

スパムメール発信源分析による TLD のガバナンス推定 Estimation of TLD governance in analyzing of spam mail source

竹下 峰弘[†]

Minehiro Takeshita

中平 勝子[†]

Katsuko T. Nakahira

三上 喜貴[†]

Yoshiki Mikami

1. はじめに

営利目的のメールを無差別に大量配信するスパムメールが社会的問題となっており、メール全体に対して約80%ものメールがスパムと判定されている^[1]。

スパムメールを防御するために多くの技術が開発されている。主な発信元詐称防止技術としては、受信側で IP アドレスから判定し、スパムメールをブロックするブラックリスト、送信時に認証するなど送信者側から認証情報を送ることによって受信者側のフィルタリングを助ける技術として、OP25B(Outbound Port 25 Blocking)^[2]、DNSSEC(Domain Name System Security Extensions)^[3]、DKIM(DomainKeys Identified Mail)^[4]などがある。

最近では、これら発信元詐称防止技術の導入が進み、Yahoo!は2010年11月16日に、googleは2011年1月7日にDKIMを導入したと発表された^[5]。また、DNSSECが2011年1月10日に「.net」「.com」に導入される^[3]など、スパム対策のために各プロバイダが発信元詐称防止技術の導入が進められている。

本稿では、最近におけるスパムメールの発信量の推移、及び発信元詐称防止技術導入による動向、スパムメール発信量の変化を調査した。また、トップレベルドメイン(TLD)別にみたスパムメールの比率などが各TLDの管理状態をあらわすものと考え、これを「TLDのガバナンス」ととらえて推定した。

2. 使用データと分析手法

2.1 使用データ

本稿では、スパムメール発信源分析をするため、本学教員のメールアカウントで受信したスパムメールのうち、2010年12月～1月、5月に送られたスパムメール約150万通をもとに、以下の5つの項目を利用して分析を行った。

- A) 送信メールサーバの TLD 名 (TLD_{ms}、不明な場合は server unknown とする)
- B) TLD_{ms} の IP アドレス (IP_{ad})
- C) メールヘッダの From: 行に記載された TLD 名 (TLD_{ad})

表1 各項目とスパムメール技術の関係

詐称困難な項目	スパムメール対策技術
TLD _{ms}	OP25B, DNSSEC
IP _{ad}	DNSSEC
TLD _{ad}	SPF 認証, DKIM
TLD _g	DNSSEC

- D) IP_{ad} から GeoIP^[6]を用いて割り出したサーバ所在国の TLD 名 (TLD_g、IP が詐称されていた場合には GeoIP unknown とする)
- E) メールの日付

2.2 分析手法

発信元詐称防止技術として、近年利用が広がりつつある OP25B, DNSSEC, DKIM, SPF 認証の4つの詐称防止技術をとりあげた。それぞれの詐称防止技術が利用されたことにより、メール発信を示す TLD_{ms}, IP_{ad}, TLD_{ad}, TLD_g の4つ項目のうち、どの項目について詐称を困難にするかを表1に示す。

OP25B はメール送信サーバにおいてスパムメール送信を防止していることから TLD_{ms} に、SPF 認証^[7]は差出人のメールアドレスにする認定を行っていることから TLD_{ad} に、DNSSEC は送信時の DNS 応答からスパムかどうかを判定していることから TLD_{ms} と TLD_g に、DKIM はメール送信時に電子署名を要していることから TLD_{ad} にそれぞれかかるものと推定している。

これらの情報と合わせ TLD_{ms}, IP_{ad}, TLD_{ad}, TLD_g の定量/定性変化から TLD におけるガバナンス推定を行った。

3 分析結果

3.1 スパムメール発信量の推移

スパムメール変化量の月次変化を見るため、ここでは2010年12月と2011年5月の集計データについて分析する。その結果を表2に示す。表2は2つの要素から構成されており、上段4データは TLD_{ms}, IP_{ad}, TLD_{ad}, TLD_g 間の関係を、下段2データはメールサーバおよびサーバ設置国が不明なものを示している。表から、設置国や IP がはっきりしているサーバからのスパムメール送信量は激減しているが、対してこれらが不明なサーバからのスパムメール送信量は絶対数は減少しているもののスパムメール総数に対する割合は純増している。

表2 スパムメール発信量比較

項目	2010年12月	2011年5月
TLD _{ms} =TLD _{ad}	153,695(31.0%)	3,224(1.59%)
TLD _{ms} =TLD _g	93,346(19.0%)	21,137(10.4%)
TLD _{ad} =TLD _g	83,995(17.0%)	727(0.36%)
TLD _{ms} =TLD _{ad} =TLD _g	55,321(11.0%)	248(0.12%)
Server unknown	270,753(55.8%)	161,023(79.6%)
GeoIP unknown	64,085(13.2%)	57,147(28.2%)

[†] 長岡技術科学大学 Nagaoka University of Technology

3.2 発信元詐称防止技術導入による TLD 別の変化

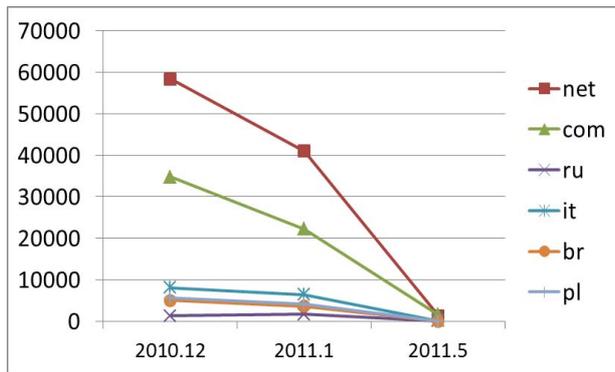


図1 TLD_{ms}=TLD_{ad}の一致状況推移

次に、図1に各TLD別のスパムメール発信量の変化を示す。ここでは、スパムメール発信量の上位6つのccTLD、gTLDを対象とした。図1より、「.com」「.net」の割合が非常に高く、2010年12月では、スパムメール総数の約2割を占めている。2010年12月から2011年1月にかけての減少は、表1において、DNSSECやDKIMが導入されたことによる効果であると推測できる。また、2011年1月から5月にかけて「.com」「.net」からのスパムメール発信量は約95%減少した。ボットネット閉鎖によるスパムメール総数の約59%減少に比べ、減少率が多いことから、特に図1のgTLD/ccTLDからは、ボットネットからスパムメールが送信されたものであると推測できる。

3.3 IP アドレスから見たスパムメール発信量推移

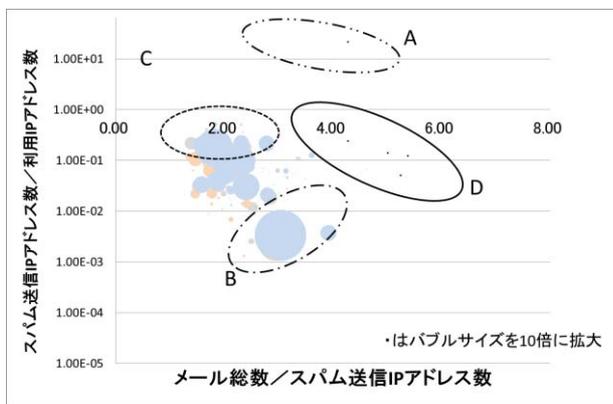


図2 IPアドレスから見たスパムメール発信推移

メール送信情報のうちIPアドレスはもっとも詐称が難しいデータ項目と言えるが、逆にこれを詐称しているTLDはIPアドレス管理が行き届いていないことを疑わせる。そこで、スパムメール発信に使われるIPアドレスを、実際に利用されているIPアドレスの数^[9]で除してスパムメール発信ドメイン比率を求めた。これは本来1以下になるはずであるが、1以上になった場合は、そのドメインには「IPアドレスを詐称して、スパムメールを発信しているものが多数存在することを意味する。図2において、この比率を縦軸に表す。横軸はccTLDごとのスパムメール総数をccTLDごとのスパム送信IPアドレスの数で除した

もので、1IPアドレスから送信されるスパムメール数の大小を示す。円の大きさは各ccTLDのスパムメール総数を示している。

領域Aはスパムメール発信比率が1を超えたものを示す。ここにはlc(セントルシア)が該当した。

領域Bは各国に割り当てられたIPアドレスからのスパムメール発信量は少ないが1つのIPアドレスからのスパムメール発信量が多い。特にus(アメリカ)が大きな割合を占めている。これは、特定のIPアドレスから大量のスパムメールが送信されていることを示している。

領域Cはスパムメールが大量に送信し、かつIPアドレスの多くをスパムメール送信に使用していることを示す。

領域Dは、少ないIPアドレスで多くのスパムを送信していることから、IPアドレスのほとんどをスパム送信に使用されていることを示す。領域C、Dともにsr(スリナム)やgu(グアム)などの島嶼国やsn(セネガル)やdo(ドミニカ)など途上国がほとんどを占めている。スパムメール発信量は少ないものの、IPアドレスのほとんどがスパムメール発信に使われており、2010年12月から2011年5月の間では大きな変化は確認出来ない。このことから、島嶼国のccTLDが悪用され続けている現状を表している。

4.TLDにおけるガバナンス推定

以上の分析から、TLDのガバナンスについて以下のように傾向を観察することができる。大規模なgTLDに対する発信元詐称防止技術の導入により、スパムメールは若干の減少が見られる。IPアドレスから見たスパムメール送信量推移は、IPあたりのスパムメール送信量及びスパムメール発信ドメイン比率から4分類される。島嶼国や途上国のccTLDのほとんどではスパムメール発信状況に改善が見られず、十分な管理がなされていないと推定できる。

5.まとめ

本稿では、各TLDの一致状況およびIPアドレスから見たスパムメール発信推移からTLDにおけるガバナンス推定を行った。今後も継続的に分析を行い、ガバナンスレベル評価に向けて各項目の詳細な分析を行っていく手法を考案する。

参考文献

- [1]シマンテック セキュリティレスポンスホワイトペーパー-2011年4月, 2011.5.28
http://www.symantec.com/ja/jp/business/security_response/whitepapers.jsp
- [2]日本データ通信協会, 2011.4.29
<http://www.dekyo.or.jp/soudan/taisaku/i1-3.html>
- [3]JPNIC, 2007.2.15; <http://www.nic.ad.jp/ja/basics/terms/DNS-cp.html>
- [4]RFC5617, 2009.8; <http://tools.ietf.org/html/rfc5617>
- [5]@IT;
<http://www.itmedia.co.jp/enterprise/articles/1101/07/news019.html>
- [6]GeOIP; <http://www.maxmind.com/app/ip-location>
- [7]RFC4871, 2007.5; <http://www.ietf.org/rfc/rfc4871.txt>
- [8]JPCERTCC, 2006.7;
http://www.jpCERT.or.jp/research/2006/Botnet_summary_0720.pdf
- [9]迷惑メール対策, 2011.6.25
http://www.quia.jp/spam/CountryList_ja.html