CHAPTER 6

CONCLUSION AND FUTURE ENHANCEMENTS

6.1 CONCLUSION

The initial experiments with the proposed moving average based predictors have proved that they are simpler, and need just a single parameter to be updated at every prediction instance, while the performance being much better than the conventional NLMS algorithm for online predictions. The proposed $\alpha$-SMA predictors proved to be simpler and efficient and thus they can be used for on-line predictions. A performance improvement of about 11% is achieved. The significant contribution in this work involves, highlighting the importance of using *average characteristics* of the traces as one of the prediction parameters.

The second part of the work used an ARIMA representation of the GOP trace, over which predictions have been carried out using VSSNLMS. A new update equation has been given for improved performance. Moreover, to control the variability in the error, a simple algorithm has been proposed, which traces back to a few previous instances to check for the right value of step-size. The average prediction performance is achieved upto 16% when compared with the standard ARIMA.

The later parts of the work focused on fuzzy based scene-change characterization for carrying out predictions as accurately as possible during scene-changes. A hybrid FuzzyNLMS mechanism has been proposed and the
This methodology includes a little bit of training at every prediction step for improved accuracy during scene-changes. FuzzyNLMS showed an improvement of upto 24%. A completely new methodology involving preprocessing the data set, prior to prediction, has been proposed. The goal is to investigate the impact of preprocessing on the prediction performance of training based models. The two techniques proposed for data preprocessing were data reduction (by removing anomalies) and data enhancement (by adding more noiseless instances). Predictions using Support Vector Regression, Neural Networks and Linear regression have been experimented with. It can be concluded with the obtained results that the performance of linear regression has been improved with data preprocessing, and so can be preferred for real-time predictions.

The final part focused on a new proposed multiplexing strategy called ERA targeting burstiness factor and resource utilization. The performance of ERA is compared with simple frame-lag and frame-aligned techniques. Even with a completely heterogeneous mix, ERA gave better performance.

Attempts were also made to investigate the applicability of exponential smoothing methods for real-time traffic prediction by predicting I-frame traffic using simple exponential smoothing, while P and B with recursive equation that takes into account, the GOP length as the basis for the order of the predictor.

6.2 FUTURE ENHANCEMENTS

Linear techniques are limited, in the strict sense, to traces with SRD characteristics. Moreover, these techniques fail miserably, when traces show
abnormal fluctuations and trend. With upcoming advanced codecs for video, non-linearity and non-stationarity have become quite common.

In general, video traffic prediction techniques perform badly at shot/scene-change transitions. A promising direction of work would be to make the proposed predictors resilient to shot changes. Furthermore, a wide set of experiments are to be conducted on video traces of various genres as well as traces of longer duration.

As an enhancement, the effectiveness of data preprocessing on other prediction techniques may be analyzed. While the current work focuses on short-term video traffic prediction, an extension may be to work on long-term traffic prediction.

Determining the threshold value either for scene-changes or for prediction error is itself a typical research problem. The work reported in this thesis determines threshold on trial and error basis. It may be advisable to use a mathematical model to derive the optimal threshold value from the attributes of the input trace.

Throughout this work, the focus was on prediction of I frames in particular, for obvious reasons. Prediction of the other type (P, B) frames assumes significance, when the primary goal of the work is to optimally utilize the network resources. Moreover, P and B frames assume equivalent significance as that of I, in the recent trace sequences encoded in the HD format (1920 x 1080), as the I, P, B size relationship (I>>P>>B) is no more true. The current work could be extended using HD traces.

Composite predictions are more challenging than component predictions. Even if P and B frame subsequences exhibit strong intra-correlations, the correlations of these frames with ‘I’ cannot be ignored. A
stochastic model that takes into account the randomness and the cyclic nature in the signal, while preserving the correlation structure should be devised.

The primary goal of traffic prediction is to estimate the future resource requirements of real-time traffic. Many algorithms have been proposed in the literature to estimate the effective bandwidth usage, when a single source is admitted. Practically many sources are often multiplexed at a link, for efficient usage of resources. SMG and utilization improve with multiplexing. It’s proven in many past works that dynamic bandwidth allocation schemes only would improve utilization, for real-time traffic. The current work should be augmented with dynamic bandwidth allocation schemes that greatly reduce renegotiation frequency and improve utilization, while considering delay and other QoS parameters. Single frame-ahead traffic predictions, often lead to increased renegotiation frequencies and also contribute to computational overhead. To address these issues, multi-frame ahead predictions should be preformed for efficient management of network resources for real-time traffic.

Sang and Li (1999) have tried to address the predictability of a traffic by estimating the prediction interval limit within confidence and minimum prediction error over the prediction interval. ARMA and MMPP models were tested. The study concludes by justifying that the ‘prediction accuracy deteriorates with increasing prediction interval’ and ‘traffic smoothing or statistical multiplexing also improves predictability’. The work reported in this thesis could make use of the predictability analysis for VBR video traffic for multi-step ahead predictions.

ARIMA models are particularly suitable for traffic that has no seasonality or trend. MPEG encoded traffic may not exhibit any trend, but certainly seasonality, because of the GOP cyclic order. Works involving ARIMA models for VBR video traffic, tried to eliminate this seasonality by
differencing or smoothing or using other equivalent procedures. Seasonal ARIMA models would make better sense in the context of modeling MPEG encoded VBR traffic. Some recent works (Al Tamimi et al 2010) applied seasonal ARIMA (SAM) or VBR traffic, by adding the seasonal components on standard ARIMA process represented as ARIMA $(p,d,q) \times (P,D,Q)^S$ where $P, D, Q$ are the seasonal counterparts of $p,d,q$ and ‘$S$’ being the period size. The results were compared with AR and ARIMA models. The current work could be extended to include a study on multiplexing traces from different encoders and applying SAM. Another alternative to understanding the seasonal aspect at depth is to apply models used in econometric analysis. Vector Auto Regression (VAR), Structural Vector Auto Regression (SVAR) and Vector Auto Regression Moving Average (VARMA) models are popularly used for econometric analysis. Here, GOP is modeled as a vector and multi-step ahead predictions are done over the chosen prediction horizon. But the modeling complexity is directly related to GOP length. The larger the GOP length the higher is the modeling complexity. Added to this, the current traces have GOP lengths not less than 12. It is therefore, important to find an average solution to modeling traffic using VAR and related models. It is hard to find a work in the existing literature on using VAR models for VBR video traffic. In the work reported by Shin et al (1995), VBR video traffic sources at slice level were characterized as VAR processes. But that work assumes traces without scene-changes and deals with H.261 traffic. This may not quite go well with the traces encoded using latest compression standards. So a detailed study is required to address traffic modeling and prediction using VAR and its variants. To reduce computational complexity, it is proposed to model GOP as a vector with 3 elements, viz., the ‘I’ frame and the average of the ‘P’ and the ‘B’ frame sub sequences within the GOP. Experiments in this direction are being carried out.