LL-008 SMTPセッションの強制切断によるspamメール対策手法
An Anti-Spam Method with SMTP Session Abort

山井 成良 † Nariyoshi Yamai†
漣 一平 ‡ Ippei Sazanami‡
岡山 聖彦 † Kiyohiko Okayama†
河野 圭太 † Keita Kawano†
中村 素典 § Motonori Nakamura§
丸山 伸 § Shin Maruyama§
宮下 卓也 * Takuya Miyashita*

1. はじめに

電子メールはWWWと並んでインターネットにおいて最も普及しているサービスの1つであり，社会的な活動を支える通信手段としてもはや必要不可欠な存在となっている．一方，電子メールはセキュリティ上最も問題の多いサービスの1つである．特に，広告やphishing詐欺などを目的に不特定多数の利用者に一方的に送信されるspamメールの蔓延は大きな社会問題にまでなっており，その対策は重要である．

spamメールへの対策方法としてブロッキング，フィルタリングなどの方式が多くの組織で用いられている．このうちブロッキングはSMTPセッション開始時あるいはセッション中に送信元の判別などによりspamメールの受信を拒否する方式で，代表的な手法としてtempfailingが知られている．tempfailingでは，信頼できないMTA（Mail Transfer Agent）からの電子メール配送を一時的に拒否することにより，再送処理を行わないMTAからの電子メール（spamメールの可能性が非常に高い）を排除することが可能であるが，たとえば異なるMTAから再送するような一部のドメインについては管理者がMTAを手作業で登録する必要があるなどの問題がある．一方，フィルタリングは電子メールの受信後にヘッダや本文などの解析に基づきspamメールを分離する方式で，代表的な手法として分散協調spamデータベースが知られている．分散協調spamデータベースは，多数の利用者が協調してspamメールをデータベースに登録するようにし，データベースへの登録の有無によりspamメールの判定を行うもので，誤認識率（非spamメールをspamメールと判定する率）が比較的小さいという利点を有する反面，認識率（spamメールを正しくspamメールと判定する率）が低いという問題を有する．

そこで本稿では，上記のような問題を軽減するためのspam対策手法を提案する．本方法ではSMTPコネクションを受信側MTAで途中で強制切断することにより送信側MTAの再送処理を促す．これによりtempfailingと同様の効果を得ながら，電子メールのヘッダや本文を取得することが可能になる．また，取得したヘッダや本文を活用して管理コストの低減や分散協調spamデータベースの認識率向上も図ることができる．

†岡山大学, Okayama University
‡株式会社日立製作所, Hitachi, Ltd.
§京都大学, Kyoto University
*津山工業高等専門学校, Tsuyama National College of Technology

2. 従来のspamメール対策方法とその問題点

前節で述べたように，spamメールへの対策方法としてこれまでに様々な手法が提案されている．そのうち，本節では提案手法との関連が深い2つの手法についてその概要ならびに問題点を述べる．

2.1 tempfailing

RFC2821[1]によると，SMTPセッション中に受信MTAから一時的なエラーを表す400番台の応答を受け取った場合，送信MTAは一定の期間待った後に再送処理を行なわなければならない．ところが，spamメールの発信に用いられるMTA（spam発信MTA）は一般にspamメール配送の確実性よりもスループットを重視するため，受信MTAが一時的なエラーを返しても再送処理を行なわないものが多い．tempfailingでは，このような挙動の違いを利用してspam発信MTAの判定を行ない，spamメール送信を拒否する．具体的には，送信MTAと受信MTAとの間でSMTPセッションが開始されると，受信MTAはこれが1回目の配送であるか2回目以降の配送（再送）であるかを判別する．もし，1回目の配送であれば，受信MTAは敢えて400番台のエラー応答を返すことで再送を促し，逆に2回目以降の配送であれば通常通りの受信処理を行うようにする．その際の再送判定に用いる情報として，greylisting[2]では（送信MTAのIPアドレス，エンベロープFromアドレス，エンベロープToアドレス）の3つ組が用いられる．また，再送を行ったMTAは自動的にwhitelistと呼ばれるデータベースに登録され，このデータベースに登録されているMTAからの電子メール配送は信頼できるものとして一時的なエラー応答を返さずに通常の受信処理を行う．

tempfailingは比較的簡単な仕組みであるにもかかわらず，高い効果が期待できる点が優れている．一方，tempfailingの問題点としては，通常メールの配送遅延が挙げられる．すなわち，RFC2821では再送間隔は30分以上とすることが推奨されているため，whitelistに登録されていないMTAからの配送はたとえ実際には配送に支障がなくても30分以上の遅延が生じることが予想される．また，再送が初回配送時とは異なるMTAから行われる場合には，さらに以下のような問題が生じる．大規模なドメインでは，負荷分散のために複数のMTAを用意し，前回とは異なるMTAを用いて再送を試みることがある．このような場合では再送の度に異なるMTAが用いられるため，受信MTAでは毎回再送でないと判定され，その都度一時的なエラーが返されることになる．これに対処するには，このような挙動をする送信MTAを管理者

が手作業で whitelist に登録する必要があり，管理コストの増大につながる．

2.2 分散協調 spam データベース

分散協調 spam データベースは，同一内容の電子メールが多数の利用者に送信されるという spam メールの性質を利用した方式である．この方式では利用者間で共有する spam メールのデータベースを導入し，このデータベースを用いて spam メールの判定を行う．すなわち，利用者は協調して各自が受信した spam メールのデータベースへの登録を行い，各 MTA は電子メール受信時にこのデータベースへの登録の有無により電子メールが spam メールであるかどうかを判定する．その際，高速化を図るため，spam メールの本文全体ではなく一種のチェックサム（本文中の部分的な改変に対応するために工夫をしているものが多い）を計算してこれをデータベースへの登録・照合に用いる．代表的な分散協調データベースには DCC （Distributed Checksum Clearinghouse）[3] がある．

この方式は，データベースへの登録が利用者の判断に基づいて行われるため，誤認識率は事実上無視できるほど小さい点が優れている．一方，この方式では spam メール判定時までに他の利用者がデータベースに登録していないと spam メールを認識できないため，認識率が比較的低いという欠点がある．これに対して，我々は宛先不明メールを spam メールと見なしてデータベースに登録する方法を提案し [4]，一定の成果を得た．しかし，この方法では，特にメーリングリストにおいて異動等により無効になったアドレス宛に配送された電子メールなど，一部の非 spam メールについても誤ってデータベースに登録する危険性がある点が指摘されている．

3. SMTP セッションの強制切断による spam メール対策手法

前節で述べたように，tempfailing や分散協調 spam データベースのような従来の spam 対策方法では，いずれも運用上あるいは性能上の問題があった．そこで本節では，両者を組み合わせ，さらに SMTP セッションの強制切断を導入することにより，これらの方式の問題点を軽減する手法を提案する．

3.1 システム構成と提案手法の概要

本手法では，従来の MTA をそのまま用いながら tempfailing と分散協調 spam データベースの両方の機能を導入するため，新たに同機能を持つメールゲートウェイを導入する．システム構成を図 1 に示す．この図に示すようにメールゲートウェイは少なくとも Primary Mail Gateway（PMG）と Secondary Mail Gateway（SMG）の 2 台から構成されている．また spam メールを登録するためのデータベースが別途導入されている．

PMG と SMG は，同図における各末端 MTA（図中右端の Server）のそれぞれプライマリ MX，セカンダリ MX として指定されるように DNS の設定を行う．またこれらは RST フラグ付きパケット送出による SMTP セッションの強制切断機能を持ち，初回配送時にはヘッダや本文を受信した時点で強制切断を行うように動作する．したがって，末端 MTA 宛の電子メールを配送する場合，

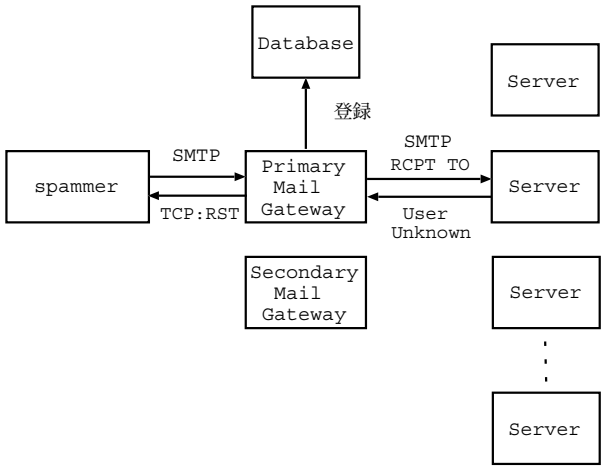

図 1: システム構成

送信 MTA はまず MX の優先順位に従って PMG への配送を試みるが，SMTP セッションの強制切断により配送に失敗する．すると送信 MTA は複数の MX レコードが存在するため直ちに SMG への配送を試み，SMG ではこれを受理する．これにより，tempfailing における配送遅延を軽減することが可能となる．

また，SMTP コネクションの強制切断機能の導入により，初回配送時に取得したヘッダや本文をもとに高度かつ迅速な受信制御を行うことが可能になる．初回配送時に必ず電子メール全体を取得する場合には本文受信後に一時的なエラー応答を返す技法により実現できるが，この場合には同一内容の通信が 2 度行われるため，特に大きいサイズの電子メールを受信する場合には通信量が増加するという問題が生じる．これに対して本手法では，強制切断の早期実行により改善することが可能である．たとえばアンチウイルスサーバを併用する環境では，添付ファイル付き電子メールはアンチウイルスサーバに配送の可否を判定させるようにして本手法では本文を spam メール判定に用いないようにすれば，ヘッダを受信した時点で添付ファイルの存在を確認して直ちに SMTP セッションを強制終了させるような動作が可能になり，通信量を削減するとともに再送判定を迅速に行うことができる．

3.2 再送判定処理の改善

従来の tempfailing では，再送判定において電子メールのヘッダや本文を全く利用せず，送信 MTA の IP アドレスやエンベロープ From/To のような SMTP セッションの情報のみに基づいて行っていたため，再送の度に異なる MTA を用いて配送が試みられる場合に正しく判定できなかった．そこで提案手法では，whitelist に登録されていない MTA からの配送の際には，ヘッダあるいは本文を受信した後に，SMTP セッションを強制的に切断する．これにより，従来のように異なった MTA からの再送に対してもたとえば Message-ID のようなヘッダ情報をもとに正しく再送判定を行うことが可能になり，管理コストの減少が可能になる．

3.3 spam メールの自動登録

従来の分散協調 spam データベースでは，利用者に届いた spam メールのみがデータベースへの登録対象となり，利用者に配送されなかった spam メールはそのまま破棄されていた．特に tempfailing において再送されない電子メールは spam メールの可能性が非常に高いにもかかわらず全く利用されず，これが分散協調 spam データベースの認識率の低下を招く要因のひとつであったと推測される．

一方，提案手法では 3.1 節で述べたように再送されなかった電子メールについてもヘッダや本文を取得することが可能であるため，これを登録することにより認識率の向上を図ることが可能である．また，その際，基本的には宛先アドレスの存在の有無ではなく再送の有無（あるいはその両方）によって spam メールの判定を行うため，2.2 節で述べた文献 [4] の手法における問題点についても軽減することが可能である．具体的には，再送された電子メールについては受信時にデータベースからの削除を行って登録を取り消すことにより，他の利用者への配送に対する影響を軽減することが可能となる．

3.4 全体の処理手順

これまでに述べた方法をまとめた，システム全体の典型的な処理手順の詳細を以下に示す．なお，以下の手順においては PMG の動作を記述しているが，spam 送信 MTA の中には MX の優先度の指定に従わないものも数多く存在するため，実際には PMG, SMG とも同じ動作を行う．

1. 送信 MTA は，プライマリ MX として指定されている PMG との間で SMTP セッションを開始する．PMG は送信 MTA の IP アドレスを記録する．
2. 送信 MTA は PMG に対して MAIL FROM コマンドおよび RCPT TO コマンドを発行し，発信者アドレスおよび受信者アドレスを指定する．PMG は末端 MTA 宛の配送であればこれらのコマンドに対して肯定応答（250 OK）を返し，これらのアドレスを後の再送判定で用いるために記録しておく．
3. 送信 MTA は DATA コマンドを発行する．PMG は同コマンドを受け付けた後，ヘッダの内容を受信する．
4. PMG はヘッダ中の Message-ID フィールドを抜き出し，たとえば（エンベロープ From アドレス，エンベロープ To アドレス，Message-ID）の 3 つ組を用いて再送判定を行う．また，送信 MTA の IP アドレスに基づいて whitelist への登録の有無を判定する．
5. 判定の結果，whitelist に登録された MTA からの配送であると判断された場合には，通常の配送処理を行い，終了する．また，初回配送であると判断された場合には，6. へ進む．それ以外の場合には，7. へ進む．
6. PMG は本文を受信した後 SMTP コネクションを強制終了させ，送信 MTA に再送を促す．また，受信したヘッダや本文をデータベースに登録して終了する．
7. PMG は再送されてきた電子メールに対して通常の配送処理を行う．また，初回配送時に行ったデータベースへの登録を無効化する．

4. 試作システムの実装と動作試験

前節で述べたシステム構成および動作手順に基づき，我々は spam メール対策システムの試作を行った．試作システムでは PMG, SMG として FreeBSD 上で sendmail を稼動させ，分散協調 spam データベースとしては DCC[3] を用いた．本システムの動作では，初回配送時にヘッダおよび本文を全て受信した後に SMTP セッションの強制切断を行うようにした．SMTP セッションの強制切断には外部プログラムにより RST フラグ付きパケットを生成し，これを送信 MTA，受信 MTA の両方に送るようにした．再送判定方法としては，3.4 節の手順 4. で述べた 3 つ組を用いるようにした．また，データベースへの登録は初回配送時に取得した本文全体を対象としたが，後述する動作試験において受信者の同意を事前に得るまでには至らなかったため，3.4 節の手順 2. の段階で宛先アドレスの存在確認を行い，宛先不明メールのみを登録の対象とした．

次に，試作システムの動作試験を行った．

まず，外部ネットワークから隔離された実験ネットワークにおいて，試作システムの基本動作を調査した．その結果，SMTP セッションの強制切断機能，再送判定機能，宛先不明メールのデータベース登録機能，再送された電子メールのデータベース登録削除機能のいずれもが正しく動作することが確認された．

次に，動作試験の一環として試作システムを外部のネットワークに繋ぎ，実際に外部から送られてくる spam メールの処理を行った．この試験運用では，近々廃止される予定の岡山大学内のあるドメイン宛の電子メールを試作システムが一旦受け取るように MX レコードを書き換え，その後試作システムが受け取った電子メールを本来の末端 MTA に配送するようにするようにした．このドメイン宛に送信される電子メールの大半は宛先不明メールで，また，その多くが spam メールであると思われるが，その確認には受信者の同意が必要となるため行うことができず，実際の spam メールの割合は不明である．この試験運用では，このような環境において，2006 年 1 月 29 日から 2 月 5 日までの 7 日間連続して試作システムを運用した．

試験運用では，まず提案手法のブロッキング機能の有効性を確認するため，試作システムのログを解析し，動作試験の期間中に試作システムで処理したメールのうち，初回の配送に失敗した後に再送されたメールと再送されなかったメールの数を算出した．その結果を表 1 に示す．

この表から，試作システムが処理した全メール 54,719 通のうち，81% にあたる 44,303 通のメールを試作システムにおいてブロックできたことが分かる．これは従来の greylisting と同等の性能であり，これより提案手法の SMTP セッション強制切断機能がブロッキング機能として十分に動作するといえる．

表 1: 試験運用の結果

	再送あり	再送なし	合　計
PMG	5,340	34,415	39,755
SMG	5,076	9,888	14,964
合計	10,416	44,303	54,719

一方，表 1 に示すように，全メールの 19%にあたる 10,416 通のメールが再送されたが，今回の実験環境では，上記の理由から再送されたメールにも多数の spam メールが含まれていたと推測される．そこで，再送されたメール 10,416 通のうち，分散協調 spam データベースにより spam メールと判定できるものの数を調べたところ，2,180 通の電子メールが該当することが判明した．正確な spam メール認識率については，残念ながら 10,416 通のうち何通が spam メールであるかが不明であるため算出できなかったが，提案手法による再送されない電子メールのデータベースへの登録がある程度の効果を持つことが確認できた．分散協調 spam データベースへの登録が試験運用期間中（7 日間）の宛先不明メールだけであったことを考慮すると，この数は長期間の運用ではさらに増加することが期待される．

5. むすび

本研究では，従来の tempfailing と分散協調 spam データベースを組み合わせ，さらに SMTP セッションの強制切断機能を導入することにより，従来の手法の問題点を軽減する spam 対策手法を提案した．また，提案手法に基づいて試作システムを実装して試験運用を行い，提案手法の有効性を確認した．

今後の課題として，長期に渡る実運用を通じての提案手法の性能評価が挙げられる．

謝辞

本研究の一部は平成 17〜19 年度科学研究費補助金（基盤研究（B），課題番号 17300038）の補助を受けている．

参考文献

[1] J. Klensin: “SIMPLE MAIL TRANSFER PROTOCOL”, RFC 2821 , April 2001.

[2] Evan Harris: “Greylisting”, `http://projects.puremagic.com/greylisting/index.html`.

[3] Rhyolite Software: “Distributed Checksum Clearinghouse”, `http://www.rhyolite.com/anti-spam/dcc/`.

[4] 漣一平，山井成良，岡山聖彦，宮下卓也，丸山伸，中村素典：“宛先不明メールを利用した分散協調型 spam フィルタの認識率向上”，情報処理学会分散システム/インターネット運用技術研究会研究報告，2005-DSM-37-15，pp.79–84，平成 17 年 5 月．