

氏 名 Le Van An

学位(専攻分野) 博士(情報学)

学位記番号 総研大甲第 2360 号

学位授与の日付 2022 年 9 月 28 日

学位授与の要件 複合科学研究科 情報学専攻
学位規則第6条第1項該当

学位論文題目 Traffic Matrix Prediction and Traffic Engineering Leveraging
Machine Learning Techniques

論文審査委員 主 査 計 宇生
情報学専攻 教授
栗本 崇
情報学専攻 准教授
金子 めぐみ
情報学専攻 准教授
阿部 俊二
情報学専攻 准教授
福田 健介
情報学専攻 准教授
John C.S. Lui
香港中文大学 計算機学科 教授

Summary of Doctoral Thesis

Name in full Le Van An

Title Traffic Matrix Prediction and Traffic Engineering Leveraging Machine Learning Techniques

The Cisco Annual Internet Report projects that there will be 5.3 billion total Internet users by 2023, up from 3.9 billion in 2018. The global fixed broadband speeds will reach 110.4 Mbps, up from 45.9 Mbps. The great demand for Internet services (e.g., video streaming, VoIP, etc.) leads to an exponential growth in the backbone network traffic and creates a huge challenge for traffic monitoring and traffic engineering. In this dissertation, we focus on solving the traffic engineering problem (TE) in the backbone network by leveraging Machine Learning (ML) techniques. We consider a network system managed by a centralized controller. The controller has the role of collecting network states (e.g., measuring the traffic volume) and generating the routing rules. Our main objectives are to optimize the network resources and increase the Quality of Service (QoS) by minimizing the network monitoring overhead and avoiding/reducing traffic congestion.

In traffic prediction-based TE approach, we first utilize machine learning models to have accurately predicted the values of future traffic demands. Then, based on the prediction results, we calculate the routing rules that can adapt to the dynamic traffic and avoid congestion. However, directly applying machine learning techniques suffers from several problems. The first and second problems are related to the low prediction accuracy caused by the missing data in both model's training and testing phases. The first problem is the traffic matrix prediction with partial information. To reduce traffic monitoring cost, the controller may not measure all the network information. It leads to the low performance in the traffic prediction as well as the TE tasks. The second problem is the missing values in the network datasets. Although missing values is a common problem in machine learning, without being properly handled, it can degrade the performance of the trained deep learning model. Then, we face the problem of frequent traffic rerouting. To optimize the network routing, most of the proposed TE solutions only address the optimization problem in a single snapshot by frequently changing the paths of traffic flows which leads to the degradation in the overall network's QoS. Finally, when applying ML techniques, many prior studies assume that network information such as traffic matrix or link utilization is available. However, with the explosion of traffic and the expansion of the physical network, obtaining all the network statistics imposes high monitoring overhead. In addition, it is required a huge amount of data for model training/predicting processes.

In this work, we tackle the aforementioned problems when applying ML for traffic prediction and traffic engineering: (1) improving the prediction accuracy under partial traffic monitoring, (2) inferring the missing values in the network datasets, (3) suppressing the number of traffic rerouting flows, and (4) reducing the network monitoring overhead.

Firstly, to improve the performance of the traffic prediction model under partial traffic monitoring, we utilize the Convolutional Long Short-Term Memory (ConvLSTM) model and introduce the data correction algorithm to correct the outputs of the prediction model. Then, we propose an algorithm to decide the set of monitored flows in the partial monitoring flow approach. By using the proposed algorithm, we not only reduce the monitoring overhead but also achieve higher prediction accuracy. Secondly, we address the missing values problem in network datasets. Based on Recurrent Neural Network (RNN) and Graph Convolutional Neural Network (GCN), we propose a novel deep learning model to efficiently learn the dynamic correlations in partially observed data and estimate the missing values. Then, we mitigate the traffic rerouting issue in the third problem by optimizing the routing over a long-time horizon.

Specifically, we introduce a routing scheme called Multi-time-step Segment Routing (MTSR) by taking the advantage of the DNN models in traffic prediction of multiple time-steps ahead. Based on the concept of source routing, Segment Routing (SR) allows the source or ingress node to inject a sequence of segment labels into the packet header and specify the routing path. Due to the routing flexibility, SR has been widely used to solve TE problems such as minimizing the maximum link utilization of a network. In addition, to lower the cost of network monitoring, we propose an approach that combines traffic prediction with compressive sensing techniques.

Specifically, we first leverage the DNN to predict a partial future traffic matrix using a small amount of the observed traffic and then utilize the compressive sensing technique to reconstruct the whole traffic matrix.

To evaluate the performance of the proposed approaches, we conduct experiments using different backbone network traffic datasets. Through extensive experiments, we show that our deep learning model and the proposed multi-time-step routing algorithm can overcome the impact of the missing data problem, achieve near-optimal performance in network traffic routing, and significantly reduce the number of traffic rerouting.

博士論文審査結果

Name in Full

氏 名

Le Van An

Title

論文題目 Traffic Matrix Prediction and Traffic Engineering Leveraging Machine

Learning Techniques

本学位論文は、「機械学習を駆使したトラフィックマトリクス予測とトラフィックエンジニアリング」と題し、全6章から構成されている。

第1章と第2章では研究の背景を紹介し、ニューラルネットワーク(NN)を用いたネットワークトラフィック予測とトラフィックエンジニアリングの課題として、トラフィック計測のコスト、データの欠損、および経路変更による品質への影響について問題提起し、既存研究について説明した。

第3章では部分情報によるトラフィックマトリクス予測方法について述べている。トラフィックを正確に予測することは経路制御、容量設計、品質保証にとって重要であるが、ネットワークすべてのノード間に流れるトラフィック（すなわち、すべてのフロー）を計測することは大きなコストを要する。研究では前後方向の補正ができる回帰型 NN を利用して、一部のフローのみの計測で全フローを表すトラフィックマトリクス予測を行い、予測の精度を高めるために計測すべきフローの指定方法の指針を示し、実データを用いた評価で有効性を明らかにした。

第4章では欠損したネットワークトラフィックデータの補間について説明している。トラフィックデータに欠損が生じる場合、それを用いた高度な解析やトラフィック制御の性能に影響を与えることが考えられる。研究ではグラフ畳込み NN と双方向回帰型 NN を組合せて作られたグラフ畳込み回帰型 NN を用いてトラフィックデータの時空間関係を学習し、欠損したトラフィックデータの補間を行った。また、各種の欠損パターンについて評価を行い、提案方法による補間の有効性と、補間したデータを用いたトラフィック制御の効果を示した。

第5章ではトラフィック予測の結果を利用したトラフィックエンジニアリングについて述べている。トラフィックエンジニアリングはネットワークの輻輳を回避するための経路制御手段であるが、トラフィックのダイナミックな挙動による頻繁な経路変更は通信品質への影響からそれを避ける必要がある。研究では、一定時間間隔内の最大トラフィックを予測し、輻輳回避性能を維持しながら、経路変更の回数を著しく減らせるようなセグメントルーティング（経路を複数のセグメントに分けるための迂回点を探索する）アルゴリズムを示した。また、圧縮センシングを利用して一部のトラフィックを計測するだけでトラフィックマトリクスを復元して将来の最大トラフィックを予測し、トラフィックエンジニアリングを行う方法を紹介した。

最後に第6章では研究結果をまとめ、未解決課題と今後の方向性を提示した。本博士研究は伝統的なトラフィックエンジニアリングが抱える諸問題を解決すべく、機械学習技術

適用の可能性を探究し、その有効性を示したものであると結論づけられた。

公開発表会では博士論文の章立てに従って発表が行われ、その後に行われた論文審査会及び口述試験では、審査員からの質疑に対して適切に回答がなされた。

質疑応答後に審査委員会を開催し、審査委員で議論を行った。審査委員会では、出願者の博士研究の学術レベルが十分に高く、実用的な意義も有していることが評価された。

以上を要するに本学位論文は、機械学習技術を駆使してネットワークトラフィックの予測、補間、およびトラフィックエンジニアリングにおける諸問題の解決方法とその有効性を示したものであり、情報通信研究分野の発展に寄与しているという点で学術的価値が大きい。また、本学位論文の成果は、情報学専攻の定めるトップ国際会議論文3件として発表され、社会的な評価も得ている。以上の理由により、審査委員会は、本学位論文が学位の授与に値すると判断した。