Master's Thesis

Title

Predictive traffic engineering incorporating real-world information inspired by human brain cognition process

Supervisor

Professor Masayuki Murata

Author

Kodai Satake

February 9th, 2018

Department of Information Networking

Graduate School of Information Science and Technology

Osaka University

Master's Thesis

Predictive traffic engineering incorporating real-world information inspired by human brain cognition process

Kodai Satake

Abstract

The amount of traffic in the Internet has been increasing both in quantity and in fluctuation as the devices connected to the Internet and the services on the Internet become popular. Predictive Traffic Engineering (TE) is one approach to accommodating such fluctuating traffic without congestion. In the predictive TE methods, a controller collects the traffic information, predicts the future traffic, and changes the routes or resource allocation based on the predicted traffic.

The accuracy of the prediction is important for the predictive TE; if the predicted traffic is inaccurate, the resources may not be allocated properly and congestions may occur. Many methods to predict the future traffic have been proposed. Most of them model the traffic changes based on the time series of monitored traffic and predict the future traffic using the model. However, it is difficult to accurately predict traffic only from the previously monitored traffic, if the signs of the fluctuation are not included in the previously monitored traffic.

The real-world information is useful for the prediction of future traffic, because the real-world information may include the signs of the traffic fluctuation, which are not included in the previously monitored traffic data. For example, the number of people in each area can increase the accuracy of the prediction, because the traffic in an area becomes large if the number of people in the area increases. In addition, if the number of peoples in the nearby areas increases, we can easily predict that the number of people in the area will also increase, which causes the increase of the traffic from the area.

In this thesis, we propose a predictive traffic engineering method which predicts future traffic using the information monitored in the real-world. Though the real-world information may contribute to the accurate prediction of future traffic, it is difficult to model the relation between future

traffic and the real-world information. That is, we need a new method to predict future traffic using such information whose relation to the future traffic cannot be clearly modeled.

Therefore, we propose a prediction method inspired by the human-brain cognition process which makes decisions from uncertain information. In this model, a human brain has stochastic variables and updates the variables by Bayesian inference every time a new observation is obtained. Then, a human brain makes a decision based on the stochastic variables.

In our method, we define multiple states by the monitored information including both of traffic and real-world information. In addition, our method learns the future traffic corresponding to each state. Then, our method predicts future traffic by deciding the current state from the traffic and real-world information by the process inspired by the human-brain cognition process; our method has the stochastic variables indicating the confidence about that the current traffic and real-world information belong to the corresponding state, and updates the variables every time a new observation is obtained. Finally, our method allocates the resources based on the future traffic corresponding to the states whose confidences are high.

We evaluate our method by simulation. The results demonstrate that our method avoids congestions without requiring a large amount of extra resources; the amount of resources required to avoid congestion is reduced by 25 % compared with the predictive TE using only the past traffic information.

Keywords

Traffic Engineering

Resource Allocation

Real-World Information

Human Brain Cognition

Contents

List of Figures

List of Tables

1 Introduction

The amount of traffic through networks has been increasing both in quantity and in fluctuation as the devices connected to the Internet and the services on the Internet become popular. Network operators need to accommodate such fluctuating traffic without congestion. Traffic Engineering (TE) is one approach to accommodating such fluctuating traffic without congestion [1–5]. In these methods, routes and/or resource allocations are changed dynamically so as to accommodate the traffic without congestion.

Most of the methods that dynamically control network resources use the observed network traffic. However, the resource allocated based on the observed traffic does not suit the actual traffic when significant traffic change occurs, but the configured resource allocation is not changed until the next control cycle. This problem may be solved by setting the short control interval. However, the short control interval caused the network stabilization.

One approach to allocating the fluctuating traffic without causing the network stabilization is the predictive TE [6]. In the predictive TE, a controller collects the traffic information and predicts the future traffic. Then, the controller allocates the resources based on the predicted traffic. The predictive TE allocates a sufficient amount of resources to avoid the congestion without setting the control interval short.

The accuracy of the prediction is important for the predictive TE; if the predicted traffic is inaccurate, the resources may not be allocated properly and congestions may occur. Many methods to predict future traffic have been proposed [7–10]. For example, Yu et al. proposed the traffic prediction that combines ARIMA and FARIMA based on the multifractal spectrum for mobile networks, and Feng et al. compared the prediction models such as IMA, FARIMA, ANN and wavelet-based prediction and demonstrated that the optimal model depends on the network.

Most of the traffic prediction models the traffic changes based on the time series of monitored traffic and predict the future traffic using the model. However, it is difficult to accurately predict traffic only from the previously monitored traffic, if the signs of the fluctuation are not included in the previously monitored traffic.

The real-world information can contribute the accurate prediction of future traffic, because the real-world information may include the signs of the traffic fluctuation, which are not included in the previously monitored traffic data. For example, the number of people in each area can improve the accuracy of the prediction, because the traffic in an area becomes large if the number of people in the area increases. In addition, if the number of peoples in the nearby areas increases, we can easily predict that the number of people in the area will also increase, which causes the increase of the traffic from the area.

In this thesis, we propose a predictive traffic engineering method which predicts future traffic using the information monitored in the real world. Though the real-world information may contribute the accurate prediction, it is difficult to model the relation between the future traffic and the real-world information. That is, we need a new method to predict future traffic using such information whose relation to the future traffic cannot be clearly modeled.

In this thesis, we propose a prediction method inspired by the human-brain cognition process which makes decisions from uncertain information. Bayesian decision-making theory is one of the theoretical models that explain the process human brain makes decisions based on uncertain information. Bayesian decision-making theory treats observed information and the confidence of cognitive objects as stochastic variables. Then, the variables are updated by Bayesian inference every time a new observation is obtained. Finally, a human brain makes decisions based on the stochastic variables.

Bayesian Attractor Model (BAM) is one of the cognitive models of brain based on Bayesian decision-making theory [11]. In this model, the cognitive options are embedded as attractors. Then, a brain has stochastic variables related to the options, and recognize which option is suitable by updating the variables by Bayesian inference.

In our method, we define multiple states by the monitored information including both of traffic and real-world information. In addition, our method learns the future traffic corresponding to each state. We embed the defined states as attractors. Then, our method predicts future traffic by deciding the current state from the traffic and real-world information by the process inspired by BAM; our method has the stochastic variables indicating the confidence about that the current traffic and real-world information belong to each state, and updates the variables every time a new observation is obtained. Finally, our method allocates the resources based on the future traffic corresponding to the states whose confidences are high.

The rest of this thesis is organized as follows. Section 2 explains the related work. Section 3 explains Bayesian Attractor Model (BAM). Section 4 proposes the predictive traffic engineering incorporating real-world information. Section 5 evaluates our method. Section 6 concludes this thesis.

2 Related Work

2.1 Traffic Engineering

The amount of traffic through networks has been increasing both in quantity and in fluctuation. A network must accommodate such fluctuating traffic without congestion. Dynamically changing the routes and/or resource allocations is one of the promising approaches to accommodating such fluctuating traffic, and many methods to change the routes and/or resource allocations, which are called *traffic engineering (TE)* have been proposed.

Many papers on the TE focus on the routes in the backbone networks or data center networks. For example, Chiesa et al. discussed algorithms to set weights for routing for the case of Equal-Cost-Multi Path (ECMP) [2]. Akyildiz et al. discussed the traffic engineering methods for the software defined networks (SDN), where a route of a flow is set by setting the flow tables of the switches.

The methods to allocate the resources for mobile networks have also been proposed. Ramanathan et al. proposed a method to dynamically allocate resources so as to provide continuous service to mobile users by estimating the resource requirements of potential handoff connections [12]. Zulhasnine et al. discussed the problem on the resource allocation between D2D communication and the cellular networks [13]. They formulated the resource allocation problem by a mixed integer nonlinear programming and proposed a heuristic method to solve the problem. Lopez et al. proposed a distributed and coordinated radio resource allocation algorithm for cellular networks [14]. This method dynamically allocates the modulation and coding scheme (MCS), resource block (RB), and transmit power so that the users ' demands are satisfied.

Most of the methods that dynamically control network resources use the observed network traffic. However, the resource allocated based on the observed traffic does not suit the actual traffic when significant traffic change occurs, but the configured resource allocation is not changed until the next control cycle. This problem may be solved by setting the short control interval. However, the short control interval caused the network stabilization.

One approach to allocating the fluctuating traffic without causing the network stabilization is the predictive TE [6]. In the predictive TE, a controller collects the traffic information and predicts the future traffic. Then, the controller allocates the resources based on the predicted traffic. The predictive TE allocates a sufficient amount of resources to avoid the congestion without setting the control interval short.

The accuracy of the prediction is important for the predictive TE; if the predicted traffic is inaccurate, the resources may not be allocated properly and congestions may occur. That is, we need the method to predict the future traffic accurately.

2.2 Traffic Prediction

The prediction of future traffic is required by the predictive TE. There are many papers on traffic prediction.

Rutka proposed a method to predict future traffic by using the self-similarity model [8]. He applies a neural network to learn the predictive model using the self-similarity and predicts the future traffic from the history of the traffic data by the proposed model. Lu proposed a prediction method using the RBF neural network and optimized its hyperparameters by the Genetic Algorithm [9]. Yu et al. proposed the traffic prediction that combines ARIMA and FARIMA based on the multifractal spectrum for mobile networks [15]. Feng et al. compared the prediction models such as IMA, FARIMA, ANN and wavelet-based prediction and demonstrated that the optimal model depends on the network [16].

The above methods predict the future traffic based on the time series of monitored traffic and predict the future traffic using the model. However, it is difficult to accurately predict traffic only from the previously monitored traffic, if the signs of the fluctuation are not included in the previously monitored traffic.

The real-world information may contribute the accurate prediction of future traffic, because the real-world information may include the signs of the traffic fluctuation, which are not included in the previously monitored traffic data. For example, the number of people in each area may improve the accuracy of the prediction, because the traffic in an area becomes large if the number of people in the area increases. In addition, if the number of peoples in the nearby areas increases, we can easily predict that the number of people in the area will also increase, which causes the increase of the traffic from the area.

Therefore, in this thesis, we propose a method that predicts the future traffic by using not only the traffic information but also the real-world information.

2.3 Bayesian Decision Making

In this thesis, we predict the future traffic by the method inspired by the cognitive process of the human brain.

There are many papers that formalize the human decision making. One of the models of the human decision making is the Bayesian decision-making theory. There is converging evidence from various communities that Bayesian approaches can serve as a coherent description of human decision making [17].

Bayesian decision-making theory treats observed information and the confidence of cognitive objects as stochastic variables. Then, the variables are updated by Bayesian inference every time a new observation is obtained. Finally, a human brain makes decisions based on the stochastic variables.

The mechanisms of the neurons that make decisions based on Bayesian decision-making theory have also been being investigated. Ma et al. have investigated how the neuron encodes the probabilistic distribution [18]. They argued that the high variability in the responses of cortical neurons implies that populations of neurons automatically represent probability distributions over the stimulus. They demonstrated that the Poisson-like variability observed in cortex reduces a broad class of Bayesian inference to simple linear combinations of populations of neural activity.

In this thesis, we use one of the models of the human decision making, which was proposed by Bitzer er al. [11]. The detail of this model is explained in Section 3.

3 Bayesian Attractor Model(BAM)

The Bayesian Attractor Model (BAM) models the process by which the brain makes decisions based on uncertain sensing information [11]. The BAM encodes the predefined i options $\phi_1, \cdots, \phi_i$ called attracter, and makes decisions that which options the current status is. The BAM has the decision state $\boldsymbol{z_t}$ as its internal state, and updates $\boldsymbol{z_t}$ based on the observation value $\boldsymbol{x_t}$ obtained from the outside by performing the Bayesian inference. The rest of this section explains how the states are updated and decisions are made in BAM.

3.1 Update of decision state

BAM has the following generative model of the decision state z_t and observation $\boldsymbol{x_t}$.

$$\boldsymbol{z_t} - \boldsymbol{z_{t-\Delta_t}} = \Delta_t f(\boldsymbol{z_{t-\Delta_t}}) + \sqrt{\Delta_t}\boldsymbol{w_t} \tag{1}$$

$$\boldsymbol{x_t} = M\sigma(\boldsymbol{z_t}) + \boldsymbol{v_t} \tag{2}$$

where $f(z)$ is the Hopfield dynamics, $\boldsymbol{w_t}$, $\boldsymbol{v_t}$ are Gaussian noise variables, and $M = [\mu_i, \cdots, \mu_N]$ is a matrix indicating the observation values and μ_i is the observation value corresponding to the state ϕ_i, which is the i-th predefined attractor. $\sigma(x)$ is a sigmoid function $\frac{tanh(ax/2)+1}{2}$ where a is slope of sigmoid function.

In the BAM, the decision state z_t is updated every time the observations $\boldsymbol{x_t}$ is obtained by inverting the generative model using Bayesian inference. The generative model is nonlinear, Bitzer et al. uses the Unscented Kalman Filter [19] to update the mean decision state of z_t. In addition to updating of the mean decision state, the posterior distribution $P(z_t|x_t)$ over the decision state is also obtained.

3.2 Decision making

The above state estimation outputs the posterior probability $P(z_t|x_t)$. Thus, the decision is made by handling the probability. Bitzer et al. introduced the threshold λ. When $P(z_t = phi_i) > \lambda$, it selects the option ϕ_i. When $P(z_t = \phi_i) \leq \lambda$ for all i, the decision is not made until a new observation is obtained.

4 Predictive traffic engineering incorporating real-world information

4.1 Overview

In this thesis, we propose a predictive traffic engineering incorporating real-world information for mobile networks. In mobile networks, the traffic from each area may change in time due to the change in the traffic generated by each user and/or the change in the number of people in the area. Network operators must prepare a sufficient amount of resources for each area to accommodate the traffic without congestion. In this thesis, we discuss the method to predict the future traffic from each area and determine the amount of resources required for each area.

In mobile networks, the number of people in each area is useful information to predict the traffic from the area, because the traffic in an area becomes large if the number of people in the area increases. In addition, if the number of peoples in the nearby areas increases, we can easily predict that the number of people in the area will also increase, which causes the increase of the traffic from the area. Therefore, our method uses the real-world information such as the number of users in each area in addition to the traffic volume of each area.

Though the real-world information may contribute the accurate prediction of the future traffic, it is difficult to model the relation between future traffic and the real-world information. That is, we need a new method to predict future traffic using such information whose relation to the future traffic cannot be clearly modeled.

In this thesis, we propose a prediction method inspired by the human-brain cognition process which makes decisions from uncertain information. In this model, a human brain has stochastic variables and updates the variables by Bayesian inference every time a new observation is obtained. Then, a human brain makes a decision based on the stochastic variables. Figure 1 shows the overview of our method.

The rest of this section explains how our method predicts future traffic and how our method allocates resources based on the prediction.

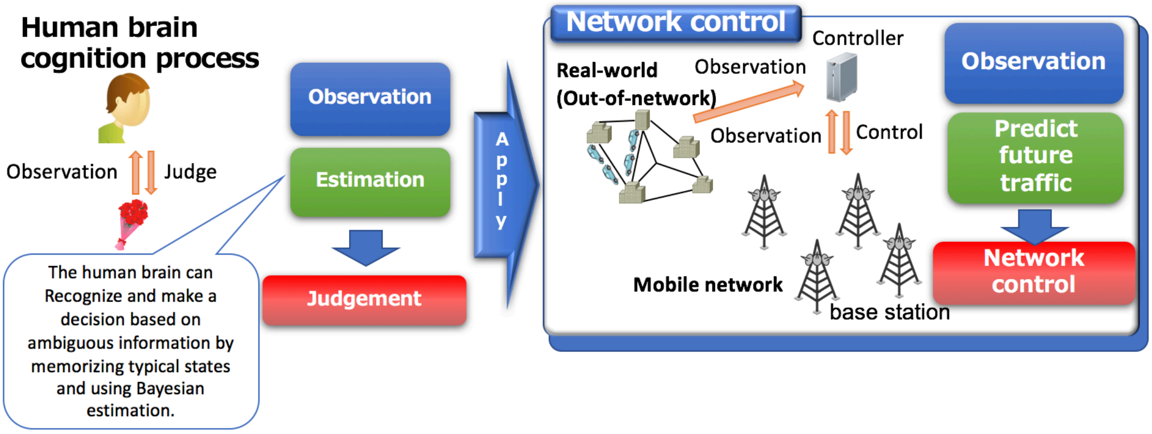

Figure 1: Application of human brain cognition process to TE

4.2 Traffic prediction based on human brain cognition process

Our method predicts the future traffic from each area and allocates the resources based on the prediction for each area. In this section, we focus on an area and predict traffic in the area at time slot $t + p$ by using the information monitored at time slot t.

In our method, we define the state of the network by the observation information. We also assign the future traffic amount for each state by using the observed information. By doing so, we can predict the future traffic by deciding the state of the network from the observation information. To decide the state of the network we use the Bayesian Attractor Model (BAM) [11], which is one of the cognitive models of the human brain.

The rest of this subsection explains the observation information used for prediction, the definition of the state of the network, and how to apply BAM to make decisions of state.

4.2.1 Observed information

In this thesis, we use the number of users in each area in addition to the amount of traffic from the area. To predict the number of users in each area, the numbers of users in the nearby areas are also useful information; if the number of peoples in the nearby areas increases, we can easily predict that the number of people in the area will also increase, which causes the increase of the traffic from the area. Thus, we also use the information of nearby areas. In addition to the absolute values of the number of users and the amount of traffic, whether the values are increasing or not is useful; the increase of the number of people in nearby areas may be a sign that the number of people and/or the amount of traffic will increase. Therefore, we also use the increase rates of the traffic and the number of people.

In this thesis, our method uses the following information in the areas whose distance from the area whose future traffic is to be predicted is less than m when predicting the traffic at the time slot $t + p$

- Traffic amount at the time slot t
- Difference between traffic amounts at the time slot $t - p$ and the time slot t
- Number of users at the time slot t
- Difference between the number of users at the time slot $t - p$ and the time slot t

4.2.2 State

If the information observed at the time slot t is similar to that observed at the time slot t', the traffic amount at the time slot $t + p$ is similar to the traffic amount at the time slot $t' + p$. Thus, we define the state by clustering method; we divide the observation information that is collected in advance into k clusters $C_1, C_2, \cdots, C_k$ so that each cluster includes the similar information. Each cluster indicates the state to be determined by the decision making.

We also assign the future traffic for each cluster. The future traffic for the cluster C_n is determined by

$$T_n^{\text{future}} = \max_{t \in C_n} T_{t+p} \tag{3}$$

where T_{t+p} is the traffic amount at the time slot $t + p$. By defining the future traffic by Eq. (3), we can avoid the case that the traffic amount at the time slot $t' + p$ for $t \in C_n$ becomes larger than the predicted traffic for the cluster C_n. That is, we can allocate a sufficient amount of resources to avoid congestion by using the predicted traffic.

In this thesis, we define the k cognitive states based on observation information using k-means method.

4.2.3 Application of BAM to state cognition

In this thesis, we use k decision makers so that ith decision maker decides whether the current state belongs to the ith option (i.e. ith cluster defined in the previous subsection). Each decision maker performs decisions based on BAM.

Hereafter, we explain how to use the observation information in each decision maker and definition of attractor in BAM.

How to use observed information In each decision maker, observations are used to determine whether the current state belongs to the target cluster or not. In this thesis, we calculate the input of the BAM model $\boldsymbol{x_i}$ in each decision maker so that the BAM model can easily obtain the information on whether the current state belongs to the target cluster or not

$$\begin{aligned} \boldsymbol{x_i} &= \sigma(\frac{a}{a+b}, \frac{b}{a+b}) \\ a &= D(\boldsymbol{X}, \boldsymbol{M_i}) \\ b &= min(D(\boldsymbol{X}, \boldsymbol{M_1}), D(\boldsymbol{X}, \boldsymbol{M_{i-1}}), \cdots, D(\boldsymbol{X}, \boldsymbol{M_{i+1}}), D(\boldsymbol{X}, \boldsymbol{M_k})) \end{aligned} \tag{4}$$

where $\boldsymbol{x_i}$is the input for BAM to make a decision whether the current state belongs to the cluster $C_i (1 \leq i \leq k)$ or not, $\boldsymbol{X}$ is the vector of the current observation value, $D(\boldsymbol{y_1}, \boldsymbol{y_2})$ is the is the Euclidean distance between vectors $\boldsymbol{y_1} and \boldsymbol{y_2}$, σ is a sigmoid function. $\boldsymbol{y_i}$ is the centroid of the observation information belonging to ith Cluster.

$\boldsymbol{x_i}$ become close to $0, 1$ when observation value is close to the centroid of observation information in the cluster C_i. On the other hand, $\boldsymbol{x_i}$ becomes close to $1, 0$ when the observation value is close to the other clusters.

Definition of attractor Each decision maker decides whether the current state belongs to the target cluster or not. That is, each decision maker has two attractors z_{yes} and z_{no}; $z_y es$ corresponds to the case that the current state belongs to the target cluster, and z_{no} corresponds to the case that the current state belongs to the other cluster. Then, we define the observed values for z_{yes} and z_{no} as follows.

$$\begin{aligned} \mu_{yes} &= (0, 1) \\ \mu_{no} &= (1, 0) \end{aligned} \tag{5}$$

4.3 Resource allocation based on prediction

The above procedure outputs the posterior probability $P(z_{yes}|x)$ (hereinafter we call confidence) and the predicted traffic amount for each cluster. In this thesis, we allocate resources based on this confidence.

If the current observation information clearly indicates that the current state belongs to a certain cluster, $P(z_{yes}|x)$ only for the cluster becomes high. However, there may be the case that $P(z_{yes}|x)$ for multiple clusters become high. In this case, we allocate the resources based on the maximum value of the predicted traffic whose corresponding $P(z_{yes}|x)$ is high so as to avoid the risk of congestion. That is, we allocate resources to accommodate the traffic whose volume is larger than

$$T^{\text{allocate}} = \max_{n \in \{n | P_n(z_{yes}|x) > \lambda\}} T_n^{\text{future}}$$

where $P_n(z_{yes}|x)$ is the confidence of the decision maker corresponding to the nth cluster, and λ is the threshold.

5 Evaluation

In this evaluation, we demonstrate the effect of using the real-world information, and the effect of the prediction inspired by the cognitive process of the human brain.

5.1 Evaluation method

5.1.1 Evaluation environment

For this evaluation, the data of the people movement and the traffic generated by the users are required. However, no actual data of the movement of people is available.

In this thesis, we synthetically generate the data of people movement and the traffic generated by the people by using the pseudo-generated GPS trajectory dataset called Open PFLOW [20] (University of Tokyo CSIS-JoRAS), and Synthetic Traffic Generator [21].

Open PFLOW includes a typical movement pattern of people in the metropolitan area for one day. In this data set, a pair of time and GPS corresponding to a user is recorded every 5 seconds. However, the number of people included in this data is 617,040 and does not include the data corresponding to all users in the metropolitan area.

Therefore, assuming that there are multiple users who move in a similar way, we generate the number of people in each area by assigning a scale factor to each user in the Open PFLOW and summing the scale factors of the users in the area. By randomly changing the scale factors, we can generate multiple data; we use one of them as the training data for the prediction, and the others for the evaluation.

Synthetic Traffic Generator reproduce the amounts of traffic, the number of requests and the interval time between requests on from $x:00:00$ to $(x+1):00:00$ ($0 \leq x \leq 23$) based on the real data. This simulator generates most of the requests on $x:00:00$, but the request of each user occurs in more various time zones in the actual network. Therefore, in this evaluation, we regenerate the request time so that the request uniformly distribute between $x:00:00$ and $(x+1):00:00$. The Synthetic Traffic Generator generates only requests and does not generate the information on the amount of traffic in shorter time granularity. In this evaluation, traffic of each user was generated assuming that the traffic rate is constant from the beginning to the end of the request.

In this evaluation, we define the area by partitioning around Chiyoda-ku, Tokyo, into areas

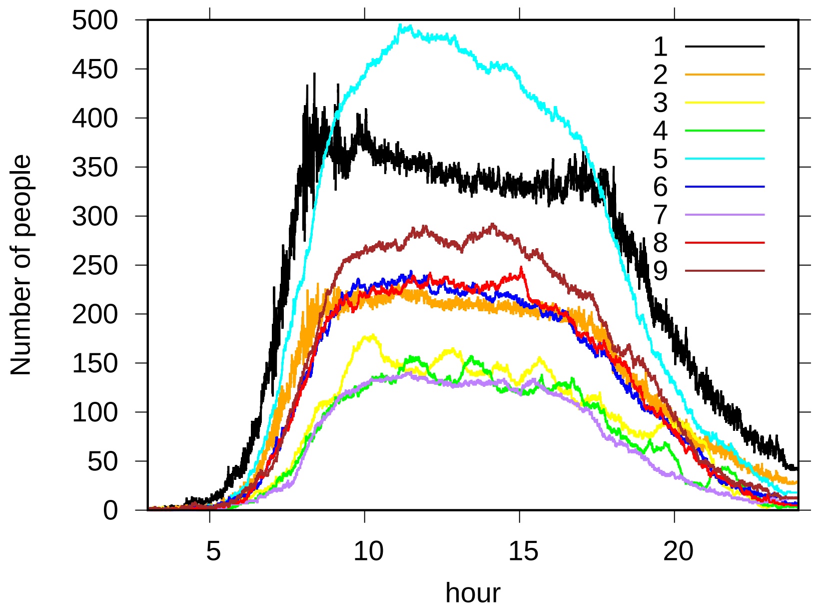

Figure 2: Time series of the number of users at prediction target area and its surroundings

of 0.0036 (about 350 m) in both latitude and longitude. We focus on one of the areas as the target area. Figure 2 shows the time series of the number of people in the target and nearby areas. Figure 3 shows the time series of the traffic in the areas. In these figures, sequence 1 corresponds to the target area.

In this evaluation, for simplicity, the unit of the allocated resources set so that one unit can accommodate 16 Mbps. In our evaluation, we set p to 40 minutes. That is, we predict 40-minute future traffic and allocate resources so as to avoid congestion for 40 minutes.

5.1.2 Compared method

In this evaluation, we evaluate the effect of using real-world information and the effect of prediction inspired by the cognitive process of the human brain. In order to evaluate the above effect, in this evaluation, we compare the following method with the proposed method (hereinafter referred to as cognitive TE method with real-world information).

Cognitive TE without real-world information This method predicts the future traffic by the same way as our cognitive TE with real-world information but uses only the information on traffic volumes to predict future traffic. The information used by this method at the time slot t is the

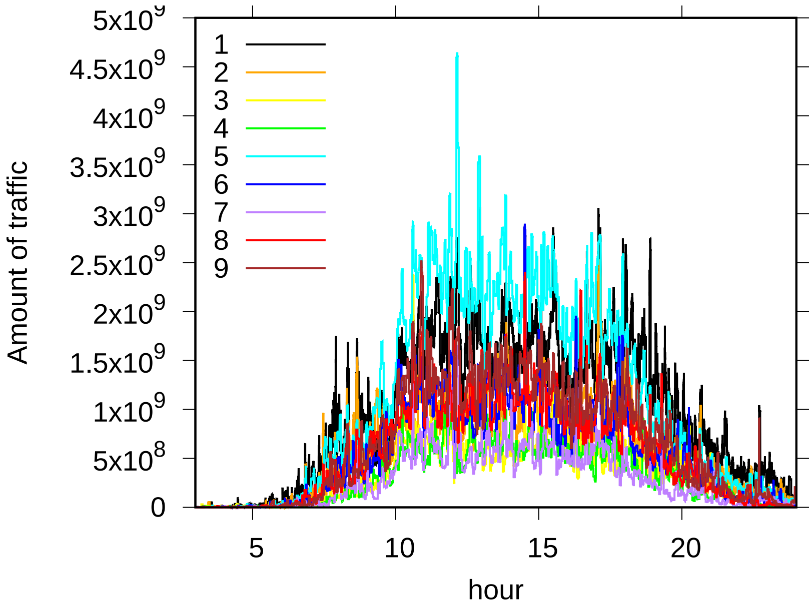

Figure 3: Time series of the amounts of traffic at prediction target area and its surroundings

following information of the areas whose distances from the target area are less than m

- Traffic volume at time slot t
- Difference between the amount of traffic at the time slot $t-p$ and the time slot t

This method is the same as the proposed method except for the information used for prediction. By comparing our method with this method, we demonstrate the effect of using the information of the number of people.

Deterministic TE With real-world information This method predicts the future traffic by using the same information as our method. However, this method does not use the process inspired by the cognitive model of the human brain. Instead, this method determines the cluster of the current status as the cluster whose centroid is the nearest to the current observation. By comparing with this method, we demonstrate the effect of prediction based on the cognitive process.

Deterministic TE Without real-world information This method predicts the future traffic in the same way as the deterministic TE with real-world information but uses only the information on the traffic volumes.

5.1.3 Parameter settings

In this evaluation, we set $m = 2$ and $p = 480$. That is, all methods use the observation information of the areas whose distances from the target area are less than two areas (about 700 m) and use the difference from the observed value 40 minutes ago as the rate of increase of the traffic amounts and the number of people.

We also set the sensory uncertainty to 0.42, and dynamic uncertainty to 0.3 in the BAM. We set the slope of the sigmoid functions a to 2.0. The other parameters in the BAM are set to the same values as the values set by Bitzer et al. [11]

5.1.4 Metrics

In this thesis, we allocate resources to the target area using each method so as to avoid congestion. We allocate the resources that can accommodate $T^{\mathrm{future}} + \alpha$ Mbps of traffic, where T^{future} is the predicted traffic and α is a margin. Setting a large margin avoids the congestion but requires more resources. It is preferable to avoid congestion with a limited amount of resources

Therefore, we investigate the number of time slots where the congestion occurs and the total amount of allocated resources.

5.2 Evaluation results

Figure 4 shows the results. The horizontal axis indicates the number of time slots when congestion occurs due to lack of resources and the vertical axis indicates the total amounts of resources allocated when we set α so as to make the number of time slots when congestion occurs less than the value on the horizontal axis. The total amounts of resources indicates the total amounts of traffic volume (KByte) that can be relayed in all time slots by the allocated resources. We also plot the sum of traffic that cannot be accommodated, the number of time slots when congestion occurs and the sum of extra resource in Figures 5, 6, and 7.

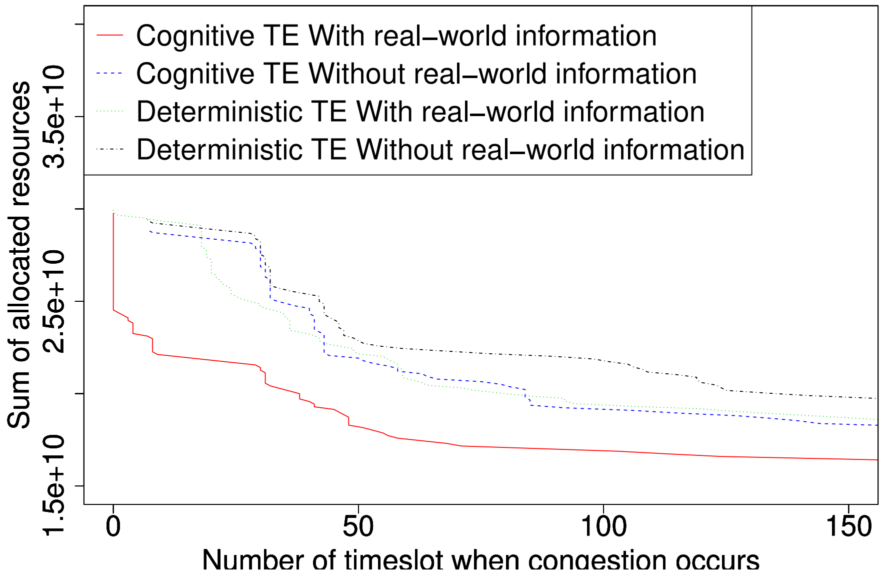

Figure 4: Sum of allocated resources necessary to keep number of timeslot when congestion occurs below a certain level

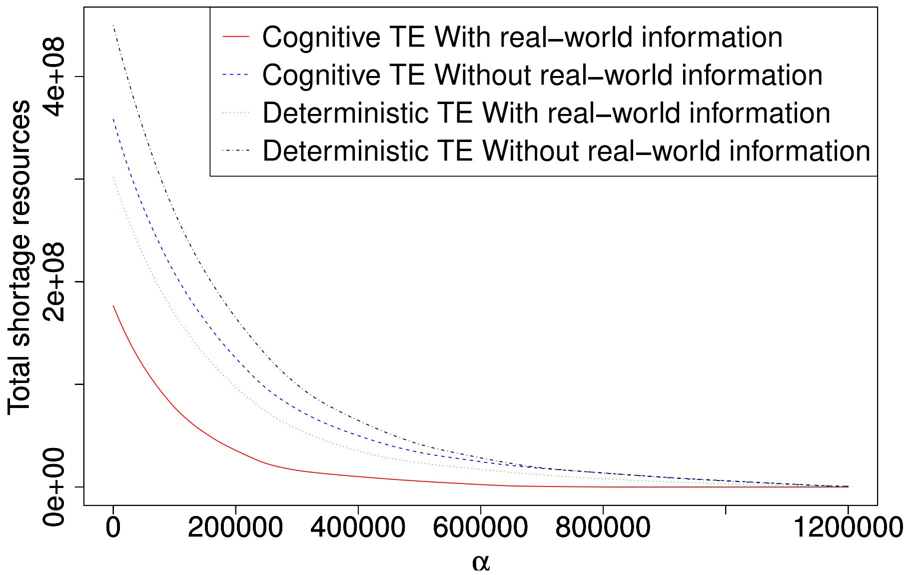

Figure 5: Sum of shortage resource

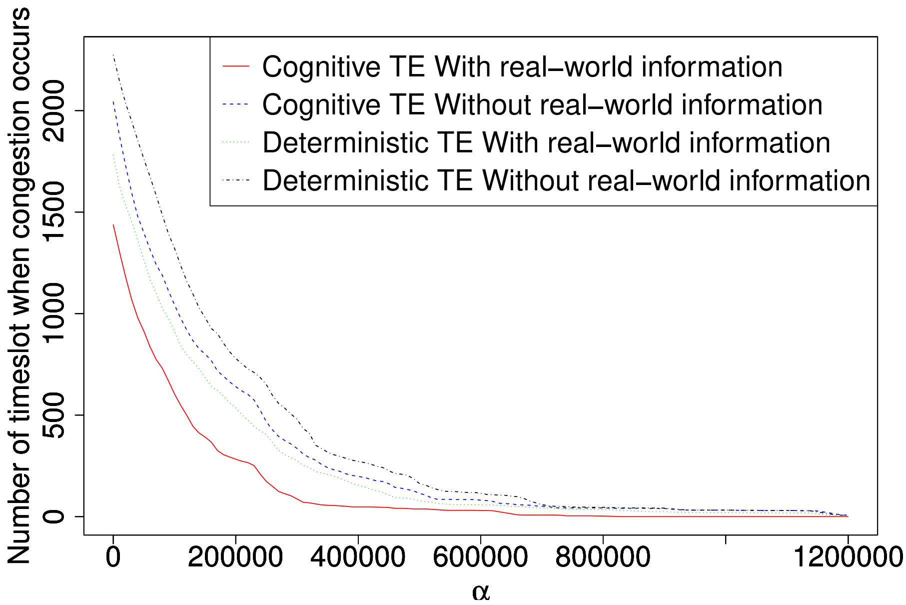

Figure 6: Number of timeslot when congestion occurs

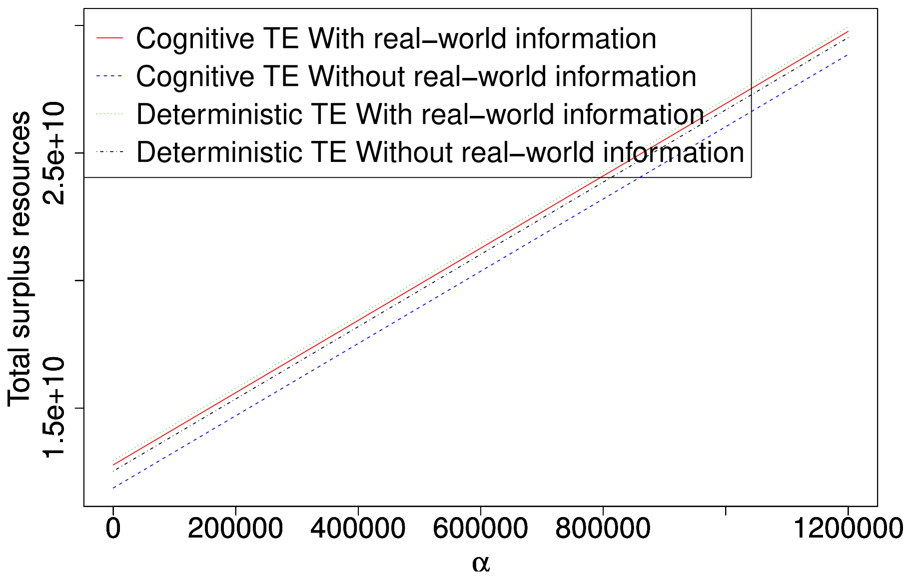

Figure 7: Sum of surplus resource

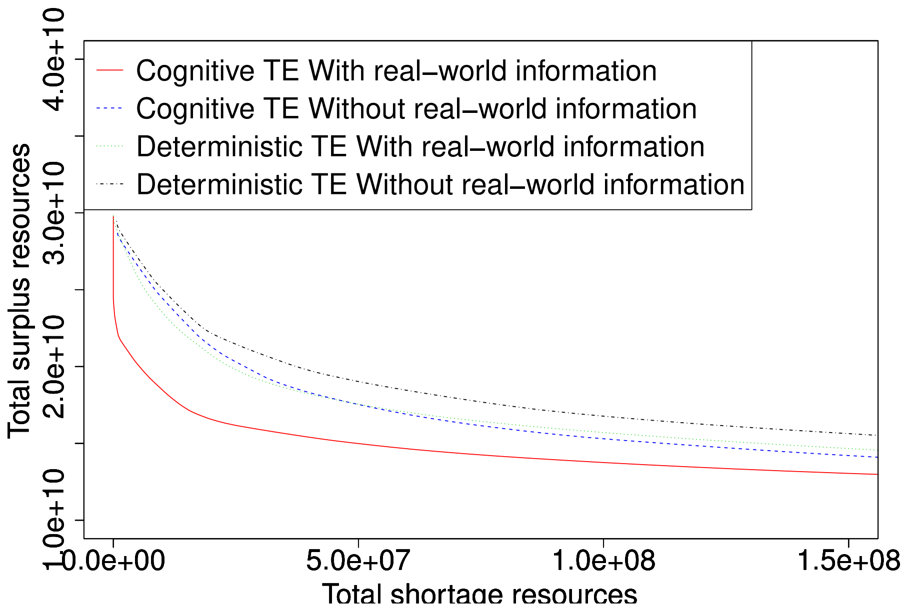

Figure 8: Sum of allocated resources necessary to keep sum of shortage resource below a certain level

Comparing the cognitive TE with real-world information and the cognitive TE without real-world information, the cognitive TE with real-world information requires a smaller amount of resources to keep the number of time slots when congestion occurred to less than a certain value. This is because the real-world information enables us to capture the difference in the states which could not be distinguished by using only traffic volume information, which leads accurate prediction of future traffic. As shown in Figures 2 and 3, the traffic volumes has a strong correlation with the number of users. That is, the number of users is useful information for the prediction of the traffic. In addition, the fluctuation of the number of users does not include large noises, compared with the fluctuation of the traffic. The cognitive TE without real-world information cannot accurately decide the status of the network only from the traffic information due to their large noise. On the other hand, the cognitive TE with real-world information accurately decides the states even in such cases.

We next compare the cognitive TE with real-world information and the deterministic TE with real-world information. Figure 4 shows that the cognitive TE requires a smaller amount of resources. This is because the cognitive TE with real-world information controls based on confidence. The cognitive TE request extra resources when confidences for multiple candidates become large. As a result, small α is enough to avoid congestion. On the other hand, the deterministic TE does not consider the confidences. As a result, large α is required to avoid congestion, which causes a large amount of extra resources.

6 Conclusion and future work

In this thesis, we proposed a predictive traffic engineering method which predicts future traffic using the information monitored in the real world. Though the real-world information may contribute the prediction of the future traffic, it is difficult to model the future traffic as the function of the real-world information. That is, we need a new method to predict future traffic using such information whose relation to the future traffic cannot be clearly modeled.

Therefore, we proposed a prediction method inspired by the human-brain cognition process which makes decisions from uncertain information. In this model, a human brain has stochastic variables and updates the variables by Bayesian inference every time a new observation is obtained. Then, a human brain makes a decision based on the stochastic variables.

In our method, we define multiple conditions by the monitored information including both of traffic and real-world information. Our method learns the future traffic corresponding to each condition. Then, our method predicts future traffic by estimating the condition of the current traffic and real-world information by the process inspired by the human-brain cognition process; our method has the stochastic variables indicating the confidence about that the current traffic and real-world information belong to the corresponding condition, and updates the variables every time a new observation is obtained. Finally, our method allocates the resources based on the future traffic corresponding to the conditions whose confidences are high.

We evaluated our method by simulation. The results demonstrate that our method avoids congestions without requiring a large amount of extra resources; the amount of resources required to avoid congestion is reduced by 25 % compared with the predictive TE using only the past traffic information.

Our future work includes the parameter settings of our method. Especially, the number of clusters k may have a large impact on the accuracy of the prediction. The evaluation of our method in a different environment is also included in our future work. For example, we will evaluate our method in the case of the different p.

Acknowledgements

This thesis would not accomplish without a lot of great supports of many people. First, I would like to express my deepest gratitude to Professor Masayuki Murata of Osaka University, for his valuable comments, insights and continuous encouragement. Furthermore, I would like to show my greatest appreciation to Assistant Professor Yuichi Ohsita of Osaka University. He devoted a great deal of time for me and gave me a lot of advices about my research. Without his support, all results in my research life so far could not be achieved. Moreover, I would like to show my appreciation to Associate Professor Shin'ichi Arakawa, Assistant Professor Daichi Kominami and Specially Appointed Assistant Professor Naomi Kuze and Tatsuya Otoshi of Osaka University. I would like to thank all the members of Advanced Network Architecture Research Laboratory at the Graduate School of Information Science and Technology, Osaka University, for their support and advices.

References

[1] P. Kumar, Y. Yuan, C. Yu, N. Foster, R. Kleinberg, and R. Soulé, "Kulfi: Robust traffic engineering using semi-oblivious routing," *arXiv preprint arXiv:1603.01203*, 2016.

[2] M. Chiesa, G. Kindler, and M. Schapira, "Traffic engineering with equal-cost-multipath: An algorithmic perspective," *IEEE/ACM Transactions on Networking*, 2016.

[3] H. H. Liu, S. Kandula, R. Mahajan, M. Zhang, and D. Gelernter, "Traffic engineering with forward fault correction," *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 4, pp. 527–538, 2015.

[4] I. F. Akyildiz, A. Lee, P. Wang, M. Luo, and W. Chou, "A roadmap for traffic engineering in SDN-OpenFlow networks," *Computer Networks*, vol. 71, pp. 1–30, 2014.

[5] A. Elwalid, C. Jin, S. Low, and I. Widjaja, "Mate: Mpls adaptive traffic engineering," in *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3, pp. 1300–1309, IEEE, 2001.

[6] T. Otoshi, Y. Ohsita, M. Murata, Y. Takahashi, N. Kamiyama, K. Ishibashi, K. Shiomoto, and T. Hashimoto, "Traffic engineering based on model predictive control," *IEICE Transactions on Communications*, vol. 98, no. 6, pp. 996–1007, 2015.

[7] L. Lu and C. Lianglun, "NETWORK TRAFFIC PREDICTION MODEL BASED ON OPTIMISING SVM WITH IMPROVED CUCKOO SEARCH ALGORITHM," *Computer Applications and Software*, vol. 1, p. 033, 2015.

[8] G. Rutka, "Neural network models for Internet traffic prediction," *Elektronika ir Elektrotechnika*, vol. 68, no. 4, pp. 55–58, 2015.

[9] W. Lu, "Parameters of Network Traffic Prediction Model Jointly Optimized by Genetic Algorithm," *JNW*, vol. 9, no. 3, pp. 695–702, 2014.

[10] G. Rutka and G. Lauks, "Study on internet traffic prediction models," *Elektronika ir Elektrotechnika*, vol. 78, no. 6, pp. 47–50, 2015.

[11] S. Bitzer, J. Bruineberg, and S. J. Kiebel, "A Bayesian Attractor Model for Perceptual Decision Making," *PLoS Comput Biol*, vol. 11, no. 8, p. e1004442, 2015.

[12] P. Ramanathan, K. M. Sivalingam, P. Agrawal, and S. Kishore, "Dynamic resource allocation schemes during handoff for mobile multimedia wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 7, pp. 1270–1283, 1999.

[13] M. Zulhasnine, C. Huang, and A. Srinivasan, "Efficient resource allocation for device-to-device communication underlaying lte network," in *Wireless and Mobile Computing, Networking and Communications (WiMob), 2010 IEEE 6th International Conference on*, pp. 368–375, IEEE, 2010.

[14] D. López-Pérez, X. Chu, A. V. Vasilakos, and H. Claussen, "On distributed and coordinated resource allocation for interference mitigation in self-organizing lte networks," *IEEE/ACM Transactions on Networking*, vol. 21, no. 4, pp. 1145–1158, 2013.

[15] Y. Yu, M. Song, Y. Fu, and J. Song, "Traffic prediction in 3g mobile networks based on multifractal exploration," *Tsinghua Science and Technology*, vol. 18, no. 4, pp. 398–405, 2013.

[16] H. Feng and Y. Shu, "Study on network traffic prediction techniques," in *Wireless Communications, Networking and Mobile Computing, 2005. Proceedings. 2005 International Conference on*, vol. 2, pp. 1041–1044, IEEE, 2005.

[17] K. P. Körding and D. M. Wolpert, "Bayesian decision theory in sensorimotor control," *Trends in cognitive sciences*, vol. 10, no. 7, pp. 319–326, 2006.

[18] W. J. Ma, J. M. Beck, P. E. Latham, and A. Pouget, "Bayesian inference with probabilistic population codes," *Nature neuroscience*, vol. 9, no. 11, p. 1432, 2006.

[19] S. J. Julier and J. K. Uhlmann, "New extension of the kalman filter to nonlinear systems," in *Signal processing, sensor fusion, and target recognition VI*, vol. 3068, pp. 182–194, International Society for Optics and Photonics, 1997.

[20] T. Kashiyama, Y. Pang, and Y. Sekimoto, "Open pflow: Creation and evaluation of an open dataset for typical people mass movement in urban areas," *Transportation Research Part C: Emerging Technologies*, vol. 85, pp. 249–267, 2017.

[21] E. M. R. Oliveira, A. C. Viana, K. P. Naveen, and C. Sarraute, "Measurement-driven mobile data traffic modeling in a large metropolitan area," in *Pervasive Computing and Communications (PerCom), 2015 IEEE International Conference on*, pp. 230–235, IEEE, 2015.