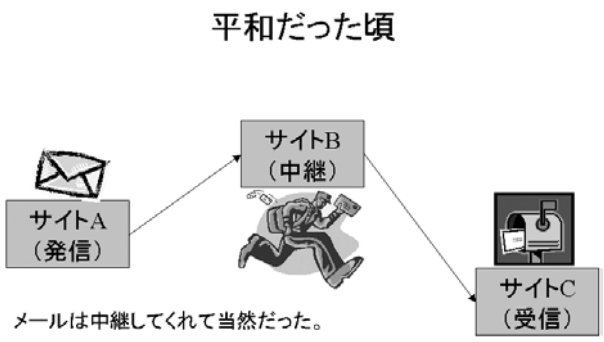
迷惑メール（スパム）対策技術の変遷

株式会社IRIコミュニケーションズ
安藤　一憲

メールボックスが頼んでもいないのに送られてくる広告メールで埋め尽くされている、という現象は昨今珍しくない。この招かざる客は一般に迷惑メールとかスパムと呼ばれる。いまやスパムの送信元は世界中に散らばっており、国内法だけで取締りが十分にできる保証はない。恐ろしいことに、たかがスパムに立ち向かうにも国を超えた対処が必要なのである。さらに言えば、スパムの受信量には著しい個人差がある。自分のメールボックスが無事でも会社の他の人間のメールボックスが無事だとは限らない。いまやスパム対策は新旧入り交じり、どの対策がどういう利点を持ち何に弱いのかを掴むのもひと苦労という状況になりつつある。本稿ではそれぞれのスパム対策技術がどのような経過で登場してきたかをひととおり振返ることで、最新の動きである送信ドメイン認証やレピュテーションといった技術がどういう意味を持っているのか改めて考えてみたい。

1. 平和だった頃

まだ世の中にWWWが存在せず、メール配送先の決定がDNSに依存していなかった時代、メール配送は中継によって成り立っていた。上流のサイトは善意でバケツリレーのようにメールを運んでくれたものである。よもやメールの中継そのものがスパム配送の温床になることを誰が想像しただろうか。しかし、善意に基づいて作られたシステムは悪用する者が現れると壊れる。スパム対策はメール配送設定の一部として不正中継防止策という形でスタートした。初期のスパム対策は、発信者のメールアドレスのドメイン部分によって中継の可否を判定するというものであった。考えるに自分のメールアドレスが流通しているからこそこれだけの数のスパムが届くのだが、いまどきワームまでもが検索エンジンを使ってメールアドレスを探す時代である。便利になる反面、悪用される度合いも増しているといえる。

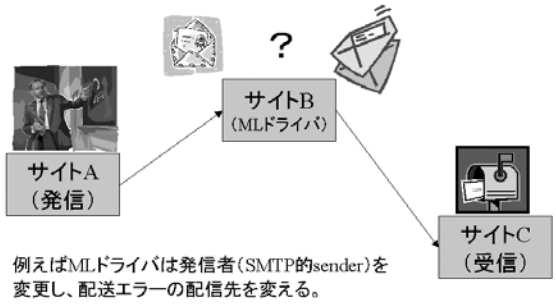


2. なりすまし

メール送受信の基本的な枠組は通信手順（RFC2821）と送信されるメッセージの形式（RFC2822）で規定されている。当初そこには発信者の認証という概念はなかった。極論すれば仕組みさえ知っていれば誰でも誰にでもなりすますことが可能であった。この便利ななりすましの枠組は現在においてもメーリングリストのアドレス展開や、メール転送（aliases）な

どでほぼ日常的に使われている。善意に基づいて作られたシステムゆえにこうなっているが、スパマー(SPAMer)は初期の不正中継対策をかいくぐるために、まずこのなりすましを利用し始めた。ISPではこれに対応して、認証方式としてはいささか邪道ではあるが、他のプロトコルであるPOPのユーザ認証を以てSMTPの認証のかわりをさせ、一定時間そのIPアドレスからのメール発信を許可する「POP before SMTP」という対策をとることになる。

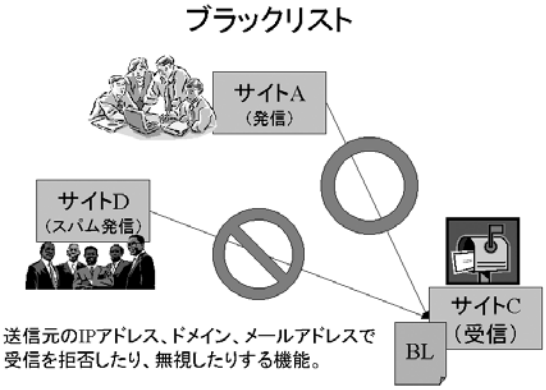
なりすまし



3. ブラックリスト

送信者のメールアドレスが詐称可能であるという事実は、送信者のメールアドレスだけに基づいた一切の不正中継対策が不完全であることを意味する。ここに至って不正中継の防止は発信サイトのIPアドレスに基づいてルールが書かれるようになった。同時に、確信犯でスパムを配信してくるサイトに対して発信サイトのIPアドレス、ドメイン、発信メールアドレスのいずれかを指定してピンポイントで配信を止める仕組みが登場する。これがブラックリストである。

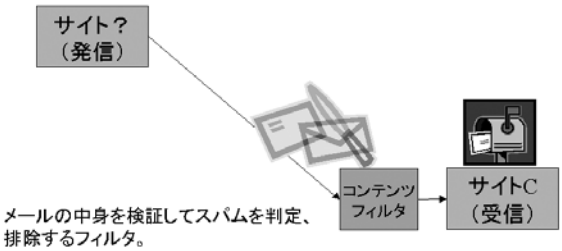
ブラックリスト



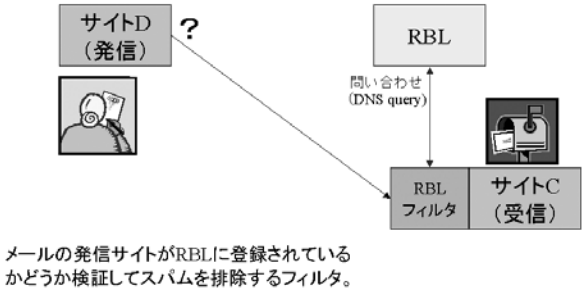
4. コンテンツフィルタとRBL (リアルタイムブラックホールリスト)

ブラックリストが普及すると、やがてスパムの発信サイトは次々に新しいアドレスを利用するようになる。その目的はブラックリストに載っていないアドレスからメールを送信することにある。その結果ブラックリストのメンテナンスコストが増大し、スパム対策は大きく2つの系統に分かれることになる。1つは送信サイトのIPアドレスには依存せず、メールの文面からスパムかどうかを判定するコンテンツフィルタのアプローチ、もうひとつは、ブラックリストを共有することでメンテナンスコストを下げるRBLのアプローチである。当初のコンテンツフィルタは単純なパターンマッチを基本としていたが、やがて正規表現が使えるようになった。しかし、スパマーはわざと単語のスペルミスをしたり、適当に「＊」をまぶしたり、HTMLのコメントを単語の途中に入れたり、ありとあらゆる手法で文面のバリエーションを増やしていくことになる。一方、RBLに待っていたのはブラックリスト情報の授受の通信そのものを妨害するサーバへのDoSであった。この頃からスパム送信へのすさまじい執念があちこちで感じられるようになる。

コンテンツフィルタ



RBLによるフィルタリング



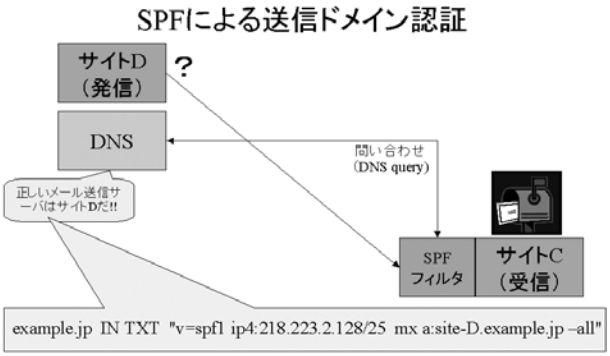
5. ベイジアンフィルタ

もともとは、文面に現れる単語の出現傾向から自然言語で書かれた文章を「区分け」する研究が行なわれていた学習型のフィルタ技術である。この技術が一斉にスパム対策に応用されたのは、Paul Grahamの「A Plan for Spam」という文章に触発されたからとされている。メールをスパムとハム（スパムでないメール）に区分けするため、単一のメールボックスしか持たないPOP（Post Office Protocol）との整合性が悪く、今まではMUA※やPOPサーバとMUAの間に入るPOP proxyの形で実装されてきた。各個人によって迷惑メールの定義が微妙に異なることもMUAへの実装が進んだ原因のひとつであろう。言うまでもなく学習型のベイジアンフィルタが登場した背景には、スパマーによる文面バリエーションの増大によりパターンマッチルールのメンテナンスコストが増大した事実がある。ベイジアンフィルタは辞書ベースであるため性能に言語依存性が存在する。不幸なことに最も知られた応用がスパム対策になってはいるが、メールを文面で仕分けする仕組みとして複数のメールフォルダを扱えるMUAかIMAPサーバで利用できるようになれば、それはそれで素晴らしい技術の応用になると言えるだろう。

6. 偽装単語列と人間の誤り訂正能力

ベイジアンフィルタは単語ベースのフィルタである。とすれば、当然スパマーは単語の出現傾向を変えてこのフィルタを突破しようとする。そこで登場してくるのが単語をランダムに並べたものや、全然関係のない文書をメールに混入するという手法である。その後の研究によりベイジアンフィルタで学習を続けると、辞書中にフィルタとして弱点となる単語が出てくることがわかっている。そのため、学習済の辞書を搭載して学習させない実装や、辞書に掲載される単語数を一定に限定する実装が出てきた。だが、困ったことに人間の誤り訂正能力はベイジアンフィルタをはるかに超越しており、例えば「V=i=a=g=r=a」のように「Viagra」と読める綴り方だけで60強通りもバリエーションがあるとする報告がある。これらをすべて辞書に持つのは現実的ではない。このバリエーションを吸収するため、辞書に正規表現のパターンを持たせたベイジアンフィルタを考えている向きもあるようだ。きっと今度は学習するうちに弱点となる正規表現のパターンが出てきて、そこを突く方策をスパマーが打ってくるという展開が待っているのだろう。コンテンツフィルタをめぐるスパマーと研究者の闘争は本当に果てしない...。だが、単に研究者とスパマーがベイジアンの木の下で鬼ごっこをやっているだけのようにも見える。

※MUA：Mail User Agent, いわゆるメーラー



7. 特徴抽出型(ヒューリスティック)フィルタ

簡単に言えば多数の判定アルゴリズムを用いて○×表を作り、○×の出現パターンによってスパムかどうかを判定するシステムである。コンテンツフィルタの一種と言って良いが、多くのシステムで判定項目の構成がブラックボックスになっており、どうやれば引っかからずに済むかを調べるのは非常に困難になっている。スパマーはこれをくぐり抜けるための努力も惜しんでいないので、固定的に判定アルゴリズムを集めただけでは、いつかはくぐり抜けられてしまうであろう。

前章でベイジアンフィルタのはまっている迷路についても書いたが、究極のフィルタは動的に判定項目が変えられるタイプのものになるであろうことが予測できる。アルゴリズムの構成が適度に変動することがスパマーによる解析を事実上不可能にする。実はコンテンツフィルタは本質的にオープンソースには向かず、少なくともプロプライエタリな部分を技術コアに持たないと成立しないのかも知れないと筆者は最近考えている。鬼がフィルタのソースを読んでしまうような鬼ごっこは、フィルタを書く側に圧倒的に不利に思えるがあなたの判断はいかがだろうか？

8. 送信者認証(SMTP AUTH)

一方、スパマーは「POP before SMTP」という陳腐化してしまった障壁をかいくぐる努力も忘れていないようだ。世の中には接続する毎にIPアドレスの変わるインターネット接続サービスがある。ということは、運が良ければ、直前にそのIPアドレスを利用していた人がPOPで認証していて、自分が認証せずともたまたまメールが送信できる状態のIPアドレスが当たるかも知れない。かくして、接続してはSMTPを叩き、接続してはSMTPを叩き、果てしなくIPアドレスを変えまくってスパム送信を試みる輩が出てくる。もはや執念と言って良いだろう。このアタックは送信者の認証がSMTPセッション自体に含まれていないことに根本的な原因がある。そこで、SMTPセッションの中でユーザ認証をする仕組みが考えられた。それがSMTP AUTHである。実は日本はISPレベルでのSMTP AUTHの普及が進んでいる国のひとつである。これはメール系の技術者の見識の高さを裏付ける事実のひとつであろう。次の段階では、認証されているという事実をいかにして送信先に伝えるかという課題が待っているわけだが、せっかく認証がSMTPセッションの中で閉じているのに、いくらでも詐称が可能なメールヘッダに認証結果を残すことが課題の解決にどれだけ貢献するかは不明である。

9. ゾンビPCとbotNET(ゾンビクラスタ)

スパマーは認証の壁をいかに破るか、RBLの壁をいかに破るかを追求してきた。その結果、ほぼ最終兵器ともいえる解答に到達する。現状で世界最大のクラスタリングマシンであろう「botNET」と呼ばれるメール送信用のPCクラスタである。インターネット上にランダムに分布する数十万とも数百万とも言われている数のPCがスパマーの出す指令の通りにスパムを送出するのである。PCはワームに感染することで、ゾンビクラスタの一部(ゾンビPC)となり、スパ

マーの指令を待つのである。ワームはOSやアプリケーションの脆弱性を利用して感染を試みるので、インターネットにマシンを繋ぐならくれぐれも発覚している脆弱性へのケアを忘れてはいけない。スパム受信側から見ると、数百万ものランダムなIPアドレスからスパムが届くという悪夢のような状態になる。結果、発信サイトのIPアドレスベースのスパム対策は実質的に無効化されてしまう。実際、これ以上発信者の隠蔽に適したシステムはない。これがフィッシングと呼ばれる詐欺メールが発生する温床のひとつとなった。

10. フィッシングメール

カード番号や暗証番号の詐取を試みるメール。差出人を詐称し、WWWブラウザの脆弱性を利用してURLをも偽装して偽のサイトへ被害者を誘導する。目的が目的だけに金融機関、カード会社、ショッピングサイト、ISPの名をかたるものが多い。実際に金銭被害が発生する犯罪であるため、各国の消費者保護を担当する役所が対策に乗り出してきており、日本の経済産業省、総務省、警察庁もその例外ではない。少なくとも初動では他の国に負けていないので、被害を最小限に押えることができたなら、彼らの対策は世界的に見て十分に賞賛に値すべきものになる。目下最も対策の緊急度が高いターゲットはこれである。メールサーバの話ではないが、フィッシングの誘導先に使われる偽のWWWサイトの7割以上がapacheだという数字がある。どうもWindowsに限った話ではないようなので、WWWサーバの管理者もサイト上にページをでっち上げられて悪用されないように注意しなければいけない。油断は禁物である。

11. 送信ドメイン認証（SPF,Sender-ID,DomainKeys）

「そのメールが送信者アドレスのドメインの正規のメール送信サーバから送出された」ことを確認するための認証技術。DNSのそのドメインに対応するエントリーにそのドメインの正規のメール送信サーバのIPアドレスを記述したり、正規のメール送信サーバに対応する公開鍵を記述することで、送信されたメールがそのドメインを持つ組織から正当に発信されていることを検証する。送信したユーザ個人を認証する技術はもともとあるが、この用途で個人を特定するのは明らかにオーバースペックなためか、普及してこなかったという背景がある。送信ドメイン認証は簡単に言えば、「この組織から発信されました」という事実を検証する枠組であり、ゾンビクラスタが存在する環境下でも発信者の詐称を著しく困難にする効果があるため、スパム及びフィッシング対策として導入が推奨されるべきものである。例えば、米国の組織が先にこの対策を導入した場合、DNSに対応するエントリーを書いておかないと日本からのメールが受けとってもらえなくなる可能性がある。その前に、最低でもどれか1種類、自社の正規のメール送出サーバの情報をDNSのエントリーとして設定しておきたい。

12. Port 25 blocking

送信ドメイン認証がメールを受信する際の対策とすれば、ポート25ブロッキングは自社のアドレスブロックにあるゾンビPCからのメール送信を遮断するための対策である。ISPの立場でいえば、コンテンツフィルタのようなサーバの高負荷を伴う対策よりはるかに小さな投資額で自分たちのアドレスブロックから発信されるスパムへの苦情に対応する労力を激減させることができるという意味で効果的な対策の1つである。仮にこの対策を怠った場合、後述するレピュテーション(ドメイン信用評価サービス)において、信用度が地に堕ちるリスクも考えられる。ポート25がブロックされた場合、ユーザは他のプロバイダからメールASPを使ってのメール送信ができなくなると考えるかも知れないが、他のプロバイダのメールサーバ

を使ったメール発信には、ポート25以外の通信を利用することで現状の使い勝手を維持することとなろう。例えば、メール発信はポート587(Message submission)を用いれば良いし、さらに、Message submissionをTLSで暗号化したり、SMTP/SSL(ポート465)を使用したりする手もある。心あるISPはこういった回避策をユーザに提供すべく必死に準備をしている。変化は発信と中継を完全に分離する方向に着実に進行している。

13. レピュテーション

あるドメインやIPアドレス、ドメイン保持者やIPアドレスブロックのオーナー情報とスパム発信履歴をもとにそのサイトの「信用度」を算出し提供するサービス。情報の取得手段はRBLに類似しているが、中身は単純なブラックリストではない。ISPが自社のアドレスブロックからのスパム発信を放置した場合、このサービスを利用しているサイトに受信を拒否される可能性が出てくる。モデルは実社会の信用調査と似ており、今後は信用度判定の情報源がどんどん高度化する方向へ進化することによりさらに似てくるものと思われる。

14. むすび

インターネット上に電子メールというメディアができて普及していく際にたどってきた過程は、使い方の自由度と対策の難しさの差こそあれ、他の社会インフラがたどってきた過程とそれほど違ったものではないように思う。この壮大なスケールでの実験はまだ続いており、メールサーバ管理者は否応なくその最前線に立たされていると思って間違いない。メールは依然としてインターネット上で最も利用されているサービスであり、ユーザの母数が多いだけに問題も最初に顕在化するからである。スパムの問題はNetnewsはもとより、今やブログのトラックバックにまで広がっており、あらゆるテキストを利用する通信サービスがその被害にあっている。問題はメールだけに留まらず、今後もユーザの多い順に次々と顕在化していくことが予想される。ユーザ側も、いつも使えているからといって、たかがメールとは思ってはいけないのだ。サーバ管理者の方は現状では高いスキルと低い労働単価を要求されているかも知れないが、気概を持って対策を検討されたい。メールサーバが機能を停止すると会社の機能がほぼまるごと停止するという事象はいまどき珍しいことではない。これは管理職もシステム管理者も必ず一度は考えておくべきリスクである。

送信者認証(SMTP AUTH)と送信ドメイン認証(SPF,DomainKeys等)は全く異なる技術であるが、あまり他の方の講演を聞く機会のない私でさえ、この2つを混同した講演に遭遇したことがある。300名近い参加者が間違った知識を持ったまま家路につくというのは痛恨の極みであり、本稿を書く動機としては十分なものであった。日本にインターネットが上陸して20年、私が初めてネットワークにIP接続されているマシンに触ってからもう17年ほどになる。ユーザの裾野は桁違いに広がり、メールなんざ誰でも使っているという時代になりつつあるが、昔も今も正しい知識の必要性は何ら変わっておらず、次の議論をするには以前の経緯を正確にふまえておくべきである。本稿がそのための何らかの助けになれば幸いである。