

Hierarchically Structured Spatio-Temporal Traffic Measurement utilizing Compressive Sensing

圧縮センシングを用いた 時空間階層化トラフィック観測

大阪大学 大学院情報科学研究科
情報ネットワーク学専攻 村田研究室
辻 喜宏

2015/2/17

1

トラフィックエンジニアリングのための トラフィック観測

- 変化量が大いネットワークのトラフィックを収容するためには短い時間粒度での TE (Traffic Engineering) が有効
 - 秒オーダーの動的な経路制御がネットワークの性能を向上を示す例 [1]



- TE では制御周期より短い周期でのトラフィック観測情報が必要

リアルタイムにネットワークの状況を把握することが必要

[1] T. Benson, A. Anand, A. Akella, and M. Zhang, "MicroTE: Fine Grained Traffic Engineering for Data Centers," in Proceedings of ACM CoNEXT, 2011, pp. 8:1-8:12.
2015/2/17

2

短い時間間隔でのトラフィック観測の課題

管理サーバは各観測器から定期的に観測情報を収集



2015/2/17

3

短い時間間隔でのトラフィック観測の課題

大規模なネットワークを短い時間間隔で観測する場合、単位時間当たりの到着量は増大



特に管理サーバ付近のリンクの帯域を圧迫する

研究目標
低負荷で収集可能なトラフィック観測手法の提案

2015/2/17

4

提案手法のアプローチ

トラフィックモデルとの差分情報のみを圧縮して送信

- 手順1: 観測器と管理サーバ間で事前に各フローのモデルを生成し共有
- 手順2: モデルからの誤差が存在する場合のみ差分データを生成
- 手順3: 事前に設定した木構造の葉から根までの経路にそって、圧縮を行いながら上位ノードの管理サーバまで差分データを転送

モデル予測値
観測値
差分データ
圧縮差分データ



差分データの圧縮方法として圧縮センシングの技術を応用

2015/2/17

5

圧縮センシング [2]

- データに含まれる冗長性を利用したサイズ削減手法
 - 次の式で生成されたより低次元なベクトル y から元のベクトル x を推定
 $y = Ax$ ($y \in R^M, x \in R^N, M < N$)
 - 次の条件を満たせばわずかな誤差で復元可能
 - ベクトル x のスパース性が高い (非ゼロの要素が少ない)
 - 観測行列 A のコヒーレンスが十分に小さい
- 圧縮 ($x \rightarrow y$) $y = Ax$
- 再構成 ($y \rightarrow x$) $\hat{x} = \arg \min_x \|y - Ax\|_2 \text{ s.t. } \|x\|_0 \leq \theta$

[2] J. Tropp and A. Gilbert, "Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit," IEEE Transactions on Information Theory, vol. 53, no. 12, pp. 4655-4666, Dec 2007.
2015/2/17

6

トラフィック差分データの生成

- 事前に共有したトラフィックモデルからの差分を量子化することで生成
- モデルからの差分データ $\epsilon_f = t_f - z_f$
- 差分データの量子化 $x_f = \text{trunc}(\frac{\epsilon_f}{\rho})$

データのスパース性を向上 (非ゼロの要素数を削減)

ϵ_f : フロー f のモデル誤差
 t_f : 実トラフィック量
 z_f : モデル予測値
 x_f : 差分データ
 ρ : 量子化サイズ

2015/2/17 7

圧縮差分データの収集

x^{Node_i} : 観測器 $Node_i$ で生成された差分データ
 y^{Node_i} : 観測器 $Node_i$ で生成された圧縮差分データ
 A : 観測行列
 C_{Node_i} : 観測器 $Node_i$ の子ノード集合

- 観測器 $Node_i$ では次の y^{Node_i} を上位ノードへ送信

$$y^{Node_i} = Ax^{Node_i} + \sum_{Node_j \in C_{Node_i}} y^{Node_j}$$

地点 $Node_i$ で生成された差分データに圧縮センシングを用いた圧縮を行い得られた値

子ノードから受信した差分データの圧縮値の和

モデル予測値
 差分データ
 圧縮差分データ

- 管理サーバで受信する観測値は差分データ x の圧縮値 y に相当

$$\sum_{i=1, \dots, k} Ax^{Node_i} = A \sum_{i=1, \dots, k} x^{Node_i} = Ax = y$$

2015/2/17 8

トラフィック観測情報の取得

- トラフィック差分データの再構成

$$\hat{x} = \arg \min_x \|y - Ax\|_2^2 \text{ s.t. } \|x\|_0 \leq \theta$$

- 差分データの逆量子化 $\hat{\epsilon}_f = \rho \hat{x}_f$
- 該当トラフィック量の算出 $\hat{t}_f = \hat{\epsilon}_f + z_f$

モデル予測値
 圧縮差分データ

\hat{x}_f : 再構成後のフロー f の差分データ
 $\hat{\epsilon}_f$: モデル誤差の取得値
 ρ : 量子化サイズ
 z_f : モデル予測値
 \hat{t}_f : 該当トラフィックの取得値

2015/2/17 9

評価環境

- 圧縮方法
 - 観測行列 A [3] $A = \arg \min_A \|A^t A - I\|_2^2$
 - 量子化サイズ ρ [0.01, 2.50] [Mbps]
 - 圧縮サイズ 2倍、4倍 (それぞれ $|y| = 50, 25$)
- 観測対象フロー x
 - 観測フロー数 $|x| = 100$
 - 平均フロー量 10 [Mbps]
- モデル誤差 ϵ_f : 各フローはガウス分布 $N_1(0, \sigma_1), N_2(0, \sigma_2)$ のいずれかに従う ($\sigma_1 = 100$ [Kbps], $\sigma_2 = 3$ [Mbps])
- 評価指標
 - モデル誤差が大きいフロー数が占める割合を変化させ、各圧縮サイズに対する観測情報に含まれる最悪誤差 $\max_{i=1, \dots, n} (|\hat{\epsilon}_f - \epsilon_f|)$ の値を確認

y : 収集データ
 x_f : 差分データ
 ϵ_f : フロー f のモデル誤差
 ρ : 量子化サイズD

[3] "On Optimization of the Measurement Matrix for Compressive Sensing," in Proceedings of European Signal Processing Conference, 2010, pp. 427-431.

2015/2/17 10

観測情報の最悪誤差 $\max_{i=1, \dots, n} (|\hat{\epsilon}_f - \epsilon_f|)$

2倍に圧縮した場合

4倍に圧縮した場合

- 圧縮サイズに従って誤差は大
- モデル誤差が大きいフロー数を 25% に抑えれば最悪誤差 1 [Mbps] で収集情報量を半分に削減可能

2015/2/17 11

まとめと今後の課題

- まとめ
 - 時間モデルからの差分データを圧縮センシングを用いて圧縮し、観測器間で集約していくことで、観測情報収集のオーバーヘッドを削減できる手法を提案した
 - 評価の結果
 - 誤差が大きいフロー数を 25% に抑えることができれば、観測情報の誤差を 1 [Mbps] 以内で収集情報量を 50% 削減可能なことを示すことができた
- 今後の課題
 - トラフィックモデル生成手法と連携させた評価
 - フローの空間相関性を利用した、さらなる観測負荷の削減手法の検討

2015/2/17 12