# TCPフィンガープリントによる悪意のある通信の分析

早稲田大学大学院　基幹理工学研究科
木佐森幸太、下田晃弘、森達哉、後藤滋樹

# 研究の背景

- ボットの脅威の拡大、検出の難しさ

- カーネルマルウェアの増加
  - 最も権限の高いレベル（Ring0）で動作し、メモリ、CPU 命令、すべてのハードウェアデバイスへのフルアクセスが可能なマルウェア
  - すべてがカーネルモードドライバで実装され，コードのすべてが Ring0 で実行されるものをフルカーネルマルウェア(FKM) と呼ぶ
  - FKMは既存OSのTCP/IP実装とは異なる独自のネットワークドライバを実装

※Ring0：CPU の動作モードの中で最も高いレベル。一般にOSのカーネルはRing0で動作する。

# 研究の目的

- TCPヘッダを分析することにより、FKM（フルカーネルマルウェア）の可能性がある感染ホストを検出する手法を提案 する

- 今回は、CCC DATAsetからFKMの可能性があるシグネチャを抽出し、その有効性をCCC DATAset およびその他の実計測データで示す

※CCC = Cyber Clean Center

# 分析手法

- FKMは既存OSとは異なる独自のネットワークドライバを実装しているため、TCP/IP ヘッダを分析することで識別できるケースがある
  - cf. The Rise and Fall of Reactor Mailer http://projects.csail.mit.edu/spamconf/SC2009/Henry_Stern/
- CCC DATAsetの攻撃通信データを分析することで、既存OSと異なる実装による（FKMの可能性がある）通信を抽出、分析する
- 分析の手段としてPassive TCP fingerprintingを用いる

# Passive TCP fingerprinting

- TCP/IP の仕様はRFC で定義されているが，OS 毎にその実装は異なる
- fingerprintingとは、通信の特徴から対象システムのOSを推定する技術
  - active：対象システムに対して通信を行い、得られたデータからOSを推定する（nmapが有名）
  - passive：対象システムに対して通信を行わず、取得済みの通信データを分析してOSを推定する
- 今回はp0fというツールを用い、Passive fingerprintingを行った

# p0f

- p0fにはいくつかのモードがあるが、今回用いたのはSYNパケットを分析対象とするモード
- 判定に用いるのは以下のデータ
  - ウィンドウサイズ
  - TTL の初期値
  - Don't Fragment ビット
  - SYN パケット全体のサイズ
  - TCP オプション（NOP、EOL、ウィンドウスケール、最大セグメントサイズ、SACK、タイムスタンプ等）
  - その他特徴的な点など
- これらを集約し、シグネチャとしてデータベース化している

# CCC DATAsetの分析

- 分析対象：CCC DATAset 2008，2009の攻撃通信データ

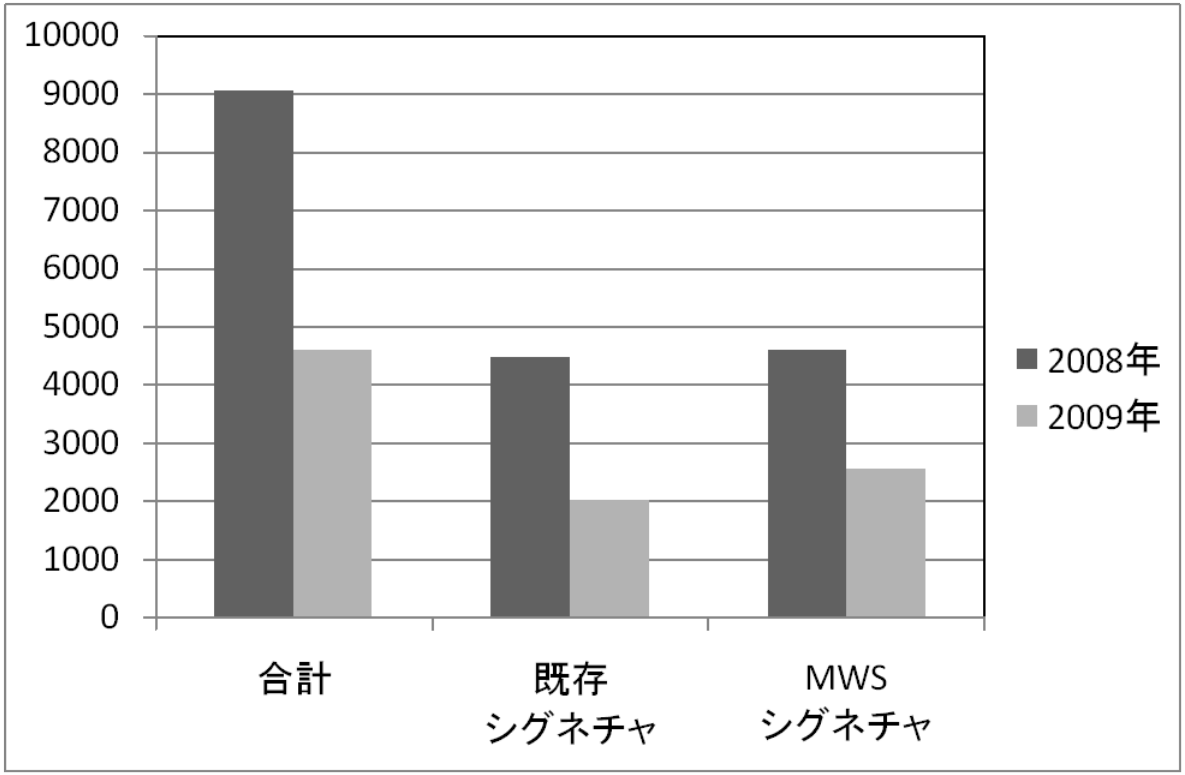
- p0fを適用した結果、既存のOSではない、UNKNOWNであると判定されたシグネチャが多数得られた
  - TTLを2の累乗の値に切り上げ、初期値として推定、集約
  - アウトバウンド1種、インバウンド43種のシグネチャが得られた
- UNKNOWNのシグネチャを総称して、MWSシグネチャと呼ぶこととする
- 以後、インバウンドの通信のみを分析対象とする

# MWSシグネチャの例

- 60352:64:0:52:M1240,N,W2,N,N,S:.:MWS:60352_I
  - ウィンドウサイズが60352バイト
  - TTLの初期値が64
  - Don't Fragmentビットが0
  - SYNパケット全体のサイズが52バイト
  - 以下のオプションが設定されている
    - 最大セグメントサイズ、NOPオプション、ウィンドウスケールオプション、SACK
  - その他の特徴はなし
  - OSの名称、詳細（バージョン等）
    - 今回は、ウィンドウサイズの値により分類し、名称を付けた

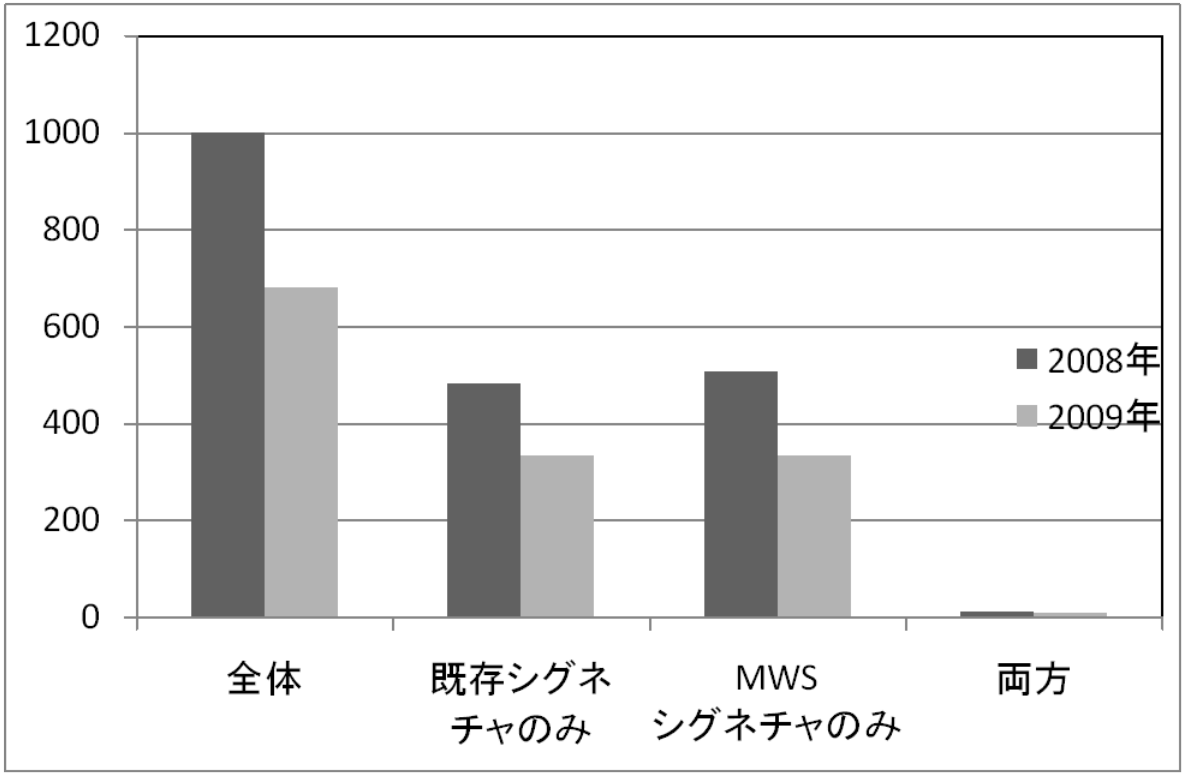
# CCC DATAsetの分析：全体(1)

- 既存シグネチャ、MWSシグネチャのSYNパケット数の比較
  - 2009年度は2008年度に比べ全体的に通信量が減少している
  - 両年度とも、SYNパケットの半分以上がMWSシグネチャによるもの



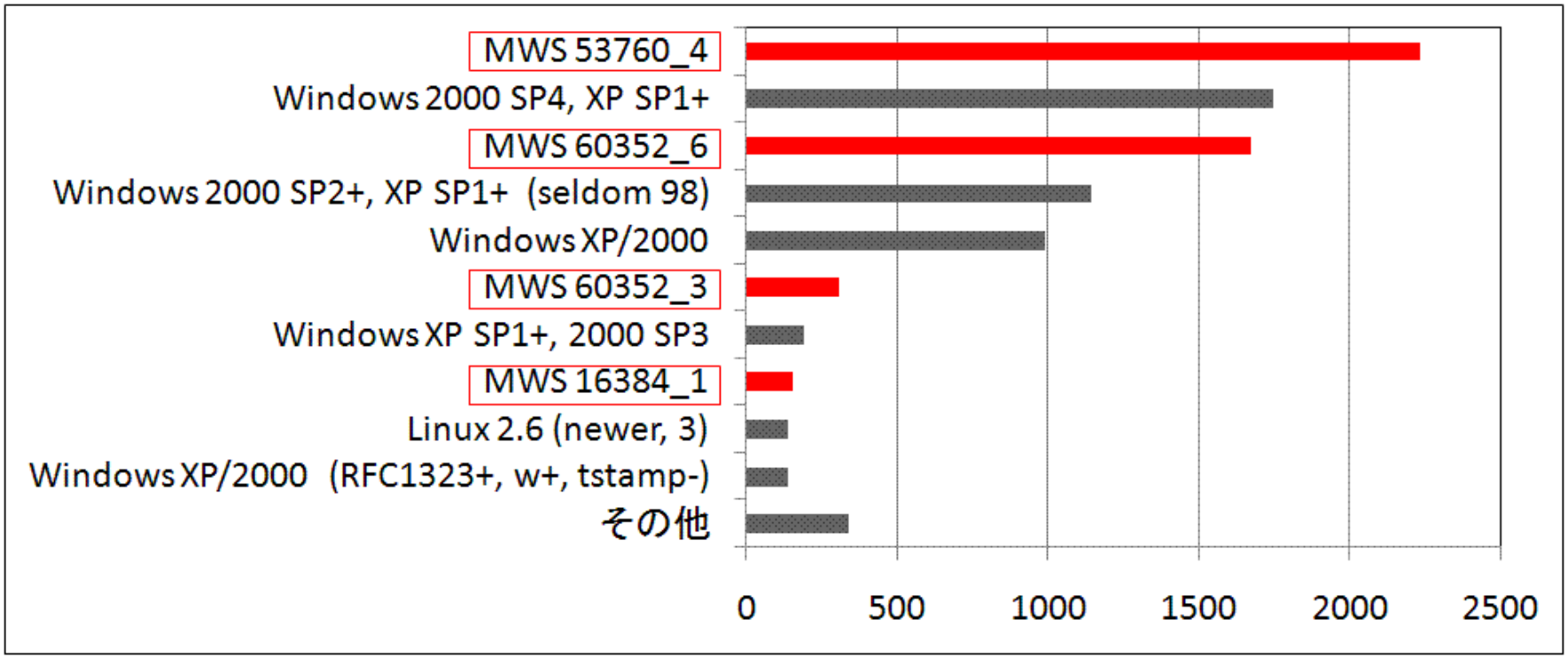
# CCC DATAsetの分析：全体(2)

- シグネチャ毎の送信元IPアドレス数
  - SYNパケット数ほどではないが全体的に減少している
  - 送信元ホストの半分以上がMWSシグネチャによる通信をしている



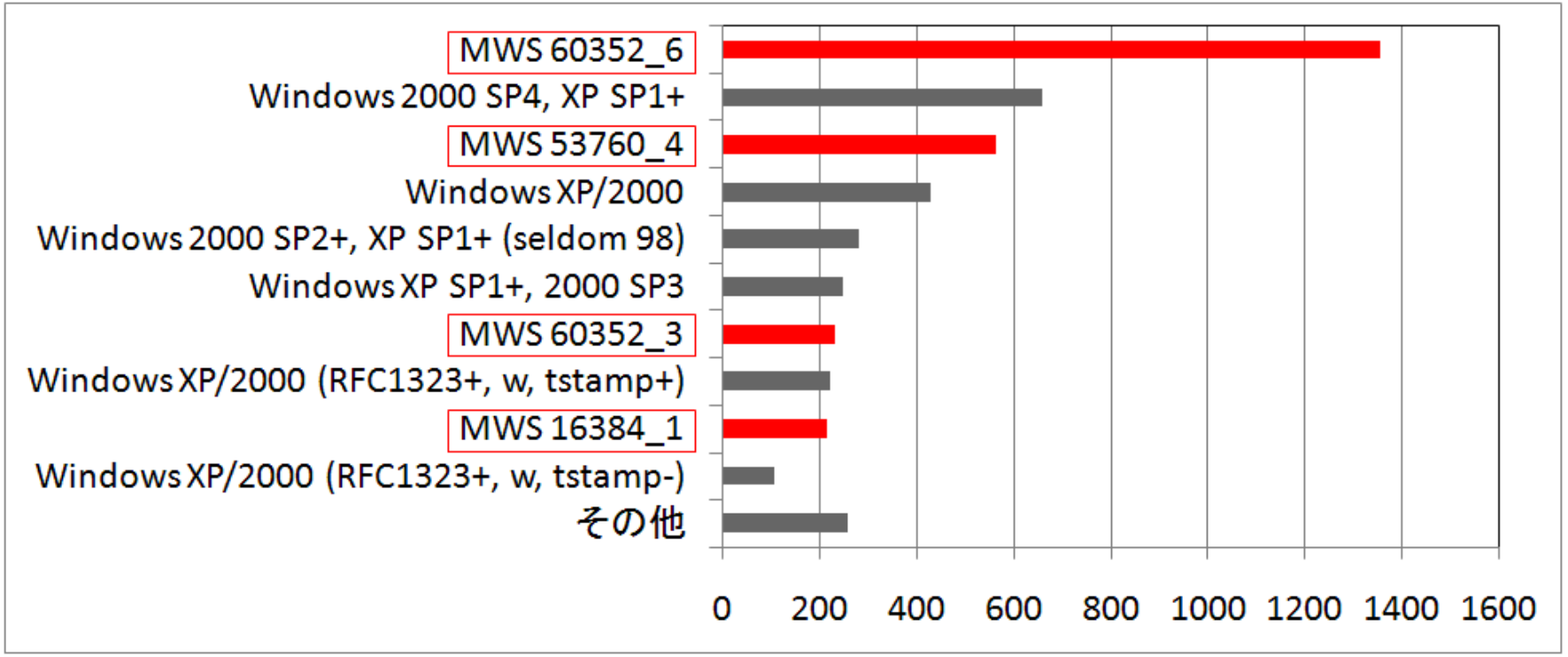
# CCC DATAsetの分析：シグネチャ別SYNパケット数(1)

- シグネチャ毎のSYNパケット数(2008)



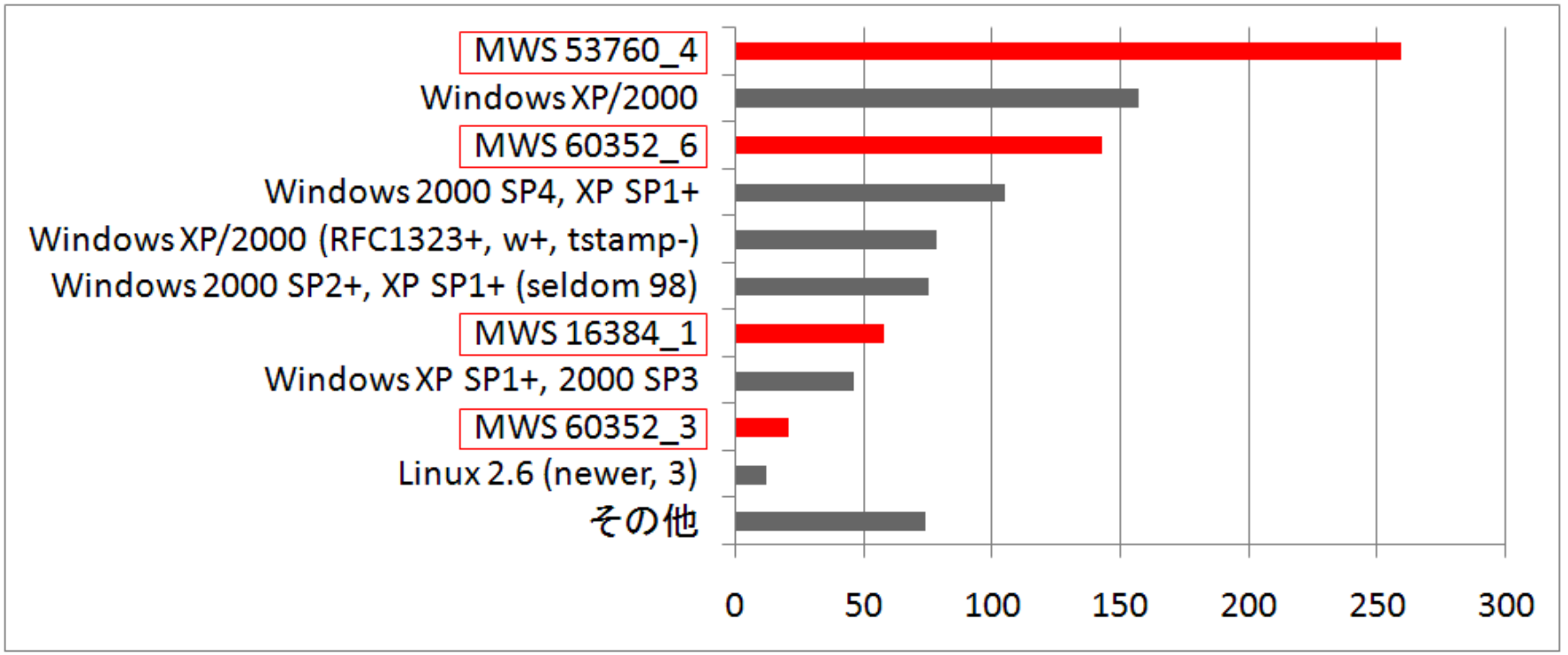
# CCC DATAsetの分析：シグネチャ別SYNパケット数(2)

- シグネチャ毎のSYNパケット数(2009)



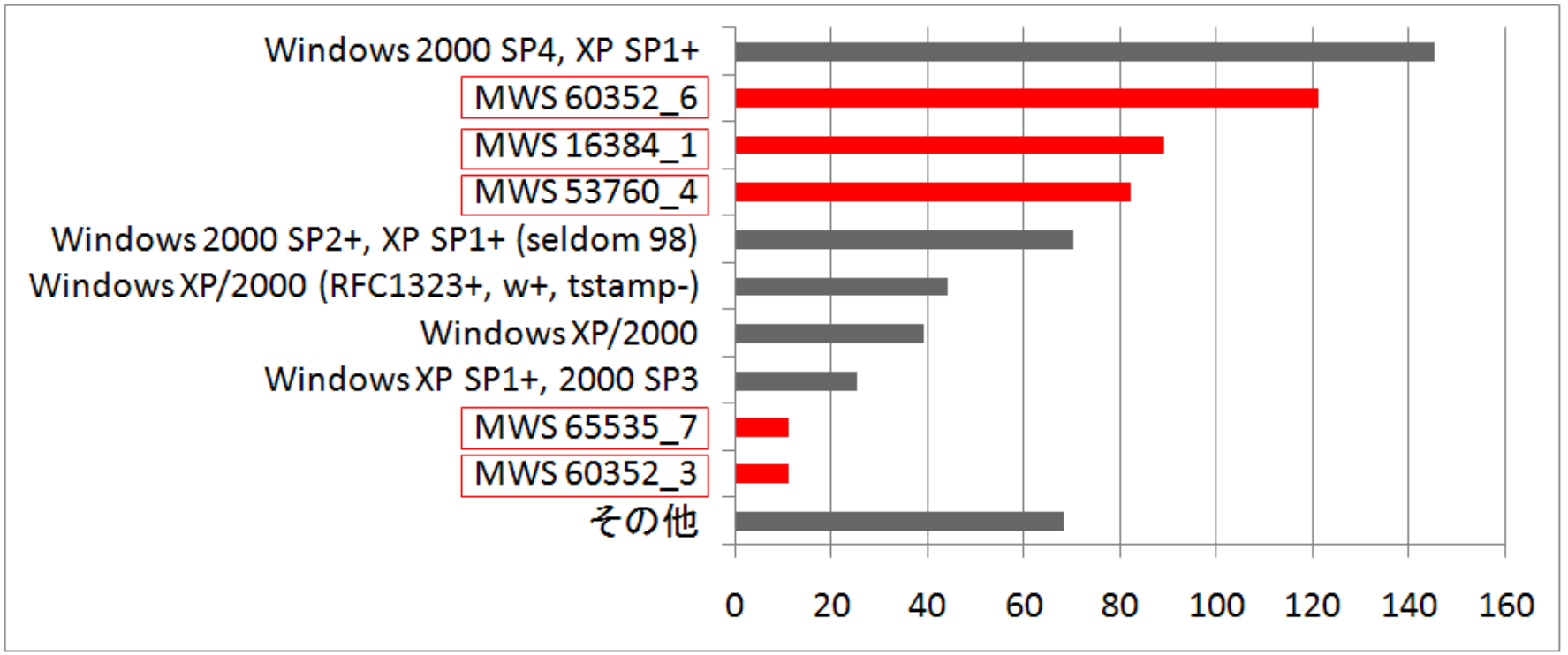
# CCC DATAsetの分析：シグネチャ別送信元IP数 (1)

- シグネチャ毎の送信元IP数(2008)

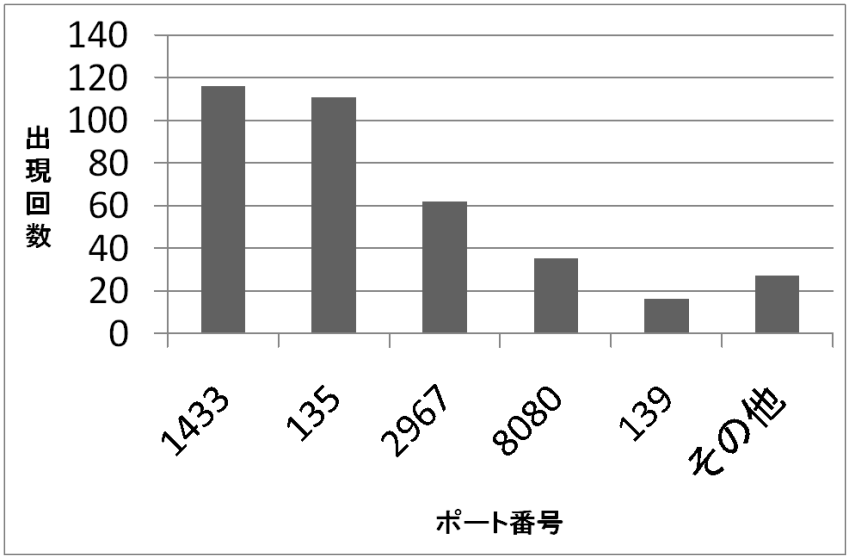


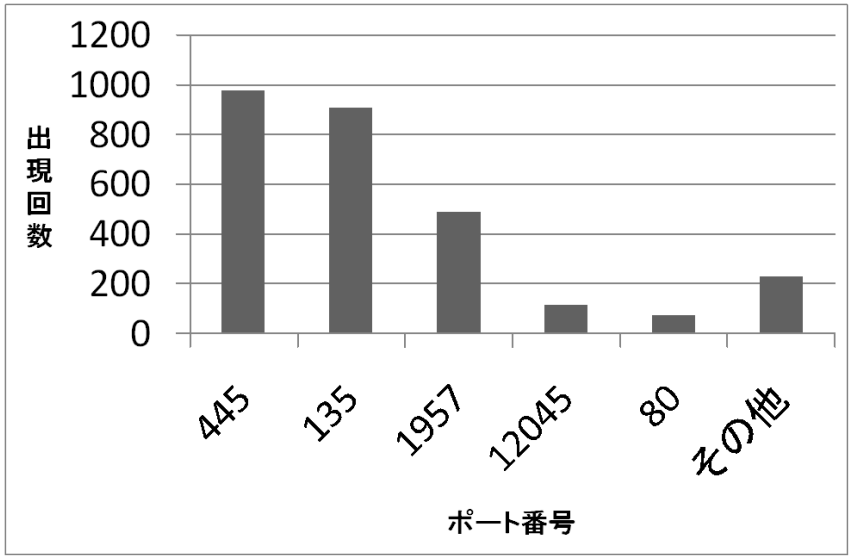
# CCC DATAsetの分析：シグネチャ別送信元IP数 (2)

- シグネチャ毎の送信元IP数(2009)



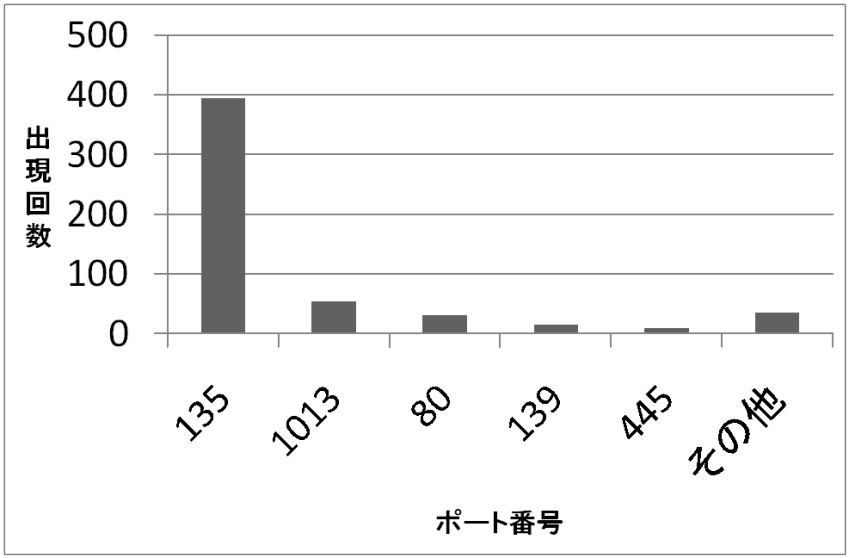
# CCC DATAsetの分析：MWSシグネチャの送信先ポート(2009) (1)

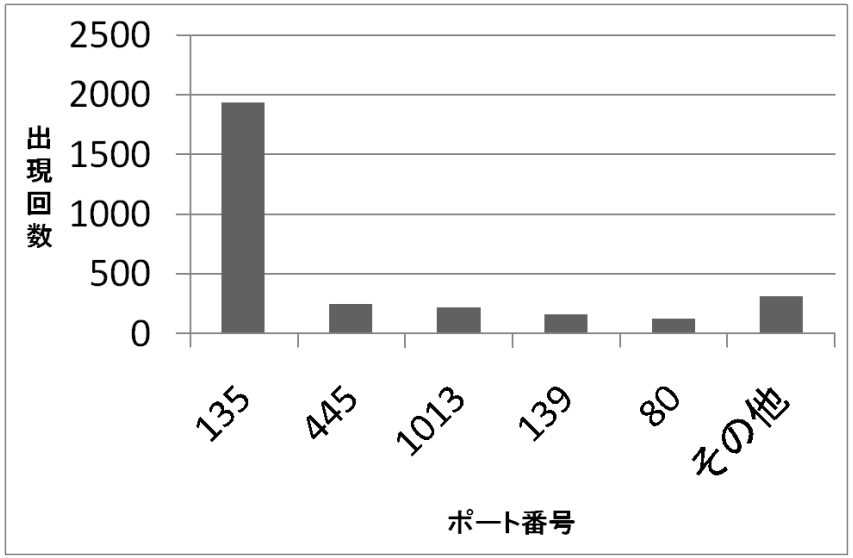



- 両年度を通じて登場頻度の高いシグネチャについて、2009年のデータにおける送信先ポートを分析
- 左：16384_1、右：53760_4

# CCC DATAsetの分析：MWSシグネチャの送信先ポート(2009) (2)




- 左：60352_3、右：60352_6
- 135、139、445、1433、2967の各ポートについては、それぞれに応じた脆弱性があることが知られている

# CCC DATAsetの分析：
# 頻度の高い通信パターンの分析(1)

- ホストAはMWSシグネチャ60352_6で通信
  - 以下SYNパケットのみ

21:26:41 ホストA:9109 -> ハニーポットA:135 (scan)
21:26:41 ホストA:9110 -> ハニーポットA:135 (rpc)
21:26:43 ホストA:9197 -> ハニーポットA:135 (rpc)
21:26:43 ホストA:9203 -> ハニーポットA:1013 (シェルコード送信)
21:26:43 ハニーポットA:1028 -> ホストA:3450 (malware 要求)
21:26:43 ハニーポットA:1028 -> ホストA:3450 (malware 要求)

※各行冒頭はタイムスタンプ

# CCC DATAsetの分析：
# 頻度の高い通信パターンの分析(2)

- ホストBはMWSシグネチャMWS 53760_4で通信

  00:35:11 ホストB:56101 -> ハニーポットB:135 (rpc)
  00:35:13 ハニーポットB:1027 -> ホストB:47602 (malware 要求)
  00:35:13 ハニーポットB:1027 -> ホストB:47602 (malware 要求)

- CCC DATAsetの攻撃元データに、ダウンロードが成功した記録が残されている

  2009-03-13 00:35:13, ハニーポットB,1027, ホストB,47602,
  TCP,c925531e659206849bf7**********,
  PE_VIRUT.AV,C:¥WINNT¥system32¥csrs.exe

# CCC DATAsetの分析：頻度の高い通信パターンの分析(3)

- ホストCは、最初のSYNパケットのみMWSシグネチャ16384_1で通信
  - 2つ目・3つ目のSYNパケットはWindowsのシグネチャであった

00:57:09 ホストC:6000 -> ハニーポットB:135 (scan)
00:57:13 ホストC:3197 -> ハニーポットB:135 (rpc)
00:57:15 ホストC:4139 -> ハニーポットB:135 (rpc)

# CCC DATAsetの分析：シグネチャ毎の通信内容

- 一度以上通信が成立したシグネチャについてまとめた表
- MWSシグネチャではftpとshellコード送信が多く、他はほとんどないことがわかる

| シグネチャ | ftp | http | irc | shell | smb | sql |
|---|---|---|---|---|---|---|
| MWS 60352_6 | 232 | 0 | 0 | 558 | 0 | 0 |
| MWS 53760_4 | 50 | 1 | 0 | 307 | 0 | 0 |
| MWS 60352_3 | 38 | 0 | 0 | 66 | 0 | 0 |
| MWS 65535_7 | 12 | 0 | 0 | 21 | 0 | 0 |
| MWS 60352_2 | 0 | 0 | 0 | 18 | 0 | 0 |
| MWS 60352_1 | 0 | 0 | 0 | 6 | 0 | 0 |
| **すべての通信** | **694** | **563** | **202** | **1660** | **9234** | **723** |

# 他のネットワーク通信データの分析：早稲田大学(1)

- 早稲田大学の対外接続回線におけるすべての通信データ（SYNパケットのみ、7/1～7/7）を分析
- 今回抽出したシグネチャを適用したところ、44種のうち20種が検出された
- 全送信元IPアドレス954,100 に対し、MWSシグネチャを有するものは280（約0.03%）程度
- 多く発見されたシグネチャやその送信先ポートなどの傾向はCCC DATAsetとは大幅に異なった
  - ファイアウォールの内側でデータの収集を行ったためと考えられる

# 他のネットワーク通信データの分析：早稲田大学(2)

- MWSシグネチャはすべてDFビットが0であったが、これを1に変更して再度p0fを適用
- 44種のうち31種が検出された
- 全送信元IPアドレス954,100 に対し、これらのシグネチャを有するものは5300（約0.5%）程度に増加した

- CCC DATAsetにて検出されたシグネチャは、既存OSのものも含めすべてDFビットが0であった
- DFビットは経路上のルータ等で書き換えられる可能性があるため、もともとはDFビットが1であった可能性もある
  - DFビットが1だとしても、MWSシグネチャと既存のシグネチャとはやはり異なる

# 他のネットワーク通信データの分析：企業のsmtpデータ

- ある企業網の電子メールサーバに接続したネットワークセグメントで収集したTCP ヘッダデータ（SYNパケットのみ、3/1～3/31）を分析
- 全通信元IPアドレス1,230,830に対し、MWSシグネチャを有するものはわずか53
- これらのIPから発信されたものはほとんどスパムメールであった
  - マルウェアの構成によってはスパム送信モジュールを搭載するものもあると考えられる
- MWSシグネチャのDFビットを1にして再度p0fを適用したところ、これらのシグネチャを有するIPは2877（約0.23%）まで増加した

# まとめ

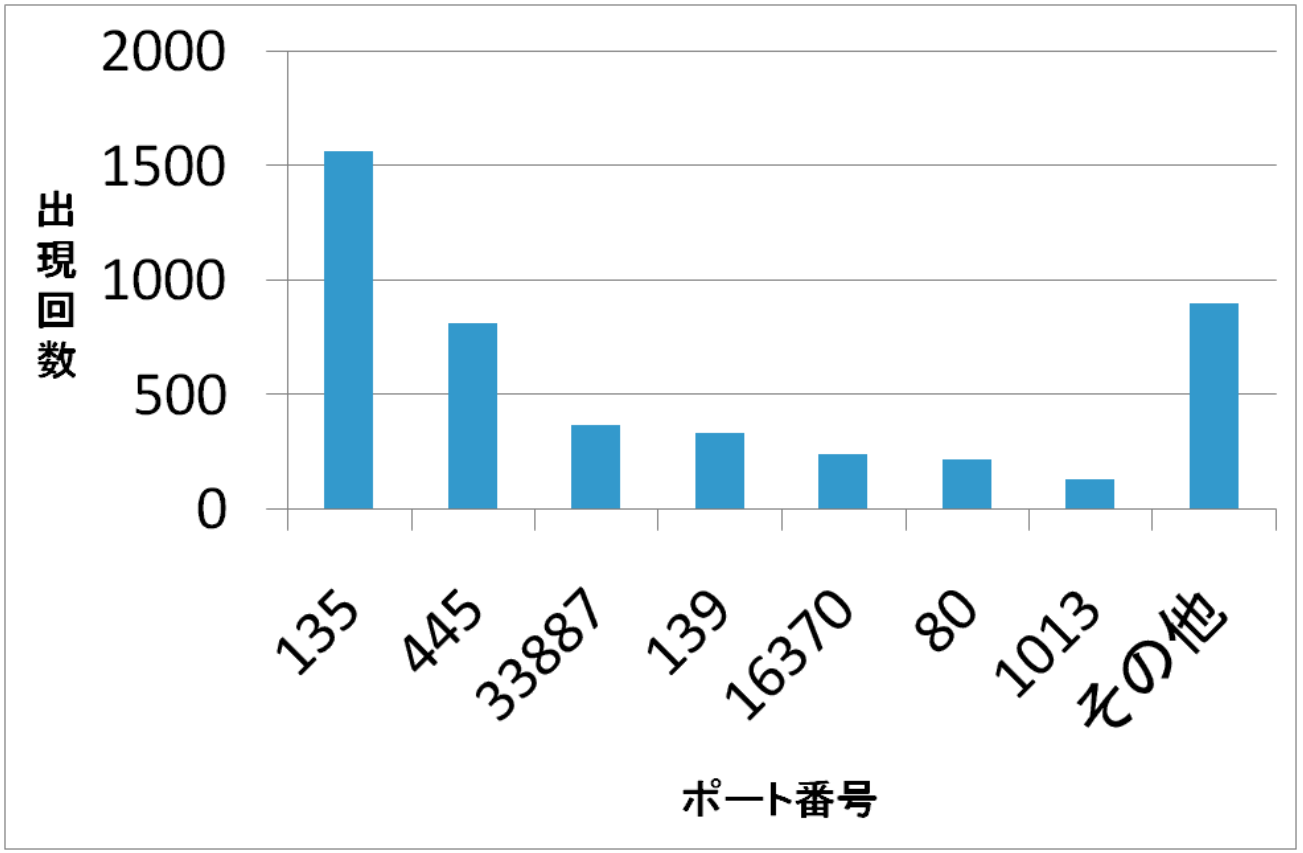
- CCC DATAsetからFKMの可能性がある（既存のTCPスタックと異なる実装である）シグネチャを抽出、通信の分析を行った
- これらのシグネチャによる攻撃通信が多数見られた
- 他のネットワーク上でもこれらのシグネチャが見られることが確認できた
  - 割合としては、各々送信元IPの1%未満
  - CCC DATAsetは、実ネットワーク上の通信に比べ、これらのシグネチャによる通信の割合が非常に高い

- 今後の課題
  - シグネチャの詳細な検討、さらなる集約
  - ハニーポットを利用したFKMの収集と分析
  - 他のネットワーク通信データの詳細分析

- ご清聴ありがとうございました。

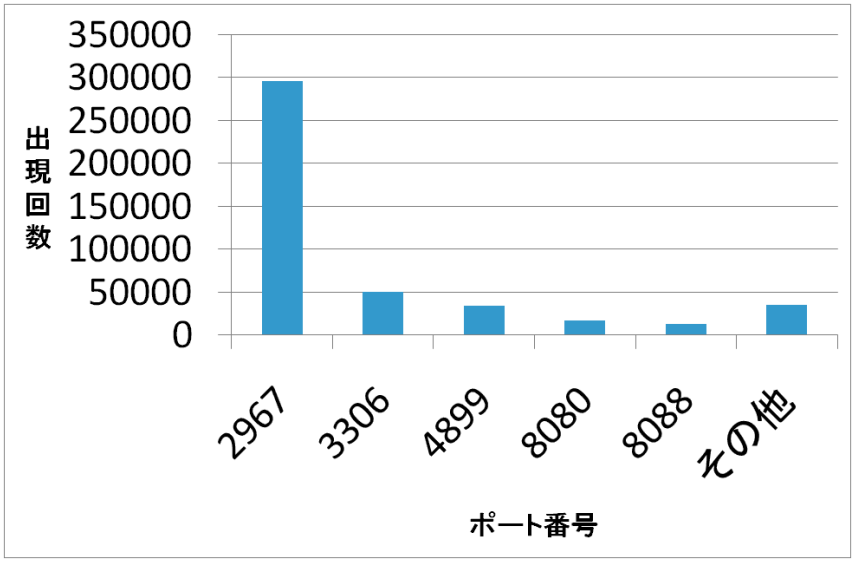
# 補足資料

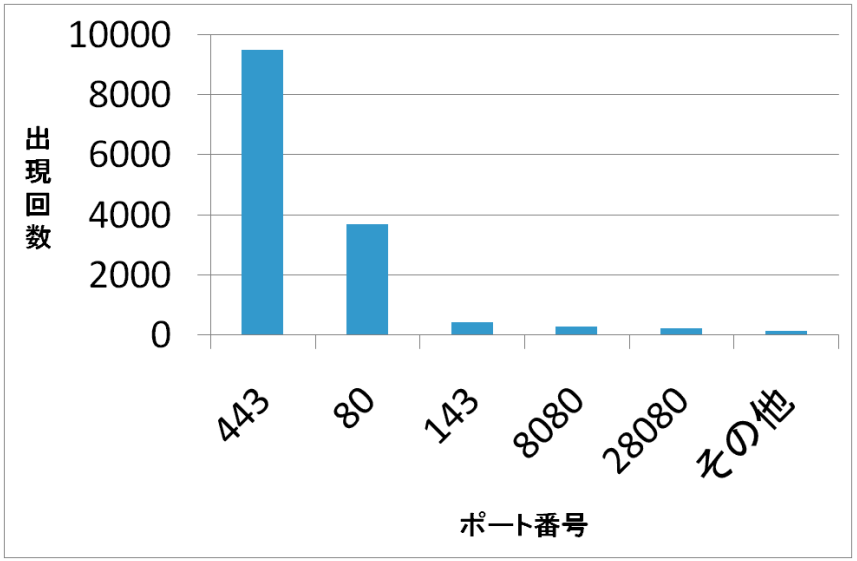
# CCC DATAsetの分析：送信先ポート(2009、全シグネチャ)




- MWSシグネチャは、135番ポートへの通信の8割以上、445番ポートへの通信の4割弱を占めていた

# 早稲田大学のデータで出現数が多かった シグネチャの送信先ポート分布




- 左(1位): 16384_1、右(2位): 65535_13
  - 16384_1はCCC DATAsetでも上位にいたが、65535_13はCCC DATAsetでは少数だった
  - 16384_1の送信ポート内訳はCCC DATAsetとかなり異なる

# 企業のsmtpデータにおける MWSシグネチャの送信メール内訳

| シグネチャ | スパム | ハム | IPアドレス数 |
|---|---|---|---|
| [65535:64:0:52:M1414,N,W3,N,N,S:.] | 29 | 0 | 9 |
| [65535:64:0:52:M1414,N,W0,N,N,S:.] | 252 | 0 | 8 |
| [65535:32:0:64:M1414,N,W3,N,N,T0,N,N,S:.] | 188 | 0 | 3 |
| [65535:64:0:52:M1400,N,W2,N,N,S:.] | 90 | 0 | 4 |
| [65535:64:0:52:M1414,N,W2,N,N,S:.] | 64 | 0 | 4 |
| [16384:128:0:60:M1414,N,N,T0,N,N,S:.] | 25 | 0 | 3 |
| [16384:16:0:40:.:.] | 21 | 0 | 1 |
| [65535:64:0:52:M1412,N,W2,N,N,S:.] | 16 | 0 | 7 |
| [53760:64:0:64:M1414,N,W3,N,N,T0,N,N,S:.] | 16 | 0 | 2 |
| [65535:64:0:64:M1414,N,W2,N,N,T0,N,N,S:.] | 9 | 9 | 1 |