Vol.2010-IOT-11 No.7
2010/10/15

Greylisting に基づく迷惑メール対策の運用評価

松原義継[†1]　大谷誠[†1]
日永田泰啓[†1]　只木進一[†1]

電子メールサービス運用において，迷惑メール対策は避けて通れないものとなっている．より良い迷惑メール対策を行うために，運用中の迷惑メール対策の効果を評価することは大事である．今回，greylisting 方式による迷惑メール対策効果を 3 年間分のデータを基に分析する．利用者が受信した電子メールの中における迷惑メールの可能性の高い電子メールの割合は減少傾向となった．

Evaluation of Anti-spam Practical Use Using by Greylisting

Yoshitsugu Matsubara ,[†1] Makoto Otani ,[†1]
Yasuhiro Hieida [†1] and Shin-ichi Tadaki [†1]

Anti-spam filtering mechanisms are mandatory for providing e-mail services. We have employed greylisting filters in our e-mail services for 3 years. The effects are evaluated. We found that the number of e-mails which are potentially spam ones had been decreasing.

1. はじめに

インターネットは，現在の情報通信における不可欠な基盤の 1 つである．その上で提供されるサービスの中でも電子メールは基本的なコミュニケーションツールの 1 つである．一方，増え続ける迷惑メールは，コミュニケーションの質の低下を招いている．そのため，電子メールサービスの基本要素として，迷惑メール対策は必須となっている．

†1 佐賀大学 総合情報基盤センター
Computer and Network Center, Saga University

佐賀大学総合情報基盤センター (以下,「**本センター**」で表記) は，2004 年秋から greylisting 方式[1] による迷惑メール対策を運用している．これは，迷惑メールを学内に入れないという方針に基づくものである．2010 年 9 月には，教職員用電子メールサービスを外注化し，これまでのサービスを停止する．また，一月遅れで，学生用電子メールサービス及び学内部局向けの迷惑メール対策及びウィルス付電子メール対策も外注の予定である．

本稿では，これまでの迷惑メール対策の効果を評価し，今後の対策の質の向上に資することを目的とする．

2. 運用環境

佐賀大学は，2003 年に旧佐賀医科大学と旧佐賀大学の統合により 5 学部構成となった総合大学である．現在の教職員数は約 2500 名，学生数は約 7300 名である．旧佐賀大学では，2000 年より，全学生及び全教職員にアカウントを付与し，電子メールサービスを運用してきた．

本センターでの greylisting 方式は，sendmail[2] 用である milter-greylist[3] に独自のホワイトリスト自動管理機能を加えたものである．迷惑メール対策の開始前に，様々な方法を検討した結果として，本センターの迷惑メール対策は greylisting 方式のみに落ち着いた．他の組織では，greylisting 方式と他の対策との同時運用や，greylisting の運用時間に工夫，等がある[4],[5]．迷惑メール対策サーバ機器は，Sun Microsystems 社製 SunFire V120 である．運用中の障害発生を考慮して，機器は 2 台構成にしている．佐賀大学外から送信される電子メールは，一部の学内部局を除いて，これら迷惑メール対策サーバ群を経由して学内に送信される．

2.1 greylisting 方式

greylisting 方式は,「迷惑メール送信元は再送要求に応じない」という仮定に基づいた迷惑メール対策である．初めて電子メールが送信されて来た場合，greylisting がインストールされている迷惑メール対策サーバは，これを受信せずに送信元に再送要求を行う．この際，greylisting の管理用データベースは，エンベロープの送信者メールアドレス (以下,「**送信者メールアドレス**」で表記), 送信元 IP アドレス, エンベロープの受信者メールアドレス (以下,「**受信者メールアドレス**」で表記) の三つ組をデータベースに格納し，この三つ組を再送待ち状態として扱う．再送元が再送した場合，この迷惑メール対策サーバはその電子メールを受信する．Greylisting の管理用データベースは，この再送して受信した電子メールの三つ組を有効期限付きのホワイトリストとして扱う (以下，この有効期限付きのホワイトリス

トを「**動的ホワイトリスト**」で表記)．受信後の有効期限までの期間は，動的ホワイトリストと同じ三つ組を持つ電子メールは，再送を要求されずに受信される．

迷惑メール送信元の中には，再送してくるものもある．再送されてくる迷惑メールを全て防ぐことはできないが，ある部分を防ぐことは可能である．迷惑メール送信元の中には，急いで再送してくるものがある．そこで，再送までの時間が短すぎるものを拒否することにより，このような迷惑メールの流入を防いでいる．

本センターでの greylisting 設定では，再送待ち時間は 3 分間，受信後の再送なしの受信期間は 1 週間である．迷惑メール対策サーバは 2 台あることから，各サーバの管理用データベースには同期設定を施され，どちらの迷惑メール対策サーバで受信しても処理は同じにしている．

2.1.1 静的ホワイトリスト管理

greylisting 方式は，再送要求発行により送信遅延を引き起こす．再送間隔は，RFC5321 にて一定の目安を定めているが[6]，実際の再送間隔は送信元の運用方針により様々である．送信元によっては，再送間隔は数時間であったり，最悪の場合は再送しない場合もある．

送信元が再送しない場合，greylisting 方式は，電子メールの不達を発生させてしまう．そこで，常時再送なしに受信するホワイトリストを用意することで，不達メール発生を抑制している (以下，このホワイトリストを「**静的ホワイトリスト**」で表記)．例えば，学内の利用者からの希望に基づき，その利用者のメールアドレスを静的ホワイトリストに登録する．これにより，その利用者宛の電子メールは常時受信可能となる．

さらに，業務・研究・教育上の配慮から，再送を行わない組織の送信元を静的ホワイトリストに登録する場合もある．この場合，電子メールの送信元ホストもしくは電子メールアドレスのドメイン名を静的ホワイトリストに登録する．

更に，送信遅延の発生を抑制するため，本センターでは，一定のルールに基づき，動的ホワイトリストの登録状況を基に静的ホワイトリストを自動更新する仕組みを導入している[7]．始めに，静的ホワイトリストの管理者は静的ホワイトリストの候補となる組織のドメイン名を設定する．その組織内送信元ホストの動的ホワイトリストへの登録頻度を集計するプログラムを定期的に作動させる．送信元ホストの登録頻度が一定以上になった場合，その送信元ホストは静的ホワイトリストに自動登録される．送信元の運用方針変更に伴う送信元ホスト変更が生じた場合，そのホストからの送信はなくなることを考慮し，静的ホワイトリストに移った後もその送信元ホストからの送信頻度を調べる．その送信元ホストからの送信頻度が一定以下になった場合，その送信元ホストは静的ホワイトリストから解除される．

2.2 調 査 方 法

今回の分析の基となっているのは，迷惑メール対策サーバ群内に記録されている 2007 年 4 月 1 日 0:00〜2010 年 5 月 31 日 23:59 間の syslog ファイルである．これら syslog ファイルから，一日毎の送信リクエスト回数を集計した．集計に際して，sendmail サービスへの接続のみで電子メールを送信しなかった分は除外している．全送信リクエスト回数は 4939 万 1113 回となった．

集計データから，送信者メールアドレス, 送信元 IP アドレス, 受信者メールアドレス，の三つ組を作る．これらの三つ組を，その送信ステータスコードに応じて, 再送要求分と受信分に二分する．

Greylisting の動作原理上，静的・動的ホワイトリストにより受信した以外のメールには，同一メールの再送要求と受信がある．このことから，これら二分した各データは，両者に共通する三つ組 A といずれか一方にのみ存在する三つ組 B に分けることができる．このことから，再送要求分と受信分に二分したデータは，さらに以下の 4 項目に分類できる．これら 4 項目の和は，迷惑メール対策サーバ群の全送信リクエスト回数となる．

(1) 再送要求分
- (1-A) 再送後に受信した電子メールの再送回数．ここに分類されるデータは，再送要求が行われた三つ組のうち，A の部分である．例えば，ある電子メールが再送要求を受け，再送後に受信した場合，ここに分類される．
- (1-B) 同一日の内に受信しなかった送信リクエスト回数．ここに分類されるデータは，再送要求が行われた三つ組のうち，B の部分である．つまり，再送要求を行ったが，同一日の内に受信しなかった電子メールが分類される．ここに分類される電子メールは，迷惑メールの可能性が高いものである．

(2) 受信分
- (2-A) 少なくとも 1 回は再送されて受信した送信リクエスト回数．ここに分類されるデータは，受信された三つ組のうち，A の部分である．つまり，再送要求に応じて再送されてきたもので，迷惑メールではない，普通の電子メールの可能性が高いものがここに分類される．
- (2-B) 再送なしで受信した送信リクエスト回数．ここに分類されるデータは，受信された三つ組のうち，B の部分である．ある電子メールが静的ホワイトリスト及び動的ホワイトリストに登録されている場合，その電子メールはここに分類される．

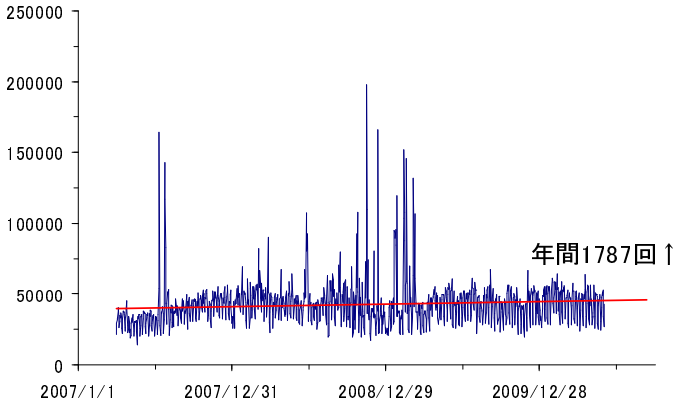

図 1 全送信リクエスト回数．横軸は年月，縦軸はリクエスト回数．集計単位は 1 日．青線は送信リクエスト回数．赤線は最小二乗法による直線．

Fig. 1 Time series of sending requests. The horizontal line is date. The vertical line is number of sending requests. The accumulation unit is one day. The blue line is the sending requests. The red line is made by least square method.

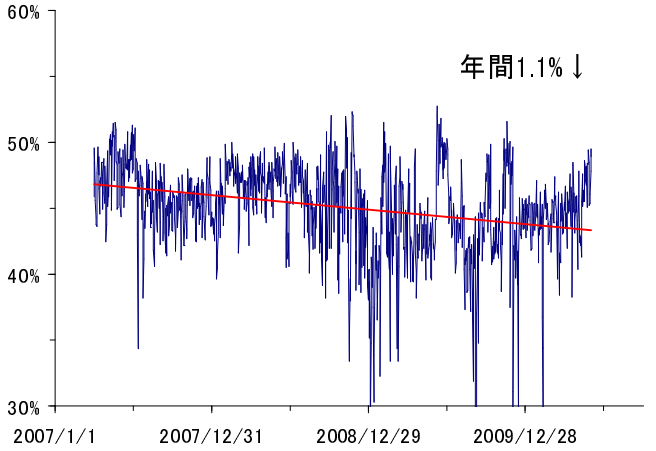

図 2 送信リクエスト回数全体における実需の割合．平均値は約 45.08 パーセント．傾向は，年間 1.1 パーセントの減少．集計単位は 1 日．

Fig. 2 Time series of rate of actual demands. The average is about 45.08 percentages. The trend is 1.1 percentages decreasing per year. The accumulation unit is one day.

3. 時系列データ分析

調査期間における全送信リクエスト回数の時系列を図 1 に示す．図 1 からは，数カ所，大量送信と思われる記録を散見できる．時系列全体の傾向を知るため，最小二乗法による直線を引いたところ，送信リクエスト回数は年間 1787 回の割合で増加していることが分かる．

もし，greylisting のない場合，再送要求は発生しなくなり，再送後の受信はなくなる．すると，受信までの再送回数である分類項目 (1-A) は存在しなくなる．同時に，再送要求後，同一日の内に受信しなかった分類項目 (1-B) の電子メールは，受信するようになる．つまり，全送信リクエスト回数から分類項目 (1-A) を除いた分は，greylisting がなかった場合の送信リクエスト回数と考えることができる．これを，「**実需**」と呼ぶことにする．全送信リクエスト回数に対する実需の割合の時系列を図 2 に示す．実需の割合は平均約 45.08 パーセントであり，全リクエスト回数の半数弱であることが分かる．また，実需の割合は年間 1.1 パーセント減少傾向している．つまり，再送回数が増加している．

先に示した 4 項目の，全送信リクエスト回数に対するそれぞれの割合の時間変化を図 3 に示す．

分類項目 (1-A) は，再送後に受信した電子メールの再送回数であり，全送信リクエスト回数から実需分を差し引いたものである．この分類項目の全送信リクエスト回数における割合の平均値は約 54.92 パーセントである．つまり，再送回数は実需を上回ることが分かる．実需の場合とは逆に，再送回数分は 1 年間に約 1.1 パーセントの上昇傾向を示している．

分類項目 (1-B) は，同一日の内に受信しなかった分であり，迷惑メールの可能性は高いものである．この分類項目の全送信リクエスト回数における割合の平均値は約 18.04 パーセント，実需の約 40 パーセントである．分類項目 (1-B) は，1 年間に約 0.07 パーセントの減少傾向を示している．

分類項目 (2-A) の全送信リクエスト回数における割合の平均値は約 22.55 パーセントである．分類項目 (2-A) は，1 年間に約 0.07 パーセントの減少傾向を示している．

分類項目 (2-B) の全送信リクエスト回数に占める割合の平均値は約 4.49 パーセントである．運用経験上，静的ホワイトリストもしくは動的ホワイトリストに登録されている所からの迷惑メール送信を確認しており，運用上の問題点を浮き彫りにしている．全体傾向は 1

Vol.2010-IOT-11 No.7
2010/10/15

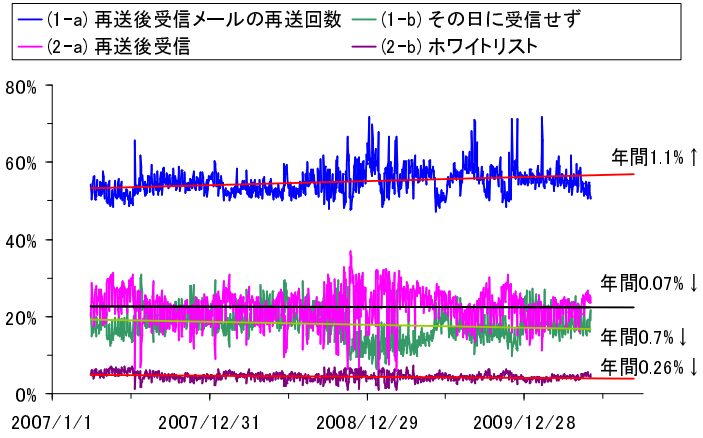

図 3 各項目の全送信リクエスト回数に占める割合．集計単位は 1 日．
Fig. 3 Time series of rates per items. The accumulation unit is one day.

年間に約 0.7 パーセントの減少傾向を示している．再送もしくは静的・動的ホワイトリストにより受信した電子メール数は下降している．

4. 考　　察

静的ホワイトリスト管理において，greylisting を解除した利用者数は 17 名である．この人数は，全利用者の 1 パーセントに満たない．再送状況が悪く，必要があるため，静的ホワイトリストに手動登録した送信元は 11 個である．その内訳は，学術系サイト 5，政府系サイト 1，送信者メールアドレス 3，送信者メールアドレスのドメイン名 2 である．

電子メールで頻繁に連絡する利用者にとって，送信遅延が利用者に与えるストレスは大きいと思われる．そのため，著者らは，greylisting 方式で運用するに際し，greylisting 解除を希望する利用者の数及び手動で静的ホワイトリストに登録する送信元の数は多いと予想していた．しかし，上述のように，非常に少数の対応で運用を行うことができた．

図 3 より，分類項目 (1-A) の再送回数は，実需の場合とは逆に上昇傾向を示している．これは受信までに何回も再送する電子メールが上昇しているからだと思われる．

本センターで採用している greylisting そのものは迷惑メール判定は行わないが，今回集計したデータを基に greylisting を通り抜けた迷惑メール数を推測することは可能である．

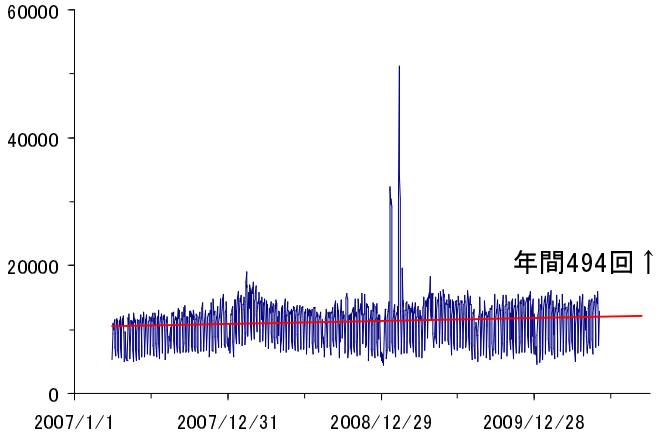

図 4 迷惑メール対策サーバ群が受信した電子メール数．集計単位は 1 日．赤線は最小二乗法による直線．
Fig. 4 Time series of received mails by re-sending or whitelist. The accumulation unit is one day. The red line is made by least square method.

推測のために注目するのは，DNS 未登録の送信元ホスト及び動的 IP アドレスと思われる送信元ホストである．これらは，組織等の電子メールサーバを経由せずに送信されていると考えられる．つまり，迷惑メールの可能性の高いものである．

再送もしくは静的・動的ホワイトリストにより受信した電子メール数は分類項目 (2-A) 及び (2-B) の和であり，これらをここでは「**利用者受信電子メール**」と呼ぶことにする．利用者受信電子メールの時系列を図 4 に示す．時系列全体の傾向を知るため，最小二乗法による直線を引いたところ，利用者受信電子メール数は年間 494 回の割合で増加していることが分かる．

DNS 未登録の送信元ホスト及び動的 IP アドレスと思われる送信元ホストからの送信リクエスト回数は，利用者受信電子メールの中から抽出した．抽出した電子メールの利用者受信電子メールに対する割合を図 5 に示す．抽出した電子メールの利用者受信電子メールに対する割合の平均値は 18.06 パーセントである．抽出した電子メールの全体傾向は，年間約 1.83 パーセントの減少傾向を示している．

Vol.2010-IOT-11 No.7
2010/10/15

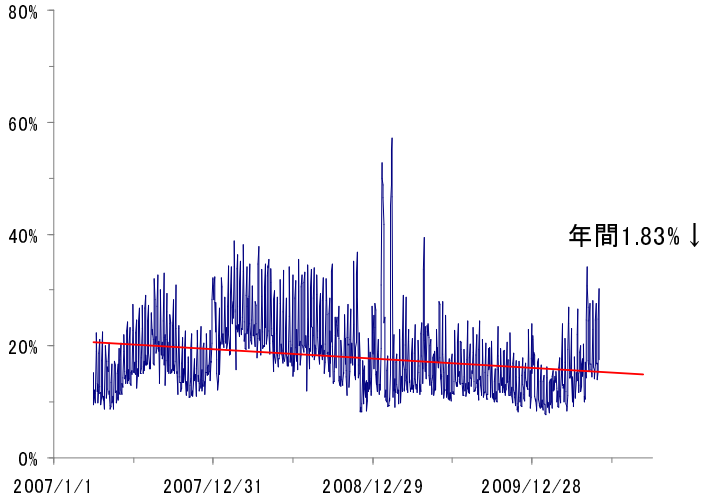

図 5 再送もしくは静的・動的ホワイトリストにより受信した迷惑メールの可能性の高い電子メールの全受信に占める割合．全体の平均値は，約 18.06 パーセント．集計単位は 1 日．

Fig. 5 Time series of rate of mails which are spam possibilities in all received mails. The accumulation unit is one day. The average is about 18.06 percents. The trend is 1.83 percentages decreasing per year.

利用者受信電子メールから抽出した迷惑メールの可能性の高い電子メール数と同一日の内に受信しなかった電子メールである分類項目 (1-B) の和は，佐賀大学宛に送信されて来た全電子メールの中で迷惑メールの可能性の高いものである．この迷惑メールの可能性の高い電子メール数の実需に対する割合の時系列を図 6 に示す．この割合の全体平均値は約 50.87 パーセントである．実需の約半数が迷惑メールと疑われるものであることがわかる．ただし，年 1.46 パーセントで減少している．これが，迷惑メールそのものの減少を示しているか，greylisting を正常電子メールとして通過する迷惑メールが増加しているのかを判定するには，更に詳しい解析が必要である．

5. ま と め

本センターにて行っている greylisting 方式に基づく迷惑メール対策を，過去 3 年間分の記録を基に評価した．送信状態に基づき送信記録は 4 種類に分類できる．

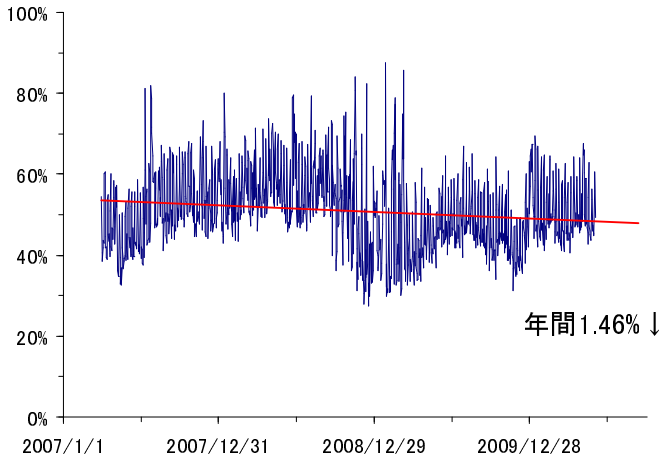

図 6 実需における迷惑メールの可能性の高い電子メールの割合．平均値は約 50.87 パーセント．傾向は，年間 1.46 パーセントの減少．集計単位は 1 日．

Fig. 6 Time series of rate of spam possibilities in actual demand. The accumulation unit is one day. The average is about 50.87 percents. The trend is 1.46 percentages decreasing per year.

全送信リクエスト回数の内，再送要求の割合は約 54.92 パーセントである．全体の約 18.04 パーセントは受信に至らず，少なくとも 1 回は再送して受信に至ったのは全体の約 22.55 パーセントである．残り 4.49 パーセントは，静的・動的ホワイトリストとして再送要求なしに受信した．

迷惑メール対策としての greylisting を行わなかった場合，再送要求は発生せず，同一日の内に受信しなかった電子メールは全て受信される．これを，本稿では実需と呼んだ．実需は，全送信リクエストの半数に当たる．一方，greylisting を用いない場合には，利用者に到達する電子メール数は 1.67 倍になる．増分は，迷惑メールである可能性が高い．

再送もしくは静的・動的ホワイトリストにより受信した迷惑メールの可能性の高い電子メールの数は，受信者が受信した電子メール数の約 18.06 パーセントであり，年間 1.83 パーセントの減少傾向である．同一日の内に受信しなかった電子メールと併せて見ると，佐賀大学宛に送信された迷惑メールの可能性の高い電子メールの数は，実需の約半数であり，実需

に対する割合は年間 1.46 パーセントの減少傾向にあることが分かった．実際に送信されている迷惑メール数はこれより多いと思われる．

静的ホワイトリストに登録されている組織からの迷惑メール送信も確認しており，このような組織の扱いは greylisting 方式の運用上問題となっている．迷惑メール対策上，このような組織は静的ホワイトリスト候補から除外すべきだが，業務・研究・教育上の理由で実際に除外することは難しい．

一部利用者については，本人希望に基づき，その本人宛電子メールを greylisting から外している．この場合，その本人は，送信遅延の発生しない代わりに，迷惑メールも含めてその本人宛電子メールは遅延なしに全て受信する．

以上より，greylisting 方式は再送する迷惑メールには無力という問題や送信遅延という問題を抱えているが，迷惑メールを学内に入れないという観点からは，一定の効果はあったと考えている．

今までの運用経験を基に，著者らは，外注先での迷惑メール対策を次のように考えている．利用者の望む迷惑メール対策は，迷惑メールのみを防ぐことであり，普通の電子メールの送信遅延もしくは送信拒否ではない．コンテンツフィルタによって内容に応じた対応が望ましい．しかし，あらかじめ迷惑メールを送信してくる可能性の高い送信元への対応は必要である．コンテンツフィルタの前段階として，例えば，taRgrey[8] のような迷惑メール送信元と思われるところのみの応答遅延と greylisting が必要であろう．送信拒否は，電子メール不達の危険性があるため，望ましくないと考えている．

電子メールリーダのコンテンツフィルタ機能が充実していることから，迷惑メール対策サーバ側と電子メールリーダ側の協調の可能性も高まっている．迷惑メール対策サーバ側での迷惑メール隔離の方法と共に，迷惑メール対策サーバ側では迷惑メールフラグを付ける方法も検討している．

参 考 文 献

1) Greylisting.org Home Page: http://www.greylisting.org/.
2) sendmail Home Page: http://www.sendmail.org.
3) milter-greylist home page: http://hcpnet.free.fr/milter-greylist/.
4) 飯田隆義，松竹俊和，吉田和幸：迷惑メール対策用 whitelist を一元管理できるメールシステムとその運用について，情報処理学会研究報告 14，IOT (2010).
5) 石島　悌，平松初珠，林　治尚：適用時間限定型 greylisting を用いた迷惑メール対策における配送遅延の改善，情報処理学会論文誌，Vol.51, No.3, pp.989–997 (2010).
6) Klensin, J.: Simple Mail Transfer Protocol, RFC 5321 (Draft Standard) (2008).
7) 松原義継，只木進一：milter-greylist のための静的 whitelist 自動生成，信学技報 18, IA (2006).
8) taRgrey Home Page: http://k2net.hakuba.jp/targrey/.