VBR Video Traffic Prediction Using Neural Networks with Multiresolution Learning

YAO LIANG
Alcatel USA
Raleigh, NC 27609, USA

EDWARD W. PAGE
Department of Computer Science
Clemson University
Clemson, SC 29634, USA

Abstract: - In this paper, we present our work on multi-step prediction of actual video traffic over ATM networks using neural networks with multiresolution learning paradigm. We demonstrate that with multiresolution learning, neural networks can successfully predict actual VBR video traffic hundreds of steps (frames) into the future. Our results show that neural networks with multiresolution learning are able to effectively achieve robust multi-step prediction of network traffic (as well as other difficult signal prediction applications) in real-world situations.

Key-Words: - Multi-step Prediction, Network Traffic, Neural Networks, Multiresolution Learning, Wavelet Theory

1 Introduction

Technological advances in high-speed networks have opened the possibility of providing switching and transport for diverse services such as voice, video, and high-speed data communication in a single integrated network. Network traffic management approaches must support different types of traffic classes at their required Quality of Service (QoS) and, at the same time, efficiently utilize the network resources. Network traffic prediction will play a very important role in future network traffic management systems. Studies have shown that traditional reactive congestion control is not suitable for broadband integrated networks due to the effects of high-speed channels [1]. An alternative approach is to develop adaptive traffic control techniques based on predicting the future behavior of the network traffic. Hence, network traffic prediction will be a key requirement in designing adaptive congestion control schemes that can learn “on-line” the complex nonlinear relationship among traffic characteristics, network resources, and the guaranteed QoS [2].

In broadband integrated networks, such as ATM, a major part of the traffic will be Variable-Bit-Rate (VBR) video traffic produced by multimedia sources including video-on-demand and teleconferencing. However, in recently published works on network video traffic prediction using neural networks, studies were focused on one-step traffic prediction [3, 4, 5, 6]. In order to evaluate the benefits and limitations of applying traffic prediction to actual network traffic control and bandwidth allocation, it is critical to investigate multi-step traffic prediction for real-world network traffic. In this paper, we describe our studies on multi-step prediction of actual VBR video traffic, using neural networks with multiresolution learning paradigm. Our work demonstrates that neural networks with multiresolution learning proposed in [7, 8] can significantly improve the traffic prediction performance. Furthermore, our work also demonstrates that multiresolution learning can improve the robustness of constructed neural networks dramatically [9]. Therefore, Our work illustrates the important progress in building neural networks of good generalization and robustness for real-world high-speed network video traffic prediction, and in
general, for other difficult time-series forecasting problems in real-world situations.

The rest of the paper is organized as follows. Section 2 describes the general scheme of neural networks for traffic prediction. Section 3 reviews multiresolution learning for neural networks, based on [7, 8]. In section 4, experimental results, on both the prediction performance of video traffic and the robustness of constructed neural networks, are provided. Finally, conclusion remark is given in Section 5.

2 Neural Networks for Signal Prediction

Since Lapedes and Farber first proposed using feedforward neural networks for nonlinear signal prediction in 1987 [10], neural network models for time series prediction have been developed rapidly, providing insights which are not available with the traditional models. Research on approximation capabilities of multi-layer feedforward neural networks has justified and motivated their use for nonlinear time-series forecasting [11, 12], as a feedforward neural network with a single hidden layer is capable of approximating uniformly any continuous multivariate function to any desired degree of accuracy.

The basic structure of a three-layer feedforward neural network for signal prediction, in general, is shown in Figure 1.

In Figure 1, $\Delta$ denotes a time (unit) delay. The input layer employs $(x(t), x(t-1), \ldots, x(t-n+1))$ as inputs. The hidden-layer neurons employ the typical sigmoid activation function, while the output neuron employs a linear activation function. As we know, a neural network trained for single-step forecasting can forecast multiple steps into the future by using the predicted output for a given step as an input for computing the time series at the next step, and all other network inputs are shifted back one time unit. This is sometimes referred to as iterated multi-step prediction. With iterated multi-step prediction, the neural network can predict as many steps ahead as needed. Iterated multi-step prediction is obviously more difficult than one-step prediction, and requires that the network trained exhibit the best possible generalization.

Generalization is a key requirement and yet very difficult problem in neural network research and applications. Another key issue is the robustness of neural networks constructed. Multiresolution learning is a new and effective learning paradigm to address these critical issues. Unlike traditional learning paradigm, the multiresolution learning paradigm exploits the correlation structures in the training data at multiple resolutions, which otherwise could be obscured at the original resolution of the training data.

3 Multiresolution Learning Review

Multiresolution learning is based on multiresolution analysis [13, 14] in wavelet theory. The multiresolution analysis framework is employed for decomposing the original signal and approximating it at different levels of detail. Unlike traditional neural network learning which employs a single signal representation for the entire training process, multiresolution learning exploits the approximation sequence resolution-by-resolution, from the coarsest version to finest version during the neural network training process. In this way, the original signal, the finest resolution version in the approximation sequence, will finally be used in the learning process.

Assume that a given sampled signal $s^m$ is to be learned. Let $m, M \in \mathbb{Z}$ and $0 < M < m$. Let a learning activity $A_j(r_j)$ denote a specific training phase conducted on the representation $r_j$ of training data with any given learning algorithm. Let $\rightarrow$ be the learning dependency operator by which $A_j \rightarrow A_i$ means that the learning activity $A_j$ should be conducted before the learning activity $A_i$. Multiresolution learning then can be defined as a sequence of learning activities $\{A_j(r_j)\}_{j \in \mathbb{Z}, j \geq M}$ associated with the sequence of approximation subspaces $\{V_j\}$ in
multiresolution analysis such that the following requirements are satisfied.

(1) The representation \( r_j \) is associated with the approximation \( s_j \) of the original signal \( s^m \) in the approximation subspace \( V_j \);

(2) \( A_j(r_j) \rightarrow A_{j+1}(r_{j+1}) \).

From this definition, it can be seen that the multiresolution learning paradigm generates an ordered sequence of learning activities as

\[
A_M(r_M) \rightarrow A_{M+1}(r_{M+1}) \rightarrow \ldots \rightarrow A_m(r_m)
\]

where the parameter \( M \) indicates the approximation level of the original signal \( s^m \) used to initiate the learning process. The first learning activity \( A_M(r_M) \) starts with randomly initialized network weights, and each subsequent learning activity \( A_j(r_j) \) starts with the connection weights resulting from the previous learning activity. The approximation \( s_j(j<m) \) will contain fewer signal samples than the original signal \( s^m \). However, the training vectors in each learning activity \( A_j \) should occupy the full dimension of the neural network inputs in \( V_m \) to guarantee the smooth transition between subsequent learning activities. Therefore, a method is needed for constructing the representation \( r_j \) of training data for \( A_j \) based on \( s^j \). This is achieved by setting the signal details \( d_k(k>j) \) to zero and reconstructing \( s^j \) in \( V_m \).

**4 Experimental Results**

MPEG (ISO Moving Picture Expert Group) coding standard is widely used for video applications [15] over broadband networks. MPEG employs a data compression algorithm, and three types of frames (I-frames, P-frames and B-frames) are considered in MPEG. The three types of frames are arranged in a deterministic periodic sequence called Group of Pictures (GOP). An actual VBR video traffic trace used in our traffic prediction studies is a cable TV news video trace which was encoded at University of Wurzburg in Germany [16] using MPEG-I. The MPEG encoding scheme has 12 frames in GOP, and the detail traffic characteristics and encoding parameter set of VBR video trace can be found in [16].

**4.1 Experiments**

In our experiments on multi-step traffic prediction, the original video traffic training data \( s^m \) is decomposed as

\[
s^m = s^{m-2} + d^{m-2} + d^{m-1} .
\]

The Haar wavelet basis is used for the decomposition. From this decomposition, two approximation versions at coarser resolutions of training data \( s^{m-2} \) and \( s^{m-1} = s^{m-2} + d^{m-2} \) are obtained. The corresponding multiresolution learning process for the video traffic then will contain \( A_{m-2}(r_{m-2}) \), \( A_{m-1}(r_{m-1}) \), and \( A_m(r_m) \), where

\[
r_j = \begin{cases} 
  s^j : & j = m \\
  s^j + \sum_{k=j}^{m-1} d^k : & d^k = 0 \quad j = m-2, m-1
\end{cases}
\]

Fig.2  Iterated Multi-step Video Traffic Predictions (Multiresolution Learning Approach)
The multiresolution learning process is
\[
A_{m-2}(r_{m-2}) \rightarrow A_{m-1}(r_{m-1}) \rightarrow A_m(r_m).
\] (5)

Backpropagation algorithm was employed. The identical 24-5-1 feedforward neural network structure is used for both multiresolution learning and traditional learning. Here the commonly used notation 24-5-1 denotes a three-layer network having 24 input nodes, 5 neurons in the hidden layer, and one output neuron. In both cases, the training process was started with identical initial random connection weights in the backpropagation procedure, and the same learning rate \(\eta = 0.01\). No momentum term was used.

In the experiments, a collection of 512 consecutive frames from the video traffic were used for training, and the subsequent 512 frames (not included in training) were used for evaluating multi-step prediction performance.

### 4.2 Results

The iterated multi-step predictions for the first 200 frames (time steps) of the test set of the video traffic are shown in Figures 2 and 3 for multiresolution and traditional learning, respectively.

The normalized mean squared error (NMSE) is used to assess forecasting performance [17]. The NMSE is computed as

![Figure 3. Iterated Multi-step Video Traffic Predictions (Traditional Learning Approach)](image)

![Figure 4. Multi-step Prediction Error for Video Traffic](image)
where

\[
NMSE = \frac{1}{\sigma^2} \frac{1}{N} \sum [x(t) - x^*(t)]^2
\]

(6)

\(x(t)\) is the observed value of the time series at time \(t\);

\(x^*(t)\) is the predicted value of \(x(t)\); and

\(\sigma^2\) is the variance of the time series over the prediction duration.

The result of NMSE in forecasting is shown in Figure 4. As we can see in Figures 2, 3 and 4, the forecasting (generalization) performance of the network employing multiresolution learning is significantly better than that of the network using traditional learning. Multiresolution learning yielded NMSE's much less than 1 even for the entire 512-frame testing period.

To further study the robustness of neural networks constructed with multiresolution learning vs. traditional learning, two additional network structures 24-15-1 and 36-3-1 were also investigated. Figures 5 and 6 show the results of NMSE for iterated multi-step predictions of the video traffic using three different neural network structures with

![Fig.5 Iterated Multi-step Prediction Error Using Multiresolution Learning](image1)

![Fig.6 Iterated Multi-step Prediction Error Using Traditional Learning](image2)
multiresolution and traditional learning, respectively. While the forecasting performance is highly dependent on the adopted network structure when using traditional learning, multiresolution learning is able to significantly reduce such structure sensitivity and achieve much more robust prediction performance. The results clearly indicate that multiresolution learning can also greatly improve the robustness of neural networks constructed.

5 Conclusion Remark

In this paper, we present our work on multi-step prediction of actual video traffic over ATM networks using neural networks. Multi-step prediction is essential to the practicality of adaptive network control based on traffic prediction. Our results show that with multiresolution learning, neural networks can predict actual video traffic hundreds of frames (steps) into the future with reasonable accuracy. Furthermore, forecasting performance is much less sensitive to the adopted neural network structure for the task, when multiresolution learning rather than traditional learning is employed. We believe that multiresolution learning paradigm not only provides new insight on neural network learning theory, but also opens new dimension and opportunities for applying neural networks to many difficult tasks in real-world situations, such as network traffic prediction, where good generalization and robustness of neural networks are highly required.

References