2004－DSM－34　(11)
2004／7／30

MTAによるspam対策の実践報告

鈴木常彦[†1,†5] 後藤邦夫[†2,†5]
山口榮作[†3,†5] 石川雅彦[†4,†5]

NPO 法人 東海インターネット協議会 (TIC) において半年続けて来た spam 対策の手法と効果を報告する．本手法では，従来の spam 対策のように受信して振り分けるのではなく，spam は受信しない方針で対策を行っている．これには，spammer のセッションの特異性に基づいてセッションを制御する手法を用いており，tempfailing や throttling と呼ばれる手法を組み合わせている．現在，好ましくないメールの 100% 近くを遮断することに成功している[1)]．

Reports on an experimental anti-spamming MTA with SMTP session control

Tsunehiko Suzuki,[†1] Kunio Goto,[†2] Eisaku Yamaguchi[†3]
and Masahiko Ishikawa[†4]

This paper reports the implementation and evaluation of an anti-spamming MTA for Tokai Internetwork Council, which has been operational for about 6 months. The MTA refuses reception of a spam before receiving a message while most of the widely used anti-spam MTA's filter messages after having received a message. The proposed method is a kind of SMTP session control with so-called tempfailing and throttling based on the characteristics of SMTP sessions originated by spammers. Almost 100% of unwanted messages are successfully filtered out so far with this method[1)].

1. はじめに

長期間使用している個人，メイリングリスト，webmaster，その他の問い合わせ受付のために公開している電子メイルアドレスには毎日数 100 通もの迷惑メイル（以下，spam）が送信され，ネットワーク管理者の一日は spam に埋もれた重要なメッセージを発掘することから始まる．

1990 年代まで，ほとんどの spam は第三者からの中継を許す MTA，すなわちオープンリレーあるいは不正中継を許す MTA を経由して発信されていた．したがって，例えば受信 MTA で DNS などで提供されるオープンリレーのリスト (ORBL:Open Relay Black List) を参照し，オープンリレーからの電子メイルメッセージ受信を拒否することで効率よく不正中継されたメッセージを拒否することができた☆．

しかし，2000 年以後，エンドユーザのインターネット接続速度が向上し，PC から直接短時間で多数の spam を発信することが容易になり，現在では，不正中継による送信件数をはるかに上回っていると思われる．また，ウイルスに感染した PC からの spam☆☆は，spammer 向け，またはウイルスやワームに含まれる SMTP クライアントから送信される．これらの SMTP クライアントは，独自に MX あるいは A レコードを問い合わせて各ドメインに直接 SMTP で送信する機能を持っている．

現在普及している spam 対策には，メッセージの内容によるフィルタリングがある．電子メイルに含まれるウイルスは，PC で普及している既知ウイルスのパターン (signature) の照合と同じ処理を MTA に追加することで，排除できる．宣伝などの迷惑メイルについては，Subject: やメッセージ本文に含まれる単語を既知の迷惑メイル中のそれらの出現頻度から算出した

†1 中京大学 情報科学部 情報科学科
School of Computer and Cognitive Sciences ,Chukyo University
†2 南山大学 数理情報学部 情報通信学科
Department of Information and Telecommunication Engineering, Nanzan University
†3 愛知県立大学 情報処理教育センター
Center for Information Education, Aichi Prefectural University
†4 SRA
Software Research Associates,Inc.
†5 東海インターネット協議会 (TIC)
Tokai Internetwork Council

☆ 設定ミスなどで，ドメインが ORBL に登録されてしまうと，設定を直してもすぐにリストが更新されず，しばらくそのドメインからのメイルを受け取れないというデメリットや，ORBL のサービス自体の継続性が保障されていない等の潜在的問題がある．

☆☆ 本稿では，ウイルスをばら撒くためのメイルも spam として扱っている

確率と比較して判定する Bayesian フィルタが普及している[5]．Bayesian フィルタも MUA に組み込むか，POP や IMAP を利用したメッセージ取り込みにフィルタつき proxy を使用することで，一般利用者が設定しカスタマイズできる[6),7)]．配送経路や方法については制限する必要はないが，MTA で一旦メッセージを受信しなければならない．また，メールを処理すればするほど spam のデータベースは肥大化するため，フィルタにおけるデータベースの管理も必要となる．したがって，spam の件数，メッセージ長が非常に大きくなると MTA や MUA の資源の浪費が顕著になる．

資源浪費を避けるためには，メッセージを受け取る前に受信を拒否することが望ましい．この要求を満たす抜本的な対策は，発信者の身元証明とメイル配送ホストの限定である．前者は，SMTP にユーザ認証がない点を補うために，各ドメインが運営する MTA において利用者認証に基づく利用者の真正性に関する情報を追加することで実現でき，いくつかの方法が提案されている[10)~12)]．この対策では，SMTP の “MAIL FROM:” または，(メッセージ受信後であっても) メイルヘッダの “From:” が正しいか判定できる．後者は，“MAIL FROM:” の return path（発信者）に自分が利用する権利がないドメインを指定できない（すなわち嘘を指定できない）ようにすることである☆1．これについては，DNS に RMX(Reverse Mail eXchanger) 情報を追加する手法が提案されている[8)]．これらの対策は広く普及すれば大きな効果が期待できるが，対策手法がある程度統一され広く普及するまでには時間がかかる☆2．

そこで本研究では，抜本的対策ができない現状における best current practice として，spammer から発せられる SMTP セッションの特異性を利用したメッセージ受け取りの一時拒否 (tempfail)，永久拒否の機能を追加した MTA を qmail[2)] を利用して構成し，運用実験を通して，その手法の妥当性と性能を評価する．spammer のホストと SMTP セッション（メイルを直接発信するという意味で MTA と呼んでもよい）の主な特徴は，重要な順に

(1) 受信拒否されると再送しない[3),4)]
(2) DNS の逆引き情報が定義されていない
(3) 逆引き結果のドメイン名から動的割り当て IP アドレスか，個人利用者
(4) SMTP 応答を遅くするとあきらめる (throttling)
(5) 発信だけで SMTP 接続を受けない
(6) その他，しつこく再送するなど

となる．これらの情報の組合せで SMTP TCP 接続の拒否，SMTP セッション中で受信を一時拒否，受け取り許可を判定する．2004 年 2 月 7 日 0 時から 2004 年 7 月 9 日 24 時までの，154 日間☆3に受けた SMTP セッション数，SMTP Deny，永久拒否，一時拒否数の合計は，それぞれ 79476，23997(30%)，684(0.86%)，50280(63%) である．これまでに得た結果からは spam は，ほぼ 100% 排除でき，とても有効な対策と言える．

2. システムの構成

本節ではシステムの構成と spam 判定ルールを説明する．システムの構成を図 1 に示す．本システムの処理は qmail 以外の MTA ソフトウェアでも実現可能であるが，tcpserver，rblsmtpd から渡されるコマンド引数や環境変数を利用可能な MTA は現在 qmail だけである☆4．

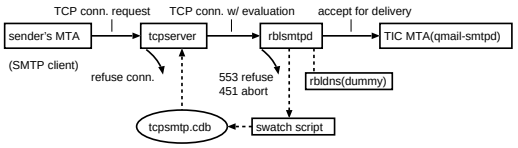

図 1 spam 対策 MTA の構成

2.1 処理の流れ

処理の流れは以下のとおりである．

(1) SMTP(25/tcp) への TCP 接続を tcpserver で受ける．SMTP クライアントの IP アドレスとパラノイド検査☆5した PTR レコードからルール (tcp.smtp.cdb) にしたがって，spam か判定する．spam 常習ホストと判定した場合は接続拒否 (TCP 接続を切る)．

(2) tcpserver は接続拒否しない場合，rblsmtpd を起動し，spam 判定結果をつけて SMTP 接続を渡す．

(3) rblsmtpd（SMTP サーバフロントエンド）は spam 判定結果が永久拒否（553），一時拒否（SMTP 応答コード 451）であればそれぞれの応答コードを返す．判定が受信可能であれば，qmail-smtpd（SMTP サーバ）を起動して，SMTP 接続を渡す．

(4) rblsmtpd の記録を swatch スクリプトで監視し，必要が生じたらルール (tcp.smtp.cdb) を動的に書き換える．

rblsmtpd は spam 判定結果を利用できるように少し改変を加えた．rbldns は rblsmtpd の本来の機能である Relay Blackhole List(RBL) 参照先 DNS の代替として空の情報を提供するダミー DNS サーバである．

☆1 例えば，他大学を訪問し，接続させてもらったノート PC の postfix で MAIL FROM: <goto@nanzan-u.ac.jp> で発信すると嘘つきと判定される．

☆2 ウイルス対策，spam 対策ソフトウェアビジネスの観点から，普及が積極的に推進されないという説もある．

☆3 運用実験は，2003 年 11 月に開始したが，前半のログを消失してしまっている．

☆4 tcpserver は inetd と同様のインターネット スーパサーバでアクセス制限と情報取得機能を強化したもの．tcpserver, rblsmtpd などは (http://cr.yp.to/ucspi-tcp.html) の ucspi-tcp に含まれる．

☆5 PTR レコードを正引きした結果の中に，元の IP アドレスが含まれているか確認すること

2.2 spam 判定ルール

以下の 1 から 6 のルールファイルを結合し，tcprules で tcp.smtpd.cdb を生成する．

(1) deny.smtp : 常習 spammer ホストのリスト（tcpserver で TCP 接続を拒否）
(2) black.smtp : くどく再送する spammer ホスト（永久拒否 553）
(3) dark.smtp : spammer の可能性大 （一時拒否 451）
(4) moderate.smtp : 再送があれば受け取るホストのリスト （一時リスト）
(5) white.smtp : ホワイトリスト （許可リスト，ML メンバなど）
(6) grey.smtp : 以上どれにもマッチしないホスト（451 で一時拒否）パラノイド検査に合格したら次回再送を受け取る．不合格は保留．

whitelist.smtp は手作業で管理する必要があるが，deny.smtp，black.smtp，dark.smtp はアクセス記録から自動生成することができる．grey.smtp と一時的な再送受け取りリスト moderate.smtp は自動生成する．

なお，一時拒否し受け取りを保留したセッションは，管理者に届く 2 種類のメールで目視検査を行うことができる．一つは一時拒否がある回数 (現在は 10,50,100 回,...) になると送信される警告メールで，Subject: が `(150):Paranoid_Check=Spammers_ISP(suspend);:212.123.84.93 mail-relay-3.tiscali.it` となっている．(150) は拒否回数である．もう一つは，crontab (一日 3 回朝昼) によりだされる以下のようなメールである．

```
Jul 10 00:28:42 217.6.22.65 helo: mail.stadt.exa
mple.de;
Jul 10 00:28:52 217.6.22.65 mail FROM:<postmaste
r@mail.stadt.example.de>;
Jul 10 00:29:03 217.6.22.65 rcpt TO:<info@ tokai-
ic.or.jp>;
--
Jul 10 01:45:19 67.15.70.72 helo: dm1.example.com;
Jul 10 01:45:29 67.15.70.72 mail FROM:<admin@dm1.
example.com>;
Jul 10 01:45:40 67.15.70.72 rcpt TO:<webmaster@t
okai-ic.or.jp>;
--
```

これは，rblsmtpd が拒否したセッションのログを “`grep -1 mail | tail -100`” したものである．管理者はこれらをみて，必要なものはホワイトリストにいれることになる．

2.3 再送の受け取り

応答コード 451 で一時拒否したホストから再送があった場合，1 回目の再送ですべて受け取っても 80% 以上の spam は排除できると思われるが，SMTP クライアント（送信 MTA）に関する付加的な情報を用いて細かく制御する実験を行っている．使用する情報は以下の通りである．

(1) DNS 逆引き結果: PTR なし，数字 4 桁以上を含む動的割り当てとみなせる名前，JP ドメインか．
(2) SMTP コールバック: SMTP サービスの有無，ある場合のバナーに JP が含まれるか．
(3) IP アドレスの WHOIS 情報: ISP の所属国，会社名などの情報

当初，ここで JP からのアクセスは比較的信用できると仮定したが，この仮定が正しくないことが判明し，2004 年 6 月 23 日から JP ドメインの優遇措置を廃止した．

3. 実験結果

2004 年 2 月 7 日 0 時から 2004 年 7 月 9 日 24 時までの 154 日間のデータを分析した結果を，表 1 から表 6 に示す．

3.1 96% の接続元を spam として拒否

表 1 に示したように，この間の全セッション数 79476 のうち，SMTP を tcpserver で接続拒否したもの (deny) が 30%，rblsmtpd で拒否したもの (blocked) が 64% である．これらを合わせると，94% の SMTP セッションを拒否したことになる．このなかには 2 回目以降でセッションを受けたものも含まれる．blocked から received を引いたものを全セッション数で割ると，spam として拒否したセッションの割合は 86% となる．さらに表 2 に示すように，重複を省いた IP アドレスの数でみると，96% の接続元を spam として拒否したことになる．なお拒否の内訳を表 3 に示す．

3.2 一時拒否の効果

表 4 に示すように，rblsmtpd で一時拒否を返したのち再送してこない相手は 86% にのぼる．これは一時拒否が spam に対して極めて有効であることを示すものである．

3.3 パラノイド検査

表 5 に示すようにパラノイド検査に合格しないものが 16% ある．これらの大半は spam であるが，受け取りたい相手の一部に DNS を正しく設定していないケースが 10 サイト程度あった．これらは手動でホワイトリストに登録せざるをえなかった．

3.4 helo の傾向

表 6 に示すように，SMTP セッションで spam が送信する helo として，受信側の IP アドレスが使われるケースが 25% もあることがわかった．理由は不明であるが，spam の判断に有効であることがわかる．

表 1 SMTP セッションログ内訳 (1)

result	session	rate
deny	23997	30%
ok	55479	70%
blocked	50964	64%
received	4515	6%
All	79476	

4. おわりに

これまでに得た結果からは本報告で述べた方法によ

表 2　SMTP セッションログ内訳 (2)

result	IP	rate
deny	2898	
ok	15365	
blocked	14865	
received	500	4%
All	17707	

表 3　rblsmtpd による拒否内訳

result	session	IP
(553)	684	212
Black	684	212
(451)	50280	14653
Dark	8131	3351
Grey	36173	6997
Pass	5976	5055
All	50964	14865

表 4　一時拒否の効果

behaviour	IP	rate
repeating	2423	16%
only once†	12942	87%
(553)‡	212	1%
† − ‡	12730	86%
All	14865	

表 5　バラノイド検査不合格率

result	IP	rate
paranoid OK	2691	53%
paranoid NG	2420	47%
All	5111	

り，spam はほぼ 100% 排除できるので，現状では有効な対策と言える．メッセージ内容を検査する方法と比較して MTA の負荷が低いことがこの手法の大きな利点である．

greylisting[4] で提案されているように SMTP エンベロープの “MAIL FROM:”，“RCPT TO:” の情報も併用するとより詳しいセッション制御が可能となるが，本格的なデータベースを使用する必要が生じる点で本方式ほど手軽とは言えない．本運用実験結果からは，IP アドレスだけで初回送信かを判定しても十分な spam 判定性能が得られると言える．

しかし再送を促す方式に共通した若干の不具合もある．利用者からみると，先に送ったメッセージが届かないときに，次に同じホストから同じ宛先に送ったメッセージが再送とみなされ先に届く，すなわちメッセージ到着順が入れ換わる問題がある．

さらに，発信人を偽ったウイルスを含む迷惑メイルを受信した MTA から偽られた発信人の MX となる MTA に大量にバウンス メッセージが送信される問題，すなわち二次的な spam がある．現在，TIC のサーバでは自らの送信先を記録して，バウンスの送信元と照合するという実験を行っている．しかし，

(1)　TIC ドメインのユーザが TIC のサーバからメイルを送信するとは限らない

表 6　SMTP HELO の内訳

domain	num.	rate
210.199.2.54	699	25%
other	2114	75%
All	2813	

(2)　MX のあがっているサーバからバウンスが返ってくるとは限らない

などにより，十分な照合にはなっていない．

根本的な対策としては，spam はバウンスしてはいけないという原則を徹底し，他の MTA でも spam 対策を実施すること．また，無駄な，あるいは spam となるバウンスを抑制する合意ならびに実施の普及が望ましい．これについては，バウンスは必須であるとする RFC2821 の改定[9] も提案されている．

そして最大の問題は，ほとんどの MTA でこの対策が普及したら，spammer が使用する発信 MTA に再送機能が追加され，この対策が無意味になることである．しかし，再送は spammer に多大なコスト負担をかけさせることが期待できる．また，spammer の手口を分析し，常に一歩進んだ対策を実施するとともに，最終的には抜本的な対策の開発，普及が必要である．

しばらくの間は，sendmail，qmail，postfix，exim などの既存 MTA ソフトウェアに対策が簡単に組み込める方法を普及させるか，spam 対策専用 MTA ソフトウェアを開発し，既存の MTA の前段に置いて手軽に利用できるようにすることが有効であろう．

参　考　文　献

1) 東海インターネット協議会: MTA における spam 対策，http://www.tokai-ic.or.jp/spam (2003-2004).
2) http://www.qmail.org/.
3) http://spam.qmail.jp/onazimi/index.html.
4) http://projects.puremagic.com/greylisting/.
5) http://www.paulgraham.com/spam.html.
6) http://popfile.sourceforge.jp/.
7) http://bsfilter.org/.
8) Danisch, H., The RMX DNS RR and method for lightweight SMTP sender authorization, Internet Draft, draft-danisch-dns-rr-smtp-04.txt, (May 2004).
9) Zinn, BZ., Relaxing SMTP's "deliver or notify" rule, Internet Draft, draft-zinn-smtp-bounces-01.txt (April 2004).
10) Sender Policy Framework, http://spf.pobox.com/.
11) Sender Rewriting Scheme, http://spf.pobox.com/srs.html, http://www.libsrs2.org/.
12) Delany, M., Domain-based Email Authentication Using Public-Keys Advertised in the DNS (DomainKeys), Internet Draft, draft-delany-domainkeys-base-00.txt (May 2004).