



インターネットの「視聴率」 って計れないの？

松本 勉

横浜国立大学

● 「視聴率」はスポンサーが払う広告料の目安だ ●

テレビの視聴率は、広告料を収入源とする民放局や関連する番組制作会社や広告代理店やスポンサーにとって、経営を左右する重大な数字だ。調査会社は、各チャンネルが時間帯ごとにモニタ家庭の何パーセントによって視聴されていたかを、モニタ家庭のテレビに接続した計測器から毎早朝に電話回線を通じて報告される前日の視聴状況をベースに統計処理して求める。

モニタ家庭数は、1万世帯に1世帯程度であり、毎月少しずつ入れ替えが行われている模様だ。家族の誰がテレビを視聴したかを当人に計測器上のボタンを押してもらうことで記録できるようにもしている。

インターネットでも広告や情報公開の効果を計る指標が必要だということで、情報提供者自身や調査機関による、ウェブページの視聴数や視聴ユーザの個人情報の調査が求められるようになってきた。WWWは双方向通信路であるため(図-1)、いろいろな方法でアクセス動向を調査でき、さまざまな調査目的に対応しやすい。視聴度、視聴数、広告効果指標などの用語も使われる。

● 何を調査するのか ●

「アクセス」ないし「視聴」というイベントに対して、以下の種類の情報が関連付けられる。

- t 「いつ」(アクセス時間情報) :
アクセス開始時刻、滞在時間など
- u 「だれが」(利用者特定情報) :
姓名、生年月日、住所、個人電話番号、個人メールアドレス、運転免許証番号、保険証番号、クレジットカード番号、銀行口座番号、パスポート番号、公開鍵証明証など
- p 「どんな人が」(利用者特定には不足な属性) :
性別、年齢層、居住域、年収、嗜好、職業など
- q 「どんな状態で」(アクセス履歴) :
直前訪問オブジェクト、何度目の訪問かなど
- f 「どこから」(ネットワーク内の位置) :
IPアドレス、ホスト名など
- v 「どこを通過して」(アクセス経路) :
直接、プロキシ経由、匿名通信路利用など
- c 「どこにある」(キャッシュ利用の種類) :

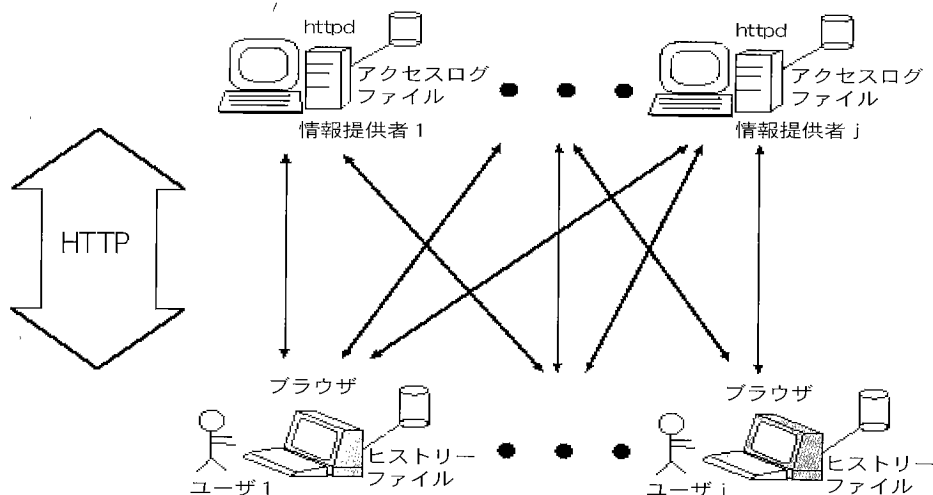


図-1 WWWにおけるアクセスと記録

オリジナル、プロキシキャッシュ、ユーザキャッシュなど

o 「なにに」(アクセス対象) :

訪問オブジェクトのURL・内容など

h 「どんな」(アクセス種別) :

読む、書く、実行するなどの種別

これらのそれぞれ、あるいはそれぞれの間の関係の調査はできるのか、できる場合、どのようにして調査しどのように集計し報告するのか、また、技術的に可能でもプライバシー上の調査を実行して構わないか、といったことが問題となる。典型的にはoが指定ウェブページに等しく、hが「読む」で、tが指定時間帯に含まれるようなアクセスの集合をAとするとき、

視聴数: Aに含まれるアクセスの総数

はいうまでもなく、マーケティング調査などに有用な

利用者属性分布: Aにおけるpの分布

や、クリックスルー(バナー広告をクリックして広告主のページに誘導する仕組み)の効果測定などに有用な

訪問経緯分布: Aにおけるqの分布

などやその推移を、なるべく正確に、かつ、誰かの都合で改変されることなどはないという信頼感を持って、調べたい。しかもなるべく低コストで調査したい。以上では、視聴とアクセスをほぼ同義としたが、視聴をどう定めるか自体についてもいろいろと議論はある。

● どのようにして調べるのか ●

データ収集ポイントにより以下のように大別できる。

C. 利用者側の記録(履歴ファイルなど)に基づく「サンプルアプローチ」

例: NPD社のMedia Metrix方式(<http://www.npd.com>)、ビデオリサーチ社のクライアントモニタ方式¹⁾、坂本らのクライアント監視方式⁵⁾などきわめて多数。基本的にはテレビの視聴率調査と同じ方法であり、利用者ごとにきめ細かい情報が収集できるが、全体の動向調査には多数のサンプルが必要となる。WWWには膨大な数のページがあるので各ページの視聴数の正確な把握は難しい。コンテンツに付加されたチップと呼ばれるデータに対して不特定多数の利用者が自身の属性証明を施して調査機関に匿名で報告するというチップ方式⁴⁾の提案もある。

S. 情報提供者側の記録(アクセスログファイルなど)に基づく「サーバセントリック法」

例: 古典的なアクセスログ解析^{1), 2)}は、視聴数はキャッシュへのアクセスを除き、正確に計数できる。しかし、アクセスログの改ざんなどによる不正への対処が必要。また利用者属性の取得もできない。利用者登

録を行いIDを発行し各サーバに溜まったアクセスログ内の利用者IDを含むデータを調査機関で解析するI/PRO社の方式(<http://www.ipro.com>)は、自己申告ではあるが利用者の属性を取得できる。ブラインド署名技術を用いて利用者側に発行された利用者属性入りクーポンをアクセス時に情報提供サーバに転送する菊池らのIPADECS方式(<http://www.ipa.go.jp/STC/senshin-dokusou-old.html>)^{1), 2), 6)}も実装されている。IPADECSでは、利用者の属性情報の取得、利用者のプライバシー保護、サーバのログ捏造防止、キャッシュへのアクセスの計数につき対応がとられている。

その他、HTTPクッキー処理を伴う方式などもある。

P. 中継者(プロキシサーバなど)でアクセス観測をする「中継者解析法」

例: 雉鳥-後藤の方法^{★1)}、Franklin-Malkhiの方法³⁾など。

解析者に信頼がかければ、情報量の多い分、さまざまな解析ができる。信頼がかけないと危険性もある。また工夫なしでは利用者属性は取得できない。

● 視聴調査のこれから ●

インターネットの「視聴率」をより効果的なものとするには、計測すべきアクセスないし視聴の定義と統計処理方法の標準化、および、利用者、情報提供者、調査機関による不正が生まれにくい仕組みの確立、さらには属性証明機能提供機関などの設立による利用者属性情報の正確性向上によって、集計結果の信用性を高めることが求められよう。また、利用者の知らないうちに利用者に関する情報が採取・処理されることがないように、プライバシー保護の仕組みが最初から組み込まれたシステムの構築も求められよう。

参考文献

- 1) 情報処理振興事業協会技術センター: インターネットにおけるプライバシー技術構築と適用に関する研究開発, 平成8年度研究成果報告書, 8技-166, 情報処理振興事業協会 (Mar. 1997).
- 2) 情報処理振興事業協会技術センター: インターネットにおけるプライバシー技術構築と適用に関する研究開発, 平成9年度研究成果報告書, 9技-76, 情報処理振興事業協会 (Mar. 1998).
- 3) Franklin, M. K. and Malkhi, D.: Auditable Metering with Lightweight Security, Conference Preproceedings of Financial Cryptography '97, Anguilla, BWI (Feb. 24-28, 1997).
- 4) 松本 勉, 清水健介, 岡本克哉: コンテンツへの付加情報を利用する視聴計数スキームとその応用, 電子情報通信学会技術研究報告, ISEC97-21 (July 1997).
- 5) 坂本泰久, 岸 見司, 本橋 健, 小野雄二: クライアント監視方式によるWWWサービス効果測定, マルチメディア, 分散, 協調とモバイル(DiCoMo) ワークショップ, pp.269-274 (July 1997).
- 6) Matsumoto, T.: Getting Users' Attributes without Violating Anonymity, Proceedings of WWCA '98 - International Conference on World Wide Computing and Its Applications, Springer LNCS, No.1368, pp.439-443 (Mar. 1998).

(平成11年1月12日受付)

★1 雉鳥雅章, 早稲田大学卒業論文, 後藤滋樹研究室(1996年度), 再構成版が文献1)の付録に掲載されている。