On the Tradeoff between Computational Efficiency and Prediction Accuracy in Bandwidth Traffic Estimation

I. Loumiotis, E. Adamopoulou, K. Demestichas, T. Stamatiadi and M. Theologou

The increasing demand for wireless broadband services poses the need for the efficient utilization of the backhaul network resources. To this end, schemes that use artificial neural networks in order to predict the forthcoming network traffic demand and proactively request the commitment of the necessary resources have been proposed. However, an up-to-date prediction model, required by these schemes, necessitates for a regularly held training process, which incurs a high computational cost. In this letter, the authors investigate the tradeoff between prediction accuracy and computational efficiency by employing evolutionary game theory and propose a novel scheme that can achieve both aspects.

Introduction: During the last years, the plethora of available wireless broadband services has led to an increasing demand for high end-user data rates [1]. Unfortunately, this increased traffic demand, which is also characterized by great fluctuations, constitutes the traditional flat commitment of the backhaul network resources inefficient. As a result, novel dynamic schemes are required in order to achieve a more efficient utilization of the backhaul network and provide enhanced QoS to the end users. To this direction, schemes that estimate the forthcoming demand by using historical data of the network traffic and proactively request the commitment of the necessary resources have been proposed. The majority of these schemes presents artificial neural network (ANN) based approaches [2], exploiting their ability to capture the non linear characteristics of network traffic. Specifically, in [3], the authors proposed a prediction model for the aggregated demand of the base station (BS), based on ANNs. It was shown that the traffic demand can be accurately predicted, and as a result, the BS can proactively request for the commitment of the necessary resources from the backhaul network. In [4], the authors studied the Internet traffic forecasting problem and they showed that ANNs outperform an autoregressive moving average model, providing accurate short-term results. In [5], the authors studied the prediction of variable bit rate traffic in asynchronous transfer mode networks using ANNs, while in [6], ANNs were used in order to forecast video streaming traffic.

Though all the above ANN-based schemes propose network traffic prediction models that can provide accurate results, they do not discuss the significance of the training process in achieving functional efficiency. It is evident that, in order to obtain accurate prediction results, an up-to-date model is required which necessitates for a regularly held training process at the expense of an increased computational overhead. In the current letter, the tradeoff between computational efficiency and prediction accuracy with respect to the frequency of the training process is studied and a novel scheme that can achieve both of its aspects is proposed. The authors treat the aforementioned tradeoff as a decision making problem and employ evolutionary game theory [7] in order to study its dynamics. Finally, a bifurcation analysis of the proposed scheme is also presented.

Problem Formulation and Relevant Assumptions: Assume a prediction model, which is responsible for monitoring the aggregated bandwidth demand of an access point (AP), storing the necessary data and using ANNs to predict the forthcoming bandwidth demand. Specifically, a set of W input-output pairs, derived from these observations, is used to train the ANN. The training process is held periodically, and it is expected that an up-to-date prediction model will provide more accurate results regarding the forecasting process. However, the training process incurs a corresponding computational overhead. Hence, there is a tradeoff between accurate prediction results and computational efficiency in the proposed scheme.

Consider that the prediction model uses a sliding window in order to select the W most recent data for the training process of the ANN. It is assumed that there exists a certain time period of N time units during which the model performs the training process $i \leq N$ times. Thus, the period of the training process is $T = {}^{N}/_{i}$, $i \in [1, \ldots, N]$. For convenience, the frequency domain is considered. Let $\mathbf{F} = [{}^{1}/_{N}, {}^{2}/_{N}, \ldots, 1]$ denote the set of possible frequencies for the training process. Hence, the prediction model has to choose a frequency $f \in \mathbf{F}$ which performs the training process in a computationally efficient and accurate way.

Let S_f denote the payoff experienced by the prediction model with regards to the accuracy of the results when the training phase is held with frequency $f \in \mathbf{F}$. It is considered that payoff S_f is monotonically increasing with respect to f. Unfortunately, the training process also incurs a computational overhead C_f which is also monotonically increasing with respect to f. It is expected that as the frequency f increases, the training process is held more often, providing a more up-to-date prediction model and, thus, more accurate results, at the expense of a higher computational overhead.

Given the payoff S_f concerning the accuracy of the prediction model with respect to the training frequency and the corresponding computational cost C_f, the net utility of the model is defined by

$$\mathcal{U}_f = \frac{S_f^{\beta}}{C_f^{1-\beta}} \tag{1}$$

where $\beta \in [0, 1]$ represents the tradeoff coefficient. For a choice of $\beta = 0$, the prediction model cares only about the computational cost, while for a choice of $\beta = 1$, the prediction model cares only about the accuracy of the results. In the following, the utility $\mathcal{U}_f$ is used as a figure of merit in order to find the optimal frequency f^* for the training process that offers both prediction accuracy and computational efficiency.

Dynamics of Training Process: The evolutionary game theoretic concept is now applied to the tradeoff problem formulated above. Consider that the prediction model consists of a population of agents, each programmed to use a certain frequency $f \in \mathbf{F}$ for the training process with a corresponding payoff (i.e. net utility) U_f. If x_f denotes the population of agents using frequency f, then it must hold that $\sum_{f \in \mathbf{F}} x_f = 1$. It is noted that the relative frequency expressed by x_f, represents the probability of the prediction model to use frequency f for the training process. In the game under consideration, it is assumed that there are certain periods in which the prediction model reviews its strategies (i.e. frequencies), and only those strategies that yield a payoff (i.e. net utility) higher than the average are favored. If this period is considered small, then the evolution of the strategies (i.e. frequencies) of the prediction model can be described by the replicator dynamics equation given by

$$\dot{x}_f = x_f(\mathcal{U}_f - \bar{\mathcal{U}}) \tag{2}$$

where $\bar{\mathcal{U}} = \sum_{f \in \mathbf{F}} x_f \mathcal{U}_f$ is the average payoff (i.e. net utility).

Based on (2), the following theorem concerning the convergence to the optimal frequency for the training process holds.

Theorem 1: The prediction model performs the training process in the optimal frequency f^* with respect to computational efficiency and accuracy of the results.

Proof: By solving (2), it can be derived that there exist N rest points in the boundary and a rest point in the interior of the surface $\sum_{f \in \mathbf{F}} x_f = 1$. Using linearization techniques for calculating the Jacobian matrix at the rest points of the system and simple block matrices algebra, it can be derived that only the rest points on the boundary experience non-zero eigenvalues described, for the k-th equilibrium point, $k \in \mathbf{F}$, by

$$\lambda_j^{(k)} = \begin{cases} \mathcal{U}_j - \mathcal{U}_k & \text{if } j \neq k \\ \mathcal{U}_1 - \mathcal{U}_k & \text{if } j = k \end{cases} \tag{3}$$

for $j = {}^{1}/_{N}, \ldots, {}^{(N-1)}/_{N}$.

Hence, there exist a training frequency $f^* \in \mathbf{F}$, $f^* = \arg\max_{f \in \mathbf{F}} \mathcal{U}_f$, for which all the eigenvalues have negative real part and, thus, corresponds to an asymptotically stable equilibrium point. ■

Theorem 1 simply states that the prediction model will perform the training process with a proper frequency that provides both computational efficient and accurate results. Hence, all suboptimal strategies (i.e. frequencies) will be eliminated resulting in the optimization of the functionality of the proposed scheme.

Bifurcation Analysis: In the previous section, it was proven that the prediction model will perform the training process with an optimal frequency. In this section, the authors perform a bifurcation analysis of the tradeoff of the proposed system in order to investigate the dependence of the optimal frequency f^* on the coefficient β of (1). Considering β as a control parameter, the following theorem is derived.

Theorem 2: There exists a threshold value $\beta_{thr}^{(f^*)}$ for each frequency in $\mathbf{F}$*, above which frequency $f^* \in \mathbf{F}$ becomes the optimal frequency for the training process. It holds that,*

$$\beta_{thr}^{(f^*)} = \begin{cases} \dfrac{ln\left(\frac{C_{f^*}}{C_{f^*-1/N}}\right)}{ln\left(\frac{C_{f^*}}{C_{f^*-1/N}}\right)+ln\left(\frac{S_{f^*}}{S_{f^*-1/N}}\right)} & if\ f^* \neq 1/N \\ 0 & if\ f^* = 1/N \end{cases} \tag{4}$$

Proof: For an optimal training frequency f^*, it holds that $\mathcal{U}_{f^*} > \mathcal{U}_j$ where $j \neq f^* \in \mathbf{F}$. Using (1) and the monotonic property of S_f and C_f, (4) can be easily derived. ∎

Theorem 2 is of great importance as it specifies the threshold values of β that render a frequency optimal for the training process. The threshold values expressed by (4) constitute the bifurcation values of the system [8].

Simulation Results: In order to validate the proposed scheme, a Matlab [9] based simulator has been developed. Assume a weekly time period which corresponds to $N = 7$ days, and consider that the normalized payoff S_f for the prediction accuracy of the model can be described by

$$S_f = \frac{\nu f}{\nu - f + 1} \tag{5}$$

where $\nu < -1$ depends on the profile of the AP and the corresponding data used for the training process. The choice of ν reflects the influence of a more frequent training process on the accuracy of the prediction model. Specifically, for $\nu \to -1$, the payoff S_f experiences a logarithmic form, and as a result, the increase in the accuracy of the prediction model becomes less significant for higher training frequencies. On the other hand, for small values of ν, a more frequent training process has a linear impact on the payoff S_f. Furthermore, as the training set has a fixed size W, it is expected that the normalized computational cost can be described by

$$C_f = f \tag{6}$$

In Fig. 1, the evolution of the training frequencies is depicted for $\beta = 0.6$ and $\beta = 0.7$, with a choice of $\nu = -1.2$. It can be seen how the tradeoff coefficient β influences the optimal frequency for the training process. Specifically, it is observed that when the prediction model cares about the accuracy of the results and the corresponding computational cost in an approximately similar way ($\beta = 0.6$), it chooses the minimum frequency for the training process. However, as β increases ($\beta = 0.7$), the prediction model cares less about the computational overhead and chooses a more often training process, and the optimal training frequency increases. Finally, when the prediction model cares least about the computational overhead, it uses the maximum training frequency in order to always be up-to-dated.

The bifurcation values of the system with respect to the optimal frequencies f^*, are depicted in Fig. 2 for different values of ν. Specifically, Fig. 2 provides the threshold values of β, above which frequency f^* becomes asymptotically stable for the system model. It can be easily seen that as ν decreases, the threshold values of $\beta_{thr}^{(f^*)}$ also decrease. In other words, for a constant β, it holds that when the prediction accuracy payoff S_f approximates a linear form, i.e. decreases according to (5) (smaller values of ν), a more frequent training process is required.

Conclusion: In this letter, the tradeoff between computational efficiency and prediction accuracy for the training process of an ANN-based bandwidth traffic prediction model was investigated. The authors proposed a novel scheme that reflects this tradeoff and employed evolutionary game theory in order to study the stability of the model. The convergence of the proposed scheme to the optimal frequency for the training process in terms of prediction accuracy and computational efficiency was proven. Furthermore, a bifurcation analysis was held and the bifurcation values of the system were derived, which correspond to the threshold values above which a frequency becomes optimal for the training process. Therefore, the proposed scheme provides the guidelines for selecting a proper frequency for the training phase of network traffic prediction models, and can be used to support any supervised learning approach.

I. Loumiotis, E. Adamopoulou, K. Demestichas, T. Stamatiadi, M. Theologou (*Institute of Communication and Computer Systems, National Technical Univerity of Athens, Greece, GR15780*)

E-mail: i_loumiotis@cn.ntua.gr

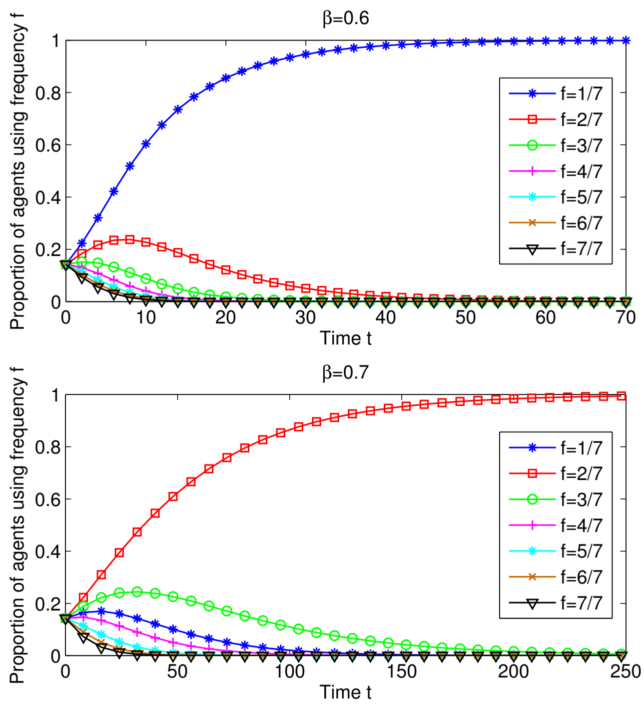

Fig. 1. Evolution of frequencies for different values of β and $\nu = -1.2$.

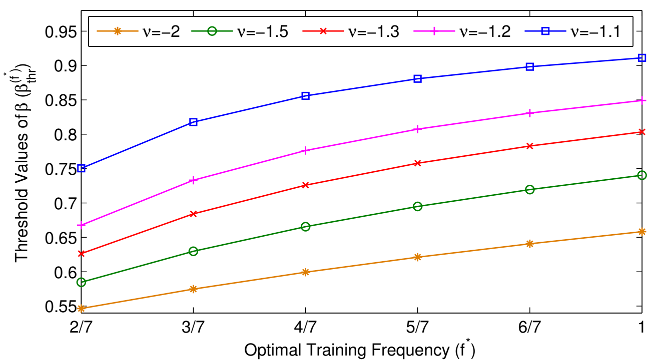

Fig. 2. Bifurcation values of the system for different values of ν.

References

1 Cisco: 'Cisco Visual Networking Index: Forecast and Methodology, 2011-2016', white paper, 2012.
2 Mitchell T.M.: 'Machine Learning', McGraw-Hill, 1997.
3 Loumiotis I., Stamatiadi T., Adamopoulou E., Demestichas K. and Sykas E.: 'Dynamic Backhaul Resource Allocation in Wireless Networks using Artificial Neural Networks', *IET Electronic Letters*, **49**, (8), 2013.
4 Cortez P., Rio M., Rocha M. and Sousa P.: 'Internet Traffic Forecasting using Neural Networks', *International Joint Conference on Neural Networks IJCNN '06* , pp.2635-2642, 2006.
5 Moh W.M., Chen M.J., Chu N.M. and Liao C.D.: 'Traffic prediction and dynamic bandwidth allocation over ATM: a neural network approach', *Computer Communications*, **18**, (8), pp.563-571, 1995.
6 Ullah F., Khan G.M., Mahmud S.A.: 'Intelligent Bandwidth Management Using Fast Learning Neural Networks', *IEEE 14th International Conference on High Performance Computing and Communication*, pp.867-872, 2012.
7 Hofbauer J. and Sigmund K., 'Evolutionary Games and Population Dynamics', Cambridge University Press, 1998.
8 Hale J. and Koçuk H.: 'Dynamics and Bifurcations', Springer-Verlang, 1991.
9 MATLAB R2012b. Natick, Massachusetts: The MathWorks Inc., 2012.