# Traffic Prediction for Dynamic Traffic Engineering Considering Traffic Variation

Tatsuya Otoshi†, Yuichi Ohsita†, Masayuki Murata†,
Yousuke Takahashi‡, Keisuke Ishibashi‡, and Kohei Shiomoto‡
† Graduate School of Information Science and Technology, Osaka University
1–5 Yamadaoka, Suita, Osaka, 565–0871, Japan
Email: {t-otoshi, y-ohsita, murata}@ist.osaka-u.ac.jp
‡ NTT Network Technology Laboratories, NTT Corporation
9–11, Midori-Cho 3–Chome Musashino–Shi, Tokyo 180–8585 Japan
Email: {takahashi.yousuke, ishibashi.keisuke, shiomoto.kohei}@lab.ntt.co.jp

***Abstract*—Traffic engineering with traffic prediction is one approach to accommodate time-varying traffic without frequent route changes. In this approach, the routes are calculated so as to avoid congestion based on the predicted traffic. The accuracy of the traffic prediction however has large impacts on this approach. Especially, if the predicted traffic amount is significantly less than the actual traffic, the congestion may occur. In this paper, we propose the traffic prediction methods suitable to the traffic engineering. In our method, we perform preprocessing before the prediction in order to predict the periodical variation accurately. Moreover, we consider the confidence interval for the prediction error and the variation excluded by the preprocessing to avoid the congestion caused by the temporal traffic variation. In this paper, we discuss three preprocessing approaches; the trend component, the lowpass filter, and the envelope. Through simulation, we clarify that the preprocessing by the trend component or the lowpass filter increases the accuracy of the prediction. In addition, considering the confidence interval achieves the lower link utilization within a fixed control period.**

***Index Terms*—Traffic Engineering, Traffic Prediction, Data Mining, Trend Component, SARIMA Model**

## I. Introduction

In recent years, the time variation of the Internet traffic becomes large due to the growth of the Internet services, such as streaming and cloud services. A backbone network has to accommodate such traffic without congestion.

Many traffic engineering schemes have addressed the problem of accommodating time-varying traffic [1–3]. In the traffic engineering methods, a control server periodically observes the traffic in a network and dynamically changes the routes so as to accommodate the observed traffic. Traffic engineering using only the observed traffic, however, cannot avoid the congestion when the traffic variation occurs. Although frequent observation and control can quickly handle such traffic variation, this may cause some problems such as network instability and high observation/control overheads. Wang *et al.* proposed a traffic engineering method that finds the routes suitable to not only the observed traffic but also any possible traffic [3]. However, this method requires more resource so as to accommodate a large number of traffic patterns.

Traffic prediction is useful for traffic engineering to accommodate time-varying traffic stably. By using the predicted future traffic variation, traffic engineering accommodates the traffic variation without frequent route changes unless unexpected traffic variation occurs. The requirements on the traffic prediction for the traffic engineering are as follows. (1) The traffic prediction should predict the traffic of the future several hours accurately. Then, by using the predicted traffic, the traffic engineering sets the routes suitable to several hours. The traffic variation for several hours is affected by the daily traffic variation. Thus, the traffic prediction should follow the daily traffic variation. (2) The traffic prediction should also consider the shorter-term time variation than the daily variation, to avoid the congestion caused by the short-term traffic variation. (3) The underprediction should be avoided, while the overprediction is acceptable. The underprediction causes the congestion due to the lack of resources assigned to the underpredicted traffic. On the other hand, the overprediction does not cause the performance degradation unless the overpredicted traffic requires too much resources and causes the lack of the resources assigned to the other traffic.

Network traffic prediction has been well studied [4–9]. There are mainly two targets for the traffic prediction, which are long-term and short-term traffic variations.

For the long-term prediction, extracting the target variation by aggregation and preprocessing is essential. Papagiannaki *et al.* [5] proposed the long-term prediction method that extracts the long-term variation by the wavelet multiresolution analysis, and predicts the variation using the ARIMA model aggregating the time series in one week. They showed the six month traffic can be predicted including upper and lower bound. However, the granularity of the prediction is too long to apply to the traffic engineering.

Predicting the short-term variation is hard due to the unpredictable traffic variation. For the short-term prediction, one-step prediction is often conducted, which predicts the only next one step every time. Balaji *et al.* proposed one-step prediction with confidence interval to avoid the underprediction for dynamic bandwidth provisioning [8]. One-step prediction is highly accurate, but traffic engineering requires to predict the traffic variation in next several hours to calculate the stable routes. The method handling the short-term variation within multiple steps (e.g. several hours) has not been sufficiently studied.

The traffic prediction for the traffic engineering should have the both aspects of the short-term and long-term prediction; the daily traffic pattern should be considered to achieve accurate prediction for several hours, while the short-term traffic variation also should be considered to avoid the future congestion.

In this paper, we address the traffic prediction for the traffic engineering. In our approach, we extract the daily traffic pattern before the prediction. By extracting the daily pattern, we accurately predict the traffic variation for several hours. Then, we consider the confidence interval of the prediction and the excluded traffic variation to avoid the congestion caused by the short-term traffic variation.

The rest of this paper is organized as follows. Section II introduces the traffic engineering method using the predicted traffic. Section III describes the prediction methods. Section IV presents evaluation of each prediction method. The conclusion and future work are mentioned in Section V.

## II. Traffic Engineering with Traffic Prediction

In this paper, we deploy a central control server which controls the network. The central control server observes and predicts the traffic rate, and calculates the routes based on the predicted traffic.

The control server observes the traffic rate at each flow in fixed intervals (e.g. ten minutes, thirty minutes or one hour) called *time slot*. The observed traffic rates of all flows in the $t$-th time slot are represented as a vector and we denote this vector as $\boldsymbol{x}_t$. The prediction of future traffic is denoted as

$$\hat{\boldsymbol{x}}_{t+1..t+f} = F\left(\boldsymbol{x}_{t-h+1..t}\right), \tag{1}$$

where $\boldsymbol{x}_{a..b} = (\boldsymbol{x}_a, \boldsymbol{x}_{a+1}, \cdots, \boldsymbol{x}_b)$ is a matrix in which each column is corresponding to each vector, $\hat{\boldsymbol{x}}_k$ is the predicted traffic in the $k$-th time slot, $f$ is the number of time slots where the traffic rate is predicted, $h$ is the length of observed time slots used in the prediction and $F$ is a prediction function defined by a prediction method.

In traffic engineering, the control server calculates the routes so as to avoid congestion for $f$ time slots. We define this $f$ time slots as the *control period*. In this paper, we consider the case that the control period is several hours. The calculated routes are represented as a matrix $A$ called *routing matrix*. The $(i, j)$-element $a_{i,j}$ in the routing matrix $A$ represents the ratio of the traffic over the flow $j$ mapped onto the link $i$. Corresponding to the routing matrix, the predicted traffic mapped onto each link in the control period is represented as

$$\hat{\boldsymbol{y}}_{t+1..t+f} = A\hat{\boldsymbol{x}}_{t+1..t+f}, \tag{2}$$

where $\hat{\boldsymbol{y}}_k$ is the vector indicating the predicted traffic on all links in the $k$-th time slot. Traffic engineering is the process to adjust $A$ so as to control $\hat{\boldsymbol{y}}_{t+1..t+f}$ in some desirable way.

In this paper, we use the simple optimization approach that minimizes the maximum utilization among all links for all time slots within the control period, though there may be more sophisticated approach using the predicted traffic. The optimization problem is formulated as the following linear programming problem;

$$minimize{:}U \tag{3}$$

$$subject\ to{:}\forall s,d,\ \sum_{p(l)=s} A^{s,d}(l) = 1 \tag{4}$$

$$\forall s,d,\ \sum_{f(l)=d} A^{s,d}(l) = 1 \tag{5}$$

$$\forall s,d,n,\ \sum_{p(l)=n} A^{s,d}(l) = \sum_{f(l)=n} A^{s,d}(l) \tag{6}$$

$$\forall l,k, \sum_{s,d} A^{s,d}(l)\hat{x}_k^{s,d}/C(l) < U, \tag{7}$$

where $U$ is the maximum link utilization, $A^{s,d}(l)$ is the ratio of traffic from $s$ to $d$ routed over the link $l$, and $p(l)$ and $f(l)$ are the start and end nodes of the link $l$, respectively, $\hat{x}_k^{s,d}$ is the predicted traffic volume of the flow from $s$ to $d$ at the $k$-th time slot and $C(l)$ is the capacity of the link $l$. $\hat{x}_k^{s,d}$ and $C(l)$ are given in this problem, and $A^{s,d}(l)$ and $U$ are the variable to be obtained. Eqs. (4–6) are the constraints for flow conservation. Eq. (7) ensures that $U$ is the maximum link utilization of all the links for all the time slots within the given control period.

## III. Traffic Prediction

### A. Overview

In the network, traffic variation has a daily pattern, and the traffic changes in several hours. The traffic prediction should follow this daily variation so that the traffic engineering calculates the routes suitable to the next several hours. However, the actual traffic variation includes noisy variation and the daily tendency is polluted. Using such polluted data, predicting the daily traffic variation becomes inaccurate. Therefore, we use preprocessing which extracts the daily periodical variation excluding the noisy variation to improve the prediction accuracy.

On the other hand, the short-term traffic variation may cause the congestion. To avoid this congestion, the short-term traffic variation should be considered. In our approach, the short-term traffic variation is considered by the confidence interval of the traffic variation excluded by the preprocessing. Moreover, we also consider the confidence interval of the prediction error to avoid the impact of the prediction error on the traffic engineering. The confidence interval causes the overprediction. However, as described in Section I, the impact of the overprediction is smaller than that of the underprediction.

Our approach is summarized in Figure 1. First, we extract the daily variation from the actual traffic variation by the preprocessing. Second, we predict the future traffic variation using the extracted variation and estimate the range of excluded variation. Finally, we obtain the upper bound of traffic variation summing up the predicted upper bound of

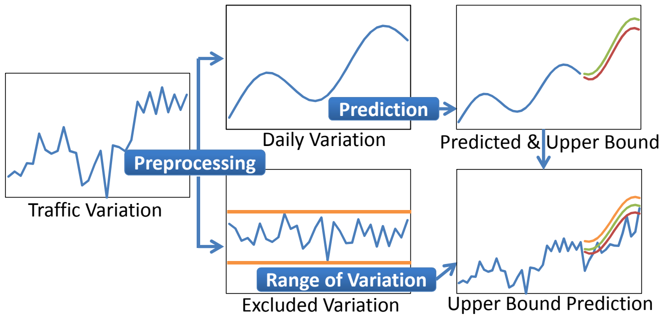



Fig. 1. prediction process

daily variation and estimated range of excluded variation. The obtained upper bound is used as input of the traffic engineering.

*B. Prediction Preprocessing*

In the preprocessing, we extract the daily periodical variation from the observed traffic. The object of preprocessing is to filter out the short-term traffic variation which is hard to be predicted. This increases the accuracy of the prediction of the daily traffic variation.

In this paper, we investigate the following preprocessing methods, the trend component, the lowpass filter and the envelope. The rest of this subsection describes the detail of the preprocessing methods.

*1) Trend Component:* One approach to extract the long-term variation of the traffic variation is to use the trend model [10]. We call the traffic variation extracted by using the trend model *trend component*. The trend component includes the daily traffic variation and longer-term traffic variation. The trend model is denoted as

$$x_k = t_k + \epsilon_k \tag{8}$$
$$\Delta t_k = \Delta t_{k-1} + w_k, \tag{9}$$

where $x_k$ is the traffic volume of a flow in the $k$-th time slot, $t_k$ is the trend component, $\Delta t_k = t_k - t_{k-1}$, $\epsilon_k \overset{\text{i.i.d.}}{\sim} N(0, \theta^2)$ is the noise of observation and $w_k \overset{\text{i.i.d.}}{\sim} N(0, \lambda^2)$ is the noise in the trend component.

Eq. (8) indicates that the original data is composed of the trend component and the noise. Eq. (9) indicates that the trend component is perturbed by Gaussian noise.

At the first step to calculate the trend component, the variances $\theta^2$ and $\lambda^2$ are found by the Maximum Likelihood Estimation (MLE). Then, the trend component $t_i (i = t - h + 1, \cdots, t)$ is determined by the conditional expectation $E\left[t_i | x_{t-h+1..t}\right]$ with the probability of transition in Eqs. (8) and (9).

*2) Lowpass Filter:* As a similar idea to the trend component, the lowpass filter by using the Fourier transform extracts the long term variation.

Using the Fourier transform, the time series of the traffic data can be represented as

$$x_k = \sum_{n=0}^{h-1} f_n \exp\left(2\pi i \frac{nk}{h}\right), \tag{10}$$

where $f_n$ is Fourier coefficient corresponding frequency $n/h$ and $i$ is the imaginary unit. Eq. (10) includes also high frequency variations such as noise. To reduce this noisy variations, the lowpass filter removes the terms with large $n$ and extracts the long-term variation as

$$l_k = \sum_{n=0}^{L} f_n \exp\left(2\pi i \frac{nk}{h}\right), \tag{11}$$

where $L$ is the threshold to remove the high frequency variations. We set $L$ so as to remove the variation of the higher frequency than the daily variation because the traffic variation has daily pattern.

*3) Envelope:* Extracting the variation of traffic upper bounds may be useful to predict the bandwidth required to accommodate the short-term traffic variation. In this paper, we extract the upper bound variation by tracing the peak value in the fixed time interval. We divide the observed values $x_{t-h+1}, \cdots, x_t$ into $l = \frac{h}{\tau}$ intervals, where $\tau$ denotes the length of the intervals. The set of the time slots in the $k$-th interval is denoted as

$$I_k = \{(k-1)\tau + t - h + 1, \cdots, k\tau + t - h\}. \tag{12}$$

We set the interval length $\tau$ to 12 hours considering the daily variation.

The peak value in $I_k$ is represented by $x_{p_k}$, where $p_k$ represents the peak time slot denoted as

$$p_k = \arg\max_{i \in I_k} x_i. \tag{13}$$

In this paper, we extract the envelope by connecting the peak values $x_{p_1}, \cdots, x_{p_l}$ and the latest value $x_{p_{l+1}} = x_t$ with lines. Including the latest value $x_t$, the prediction can reflect the latest data. We perform the linear interpretation for points between $x_{p_{k-1}}$ and $x_{p_k}$, and each point is interpreted as

$$x_i = x_{p_k} + \frac{x_{p_{k+1}} - x_{p_k}}{p_{k+1} - p_k}(i - p_k) \tag{14}$$
$$i = p_k, p_k + 1, \cdots, p_{k+1}, \quad k = 1, \cdots, l.$$

*C. Prediction*

The traffic prediction is performed based on the prediction model after each preprocessing. The model-based prediction learns the model parameters from inputted data, and then predicts the future values based on the obtained model. Focusing on the effect of considering the periodicity in the prediction, we take the SARIMA model and the ARIMA model as examples of periodic and non-periodic prediction models, respectively. The rest of this section gives an overview of prediction with the ARIMA and the SARIMA model.

*1) ARMA model:* Before describing the ARIMA and the SARIMA model, we give a short explanation of the ARMA model which is the base model for the ARIMA and the SARIMA model.

The ARMA model represents data at each time slot using the previous data and errors as

$$x_n = \sum_{i=1}^{p} a_i x_{n-i} + \sum_{i=0}^{q} b_i \epsilon_{n-i} + c \quad (15)$$
$$b_0 = 1,$$

where $p$ and $q$ denote the numbers of past data and error which data at each time slot depends on, respectively. $a_i$ and $b_i$ are the coefficients, $\epsilon_i$ is the error at the $i$-th time slot and $c$ is a constant.

*2) ARIMA model:* The ARIMA model is an extension of the ARMA model so as to model the non-stationary data, such as the data whose mean value fluctuates over time. In order to apply the ARMA model to such data, removing the non-stationarity is performed. When the variation of the mean has linear characteristic, the differenced data $\Delta x_n = x_n - x_{n-1}$ excludes the variation of the mean. In this manner, $d$ times differencing operation $\Delta^d$ can remove the mean variation following a polynomial of degree $d$. In the ARIMA model, ARMA model in Eq. (15) is applied to the differenced data $\Delta^d x_n$.

*3) SARIMA model:* The SARIMA model is a generalization of the ARIMA model. Considering the periodicity, the SARIMA model applies a periodical differencing to the data as $\Delta_s x_n = x_n - x_{n-s}$ where $s$ is a period length. After applying the $D$ times of the periodical differencing $\Delta_s^D x_n$, the differencing method in the ARIMA model is also applied. Therefore differenced data is finally denoted as $\Delta^d \Delta_s^D x_n$. Considering the daily variation and weekday/weekend difference, we set $s$ to the weekly length.

The differenced data is fitted to the following model which expands the ARMA model including the data and error at previous periods as

$$\begin{aligned} x_n =& \sum_{i=1}^{p} a_i x_{n-i} + \sum_{i=0}^{q} b_i \epsilon_{n-i} + c \\ &+ \sum_{j=1}^{P} A_j \sum_{i=1}^{p} a_i x_{n-sj-i} + \sum_{j=1}^{Q} B_j \sum_{i=0}^{q} b_i \epsilon_{n-sj-i} \quad (16) \\ b_0 =& 1, \end{aligned}$$

where $P$ and $Q$ denote the numbers of previous periods for depended data and error, respectively. $A_i$ and $B_i$ are the coefficients.

*4) Model Fitting:* A SARIMA model is fitted to the data by the following steps.

First, differencing the data is repeated until the data become stationary. Stationarity test is performed by examining whether the data follows non-stationary process $x_t = x_{t-1} + \epsilon$ called unit route process. We use the KPSS test [11] and Canova-Hansen test [12] for determining $d$ and $D$, respectively. The KPSS test examines the null hypothesis $\epsilon = 0$ which means the data is stationary. The Canova-Hansen test applies the null hypothesis test to the Fourier coefficients variation of the each period.

After the number of differencing the data is determined, the feasible $p, q, P$ and $Q$ are searched so that the MLE method which decides the coefficients $A_i, B_i, a_i, b_i$ obtains the good model. The goodness of a model is defined by the Akaike Information Criterion (AIC) [13] which is defined by

$$\text{AIC} = -2 \log L + 2k, \quad (17)$$

where $L$ is the maximized likelihood with the MLE, $k$ is the number of parameters. In the SARIMA model, $k = p + q + P + Q$. A model with a large number of parameters can fit the data well, but may fit the incidental variation such as noise. Penalizing $k$, AIC can select the best model avoiding overfitting the data. Using the method by Hyndman *et al.* [14], the search process is performed by changing $p, q, P$ and $Q$ by one until no new model can improve AIC.

An ARIMA model fitting can be also performed by fixing the parameter $D = P = Q = 0$ in the SARIMA model fitting.

*5) Prediction with Fitted Model:* After fitting a SARIMA model, the future traffic is predicted according to the obtained model. The predicted traffic in the next $k$-th time slot is calculated as

$$\bar{x}_{t+k} = E[x_{t+k} | x_{t-h+1..t}]. \quad (18)$$

*6) Confidence Interval:* The SARIMA model can calculate the confidence interval for the prediction error. The upper confidence bound for the prediction can be calculated by $\bar{x}_{t+k} + \alpha \hat{\sigma}_{t+k}$, where $\bar{x}_{t+k}$ is the predicted traffic volume at the next $k$-th time slot, $\alpha$ is a parameter indicating the considered confidence level and $\hat{\sigma}_{t+k} = \sqrt{V[x_{t+k} | x_{t-h+1..t}]}$ is the estimated standard deviation of prediction error.

### *D. Range of Excluded Variation*

The traffic variation excluded by the preprocessing should also be considered, because it may cause the congestion. In this paper, we consider the excluded traffic variation by using the standard deviation of the excluded traffic variation. The standard deviation is calculated as

$$\sigma = \sqrt{V\left[x_{t-h+1..t} - x'_{t-h+1..t}\right]}, \quad (19)$$

where $x_{t-h+1..t}$ is the original time series of traffic on a flow and $x'_{t-h+1..t}$ is the extracted variation. Using $\sigma$, we compensate for the excluded variation in the predicted traffic as $\bar{x}_t + \beta\sigma$ where $\beta$ is a parameter indicating the confidence level for the upper bound prediction of the excluded variations.

Finally, the upper bound prediction including both the prediction error and the excluded variation in the preprocessing can be calculated as

$$\hat{x}_i = \bar{x}_i + \alpha \hat{\sigma}_i + \beta \sigma. \quad (20)$$

## IV. Evaluation

### *A. Evaluation Methodology*

*1) Used Data:* We use the actual traffic traces in the backbone network of Internet2 [15] which is the research and education network in the United States. This traffic data is collected by Netflow protocol at each of the nine Point of

Presence (PoP) routers. The sampling rate is one packet in every one hundred packets and aggregated data is exported every five minutes. The large daily variation between day and night is mainly observed in the traffic variation. Focusing on such traffic variation within several hours, we aggregate the observed data within each one hour, that is, we set the length of the observation time slot to one hour.

We use four week data from 11/28/2011 to 12/25/2011 aggregated into the flows between PoP routers using the BGP information.

*2) Evaluation Process:* In our evaluation, we use the data of the previous two weeks as the observed data. By using this data, we perform the preprocessing and prediction processes. Then, we calculate the optimal routes for a given control period by using the predicted traffic. After the calculated routes are set, we investigate the link utilization calculated by using the actual traffic.

We perform the prediction 24 times, changing the start time of the prediction, because the traffic variation at the start time of the prediction has large impact on the accuracy of the prediction. We compute the optimum routes by solving the linear programming problem in Eqs. (3–7) using CPLEX [16].

## B. Evaluation Result

*1) Traffic Prediction Result:* To investigate the characteristic of the prediction methods, we show the mean prediction error of each traffic prediction method in Figure 2. The prediction error is calculated as the relative error $|\bar{x}_t - x_t|/\mu$ where $\mu$ is the average value of the traffic.

In Fig. 2, "non-preprocess" means prediction using original data without preprocessing. "trend", "envelope" and "lowpass" mean prediction with each corresponding preprocessing. "arima" and "sarima" mean prediction by ARIMA and SARIMA model, respectively.

Fig. 2 indicates that prediction with the trend component or the lowpass filter achieves the lower prediction error than prediction without preprocessing. This result indicates that preprocessing is effective for improving the prediction accuracy. However, since the envelope extracts upper bound of the traffic variation, the method with the envelope usually overpredicts and the mean prediction error is high.

Fig. 2 also indicates that the prediction method using the SARIMA model is more accurate than the ARIMA model when preprocessing is performed, though the both SARIMA and ARIMA prediction without preprocessing have almost same accuracy. The SARIMA model can predict the daily variation more accurately than the ARIMA model because of considering the periodicity. Using the preprocessing, the daily tendency of the traffic variation becomes clear and leaning the periodicity becomes easy. According to this result, the SARIMA model with preprocessing will be better choice for traffic engineering which needs to handle the daily traffic variation.

Focusing on the difference between the trend component and the lowpass filter, we investigated the both prediction results in detail and found an interesting difference. Figure 3

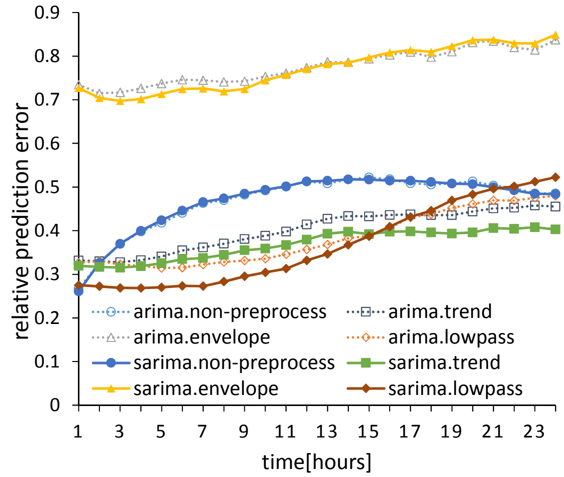



Fig. 2. relative prediction errors

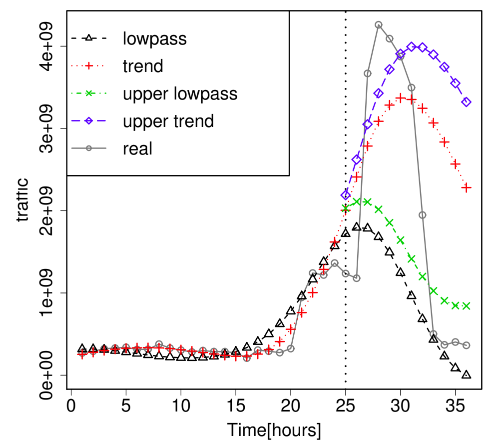



Fig. 3. example of the SARIMA prediction using the trend component and the lowpass filter

shows the SARIMA prediction result of a flow. In Fig. 3, the real traffic variation and the predicted variation are plotted. The vertical dotted line indicates the start point of the prediction. "upper lowpass" and "upper trend" mean the upper bound of 80% confidence interval for the lowpass filter and the trend component, respectively. In Fig. 3, the method using the lowpass filter significantly underpredicts the rapidly increasing traffic. Because the lowpass filter removes the all shorter term variation than daily variation, the lowpass filter cannot extract the rapidly increasing tendency and the prediction using the lowpass filter sometimes underpredicts the increasing traffic. The trend component more sensitively extracts the increasing tendency than the lowpass filter. Therefore prediction using the trend component more sensitively follows the variation though the mean prediction error becomes higher.

*2) Traffic Engineering Result:* First, to investigate the effect of considering the confidence interval to avoid the underprediction, in Figure 4 we show the complement cumulative distribution function (CCDF) of the maximum link utilization with different confidence levels especially for the trend component and the lowpass filter method with the SARIMA model. In Fig. 4, we present only the evaluation results setting the both confidence level $\alpha$ and $\beta$ to the same value in order to show the effect of the confidence level briefly. "mean" is the result using mean prediction without confidence interval and "k %" means that confidence level is corresponding to $k\%$. To clarify the response to the sudden traffic increase which usually causes higher link utilization, we focus on the higher maximum link utilization.

In the both preprocessing methods, we can observe that considering the confidence interval can reduce the higher link utilization except one point of the result in Fig. 4(a), where the unexpected traffic increase occurs. Considering the confidence interval, the traffic engineering allocates the large resource to a flow which has the large short-term variation. Thus, the impact of the short-term variation is reduced and the link utilization becomes low.

From Fig. 4, the effect of using the confidence interval becomes more clear for the long control period. Generally, the longer control period becomes, the more unexpected variation occurs. Therefore considering the confidence interval is more effective to calculate a stable route for the long-term control period.

Fig. 4 also indicates that considering the confidence interval is more effective for the lowpass filter than the trend component. As Fig. 3 shows, the lowpass filter sometimes underpredicts the traffic volume significantly while the trend component follows the traffic variation well. Thus, considering the confidence interval is more useful to complement the predicted value for the lowpass filter.

In all results shown in Fig. 4, using too high confidence level causes the large link utilization. This is caused by the allocation of too much resources to a particular flow. Allocating too much resources causes the lack of resources assigned to the other flows. The decision of the suitable confidence level may be performed by using the feedback from the observation after the traffic engineering, which is one of our future work.

We also compare the results of the traffic engineering using the traffic predicted by the prediction method using different preprocessing approaches. Figure 5 shows the CCDF of the maximum link utilization. Figures 5(a)–5(c) show the results with different control periods. "ideal" means results of the routes calculated by using the actual traffic variation. "previous 1hour" and "previous 2week" mean calculating routes using the previous one hour and two weeks data instead of predicted traffic, respectively. The others are the result using the SARIMA model for prediction with each preprocessing method.

We configure the confidence level of each prediction method so that the maximum link utilization at the peak time slot is minimized. This means that we evaluate the performance of each prediction method overall time slots within the control period while guaranteeing the performance at the worst case.

In all cases of the control period, the traffic engineering with the prediction is useful to achieve the lower link utilization. Especially, the results of "trend" and "lowpass" are better. This is because the preprocessing methods can improve the prediction accuracy by extracting the periodical variation.

When the control period is 12 hours, "lowpass" does not keep the link utilization low. This is because the prediction using the lowpass filter sometimes underpredicts the suddenly increasing traffic. The impact of this underprediction appears more clearly in the longer control period, because the longer control period has the higher probability of the occurrence of unexpected traffic variation.

"trend" achieves low link utilization even though the control period becomes long. Therefore, for the traffic engineering that aims to set the stable routes, the prediction with the trend component is suitable.

## V. Conclusion

In this paper, we proposed the traffic prediction method for traffic engineering using preprocessing and confidence interval. In our method, we extract the long-term variation before the prediction, so as to improve the prediction accuracy of the daily traffic variation. The short-term traffic variation is also handled by considering the confidence interval of the prediction and the traffic variation excluded by the preprocessing. Through evaluation, we clarified that the preprocessing improves the accuracy of the daily traffic variation. In addition, the results show that considering the confidence interval avoids the large link utilization caused by the traffic variation excluded by the preprocessing. The results also indicates the SARIMA model with the preprocessing of the trend component achieves the low link utilization for the long control period.

Our future work includes the setting optimum confidence levels for the upper bound prediction and the investigation of the suitable traffic engineering method to use the predicted traffic.

## Acknowledgement

This work was supported in part by the Strategic Information and Communications R&D Promotion Programme of the Ministry of Internal Affairs and Communications, Japan.

## References

[1] N. Wang, K. H. Ho, G. Pavlou, and M. Howarth, "An overview of routing optimization for Internet traffic engineering," *IEEE Communications Survey & Tutorials*, vol. 10, no. 1, pp. 36–56, first quarter 2008.

[2] D. O. Awduche and U. (mci Worldcom), "MPLS and traffic engineering in IP networks," *IEEE Communications Magazine*, vol. 37, pp. 42–47, Dec. 1999.

[3] H. Wang, H. Xie, L. Qiu, Y. R. Yang, Y. Zhang, and A. Greenberg, "COPE: traffic engineering in dynamic networks," in *Proceedings of SIGCOMM*, vol. 36, no. 4. New York, NY, USA: ACM, Aug. 2006, pp. 99–110.

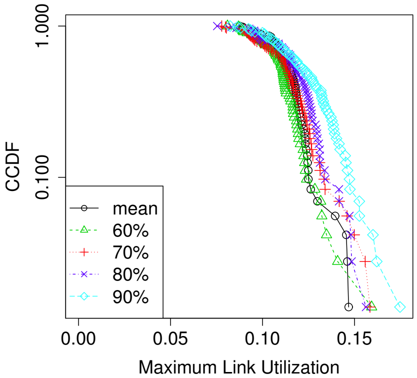
[4] M. F. Zhani, H. Elbiaze, and F. Kamoun, "Analysis and prediction of real network traffic," *Journal of Networks*, vol. 4, no. 9, pp. 855–865, Nov. 2009.

[5] K. Papagiannaki, N. Taft, Z.-L. Zhang, and C. Diot, "Long-term forecasting of Internet backbone traffic: Observations and initial models," in *Proceedings of INFOCOM*, vol. 2, Mar. 2003, pp. 1178–1188.

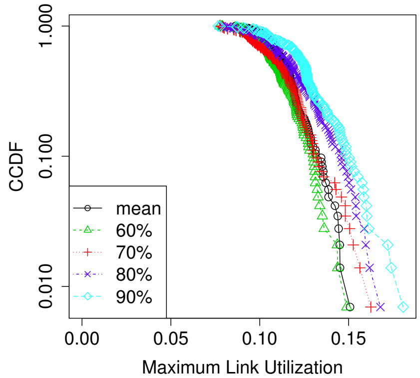
[6] B. Zhou, D. He, Z. Sun, and W. H. Ng, "Network traffic modeling and prediction with ARIMA/GARCH," in *Proceedings of HET-NETs*, Sep. 2006, pp. 1–10.

[7] L. Xiang, "A new hybrid network traffic prediction method," in *Proceedings of GLOBECOM*, Dec. 2010, pp. 1–5.

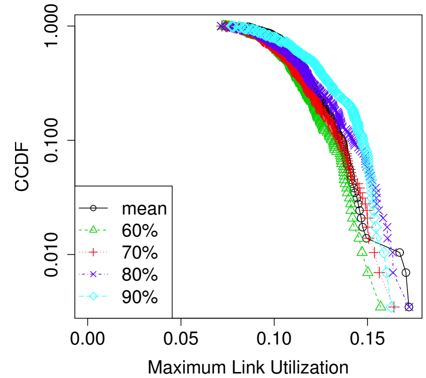
[8] B. Krithikaivasan, T. Zenf, K. Deka, and D. Medhi, "ARCH-based traffic forecasting and dynamic bandwidth provisioning for periodically measured nonstationary traffic," *IEEE/ACM Trans. On Networking*, vol. 15, no. 3, pp. 683–696, Jun. 2007.



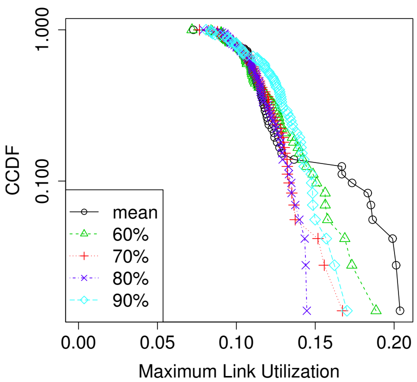

(a) Trend Component(control period:3hours)



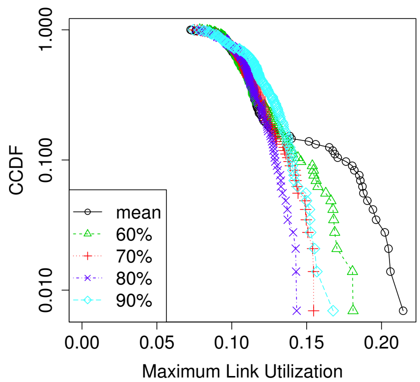

(b) Trend Component(control period:6hours)



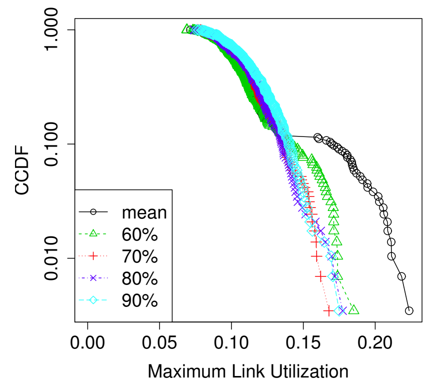

(c) Trend Component(control period:12hours)




(d) Lowpass Filter(control period:3hours)

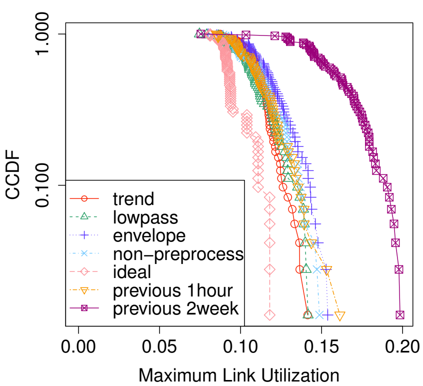



(e) Lowpass Filter(control period:6hours)

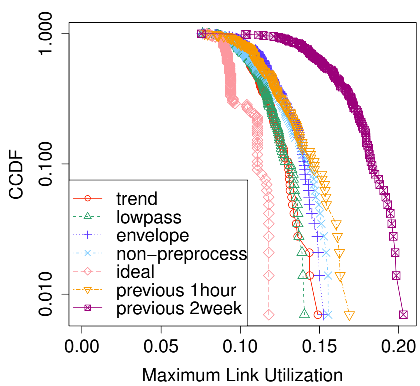



(f) Lowpass Filter(control period:12hours)

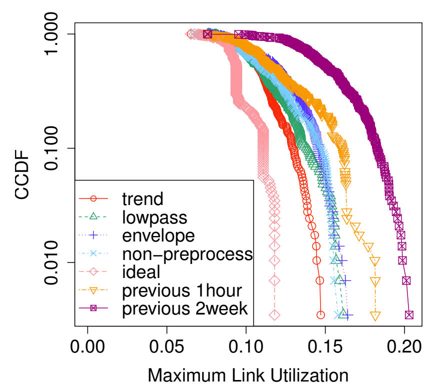
Fig. 4. complement cumulative distribution of maximum link utilization with different confidence levels




(a) control period : 3hour




(b) control period : 6hour




(c) control period : 12hour

Fig. 5. complement cumulative distribution of maximum link utilization with different prediction method

[9] M. L. F. Miguel, M. C. Penna, J. C. Nievola, and M. E. Pellenz, "New models for long-term Internet traffic forecasting using artificial neural networks and flow based information." in *Proceedings of NOMS*. IEEE, Apr. 2012, pp. 1082–1088.

[10] 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.

[11] D. Kwiatkowski, P. C. Philips, and P. Schmidt, "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?" *Journal of Econometrics*, vol. 54, no. 1–3, pp. 159–178, Oct–Dec 1992.

[12] F. Canova and B. E. Hansen, "Are seasonal patterns constant over time? A test for seasonal stability," *Journal of Business & Economic Statistics*, vol. 13, no. 3, Jul. 1995.

[13] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Proceedings of the Second International Symposium on Information Theory*, 1973, pp. 267–281.

[14] R. J. Hyndman and Y. Khandakar, "Automatic time series forecasting: The forecast package for R," *Journal of Statistical Software*, vol. 27, no. 3, pp. 1–22, Jul. 2008.

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

[16] "IBM ILOG CPLEX Optimizer," optimization software : http://www-01.ibm.com/software/integration/optimization/cplex-optimizer.