社団法人 情報処理学会 研究報告
IPSJ SIG Technical Report

2006－CSEC－35（13）
2006／12／8

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

李 鎮† 黄 楽平†† 松浦 幹太†††

ユーザのメールデータから社会ネットワークを抽出し，その信頼的なネットワークからユーザの固有な独特の性質を以って正確な選択を選ぶのである．しかし，一人のユーザにこの手法を用いた場合，識別率が低いという欠点が伴っている．本稿ではこの社会ネットワーク手法に注目し，実験を通じて，複数のユーザが各自のメールの社会ネットワークを共有した場合，全体的の識別率が高まる結果を示す．

Exchanging Social Networks to Combat Spam Emails

Zhen Li† Leping Huang††
Kanta Matsuura†††

We introduce a collaborative anti-spam technique, which shares users' email social network(SN). In a single user case, although the detection rate (DR) is nearly 50%[1], the false positive rate (FPR) is 0. In the paper, we focus on the case when several users share their SN filter, and our experiment showed us a higher DR could be generated.

1. はじめに*

スパムメール問題は世界中のインターネットユーザを悩ましている．米国ユーザ宛のメールの 87%はスパムメールであり，更に今日のスパムメールの流量はメールトラフィックの 6 割に及ぶほどとなっている．その原因にはスパムメール発信者が，非常に少ないコストで大量のユーザ宛に一気に広告メールを送信することができることである．米国や欧州では，このような悪質な業者に対して，厳しい取り締まる法則が出ているが，業者側には送信サーバなどを法則力が弱い発展途上国に移ることで回避され，業者増加の勢いが更に深刻となっている．このような状況で，世界規模の法則が組み立てるまでに，技術による対策を中心として各団体が研究を進めている．

このような対策は，Email サービスプロバイダー(ESP)側とユーザ側で採用する対策の二つに大きく分けられている．前者には，発信者認証をメインに，SenderID や Domain Key などの研究が進んでいる．一方，ユーザ側では，受信するメールをフィルタリングする技術をメインにルールベース手法，コンテンツベース手法などが挙げられる．ルールベース手法とは，メールの特徴を抽出し，その特徴を集めてルール策定し，それらのルールによる判定を行う最も基本なフィルタリング手法である．逆にスパムメール業者側がそれらのルールの穴を通過するようにメールを修正すればこの手法を簡単に破れる．一方，コンテンツベース手法として，ベイジアンフィルタが近年に注目されている．それは，メールの本文から，形態素解析などによる単語を抽出し，代表的な単語群のスパムメール確率で個々のメールを判定する手法である．メール受信する度に学習と判定を行うので，常にメールの変化に追いつくが，コンテンツベース手法にある一定の誤判定率があり，それによって，ユーザの受信したいメールをスパムメールと判定してしまう欠点が伴っている．

本稿では，最近大きく取り上げられている社会ネットワークを利用した手法に注目している．この手法では，ユーザのメールデータから送受信関係の社会ネットワークを抽出し，それを分析してから，メールアドレスのホワイトリストとブラックリストを構築する手法である．この手法の大きく注目されている所は，誤判定率が 0 である．一方，判定精度が高いわりに，一部のメールしか判定できないという低判定率が伴っている．本稿では，一人のユーザが構築できた社会ネットワークを複数個集めて，複数のユーザ間で共有し合うことで，実験によって判定率が高くなることを示し

* †東京大学 生産技術研究所
Institute of Industrial Science, University of Tokyo
††ノキア・ジャパン
Nokia Japan
†††東京大学 生産技術研究所

ている．

次の２章では，個人ユーザの Email 社会ネットワークを構築する手法について説明する．３章では，複数のユーザ間で各個人の Email 社会ネットワークを共有しあう手法について述べる．４章，５章では，今回行った実験とその結果及び評価について説明する．

2. 社会ネットワークを利用したフィルタリング技術

ユーザのメールデータから送受信の社会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・ブラックリストを構築する手法である．メールアドレスはある意味では，一つのアイデンティティであり，それでユーザの社会ネットワークを構築すれば，そのネットワークにいる相手しか信頼できない．その前提として，構築された社会ネットワークが信頼できるものでなければならない．そのため，信頼できるネットワークを構築する手法が求められている．信頼性を示す一つの常識ルールとして，自分が相手をどれほど信頼するかだけでなく，ほかの人が相手をどれほど信頼するかにも依存する．これを利用すれば，信頼できる社会ネットワークを構築することができ，それを利用するスパムメールフィルタリング手法では，驚くことに，0 の誤判定率が示されている．

この手法では，ユーザのメールデータからのヘッダー情報に注目し，各 From，To，Cc ヘッダーからメールアドレスを抽出し，ユーザの社会ネットワークを構築する．まず，ユーザメールデータの各メールヘッダーから各アドレスをノードとして抽出する．次に，同じメールヘッダーに現れるアドレス同士間に枝を繋げる．これらは，このユーザを介して通信しあうユーザ同士を代表する．この手法ではこのユーザ同士間のコネクションにしか注目しないため，次にそのユーザ自身のアドレスを代表するノードを全部除く．これで，そのユーザの周りのメールアドレスネットワークが構築できる．

構築できたネットワークには，複数の独立したネットワークと個々の独立したノードが含まれている．この独立したネットワークに対して，ネットワークの親密度，つまり，クラスタリング係数を計算する．ユーザの知人同士はお互い知り合う可能性が高く，その間にメールのやり取りが存在することで，そのような独立したネットワークのクラスタリング係数が高くなる．一方，スパム業者はお互いにスパムメールを送信しなく，且つ大量の被害者同士の間に知り合う可能性が低いため，そのような独立したネットワークのクラスタリング係数が非常に低くなる．

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

式(1)では，あるノード i の周りに ki 個の隣接ノードが存在することを表す．Ei は実際の隣接ノード間のコネクションの数を表す．N はその独立したネットワークにあるすべてのノードの数を表す．但し，$N2$ は少なくとも２個以上の隣接ノードを持つノードの数を表す．

学習のプロセスでは，式(1)よりクラスタリング係数 C を計算して，0.1 以上ならば，Ham(送受信可)ネットワークとし，そこにあるアドレスを全部ホワイトリストに登録する．0.01 以下ならば，Spam(送受信不可)ネットワークとみなし，ブラックリストに登録する．その間にいるネットワークのアドレスをグレイリストに登録する．次に判定のプロセスでは，各リストにあるアドレスと照合して新しく受信したメール判定する．

この手法では，全メールデータの一部しか判定できないという低判定率がともなっている．

3. 複数ユーザによる Email 社会ネットワークの共有

個々のユーザが構築できた社会ネットワークを複数集めて，それを複数のユーザ間で共有し合うことによって，協力的なフィルタが構築でき，全体の判定率を高くすることが実現できる．つまり，一つの薪は簡単に折れるが，三つ集めれば簡単に折れなくなることと相当する．

協力的なフィルタを構築する場合，どんなユーザが参加するか，どんなユーザと共有するかの問題を抱えている．信頼できるユーザと信頼できないユーザ，そして関わりの強いユーザと関わりの弱いユーザ，のように二分岐で考えるとする．信頼できないユーザと共有し合う場合には，更に信頼性を評価するプロセスが必要で，何より，そのようなユーザと共有し合うのを諦める傾向が一般ユーザが取ると考えられる．一方，社会ネットワークはある種の信頼のネットワークであり，その前提として，信頼性が求まる．そのため，参加するユーザを個個のユーザの Email 社会ネットワークに存在するユーザとし，更に，そのようなユーザ間でのみ共有し合うことを前提とする．

関わりの強いユーザというのは，日常によく合える友達や仕事同士のように，メールのやり取りが多いユーザのことを意味する．そのようなユーザの社会ネットワークには，お互いの知り合いが出現する可能性が

高く，それぞれの Email 社会ネットワークにある個々の独立した Ham ネットワークが共通の知り合いにより，より大きく成長できることがある(図 1).

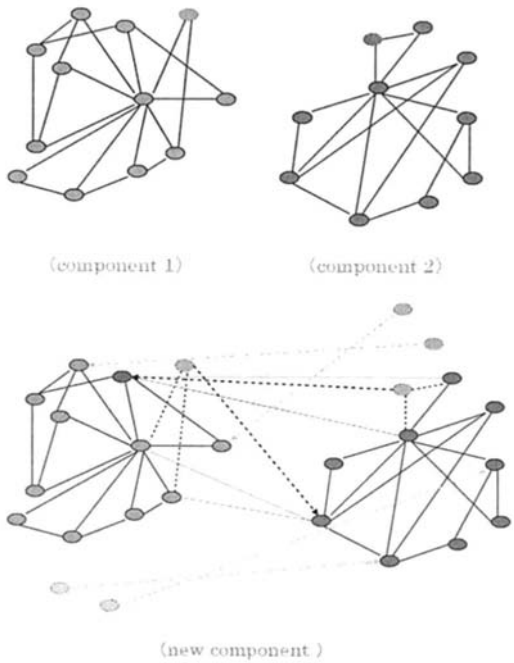

図 1，薄色のノードは，それぞれの Email 社会ネットワークに存在した独立したノードを表す．共有後の Email ネットワークに独立した Ham ネットワークに現れた.

図 1 から見えるように，更に，個々のユーザの Email 社会ネットワークに存在していた独立したノードも，それぞれのネットワーク内で対応しないのに比べて，共有し合う他のユーザの Ham ネットワークに存在する場合，新しくできた共有の Email 社会ネットワークに独立した Ham ネットワークに付け加えることができる．これによって，個々で判定できないアドレスを判定することができるようになり，全体の判定率があがることとなる.

4. 実験

表 1 に示すように，実験対象として，2006 年 5 月から 9 月までの 4 つのアカウントから集めたメールデータを利用した.

account	Lab1	Lab2	Hotmail	kc
All	191	220	852	174
Ham	158	172	445	93
Spam	33	48	407	81

表 1，実験用メールデータ数

この 4 つのアカウントの中に，Lab 1 と Lab2 は同じ研究室のメンバーのアカウントであり，Hotmail は Lab1 と同じユーザのアカウントである．更に，kc は他研究室のアカウントである.

実験として，関わりの強いユーザ間で共有し合う場合と，関わりの薄いユーザ間で共有し合う場合の二つに分かれている．更に，その日常生活の関わり度によって，高いほうから低いほうに，4 つの実験を行い，結果を観測した．観測対象として，判定率(Detection Rate)をメインに，誤遮断率（FPR:false-positives rate=正当な電子メールが迷惑メールとして遮断されてしまう確率），誤通過率（FNR:false-negatives rate=迷惑メールが正当なメールとみなされフィルタを通過してしまう確率）をも観測した.

5. 結果と評価

まず始めに，同じユーザの二つのアカウント(Hotmail, Lab1)から集めたメールデータから共有し合う社会ネットワークを構築した(表 2). 両方にも日常生活でよく利用されるアカウントであり，一部両方に来る同じ内容のメールデータが存在したが，ヘッダー情報が違い，それぞれ違うメールだとみなされる.

	Hotmail	Lab1	Com	*
Mail(ham)	852(445)	208(175)	1060(620)	
Ham(%)	40.4 (180/445)	47.3 (83/175)	47.6 (295/620)	(32)†
Spam(%)	51.5 (210/407)	31.2 (10/33)	50.0 (220/440)	
DR (%)	36.4	44.7	48.6	

*FNR, FPR both is 0%

表 2

次に，同じ研究室の 2 つのアカウントから集めたメールデータから共有し合う社会ネットワークを構築した(表 3).

	Lab2	Lab1	Com	*
Mail(ham)	220(172)	208(175)	428(347)	
Ham(%)	32.0 (55/172)	47.3 (83/175)	43.5 (151/347)	(13)†
Spam(%)	70.8 (34/48)	31.2 (10/33)	59.3 (48/81)	(4)†
DR(%)	40.5	44.7	46.5	

*FNR, FPR both is 0%

表3

そして，より関わりの弱い 2 つのアカウント Lab2 と Hotmail から集めたメールデータから共有し合う社会ネットワークを構築した(表4).

	Lab2	Hotmail	Com	*
Mail(ham)	220(172)	852(445)	1072(617)	
Ham(%)	32.0 (55/172)	40.4 (180/445)	38.2 (236/617)	(1)†
Spam(%)	70.8 (34/48)	51.5 (210/407)	53.6 (244/455)	
DR(%)	40.5	36.4	44.8	

*FNR, FPR both is 0%

表4

最後に，関わりの強い 2 つのアカウント Lab1 と Hotmail に関わりの弱いアカウント kc の 3 つから集めたメールデータから共有し合う社会ネットワークを構築した(表45).

	Hotmail	Lab1	Com	kc	Com †	
Mail(ham)	852(445)	208(175)	1060(620)	174(93)	1234(713)	
Ham(%)	40.4 (180/445)	47.3 (83/175)	47.6 (295/620)	32.0 (39/93)	49.5 (353/713)	(19)†
Spam(%)	51.5 (210/407)	31.2 (10/33)	50.0 (220/440)	22.2 (18/81)	45.7 (238/521)	
DR (%)	36.4	44.7	48.6	40.5	47.9	

表5

表2から表5まで，いずれの場合にも判定率が上がることが分かる．特に，表5から分かるように，関わりの弱いユーザと共有し合うより，関わりの強いユーザ間で共有し合う場合の判定率がより多く上昇するができる．

6. おわりに

本稿では，複数のユーザ間で Email 社会ネットワークを共有し合う場合に，全体の判定率が高くなることを示した．ユーザの社会ネットワークに存在するユーザ，即ち信頼できるユーザと共有し合う前提として，関わりの強いユーザと共有し合うほうが全体の判定率がより高く上昇する．

一方，信頼できるユーザ間でも，個人差によって，同じメールを違う定義する場合も考えられる．そのような場合の対策も必要である．更に，共有し合うプロセスにユーザのプライバシーが保護された前提で行わなければならないので，プライバシー保護の対策も今後取り入れる必要がある．

参考文献

1) P.O.Boykin and V.Roychowdhury. Leveraging social networks to fight spam. IEEE Computer, Vol.38, no. 4, pp. 61-68, 2005

J. S. Kong, P. O. Boykin, B. Rezaei, N. Sarshar,and V. Roychowdhury. Let your cyberalter ego share information and manage spam. 2005. pre-print http://xxx.lanl.gov/abs/physics/0504026

2) J.S.Kong, P.O.Boykin, B.Rezaei, N.Sarshar, and V.Roychowdhury. Let your cyberalter ego share information and manage spam. 2005. Pre-print.http://xxx.lanl.gov/abs/physics/0504026

3) Joshua T. Goodman, Robert Rounthwaite. Stopping outgoing spam. May 2004. Proceedings of the 5th ACM conference on Electronic commerce

4) www.usenix.org/events/sec05/tech/bethencourt/bethencourt.pdf

5) H. Ohfuku, K. Matsuura. Integration of Bayesian Filtering and Social Network Technique. CSS, Vol.1, pp.325-330, 2005