

5U-7

インターネットトラフィック予測システムの構築とその分散オブジェクト指向プラットフォームへの実装

小松原 重之[†] 松重 憲一[‡]
 comacie@slab.ntt.co.jp mken@po.ntts.co.jp
[†]NTTソフトウェア研究所 [‡]NTTソフトウェア

1.はじめに

インターネットにおけるトラフィックの特徴を、従来の電話などの通信網のトラフィックと比較すると、

- よりパースト的な挙動
- 状態遷移が早い
- 定常的な変化は顕著ではない

といった点が挙げられる。従来の通信需要の予測では、ARIMAモデルなどの線形回帰モデルを応用したものが中心であった。しかしインターネットのようなコネクションレス型大規模ネットワークでは、統計的多重効果以上に、個人の振舞いが全体のトラフィックを決定する場合がある。こうしたアーキテクチャにおいて効果的に予測をおこなうためには、より繊細にネットワーク挙動に着目し、かつ多様な観点から解析をおこなうことが必要である。本研究は以前より、決定論的非線型モデルを用いた予測方式の可能性を示してきた。^[1]今回構築したインターネットトラフィック予測システムについて報告し、分散オブジェクト指向プラットフォーム上への実装例について紹介する。

2.決定論的非線型モデルによる予測

予測値の計算には、決定論的非線型モデルを基調とする解析理論を応用している。一本の観測時系列データをTakensの多次元埋込み定理^[2]に従い多次元化、そのアトラクタの状態空間における座標から、ヤコビ行列式を最小二乗法により求め、極小時間後までのトラフィック変化を計算する。従来の方法に比べこの計算をおこなう利点は、系の不安定な挙動を人為的に解釈することなく、予測結果を導出できる点にある。また欠点は、長時間経過した系の状態が予測できないことである。本研究ではこれらのことをふまえ、近い将来のトラフィック値を、実時間内でリアルタイムに予測できるシステムを構築することをコンセプトとした。

3.システム構成

システムは、データや各方式の制御をおこなうシステム本体と、予測値を計算するための複数の方式(予測方式群)からなる。

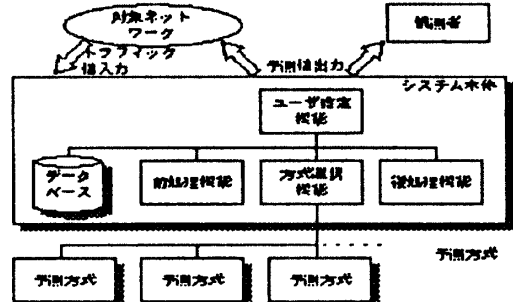
3.1.システム本体

システム本体は、5つの機能から構成される(図1)。

Shigeyuki KOMATSUBARA[†]
 Kenichi MATSUSHIGE[‡]
[†]NTT Software Laboratories
[‡]NTT Software Corporation

- ① ユーザ設定機能は、本システムのコンフィグレーション情報を保持する。
- ② データベース機能は、様々な形式のトラフィックデータを収集し格納をおこなう。新しいデータを受け取ることに、ユーザ設定機能のコンフィグレーションを参照し、予測用データを作成する。
- ③ 前処理機能は、予測用データを受け取ると、ユーザ設定機能のコンフィグレーションを参照し、その設定ファイル中で、埋込み整数次元数が最大、過去の利用データ期間が最長、平均化時間の間隔が最小なデータをメモリ上に作成、保持する。これはデータアクセスの回数を減らし、処理の高速化を実現するためである。
- ④ 方式選択機能は、予測値の計算に参加する方式を管理し、予測的中率テーブルから各方式的の中率を求める。的中率の最も優れた方式(Distinguished Algolysm 以下DA)を決定し、その方式に応じたトラフィックデータを作成する。整形後のデータをDAに渡し、予測結果を受け取る。予測的中率 J_i は、予測値を P 、実測値を R_i として次式により定義する。

$$J_i = \left(1 - \frac{|P_i - R_i|}{R_i}\right) \times 100$$
 (ただし $P_i \geq R_i$ の場合は、 $J_i = 0$)
- ⑤ 後処理機能は、DAに選ばれなかった各方式についてトラフィックデータを作成し、各方式に整形データを渡し、予測値の計算結果を受け取る。それぞれの予測的中率を求め、予測的中率テーブルを更新する。また、実測値および予測値をグラフィに上に出だし、必要であればそれらに条件演算を行い、警報を発する機能がある。



(図1: システム構成)

3.2.予測方式群

予測方式群は、別の、複数の計算機上に分散させる

ことが出来る。方式選択機能との間で通信をおこない、整形データと予測値のみをやりとりする。予測方式群は処理の性質により、2つのカテゴリに分けられる。一つは、決定論的非線型モデルによって、アトラクタの状態空間における挙動から、次点の予測を行うもの。もう一つは、時間の概念によって予測に用いる探索空間を限定することで、特定の時系列挙動を抽出し、予測に反映させる方法である。以下に、それぞれの方式群について述べる。

3.3.状態空間におけるアトラクタ挙動解析による予測方式群

状態空間におけるアトラクタの軌跡について計算を行う方式として、以下の4つの方法がある。

- ① 直近座標の推移ベクトルを求めて写像を計算する方法
- ② 近傍空間にある距離で探索し、存在する座標データから、ヤコビ行列式を求めて写像を計算する方法
- ③ 近傍空間にある距離で探索し、存在するアトラクタに対し、ベクトル比重計算を行うことで推移ベクトルを求め、写像を計算する方法
- ④ 解析次元数(埋込み次元数)を整数次元毎に動的に変化させ、計算する方法

これらの方式によって、系の状態に関して定性的根拠を持たない振舞いをシミュレートさせることが出来る。それにより従来予測が困難であった挙動を持つ系への適応を可能としている。

3.4.時系列類似性に着目した予測方式群

予測に用いるデータを時間パラメータにより抽出し、ネットワーク上の特定のイベントを重点的に学習させることが出来る。具体的には、以下の時間パラメータを用いる。

- ① 週日と週末・祝祭日
- ② 直近1週間
- ③ 直近2週間
- ④ 直近1カ月
- ⑤ 月末の3日間
- ⑥ 週日の午前10時から12時
- ⑦ 週日の午後6時から8時

また、以下のイベントについて既知の場合、別に時間パラメータを作成する。

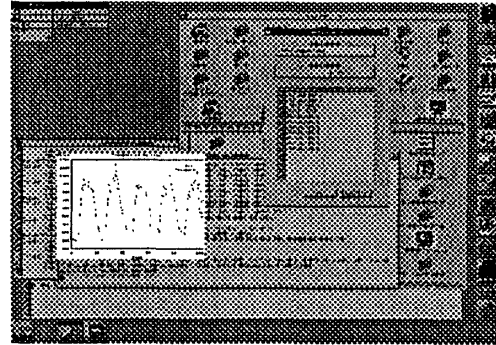
- ⑧ ネットワーク実験
- ⑨ 遠隔会議
- ⑩ 大規模災害

これらの方式によって、既知の定性的な事象について、決定論的非線形モデル上による効率的な予測を可能としている。

4.分散オブジェクト指向プラットフォームへの実装

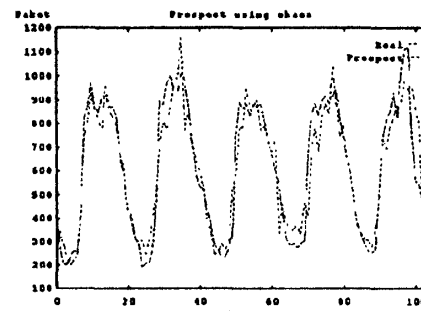
以上のシステム本体および方式群について、分散オブジェクト指向プラットフォームへ実装をおこなっている。

(図2)この技術により、予測方式の数を変化させたり、複数の実行マシンに負荷を分散させることが可能である。トラフィックデータは、MIB(Management Information Base,RFC1156)のOctetsとPacketsの値を用いた。また、実ネットワークで管理されているErrorやDiscardなどのMIB値についても本システムで予測可能とした。



(図2:システム画面の例)

図3は、実ネットワークのトラフィックと、本システムによる予測値を示したものである。平均の予測的中率として87.5%の性能を得た。



(図3:実測値と予測値)

5.おわりに

インターネットの多様性とユーザの広がり、急激な技術進化により、インターネットトラフィックの予測は難しくなりつつある。この問題に対し、決定論的非線型モデルによる予測方式を作成、その派生する方式を性能評価指数によって、フィードバック選択可能な構成を作成した。これにより、刻々と振舞いを変える、様々なネットワークへの適応する予測方式を実現した。現在検討が進められている次世代インターネットプロトコル(IPv6)におけるリアルタイム制御に、この予測システムを応用していく計画である。

文献

- [1] 小松原重之,鈴木亮一,三上博英."インターネットにおけるトラフィック分析".情報処理学会第54回全国大会発表予稿集,1997
- [2] F.Takens,"Detecting Strange Attractors in Turbulence, In Dynamical System and Turbulence",SpringerVerlag,Berlin,pp.366-381,1981