インターネットにおける不要トラフィックの解析

Analysis of Unnecessary Traffic on the Internet

安田 歩　　須藤 年章　　水越 一郎
NTTコミュニケーションズ株式会社

概 要

現在インターネットでは，明らかに通信として成り立っていない不必要なトラフィックが増大している．本論文では，インターネット上に存在しない経路を宛先に持つトラフィックが実際にどの程度あるのかをOCNのバックボーンで計測した．さらにこれらのトラフィックのパケットを解析し，インターネットで不要となっているトラフィックの傾向と特性を明らかにした．これにより，不必要なトラフィックが発生する原因を追求した．この解析結果に基づき，インターネット上から不要なトラフィックをなくし，ネットワークをより効率的に設計・運用することが可能となった．

1　はじめに

1.1　研究背景

インターネット上でのトラフィックは増加の一途であり，インターネットを構成するネットワークも拡大を続けている．ISPなどの回線所有業者は，ネットワークの設計に際して，バックボーンおよびユーザがいる各地域ごとにトラフィックの流量を把握し，今後の動向を想定する．トラフィックが増加しているのであれば，それに対応して回線容量を増速させる．今後増加が見込めなければ，増速は行なわない．このように，実際にインターネットに流れているトラフィック量が，ネットワークの規模や構成を決定している大きな要因となっている．

しかし，インターネットを流れるトラフィックの全てが，正当なユーザ通信あるいはネットワーク監視のために必要なものという訳ではない．ワームやDoS攻撃等，特定のサーバや不特定多数の相手への悪意のある攻撃を目的としたトラフィックは年々増加する一方である．それ以外にも，正当な宛先を持たずに発信され，結果として途中で廃棄されるパケットも少なくない．

ネットワークの帯域は貴重な資源であり，不要なトラフィックに占有されることは好ましくない．コストを低減し，不要なトラフィックによる帯域への圧迫から正当なユーザ通信を保護するためにも，不要なトラフィックは可能な限り排除するべきである．このためには，実際のトラフィックのうち，どの程度が全くの無駄なトラフィックなのかを把握することが重要である．

1.2　目的

本論文では，インターネット上に存在しない経路を宛先としたパケットがどの程度存在しているかを把握し，その中身を解析する．

不要なトラフィックの性質を見極めることで，これらのトラフィックがユーザやアプリケーションのミスで発生したものなのか，あるいはそれ以外の原因によるものなのかを特定できる．さらに，これらのデータを基に，不要なトラフィックを排除することができれば，より効率的なネットワークの構築・運用が可能となる．

また，不要トラフィックが多ければ，より多くのルータで full route を持つことがバックボーンネットワークを設計する上で効果的となる．一方，full route に乗るトラフィックと，default route に乗るトラフィックに差分がなければ full route を持つ必要があるルータの台数を少なくすることが有用であることを証明できる．

1.3　対象とするネットワーク

本研究は，OCN[1] のバックボーンを対象として行なう．OCNは，2004年10月現在，個人・法人に偏りなく合わせて300万以上のユーザがいる．バックボーンはTier1である上位ISPと48Gbpsで接続し，国内の基幹である東京 - 大阪 間は80Gbpsの回線を持つ．また国内では多数のISPと接続している。これらより，国内でも最大規模であるOCNは，日本のインターネット全体のサンプルとするに充分な値かつ構成であると言える．

2　現状のインターネットのトラフィック特性

2.1　トラフィックの性質

不要となるトラフィックの存在がインターネットにおいてどの程度の影響があるのかを考察するには，現状のインターネットのトラフィック特性を把握する必要がある．

しかし，ユーザのネットワークの利用目的や用途，利用するアプリケーションを明確に知ることはできない．ビジネスで利用するか個人でプライベートに利用するのかによっても，障害が発生した時の影響は大きく異なる．また，オンラインゲームやIP電話などの普及により，ネットワークに対するリアルタイム性が強く要求されてきている．

このように，通信に絶対的な信頼性が求められるものとそうでないもの，遅延を最小に抑えたリアルタイム性が要求されるものとそうでないもの，といったトラフィッ

クが，それぞれどの程度あるのかを測定できるのが望ましい．

しかし，ユーザのトラフィックをキャプチャすることはできない．今回は，ISP の運用者として判断可能な以下の項目のみ，実データに基づいてトラフィックの性質を分析した．

1. 時間帯によるトラフィックの流量
2. アクセス回線の種別による格差
3. 地域による格差

OCN バックボーンを対象とし，それぞれの項目について以下に述べる．

2.1.1 時間帯によるトラフィックの流量

ここ 1,2 年は，早朝の時間帯にはトラフィックは減るものの，昼夜のトラフィックの流量はほとんど変わらない．これは，昼間のビジネスでの利用量と，個人が夜間で利用するトラフィックの量が変わらなくなってきていることを示している．ADSL, 光アクセスによる常時接続が一般的となったことも原因として挙げられる．

2.1.2 アクセス回線の種別による格差

ユーザが利用するアクセス回線として，OCN では ISDN, ADSL, 光アクセスの 3 種類が挙げられる．利用目的によりアクセス回線が異なることが想定されるため，このアクセス回線の違いによる 1 の項目について検証した．

ここで，ある県の ISDN ユーザに限定した一日のトラフィック量の推移を図 1に示す．

ISDN のユーザは，深夜から早朝にかけては利用量が明白に落ち込んでいるが分かる．これとは対照的に，ADSL, 光アクセスのユーザは，深夜から早朝にかけてもこれほどの落ち込みはない．

このことより，常時接続のユーザは時間帯に関わらずインターネットを利用していることが分かる．

2.1.3 地域による格差

各都道府県ごとの地域により，利用形態が異なることが考えられる．地域の違いによる 1 の項目について検証した．

地域により，トラフィック量の差はあるものの，ISDN のユーザは深夜から早朝にかけて利用量が相対的に少なくなり，ADSL, 光アクセスのユーザは一日を通してそれほど利用量に変化がないという傾向は変わらなかった．

2.2 不要トラフィックの定義

IANA[2](ICANN[3]) および RIR から ISP などの運用組織に割り当てらていない，プライベートアドレスおよび特別な用途に予約された IP アドレス [4] を bogons と呼ぶ．これらの IP アドレスは，グローバルのネットワーク上に経路として存在してはならないし，実際に存在しない．プライベートアドレスが存在するのは，プライベートなネットワーク内だけである．

bogons を含む，full route に存在しない経路の IP アドレスを宛先としたパケットは，その宛先に到達することはなく，廃棄される．このため，インターネットではこれらのトラフィックは，全く不必要となる．本論文では，full route に存在しない IP アドレスを宛先としたパケットを，不要トラフィックと定義する．パケットのペイロードには関知せず，ヘッダの宛先 IP アドレスのみで不要かどうかを判断する．

2.3 不要トラフィックの問題点

インターネット上での不要トラフィックの存在は，以下の点で問題となる．

- ネットワーク帯域の浪費
- ルータ・スイッチへの余計な負荷

full route を持つルータや多くの処理を実行するルータでは，メモリや CPU の負荷が問題となる．本来なら必要のない負荷を掛けることは，高負荷を起因とする reload を引き起こすなど，ネットワークに多大な影響を及ぼす障害をの原因となりかねない．ネットワークを運用する立場としては，結果として正当なユーザ通信を阻害することになる不要トラフィックは，可能な限り抑制する必要がある．

3 データの取得

3.1 OCN バックボーンの構成

データを取得する前提として，実際にデータを取得するネットワークの構成について述べる．

OCN のバックボーンは，東京と大阪それぞれに 2ヶ所づつある NOC を中心として，その配下に全国 10 数ヶ所の NOC が接続されている．大まかな構成図を，図 2に示す．

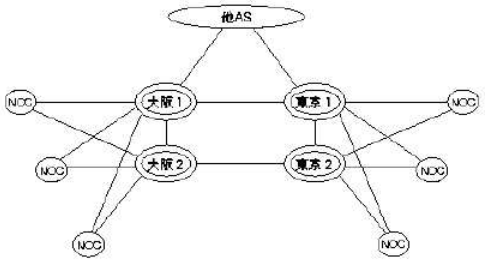

図 2: OCN バックボーンの構成図

東京と大阪を含む全国の NOC の下では，それぞれユーザを収容している．今回，東京と大阪の 2ヶ所にデータ取得サーバを設置した．今回設置した 2ヶ所の NOC を通過するユーザからのトラフィック量の，全国における割合はおよそ 36% である．

一方，各 NOC の下で折り返すトラフィック量を実際に計測したところ，各 NOC および時刻の平均として，およそ 3.8%から 4.0% という値が出た．

これにより，OCN のユーザが発生させる全トラフィックのおよそ 34% のトラフィックについて，不要トラフィックを計測できた．これは，サンプルとしては充分な値である．

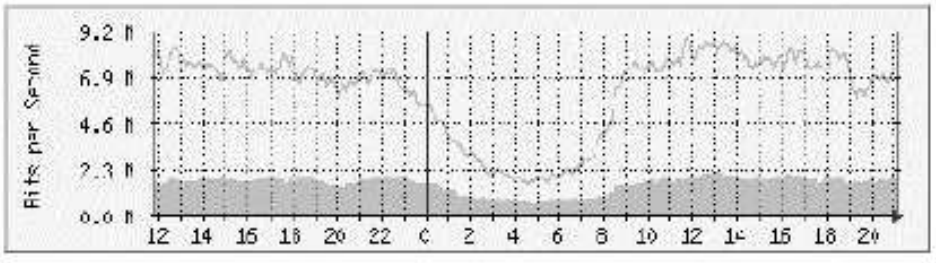

図 1: ISDN ユーザのみのトラフィック量を示すMRTG

3.2 データ取得の環境

次に，不要トラフィックを取得した際のネットワーク構成について述べる．

OCN バックボーン内にデータ取得サーバを設置し，不要トラフィックのデータ収集を行なった．トポロジを図3に示す．

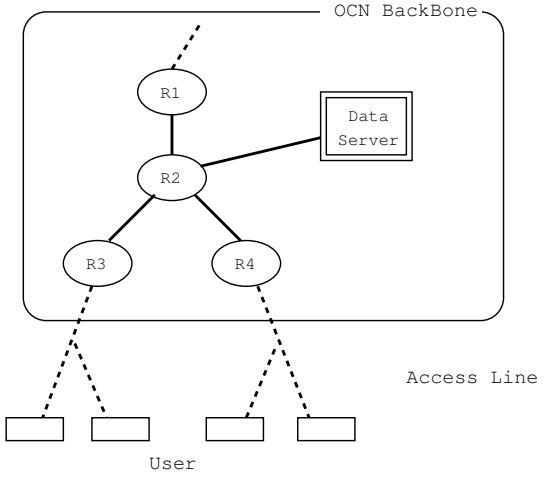

図 3: OCN バックボーンとデータ取得サーバの構成図

一般的なユーザはFLET'S[5] などのアクセス網を介して，OCNバックボーンに接続される．図3において，R3, R4 がアクセス網と接続してユーザを収容しているルータである．なお，図の点線部では，ネットワークの構成図を一部省略している．ユーザからのトラフィックは，R3, R4 を通り，ネットワークの上流である R2 に到達する．R3, R4 は full route を持っていないが，R2 は full route を持っており，full route にない経路への宛先を持つパケットはR2で廃棄している．本研究では，このルータ R2 にデータ取得サーバを接続した．本来R2で廃棄されるトラフィックをデータ取得サーバに向けることにより，データの収集を行なった．

なお，不要トラフィックの解析にあたっては，10 分間隔で5分間収集したデータを，対象とした．

3.3 取得対象としたデータの種類

収集したパケットを以下の項目について観測した．

- 送信元IP アドレス
- 宛先 IP アドレス
- 送信元ポート番号
- 宛先ポート番号
- プロトコル
- パケットサイズ

送信元ポート番号については解析意義が低いため，本論文では解析対象から除外した．

4 データの解析と考察

4.1 不要トラフィックの量および割合

実際にデータ取得サーバで収集したデータについて解析した結果を，以下に述べる．

東京で収集している不要トラフィックの，ある一日のトラフィック量の推移のグラフを図4に示す．

あるワームの急激な広がりがない限り，一日における不要トラフィック量の推移はほぼ図4が示すように15Mbps 前後の値を示す．

OCN 全体に換算すると，80Mbps から 100Mbps のトラフィックが宛先が不正のトラフィックであることが分かる．

4.2 宛先ポート番号別の集計

2004 年10 月のある1週間での不要トラフィックである，不正な宛先を持ったパケットの宛先ポート番号別の集計を，図5に示す．縦軸はパケット数を表す．

上位を占めているポート番号は，いづれも著名なワームで使用されているものである．発生し流行してからかなりの月日が経つものの，ユーザがワームを駆逐しきれていない現実が分かる．

この結果より，不要トラフィックは，近年のワームによる影響が大きいことが分かった．

4.3 送信元IP アドレス別の集計

不要トラフィックを送信しているトラフィックを，送信元IP アドレス毎に集計した．

この結果，送信元IP アドレスのほとんどが正規のユーザとしてのIP アドレスからであった．

意図的に不正な攻撃行為を行なうとき，多くの場合で送信元IP アドレスを詐称する．このとき，数値的に近くのIP アドレスを詐称することは考えにくい．つまり，実際には検知した，送信元通りのIP アドレスから送信され

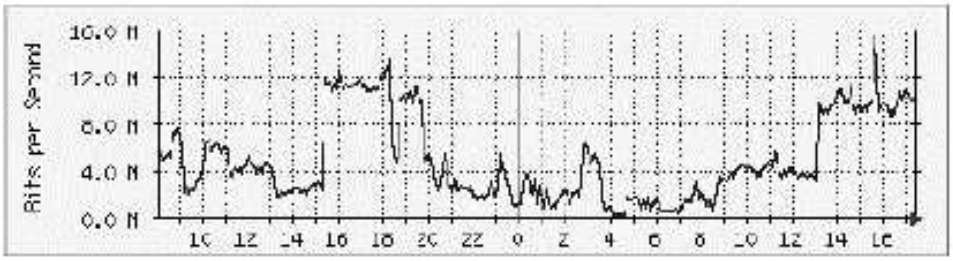

図 4: 一日の不要トラフィックの量

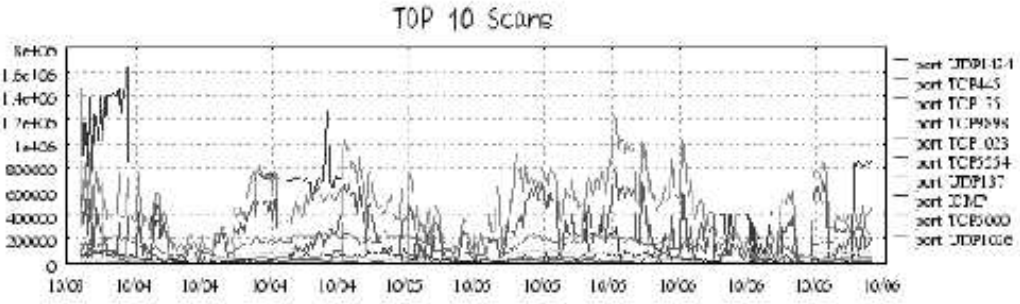

図 5: 不正宛先向けパケットの宛先ポート番号別集計

ていると考えられ，意図的な不正の攻撃行為は極めて少ないことを示している．

また，不要トラフィックを発信しているIPアドレスのうち，トラフィック量の多い上位10ユーザ の，全体における割合はおよそ 37% であった．送信元IPアドレスの総数もかなりあり，多数のユーザがワームに感染している結果であることが考えられる．

4.4 宛先IPアドレス別の集計

不要トラフィックを，/16 単位で宛先IPアドレスごとに集約し集計した．

この結果，宛先IPアドレスの傾向として，次の2つ場合が混在していることが明らかになった．

4.4.1 宛先IPアドレスに一定の規則性がある場合

ワームの挙動として，ワームのアルゴリズムに基づき宛先IPアドレスをインクリメントしてパケットを送信するものがある．この例としてMS Blaster などが挙げられる．また，宛先ポート番号を固定したり，1つの宛先IPアドレスごとに wellknown port を順々に探索していたりするものもある．これらは，ツール等を使用した意図的なポートスキャンによるものと想定される．これらのパケットのうち，宛先IPアドレスが full route に存在しないIPアドレスになったものが，今回観測されたと想定される．

4.4.2 宛先IPアドレスがランダムな場合

ワームの挙動として，ワームのアルゴリズムに基づき宛先IPアドレスをほぼランダムにパケットを送信するものがある．例として Slammer が挙げられる．これらのパケットのうち，宛先IPアドレスが full route に存在しないIPアドレスになったものが，今回観測されたと想定される．

4.5 パケットサイズ別の集計

不要トラフィックを，パケットサイズ別に集計した．値は，10分間隔に5分間収集したものである．パケットサイズごとの，パケット数とその割合を表1に示す．

表 1: パケットサイズ別の集計表

パケットサイズ(byte)	パケット数	割合
0 - 31	23919	1.63 %
32 - 127	1579867	88.84 %
128 - 255	90324	5.08 %
256 - 511	435	0.02 %
512 - 1023	25963	1.46 %
1024 -	52858	2.97 %

32-127Byte のパケットが大半を占めている．これは，TCP を用いるワームは最初に TCP のコネクションの確立を試みるために Syn パケットを送信する．また TCP ポートへのポートスキャンにおいても TCP の Syn パケットを用いる．ところが，ここで観測されているパケットは不正宛先向けであるため廃棄されてしまうため TCP シーケンスは進行しない．したがって TCP パケットはほぼすべてが Syn パケットであり 48Byte になる．このようなワームのパケット送信量の多さ，感染者数の多さからこのような結果が得られていると想定される．UDP を用いるワームについては，そのパケットサイズがワーム毎に特徴をもっているため，パケットサイズによりワームの分布を想定することも可能である．

4.6 ワームとの具体的な関連性

ここで，不要トラフィックの中で観測された宛先プロトコル／宛先ポート番号の組合せのうち，トラフィック量の多い上位10組について，それぞれに対応するワームの一覧を示す．

表 2: 不要トラフィックを占めるワームの一覧

protocol / port	service	warm
UDP 1434	MS-SQL-Monitor	Slammer
TCP 445	microsoft-ds	Sasser, Dabber
TCP 135	Location Service	Blaster
UDP 137	NetBIOS	Blaster, etc
TCP 9898	N/A	Dabber
TCP 1023	N/A	Dabber
TCP 5554	N/A	Dabber
ICMP	N/A	N/A
UDP 1027	messenger service	spam広告
UDP 1026	messenger service	spam広告

表2に示す通り，不要トラフィックのほとんどが著名なワームであることが明らかになった．

4.7 考察

不要トラフィックの量は，正常な通信に比較してわずかであることが分かった．逆に不正トラフィックが急増した場合，ワーム被害の急激な拡大や，大規模なDDoS攻撃が発生していることが考えられる．

また，ポート番号による統計より，不要トラフィックの多くがワームによるものであることが判明した．さらに，IPアドレスによる統計より，ワームの挙動として既知のものである通り，宛先IPアドレスに対して体系だってトラフィックを出していくものと，ランダムに出していくものの双方を観測することができた．これらのワームは出現してから月日が経つものの，インターネットにはまだ蔓延している実態が明らかになった．

5 不要トラフィックがインターネットに及ぼす影響

5.1 一般的なネットワーク構成

現在，full route は /24 未満の経路を除いても15万弱存在し，この数字は日に日に増加している．これだけの数の経路を処理するためには，ルータにある程度の性能が要求される．また，各経路制御プロトコルでは，経路情報が更新される度にその情報が伝搬される．このため，full route を持つルータが数多くなると，それに比例して経路情報を交換するためのトラフィックが増大する．さらに，経路表作成のための処理などにより，ルータに大きな負荷がかかる．これらのことより，ネットワーク上の全てのルータがfull route を持つことは現実的ではない．

ネットワーク的に下流となるユーザ側のルータでは，default route が上位のルータに向いていることで，結果的に経路表のエントリ数を抑制している．一方，上位のルータは full route を持つことで経路選択の正確性を維持している．

ここで，階層化されているネットワークにおいて，どこまで full route を持つ必要があるか，ということが問題となる．グローバルASを持つ組織での一般的なネットワーク構成では，組織内のネットワーク的階層の上位，つまり他ASと接続するルータか，あるいはその1階層下のルータまでが full route を持つことが多い．

full route を持つルータのみが，トラフィックが不要かどうかの判断が可能であり，また不要トラフィックを廃棄することができる．

5.2 不要トラフィックを抑制する必要性

不要なトラフィック量が少なく，割合としても小さな場合は，通常の運用に支障をきたさない．しかし，突発的なワームの発生などにより不要トラフィックが急増した場合，ネットワークの帯域およびルータ，スイッチへの高負荷といったハード的資源へ悪影響を及ぼす．また，多量ではなくても慢性的に不要トラフィックがある場合も，ネットワーク帯域資源の恒常的な無駄となる．

ハード的トラブルを未然に防ぎ，ルータやネットワークの資源を有効に活用するためにも，不要トラフィックは，通過するリンクやノードが少ないうちに廃棄するべきである．

6 まとめ

6.1 結論

本研究では，インターネット上での不要なトラフィックの量を調査し，その実態を把握することを目的とした．この目的を実現するために，OCNのバックボーンにデータ取得サーバを設置した．このサーバで，インターネット上に存在しない経路を宛先としたトラフィックをキャプチャし，その中身を解析した．複数の項目について解析したところ，不要トラフィックの多くがワームであることが分かった．また，通常時において，実測した不正なトラフィック量は，正常な通信に比較して微小であることが分かった．

今までは，大規模なネットワークにおいてネットワークに多大な影響を及ぼすワーム，DDoSの状況の全体を把握するためには，大量のトラフィックに対するモニタリング，または高性能なモニター機器の多くの場所へ分散的な設置が必要であった．

本研究により，全トラフィックに対して微小である不要トラフィックについて，その総量の変動を観測することだけで，ワーム等の活動の活発化を発見することができることが示された．そのパケットを解析することによりネットワーク全体における，各種ワームの活動状況，新種のワームの発生を早期に容易に発見し，その状況を容易に想定することが可能になった．

6.2 今後の課題

不要トラフィックを抑制した際の効果を測定するには，不要トラフィックの発信を何らかの手法により抑制する必要がある．これについては，ユーザのトラフィックを制限することになるため，具体的な実施方法を含めて検討する必要がある．

また，今回は東京と大阪の2ヶ所のみに設置したが，今後全国のNOCに同様のサーバを設置していく．これにより，より精度を上げて各地域に即したデータを取得して状況を把握することが可能となる．

参考文献

[1] Open Computer Network, NTT Communications. http://www.ocn.ne.jp/

[2] Internet Assigned Numbers Authority. http://www.iana.org/

[3] Internet Corporation For Assigned Names and Numbers. http://www.icann.org/

[4] IANA. Special-Use IPv4 Addresses. RFC 3330, IETF, September 2002.

[5] 東日本電信電話株式会社 http://www.flets.com/ 西日本電信電話株式会社 http://flets-w.com/