迷惑メール対策手法総論

山井　成良（岡山大学）

IAjapan第8回迷惑メール対策カンファレンス

2010年5月31日 コクヨホール

Contents

- 受信対策
- 送信対策
- 送信ドメイン認証

受信対策

受信対策の性能評価基準

- 代表的な2つの評価基準
  - 見逃し(FN: false negative)率
    - 迷惑メールを通常メールと判断する割合
    - 検出率（迷惑メールを正しく判断する割合）と等価
  - 誤検出(FP: false positive)率
    - 通常メールを迷惑メールと判断する割合
- 重要なのは誤検出率
  - 見逃した迷惑メールは単に削除すればよい
  - 重要なメールが迷惑メールと判定されると影響大

望ましい受信対策

- 様々な観点が存在
  - 見逃し(FN)が少ない
  - 誤検出(FP)が少ない
  - スループットが高い
  - 遅延時間が短い
  - サービス不能(DoS)攻撃に強い
  - 外部のサービスに依存しない
  - 管理コストが低い
- どの観点が重要かは状況により異なる

代表的な受信対策

- ブロッキング・スロットリング
  - 迷惑メール受信時の対策
- フィルタリング
  - 迷惑メール受信後の対策

ブロッキング(1)

- 定義
  - 送信側IPアドレス，エンベロープFromアドレス等に基づいて迷惑メールかどうかを判定し，迷惑メールの本文を受信せず拒否する方法
- 代表的なブロッキング技法
  - IPアドレスの逆引き
  - ブラックリスト/ホワイトリスト
  - Tempfailing
  - 自動認識付きホワイトリスト

ブロッキング(2)

- IPアドレスの逆引き
  - 失敗すればペナルティ
    - 例：恒久拒否，一時拒否，応答遅延
  - 成功すればホスト名を検査
    - 例：多くの数字を含むようなホスト名なら拒否
      - 動的IPアドレスからの発信と判断
    - 例：特定の国であれば一時拒否or応答遅延

ブロッキング(3)

- IPアドレスの逆引き(続き)
  - 長所
    - 単純で効果的
  - 短所
    - 誤検出率が高い
      - PTRレコードがない正当なMTAもかなり多い
      - ネットワークやDNSサーバのトラブルで受信拒否
    - 本来であれば不要な通信が発生
      - 特にパラノイド検査(正引きとの整合性検査)実施時

ブロッキング(4)

- ブラックリスト(DNSBL: DNS Black List)
  - 迷惑メール発信ホスト，不正侵入ホスト，迷惑メールに含まれるURL等を登録
  - 代表例
    - Spamhaus ZEN (http://www.spamhaus.org/zen)
    - SpamCop SCBL (http://www.spamcop.net/bl.shtml)
    - SORBS (http://www.us.sorbs.net/)
    - ORDB (http://ordb.org/) ※2006年12月サービス中止
  - 使用例(Spamhaus ZENの場合)
    - IPアドレスがA.B.C.DのMTAからSMTP接続
    - D.C.B.A.zen.spamhaus.orgのAレコードを検索
    - Aレコード(127.0.0.x)が得られれば，接続を拒否

ブロッキング(5)

- ブラックリスト（続き）
  - トラブルも多い
    - 登録ホストからは通常メールも（ある日突然）拒否
    - 対策完了後も復旧に時間を要するものもある
    - 一部は訴訟にまで発展
  - 効果も疑問（CEAS 2006の論文†における調査）
    - 登録ホストはbot感染ホストの6%程度
    - 検出後に直ちに登録されるホストは少ない

† A. Ramachandran, *et al.*: Can DNS-Based Blacklists Keep Up with Bots? http://www.ceas.cc/2006/14.pdf

ブロッキング(6)

- ホワイトリスト (DNSWL: DNS White List)
  - 信頼できるホストを登録
    - 評価(reputation)サービス
      - 例：Sender Score(Return Path社)
    - 認定(accreditation)サービス
      - 例：Sender Score Certified(Return Path社)

ブロッキング(7)

- Tempfailing
  - 「迷惑メール発信MTAは再送をしない」との仮説に基づく方法
    - 通常MTAは信頼性重視
    - 迷惑メール発信MTAは配送効率重視
  - 一時的に受信を拒否
    - 再送されれば受信
  - 代表例
    - お馴染みさん方式
    - Greylisting

ブロッキング(8)

- Tempfailing(続き)
  - 利点
    - かなり効果的（80%程度排除）
  - 欠点
    - 配送遅延が結構大きい
      - 再送まで1時間のものもある
      - 別MTAからの再送も一時拒否
      - 再送間隔が短すぎるものは再送と見なされないことも
    - 誤検出（再送しない通常MTA）も多い
      - 一部のファイアウォール・オンライン予約システムなど
    - ホワイトリスト（除外MTAリスト）の管理が必須

ブロッキング(9)

- 自動認証つきホワイトリスト(challenge&response)
  - プログラム（bot）が迷惑メールを送信する点を利用
    - 初めての相手には再送要求メッセージを返送
    - 再送されればホワイトリストに登録して配送
  - 長所
    - 人間相手には非常に効果的
  - 短所
    - 送信元が正当なプログラムな場合には適用不可
      - オンライン予約システムなど
    - 第三者(詐称された送信者)に再送要求メッセージを配送
      - バウンスメール（エラーメール）による攻撃と同じ

スロットリング(1)

- 定義
  - 通信速度などを意図的に低下させることにより，迷惑メールの大量送信を妨害する方法
- 代表的なスロットリング技法
  - 同時接続数・確立頻度・帯域の制限
  - 配送不能宛先数の制限
  - Tarpitting

スロットリング(2)

- 同時接続数・接続頻度・帯域の制限
  - サービス不能(DoS)攻撃に対する防御
  - Bot等からの大量配送の防止
- 配送不能宛先数の制限
  - 一部の迷惑メールに対して効果的
  - アドレス収集の防止

※ いずれも一部の正常メール配送（特にメーリングリスト）に影響

スロットリング(3)

- Tarpitting
  - 意図的に応答を遅延
    - 迷惑メール送信側でのタイムアウトを誘発
    - あるいは配送効率を抑制
  - ブラックリスト/ホワイトリストとの併用が多い
    - ブラックリスト登録MTAに対して遅延挿入など
  - 代表的な技法
    - Greet pause
    - TCP damping

スロットリング(4)

- Greet pause
  - コネクション確立時の応答(220 ...)を遅延
    - RFC5321では送信側は5分間待つべきと規定
    - 多くの迷惑メール送信MTAは15秒程度で切断
  - MAIL/RCPTの応答を遅延する方法も
    - 例: 宛先不明の場合には遅延挿入
  - 応答を待たずに送信するMTAも拒否
    - 本来はPIPELININGが指定されている場合のみ可

スロットリング(5)

- Greet pause（続き）
  - 長所
    - Tempfailingより設定が簡単
      - 再送かどうかの判定が不要
    - 配送遅延が小さい
    - 誤判定が少ない
  - 短所
    - 性能はtempfailingのほうがよい(?)
      - 併用の場合，greet pauseで拒否できず tempfailingでは拒否できるものがかなりある
    - サービス不能攻撃に弱い

スロットリング(6)

- TCP damping†

† K. Li, C. Pu, M. Ahamad: Resisting SPAM Delivery by TCP Damping, http://www.ceas.cc/papers-2004/191.pdf

  - セッション中での迷惑メール判定
    - SMTPコマンド引数，ヘッダ情報などを判定に利用
  - TCPレベルで遅延挿入
    - ACKの送信を遅延
    - 広告ウィンドウサイズを減少
    - 輻輳状態を偽装
  - 広範囲な遅延時間の調整が可能

フィルタリング(1)

- 基本方針
  - メール受信後に迷惑メールかどうかを判断
  - 迷惑メールは削除あるいは別に格納
- 代表的な方法
  - ルールベースフィルタ
  - ベイジアンフィルタ
  - 分散協調フィルタ(シグネチャベースフィルタ)

フィルタリング(2)

- ルールベースフィルタ
  - 迷惑メールの特徴をルールとして記述
    - 単純なパターンマッチング
      - 本文中に「$」「Viagra」など特定のキーワードを含む
    - ヒューリスティック
      - 長い英単語がある，FromとToが同じアドレスなど
  - マッチした場合，ルールに対応したスコアを加算
  - 一定のスコア以上のものを迷惑メールと判定
  - 欠点＝柔軟性の欠如
    - スコアの調整は可能だが限界が存在
    - 新たな手口には新たなルールが必要
    - 誤検出が比較的多い

フィルタリング(3)

- ベイジアンフィルタ(Bayesian filter)
  - キーワード(単語，3字組等)の出現率を学習
  - キーワードの種類に応じて迷惑メールを判定
  - ベイズ則 P(A|B)= P(A)P(B|A)/P(B)を利用
    - 事象A･･･メッセージが迷惑メールである
    - 事象B･･･メッセージがキーワードを含む
  - 有効なキーワードの例
    - ff0000 … HTMLメールにおける赤色指定
  - 新しい手口にもある程度対応可能
    - 但し，学習が必要
    - 最近は対応できないような回避策がいろいろ使われている

フィルタリング(4)

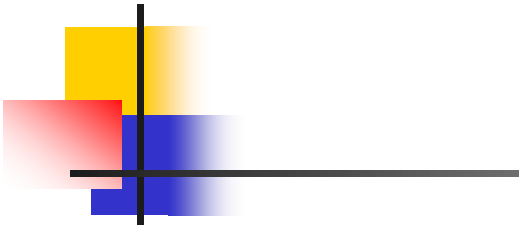
- 分散協調フィルタ(シグネチャベースフィルタ)
  - 判定済みの迷惑メールの再受信を排除
    - 同一内容の迷惑メールが大量配送される点を逆利用
  - 利用者が迷惑メールをデータベースに登録
  - メール受信時に同一メッセージの存在を問合せ
    - 一定数以上の登録があれば迷惑メールと判定
  - 迷惑メールの認識率が低い点が問題
    - 大量の迷惑メールを登録する必要あり
    - 登録までのタイムラグあり
    - 内容の一部変更に弱い ⇒ URIブラックリストの活用

フィルタリング(5)

- Spammer側のフィルタリング回避策
  - 十分にフィルタリング技法を研究
    - 単語の加工/挿入
    - 背景と同じ色での単語埋込み
    - 一部のWebサイトが提供するredirect機能の利用
    - サーチエンジン検索URLの埋込み
      - 検索結果の先頭に誘導先URLが表示されるようなリンク
    - ファイルへの埋込み（PDF, MS Word等）
    - 画像ファイルの添付+宛先毎の変形

フィルタリング(6)

- フィルタリング側での対策
  - 複数の技法の組合せ
    - 多いものでは10種類の技法を利用
  - フィルタリング技法の秘匿化
    - 迷惑メール送信者に回避策のヒントを与えない
  - ハニーポットの活用
    - おとりのアドレスに迷惑メールを誘導
    - ゾンビPCのハニーポットを仕掛けることも



送信対策

代表的な送信対策

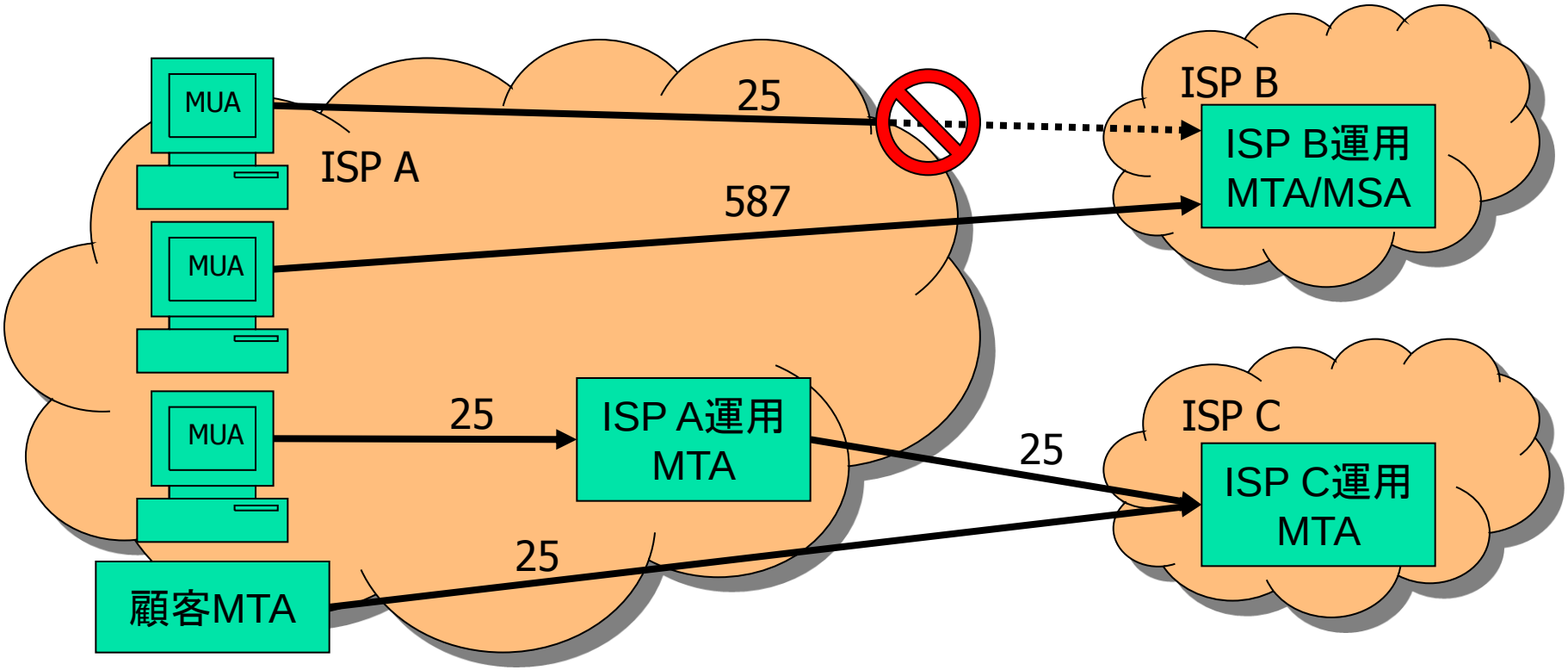
- ISPでのブロッキング
  - 迷惑メールに限らず送信を原則的に禁止
- ISPでのスロットリング
  - 迷惑メールに限らず大量送信を抑制
- 送信者認証
- 法的対策

ISPでのブロッキング(1)

- Outbound Port 25 Blocking (OP25B)
  - 自網からの迷惑メール送信防止が目的
    - 迷惑メール配送業者やbotが対象
    - 普通の電子メール発信は対象外
  - 方法
    - 自網→外部MTAへのSMTP(25番)をブロック
    - 他社MTAの利用者には発信ポートの利用を推奨
      - Submission(587番), SMTP/SSL(465番)
    - 一般利用者は自社ISP運用のMTAを利用
    - 自網内の顧客MTAは固定IPアドレスで対応
      - 当該IPアドレスのみブロックを解除

ISPでのブロッキング(2)

- Outbound Port 25 Blocking (続き)



ISPでのブロッキング(3)

- Outbound Port 25 Blocking (続き)
  - 既に多くのISPが導入
  - 十分な効果
    - 国内宛迷惑メールの送信拠点が国外に移行
  - 問題点
    - 発信ポートを提供していない組織もまだ多い
      - 特に大学，中小企業が問題かも

ISPでのスロットリング

- 外部MTAに対するメール発信を制限
  - 同時送信数・送信頻度・帯域などを制限
    - 基本的には受信対策の場合と同じ
  - OP25Bとの併用
    - 自網内からのメールの大量送信を直接的にも間接的にも防止

送信者認証(1)

- 目的
  - 第三者による不正発信の拒否
  - 問題発生時の発信者特定
- 送信者アドレスの正当性は対象外
  - 他のISPから発信する場合などを考慮
  - 利用者レベルでデジタル署名を利用可能
    - PGP (Pretty Good Privacy), S/MIMEなど
  - 送信者アドレスの正当性保証も可能
    - 強制的にSender:ヘッダを挿入/置換するなど

送信者認証(2)

- POP before SMTP
  - 前提条件
    - POPサーバと発信用MTAが同じ
  - 方法
    - 受信時に先立ってPOPで認証
    - 認証に成功したIPアドレスを一定時間（例えば10分）登録
    - 登録IPアドレスからは任意の宛先への配送を許可
  - 利点・欠点
    - MUAを選ばない
    - IPアドレスが同じMUA/MTAから他の利用者も発信可能
      - 特にNATを利用するISPから発信する場合に問題

送信者認証(3)

- SMTP-AUTH
  - SMTPの拡張(RFC2554⇒RFC4954)
    - 新しいコマンドAUTHの追加
    - メール送信時に利用者を認証
    - いくつかの認証方法がサポート(SASL:RFC4422)
      - CRAM-MD5, DIGEST-MD5, PLAIN, LOGIN, etc.
  - 対応したMUAが必要
    - 最近は殆どのMUAがどれかの認証方式に対応

法的対策(1)

- 技術的な迷惑メール対策の限界
  - ブロッキング・フィルタリングでは
    迷惑メール発信者は不利益を被らない

  ⇒何らかの法的な対策が必要
- 法的な迷惑メール対策の実施国
  - 日本
  - アメリカ合衆国
  - EU
  - オーストラリア
  - 韓国など

法的対策(2)

- 日本における法律
  - 最初の迷惑メール対策法(2002/7施行)
    - 特定商取引に関する法律の一部を改正する法律
      （特定商取引法）
    - 特定電子メールの送信の適正化に関する法律
      （特定電子メール法）

※ 広告メールを全面的に禁止するものではない
（オプトアウト方式）

- 広告は企業活動にとって必要
- CM, ダイレクトメールとの比較

法的対策(3)

- 迷惑メール対策法の比較

法律名	特定商取引法	特定電子メール法
担当官庁	経済産業省	総務省
規制対象	事業者	メール発信者
表示義務	1. メールアドレス 2. 未承諾広告※ 3. オプトアウト方法	1. 未承諾広告※ 2. 氏名・住所 3. 発信アドレス 4. 受信アドレス
禁止事項	拒否者への送信	・拒否者への送信 ・架空アドレスへの送信
罰則	・2年以下の懲役 ・300万円(法人は3億円)以下の罰金	50万円以下の罰金

法的対策(4)

- 迷惑メール対策法の効果
  - 殆どの広告メールは表示義務に違反
    - 違反者の調査が困難
      - 発信者情報の欠落
      - 多くの場合，ゾンビPCから発信 ⇒ 追跡が困難
    - 違反しても直ちには処罰されない
      - 措置命令に違反した場合に初めて罰金・懲役
  - 2003/10/9に初めて2社が行政処分
    - 件名に「未承諾広告※」「※未詳諾広告※」などと表示
    - 2002/8頃から2003/9頃まで発信
    - 2003/6以降は送信者情報表示義務にも違反

法的対策(5)

- 迷惑メール対策法の効果(続き)
  - 総務省・経済産業省の合計で10件程度
    - 迷惑メール発信で初の逮捕者(2005/5/16)
      - 容疑は「有線電気通信法」違反
      - メールサーバに過負荷を与えたため
    - 迷惑メール発信で初の業務停止命令(2005/6/14)
      - 同一人物が運営する2社
      - 特定商取引法違反（表示義務違反）

法的対策(6)

- 迷惑メール対策法の改定(2005/11)
  - 送信者情報詐称禁止
  - 直罰規定（詐称の場合）
  - 刑罰の引上げ（特定電子メール法）
    - 50万円以下の罰金
      ⇒1年以下の懲役または100万円以下の罰金
  - 対象の拡大
    - 私用アドレスに加えて事業用アドレスも対象

法的対策(7)

- 改定迷惑メール対策法の効果
  - 行政処分は計4件
  - 警察による摘発が増加
    - 2006/5/25 初の逮捕者（千葉県警）
      - 懲役8か月・執行猶予3年・罰金80万円
    - 2006/8/3 3名が書類送検（大阪府警）
      - 罰金100万円×1，罰金50万円×1
    - 2007/1/16 4名逮捕（千葉県警）・・・タクミ通信
      - 8か月・4年×2，6か月・5年×1，6か月・3年×1
    - 2008/2/15 1名逮捕（警視庁）
      - 懲役6か月・執行猶予3年

法的対策(8)

- 迷惑メール対策法の再改定(2008/12)
  - オプトイン方式の導入
    - 受信者の同意のない広告・宣伝メールの送信禁止
    - 直罰規定あり（特定商取引法）
  - 刑罰の強化（法人に対する罰金）
    - 100万円以下から3000万円以下に引き上げ
  - 外国執行当局との連携強化
    - 送信者の情報提供が可能に
    - 外国からの迷惑メール送信への対応
    - 海外の業者への送信委託も規制対象

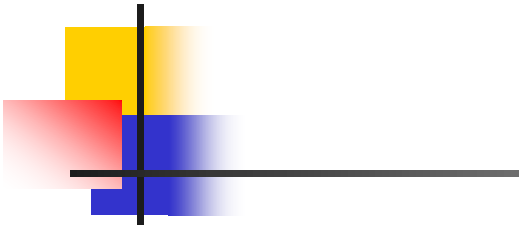
法的対策(9)

- 再改定迷惑メール対策法の効果
  - 計11件の行政処分
    - いずれもオプトイン規制違反
    - 2009/9以降は消費者庁との共管
    - 発表直後は日本語の迷惑メール（主に出会い系）が一時的に減少
  - 警察による摘発は未だなし

法的対策(10)

- 罰則は適切か
  - タクミ通信事件(2007/1/16 4名逮捕)の場合
    - 出会い系サイトを運用
    - 2ヶ月間で54億通
    - 1か月の売上高1億2000万円
    - 判決
      - 懲役8月執行猶予4年×2
      - 懲役6月執行猶予5年，懲役6か月執行猶予3年

    ※ 罰金はなし!!



送信ドメイン認証

送信ドメイン認証(1)

- 発信者ドメインの詐称を識別する手段
  - ローカルパートの詐称は対象外
    - 必要なら送信者認証を活用
  - メッセージの中身も対象外
    - 迷惑メールを受け取ることもあり得る
- 問題発生時の追跡が目的
  - Spamメールを受け取ったときの苦情先
  - フィッシング詐欺の抑止

送信ドメイン認証(2)

- 2種類の方法
  - IPアドレスに基づく認証
    - SPF 1.0 (classic) ･･･ RFC4408
    - Sender ID = SPF 2.0 + Caller ID ･･･ RFC4406
      - SPF (Sender Policy Framework) ･･･ POBOX
      - Caller ID ･･･ Microsoft
  - デジタル署名を利用した認証
    - DKIM = DomainKeys + IIM ･･･ RFC4871
      - DomainKeys (RFC4870)･･･ Yahoo!
      - IIM (Identified Internet Mail) ･･･ Cisco Systems

送信ドメイン認証(3)

- Sender ID(1)
  - 3種類の要素により構成
    - ヘッダ内の送信者の認証(PRA)
    - エンベロープFromの認証(MFROM)
    - 送信側ドメインのポリシー定義(SPFレコード)

送信ドメイン認証(3)

- Sender ID(2)
  - PRA (Purported Responsible Address)
    - RFC4407で規定
    - 責任があるとされるアドレス
      - Resent-Sender:, Resent-From:, Sender:, From:の順
      - 転送する場合には，Resent-From:などを追加
    - MAILコマンドのSUBMITTERオプションも利用可能
      - RFC4405でSMTPサービス拡張として導入
      - 本文を受け取る前に判定するため

例：MAIL FROM: <alice@example.com> SUBMITTER=<alice@example.jp>

送信ドメイン認証(4)

- Sender ID(3)
  - SPFレコード
    - DNSのTXT (SPF)レコードで送信サーバを宣言
      - +　　pass (受信許可)
      - ?　　neutral (宣言なしと同様)
      - ~　　softfail (neutralとfailの中間)
      - -　　fail (受信拒否)
    - 例：AレコードかMXレコードに対応するIPアドレスを持つMTAからのみ送信可能な場合
      - example.jp　IN TXT　“v=spf1 +a +mx –all”
      - example.jp　IN SPF　“spf2.0/mfrom,pra +a +mx -all”

送信ドメイン認証(5)

- Sender ID(4)
  - Sender IDに関する話題
    - MicrosoftがPRAに対して知的所有権を主張
      - 特許料は取らないがライセンス契約が必要など
    - IETFでのワーキンググループが余波を受け解散
      - MARID (MTA Authorization Records in DNS) WG
    - RFCにも影響
      - 強制力のあるStandardではなくExperimentalに

送信ドメイン認証(6)

- DKIM (DomainKeys Identified Mail)
  - 公開鍵暗号方式を利用
    - 送信側
      - 秘密鍵を使って署名
    - 受信側
      - DNSを用いて公開鍵を取得
      - 公開鍵を使って署名を検証

送信ドメイン認証(7)

- DKIM (続き)
  - 署名つきヘッダの例

```
DKIM-Signature: v=1; a=rsa-sha256; s=brisbane; d=example.com;
                                  ↑アルゴリズム  ↑セレクタ    ↑ドメイン
      c=simple/simple; q=dns/txt; i=joe@football.example.com;
          ↑正規化方法          ↑公開鍵入手法   ↑ユーザ名
      h=Received : From : To : Subject : Date : Message-ID;
          ↑ 署名対象に含めるヘッダフィールド ↓本文のハッシュ値
      bh=2jUSOH9NhtVGCQWNr9BrIAPreKQjO6Sn7XIkfJVOzv8=;
          ↓署名
      b=AuUoFEfDxTDkHlLXSZEpZj79LICEps6eda7W3deTVFOk4yAUoqOB
        4nujc7YopdG5dWLSdNg6xNAZpOPr+kHxt1IrE+NahM6L/LbvaHut
        KVdkLLkpVaVVQPzeRDI009SO2Il5Lu7rDNH6mZckBdrIx0orEtZV
        4bmp/YzhwvcubU4=;
```

送信ドメイン認証(8)

- DKIM (続き)
  - DNSの設定例（続き）

```
brisbane._domainkey.example.com.   IN TXT   (
          "v=DKIM1; p=MIGfMA0GCSqGSIb3DQEBAQUAA4GNADCBiQ"
          "KBgQDwIRP/UC3SBsEmGqZ9ZJW3/DkMoGeLnQg1fWn7/zYt"
          "IxN2SnFCjxOCKG9v3b4jYfcTNh5ijSsq631uBItLa7od+v"
          "/RtdC2UzJ1lWT947qR+Rcac2gbto/NMqJ0fzfVjH4OuKhi"
          "tdY9tf6mcwGjaNBcWToIMmPSPDdQPNUYckcQ2QIDAQAB"
          )
```

送信ドメイン認証(9)

- 2つの認証方式の選択
  - IPアドレスに基づく認証
    - ヘッダや本文の書換えに強い
    - 転送に弱い
      - PRA，MFROMが維持できるかどうかが問題
  - デジタル署名を利用した認証
    - 転送に強い
    - ヘッダや本文の書換えに弱い
  - ⇒相補的に利用することが重要

送信ドメイン認証(10)

- 普及後の問題点
  - 迷惑メール送信者は送信ドメイン認証に対応
    - 単に認証するだけでは問題
  - 認証後の判定が重要に
    - 認定(accreditation)サービス
      - 信頼のある機関に公的に認定してもらう
    - 評価(reputation)サービス
      - Spamメールを大量に発信すると評価が下がる
    - 新規（使い捨て）ドメインの評価が問題

おわりに(1)

- 次々に出現する新たな手口
  - フィルタリング対策
    - 画像
      - サイズ/色/位置/傾きなどの変更
    - PDF/Word/Excel等のファイル添付
      - 単純なパターンマッチングの回避
      - 暗号化されているものも存在
    - DNSBL回避
      - ランダムなサブドメインの利用
      - IPアドレスによるサーバ指定
  - 無料電子メールアカウントの悪用
    - CAPTCHA破りが横行
      - 機械解読
      - エログリッドコンピューティング

おわりに(2)

- 次々に出現する新たな手口（続き）
  - SPF対策
    - ドメインの新規取得+SPFレコードの設定
- 社会的なイベントの悪用
  - 大地震への便乗
    - UNICEFを騙った募金活動
    - 映像を見るとマルウエアに感染
  - 自動車の大規模リコールへの便乗
    - リコール情報へのアクセスでマルウエアに感染

⇒ いたちごっこはまだまだ続く．．．