M-008

アクセスログを用いた不正ホスト総数の推定に関する検討

The analysis of the number of the unauthorized computer by decentralized observation of the Internet

杉山 太一† Taichi Sugiyama　菊地 大輔† Daisuke Kikuchi　寺田 真敏†† Masato Terada

福野 直弥‡ Naoya Fukuno　田中 貴之‡ Takayuki Tanaka　菊池 浩明‡ Hiroaki Kikuchi　土居 範久† Norihisa Doi

1. はじめに

近年，インターネットの普及や，不正アクセスツールの高機能化，利用拡大に伴い，ポートスキャン，ワーム，システム侵害ならびにサイトの運用を阻害する DoS 攻撃など，インターネット上での不正アクセス活動は活性化している．インターネットにおけるこのような不正アクセス活動を観測する活動は，国内ならびに海外を含め，数多く実施されている．観測結果は，不正アクセスの活動の検出や，インターネットに与えうる脅威の度合いを推測することに利用されてはいるが，観測データから不正ホストの総数を推定するなどの検討はほとんど行われていない．

本稿では，インターネット上で実際に観測した複数のファイアウォールのアクセスログから得られた特徴を示すと共に，これら観測したアクセスログから不正ホストの総数推定の試みについて述べる．

2. アクセスログの収集環境

アクセスログから不正ホストの総数推定の試みをおこなうにあたっては，インターネットで発生している不正アクセス活動を観測するアクセスログ収集環境を使用した(図 1)．観測装置(PC)には，ファイアウォールをインストールしておき，すべてのポートを閉じておく．このようにして，到達する全てのパケットをイベントとしてロギングするようにしている．用意した観測装置のアクセスログに関する情報を表 1に示す．

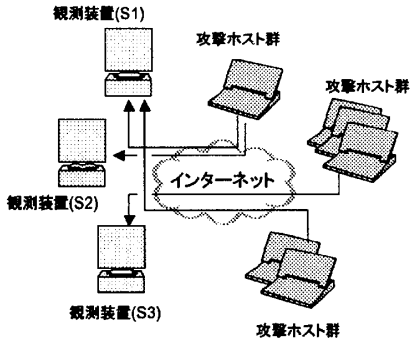

図 1　アクセスログの収集環境

このようなインターネット上のトラフィックを観測するシステムとしては表 2に示す観測システムがある．本稿において不正ホストの総数推定の試みをおこなうにあたり用意した環境は表 2の観測システムを構成するセンサと同様の機能を有すると考える．

表 1　観測装置のアクセスログの概要

	観測装置(S1)	観測装置(S2)	観測装置(S3)
観測期間	2005/05/29 - 2005/06/18		
イベント総数	499 件	2047 件	10112 件
接続するネットワーク	大学構内	大学構内	インターネット直結

注)大学構内に接続した観測装置では，特定ポート番号へのトラフィックに対してインターネットと構内ネットワークとの境界においてフィルタリングが行われている．

表 2　インターネット観測システム

名称	提供機関
ISDAS [1]	JPCERT/CC
TALOT [2]	情報処理推進機構
インターネット定点観測システム [3]	@police

3. 観測結果と不正ホスト総数の推定

3.1. 観測結果

下記前提条件の下で，観測装置で得られたアクセスログから週単位にユニークな不正ホスト数の算出をおこなった結果を表 3に示す．

【ユニークな不正ホスト数の算出に伴う前提条件】

- 観測装置と同一ネットワークからのアクセスは，算出対象から除外する．
  同一ネットワークからのアクセスとは，第 1・第 2 オクテットが同一の IP アドレスからのアクセスのことであり，そのネットワーク固有に発生しうるトラフィックを除外することを意図としている．
- 観測期間中の不正ホストの IP アドレスは変更されない．
  DHCP などにより動的に IP アドレスが割り振られる場合，不正ホストの IP アドレスが変更されてしまうが，本稿においては，観測期間中の不正ホストの IP アドレスは変更されないと仮定し推定をおこなう．
- 発信元 IP アドレスは詐称されていない．
  すべてのポートを閉じたファイアウォールでの観測をおこなっているため，詐称された発信元 IP アドレスがアクセスログに記載される場合もある．本稿においては，観測期間中の不正ホストからのアクセス

† 中央大学 理工学部 情報工学科
†† 中央大学研究開発機構
‡ 東海大学 電子情報学部 情報メディア学科

において，発信元 IP アドレスは詐称されていないと仮定し推定をおこなう．

本稿の観測期間中においては，週単位に観測されたユニークホスト数は，観測装置毎にほぼ同数を保っていることがわかった．また，各観測装置に記録された発信元 IP アドレスの分布を図 2 に示す．横軸は発信元 IP アドレス範囲(0.0.0.0～255.255.255.255)，縦軸はユニークな不正ホスト数である．いずれもほぼ類似した分布であることがわかった．

表 3　週単位に観測したユニークホスト数

	観測装置(S1)	観測装置(S2)	観測装置(S3)
1 週間目	103	354	1005
2 週間目	157	365	1085
3 週間目	128	371	1025
平均	129.3	363.3	1038.3

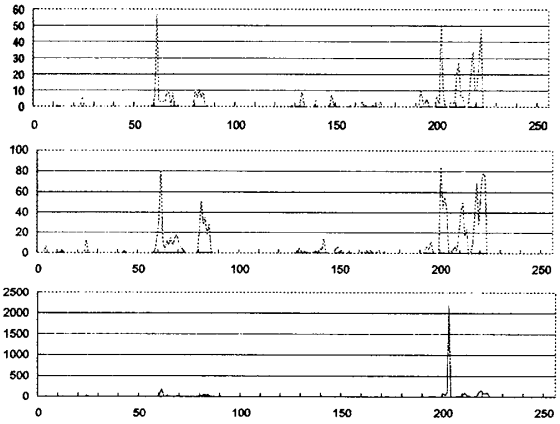

図 2　発信元 IP アドレスの分布
(上段:観測装置 S1,中段:観測装置 S2,下段:観測装置 S3)

3.2. 不正ホスト総数の推定

3 つの観測装置で得られた観測結果から，不正ホスト総数の推定方法として，観測期間を延長することで推定する方法，観測装置を増設することで推定する方法を検討する．

(1) 観測期間の延長により推定する方法

週単位にユニークな不正ホストの累積数算出をおこなった結果を表 4，図 3 に示す．インターネットにおけるIPアドレス空間は 2^{32}=4,294,967,296 であるが，実際に利用されているアドレス空間はこの理論値よりも少ない．観測結果の図 2 では，3 つの観測装置のいずれにおいても観測されないアドレス空間はAクラス 167 ブロックであった．この観測結果を用いると，不正ホストの上限は 89×2^{24} = 1,493,172,224 と仮定できる．図 3 の上段水平線は，観測結果に基づく不正ホストの上限数である．従って，観測期間の延長により，不正ホストの上限数以下のところで飽和するポイントが不正ホスト総数と推定することができる．

(2) 観測装置の増設により推定する方法

週単位にユニークな不正ホストの累積数算出をおこなった結果を表 5，図 4 に示す．観測期間の延長により推定する方法と同様に，不正ホストの上限数が存在し，観測装置の増設により，不正ホストの上限数以下のところで飽和するポイントが不正ホスト総数と推定することができる．

表 4　観測したユニークホストの累積数(観測期間)

	1 週間目	2 週間目	3 週間目
観測装置(S1)	103	244	349
観測装置(S2)	354	690	1030
観測装置(S3)	1005	2013	2944
観測装置全体	1462	2865	4214

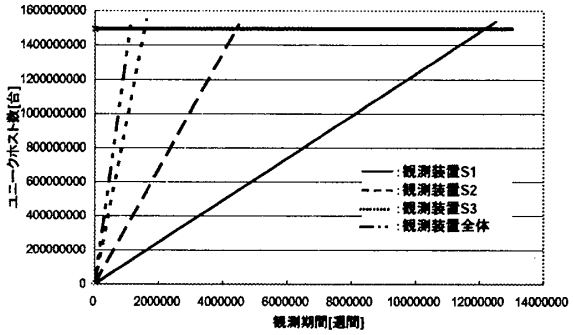

図 3　観測したユニークホストの累積数(観測期間)

表 5　観測したユニークホストの累積数(観測装置)

	観測装置(S1)	観測装置(S2)	観測装置(S3)
1 週間目	103	438	1415
2 週間目	157	497	1559
3 週間目	128	480	1481
観測期間全体	349	1331	4214

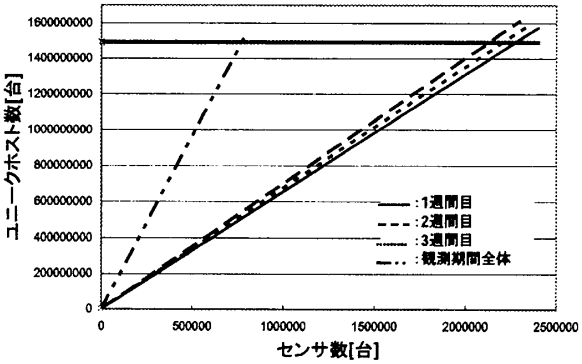

図 4　観測したユニークホストの累積数(観測装置)

4. おわりに

本稿では，インターネット上で実際に観測した複数のファイアウォールのアクセスログから得られた特徴として，各観測装置に記録された発信元 IP アドレスの分布はいずれもほぼ類似した分布となっていること，これら観測したアクセスログから不正ホスト総数を推定する方法を示した．今後の課題としては，さらに観測装置を増設するとともに，観測期間を延長しながら推定方法について評価をおこなっていきたいと考えている．

参考文献

[1] JPCERT/CC, ISDAS(Internet Scan Data Acquisition System), http://www.jpcert.or.jp/isdas/
[2] IPA, TALOT(Trend,Access,Logging,Observation,Tool), http://www.ipa.go.jp/about/press/20040511.html
[3] @police, インターネット定点観測, http://www.cyberpolice.go.jp/detect/observation.html