Vol. XVI & Issue No. 12 December - 2023
INDUSTRIAL ENGINEERING JOURNAL

ADAPTIVE PLAYOUT MODELLING OF VOICE PACKETS USING ADAPTIVE NEURO - FUZZY INFERENCE APPROACH

Dr. Priya Chandran
Dr. Suhasini Vijaykumar

Abstract

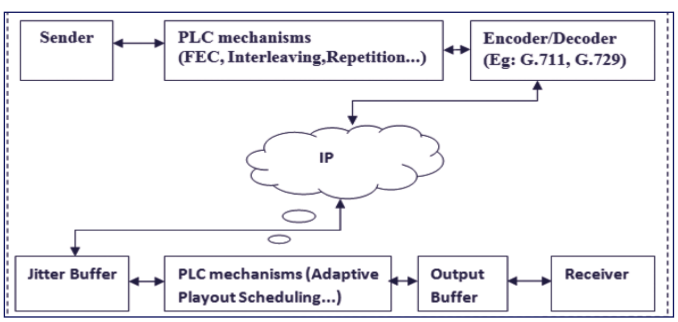
Rapid growth in the deployment of VoIP applications increases the need for assessing quality of voice communication. The end users expect the same quality of service as traditional telephone system, which is circuit switched. Since the network is originally designed for data communication, VoIP users experience degradation in voice quality due to factors like latency, jitter, packet loss, bandwidth etc. So it is important to conceal the packet loss, so that the end users can experience high quality communication. In this paper we propose a receiver- based packet loss concealment mechanism by predicting the jitter buffer delay using Adaptive Neuro-Fuzzy Inference System (ANFIS) model.

Keywords: VoIP, QoS, ANFIS, Neuro-Fuzzy system, Jitter, Packet loss

1.INTRODUCTION

VoIP is the transmission of telephone services over IP-based packet switched networks switched networks. Voice conversations are converted to digitized data and packetized for transmission across a network. Duration of the signal contained in VoIP packet are typically between 10 and 40 ms (B. H. Kim et al., 2013). Sender-side packet loss concealment mechanisms are applied on the voice conversations and then using a coder, the voice data is encoded and transmitted over IP. A VoIP packet is transmitted every 10-30 ms. On the receiving end, the packet is first stored in a jitter buffer to allocate late packets. From jitter buffer the packet goes to output buffer, from where it is taken for encoding. The architecture of VoIP data transmission is shown in figure1.

Fig1: VoIP Architecture of two end-point communication.



The quality of VoIP systems vary widely because in principle, the systems are not standardized as much as the conventional circuit-switched telephone system. Voice traffic uses the same network which was originally designed for data traffic. There are systems that are merely digitized versions of the conventional telephones with bandwidth approximately 50 to 3400 Hz. and there are systems that handle wide band speech with bandwidth up to 7 KHz and higher (Huang et al., 2008).

Many factors affect quality of VoIP. Some of the parameters are latency, jitter, jitter buffer, packet loss, bandwidth, choice of speech codec etc. Since it is a real-time communication, VoIP is sensitive to the delay due to the above factors and retransmission of lost packets cannot be done. VoIP call can tolerate up to 150ms end-to end delay in a single direction. The value beyond this affects the quality of service. End-to-end delay consists of processing delay, network delay and buffering delay.

This paper focuses on voice quality improvement by reducing jitter by adaptive playout of voice packets and thereby reducing packet loss. Jitter is the variation of packet inter-arrival time. In packet –based networks, the sender is transmitting voice packets at regular intervals and these voice packets can be delayed and not arrive at the same regular interval at the receiving side. The difference between the expected time of packet arrival and actual time is called jitter. In order to reduce jitter, a jitter buffer is in used at receiver side to store the early received packets. While waiting for the late packets, the early received packets are waiting in the buffer for playing it out. In order to reduce the buffer delay and packet loss due to discarding of late packets, calculation of additional delay is to be incorporated. This task is very difficult because of the imprecise nature of transmission impairments in voice communications. In this research we have used ANFIS for playout modelling of packets in VoIP communications. An ANFIS system is generated to model the above said system. ANFIS is a hybrid model of artificial neural network and fuzzy logic system. The primary benefit of fuzzy logic is its ability to identify non-linear relationships between the input and output. Artificial neural networks are capable to learn by changing the neuron interconnection weights and also hence improve the performance. A hybrid approach is derived by combining these two techniques to exploit the advantage of both techniques (Roger Jang, 1993). The gain of using neural network in ANFIS is that fuzzy decision-making is achieved by training the neural network algorithm using membership function parameter values. ANFIS can construct a set of fuzzy if-then rules to generate proper membership functions. Consequently, ANFIS can give better performance than the application of fuzzy logic alone.

This paper is organized as follows: Section 2 discusses literature review and section 3 describes ANFIS model. The modelling and implementation of voice packet delay prediction method is described in section 4 and result analysis is discussed in section 5. Finally the conclusion of proposed model is given in section 6.

2. LITERATURE REVIEW

Different algorithms are proposed to minimise the trade-off between buffer size and late waiting (Kim et al., 2013), (H.-G. Kim & Lee, 2012), (Annual IEEE Computer Conference et al., n.d.). An adaptive playout algorithm with delay spike detection is described in (Shallwani & Kabal, n.d.). A buffer re-synchronization algorithm based on silence skipping to prevent unacceptable increase in the buffer preloading delay and even buffer overflow is presented in. A classifier is developed that learns to select the best equalization algorithm using learning examples derived from subjective tests under limited network and conversational conditions is proposed in (Huang et al., 2008). A game theoretic approach for adjusting the jitter buffer size for non-interactive communication and hence to reduce packet loss is proposed by authors in (Chandran & Lingam, 2018a). The authors have studied the factors affecting the quality of communication using a testbed in (Chandran & Lingam, 2015 ICIP.). A statistical approach was proposed in (Chandran & Lingam, 2018b) to minimise the packet loss in real time communication.

Machine learning techniques have been widely used in a variety of applications like medical filed for prediction (Mohan et al., 2019). (Khan & Algarni, 2020) proposed heart disease prediction model and (Yadollahpour et al., 2018) proposed kidney disease prediction model using ANFIS technique. A large collection of data set is to be evaluated and then the evaluation value is to be assigned to each variant in building applications. In or der to simplify this task, the authors have applied ANFIS to prepare the expert opinions of different design variants by grouping it into four main categories (Szafranko et al., 2022). ANFIS is also used for prediction of food production (Nosratabadi et al., 2021). (Al-qaness et al., 2021) studied air quality and used ANFIS technology for forecasting air quality. The authors have also proposed an optimization method using Particle swam optimization and Slime mould algorithm to improve the results of ANFIS. Hybrid machine learning techniques are also applied in agricultural domains to

3. ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM

ANFIS is a multilayer adaptive neural network-based fuzzy inference system proposed by Jang (Roger Jang, 1993). It is a kind of artificial neural network that is based on T-S fuzzy inference system. By integrating both neural network and fuzzy logic concepts in one model, it has learning capability to adjust the member functions according to the training data provided. I.e., this model has learning capability of neural network and decision taking ability of fuzzy system.

3.1 ANFIS Architecture

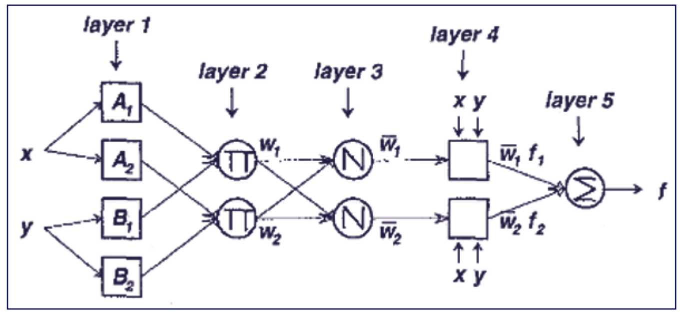
ANFIS architecture consists of five layer neural network. Each layer has different functionalities. One input layer, one output layer and multiple hidden layers. Both input layer and output layers are adaptive in nature. Each layer passes its output the successive layers. In the basic ANFIS architecture shown in fig1, there are two inputs, x and y and one output, f. The membership functions of input x is A1 and A2, and membership functions of input y is B1 and B2. Only layer 1 and layer 4 are adaptable. The two rules formulated for the model is:

Rule1: If X is A_1 and Y is B_1; then $f_1=p_1k_1 + q_1k_2+r_1$

Rule2: If X is A_2 and Y is B_2; then $f_2=p_2k_1 + q_2k_2+r_2$

Where p_1, q1, r_1, q_2 and r_2 are linear parameters, and A_1, B_1, A_2, B_2 are nonlinear parameters.

Fig 2: ANFIS 5 layer architecture (Roger Jang, 1993)



Layer1: In layer 1 is the fuzzy layer in which each node corresponds to the linguistic label, for which different member functions are generated. The crisp values given as the training data set is matched as the corresponding membership functions. The membership function for node i of first layer is given as,

$O_{1,i} = \mu_{Ai}(x_i)$ for i =1, 2 or

$O_{1,i} = \mu_{Bi-2}(x_i)$ for i =3, 4 where A_i and B_{i-2} are the fuzzy sets.

The shape of the membership function is governed by the following equation.

$\mu_{Ai}(x_i) = 1 / (1+[(x - c_i^2) / a_i]^{bi})$

Where a_i, b_i and c_i are the premise parameters. These premise parameters are updated during training process.

Layer2. Each neuron in second layer calculates the firing strength of each rule using the output membership values from layer1. This operation uses min or prod operator. The output of this layer is represented as .

$O_{2,i} = w_i = \mu_{Ai}(x)\ \mu_{Bi} * (\mu_{Ai}(y))$ i = 1, 2

Layer3: This layer is responsible for normalization of firing strength. Each neuron in this third layer finds the normalized firing strength by calculating the sum of firing strength passed from layer 3. The normalized strength is calculated using the following equation:

$O_{3,i} = \mu_{Ai}(x)\ \mu_{Bi} * (\mu_{Ai}(y))$, i = 1, 2

$O_{3,i} = w_i/(w_1+ w_2)$, i = 1, 2

Layer4: The neurons compute a parameter function on the layer 3 output and each node is adaptive in nature. Parameters in this layer are called consequent parameters. This layer produces a defuzzified output to layer 5.

$O_{4,i} = \bar{w}_i \, f_i = \bar{w}_i \, (px + q_i y + r_i)$

Where p_i, q_i and r_i are consequent parameters and are updated during the training process.

Layer5: The last layer is output layer. This consists of single neuron which finds the summation of the outputs of layer 4 neurons.

$O_{5,j} = \Sigma_i \bar{w}_i \, f_i = \Sigma_i w_i f_i / \Sigma_i w_i$

During the learning of neural network, premise and consequent modifiable parameters tune all the membership functions to the given input/output training data. Fuzzy modelling procedure of ANFIS makes the system to learn about the information given in the training data set and then it computes the corresponding membership functions parameters that best suited for the given training data. The membership function parameters are tuned using either back propagation algorithm or a hybrid learning algorithm combining gradient descent back propagation and Mean Least Squares optimization algorithm. Hybrid learning algorithm consists of forward pass and backward pass.

During forward pass, node outputs go from first to fourth layer and the premise parameters are fixed. The consequent parameters are updated with least square estimates. During backward pass the error signals propagate backward to layer1 and premise parameters are updated using gradient descent [6]. This way backward passing updates the premise parameters and helps fuzzy sets to fit to layer1. Updating consequent parameters fine tune the shape of membership function. Since ANFIS is a neuro-fuzzy approach, membership functions of initial FIS constructed is adjusted through iterations by calculating function parameters, so that it best fits the model. This learning process is implemented using the neural network part of ANFIS.

4. MODELLING OF VOICE PACKET DELAY PREDICTION

In this research we have modelled ANFIS to predict the playout time of voice packets in jitter buffer to reduce packet loss due to jitter in voice communication. The first step in modelling using ANFIS is the construction of initial approximate FIS using the training data and then improved through the learning process. The training data should include as many delay calculation parameters for a better result. Using the forwards pass and backward pass of the learning algorithm, the initial FIS learn from the given training data set and adjust its membership function parameters which fine-tune the FIS to track the given input/output data. This makes the FIS to model the given input/output data for the given set of parameters. The membership functions associated with the FIS will adjust through the iterations of the learning process, so that the FIS is tuned to have the best fit one.

In this study, total 100 experimental data were collected from different VoIP conversations and their corresponding arrival time and jitter is extracted from the data (Chandran & Lingam, 2018b). This input/output data set is used by ANFIS to train and test the system purpose for predicting the adaptive delay in jitter buffer to conceal the voice packet loss. Out of 100 data we have collected, we used first 60 data for the training the system and remaining for testing. Membership function parameters are updated and FIS model is constructed using the training data set. The test data is used to determine the termination of training by preventing over-fitting.

We have used two parameters packet arrival time, PA_i and delay of previous packet, D_{i-1} for the implementation of our model. estimates the buffer delay of current packet based on the arrival time of previous packet arrived and the delay identified for that packet in jitter buffer. Once the delay is estimated, we can calculate the time at which the packet is played.

The jitter of i^{th} packet, j_i , is calculated as,

$j_i = dn_i - dn_{i-1} = (tr_i - ts_i) - (tr_{i-1} - ts_{i-1})$

where dn_i, rn_i and ts_i are the network delay, receiving time and sending time respectively. Network delay is the time taken by a packet to travel from sender side to receiver side. It is calculated as,

$dn_i = tr_i - ts_i$

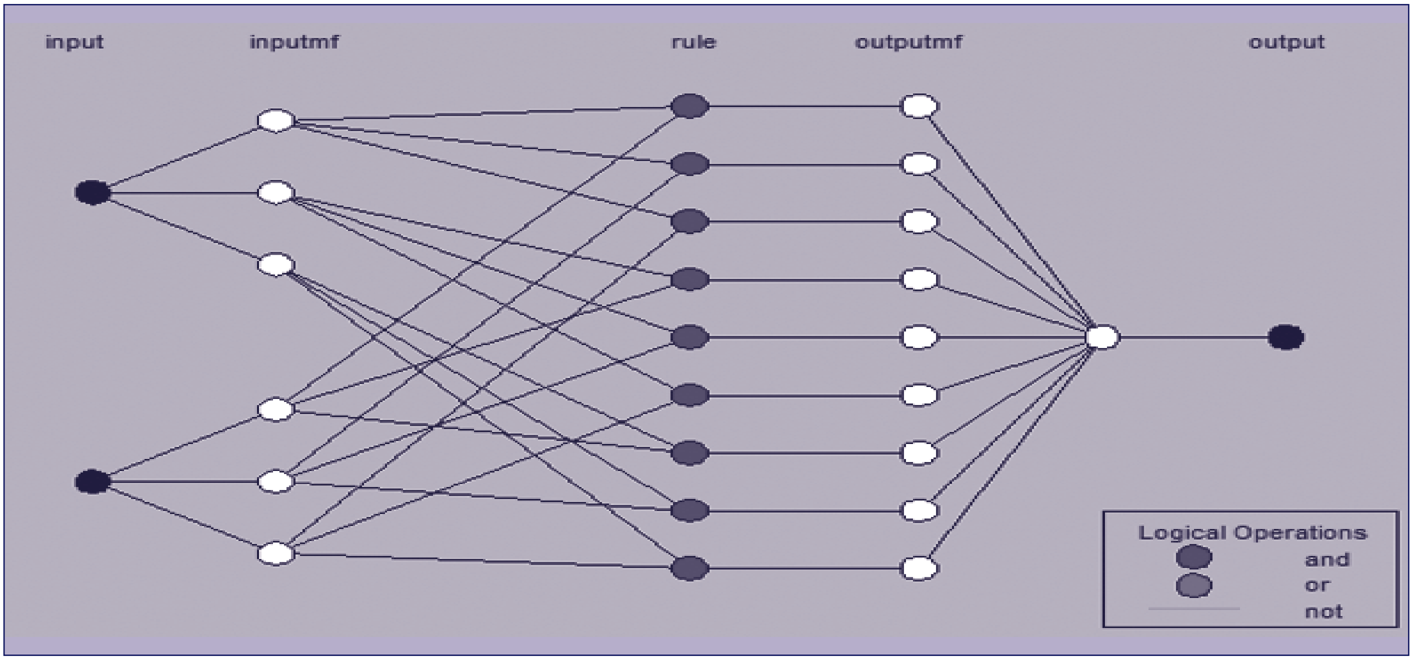
Total delay experienced by the packet, is the sum of buffering delay, dbi and network delay and is given as:

$dt_i = db_i + dn_i$

In our study, ANFIS was implemented by using MATLAB and grid partition was used for generating the fine tuned FIS through 100 epochs. Different parameters are considered for finding the adaptive playout time of a voice packet. One of the main parameters is jitter. The inter arrival time between packets are small in normal delays, but sudden large increase in the packet delay results in spike delay. In order to solve the jitter buffer overflow due to spike delays, the delay can be evenly distributed among all packets within a talk-spurt. For that we have considered the arrival time and delay experienced by the previous packet is considered as the inputs of the system to be modelled and output is the delay to be experienced by the current packet. ANFIS trains the prediction model using the given input data as part of the training data set. In order to get a more accurate model, the training data and test data are evenly distributed across the experimental data set. We have stored both training data set and testing data set in excel worksheet with first two columns as input and last column as output data. Since matlab and excel is interfaced, we can directly read data from excel to matlab using xlsread(). We have used bell shaped member functions are used to construct the FIS and grid partition method is used as the input partitioning method to create rule base. Using genfis1(), we have created the initial FIS structure with three membership functions and each membership functions represents the linguistic labels of the input / output data. After creating initial membership functions, anfis command in matlab is used to train the system and evalfis() is used to study the performance of the trained system.

The ANFIS parameters, premise parameters and consequent parameters, are optimized using hybrid learning algorithm, which is a combination two methods, gradient descent back propagation and Mean Least Squares optimization algorithm. ANFIS model structure after training the system is shown below in figure2. We have implemented the model with three membership functions for each input.

Fig 3: ANFIS model structure of the proposed system



In figure2 the inputs, PA_i , D_{i-1} and are represented in layer1 using two nodes. Each input corresponds to three linguistic labels, small, medium and high and these linguistic labels are represented using three bell shaped member functions in layer2. Each node in this layer calculates the firing strength of each rule using the output membership values obtained from layer1. Each node in layer3 finds the normalized firing strength by calculating the ratios of the rule's firing strength to the sum of all the rules firing strength. Node i in layer 4 compute the contribution of i^{th} rule toward the overall output. Layer5 consists of a single node, which computes the overall output as the summation of signals from layer4. This forward pass in learning process is implemented using least square method and during this forward pass the consequent parameters in layer 4 are computed. If the error is larger than the threshold value, then the backward pass using gradient descent learning algorithm updates premise parameters in layer1. Thus the membership function parameters are tuned using a hybrid learning algorithm during learning process of ANFIS.

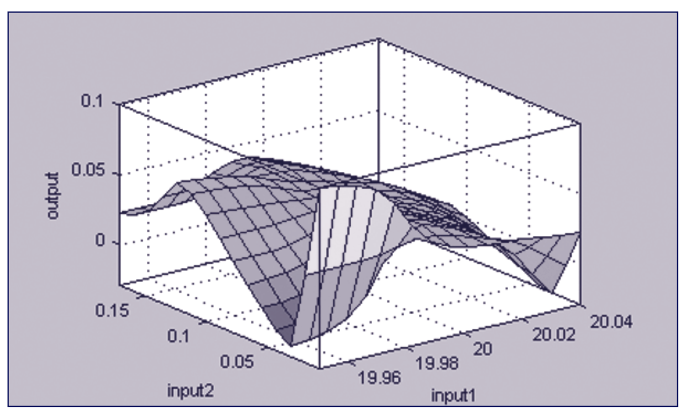
5. RESULTS AND DISCUSSIONS

Root Mean-Squared Error (RMSE) between the ANFIS prediction and system output is calculated. It is assumed that to achieve a lower RMSE during the epochs, the input selection of ANFIS model is done with smallest RMSE after first epoch of training. At each epoch, an error measure between actual and desired output is reduced. Training stops when either the predefined epoch number or error rate is obtained. ANFIS training parameters are depicted in table 1.

Table 1. ANFIS parameters

Sr. No.	ANFIS Parameter Description	Value
1	Number of nodes	35
2	Number of linear parameters	27
3	Number of nonlinear parameters	18
4	Total number of parameters	45
5	Number of fuzzy rules	9

Fig 4: Surface view of ANFIS mode



Rate of error after training the ANFIS model is calculated by finding the difference between evalfis() output and target output. This error rate is shown in figure 4 and the RMSE (Root Mean Square Error) calculated in each epoch is shown in figure 6. Figure 7 shows the error histogram compared to Normal distribution.

From the graph we can see that the actual values and ANFIS predicted values are in close agreement with each other. Thus the ANFIS is trained according to the training inputs. Now we can use this trained model to predict the jitter buffer delay of voice packets for the adaptive playout to conceal the packet loss.

Fig 5: Error Rate

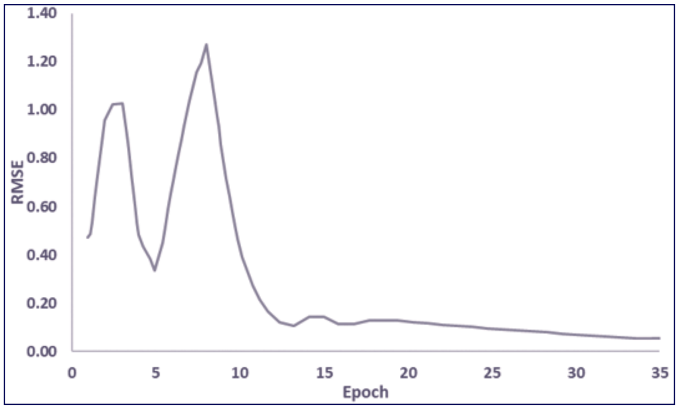
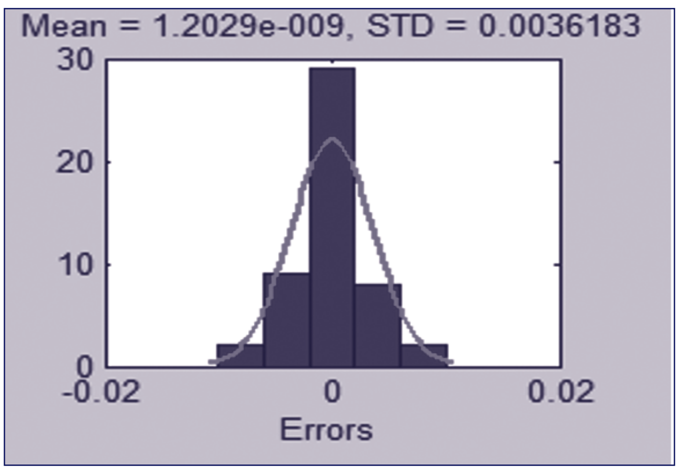


Fig 7: Error histogram compared to Normal distribution



6. CONCLUSION

This study focuses on the development of a machine learning approach to predict the jitter delay time in VoIP communications to avoid packet loss, and thereby improving the quality of communication. Since VoIP is based on IP network and it does not guarantee end-to-end delay, jitter and packet loss. Due to delay in arriving voice packets at the receiver side, the packets are considered as lost or discarded. A jitter buffer is used at the receiver side to temporarily store the packets for some time, there by delaying the playout time of other packets. This additionally imposed playout delay on voice packets by storing in the jitter buffer allows the delayed packets to playout. But this result in jitter buffer overflow and successive packets may be lost. In order to solve these issues adaptive playout mechanism is at the receiver side by considering both buffer size and buffer delay. In this paper, Adaptive Neuro-Fuzzy Inference System (ANFIS) model for training the fuzzy system is proposed to predict the adaptive playout delay time of voice packets in jitter buffer. This attempt to model prediction of jitter buffer delay calculation to reduce packet loss at the receiver side gave very good result. The results show that the ANFIS model is having stronger capability to predict jitter delay than the state of the art techniques.

ACKNOWLEDGEMENT

This study was funded by University of Mumbai, via Minor Research Grant Scheme (Project Number: 111, Ref. No. APD/237/601 of 2019).

REFERENCES

1. *Al-qaness, M. A. A., Fan, H., Ewees, A. A., Yousri, D., & Abd Elaziz, M. (2021). Improved ANFIS model for forecasting Wuhan City Air Quality and analysis COVID-19 lockdown impacts on air quality. Environmental Research, 194. https://doi.org/10.1016/j.envres.2020.110607*
2. *Annual IEEE Computer Conference, IEEE International Symposium on Circuits and Systems 2011.05.15-18 Rio de Janeiro, & ISCAS 2011.05.15-18 Rio de Janeiro. (n.d.). IEEE International Symposium on Circuits and Systems (ISCAS), 2011 15-18 May 2011, Rio de Janeiro, Brazil.*
3. *Chandran, P., & Lingam, C. (2018a). Adaptive jitter buffer management: a game theoretic approach. In Int. J. Communication Networks and Distributed Systems (Vol. 21, Issue 1).*
4. *Chandran, P., & Lingam, C. (2018b). A statistical approach to adaptive playout scheduling in voice over internet protocol communication. International Journal of Electrical and Computer Engineering, 8(5), 2926–2933.*
5. *Chandran, P., & Lingam, C. (2015, December). Performance evaluation of voice transmission in Wi-Fi networks using R-factor. In 2015 International Conference on Information Processing (ICIP) (pp. 481-484). IEEE*
6. *Huang, Z., Sat, B., & Wah, B. W. (2008). Automated learning of playout scheduling algorithms for improving perceptual conversational quality in multi-party VoIP. 2008 IEEE International Conference on Multimedia and Expo, ICME 2008 - Proceedings, 493–496.*
7. *Khan, M. A., & Algarni, F. (2020). A Healthcare Monitoring System for the Diagnosis of Heart Disease in the IoMT Cloud Environment Using MSSO-ANFIS. IEEE Access, 8, 122259–122269.*
8. *Kim, B. H., Kim, H.-G., Jeong, J., & Kim, J. Y. (2013). VoIP Receiver-Based Adaptive Playout Scheduling and Packet Loss Concealment Technique. In IEEE Transactions on Consumer Electronics (Vol. 59, Issue 1).*
9. *Kim, H.-G., & Lee, J.-H. (2012). Enhancing VoIP Speech Quality Using Combined Playout Control and Signal Reconstruction. In IEEE Transactions on Consumer Electronics (Vol. 58, Issue 2).*
10. *Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. IEEE Access, 7, 81542–81554.*
11. *Nosratabadi, S., Ardabili, S., Lakner, Z., Mako, C., &*

Mosavi, A. (2021). Prediction of food production using machine learning algorithms of multilayer perceptron and anfis. Agriculture (Switzerland), 11(5).

12. *Roger Jang, J.-S. (1993). ANFIS : Adaptive-Ne twork-Based Fuzzy Inference System (Vol. 23, Issue 3).*
13. *Shallwani, A., & Kabal, P. (n.d.). An Adaptive Playout Algorithm with Delay Spike Detection for Real-Time VOIP.*
14. *Szafranko, E., Srokosz, P. E., Jurczak, M., & Śmieja, M. (2022). Application of ANFIS in the preparation of expert opinions and evaluation of building design variants in the context of processing large amounts of data. In Automation in Construction (Vol. 133). Elsevier B.V.*
15. *Yadollahpour, A., Nourozi, J., Mirbagheri, S. A., Simancas-Acevedo, E., & Trejo-Macotela, F. R. (2018). Designing and implementing an ANFIS based medical decision support system to predict chronic kidney disease progression. Frontiers in Physiology.*

AUTHORS

Dr. Priya Chandran, Assistant Professor, Bharati Vidyapeeth's Institute of Management & Information Technology, Sector 8, CBD Belapur, Navi Mumbai – 400614, (MS)

Dr. Suhasini Vijaykumar, Professor, Bharati Vidyapeeth's Institute of Management & Information Technology, Sector 8, CBD Belapur, Navi Mumbai – 400614, (MS)

DISCLAIMER

(Forgery of Documents)

It has come to my notice that some agencies or individuals are publishing papers online under the title of "Industrial Engineering Journal" in a Clandestine way and are also issuing certificates of publication of papers with the forged signatures and stamp of the undersigned. This is to bring to the notice of all concerned that the Indian Institution of Industrial Engineering is the sole proprietor of IE Journal and does not issue/offer any certificate to authors for publishing his/her paper in the Industrial Engineering Journal. Also, Industrial Engineering Journal does not publish online papers. All are advised to be cautious of this unscrupulous practice. The IIIE National Headquarters or the undersigned do not take any responsibility for the fake publication and issue of the certificate.

Sd/-
Editor-in-chief
Industrial Engineering Journal
IIIE National Headquarters
Navi Mumbai

WARNING! WARNING! WARNING!

IMPORTANT CAUTIONARY NOTE

Dear Readers / Authors,

It has come to our notice that a website https://ivyscientific.org/index.php/journal has been operating with mala-fide intention by cloning IIIE's Industrial Engineering Journal without our knowledge. This site https://ivyscientific.org/index.php/journal is a clandestine/unauthorized site and IIIE's name is being misused for publishing IE Journal online with an ulterior motive. Appropriate action has been initiated to deal with such unscrupulous activity.

It may be noted that IIIE publishes IE Journal on behalf of INDIAN INSTITUTION OF INDUSTRIAL ENGINEERING (IIIE), NATIONAL HEADQUARTERS (NHQ), SECTOR 15, PLOT NO.103, CBD BELAPUR, NAVI MUMBAI – 400 614 and it is a monthly journal published only in hard copy form.

All are hereby cautioned not to fall prey to the above site and make any payments (Rs. 4000/- per paper) or whatsoever for publishing the paper online. IIIE NHQ shall not be responsible in any capacity for anyone making payments and falling prey to the above unscrupulous site.

All prospective authors are advised to kindly send their Manuscripts only to IIIE journal Email id: journal4iiie@gmail.com or call us at 022-27579412 / 27563837.

Sd/-
Chairman
National Council
IIIE National Headquarters