Intelligent Bandwidth Estimation for Variable Bit Rate Traffic

Article in IEEE Transactions on Evolutionary Computation · October 2013
Impact Factor: 3.65 · DOI: 10.1109/TEVC.2013.2285122

4 authors, including:

Gul Muhammad Khan
University of Engineering and Technology, P…
74 PUBLICATIONS 226 CITATIONS

Rabia Arshad
University of Engineering and Technology, P…
8 PUBLICATIONS 11 CITATIONS

S.A. Mahmud
University of Engineering and Technology, P…
64 PUBLICATIONS 146 CITATIONS
Intelligent Bandwidth Estimation for Variable Bit Rate Traffic

Gul Muhammad Khan, Rabia Arshad, Sahibzada Ali Mahmud, Fahad Ullah

Abstract—A novel Neuro-Evolutionary Cartesian Genetic Programming based Frame Size Estimator for multimedia streaming applications has been proposed in this work. The frame size obtained from the proposed estimator is used to calculate and allocate the bandwidth required for frame transmission. Bandwidth calculation and allocation is done via various probabilistic and linear regression methods. To obtain conclusive results regarding the feasibility of the proposed system, different test case scenarios have been exploited. The bandwidth allocation efficiency for the technique has been compared with previously proposed methods to evaluate its effectiveness in precise bandwidth utilization. Compared to other contemporary techniques, our technique gives approximately 98% efficient frame size prediction and bandwidth allocation.

Index Terms—Scheduling, Evolutionary Algorithm, Traffic Estimation, MPEG-4, Bandwidth Allocation.

I. INTRODUCTION

VISITING video streaming sites nowadays is a very popular trend among internet users because of ease of access to high speed broadband internet. These sites get the highest traffic after the exceptionally popular social networking websites. This excessive video streaming leads to bandwidth availability issues that need to be addressed efficiently and intelligently. It is a challenge to efficiently manage the bandwidth utilization for different users while adhering to the required standard quality of service (QoS). Real time data traffic streaming gives rise to many difficult issues. Especially in the case of Variable Bit Rate traffic, these issues end up being more pronounced and sometimes nontrivial [1]. For devices working within wireless operational frameworks it becomes quite tedious to deal with such nontrivial problems.

The issue of bandwidth scheduling and allocation has been tackled in many ways in literature. Filtering Techniques are quite popular for this purpose. One such technique includes the Kalman filtering based technique (involves self induced congestion) known as BART (Bandwidth Available in Real-Time) [2] for bandwidth estimation in a frame switched network. MR-BART (Multi-Rate BART) [3] is an extension to BART and shows a higher convergence rate than BART. A genetic algorithm-based neural fuzzy decision tree (GANFDT) [4] is one algorithm that uses a combination of algorithms to get the required bandwidth estimation and allocation. Another novel intelligent bandwidth estimation technique has also been proposed by Yuan et.al [5], which utilizes multimedia packets rather than probing traffic for available bandwidth estimation. Probing based estimations [6] [7] are also popular for bandwidth estimation and calculation. Probing techniques use probe packets to estimate the available bandwidth and then use this estimation to allocate the required bandwidth to the multimedia streaming users.

For bandwidth scheduling of MPEG-4 traffic, a traffic predictor is employed to maintain a fair amount of QoS. Variable step size normalized least mean square algorithm (VSSNLMS) is one of the techniques proposed for this purpose [1]. It is quite efficient in terms of implementation but in certain cases introduces a large prediction error especially at a scene boundary. Another predictor implementation divides the frames into I-Frame, P-Frame and B-Frame types and then utilizes the normalized least mean square algorithm (NLMS). Yang and Tien [8] devised a method that intelligently allocates bandwidth by using a neural fuzzy online traffic Predictor. It uses two phases of learning i.e. the structure learning and the parameter learning. A long term online traffic predictor has also been proposed by Lee and Chang [9], which uses the \( \rho \)-domain rate model for prediction. The technique has been compared against LMS [10] and Adaptive-LMS (ALMS) [11] and has been observed to give better results than both the methods.

In this letter the main focus is to present a new and efficient solution based on the Neuro-evolutionary technique called the Cartesian Genetic Programming Evolved Artificial Neural Network (CGPANN) [12] to develop a traffic predictor for scheduling of video streaming data. Neuro-Evolution is the phenomenon of evolving artificially all the entities of an Artificial Neural Network (ANN). In CGPANN Neuro-Evolution exploits the powerful structural properties of Cartesian Genetic Programming [13] and the functional properties of an ANN. This leads to the evolution of all the network parameters including connection weights, node inputs, node functions, network topology as well as the system outputs [12]. The research solution discussed here is specific only to the feed forward CGPANN or FCGPANN. The best evolved network uses historical frame size data to predict the size of the next frame and allocate bandwidth accordingly based on the predicted size.

II. FCGPANN COMPOSITION

Like an ANN, the FCGPANN also comprises of neurons or nodes. Each node lists the inputs, weights, an output and an activation function. A node takes its inputs from the outputs of other nodes or from the external system inputs to the FCGPANN system [12] [13]. Each input is multiplied by a node weight in the range [-1 to +1]. An activation
function operates on the sum of the products of the node outputs. This involves the following steps

1) BEGIN
2) INPUTS
R Total Nodes $N$.
S System Inputs Vector $I$.
T 1D CGPANN Output Vector (Estimated Frame Size parameters) with size $N_{out}$.
U $N_{gen}$: The total population of Genotypes produced in each evolutionary run.
V Alternate Activation Functions act_func[1...$N_{gen}$].
W Number of Inputs to a Node $N_{in}$.
X Mutation Rate $M_{mut}$.
Y Maximum Number of Evolutionary Generations $M_{gen}$.
Z Fitness Threshold $F_{thresh}$.
3) INITIAL: Create an Initial Population of $N_{gen}$ Genotypes using equation 1 and equation 2 by employing a pseudorandom generator (PRG) [14].
4) FITNESS EVALUATION: Root Mean Square Error (RMSE) [15] is calculated for the output produced by each genotype as a criterion for fitness evaluation.
5) MUTATE: Mutate the selected parent by mutating with rate $M_{mut}$ to ascertain whether the required degree of fitness $F_{thresh}$ is achieved. If a parent fulfills the required fitness criteria or the evolutionary runs for the CGPANN evolution reach $M_{gen}$, then proceed to step 6. Otherwise continue to the next step.
6) END

III. EVALUATION AND RESULTS

For the simulation setup of the FCGPANN system, genotypes of 100 nodes with 10% mutation rate have been taken into consideration based on the exemplary results with the quantitative values given in [12]. Based on the $1 + \lambda$ evolutionary strategy, 1 parent and 9 offspring were produced during each evolutionary run. Inputs to the system are the sizes of 10 most recently transmitted frames. The FCGPANN operates on these inputs to produce 10 system outputs that are averaged to obtain the size of the next future frame. Figure 3 shows the implemented simulation setup. The testing scenario involves users viewing videos in which MPEG-4 video streaming data division is taking place. For the training data set 10,000 data points are chosen for $N$ users randomly trying to view a video at different time intervals (further details are given below). The typical frame sizes for the various movies used while testing can be observed from Table I.

The system evolved during the best evolutionary run calculates the frame sizes using equation 4. Equation 4 shows that 8 of the system outputs directly form connection with the $9^{th}$ system input ($i_9$) which is the size of the most recently transmitted frame. It further shows that one output directly takes the size value of the $7^{th}$ frame ($i_7$), while the last system output connects to the output of $17^{th}$ node ($N_{17}$) or neuron. These system outputs are then averaged to estimate the size of the future frame.

$$\text{frameSize} = \frac{8i_9 + i_7 + N_{17}}{10}$$

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>FRAME SIZE PARAMETERS FOR DIFFERENT MPEG-4 SOURCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>Minimum</td>
</tr>
<tr>
<td>Movie 1</td>
<td>95</td>
</tr>
<tr>
<td>Movie 2</td>
<td>20</td>
</tr>
<tr>
<td>Movie 3</td>
<td>24</td>
</tr>
<tr>
<td>Movie 4</td>
<td>20</td>
</tr>
<tr>
<td>Movie 5</td>
<td>149</td>
</tr>
</tbody>
</table>

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
The contribution of the 17th node from equation 4 to the overall system is derived from the active system nodes. The contribution of an active system node can be seen from equation 5. Where \( \psi_i \) and \( w_{ji} (i = 1, \ldots, 5) \) represent the node inputs and connection weights respectively. Based on equation 4 and 5, a mathematical expression representing the final frame size obtained from the FCGPANN system is expressed from equation 6

\[
N_j = f_j(\psi_1, \psi_2, \psi_3, \psi_4, \psi_5) = 1 + e^{-\psi_1 w_{j1} + \psi_2 w_{j2} + \psi_3 w_{j3} + \psi_4 w_{j4} + \psi_5 w_{j5}}
\]

Frame Size = \( \frac{1}{10}(8i_0 + i_1 + f_{f1}(f_{f6}(i_4, i_8, i_2, i_5, i_6), f_{f6}), f_{f6}(i_4, i_8, i_2, i_5, i_6, i_3), f_{f6}(i_4, i_8, i_2, i_5, i_6, i_3, i_4), f_{f6}(i_4, i_8, i_2, i_5, i_6, i_3, i_4, i_5), f_{f6}(i_4, i_8, i_2, i_5, i_6, i_3, i_4, i_5, i_6)) \)

From equation 6 it can be seen that from a total of 100 nodes, 10 nodes are contributing towards estimating the frame size and are thus the active nodes.

For evaluation, different scenarios (see Table III) have been taken into consideration. For the fixed data rate scenario, 10% of the users from the total number of users have been assigned a fixed frame size and the bandwidth. The frame size estimation is carried out for the remaining 90% users and total bandwidth allocation is carried out taking into account both the 10% fixed, and the 90% estimated bandwidth. The variable data rate scenario is the simplest scenario, which simply estimates the frame size for each user and estimates and allocates bandwidth based on the estimated frame sizes. For the probability based bandwidth allocation (Figure 4), the predicted frame sizes from the FCGPANN are used to calculate the required bandwidth for transmission. If the required bandwidth is less than or equal to the available bandwidth then the frames are transmitted. Otherwise frame drop probability is calculated for all users based on their predicted frame sizes. A frame with a larger size is assigned a higher drop probability and visa versa. A random number is then generated in the range of these assigned probabilities. The frame of that user is dropped whose probability is close to the generated random number. This process is repeated until the bandwidth overhead becomes negligible and adequate bandwidth becomes available for frame transmission. The bandwidth allocation and calculation is based on more than 90% accurate frame size prediction which results in an efficient management of frame drop probability.

We performed 20 independent evolutionary runs to obtain an evolutionary model with an average prediction accuracy of 90%. This percentage was obtained for 10,000 iterations during system training. When the iterations were increased to 1 million, the average prediction accuracy became 98%.

Frame prediction performance was evaluated on 50,000 data points of 6 different movies. Figure 6 shows the estimated versus actual values of 500 points of movie 3 which gave the minimum MAPE (Mean Absolute Percentage Error) value of 0.0112 (for all points prediction accuracy 98% approx). The maximum error was for movie 6 0.0733 giving a prediction accuracy of 93%.

The prediction error of movie 3 was compared with previous frame prediction techniques in table II. The priority based scenario resembles the Fixed Data Rate Scenario in the sense that here 10% of the users are being given a fixed higher priority and their bandwidth allocation is also dependent upon this priority (see Figure 5).

<table>
<thead>
<tr>
<th>S.No</th>
<th>Technique</th>
<th>% Prediction Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LARS [10]</td>
<td>0.107</td>
</tr>
<tr>
<td>2</td>
<td>ALMS [11]</td>
<td>0.074</td>
</tr>
<tr>
<td>3</td>
<td>( p = 0.2P ) [9]</td>
<td>0.0352</td>
</tr>
<tr>
<td>4</td>
<td>Proposed Scheme FCGPANN</td>
<td>0.0112</td>
</tr>
</tbody>
</table>
### Scenarios

The authors would like to thank the National ICT R&D Fund Pakistan for providing the support and help in the completion of this research initiative.

**ACKNOWLEDGMENT**

REFERENCES


