Matias, Y. and Mayer, A.: Curbing Junk E-Mail via Secure Classification, *Financial Cryptography '98*, LNCS 1465, pp.198–213, Springer (1998).
13) Jakobsson, M., Linn, J. and Algesheimer, J.: How to Protect Against a Militant Spammer, ePrint Archive, Report 2003/07 (2003).
14) Boykin, P.O. and Roychowdhury, V.P.: Leveraging Social Networks to Fight Spam, *IEEE Computer*, Vol.38, No.4, pp.61–68 (2005).
15) 李　鎮，黄　楽平，松浦幹太：社会ネットワークを利用したスパムメール対策におけるデータ共有の手法とその効果について，情報処理学会研究報告，Vol.2006, No.129, pp.69–72 (2006).
16) Samano, V.J.Z. and Matsuura, K.: Using time to classify spam, Information Processing Society of Japan, Technical Report, Vol.2007, No.126, pp.19–24 (2007).
17) 大原泰樹，松浦幹太：ベイジアンフィルタと社会ネットワーク手法を統合した迷惑メールフィルタリングとその最適統合法，情報処理学会論文誌，Vol.47, No.8, pp.2548–2555 (2006).
18) Wasserman, S. and Faust, K.: *Social Network Analysis – Methods and Applications*, Cambridge University Press, Cambridge (1994).
19) Page, L., Brin, S., Motwani, R. and Winograd, T.: The PageRank Citation Ranking: Bringing Order to the Web, Stanford Digital Library Technologies, Working Paper SIDL-WP-1999-0120 (1998).
20) Henzinger, M.: Link Analysis in Web Information Retrieval, *IEEE Bull. of the Tech. Committee on Data Engineering*, Vol.23, No.3, pp.3–8 (2000).
21) Haveliwala, T.H.: Efficient Computation of PageRank, Stanford Digital Library Technologies, Technical Report 1999-31 (1999).
22) Nabeya, K.: bayesian spam filter. http://bsfilter.org/

(平成 20 年 1 月 21 日受付)
(平成 21 年 12 月 18 日採録)

推　薦　文

社会的に対策の必要性が認識されている迷惑メールの対策手法として，社会ネットワーク分析を用いた迷惑メールフィルタリングを取り上げ，誤判定を軽減するため，有向グラフ上の各ノードについて固有ベクトル中心性の得点に基づく方法を提案している．受信者への一

方向の送達を効果的に得点に反映させており，実験により，ハムのメールの判定において有効であることを示している．加えて，提案方式を含む社会ネットワーク分析を利用した方式をベイジアンフィルタと連動させることにより，ハムのメールの判定において精度向上を示しており，価値が高い．以上のように，本論文は，有用性に優れており，本会会員にとって有益な内容であるため，推薦論文として推薦したい．

（コンピュータセキュリティ研究会主査　寺田真敏）

白石　善明（正会員）

1995 年愛媛大学工学部卒業．2000 年徳島大学大学院博士後期課程修了．博士（工学）．2002 年近畿大学理工学部講師，2006 年名古屋工業大学大学院工学研究科助教授，現在，准教授．情報セキュリティ，コンピュータネットワーク等の研究，教育に従事．2002 年電子情報通信学会オフィスシステム研究賞，2003 年暗号と情報セキュリティシンポジウム（SCIS）20 周年記念賞，2006 年 SCIS 論文賞，2007 年，2008 年 DICOMO2007，DICOMO2008 優秀論文賞．電子情報通信学会，IEEE 各会員．

福田　洋治（正会員）

2000 年徳島大学工学部卒業．2002 年徳島大学大学院博士前期課程修了．博士（工学）．2005 年愛知教育大学助手，現在，講師．情報セキュリティ，教育支援に関する研究，教育に従事．2002 年電子情報通信学会オフィスシステム研究賞，2008 年 DICOMO2008 優秀論文賞．電子情報通信学会会員．

溝渕　昭二（正会員）

1995 年徳島大学工学部卒業．2000 年徳島大学大学院博士後期課程修了．博士（工学）．2000 年理化学研究所協力研究員，現在，近畿大学理工学部講師．自然言語処理，情報検索等の研究，教育に従事．DICOMO2007 優秀論文賞．電子情報通信学会，人工知能学会各会員．

鈴木　貴史

2007 年名古屋工業大学工学部電気情報工学科卒業．2009 年名古屋工業大学大学院工学研究科博士前期課程修了．現在，ブラザー工業（株）．2007 年 DICOMO2007 最優秀プレゼンテーション賞，優秀論文賞，FIT2007 ヤングリサーチャ賞，2008 年本会東海支部学生論文奨励賞．電子情報通信学会会員．