Smoothing-aided Support Vector Machine based Nonstationary Video Traffic Prediction Towards B5G Networks

Yingqi Li, Juan Wang, Xiaochuan Sun, Zhigang Li, Miao Liu, and Guan Gui, *Senior Member, IEEE*

***Abstract*—Video services have hold a surprising proportion of the whole network traffic in wireless communication networks. Accurate prediction of video traffic can endow networks with intelligence in resource management, especially for the forthcoming beyond the fifth-generation (B5G) networks. However, the existing approaches fail to accurately predict video traffic with all types of frames, due to the natures of strong long-range dependence, self-similarity and burstiness. Obviously, it is unable to meet the QoS and QoE requirements of dynamic bandwidth allocation. In this paper, we propose the feasibility of advanced machine learning methodology applied in nonstationary video traffic prediction, i.e., smoothing-aided support vector machine (SSVM) model. The model utilizes classical smoothing methods to preprocess video traffic by relieving the drastic fluctuation of video stream. It can provide an effective association for the subsequent support vector regression, as the preprocessed data becomes more smooth and continuous than the original unprocessed one. Experimental results show that our proposed model significantly outperforms the state of the art model, i.e., logistic smooth transition autoregressive, in prediction performance. The superior nonlinear approximation capacity is further demonstrated by visualized statistical analysis.**

***Index Terms*—Variable bit rate, video traffic prediction, resource management, machine learning, support vector machine.**

I. INTRODUCTION

Beyond the fifth-generation (B5G) networks are expected to provide rather reliable services with large-scale connectivity, superhigh transmission rate, ultra-low latency, much enhanced security, very little energy loss, and excellent quality of experience (QoE) [1]. Given these superiorities, the rising B5G networks should be intelligent enough to cope with more autonomous [2]–[4], self-organizing and dynamic situations, such as dynamic topologies, application-oriented storage, autonomic and scalable computing, and various quality of service (QoS) requirements. For the design of such smart networks, machine learning is deemed as a critical element and it has been adopted in many applications [5]–[16].

With the deployment of 5G mobile networks [17], traffic has been governed by explosive video applications including IPTV, video assisted surgery, video conference and augmented reality. For the following B5G networks, this situation can be further exacerbated. Network operators are confronted with a significant issue: if they intend to reduce the bandwidth provision for a users requirement to enable more users to access the network, it may lead to potential packet loss that has a significant impact on the communication quality; on the other hand, if they oversupply bandwidth to cope with possible bursts in data, bandwidth wastage occurs. These diversified multimedia services certainly require higher bandwidth configuration in contrast to traditional network traffic, such as web and email. Generally, excellent rate control mechanisms are design for video stream in dynamic heterogeneous network scenarios. It yields well-known variable bit rate (VBR) traffic. In this case, resource management should possess potential robustness and efficiency, aiming at guaranteeing high bandwidth utilization, so that any QoS requirement can be protected [18]–[20]. Conventional fixed allocation of network resource suggests the reservation of a great many bandwidth for a specific QoS guarantee. Nevertheless, dynamic allocation under the framework of video traffic prediction is an alternative technique to solve this intractable issue [21].

From data characteristics, VBR traffic shows the properties of strong nonlinearity, variability, and slow-decaying autocorrelations between samples. These implicit traits determine the long range dependent (LRD) of video streaming. Fixed allocation mode has to abandon the available bandwidth. To this problem, dynamic bandwidth allocation appears to prevent probable network congestion. In this dynamic mode, frame size prediction is regarded as a crucial supplementary mean. However, the one-step ahead prediction models on LRD video traffic can not respond to dynamic bandwidth allocation in high-latency links. Hence, long-term video traffic prediction should be required for ensuring flexible network control strategies [22].

A great deal of machine learning research has been focused on multiple-steps ahead video traffic prediction, mainly including statistical learning methods and neural networks (NNs). The key issue is how to improve nonlinear approximation

This work is supported in part by the National Science and Technology Major Project of China under Grant TC190A3WZ-2, National Natural Science Foundation of China under Grant 61571240 , the Natural Science Foundation of Hebei Province under Grant F2018209181, the Hebei Colleges and Universities Science Foundation under Grant QN2018115, the Implementation Program on the Innovative Development of the Qinhuangdao Marine Economy under Grant 2017-103, the Jiangsu Specially Appointed Professor under Grant RK002STP16001, the Innovation and Entrepreneurship of Jiangsu High-level Talent under Grant CZ0010617002, the Six Top Talents Program of Jiangsu under Grant XYDXX-010, the 1311 Talent Plan of Nanjing University of Posts and Telecommunications. (*Corresponding authors: Xiaochuan Sun, Guan Gui*)

Y. Li, X. Sun and Z. Li are with College of Artificial Intelligence, North China University of Science and Technology, Tangshan 063210, China, (e-mails: {liyingqi, sunxiaochuan, lizhigang}@ncst.edu.cn)

J. Wang, M. Liu, and G. Gui are with College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China (e-mail: {1219012920, liumiao, guiguan}@njupt.edu.cn).

capacity as much as possible [23]–[25]. It is demonstrated that classic AR, ARMA and ARIMA can only seizure both linearity and short range dependencies (SRD) hidden between video data, but are powerless to LRD, which leads to weak performance in video traffic prediction [26]. In [27], a logistic smooth transition autoregressive was developed for VBR video traffic prediction, where adaptive least mean square and its extensions were considered for the determination of model parameters. It could achieve a nonlinear approximation up to 300-step ahead prediction. In [28], a seasonal autoregressive method was used for high definition video stream modeling. In [29], a Markovian model was presented for the specific B-frame prediction, aiming at reducing bandwidth requirements and smoothing the video stream. In [30], a hybrid deep computing model, consisting of stacked restricted Boltzmann machine and minimum complex reservoir, was proposed for one-step video frame size prediction. In [31], six NN models were compared according to the capacities in predicting each frame type of MPEG-4, i.e., I, P and B. Moreover, this work was further extended, where more NNs were introduced for whole MPEG-4 video stream processing [32]. In spite of these positive results, the accurate multiple-steps ahead prediction of video traffic in autonomous network management remains an challenging task and the technology is far from actual deployment. This is mainly due to the hight-variability in frame size of video traffic, i.e., burstiness. This paper focuses on exploring more effective multi-steps ahead prediction method of video traffic supporting intelligent resource management towards B5G networks.

Support vector machine (SVM) is deemed as a powerful machine learning tool for prediction, whose success lies in Vapniks pioneering study in statistical learning theory [33], including robustness, restorability to over-fitting, optimal solution, and modularity of kernel mapping. Concretely, SVM has remarkable characteristics of impressive generalization capability, no local minimum and sparse representation of solution [34]. It has been used widely for multi-steps ahead prediction in numerous scenarios, such as inbound tourist arrivals, physiological abnormal signals and consumer's heat load [35]–[37]. Besides, it has been demonstrated that data smoothing is beneficial for nonlinear approximation in time series prediction [38]. This preprocessing method could handle significant fluctuations and outliers by adjusting the built-in sliding window. Inspired by these, we consider the combination of SVM and smoothing for the multi-steps ahead prediction of VBR video traffic.

In this paper, we propose a smoothing-aided SVM (SSVM) model for VBR video traffic prediction method in order to meet the requirement of dynamic resource allocation. Structurally, it can be viewed as a uniform and successive system with functional modules of smoothing and regression. The employed smoother is capable of reducing the size of leptokurtic frame, thereby alleviating drastic fluctuation of video stream. This way can facilitate obtaining a well-behaved SVM for multi-steps ahead prediction. The effectiveness of our proposal is illustrated on different video traffic prediction tasks.

The rest of this paper is organized as follows. For the dynamic resource allocation demands of B5G networks, Section II provides a detailed description on the smoothed-aided SVM model for nonstationary video traffic prediction. Section III gives some experiments to verify the performance of our proposal, including smoothing analysis, prediction accuracy and statistical validation. Finally, Section IV concludes the paper and points out the future research directions.

II. MODEL DESCRIPTION AND OUR PROPOSED METHOD

Tackling the challenges of resource management in B5G network environments, a promising machine learning model, combining with smoothing methods and SVM, is applied for VBR video traffic prediction, called SSVM. In fact, it is derived from the promising finding in [38]. In structure, our video traffic prediction model consists with functional devices of smoother and predictor. The prioritized smoothing aims at effectively eliminating the tricky abrupt fluctuations of video traffic, i.e., short-term burstiness. It is conducive to providing more suitable initial points for nonlinear approximation. The well-behaved SVM serves as a predictor, characterised by a high-dimensional feature transformation based on kernel projecting. Combining these two components, our model is expected to achieve a strong prediction performance. For a better insight, we provide a detailed description on the functionalities of both.

Generally known, significant burstiness often appears in compressed video over numerous time scales. This is largely due to the special frame structures determined by different encoding schemes and high variations related to scenes. Moreover, burstier VBR traffic can be formed from aggregations of multiple video streams attaching orchestrated text, audio and images. In our scenario, smoothing is used to decrease bit rate variability appropriately, so that it can produce a relative smooth video stream in priority for acceptable nonstationary VBR video traffic prediction. The following four smoothing methods are considered [38], given by

- Moving average (MA) is viewed as a simplest version of smoothing methods. Its readout can be generated from the average of its corresponding neighboring points in a customizable window, expressed as follows

$$Z(z_i) = \frac{1}{2n+1} \sum_{k=-n}^{n} z_{i+k} \tag{1}$$

  where z denotes a sample of this smoother input, $2n+1$ is the size of moving window, and j offers an index of the current readout value.

- Gaussian smoothing (GS) is on the basis of the well-known normal distribution, defined by the following probability density function

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \tag{2}$$

  where x is the sample value obeying Gaussian distribution, σ, σ^2 and μ denotes the standard deviation, the variance and the mean of samples, respectively.

- Robust locally weighted smoothing (RLWS) is achieved by a bisquare function w_i, given by

$$w_i = \begin{cases} (1-(r_i/\kappa)^2)^2 & |r_i| < \kappa \\ 0 & |r_i| \geq \kappa \end{cases} \tag{3}$$

where r_i denotes the i-th residual value generated form the current smoothing, and κ is the sixfold median absolute deviation (MAD) of these residuals, given by

$$MAD = median(|r|) \tag{4}$$

- Savitzky-Golay smoothing (SG) is actually regarded as a low-pass filter that has the ability of local least-square polynomial approximation. Considering a window $i \in (M, N-M)$, the smoothing has the following form

$$SG(i) = \sum_{k=0}^{n} a_k i^k \tag{5}$$

where N denotes the sample size, n denotes the polynomial degree, and a denotes the vector of coefficients, calculated by

$$E = \sum_{i=-M}^{M} [SG(i) - x_i]^2. \tag{6}$$

Once the smoothing on VBR traffic streams is completed, the LS-SVM based predictor works for the subsequent nonlinear approximation. For a given video dataset $V = \{(s_i, o_i), s_i \in R^n, o_i \in R\}_{i=1}^{l}$, the regression output is measured by a nonlinear mapping function $\varphi(\cdot)$, defined as follows

$$f(s) = w^T \cdot \varphi(s) + b \tag{7}$$

where w denotes a weight vector, and b is the corresponding bias. Given this, a separating hyperplane emerges for the fitting of o and s.

Through the rule of structure risk minimization, the regression related to LS-SVM can be converted to the following optimization issue, given by

$$\min_{w,b,e} O(w,b,e) = \frac{1}{2}\left(\|w\|^2 + C\sum_{i=1}^{l} e_i^2\right) \tag{8}$$

obeying

$$o_i = w \cdot \varphi(s) + b + e_i, i = 1, 2, 3, \cdots, l \tag{9}$$

where e_i denotes a randomly generated error, and $C > 0$ denotes a regularization measure. It can be solved by the Lagrange multiplier method, formulated as follows

$$\begin{aligned} L(w, b, e_i, \alpha_i) = O(w, b, e_i) - \\ \sum_{i=1}^{l} \alpha_i [w \cdot \varphi(s_i) + b + e_i - o_i] \end{aligned} \tag{10}$$

where α_i is a Lagrange multiplier. Furthermore, we have the following partial differentials

$$\begin{aligned} &\frac{\partial L}{\partial w} = 0 \to \sum_{i=1}^{l} \alpha_i \varphi(s_i) = w \\ &\frac{\partial L}{\partial w} = 0 \to \sum_{i=1}^{l} \alpha_i = 0 \\ &\frac{\partial L}{\partial e_i} = 0 \to C e_i = \alpha_i \\ &\frac{\partial L}{\partial \alpha_i} = 0 \to w \cdot \varphi(s_i) + b + e_i - o_i = 0 \end{aligned} \tag{11}$$

TABLE I: Model performance on the prediction tasks of NBC News video traffic traces (QP=10 and QP=34) over different prediction steps.

Video	Step	QP	Predictor	
			SSVM	LSTAR [27]
News	1	10	0.0231	0.3466
		34	0.0330	0.3424
	50	10	0.2823	0.3409
		34	0.1871	0.3960
	200	10	0.3616	0.4015
		34	0.3506	0.4822

Through eliminating the common parameters w and e_i, we have

$$\begin{bmatrix} 0 & \vec{I}^T \\ \vec{I} & \Omega + C^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{12}$$

where

$$\begin{cases} I = [1, 1, 1, \cdots, 1]^T \\ \Omega = \{\Omega_{ij} | i, j = 1, \ldots, l\} \\ \alpha = [\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_l]^T \\ y = [y_1, y_2, y_3 \cdots, y_l]^T \end{cases} \tag{13}$$

and $\Omega_{ij} = \varphi(s_i)^T \varphi(s_j) = K(s_i, s_j)$, for $i, j = 1, 2, 3, \cdots, l$ with K a Mercer kernel function. In our case, the linear kernel, the polynomial kernel, the radial basis function kernel (RBF) and multi-layer perceptron kernel (MLP) are considered, given by

$$K(s_j, s_i) = s_j \cdot s_i \tag{14}$$

$$K(s_j, s_i) = ((s_j \cdot s_i) + 1)^d \tag{15}$$

$$K(s_j, s_i) = esp\left(-\frac{\|s_j - s_i\|^2}{\sigma^2}\right) \tag{16}$$

$$K(s_j, s_i) = tanh(\eta < s_j, s_i > + \theta) \tag{17}$$

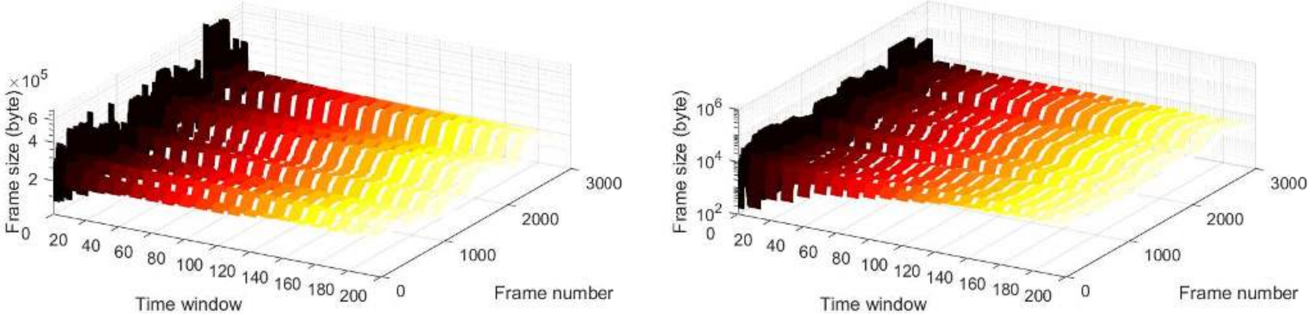
Solving the linear equation (12), α and b can be determined. It implies the completion of the LS-SVM training. Our optimal model is given by

$$f(x) = \sum_{i=1}^{l} \alpha_i K(x, x_i) + b \tag{18}$$

Subsequently, the trained model can be used for real-world VBR video traffic prediction.

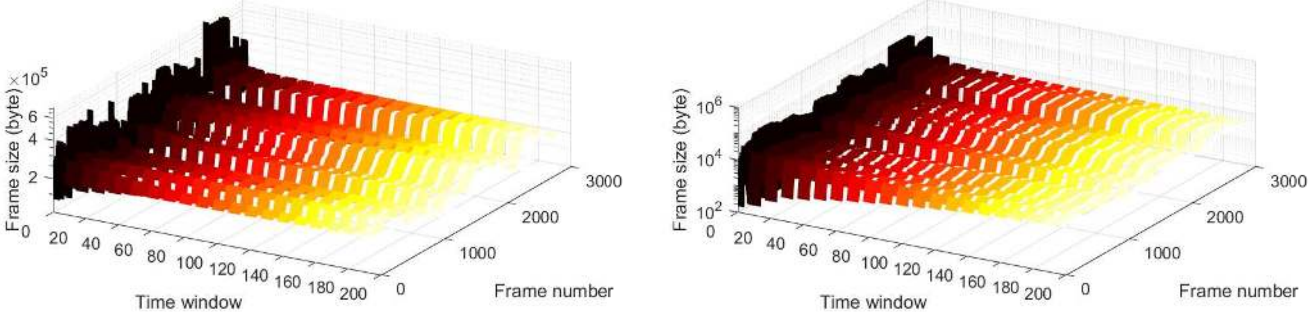
III. EXPERIMENT RESULTS AND DISCUSSIONS

In this section, an exhaustive simulation is provided for the performance assessment of our SSVM paradigm, considering real-world nonstationary VBR video traces. During evaluations, we consider different types of smoothers under user-defined sliding window, such as MA, GS, RLWS and SG, as well



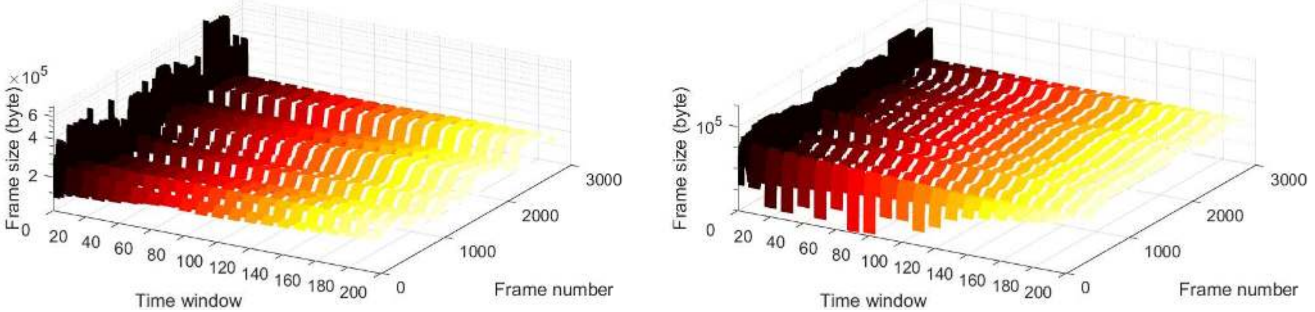
(a) QP=10, MA

(b) QP=34, MA



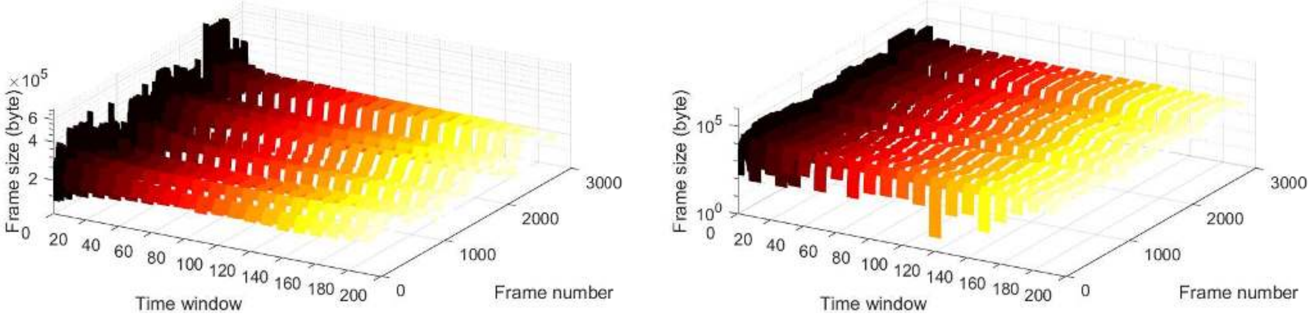
(c) QP=10, GS

(d) QP=34, GS



(e) QP=10, RLWS

(f) QP=34, RLWS

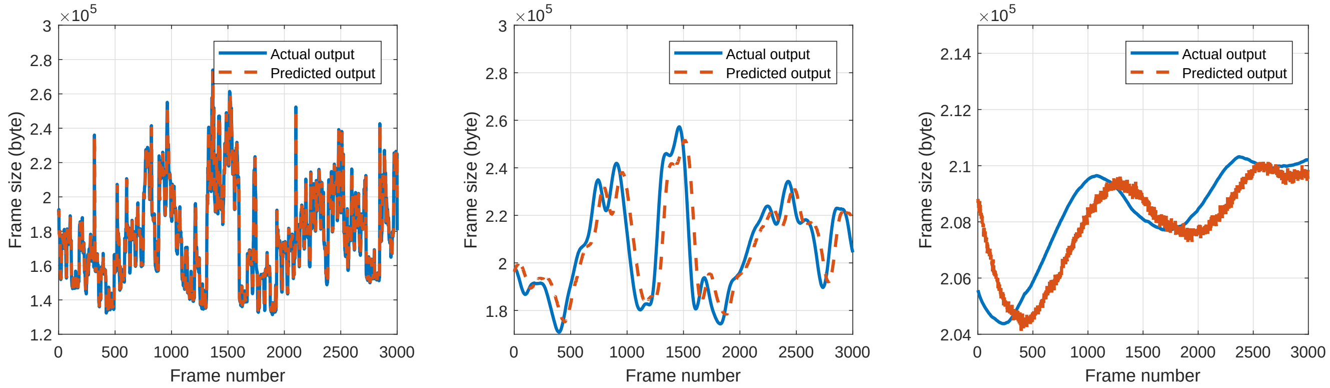


(g) QP=10, SG

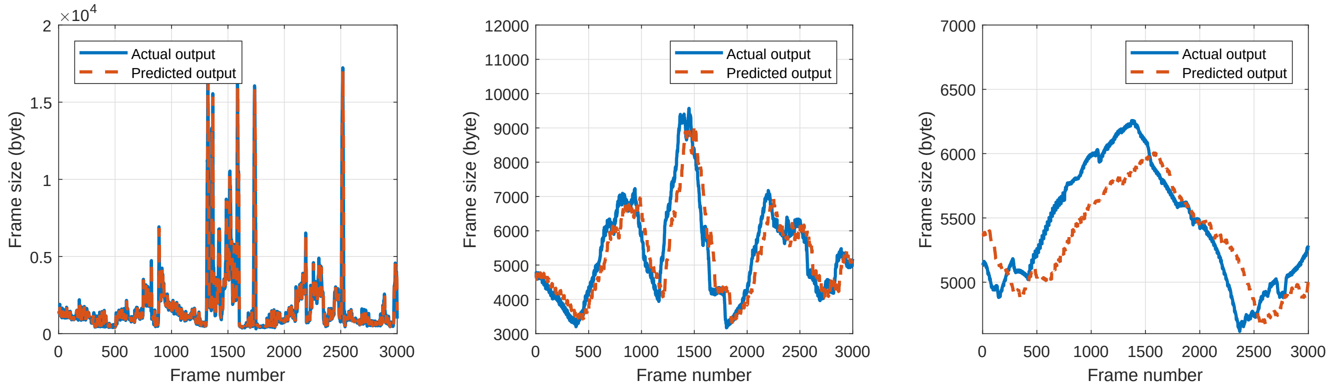
(h) QP=34, SG

Fig. 1: Comparisons of the original NBC News video traffic series (QP=10 and QP=34) and the smoothed ones via time windows for the considered smoothing methods. The time window is set to the values from 0 to 200 (from left to right), where the time step is 10.

as SVM-based predictor with the linear kernel. To prove the effectiveness of our hybrid model, we also compare the logistic smooth transition autoregressive model (LSTAR) in the same application scenario [27]. Model parameters are determined by



(a) QP=10, Step=1 (RLWS) (b) QP=10, Step=50 (GS) (c) QP=10, Step=200 (GS)



(d) QP=34, Step=1 (RLWS) (e) QP=34, Step=50 (MA) (f) QP=34, Step=200 (SG)

Fig. 2: Output comparisons of our SSVM model with optimal smoothing on the prediction tasks of NBC News video traffic traces (QP=10 and QP=34) over different prediction steps.

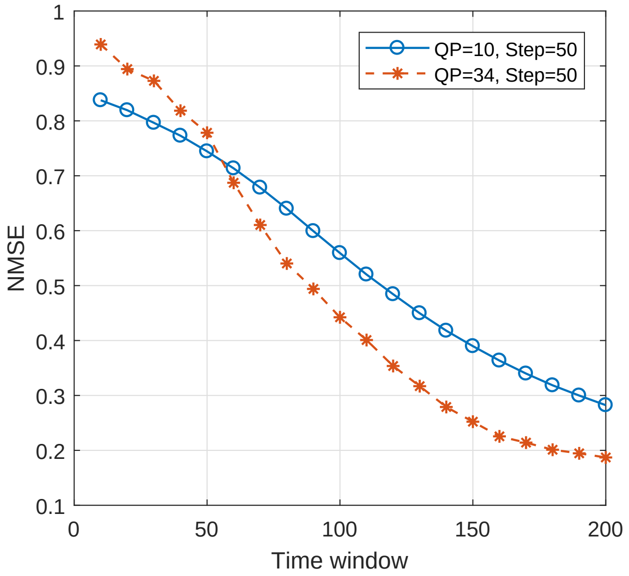


Fig. 3: NMSE versus time window size for SSVM on the 50-step ahead prediction tasks of NBC News video traffic traces.

the grid search method. All results are given over an average of the experiments repeated 10 times.

A. Data and Assessment Criteria

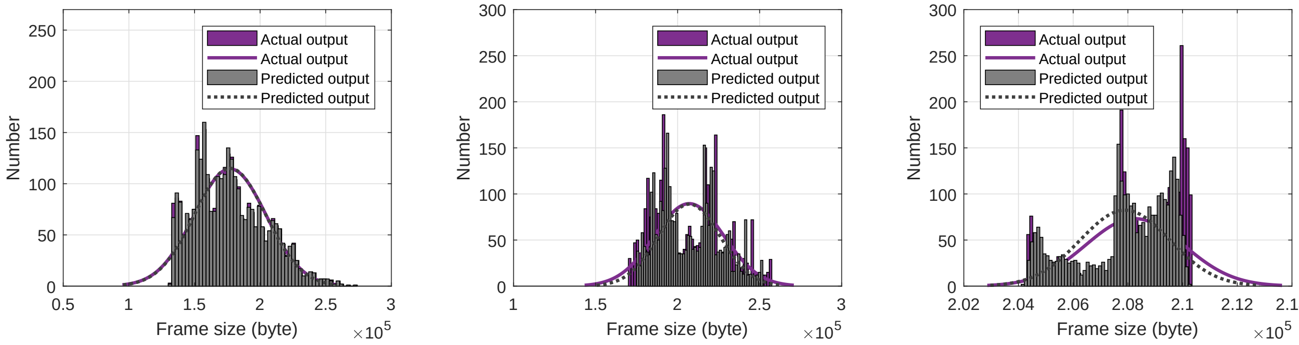
Here, NBC news dataset from the video trace library is used in experiments, found at *http://trace.kom.aau.dk/h264/index.html*. With regard to its characteristics, there exist frequent transitions among scenes. Two types of NBC news video stream (QP=10 and QP=34) are considered, and the number of frame is 49521. These video frames encoded by H.264/AVC have a structure of G16B3 GOP, in which 16 frames are introduced between adjacent I and B frames. The prediction performance of our model is measured by normalized mean squared error (NMSE), given by

$$NMSE = \frac{1}{L\sigma_x^2}\sum_{l=1}^{L}(x(n)-\hat{x}(n))^2 \tag{19}$$

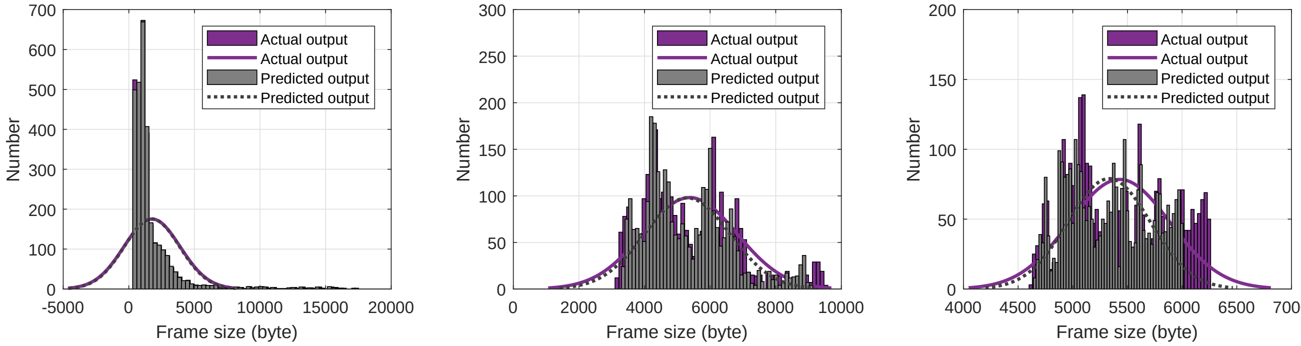
where σ_x^2 is the variance of video traffic series x over L points, $\hat{x}(n)$ denotes the predicted value of $x(n)$.

B. Smoothing Analysis

To investigate the effect of the smoothing on the series characteristics, we plot the trends of the 21 video traffic series as a function of the time window in Fig. 1, where the time window is 0, meaning the original series, while its other settings correspond to the 20 smoothed series. From



(a) QP=10, Step=1 (RLWS,7) (b) QP=10, Step=50 (GS, 200) (c) QP=10, Step=200 (GS, 2500)



(d) QP=34, Step=1 (RLWS, 7) (e) QP=34, Step=50 (MA, 200) (f) QP=34, Step=200 (SG, 2500)

Fig. 4: Histogram fitting of our SSVM model with optimal smoothing on the prediction tasks of NBC News video traffic traces (QP=10 and QP=34) over different prediction steps.

this figure, it can be seen that the original video series shows a dramatic fluctuation, since there exit a great deal of bursty sources. The specific data characteristic enables accurate multi-step ahead prediction of video traffic to be a quite challenging work. In fact, smoothing effectively reduces significant oscillations in our scenario. Moreover, as the time window increases, this smoothed series become relatively stationary for the considered smoothing methods, which is highly beneficial for prediction. However, this increase is not arbitrary, because video streams may have fundamental changes in internal and external characteristics, such as self-similarity, bursty and trend. Hence, the reasonable choice of the time window in smoothing is critical for given video traffic prediction tasks. It requires an effective tradeoff between prediction accuracy and data characteristics, but this is beyond the scope of this paper. Our study aims at exploring the feasibility in smoothing assisted video traffic prediction for dynamic resource management.

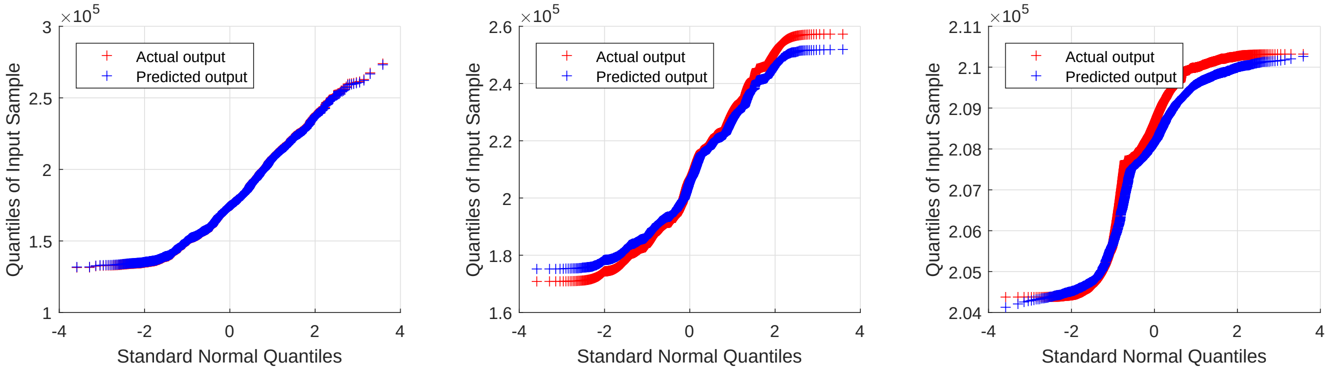
C. Prediction Performance

Fig. 2 plots comparison curves of the actual signal and the predicted output yielded by our SSVM model over frame numbers, where multiple-steps ahead prediction, such as 1, 50 and 200, is considered for NBC News video traffic traces (QP=10 and QP=34), as well as the most suitable smoothing method and its time window are suggested for each prediction mode. It can be observed from this figure that our model readout is consistent with the actual series, especially for the one-step ahead prediction. Nevertheless, this trend fitting results become relative poor as the prediction step increases, as shown in Fig. 2(c) and Fig. 2(f). The superior performance of our model over LSTAR is further listed in Table I. It is worth noting that compared with LSTAR, SSVM has NMSE decreases of 32.35% and 30.94% for one-step ahead prediction, respectively. In fact, its significant performance advantages are mainly ascribed to the prepositive smoothing mechanism controlled by time window, which can effectively alleviate the high variability of video frames, i.e, burstiness. Taking an example of 50-step ahead prediction, the relation between time window and prediction performance is shown in Fig 3. Obviously, our model is relatively sensitive to the choice of time window. The bigger time window can lead to the better prediction performance. However, the time window is not infinitely increased, because it can make the smoothed video streams lose the essential characteristics.

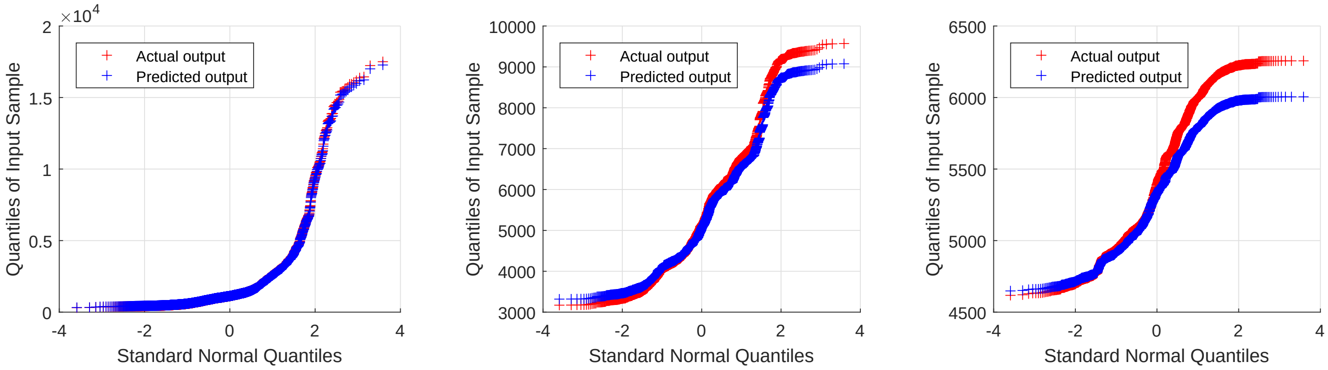
D. Statistical Validation

In the following part, from the perspective of statistical analysis, we verify the nonlinear approximation capacity of our SSVM in the video traffic prediction tasks. The methods, such as histogram fitting, QQ plot and box-plot, are used in the case.

Fig. 4 provides graphical representations of histogram plots on the shapes and distributions of the actual and predicted



(a) QP=10, Step=1 (RLWS, 7) (b) QP=10, Step=50 (GS, 200) (c) QP=10, Step=200 (GS, 2500)



(d) QP=34, Step=1 (RLWS, 7) (e) QP=34, Step=50 (MA, 200) (f) QP=34, Step=200 (SG, 2500)

Fig. 5: QQ plots of our proposed SSVM model with optimal smoothing on the prediction tasks of NBC News video traffic traces (QP=10 and QP=34) over different prediction steps.

outputs for SSVM in our video traffic prediction scenario. It is seen that for one-step ahead prediction, our model can achieve good fittings of vertical bars and data distribution for the two actual video trace. Whereas, as the prediction step is up to 200, this fitting gets worse, meaning the performance degradation of SSVM. Fig. 5 shows the QQ plots of the actual and predicted video traffic, yielded by plotting the quantiles of video data distribution versus the ones of normal distribution. Obviously, the well-fittings in Fig. 5(a) and Fig. 5(d) indicate that SSVM behaves well in the scenario of one-step ahead prediction. However, in Fig. 5(b)-(c) and Fig. 5(e)-(f), the divergences between distributions mean relatively poor multiple-steps ahead prediction performance. This case can be further demonstrated in Fig. 6 based on the popular box-plot visualization, where we gives an insight on the superiority of our model through these comparisons of the upper and lower quartiles, the upper and lower bound, the median and the outliers. Besides, we can also observe that the smoothed video streams have few outlier, meaning the decrease of bursty points. Above all, these results correspond to the findings in Fig. 2. It demonstrates the efficacy of the SSVM model in our nonstationary VBR video traffic prediction tasks.

IV. CONCLUSION

In this paper, we explore the application of smoothed support vector machines for nonstationary video traffic prediction in the B5G network environment. It is actually a hybrid data processing model, where multiple smoothing methods are considered for video stream preprocessing. The preceding smoother can effectively alleviate burstiness in video data, providing excellent initialization points for the following SVM-based approximation. Experimental results verify the utility of our SSVM model in terms of prediction accuracy and statistical comparison. Further research will concentrate on the adaptive determinations of smoothing methods for given VBR video stream, as well as the selection of the corresponding time windows.

REFERENCES

[1] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, Y. Zhang, "The roadmap to 6G: AI empowered wireless networks," *IEEE Wirel. Commun.*, vol. 57, no. 8, pp. 84–90, 2019.

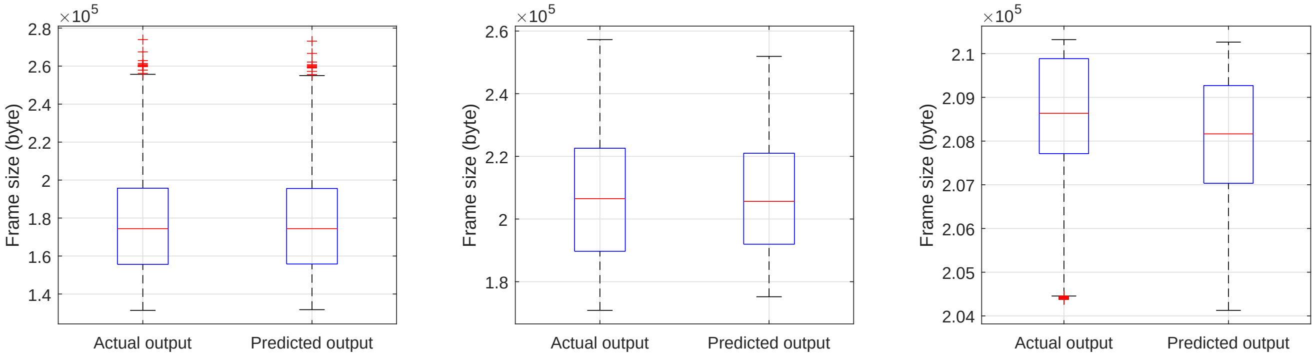
[2] N. Kato, Z. Md. Fadlullah, F. Tang, B. Mao, S. Tani, A. Okamura, J. Liu, "Optimizing space-air-ground integrated networks by artificial intelligence," *IEEE Wirel. Commun.*, vol. 26, no. 4, pp. 140–147, 2019.

[3] G. Gui, M. Liu, F. Tang, N. Kato, and F. Adachi, "6G: Opening new horizons for integration of comfort, security and intelligence," submitted to *IEEE Wirel. Commun.*, doi: 10.36227/techrxiv.11634669.v1

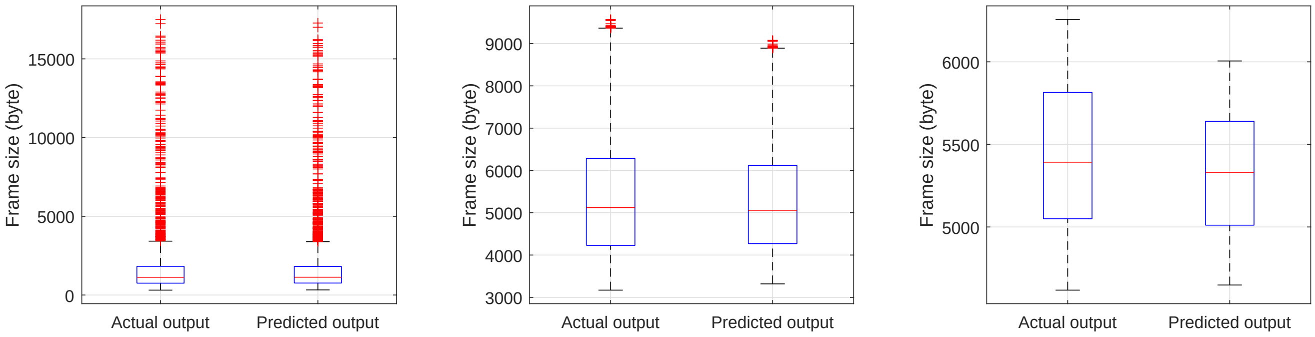
[4] H. Gacanin, "Autonomous wireless systems with artificial intelligence: A knowledge management perspective," *IEEE Veh. Technol. Mag.*, vol. 14, no. 1, pp. 51–59, 2019.

[5] G. Gui, H. Huang, Y. Song, and H. Sari, "Deep learning for an effective nonorthogonal multiple access scheme," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8440–8450, 2018.

[6] H. Huang, Y. Peng, J. Yang, W. Xia, and G. Gui, "Fast beamforming design via deep learning," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 1065–1069, 2020.



(a) QP=10, Step=1 (RLWS, 7) (b) QP=10, Step=50 (GS, 200) (c) QP=10, Step=200 (GS, 2500)



(d) QP=34, Step=1 (RLWS, 7) (e) QP=34, Step=50 (MA, 200) (f) QP=34, Step=200 (SG, 2500)

Fig. 6: Box-plots of our SSVM model with optimal smoothing on the prediction tasks of NBC News video traffic traces (QP=10 and QP=34) over different prediction steps.

[7] X. Sun, G. Gui, R. Liu, Y. Li and Y. An, "ResInNet: A novel deep neural network with feature reuse for internet of things," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 679–691, 2019.

[8] B. Mao, *et al.*, "Routing or computing? The paradigm shift towards intelligent computer network packet transmission based on deep learning," *IEEE Trans. Comput.*, vol. 66, no. 11, pp. 1946–1960, Nov. 2017.

[9] G. Gui, F. Liu, J. Sun, J. Yang, Z. Zhou, and D. Zhao, "Flight delay prediction based on aviation big data and machine learning," *IEEE Trans. Veh. Technol.*, in press, doi: 10.1109/TVT.2019.2954094.

[10] B. Mao, F. Tang, Z. Md. Fadlullah, and N. Kato, "An intelligent route computation approach based on real-time deep learning strategy for software defined communication systems," *IEEE Trans. Emerg. Topics Comput.*, in press, doi: 10.1109/TETC.2019.2899407.

[11] J. Sun, W. Shi, Z. Han, J. Yang, and G. Gui, "Behavioral modeling and linearization of wideband RF power amplifiers using BiLSTM networks for 5G wireless systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 11, pp. 10348–10356, 2019.

[12] B. Mao, Z. Md. Fadlullah, F. Tang, N. Kato, O. Akashi, T. Inoue, and K. Mizutani, "Routing or computing? The paradigm shift towards intelligent computer network packet transmission based on deep learning," *IEEE Trans. Comput.*, vol. 66, no. 11, pp. 1946–1960, Nov. 2017.

[13] F. Tang, Z. Md. Fadlullah, B. Mao, and N Kato, O. Akashi, T. Inoue, and K. Mizutani,"An intelligent traffic load prediction-based adaptive channel assignment algorithm in SDN-IoT: A deep learning approach," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 5141–5154, 2018.

[14] B. Mao, F. Tang, Z. Md. Fadlullah, and N. Kato,"A novel non-supervised deep learning based network traffic control method for software defined wireless networks," *IEEE Wirel. Commun. Mag.*, vol. 25, no. 4, pp. 74–81, 2018.

[15] N. Kato, Z. Md. Fadlullah, B. Mao, F. Tang, O. Akashi, T. Inooue, and K. Mizutani, "The deep learning vision for heterogeneous network traffic control: Proposal, challenges, and future perspective," *IEEE Wirel. Commun. Mag.*, vol. 24, no. 3, pp. 146–153, 2017.

[16] Y. Wang, M. Liu, J. Yang, and G. Gui, "Data-driven deep learning for automatic modulation recognition in cognitive radios," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 4074–4077, Apr. 2019.

[17] R. Yu, J. Ding, X. Huang, M. Zhou, S. Gjessing and Y. Zhang, "Optimal resource sharing in 5G-enabled vehicular networks: a matrix game approach," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 7844–7856, 2016.

[18] Y. Xu, G. Li, Y. Yang, M. Liu and G. Gui, "Robust resource allocation and power splitting in SWIPT enabled heterogeneous networks: a robust minimax approach," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10799–10811, 2019.

[19] M. Liu, T. Song, G. Gui, J. Hu, and H. Sari, "Deep cognitive perspective: resource allocation for NOMA based heterogeneous IoT with imperfect SIC," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2885–2894, 2019.

[20] Z. Zhou, Y. Guo, Y. He, X. Zhao, W. Bazzi, "Access control and resource allocation for M2M communications in industrial automation," *IEEE Trans. Indus. Infor.*, vol. 15, no. 5, pp. 3093–3103, 2019.

[21] C. Katris, S. Daskalaki, "Dynamic bandwidth allocation for video traffic using FARIMA-based forecasting models," *Journal of Network and Systems Management*, vol. 27, no.1, pp. 39–65, 2019.

[22] Z. Tan, Y. Wang, Y. Zhang and J. Zhou, "A novel time series approach for predicting the long-term popularity of online videos," *IEEE Trans. Broadcast.*, vol. 62, no. 2, pp. 436–445, 2016.

[23] T. Zhou, Y. Li, and G. Gui, "Noise learning based discriminative dictionary learning algorithm for image classification," *J. Franklin Inst.*, in press, doi: 10.1016/j.jfranklin.2020.01.007

[24] Y. Li, F.Dai, X. Cheng, L. Xu, and G. Gui, "Multiple-prespecified-dictionary sparse representation for compressive sensing image reconstruction with nonconvex regularization," *J. Franklin Inst.*, vol. 356, no. 4, pp. 2353–2371, 2019.

[25] Z. Yang, L. Fan, Y. Yang, Z. Yang, and G. Gui, "Generalized singular value thresholding operator based nonconvex low-rank and sparse decomposition for moving object detection," *J. Franklin Inst.* vol. 356, no. 16, pp. 10138–10154, 2019.

[26] N. Haghighat, H. Kalbkhani, M. G. Shayesteh, "Variable bit rate video

traffic prediction based on kernel least mean square method," *IET Image Processing*, vol. 9, no. 9, pp. 777–794, 2015.

[27] H. Kalbkhani, M. G. Shayesteh, N. Haghighat, "Adaptive lstar model for long-range variable bit rate video traffic prediction," *IEEE Transactions on Multimedia*, vol. 19, no. 5, pp. 999-1014, 2016.

[28] D. R. Marković, A. M. Gavrovska and I. S. Reljin, "4K video traffic analysis using seasonal autoregressive model for traffic prediction," *Telecommunications Forum*, Belgrade, 2016, pp. 1-4.

[29] A. Kalampogia, P. Koutsakis, "H.264 and H.265 video bandwidth prediction," *IEEE Transactions on Multimedia*, vol. 20, no. 1, pp. 171–182, 2017.

[30] X. Sun, T. Li, and Y. Li, "Recurrent neural system with minimum complexity: A deep learning perspective," *Neurocomputing*, vol. 275, no. 2018, pp. 1333–1349, 2018.

[31] A. Abdennour, "Evaluation of neural network architectures for MPEG-4 video traffic prediction," *IEEE Trans. Broadcast.*, vol. 52, no. 2, pp. 184-192, 2006.

[32] J. P. Kharat, "Comparative study of various neural network architectures for MPEG-4 video traffic prediction," *International Journal of Advances in Applied Sciences*, vol. 6, no.4, pp. 283–292, 2017.

[33] C. Cortes, V. Vapnik, "Support-vector networks,“ *Machine learning*, vol. 20, no. 3, pp. 273-297, 1995.

[34] K. Wang, C. Xu, Y. Zhang, S. Guo and A. Y. Zomaya, "Robust big data analytics for electricity price forecasting in the smart grid," *IEEE Trans. Big Data*, vol. 5, no. 1, pp. 34–45, 2019.

[35] H. ElMoaqet, D. M. Tilbury and S. K. Ramachandran, "Multi-step ahead predictions for critical levels in physiological time series,“ *IEEE Transactions on Cybernetics*, vol. 46, no. 7, pp. 1704-1714, 2016.

[36] M. Protić, S. Shamshirband, D. Petković, et al, "Forecasting of consumers heat load in district heating systems using the support vector machine with a discrete wavelet transform algorithm,“ *Energy*, vol. 87, no. 2015, pp. 343-351, 2015.

[37] L. J. Wu, G. H. Cao, "Seasonal SVR with FOA algorithm for single-step and multi-step ahead forecasting in monthly inbound tourist flow," *Knowledge-Based Systems*, vol. 110, no. 2016, pp. 157-166, 2016.

[38] Z. Li, N. Wang, Y. Li, X. Sun, M. Huo and H. Zhang, "Collective efficacy of support vector regression with smoothness priority in marine sensor data prediction," *IEEE Access*, vol. 7, no. 1, pp. 10308–10317, 2019.