

## トラヒックの時間変動を考慮した動的トラヒックエンジニアリングのためのトラヒック予測

大歳達也† 大下裕一† 村田正幸†  
高橋洋介†† 石橋圭介†† 塩本公平††

† 大阪大学 大学院情報科学研究科  
†† 日本電信電話株式会社  
NTT ネットワーク基盤技術研究所

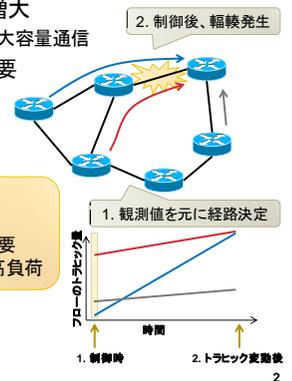
2013/1/18

EICE NS研究会

1

## 背景

- ネットワーク内のトラヒック変化量が增大
  - ・ストリーミング配信・クラウドサービス等の大容量通信
- トラヒックエンジニアリング (TE) が必要
  - ・トラヒックの定期的な観測・経路最適化



従来の TE の問題点

- ・制御以降の変動による輻輳は回避不能
- ・変動への対応を早めるには頻繁な制御が必要  
→ ネットワークの不安定化、観測・制御の高負荷

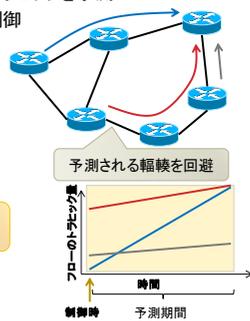
2013/1/18

EICE NS研究会

2

## トラヒック予測を用いた TE

- 方法
  - ・定期的なトラヒックの観測を元に、将来のトラヒックを予測
  - ・予測される輻輳を予め回避するよう経路制御
- 利点
  - ・経路変更後の変動にも対応可能
  - ・経路の切り替え数を削減
    - ・長期的にトラヒックを収容できる経路を設定
    - ・輻輳が予想されない限り経路を継続



TE に適した精度のトラヒック予測が不可欠

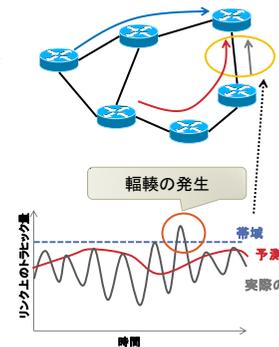
2013/1/18

EICE NS研究会

3

## TE に求められる予測の要件

- トラヒック量の過小評価を回避
  - ・不十分な資源の割り当てにより輻輳が発生する可能性がある
  - ・輻輳発生時、通信品質が大きく下がる
- 過大評価は許容
  - ・余分な資源を必要とし、資源の利用効率率が下がる
  - ・通信品質に影響はない



2013/1/18

EICE NS研究会

4

## 本研究の目的

- 従来のトラヒック予測の研究では、TE に用いるという観点での評価はされていない
  - ・過小評価、過大評価に焦点を当てていない
  - ・単純に予測誤差を小さくすることを目的としている
- TE ではトラヒックの過小評価を優先して避ける必要がある

過小評価を限りなく小さくすることに焦点を当て、TE との連携に適する予測方法を検討する

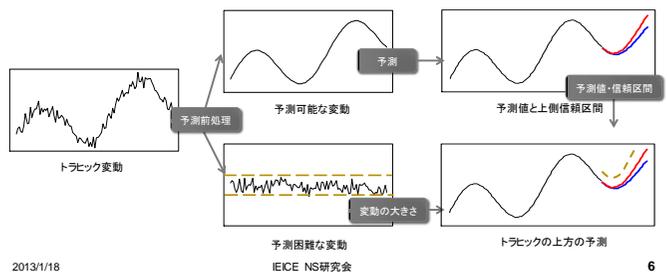
2013/1/18

EICE NS研究会

5

## 予測の流れ

- トラヒック変動では予測可能な変動と予測困難な変動が混在
  - ・予測可能な変動を取り出すことで予測精度の向上
  - ・予測困難な変動の大きさのみ見積り
- 過小評価を避けるためトラヒックの上方を予測



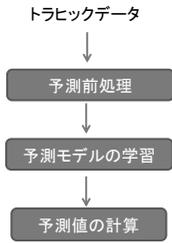
2013/1/18

EICE NS研究会

6

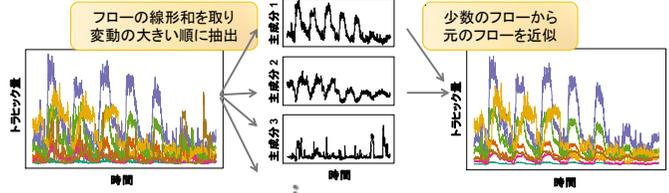
## 予測方法

- 予測前処理
  - 主成分分析：全フローから主要な変動を抽出
  - トレンド成分：各フローの短期的変動を除去
  - 包絡線：各フローの上端の変化を抽出
  - 前処理なし：元のフローをそのまま予測に使用
- 予測モデル
  - ARIMA モデル
    - データの周期性を考慮しないモデル
    - 日周変動の影響が小さい短期の予測に使用
  - SARIMA モデル
    - データの周期性を考慮したモデル
    - 日周変動の影響を考慮した長期の予測に使用



## 主成分分析

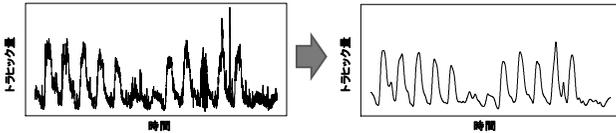
- 多数のフローの中から主要な変動(主成分)を抽出
  - 上位少数個の主成分によりフロー全体の変動を近似可能<sup>[10]</sup>
- 主要な変動に焦点を当てることで予測精度を向上
  - ノイズ等の小さな変動による影響を除去



[10] L. Anukool et al., "Structural analysis of network traffic flows," in *Proceedings of SIGMETRICS/Performance*, vol. 32, no. 1, pp.61-72, Jun. 2004

## トレンド成分

- 各フローのノイズ等の短期的変動を取り除き、長期的な変動を抽出
- 長期的な予測の精度を向上



- トレンド成分モデル<sup>[11]</sup>

$$\begin{cases} x_n = t_n + \varepsilon_n \\ \Delta t_n = \Delta t_{n-1} + w_n \end{cases}$$

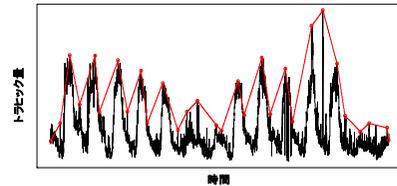
傾きの変化が正規分布に従う

$x_n$  : 観測値,  $t_n$  : トレンド成分,  $\Delta t_n = t_n - t_{n-1}$ ,  
 $\varepsilon_n \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$  : 観測誤差,  $w_n \stackrel{i.i.d.}{\sim} N(0, \tau^2)$  : トレンド内変動

[11] G. Kitagawa and W. Gersch, "A smoothness priors-state space modeling of time series with trend and seasonality," *Journal of the American Statistical Association*, vol. 79, no.386, pp.378-389, Jun. 1984

## 包絡線

- 各フローのトラヒックの上端の変化を抽出
  - 一定間隔毎にピーク点を取り出す
  - 各ピーク点の間を直線で結ぶ
- 直接トラヒックの上方を予測



## TE に適した予測方法の検証

- 過小評価を抑えた時の過大評価
  - 過小評価を、フローの平均値に対して最大 1% 以内に抑える
  - 過大評価は、フローのトラヒックの平均に対する相対誤差を用いる

$$e_{ij} = \begin{cases} \frac{\hat{y}_{ij} - y_{ij}}{\mu(y_j)} & (\hat{y}_{ij} \geq y_{ij}) \\ 0 & (\hat{y}_{ij} < y_{ij}) \end{cases}$$

$\hat{y}_{ij}$  : 予測値,  $y_{ij}$  : 実際のトラヒック,  $\mu(y_j)$  : フローのトラヒックの平均値

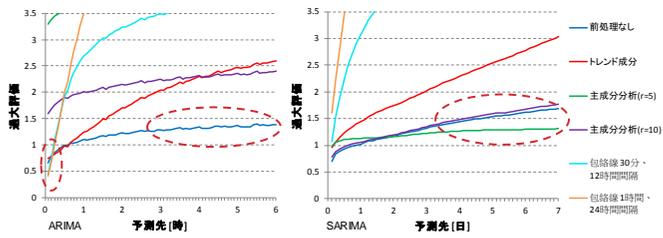
- 計算時間
  - 各前処理を含めた予測に要する時間を評価する

## 検証環境

- データ
  - 米国の研究教育ネットワーク Internet2 のバックボーントラヒックデータ<sup>[15]</sup>
  - 9 個の Point of Presence (PoP) のルータ間のフローのトラヒック
  - 2011/11/28 - 2011/12/25 の 4 週間のトラヒックデータ
- 予測
  - 予測開始時点をずらし複数回予測
    - ARIMA モデルでは 72 回, SARIMA モデルでは 56 回
  - 一点の時間粒度
    - ARIMA モデルでは 5 分, SARIMA モデルでは 3 時間
  - 過去 2 週間のデータを元にモデルを学習

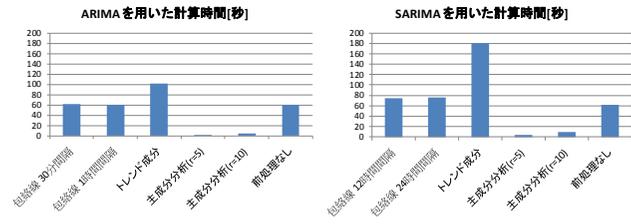
[15] "Internet2 data," available from <http://internet2.edu/observatory/archive/data-collections.html>

### 過小評価を抑えた下での過大評価



- 主成分分析により長期的な予測精度が向上 r: 予測する主成分の数
- 短期の予測では、前処理を行わない予測が高精度
- 直近の予測では、包絡線を用いた予測が高精度

### 予測に要する計算時間



- 主成分分析により計算時間を短縮
  - 全フローの予測を少数のフローの予測で実現

### まとめと今後の課題

- 主成分分析により長期予測の精度向上・予測時間の短縮
- 予測スケールによる精度の違い
  - 直近の予測では包絡線が高精度
  - 短期の予測では元のフローを直接予測する場合が高精度



#### 今後の課題

- タイムスケール毎に適した予測方法の定め方の検討
- 予測されたトラフィック変動を用いたTEの評価