

インターネットにおけるトラフィック分析

1U-2

小松原 重之 鈴木 亮一 三上 博英

comacie@slab.ntt.co.jp, ryoichi@slab.ntt.co.jp, mikami@slab.ntt.co.jp

日本電信電話株式会社 ソフトウェア研究所 広域コンピューティング研究部

1.はじめに

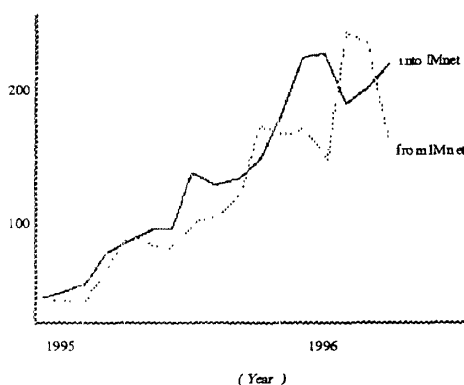
インターネットのトラフィックは、多様なユーザによる利用と、その形態の急激な変化のために複雑な振舞いをする。本研究では、その挙動をモデル化し、インターネットにおけるトラフィック予測技術を確立することを目指している。なお実装と評価に当たっては省際研究情報ネットワーク (Inter Ministry research information Network; 以下IMnet)のトラフィックデータを用いた。

2.評価の期間と変動

同じトラフィックデータであっても評価期間によって、振舞いは異なって見える。以下では年間変動を長期的、月間変動を中期的、それより短い変動を短期的なものとして定義し、それぞれの評価期間における変動の特徴について述べる。

2.1 長期的な変動の特徴

図1はIMnetと、他のNSP(Network Service Provider)の間の、トラフィック推移である。



【図1:IMnetにおける長期的変動】

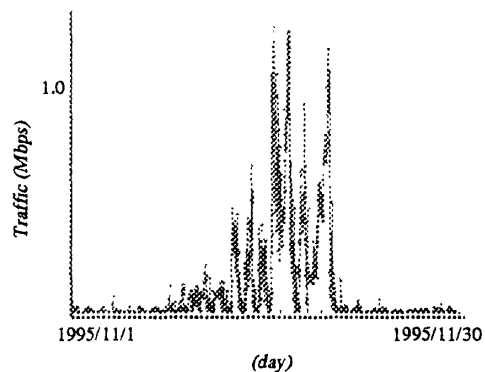
出方向入方向ともに、1年間で10倍以上に増加したことが分かる(この間にIMnetに接続する組織の

The Analysis of the Internet Traffic
Shigeyuki KOMATSUBARA, Ryoichi SUZUKI and Hirohide MIKAMI
NTT Software Laboratories

数は約2倍に増加)。年間のトラフィック変動は、ユーザ数の増加よりも、各ユーザの利用形態の変化による影響の方が大きいことが、推察できる。

2.2 中期的な変動の特徴

図2はIMnetの東京大阪間リンクでの、APEC大阪国際会議(11月14~21日開催)の影響である。

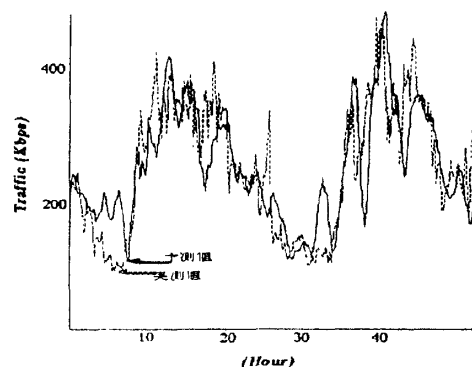


【図2:中期的な変動の例】

このようなイベントに伴うトラフィックの変動は、人気ブラウザの配信や、ネットワークを介した研究実験などにおいても見ることができる。ただしこれらの変動は、再現性が低い。

2.3 短期的な変動の特徴

図3中の実測値は、IMnetのバックボーンのある2日間のトラフィックの推移である。

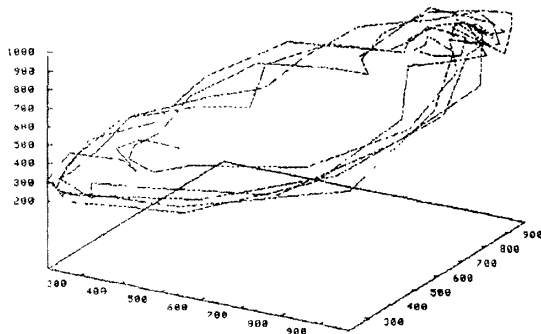


【図3:短期的な変動の例と予測値】

サンプリング・レートは10分間隔であるが、微視的な激しい振動と、低周期の増減について再現性が見られる。この性質に着目したトラフィックの解析を、次の手法により行った。

3. カオス的手法による解析^[1]

図4は、Takensの理論^[2]に基づきIMnetのトラフィックデータから作成した、多次元埋込みアトラクタである(解析次元は3次元)。週日の5日分の短期的なデータによって表現されている。



〔図4:IMnetトラフィックのアトラクタ〕

滑らかな円弧となっている部分では、トラフィックが推移的に振舞うことを示し、右上部の渦状の部分には、昼間の混雑時におけるカオス的な振舞いを示している。これらの性質を利用し、単位時間後のトラフィックを予測するシステムを作成した^[3]。図3中の予測値データは、本システムにより得た予測結果である。10分間隔の予测试行で、平均誤差率12.5%の予測精度を得ている。

4. トラフィック予測システムの実装

実行環境(Sun Sparc5, メインメモリ128MB)において、単位時間以内に、単位時間後のトラフィックを計算するプログラムを作成した。

4.1 リアルタイム性について

本システムは計算に際し、10分間隔のトラフィックデータを格納したデータベースを探索し、中間ファイルをメモリ上に作成する。ここで、データベースアクセスは1エントリあたり10.9msec程度であった。このため探索範囲を8,000時間以上とすると、中間ファイル生成が10分間で行えず、リアルタイム処理が成立しないことが分かった。また、中間ファイルを使う最も大きな処理である解析次元の変更を行っても、処理時間の増加は見られなかった。これらのことから、中間ファイルの生成が処理を律速していることが分かる。現在、その部分の高速化を検討している。

4.2 アルゴリズム

本システムの予測アルゴリズムは、軌跡上の各点について次の計算を行う。

1. 設定された探索範囲の中から、現在の最近傍値を求める
 2. 最近傍値の単位時間後までの推移を求める
 3. 現在値と、上項の推移から予測値を求める
- この処理を繰り返し、逐次的に予測を行う。また、実装にあたっては次の機能を加えている。

1. 週日のデータと週末のデータを分離、独立して計算する機能
 2. データベースの各エントリに対して、あらかじめ時間帯によるグルーピングを行う機能
- 前者は、振舞いの異なるデータを分離し、別々に計算を行うことが目的である。これにより計算量の増加はあるが、3%程度の予測精度向上を実現している。

後者は最近傍値の算出の際に、データベースの探索量を減らすことが目的である。タイムスタンプなどの情報に基づいてインデックスされた集合から探索を行うことで、計算量の削減を実現する。この機能により、データベース処理の簡略化(最大アクセス時の1/144回へ削減)を実現し、予測精度に大きな影響がないことを確認した(予測誤差上昇1%未満)。

5. まとめ

インターネットトラフィックの、長期的、中期的、短期的な変動の特徴について述べた。また、カオス的手法によりアトラクタを作成、その妥当性についてIMnetのトラフィックデータを用いて、予測可能性を示すことで検証した。予測にあたり、データベース処理が律速プロセスとなること、アダプティブな実装によってそれに対応出来ることを示した。

今後は中間ファイル生成ルーチン的高速化を行い、実行環境の強化を図る。また、計算コストと性能(誤差率)の2つの観点から、最適な実装について考察を進める予定である。

〔参考文献〕

1. 中川善富,塚本健一,山田雅之,世木博久,伊藤英則,"カオス的手法を用いた交通量解析",第52回全国大会1Z-2,May1996.
2. F.Takens,"Detecting Strange Attractors in Turbulence, In Dynamical Systems and Turbulence"(eds.D.A.Rand and L.S. Young),Springer-VerlagBerlin,1981.
3. 小松原重之,鈴木亮一,三上博英,"IMnetにおけるトラフィックオペレーションについて",情処研報Vol.96,No.111(ISSN0919-6072)96-DSM-4,1996.