応答遅延を用いたスパム対策とその運用について
On the SPAM blocking using the SMTP throttling

浜元 信州 †, 青山 茂義 †, 三河 賢治 †

Nobukuni Hamamoto†, Shigeyoshi Aoyama†, Kenji Mikawa†

hamamoto@cais.niigata-u.ac.jp

新潟大学 情報基盤センター †
Center for academic information service, Niigata University†

概要

本報告では，公開リスト，HELO コマンドの引数を利用して，選択的に到着メールへの応答遅延，または，拒否を行いスパムメールを拒否する方式を提案する．さらに，この方式の適用例として，新潟大学メールゲートウェイサーバでスパムメール対策を行なった結果について報告する．

通常，スパムメール対策では，ホワイトリストなどのメンテナンスが欠かせないが，この方式では，公開リストを利用するため，リストのメンテナンスをサーバ管理者側では行わない．また，公開リスト，HELO コマンドの引数を組み合わせて，応答遅延する接続を選択することにより，応答遅延する接続数を抑えることが出来る．

導入から現在まで，偽陽性判定の報告はなく，メンテナンスフリーでの運用が出来ている．ユーザに届くメール数は， 対策前と比較して 3 割減となり，特定のメールアドレス宛てに対する調査では，本方式導入前に届いていたスパムメールの 9 割を削除することが出来ていることが分かった．

Abstract

We proposed a new method to block SPAM mail by selecting throttling and droping smtp connections using some blacklist and the argument of HELO command sent by a remote server. We implemented the method to the mail gateway server of Niigata university. In this paper, we describe how our method blocks SPAM mail and the blocking, false-positive/negative rate observed at our server.

In the present method, server adminstrator does not need to manage the whitelist of mail servers. Instead of the whitelist, we use some of the well-maintaind public lists. Furthermore, we can reduce the number of the concurrent connections by selecting throttling connections using the argment of HELO command and some blacklist.

After implementing our method to the mail gateway server, we do not recieved any false-positive report from our user. The arrived mails to our server and arrived SPAM mails to the paticular mail address are reduced by approximately 30% and 90 %, respectively.

キーワード

ネットワーク・システム関係, セキュリティ管理関係

1 はじめに

近年，インターネットを流れるメールの 80%ものメールがスパムメールであるとも言われている [1]．この状況は，ネットワークトラフィックの無駄であると同時に，スパムメールの処理がメールサーバに負荷をかけるなど，管理者にとっても問題である．また，個々のユーザにおいてもスパムメールに紛れて大事なメールを見逃すなどの問題がある．特に，近年ではフィッシングの手

段として使われるなど，ただの迷惑だけではなく，実害もあり，社会問題になりつつある．

新潟大学では，メールゲートウェイサーバを運用している．大学内に宛てたメールは全てこのサーバを経由するように構成し，メールのウイルス対策を，このサーバ上で行っている．近年，スパムメール対策への要望が高まっていることから，2006 年 6 月より，過去にスパム送信したホストの IP を登録している公開ブラックリスト（公開リスト A）を利用してスパムメール対策サービスを新たに行った．この対策は公開リスト A に登録されている IP からの接続を拒否する方式である．これはサーバ負荷軽減など，一定の効果をあげたものの，検出率が十分とは言えなかった．このため，2008 年 8 月より，応答遅延を組み合わせた新方式によるスパム対策を行った．本報告ではその方式と実運用での評価結果について述べる．

2 提案方式と実装

今回，新潟大学のメールゲートウェイでは，下記の提案方式を利用してメールサーバからの接続を拒否，または，応答遅延することによってスパムメール対策を行った．応答遅延は，メールサーバから接続があった際，それへの応答を一定時間行わない方法である．RFC2821 では，各コマンドのタイムアウト値が SHOULD で規定されているが，スパムメール送信などの大量のメール送信を行う場合には，効率よくメールを配送するため，タイムアウトの前に接続を切るケースが多い．このような場合には，応答を遅らせることで，結果として，スパムメールを受け取らないようにすることが出来る．応答遅延を用いたスパムメール対策としては，graylisting や spamassasin 等と組み合わせた方式による報告もある [2] が，今回は，HELO コマンドの引数のみと組み合わせた単純な方式を取り，配送遅延が発生せず，メールの中身を見ることのない方法を目指した．

提案方式

判定条件 1 公開リスト A（ブラックリスト）に登録されている IP からの接続は「拒否」

判定条件 2 下記 (2a),(2b) の両方を満たす場合は「拒否」，どちらか一方の場合は「応答遅延」

- **(2a)** 公開リスト B（エンドユーザ IP リスト）に登録されている IP からの接続
- **(2b)** HELO コマンドの引数が FQDN か IP リテラル以外

判定条件 3 HELO コマンドの引数が学内のホストを示す場合は「拒否」

なお，学内からの接続に対しては，上記の判定を行わずに全て接続を許可する．

応答遅延の問題点の一つに，大量接続があった際，サーバの接続数が増えることがある．これを防ぐため，判定条件 2 に示したように応答遅延を起こす接続を減らすよう試みている．(2a),(2b) の条件をつけることにより，応答遅延の対象となる接続を減らし，(2a),(2b) の両方の条件を満たした場合に「拒否」することで，応答遅延の対象となる接続をさらに少なくして，メールサーバへの負荷を減らしている．なお，拒否はエラーコード 550 で行ない，再送を要求しない．

新潟大学では，従来から判定条件 1 に基づくスパム判定を行っていたが，これだけでは十分な対策とは言えない状況が続いていた．判定条件 2,3 が今回追加された条件となる．

上記の実装を行うため，MTA は post x2.4 系を利用した．上記のうち，判定条件 2 の部分に関しては標準の設定では実現出来ないため，上記の条件を満たすよう post x ポリシーサーバ機能を利用して実装を行った [3].

今回の実装での接続判定は下記の順番で行われる．なお，post x では RCPT コマンド受け取った段階で，下記の確認が実行される．一つの接続で複数 RCPT を実行する場合には，応答遅延が複数回行われる．

1. 接続 IP アドレスが学内なら許可，学外なら下記を実行
2. 公開リスト A へ登録されていれば拒否（判定条件 1）
3. HELO コマンドの引数が学内ホストを示せば拒否（判定条件 3）
4. 学内サーバにユーザがいなければ拒否
5. ポリシーサーバへ転送し，判定条件 2 による判定を実行

判定条件 (2a) で用いる公開リスト B は，公開リスト A とは性質の異なるものを利用している．一般に，RBL と呼ばれるブラックリスト型の公開リストの問題点の一つとして，スパム配信行為や，サーバ乗っ取りなどの検出，登録に時間がかかるため，登録前に送信されたスパムを防げないという問題が指摘されている [4]．現在，ボットネットを利用した，エンドユーザ回線からのスパムメール送信が主流であると推測される [5] ことから，今回の対策では，エンドユーザ回線の IP アドレスを登録している公開リスト (公開リスト B) を利用する．このリストは，過去のスパム配信行為や，サーバ乗っ取りなどの事実がなくともエンドユーザ回線を一律に登

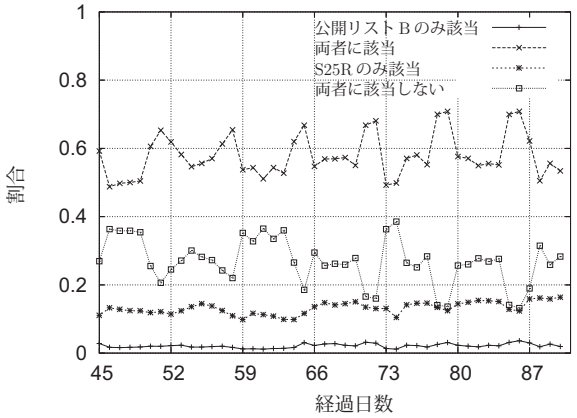

図- 1: S25R と公開リスト B の登録状況比較.

		公開リスト B	
		登録あり	登録なし
S25R	該当あり	57.8%	13.4%
	該当なし	2.1%	16.5%

表- 1: 公開リスト B，及び，S25R 方式によるエンドユーザ回線の判定結果比較表.

録しているため，エンドユーザのウイルス感染などにより送信されるスパムを防ぐことが出来る.

同様のエンドユーザ回線の判定法には，S25R 方式という簡便な方法がある [6, 7]．そこで，初めに，S25R でのエンドユーザ回線の判別と，公開リスト B での判別についてどの程度違いがあるかについて調べた結果を報告する.

図 1 には，接続があったサーバのうち，公開リスト A に該当しないサーバのホスト名について，公開リスト B に登録されているサーバからの宛先数，S25R に該当するサーバから宛先数について分類して集計を行った．公開リスト B 及び S25R の両方に該当する宛先数を×印で示した．両方に該当しない宛先数は□印で，公開リスト B についてのみ該当するホストからの宛先数は＋印で，S25R のルールのみに該当するホストからの接続数は＊印でそれぞれ示している．どの接続も日変動があるが，特異な傾向があるのは週末のみである.

図 1 の期間中にあった宛先を全て集計すると，表 1 のようになる．両者ともに該当する接続と，両者ともに該当しない接続を合わせると 74.3%となり，両者は，ほぼ共通の判定をすると考えてよい．しかし，S25R にのみ該当する宛先が 13.4%である一方，公開リスト B にのみ登録されている宛先は 2.1%である．この結果から，S25R 方式の方が多くのスパムメールを拒否できるとも考えられるが，論文 [6] にもあるように，S25R 方式の場合には，S25R に合致する宛先の 17%程度が正常なメールを出すサーバと報告されている．このため，S25R 方

BAC24ca.bac.pppool.de
Broadband-Dynamic-Central1436.connect.com.fj
D128.D-IP01.lipetsk.ru
La45c.l.pppool.de
VPN-148.PPTP-197-SA.GlobalNET.ba
aaeo30.neoplus.adsl.tpnet.pl
public8651.xdsl.centertel.pl
user-0ccemdu.cable.mindspring.com

表- 2: 公開リスト B に登録されているが，S25R には該当しないサーバのホスト名の例.

式では，ホワイトリストの整備が必須であると思われる．今回の我々の方式では，公開リスト B を利用することにより，サーバ管理者側でホワイトリストのメンテナンスをしなくとも偽陽性判定のない運用が出来る可能性があるため，検出率は落ちると思われるが，こちらを採用することにした.

一方で，公開リスト B に登録されているが S25R 方式には該当しない宛先も 2.1%ほどであるが存在する．我々の方法では，このような宛先でも検知するので，ここからのスパムメールは応答遅延により拒否できる可能性がある．これに該当する宛先数は期間中 5000 程度であった．主なものを参考までに表 2 に示す.

3　運用結果

図 2 に新潟大学メールゲートウェイにて提案方式の実装を行った結果を示す．横軸が経過日数であるが，45 日目 (2008 年 8 月 5 日) に提案方式の判定条件 2,3 について導入を行った．縦軸は，学外から受信したメールの宛先数に対する割合を示す．なお，応答遅延により接続が切れる場合には，1 接続あたり 1 宛先としてカウントしている．図 3 には，導入前，導入後 45 日間にメールゲートウェイに届いた宛先総数ついて，その処理方法別に分類し，割合で示した.

図 2 から，接続の処理について，日変動があることが分かる．これは，7 日周期を持つような週末に特異な変動を示すものではない．スパム送信行為の周期などに依存していると推測している.

図 2 の＊印で公開リスト A で拒否しているメール宛先数を示す．判定条件 2,3 の導入の前後に依らず，全体のうち 50%程度がこの条件で拒否されていることが分かる．以前より効果は薄れつつあるものの，依然として判定条件 1 の対策も無視できない．図 2 の□印に宛先不明のメールを示す．これは対策後，上昇していることが分かる．図 3 に示す総数で見ると，導入前は 16.2%，導入後は 25.6%と 10%程度上昇している．しかし，日

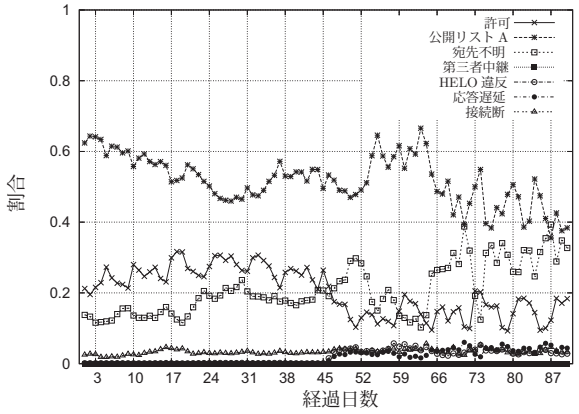

図- 2: 本方式によるスパム判定結果.

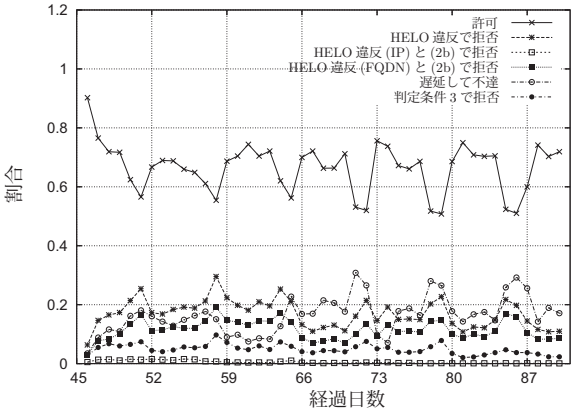

図- 4: 本方式の判定条件 2,3 によるスパム判定結果.

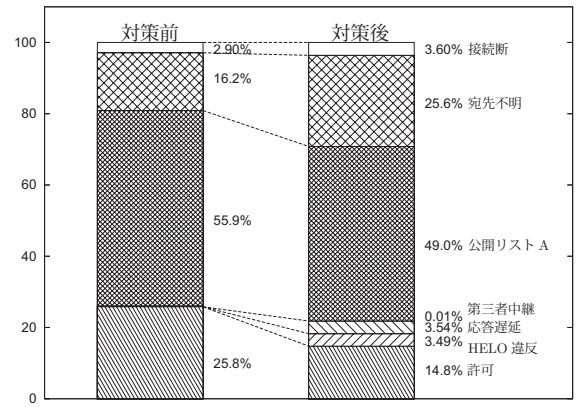

図- 3: 導入前，及び，導入後，45 日間で処理したメールのスパム判定結果の比較.

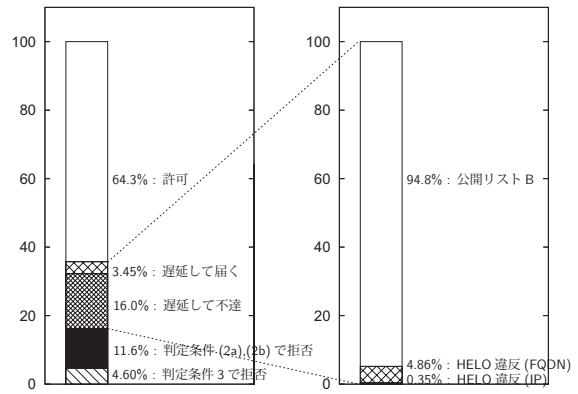

図- 5: 本方式の判定条件 2,3 による判定結果（左）．応答遅延を行った理由（右）.

変動が大きいため，本対策の影響と言うよりは，集計時期にたまたまユーザ不在のスパムを送信する行為が多く行われたものと考えている．図 2 の△印に，応答遅延とは関係なく，原因不明で送信サーバ側で接続を切られるケースについて，その接続数を示す．新対策導入の影響を特に受けることなく，常に 3%程度を占めていることが分かる．図 2 の■印には，第三者中継により拒否されるメールを示すが，若干であり，全体の 0.01%程度となっている．現在では第三者中継を利用したスパムメール配信はないと考えてよい．

図 2 の×印には，受信を許可したメールの宛先数を示す．これは，新対策導入後に落ち込んでいる傾向にあることが分かる．図 3 に示す総数では「許可」に対応し，25.8%から 14.8%に減少している．これは，新対策の導入の効果である．図 2 の○印には，判定条件 3，または，判定条件 (2a),(2b) の両方を満たしたため拒否となった宛先を示している．総数で見ると図 3 の「HELO 違反」に対応し，3.49%である．図 2 の●印には，(2a),(2b) のどちらか 1 つを満たしたため，応答遅延となり，接続が失われた数を示す．図 3 の「応答遅延」に対応し，3.54%となる．この二つが，新対策により拒否したメールの割合となる．

2008 年 8 月 5 日の導入から現在まで，学内から届くはずのメールが届かなかったという偽陽性判定の報告は一件もない．少なくとも新潟大学の環境では，現在まで公開リスト B の他にホワイトリストのメンテナンスをすることなく運用が出来ている．

今回，新たに対策を入れた，判定条件 2,3 の部分についての効果をみるため，図 4 では，許可した宛先数，遅延して不達となった宛先数，判定条件 (2a),(2b),3 により拒否した宛先数の和を 1 として，各宛先数の割合を示す．これら全てのメールは，新対策を入れる前には，全て許可となり，ユーザに届いていたはずのメールである．横軸は経過日数である．拒否理由に依らず，ほぼ 7 日周期での変動が見られるのは，休日か平日かの違いである．応答遅延により不達となったメールを○印で示すが，これについては，例外的に，70 日目を境にして多くなっている．特に設定変更等もなかったことから，送信されるスパムの性質により上昇していると考えている．

図 5 の左図は，対策導入後 45 日間での接続総数について，許可，遅延して許可，遅延して不達，拒否に分けて分類し集計した図である．

ユーザに届いたメールは，本方式による対策前と比

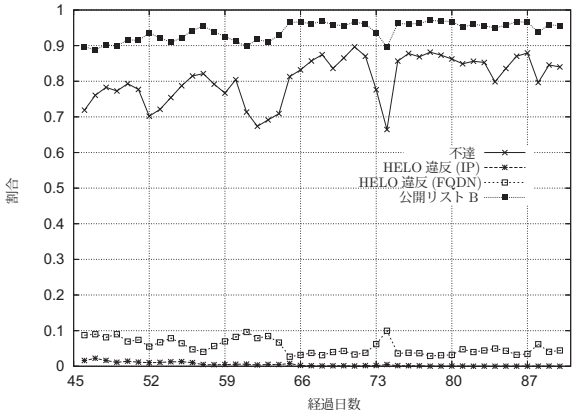

図- 6: 応答遅延を行った理由.

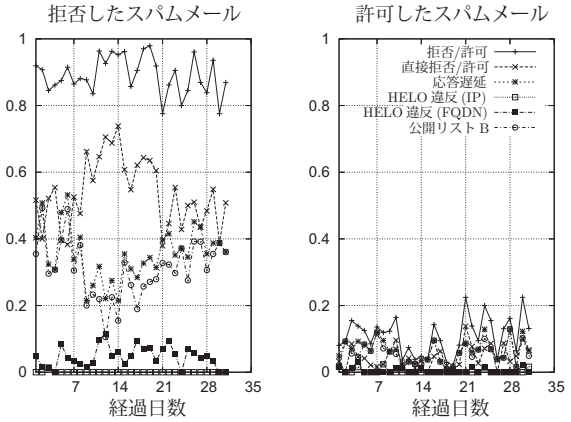

図- 7: 拒否したスパムメールとその理由（左），許可したスパムメールとその理由（右）.

べて 64.3%となった．残りの 35.7%が本方式により届かなくなったメールである．その原因を見ると応答遅延によって届かなかったメールが 16.0%，届いたメールが 3.45%である．また，判定条件 (2a) と (2b) の両方を満たすため拒否した接続が，11.6%，HELO コマンドの引数が学内ホストのため拒否した接続が 4.60%である．

判定条件 (2a),(2b) の両方を満たしたため拒否した宛先 (11.6%) は，判定条件 (2a),(2b) の一方に該当し，遅延した宛先 (19.5%) の半数程度なので，(2a),(2b) のどちらかに該当した場合，単純に応答遅延する方式と比較すると，応答遅延による接続数は 2/3 に減少することが分かる．

図 6 には，応答遅延を行ったメールの宛先を，原因別に応答遅延を行った宛先数に対する割合で示している．横軸には日数を表す．図 5 の右図は，45 日間に届いたメールの宛先を，応答遅延を行った原因別に分類して，割合で示している．

応答遅延を行った理由は，判定条件 (2a),(2b) のいずれか一つを満たす場合である．ただし，判定条件 (2a) については，HELO 引数が IP アドレス（IP リテラルではない）の場合（＊）と，それ以外で，FQDN でない場合（□）の 2 つに分類した．

図 6 の■は公開リスト B に登録があるサーバからの宛先数を示しており，図 5 の右図に示した総数で見ると 94.8%である．つまり，大半の遅延した接続は，エンドユーザ回線からの接続と考えられる．残りは，HELO コマンドの引数が不正な接続だが，そのうち HELO の引数が IP アドレスの接続は，総数では 0.35%とほとんどない．大部分は，FQDN でない引数のサーバからの接続で，総数では 4.86%となる．

この様に，応答遅延は，ほとんどが判定条件 (2a) にある公開リスト B に該当するため発生する．判定条件 (2b) の HELO コマンドの不正による遅延はほとんどないため，判定条件 (2a),(2b) の両方を満たすことを確認することなく (2b) の HELO コマンドの不正のみで削除する方が簡便とも考えられる．しかしながら，運用中に，(2b) の条件にのみ該当する正常なメールを送るサーバが報告された．このため，(2b) のみの条件で拒否は出来ない．

なお，応答遅延を行ったメールのうち，遅延中に接続が切れてしまい，結果としてメールを拒否した接続数は，図 6 の×で示している．総数で見る場合には，図 5 の右図から，遅延して届いたメールが 3.45%，遅延して不達となったメールが 16.0%であるから，応答遅延したメールのうち 82.3%のメールが不達となっていることが分かる．

メールがスパムメールかどうかは，実際には，人間がメールの内容を見ないととわからない．このため，情報基盤センターの特定のメールアドレスに届いたメールについて，人間がスパム判定を行ない，集計を行った．このメールアドレスに届いたメールについて，人間が内容を見て，スパムか否かを判定する．届かなかったメールについては，内容が分からないため判定が出来ないが，ログから得られた差出人の情報を用いて迷惑メールか否かを判断し，集計を行った．実際には，届かなかったメールは全てスパムメールだったと判定された．つまり，このアドレスに届いたメールについては，偽陽性判定はなかったと言える．なお，判定条件 1 で拒否したため届かなかったメールは，今回の集計から除いている．今回この集計に利用したメールアドレスには，判定条件 1 で拒否したメールを除いても，迷惑メールが一日あたり 50 から 100 通程度着信している．集計結果を図 7 に示す．

図 7 の左図は，人間がスパムと判定したメールのうち，届かなかったスパムメールを理由別に集計した図である．右図は，同様に，人間がスパムと判定したメールのうち，届いてしまったメールを理由別に集計した図である．ただし，人間がスパムと判定した全メールの数を 1 として，割合で示している．

人間がスパムと判定したメールのうち，今回の対策で拒否出来ているメールを図 7 左図の＋印で示す．これは約 9 割程度となり．効率よくスパムメールを拒否出来ていると言える．図 7 左図の×印には，判定条件 (2a),(2b)，または，3 により拒否したメールを示す．対策前に届いていたスパムメールのうち 5 割程度は，応答遅延することなく，直接拒否できていることが分かる．図 7 左図の＊印に応答遅延の結果，受信しなかった接続数を示す．これが，残りの 4 割程度となっている．応答遅延を起こした原因の一つとして，図 7 左図○印に公開リスト B に登録のあった IP からの接続を示す．これは，遅延して受信できなかった接続数とほぼ同じであることが分かる．図 7 左図の■印には，HELO コマンドの引数が FQDN でない接続を示すが，1 割程度であり，応答遅延を起こす原因としては少ない．さらに，図 7 左図の■印には，HELO コマンドの引数が IP アドレスの接続を示すが，これはほとんどない．全体集計（図 6，5 右図）で見たときと同様の傾向である．

一方，右図には，今回の対策で阻止出来なかったスパムメールを＋印で示す．対策前に届いていたスパムメール全体の 1 割程度が阻止できていないことがわかる．図 7 右図＊印には，応答遅延にもかからず届いたスパムメールを示す．届いたスパムメールの半分程度であることが分かる．図 7 右図○印に，公開リスト B に該当し，遅延したメールを，＊印に，遅延したメール全体を，それぞれ示す．両者を比べると，遅延した接続のほぼ全てが公開リスト B に該当したことが原因であることが分かる．残りのスパムメールは，残念ながら，全ての判定条件にかかることなく，ユーザに届いてしまっている．対策前に届いていたスパムメールの 5%程度となる．

なお，今回のデータには示さなかったが，正常メールのうち，応答遅延にかかって到着したメールはなかった．この結果からすると，応答遅延を行うことなく，拒否する方が適切と思われるが，この統計に利用したメールアドレスとは別のメールアドレスで，(2b) の条件に該当し遅延して届いた正常メールが報告された．このため，条件 (2b) で拒否した場合には，偽陽性判定が生じることとなる．なお，報告のあったメールの送信元は，メーリングリスト配送業者のサーバと思われる．

以上のように，今回の対策を導入した結果，特定のメールアドレスに対しては，今まで届いていたスパムメールの 9 割近くを排除する結果となった．しかしながら，これは特定のメールアドレスでの場合での結果であるため一般的ではない可能性があることに注意して欲しい．また，図 7 に見るように，日変動も大きいこともあり，実感としてはスパムメールが多いと感じる場合もある．

4 結論と今後の課題

今回の提案方式を新潟大学メールゲートウェイに実装することで，スパム対策を行った．今回の対策後，受信メールを 3 割ほど削減することに成功した．一方で，導入から現在まで，偽陽性判定が疑われるケースについては報告がない．正常メールの中でも，応答遅延の条件に該当し，メールは遅延したもの，接続が切れず受信していたケースが 1 件のみ報告された．

特定のメールアドレスについて，拒否できたスパムメールを調査したところ，対策前に比べて，9 割程度の迷惑メールを拒否しているという結果となった．また，HELO コマンドの引数による判定のみでは，偽陽性判定があることが報告されたが，公開リスト B と HELO コマンド引数を組み合わせ，「拒否」と「応答遅延」に分けて判定することで，偽陽性判定は無くなった．拒否条件の導入により応答遅延を起こす接続を約 2/3 に減らすことに成功し，サーバの負荷軽減が出来た．

以上のように，今回の方式では，サーバ管理者側で，常時，ホワイトリストのメンテナンスをしなくとも，新潟大学のメールゲートウェイとして運用可能であり，かつ，スパムメール対策として有効な設定となることが分かった．また， 本対策後，極端にスパムメールが減ったとの声も多数，よせられている．

しかしながら，スパムメールが届かなくなったわけではない．また，ボットプログラムが配送効率を多少犠牲にしても接続を待つようになり，応答遅延の方式に対応した場合には，効果が弱くなるなどの問題もある．今後は，普及が期待されている送信ドメイン認証 [8] の結果なども組み合わせるなどして，さらに有効なスパム対策を考える必要がある．

謝辞

今回の迷惑メール対策の試験運用，及び，メール集計に協力頂いた新潟大学情報基盤センター技術補佐員沢田浩氏に感謝致します．

参考文献

[1] C. Rossow: Anti-spam mearure of European ISPs/ESPs (online), available from <http://www.internet-sicherheit.de/ leadmin/docs/publikationen/anti-spam-measures-of-europian-isps-esps.pdf>

[2] 吉田 和幸: メールゲートウェイにおける spam 対策について, 学術情報処理研究, No 9 pp37-44 (2005)

[3] Post x SMTP Access Policy Delegation (online), available from <http://www.post x.org/SMTPD_POLICY_README.html>

[4] R. Ramachandran, D. Dagon and N. Feamster: Can DNS-Based Blacklists Keep Up with Bots ? (online), available from <http://www.ceas.cc/2006/14.pdf>

[5] M. Swimmer, I. Whalley, B. Leiba and N. Borenstein: Breaking Anti-Spam Systems with Parastic Spam (online), available from <http://www.ceas.cc/2006/9.pdf>

[6] H. Asami: Study Report of an Anti-spam System with a 88% Block Rate – The Selective SMTP Rejection (S25R) System – (online), available from <http://www.gabacho-net.jp/en//anti-spam/paper.html>

[7] 川田良文，山田一成，田島尚徳，拓殖明: 全学メールサービスにおける迷惑メール・ウイルスメール対策, 名古屋大学情報連携基盤センターニュース Vol7, No 3 pp290-294 (2008)

[8] 本間輝彰: JEAG テクニカルアップデート ～迷惑メールの現状と技術的課題等について～, 第 5 回迷惑メール対策カンファレンス