# 迷惑メール問題の基礎知識

国立情報学研究所
中村 素典

# 内容

- 迷惑メールの現状
- なぜ迷惑メールが可能なのか
- 迷惑メール送信の手口

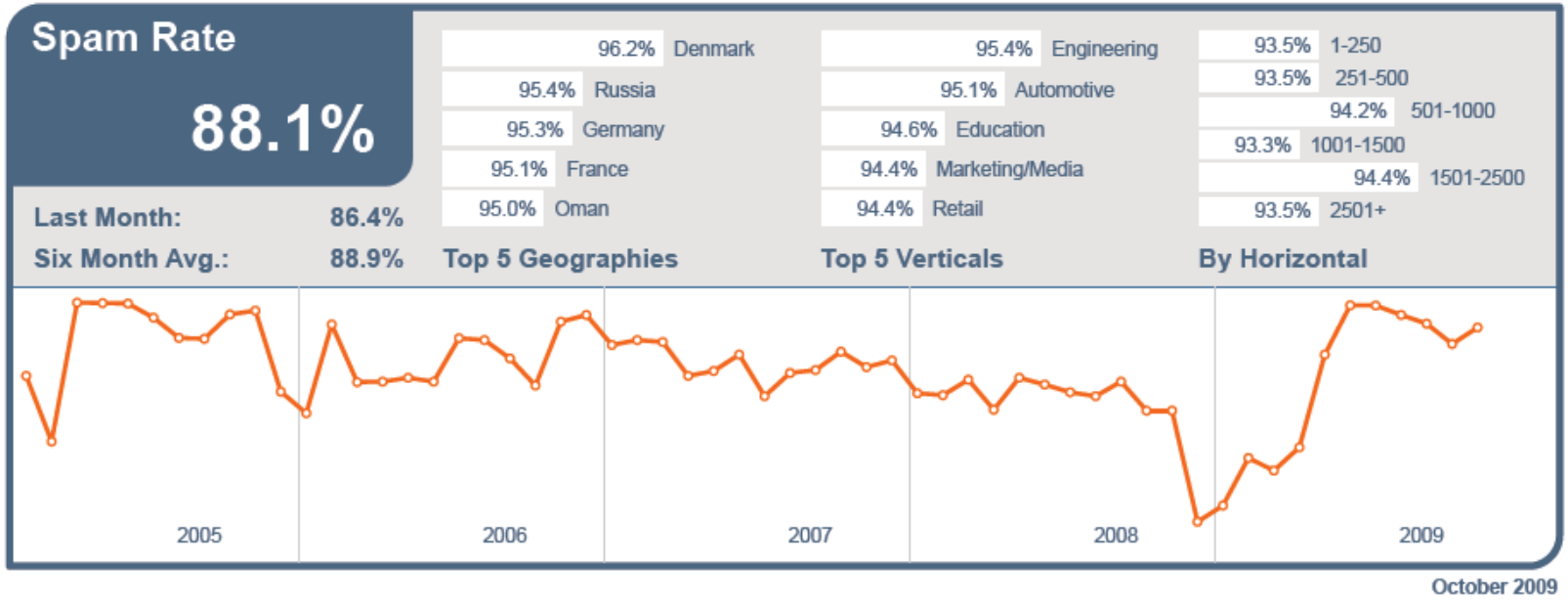
# スパムの現状

- 7割以上が迷惑メール
  - 2009年8月調査（迷惑メール対策推進協議会）
    - http://japan.cnet.com/marketing/story/0,3800080523,20401760,00.htm
- 迷惑メールは儲かる：1日40万円の売上げも
  - http://wiredvision.jp/news/200910/2009100622.html
- 迷惑メール発信地域調査 (SophosLabs, 2009/7-9)
  - 1位が米国（13.3%）、2位がブラジル（12.1%）、3位がインド（5.6%）、4位が韓国（5.5%）、5位がベトナム（4.7%）
  - http://www.asahi.com/digital/cnet/CNT200911040044.html
- その他統計情報
  - http://spamlinks.net/stats.htm#received-big-pc
  - http://www.viruslistjp.com/analysis/

# MessageLabs Report

- MessageLabs Intelligence October 2009

**Skeptic™ Anti-Spam Protection:** In October 2009, the global ratio of spam in email traffic increased by 1.7% from the previous month to 88.1% (1 in 1.1 emails).




http://www.messagelabs.com/mlireport/October%202009%20MessageLabs%20Intelligence%20Report%20FINAL_EN.pdf

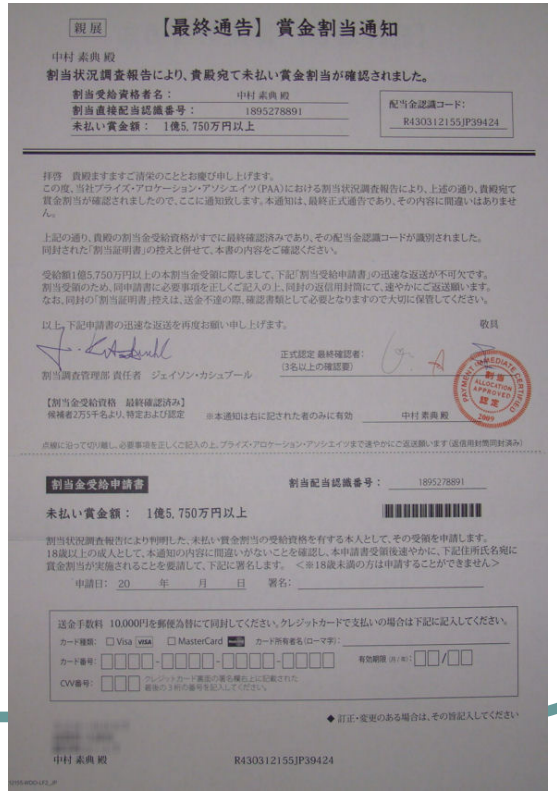
# スパムメールと二酸化炭素排出量

McAfeeによる調査報告
http://www.mcafee.com/common/media/mcafeeb2b/international/japan/pdf/p_web/CarbonFootprint_J.pdf

- 電子メール関連のエネルギー消費のうち1/3はスパムに起因（対策分も含む）。
- 2008年後半の、McColoの遮断によるスパムメールの減少は70%。再構築の間に減少したエネルギー消費量は、1日220万台分に相当。

# これもCO²排出量増加の原因

- こんなに続けて賞金は当たらない！
- 賞金をもらうのになぜ手数料を支払うの？



親展
【最終通告】賞金割当通知
中村 素典 殿
割当状況調査報告により、貴殿宛て未払い賞金割当が確認されました。
割当受給資格者名： 中村 素典 殿
割当直接配当認識番号： 1895278891
未払い賞金額： 1億5,750万円以上
配当金認識コード: R430312155JP39424
割当調査管理部 責任者 ジェイソン・カシュプール
【割当金受給資格 最終確認済み】
候補者2755千名より、特定および認定
※本通知はお宛に記された者のみに有効
中村 素典 殿
割当金受給申請書
割当配当認識番号： 1895278891
未払い賞金額： 1億5,750万円以上
申請日： 20 年 月 日 署名：
送金手数料 10,000円を郵便為替にて同封してください。クレジットカードで支払いの場合は下記に記入してください。
カード種別: Visa MasterCard カード所有者名(ローマ字):
カード番号: - - - 有効期限(月/年): /
CVV番号:
◆訂正・変更のある場合は、その旨記入してください
中村 素典 殿
R430312155JP39424

# インターネット裏社会の実態

## インターネット裏社会で行われている犯罪の料金

| 犯罪サービス | 料金 |
|---|---|
| マルウェア*を誰かに1度インストールさせる | アメリカ: 30セント、カナダ: 20セント、イギリス: 10セント、その他の地域で2セント |
| マルウェアの基本セット | 1,000ドルから2,000ドル |
| マルウェアのオプショナル・サービス価格 | 20ドルから |
| エクスプロイト・キット*のレンタル(1時間) | 0.99ドルから1ドル |
| エクスプロイト・キットのレンタル(2.5時間) | 1.60ドルから2ドル |
| エクスプロイト・キットのレンタル(5時間) | 4ドルから |
| AVベンダーに未検出の情報を盗めるトロイ* | 80ドルから |
| DoS攻撃* | 1日あたり100ドルから |
| 10,000台の、既に仕掛けが施されたPC | 1,000ドルから |
| 盗んだ銀行アカウント情報 | 50ドルから |
| 最新のe-mailリスト(有効かの確認は未だ) | 8ドルから。有効確認済の場合は値段が高い |

**出典;Sample data from research on the underground digital economy in 2007**

＊ マルウェア･･･ウイルスやスパイウェアなどの不正プログラム
＊ エクスプロイト・キット･･･脆弱性を攻撃するための小さなプログラム、また脆弱性の存在を実証するためのプログラム
＊トロイ･･･トロイの木馬のこと。
＊ Dos攻撃･･･大量のデータや不正パケットを送りつけるなどして、不正にサーバなどへ攻撃すること

トレンドマイクロ社のWebより: http://virusbuster.jp/vb2010/underground/

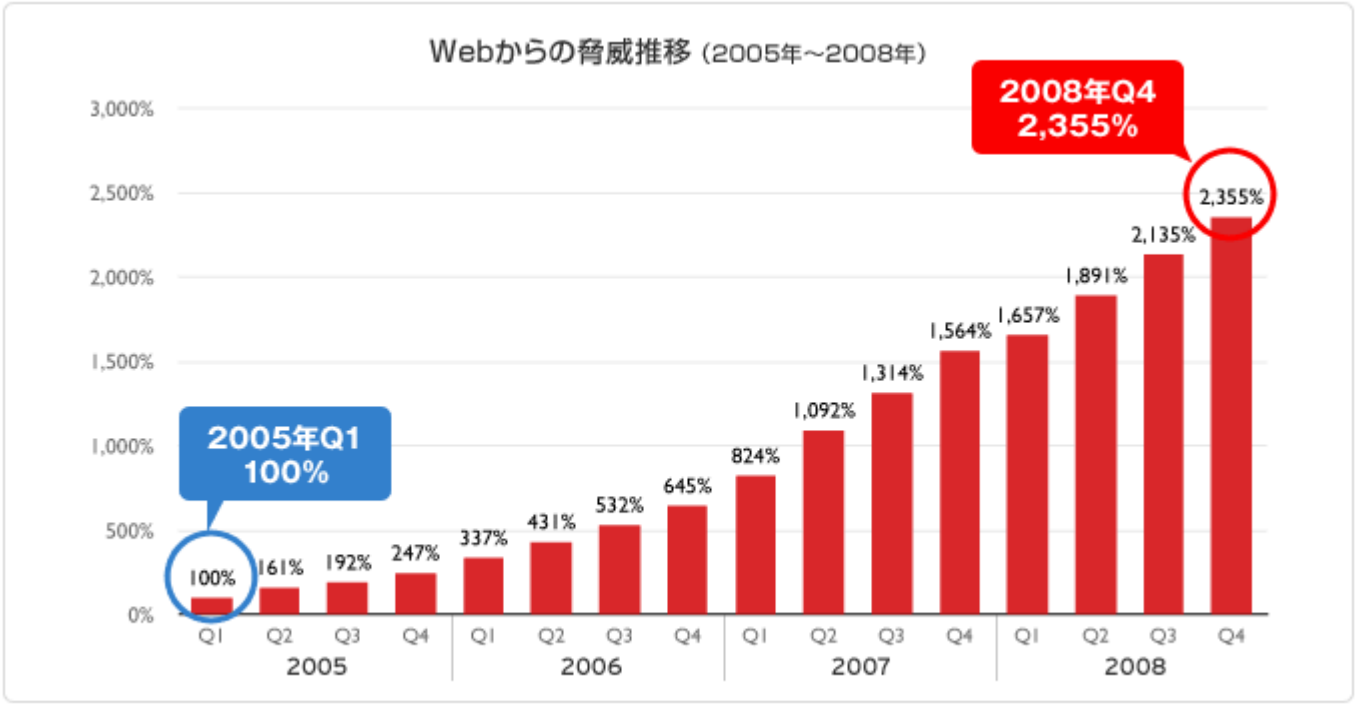
# 組織ぐるみで行われる巧妙な手口


近年の金銭や情報収集を目的としたインターネット犯罪は、情報利用者、攻撃者、仲介者など各人の役割が分業化されており、まるで犯罪シンジケートのような形をなしています。

トレンドマイクロ社のWebより: http://virusbuster.jp/vb2010/underground/

# 激増するWebサイトからのインターネット犯罪

- 2008年末には2005年はじめの23倍




＊ 2009年トレンドマイクロ調べ

トレンドマイクロ社のWebより: http://virusbuster.jp/vb2010/underground/

# 電子メール環境の変遷

- ホスト計算機によるネットの時代
  - ユーザ権限と管理者権限の分離で秩序を保つ
- クライアント・サーバシステムへ
  - UUCPからSMTP (TCP/IP) へ
    - telnet HOST 25 という技
  - 計算機の低価格化、パーソナル化
    - MS-DOSからWindowsへ
    - 個人UNIXシステム (*BSD、Linux)

# （設計当初の）SMTPの問題

- 誰でも（どこからでも）メールを送信することが可能
- 誰に対しても送信することが可能（中継可）
- 送信者アドレスの正当性を確認しない

SMTPの手順

S: MAIL FROM:<Smith@Alpha.ARPA>
R: 250 OK
S: RCPT TO:<Jones@Beta.ARPA>
R: 250 OK
S: RCPT TO:<Brown@Beta.ARPA>
R: 250 OK
S: DATA
R: 354 Start mail input; end with <CRLF>.<CRLF>
S: ...etc. etc. etc.
S: <CRLF>.<CRLF>
R: 250 OK

故き良き時代の産物

# 電子メールに関する インターネット標準規格

- RFC821: SIMPLE MAIL TRANSFER PROTOCOL (1982/8)
  - RFC772 (1980/9), 780 (1981/5) Mail Transfer Protocol, 788 (1981/11) SMTP の更新
  - RFC2821 (2001/4), 5321 (2008/10) へと更新
- RFC822: STANDARD FOR THE FORMAT OF ARPA INTERNET TEXT MESSAGES (1982/8)
  - RFC2822 (2001/4), 5322 (2008/10) へと更新
  - RFC733  Standard for the format of ARPA network text messages (1977/11) の更新

# RFC722: Mail Transfer Protocol (1980/9)

- TELNET の拡張として規定
- コマンド
  - MAIL
    - MAIL FROM: [TO: ]（宛先も指定できた）
  - MRSQ (MAIL RECIPIENT SCHEME QUESTION)
    - MRSQ R – recipients first
    - MRSQ T – text first
      - because neither by itself is optimal for all systems
  - MRCP (MAIL RECIPIENT)
    - MRCP TO:（複数の宛先を指定する場合）
  - HELP
  - QUIT
  - NOOP

# MASQ Tという手順があった

- If these receiver-MTPs tried to emulate MRSQ Rs bulk mailing, they would have to ensure that a success reply to the MAIL indeed meant that it had been delivered to ALL recipients specified -- not just some.

S: MRSQ T <CRLF>
R: 200 OK, using that scheme
S: MAIL FROM:<WALDO@A><CRLF>
R: 354 Type mail, ended by <CRLF>.<CRLF>
S: Blah blah blah blah....etc. etc. etc.
S: <CRLF>.<CRLF>
R: 250 Mail stored
S: MRCP TO:<Foo@Y> <CRLF>
R: 250 Stored mail sent
S: MRCP TO:<Raboof@Y> <CRLF>
R: 553  No such user here
S: MRCP TO:<@Y,@X,fubar@Z> <CRLF>
R: 200 OK

ユーザ数が確定するのが最後なので、やはりこれでも困ったか

# TELNET port 25 問題解決の試み

- RFC1413 Identification Protocol (113/tcp) を用いてユーザIDを取得しトレース可能にしよう
  - ヘッダに記録を残す
    - Received from: site (user@host [192.168.1.135])
  - Sendmail 6.51 (1993/4) でコード追加
- ファイアウォールとの相性問題（メールが届かない）
  - Sendmail 8.6.5 (1994/1) で簡単にdisableできるように (-DIDENTPROTO=0)
  - Sendmail 8.10 (2000/3) でタイムアウトを30秒から5秒に変更

# 認証の仕組みの提案

- POP XTND XMIT
  - SMTPで送信すべきとして非推奨
- POP before SMTP
  - POP認証との連携
    - IPアドレス再利用やNATからのアクセスに問題
- Message Submission の分離 (RFC2476)
  - 587/TCP を利用
- SMTP ext for Authentication (RFC2554)
  - AUTH LOGIN/PLAIN はパスワードが暗号化されずに流れるので、通信路の暗号化が別途必要
    - SMTP/SSL (465/TCP)
    - SMTPTLS (同一ポートで途中からTLSにスイッチ: RFC3207)

# RFC5321 (2008/10) での主な変更点

- メールクライアントにおける Message Submission Protocol (RFC4409, 587/TCP) の利用推奨
- 迷惑メールに関連した記述の追加
  - 必ずバウンスするのは現実的ではない
    - 有用であると判断できるときのみバウンスすべき
  - メールを捨てると信頼性を損なう
    - 確実に無効であると判断できるときのみ捨てるべき
  - 宛先(RCPT TO)の大半が無効である場合はサーバ側から切断することもやむなし

# 迷惑メール送信の手口

1. 無料電子メールアカウントの悪用
2. 電子メールアカウントの乗っ取り
3. ISP/ASPの「渡り」
4. Webフォームの悪用
5. Botの利用

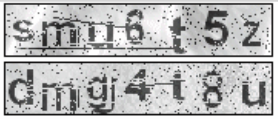
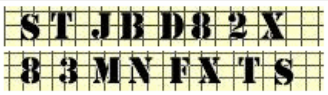
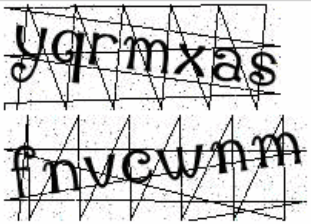
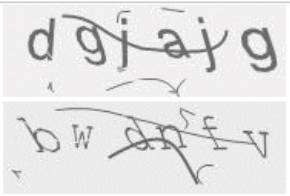
参考：http://wiki.asrg.sp.am/wiki/Taxonomy_of_spamming_techniques

# 無料電子メールアカウントの悪用

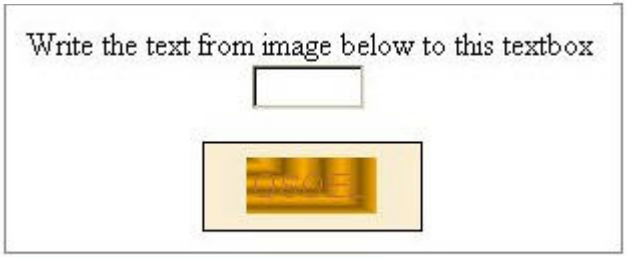
- コストをかけずに、迷惑メールの送信が可能
- 当然、アカウント取得の自動化が行われる
- 対策として、CAPTCHAが広く用いられる
  - Completely Automated Public Turing test to tell Computers and Humans Apart
  - 簡単に読めるものから、非常に難解なものまで
    - http://labaq.com/archives/50932625.html
- 対抗策
  - 別のサービスへのアクセスを目的に、CAPTCHA画像を転送して人間に読ませる（エログリッド）
  - 機械解読手法の開発
    - PWNtcha：http://caca.zoy.org/wiki/PWNtcha

# PWNtchaによる解読

http://caca.zoy.org/wiki/PWNtcha より

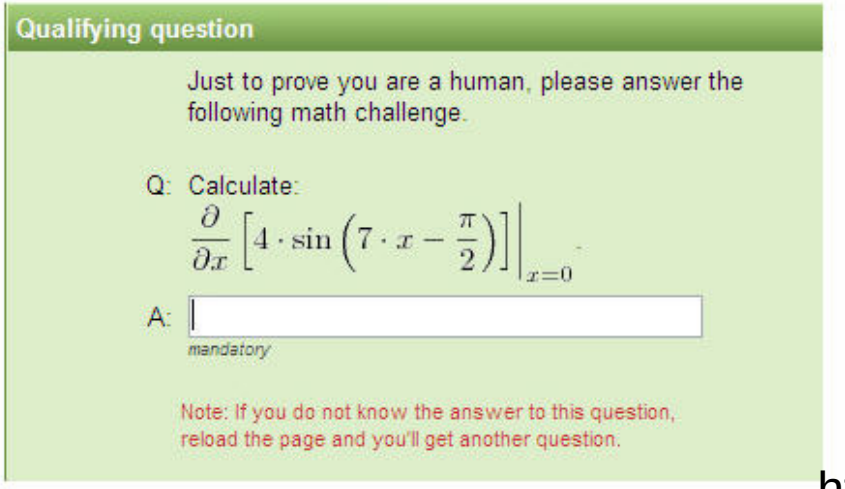
| | | | |
|---|---|---|---|
| Authimage |  | 100% | Vendor site: http://www.gudlyf.com/index.php?p=376 ⇨ Weaknesses: constant font, aligned glyphs, constant glyph position, constant rotation, no deformation, non-textured background, constant colours, no perturbation. |
| linuxfr.org |  | 100% | Weaknesses: constant font, aligned glyphs, no rotation, no deformation, non-textured background, weak colour variation, weak perturbation. |
| LiveJournal? missing! |  | 99% | Weaknesses: constant font, constant character position. |
| Paypal |  | 88% | Weaknesses: constant font, almost aligned glyphs, no rotation, no deformation, constant background, no colour variation, no additional perturbation. |
| phpBB |  | 97% | Vendor site: http://www.phpbb.com/ ⇨ Weaknesses: constant font, no rotation, no deformation, constant colours, weak perturbation. |
| Scode and derivatives |  | 100% | Vendor site: http://james.seng.cc/archives/000145.html ⇨ Weaknesses: at most 3 different fonts, no rotation, no deformation, weak colour variation, useless perturbation (separate colour key). |
| Slashdot |  | 89% | Weaknesses: constant font, no deformation, constant colours, weak perturbation. |
| Xanga |  | 49% | Weaknesses: fixed horizontal glyph position, no rotation, no deformation, constant colours, insufficient perturbation. |

# 難解なCAPTCHA










http://labaq.com/archives/50932625.html より

# 電子メールアカウントの乗っ取り

- フィッシング等により電子メールアカウント情報を入手して悪用
  - 特に多数のユーザがいるメールサーバ
- 対策:
  - メール送信の頻度規制
  - 組織外からのサブミッションに対する規制

# オープンリレー

- 第三者中継を許すメールサーバ
  - スパムメールが社会問題として顕在化する前は、オープンリレーは一般的だった
    - ソースルーティング表記も使えた（user%domain@relay等）
  - 現在は、IPアドレスの範囲制限や認証に基づいて中継させるのは常識
    - POP-before-SMTP（もはや古い）
    - SMTP AUTH (over TLS)
  - ウィルス感染によりオープンリレー化される場合も。
- Webのオープンプロキシも問題
  - CONNECTメソッドが使えるとメールの送信が可能

参考：http://www.atmarkit.co.jp/fsecurity/rensai/handling03/handling01.html

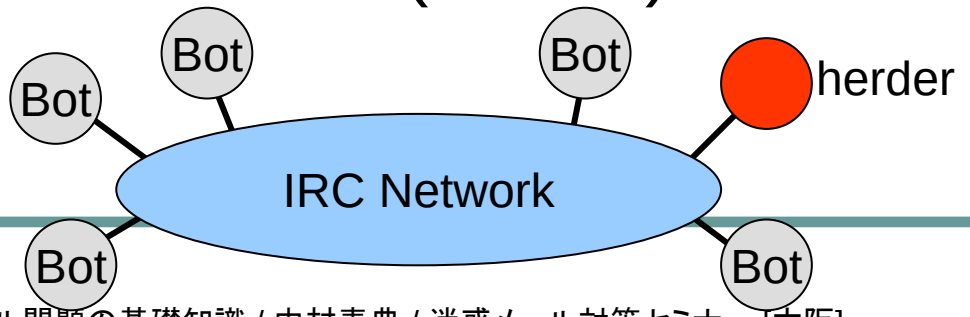
# Webフォームの悪用

- Webフォームからメールが送信されるCGI等は認証なしで利用できることが多い
- 対策:
  - 本文の内容が自由に指定できないように、変更できる範囲を限定する
  - 宛先を限定する
  - 送信頻度を規制する

  など

# BotNet

- ウィルスに感染し遠隔制御可能にした多数のPC (Bot, Zombie)によって構築されたネットワーク
- スパム送信、DDoS (Distributed Denial of Service: 分散サービス使用不能)攻撃、クリック数の獲得（Web広告収入の不正取得）、フィッシング等に利用される
- IRC (Internet Relay Chat) ネットワーク等を介して制御されており、制御者(herder)の足取りがつきにくい



# BotNetによるスパムの送信

- 送信元となるBotを順に変化させていくことによって、発信ホストに対するブラックリストをかいくぐる
  - MAPS Dial-Up List (DUL)等のDNS参照方式ブラックリスト (DNSBL) → by TrendMicro since 2005
- 詐称する送信者メールアドレスも変化させている？
- ISP内の正規な送信メールサーバが利用されるとOP25B (Outbound Port 25 Blocking)が回避される
  - 認証情報も取得されてしまう可能性
  - 送信レート制限、総量制限は有効か？

# フィッシング(Phishing)詐欺

- 2003～2004年頃から広く認識されるようになる
- スパムを送信し、偽装したWebサーバに銀行口座やクレジットカードの情報を入力するように仕向け、不正に金銭を取得するといった手口
  - 被害の大半は、メール受信の直後に発生
  - さらにその銀行口座を利用して、オークションサイトで嘘の出品を行い、代金をだまし取られる可能性も
  - さらにPCをウィルスに感染させる
    - キーロガー等のスパイウェアでさらに被害が増大
    - Bot化で被害者が加害者に
- DNSの脆弱性を用いたファーミングにも注意

# フィッシングの高度化: Fast-Flux

- フィッシングの足取りをつきにくくするために、BotにWebアクセスを中継させる（2007夏頃～）
  - 背後にいる情報収集用サーバは1台(Mother Ship)
  - スパムメールに記載されるURLに使われるドメイン名は1種類
  - DNSにおいて、ドメイン名からIPアドレスを解決する際に、毎回違うものを答える（BotのIPアドレス）
  - DNSのTTLは極端に短くしておくことで、経由したBotを特定しにくい

# Fast-Fluxの進化

- Single-Flux
  - DNSサーバが固定のもの（ISP提供？）
- Double-Flux
  - DNSサーバ機能もBotNet内で提供（中継）
- ランダム化されたURL
  - URLフィルタリングもすり抜ける
    - 同じ宣伝内容でも、アクセス先のドメインが異なる
      - http://xxx.user12345.com/login
    - 2006年夏頃から多く見られるように
- 「Rock Phish Kit」による複数サイトの偽装

# ハーベスティング

スパムメールをできるだけ多くのユーザに送信するためには、有効なメールアドレスの入手が必要

- Webからメールアドレスを収集
  - テキストで書くのは避ける（mailto:も）
  - 画像化は一つの対策（その気になれば文字認識で自動処理されてしまう）
    - コピーが面倒
  - @を“&#64”にしたり、JavaScriptで細工しても拾われるらしい
    - 参考：http://mailspam.cocolog-nifty.com/blog/
  - Code words（合い言葉が含まれていれば受け取る）や使い捨てアドレス（定期的に変更）と組み合わせるのも有効か
- メールサーバ自体からメールアドレスを収集
  - SMTP VRFY（アドレス存在確認のコマンド）
    - 今日では拒否するように設定するのが常識
  - SMTP RCPT
  - 実際にメールを送ってみてバウンスするか確認
- サービスに登録した情報の漏洩

# 送信者詐称

- 送信者アドレスのブラックリストによるフィルタリングを回避する手口
  - メールの返信が届く必要はない
    - メール本文のURLにアクセスしてもらえればOK
    - メールに添付したウィルスに感染してくれればOK
  - スパムメールが大量にバウンスした場合、詐称されたユーザに大量のバウンスメールが届くことも問題である
    - バックスキャッター
  - エラーメールに見せかけたスパムもある
- フィッシングの場合は本物に見せかけるために用いられる

# ベイジアンフィルタへの対抗

- ベイジアンフィルタ
  - 主としてスパムメールの本文を解析し、単語の出現頻度を統計的に解析することで、スパムメールかどうかを判別する手法
    - 言語ごとに処理を実装する必要がある
- スノーフレーキング
  - 関係のない単語を混ぜる等によりベイジアンフィルタを回避する手口
    - 特にHTMLの場合は、画面に表示されない部分に埋め込まれる
    - 単語内にスペースを挿入したり、Iを1にしたりといった回避方法もある
- 画像化

# まとめ（電子メール利用者の立場から）

- まず基本事項は守るとして：
  - PCのアップデートはこまめに行う
  - 怪しい添付ファイルは開かない
  - 怪しいリンクはクリックしない
    - 事前にURL（表示されていない方）を確認する
    - 最近のウィルス対策ソフトは事前にチェックしてくれるけれど．．．
  - 偽のオプトアウトにはだまされない（期待しない）
- 以下をサポートするサービスが普及するといいかも
  - 複数のメールアドレスを使い分ける
    - Tagged address (user+tag@domain)
    - アドレスごとに、どこからメールの送信元を管理できると効果的
    - 情報がどこから流出したかが把握可能
  - 公開するメールアドレスは定期的に変更
    - Code wardsが設定できる
  - アドレスの変更を必要な人にだけ通知する