

違法ダウンロード刑事罰化の影響は限定的

法的強制措置の効果を検証する意味で、2012年10月1日の違法ダウンロード刑事罰化の影響について注目していました。

施行後3ヵ月間程はトラフィックが減少しましたが、その後はもとのトレンド曲線に戻ったことから、今回のダウンロード刑事罰化は一時的な心理効果しかなかったように見えます。

3.1 概要

本レポートでは、毎年、IIJが運用しているブロードバンド接続サービスのトラフィックを分析して、その結果を報告しています^{*1*2*3*4}。今回も、利用者の1日のトラフィック量やポート別使用量などをもとに、この1年間のトラフィック傾向の変化を報告します。

図-1は、ブロードバンド全体の過去6年間の月平均トラフィックです。2010年1月のトラフィック減少は、2010年1月に施行された改正著作権法、いわゆるダウンロード違法化の影響だと考えられています。それ以降、ダウンロード量(OUT)が増えている一方で、アップロード量(IN)は横ばいとなっていて、P2Pファイル共有のトラフィック割合が減っていることが窺えます。2011年3月の東日本大震災では、被災県でこそ設備と回線の被害や停電の影響によるトラフィック減少が観測されましたが、全国レベルで見た影響は大きくはありませんでした。2012年10月には、違

法ダウンロードの刑事罰化を含む改正著作権法が施行されました。その前後でトラフィックの増減が観測されましたが、その影響は限定的で、2010年のダウンロード違法化のときのような長期トレンドの変化はなかったと言えます。この1年のトラフィック量は、INは8%増加、OUTは16%増加しています。

3.2 データについて

今回も前回までと同様に、個人及び法人向けのブロードバンド接続サービスについて、ファイバーとDSLによるブロードバンド顧客を収容するルータでSampled NetFlowにより収集した調査データを利用しています。ブロードバンドトラフィックは平日と休日で傾向が異なるため、1週間分のトラフィックを解析しています。今回は、2013年6月3日～9日の1週間分のデータを、前回解析した2012年5月28日～6月3日の1週間分と比較します。

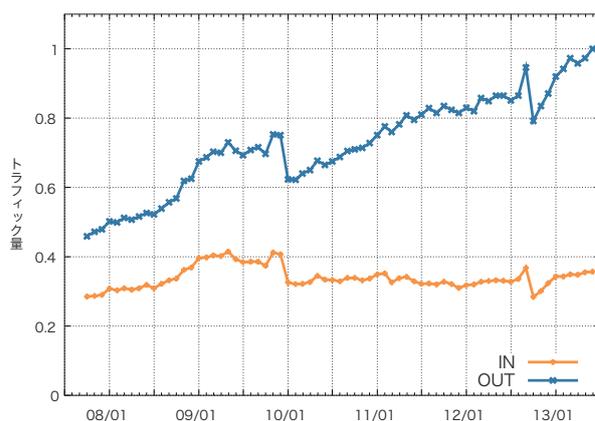


図-1 過去6年間のブロードバンドトラフィック量の推移

各利用者の使用量は、利用者に割り当てられたIPアドレスと、観測されたIPアドレスを照合して求めています。また、NetFlowではパケットをサンプリングして統計情報を取得しています。サンプリングレートは、ルータの性能や負荷を考慮して、1/8192に設定されています。観測された使用量に、サンプリングレートの逆数を掛けることで全体の使用量を推定しています。サンプリングによって、使用量の少ない利用者のデータには少し誤差がでますが、ある程度使用量のある利用者に対しては統計的に意味のある数字が得られます。

- *1 長健二郎。ブロードバンドトラフィックレポート：この1年間のトラフィック傾向について。Internet Infrastructure Review. Vol.16. pp33-37. August 2012.
- *2 長健二郎。ブロードバンドトラフィックレポート：マクロレベルな視点で見た、震災によるトラフィックへの影響。Internet Infrastructure Review. Vol.12. pp25-30. August 2011.
- *3 長健二郎。ブロードバンドトラフィックレポート：P2Pファイル共有からWebサービスへシフト傾向にあるトラフィック。Internet Infrastructure Review. Vol.8. pp25-30. August 2010.
- *4 長健二郎。ブロードバンドトラフィック：増大する一般ユーザのトラフィック。Internet Infrastructure Review. Vol.4. pp18-23. August 2009.

ここ数年でDSLからファイバーへの移行が進み、2013年には観測されたユーザ数の93%はファイバー利用者で、トラフィック量全体の96%を占めるまでになっています。

なお、本レポート中のトラフィックのIN/OUTはISPから見た方向を表し、INは利用者からのアップロード、OUTは利用者へのダウンロードとなります。

3.3 利用者の1日の使用量

まずは、ブロードバンド利用者の1日の利用量をいくつかの切口から見ていきます。ここでの1日の利用量は各利用者の1週間分のデータの1日平均です。

図-2は、利用者の1日の平均利用量の分布(確率密度関数)を示します。アップロード(IN)とダウンロード(OUT)に分け、利用者のトラフィック量をX軸に、その出現確率をY軸に示していて、2012年と2013年を比較しています。X軸はログスケールで、10KB(10^4)から100GB(10^{11})の範囲を示しています。一部の利用者はグラフの範囲外にいますが、概ね100GB(10^{11})までの範囲に分布しています。

INとOUTの各分布は、片対数グラフ上で正規分布となる、対数正規分布に近い形をしています。これはリニアなグラフで見ると、左端近くにピークがあり右へなだらかに減少するいわゆるロングテールな分布です。OUTの分布はINの分布より右にずれていて、ダウンロード量がアップロード量より、ひと桁以上大きくなっています。2012年と

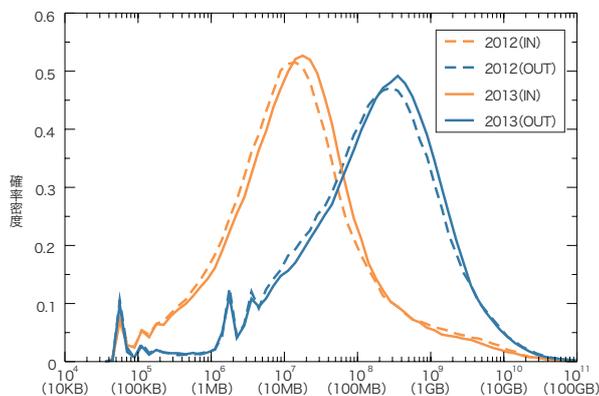


図-2 利用者の1日のトラフィック量分布
2012年と2013年の比較

2013年で比較すると、INとOUT共に分布の山が右に少し移動していて、利用者全体のトラフィック量が増えていることが分かります。

OUTの分布を見ると、分布のピークはここ数年間で着実に右に移動していますが、右端のヘビーユーザの使用量はあまり増えていないので、分布の対称性が崩れてきています。一方で、INの分布は右側の裾が広がっています。以前は、ここによりはっきりした山がINとOUT両方にあり、IN/OUT量が対称なヘビーユーザを示していました。そこで便宜上、大多数のIN/OUT非対称な分布を「クライアント型利用者」、右側の小数のIN/OUT対称なヘビーユーザの分布を「ピア型利用者」と呼んできました。今回もその慣習に従いますが、ここ数年で、ピア型利用者の山は小さくなってきていますが、これは、ヘビーユーザの割合が減少していることを示しています。グラフ左側に少しヒゲが出ていますが、これはサンプリングレートの影響によるノイズです。

表-1は、平均値と、分布の山の頂点にある最頻出値の推移を示します。分布の最頻出値を2012年と2013年で比較すると、INでは14MBから18MBに、OUTでは282MBから355MBに増えていて、各利用者のトラフィック量が、特にダウンロード側で増えていることが分かります。一方、平均値はグラフ右側のヘビーユーザの使用量に引っ張られるので、2013年には、INの平均は397MB、OUTの平均は1038MBと、最頻出値よりかなり大きな値になります。2012年では、それぞれ410MBと1026MBだったので、平均値で見ると、INが減少して、OUTが増える傾向が続いています。

年	IN (MB/day)		OUT (MB/day)	
	平均値	最頻出値	平均値	最頻出値
2005	430	3.5	447	32
2007	433	4	712	66
2008	483	5	797	94
2009	556	6	971	114
2010	469	7	910	145
2011	432	8.5	1,001	223
2012	410	14	1,026	282
2013	397	18	1,038	355

表-1 利用者の1日のトラフィック量の平均値と最頻出値の推移

図-3は、利用者ごとのIN/OUT使用量の中から5,000人をランダムに抽出してプロットしています。X軸はOUT(ダウンロード量)、Y軸はIN(アップロード量)で、共にログスケールです。利用者のIN/OUTが同量であれば対角線上にプロットされます。

対角線の下側で対角線に沿って広がるクラスは、ダウンロード量がひと桁多いクライアント型の一般ユーザです。以前は、右上の対角線上あたりを中心に薄く広がるピア型のヘビーユーザのクラスがはっきり分かりましたが、今では識別が難しくなっています。便宜上、クライアント型とピア型に分けましたが、実際には、クライアント型の一般ユーザでもSkypeなどのピア型のアプリケーションを利用し、また一方のピア型のヘビーユーザもウェブなどのダウンロード型のアプリケーションを利用しているので、その境界はあいまいです。つまり、多くの利用者は両タイプ

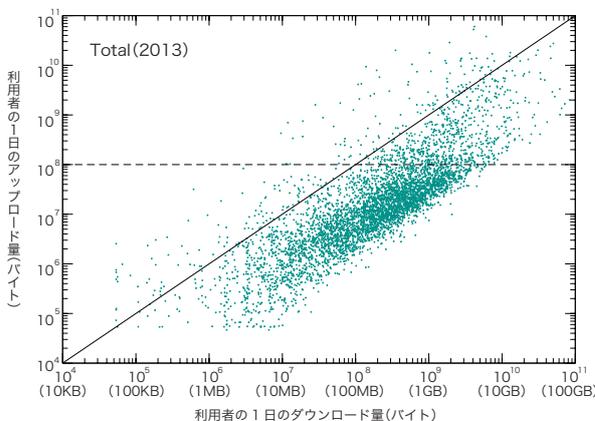


図-3 利用者ごとのIN/OUT使用量

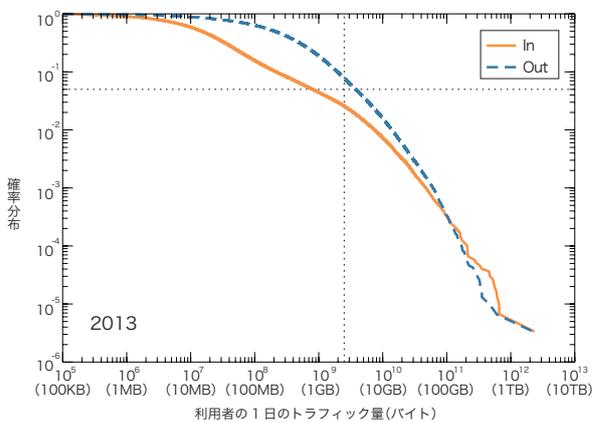


図-4 利用者の1日のトラフィック量の相補累積度分布

のアプリケーションを異なる割合で使用しています。また、各利用者の使用量やIN/OUT比率にも大きなバラツキがあり、多様な利用形態が存在することが窺えます。ここでは、2012年と比較しても、ほとんど違いは確認できません。

図-4は、利用者の1日のトラフィック量を相補累積度分布にしたものです。これは、使用量がX軸の値より多い利用者の、全体に対する割合をY軸に、ログ・ログスケールで示したもので、ヘビーユーザの分布を見るのに有効です。グラフの右側が直線的に下がっていて、ベキ分布に近いロングテールな分布であることが分かります。2012年と比較すると、グラフ右端のテール部分が若干右側に延びていて、昨年減った直線から右側に外れるような極端なヘビーユーザが再び観測されています。いずれにせよ、ヘビーユーザは統計的に分布していて、決して一部の特殊な利用者ではないと言えます。

図-5は、利用者間のトラフィック使用量の偏りを示します。使用量上位X%の利用者が、全体トラフィック量のY%を占めることを表します。使用量には大きな偏りがあり、結果として全体は一部利用者のトラフィックで占められています。例えば、上位10%の利用者がOUTの71%、INの94%を占めています。更に、上位1%の利用者がOUTの32%、INの65%を占めています。

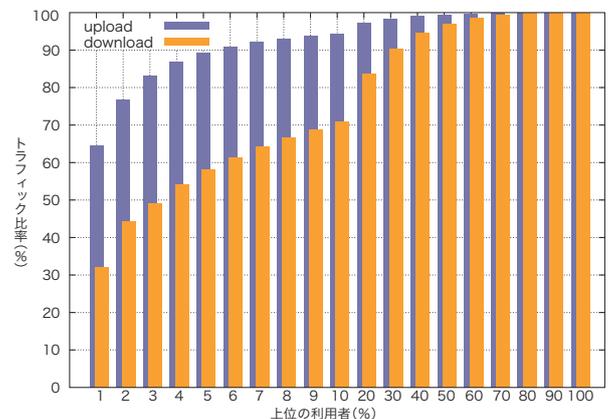


図-5 利用者間のトラフィック使用量の偏り

3.4 ポート別使用量

次に、トラフィックの内訳をポート別の使用量から見ていきます。最近では、ポート番号からアプリケーションを特定することは困難です。P2P系アプリケーションには、双方で動的ポートを使うものが多く、また、多くのクライアント・サーバ型アプリケーションが、ファイアウォールを回避するため、HTTPが使う80番ポートを利用します。大雑把に分けると、双方が1024番以上の動的ポートを使っていればP2P系のアプリケーションの可能性が高く、片方が1024番未満のいわゆるウェルノウンポートを使っていれば、クライアント・サーバ型のアプリケーションの可能性が高いと言えます。そこで、TCPとUDPで、ソースとデスティネーションのポート番号の小さい方を取り、ポート番号別の使用量を見てみます。

また、全体トラフィックは、ピア型のヘビーユーザのトラフィックに支配されているので、クライアント型の一般利用者の動向を見るために、少し荒っぽい方法ですが、1日のアップロード量が100MB未満のユーザを抜き出して、これをクライアント型利用者としてします。これは、図-3では、IN=100MBにある水平線の下側の利用者にあたります。

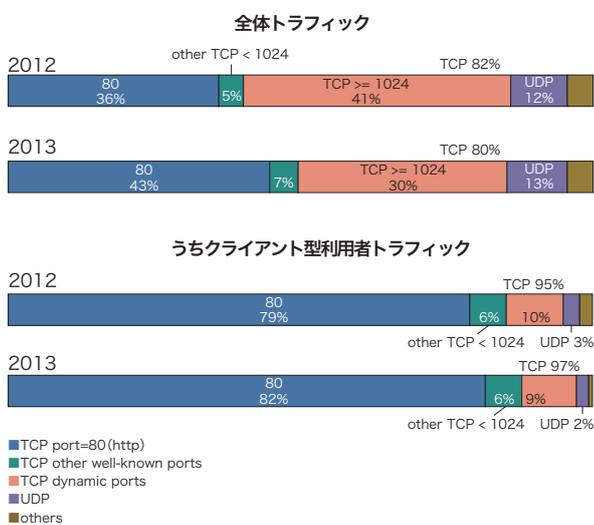


図-6 ポート別使用量概要

図-6はポート使用の概要を、全体とクライアント型利用者について、2012年と2013年で比較したものです。また、表-2にその詳細を数値で示します。

2013年のトラフィックの80%はTCPです。更に、全体で見ると、2012年には総量の41%だったTCPの動的ポートが、2013年には30%にまで減少しました。動的ポートでの個別のポート番号の割合は僅かで、Flash Playerが利用する1935番が最大で総量の約2%ありますが、後は0.5%未満となっています。逆に、80番ポートの割合は、2012年の36%から43%に増加しています。TCP以外のトラフィックのほとんどはVPN関連で、増加傾向にあります。

一方、クライアント型利用者に限ると、2012年には79%を占めていた80番ポートが、2013年には82%に増加しています。2番目に多いのは、HTTPSで使われる443番ポートで、2012年の3%から5%に増えています。これに対して、動的ポートの割合は、10%から9%に減少しています。

これらのデータから、TCP80番ポートのトラフィック増加傾向が一般利用者だけでなく、ヘビーユーザでも引き続き進行していることが確認できます。80番ポートにはビデオ

protocol port	2012		2013	
	total (%)	client type	total (%)	client type
TCP	81.86	95.09	79.79	96.91
(<1024)	41.23	85.25	49.57	88.15
80(http)	36.22	79.39	43.44	81.61
443(https)	2.45	3.43	3.90	4.80
554(rtsp)	0.77	1.01	0.51	0.58
22(ssh)	0.22	0.06	0.24	0.04
(>=1024)	40.63	9.84	30.22	8.76
1935(rtsp)	2.12	3.91	2.39	3.60
7144(peercast)	0.44	0.04	0.40	0.04
8080	0.30	0.17	0.34	0.19
UDP	12.38	2.94	13.21	2.12
ESP	5.29	1.79	6.54	0.88
GRE	0.16	0.14	0.20	0.06
IP-ENCAP	0.09	0.01	0.13	0.00
L2TP	0.14	0.00	0.09	0.00

表-2 ポート別使用量詳細

コンテンツやソフトウェアアップデートなども含まれているため、コンテンツタイプの特定はできませんが、クライアント・サーバ型の通信量が増えているのは明らかです。

図-7は、全体トラフィックにおけるTCPポート利用の週間推移を、2012年と2013年で比較したものです。ここでは、TCPのポート利用を80番、その他のウェルノウンポート、動的ポートの3つに分けてそれぞれの推移を示しており、ピーク時の総トラフィック量を1として正規化して表しています。2012年と比較すると、全体でも80番ポートの割合が更に増え、動的ポートの割合より大きくなってきているのが確認できます。全体のピークは21時～1時、土日に

は昼間のトラフィックが増加していて、家庭での利用時間を反映しています。

図-8と図-9は、同様にTCPポート利用の週間推移について、クライアント型利用者とピア型利用者に分けて、それぞれ2012年と2013年を比較しています。クライアント型利用者では、大半が80番ポートで、ピーク時間は21時～23時となっており、昨年とほとんど違いは見られません。一方、ピア型利用者については、動的ポートの利用割合が減少しており、2013年には動的ポートが80番ポートより若干多い程度になっています。

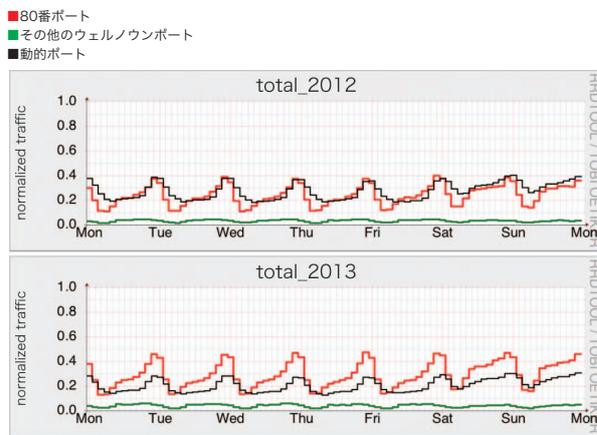


図-7 TCPポート利用の週間推移
2012年(上)と2013年(下)

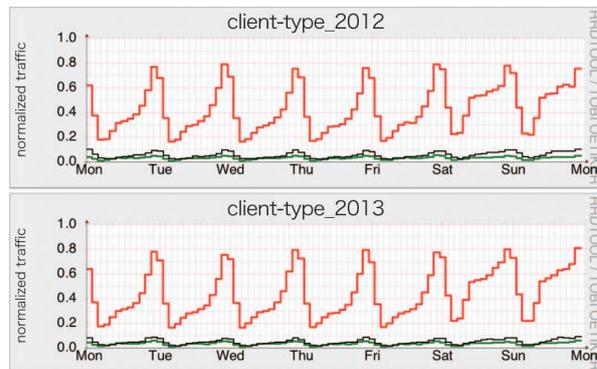


図-8 クライアント型利用者のTCPポート利用の週間推移
2012年(上)と2013年(下)

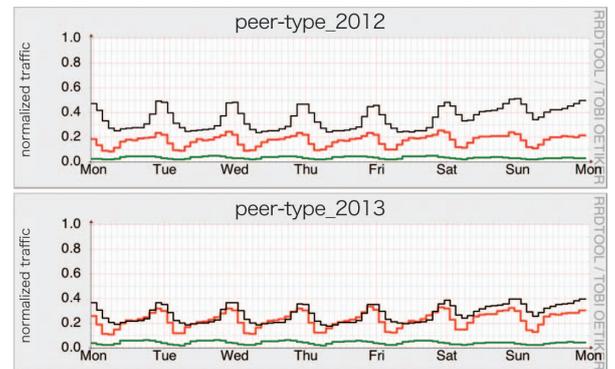


図-9 ピア型利用者のTCPポート利用の週間推移
2012年(上)と2013年(下)

3.5 まとめ

これまで見てきたように、この1年間のブロードバンドトラフィックは、2012年10月前後に増減があるものの、全体のトレンドには大きな変化はなかったと言えます。全体として、ダウンロード量は16%増えました。アップロード量も2010年以降横ばいだったのが、8%の増加に転じました。また、TCP80番ポートの割合が更に増えていて、以前から報告しているWebサービスへの移行が一層進んだことが確認できます。

2012年10月1日に施行された違法ダウンロードの刑事罰化を盛り込んだ改正著作権法の影響は限定的でした。2010年1月のダウンロード違法化のときは、明らかにトラフィックの長期トレンドに影響が出ました。しかし、これは以前のレポートで議論したように、既にあった流れを加速するトリガーとなったに過ぎないという見方もできました。我々は、法的強制措置の効果を検証する意味で、今回の違法ダウンロードの刑事罰化の影響について注目していました。

結局、今回は、施行の前に駆け込みと思われるトラフィック増加が見られ、施行後3ヵ月程トラフィックが減少しま

したが、その後はもとのトレンド曲線に戻っています。マクロなトラフィック傾向だけで違法ダウンロードの増減を推測することはできませんが、今回の施行前後の増減は、一部ユーザの違法ダウンロード行動の反映と見ていいでしょう。しかし、3ヵ月で元のトレンド曲線に戻っていることや、他の数値からも特に傾向に大きな変化が見られないことを併せて考えても、今回のダウンロード刑事罰化は一時的な心理効果しかなかったように見えます。

法整備の役割には、技術的あるいは社会的な背景を持つユーザの行動変化の流れをうまくガイドする、という側面があります。その意味で、2010年の著作権法改正はWebサービスへの移行という流れに沿っており、対象ユーザにも受け入れられたため効果があったと考えられます。その一方で、2012年の改正は対象ユーザに受け入れられなかったように思われます。

IJでは、ユーザの利用形態の変化に迅速に対応できるよう、継続的なトラフィックの観測を行っています。今後も、定期的にレポートをお届けしていく予定です。

執筆者:



長 健二郎 (ちょう けんじろう)
株式会社IJ イノベーションインスティテュート 技術研究所所長。