1. Adil BAKRI[1], 2. Ahmed KROBBA [2], 3. Abderrahmane Amrouche[2],

Scientific and Technical Research Centre on Arid Regions CRSTRA, Algeria (1),
Speech Communication and Signal Processing Laboratory, LCPTS Faculty of Electronics and Computer Science, USTHB. Algeria (2)
ORCID: 1. 0000-0002-2146-650X

doi:10.15199/48.2024.03.22

A New Regression Model Applied for Packet Loss Recovery Technique in Real-Time Communications

***Abstract**. In the present paper, a new regression model adapted to the G.711-PLC was proposed to avoid deterioration of the synthetic signal caused by the use of the attenuated part of the scale-speech. The method of regression consists of discovering a polynomial of a set of data points to estimate values between points. This method is very useful in applied mathematics for approximations and estimations. The main objective is to investigate methods for enhancing the performance of the G.711-PLC, which is the most commonly used technique for speech transmission in real time. This technique is less efficient due to the fact that the precision of attenuation is insufficient because of the variability of speech signal. For this reason, we implement regression polynomial algorithms into attenuation block. The regression polynomial is performed between the amplitudes and the positions of the peaks / valleys to provide a polynomial model of the signal on the interval in the last good packet before the loss. The results revealed that the proposed algorithm has better performance for the speech quality.*

***Streszczenie.** W artykule zaproponowano nowy model regresji dostosowany do sterownika G.711-PLC, aby uniknąć pogorszenia sygnału syntetycznego spowodowanego użyciem tłumionej części mowy skali. Metoda regresji polega na odkryciu wielomianu zbioru punktów danych w celu oszacowania wartości pomiędzy punktami. Metoda ta jest bardzo przydatna w matematyce stosowanej do przybliżeń i szacunków. Głównym celem jest zbadanie metod zwiększania wydajności sterownika G.711-PLC, który jest najczęściej stosowaną techniką transmisji mowy w czasie rzeczywistym. Technika ta jest mniej efektywna ze względu na niedostateczną precyzję tłumienia ze względu na zmienność sygnału mowy. Z tego powodu implementujemy algorytmy wielomianu regresji do bloku tłumienia. Wielomian regresji wykonuje się pomiędzy amplitudami i pozycjami szczytów/dolin, aby zapewnić wielomianowy model sygnału w przedziale ostatniego dobrego pakietu przed utratą. Wyniki wykazały, że proponowany algorytm ma lepszą wydajność w zakresie jakości mowy. (**Nowy model regresji zastosowany do techniki odzyskiwania utraty pakietów w komunikacji w czasie rzeczywistym**)*

Keywords: PLC, Packet Loss, Regression Polynomial, PESQ.
Słowa kluczowe: PLC, Utrata pakietów, Wielomian regresji, PESQ.

Introduction

Voice over IP (VoIP, Voice over Internet Protocol), represents a recent technology, which is rapidly establishing itself in the field of voice communication. It also makes it possible to introduce new communication methods using multimedia interfaces [1,2]. In VoIP systems, real-time speech transmission primarily uses UDP (User Datagram Protocol) as a transmission protocol instead of TCP (Transmission Control Protocol) [3,4]. In UDP, the speech coded by a Codec is then put in the form of data packets (of 10 ms to 40 ms), and then sent from the source to the destination via the Internet [5]. During packet transmission over IP networks, packet loss can occur by exceeding buffers timeout limit in routers or in end systems (receiver) [6]. The distribution of losses is a very important metric for evaluating protocols dealing with the congestion problem (such as; TCP) [7]. A network link can be characterized by its error rate, which is calculated over the transmission time interval [8]. It corresponds to the number of erroneous packets divided by the total number of packets sent. Because of the timeout limit, the retransmission protocols are then not well-suited for resolving this problem [9]. Also, packets that arrive late due to delay are discarded at the receiver and are therefore lost [10]. The packet loss causes a choppy and sometimes unintelligible speech, and thus the quality of speech at the receiving end of a voice transmission system is degraded [11,12]. Since the voice transmission is a real-time process, the receiver cannot request retransmission of the missing packets [13, 14]. To mitigate the effect of these losses on the voice quality, the masking mechanisms of packet loss or PLCs are introduced at the level of the transmitter and the receiver or at the level of the receiver only. This paper proposes a regression model algorithm for G.711-PLC to improve its performance when using real-time in VoIP. In order to avoid deterioration of the synthetic signal caused by the use of scale-speech part in PLC [15], we have developed the algorithm of PLC using the regression polynomial method through we have changed a part of this technique by modifying the scale speech of attenuation in PLC. When few consecutive frames are lost, repeating a single pitch period can result in noticeable harmonic artifacts. In this case, a few previous pitch periods are used to synthesize the replacement signal. This increases the variation in the signal, which is more consistent with real speech. It is necessary, with consecutive packet losses, to attenuate the signal as it progresses because the synthesized signal is more likely to diverge from the real signal. Without attenuation, artifacts are created by holding certain types of sounds for a long-term period. To ensure the subject herein, there is a smooth transition between the synthesized signal and the real speech signal. Synthesis is continued beyond the end and beyond the beginning (by keeping a short delay) of the lost packets and is mixed with the real signal. The proposed algorithm uses the peaks and valleys of pitch segment to generate the best matched signal for the lost segment. Thus, the synthetic signal is increased or attenuated linearly by the polynomial model at the beginning of the 10 ms (80 samples) deletion and, accordingly, the polynomial model is considered as a useful tool for representing a set of data in a linear fashion. To demonstrate the merits of this algorithm, we apply it to attention of PLC and compare it to this technique and with others techniques (WOSLA [31], HMM [32]). Therefore, we conducted tests of objective perceptual evaluation of speech quality (PESQ). The following are the key contributions that we make:

- Identifying the effect of packet loss and overview of several PLC techniques for speech transmission in over IP networks;
- Implementing the PLC to overcome the effects of packet loss and improve the speech quality in over IP networks;
- Proposing a modification on the attenuation block of PLC by replacing the regression polynomial; and

- Providing an experimental evaluation with the proposed algorithms to improve the PLC performance. Then, the differences between the analytical and experimental results are discussed.

This paper is organized as follows; after a brief introduction, in Section 2, we will review some existing techniques that are also used for PLC with different models. We present in section 3 the components of the system blocks of PLC technique which have a role in enhancing the speech quality in VoIP. The proposed method is illustrated in section 4 and we also describe the proposed method integrated into the PLC technique. Performance results obtained in section 5, and section 6 presents the main conclusion of this paper.

Related works

There are several transmitter-based PLCs, these techniques are generally more efficient but more complicated [16]. In this context, we can distinguish two categories of techniques. Regarding the first category, we can cite Forward Error Correction (FEC) [17]. This technique adds inter-packet redundancy at the transmitter level. Then the receiver uses the redundancy information to reconstruct the lost packets. The FEC technique imposes an additional delay, an increase in bandwidth and an implementation difficulty at the decoder level [18,19]. The second category is interleaving, this technique disperses small units of a packet to distribute bursts of packet loss into a series of small losses, and then put back in their original order at the receiver. Interleaving helps disperse the effects of packet loss on the stream, to make them less apparent [20].

Receiver-based PLC techniques replace lost packets with a signal that comes as close as possible to the missing signal [21,22]. The lost signal is often reconstructed based on the information of the last packets received. This is the important because, in general, the statistical properties of the speech signal evolve relatively slowly from one packet to the next [23,24]. Generally, there are two categories of receiver-based PLC. The first category of insertion consists in replacing the lost packets either by noise, or by the repetition of the last packet received [25]. These methods are simple to implement, but when the packet length increases, the quality of the synthesized signal degrades rapidly [26]. The second category, based on interpolation, consists in interpolating certain parameters of past and future good packets in order to create a substitute for the lost packet [27]. The time-scale modification (TSM) technique consists of modifying the time scale (changing the duration of the speech signal without changing the period of the fundamental "pitch" of the speech signal based on the waveform-similarity-based synchronized overlap-add method [28]. Timescale modification refers to the processing of timescale dilation of a speech signal in order to supplement lost packets. The Interpolation by the technique of G.711-PLC consists in producing a synthetic voice signal intended to replace the missing data. The G.711-PLC is certainly the most used for speech transmission in VoIP [29]. This technique estimates the data that has been lost and generates synthetic speech to cover the missing data. The synthetic frame is created by finding the last pitch period in the previous packet and playing this pitch period over and over until a whole frame is created. According to the PLC algorithm, up to 3 previous pitch periods are used for generating the synthetic signal. Smooth transition is accomplished by mixing 0.25 pitch period beyond the end and 0.25 pitch period beyond the beginning of the lost packets with the real signal using an overlap-add (OLA) technique. After 20 ms, the synthesized signal is linearly attenuated with a ramp at the rate of 20% per 10 ms, so after 60 ms the synthesized signal is centred on zero.

In the context of enhancing PLC proposed in [30], implemented a combination of linear prediction LP model in G.711-PLC. This approach depends on the well-known linear prediction model in estimating the missing speech waveform. This proposed algorithm combines a linear prediction model and reverse-order replicated pitch period technique, and it is designed to work with the conventional sampling rate of 8 kHz and frame sizes of 10 ms. The PESQ scores obtained for random loss tests show that the algorithm performs better than existing method .The algorithm has a much higher complexity compared to the G.711-PLC. This complexity is approximately 5 times higher than the complexity of G.711-PLC.

A PLC approach is proposed in [31], based on the combination of Waveform Similarity Overlap-Add (WSOLA) and Deep Neural Network (DNN). In this method, the log-power spectrum of the lost packet is estimated by DNN, and the speech signal without packet loss is extended by the WSOLA method until the lost packet is covered. The methods effectively improve the performance of PLC in continuous packet loss scenarios. Experimental results demonstrate that the proposed PLC methods provide better speech quality compared to reference methods, however this PLC technique is not suitable for real-time processing of packet-loss as an example, the models should not use bidirectional contexts to make predictions, because the backward context in real-time applications will not be available at the time of prediction using such models will make the predictions more costly, which is typically not present in real-time processing.

In [32], the author's proposed hidden Markova model (HMM) on the G.722.2-PLC, each packet is analyzed to give the observation parameter vector. As soon as a sequence of packets is missing, an estimate of the missing parameters vectors is produced from an acoustic HMM with probability density functions (pdfs). HMM-based PLC methods have shown improvement over traditional methods in reconstructing missing speech packets in Voice over IP (VoIP) communications. However, these methods may result in perceptually disturbing artifacts due to the structure of the HMM model and it was not an omnipotent one, which underlines that further enhancements are still needed to improve the real-time quality speech performance.

PLC algorithm overview

The PLC relies on the direct insertion of a waveform into the gap. The objective of this technique is to generate a synthetic speech signal to cover missing data in a received bit stream as shown in Fig 1. Ideally, the synthesized signal will have the same timbre and spectral characteristics as the missing signal, and will not create unnatural artifacts. Since speech signals are often locally stationary, it is possible to use the signal's past history to generate a reasonable approximation to the missing segment [29]. When a packet is received, a copy of the decoded signal is recorded in a 48.75 ms (390 samples) buffer to calculate the pitch. The result (output signal) is delayed by 3.75ms (30 samples) to perform an overlap and add OLA at the end of the loss episode and thus ensure a cross fade between the speech received and the speech from concealment [33].

A. Overlap-Add (OLA) method

A method provides smoothing between real signals and synthetic signals. The OLA operation replaces the terminal part of the pitch buffer. The weighting window and location is done by meanst of a triangular window at a quarterwavelength at the fundamental period (Cf. Fig 2). The OLA window found at the end and beginning of the synthetic frame of the expanded error frame in the previous and subsequent correct frames. The length of the latter depends both on; the fundamental period and the length of the erasure

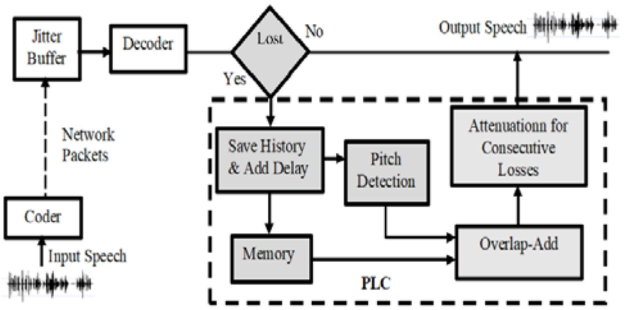

Fig.1. Flowchart of packet loss concealment in in real-time communications

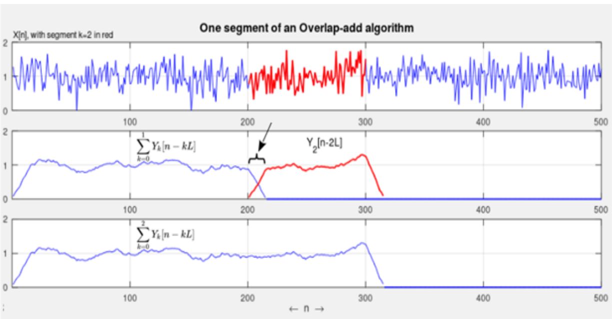

Fig. 2. The Overlap Add operation (OLA).

B. Computations for the first lost frame

During the first lost frame, the pitch period of the history buffer is estimated by finding the peak using an autocorrelation. The synthetic signal is then generated based on the most recent 1.25 pitch periods of the pitch buffer that are used during the first 10 ms . These periods are then repeated with a window operation to synthesize the first 10 ms of the missing signal.

C. Synthetic signal generation after the first lost frame

If the erasure is 20 ms long, the number of pitch periods used to synthesize the speech is increased to two, and if the erasure is 30 ms long, the third pitch is added and the synthesized signal is attenuated by 20%. Beyond 30 ms of erasure, no changes are made to the history buffer. The number of pitch periods used to synthesize the speech is the third pitch periods. Besides, an Overlap Add (OLA) is performed using a triangular window on one quarter of the pitch period between the last and the next to last period.

D. Attenuation

The attenuation by scale-speech is occurred as the erasure continues. The signal is not attenuated during the first 10 milliseconds, it will linearly attenuate at a rate of 20% per every 10ms until the synthetic signal disappears after 60ms.

E. First good frame after good frame

At the first good frame (10 ms) after an erasure, a smooth transition is needed between the synthesized erasure speech and the real signal. To do this, the synthesized speech from the pitch buffer is continued beyond the end of the erasure, and then mixed with the real signal using an OLA.

Proposed regression model for PLC

In the G.711-PLC, it turns out that after the start of the second lost packet (two consecutive packets), the synthetic signal is attenuated with a linear ramp [29]. This attenuation occurs at the rate of 20% per 10 ms (packet) and begins at the start of the second lost packet, so the precision of the attenuation is insufficient due to the variable of the speech signal. For this reason, we incorporate the regression model algorithm into the PLC, in order to improve the efficiency of this technique in the face of packet loss; Figure 3 shows the modified attenuation scheme in the PLC blocks. The proposed algorithm uses only the peaks and valleys of the OLA output block to generate the best approximation of the synthetic signal such as the original signal, using the regression polynomial method which associated with the blocks of PLC.

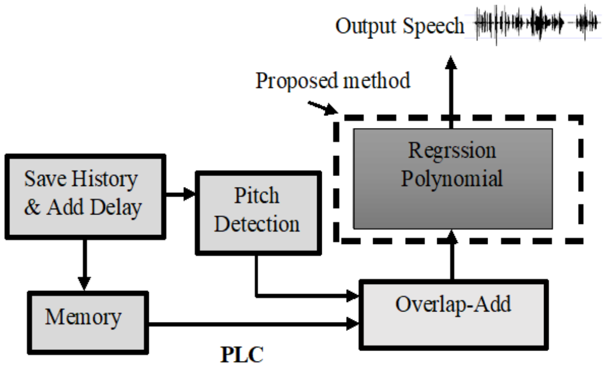

Fig. 3. The regression model in PLC.

For illustration purposes, we suppose that each frame (packet) contains *N*-samples and an *i*-frame number. Also, we assume that it includes *n*-position and *e(n)*-amplitude. Hence, we have already represented a speech signal frame by an expression of the following as:

$$F(i) = (e(1), e(2).......e(n).......e(N)) \quad (1)$$

Fig. 4 below presents the overall architecture of the proposed PLC system based on regression model

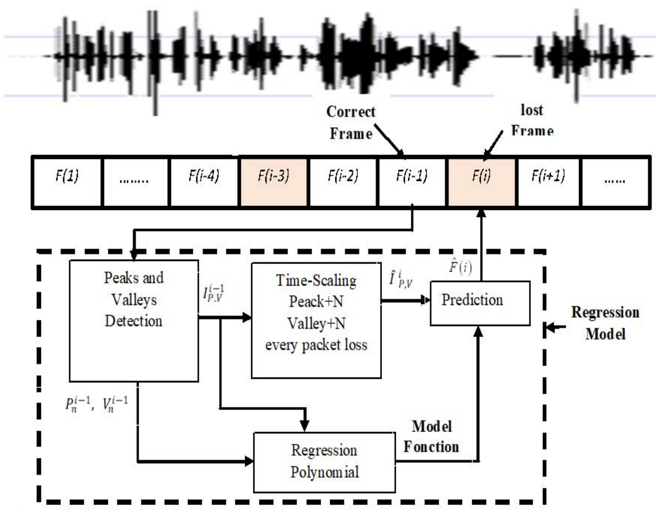

Fig.4. Blocks of the proposed method.

A. Peaks/Valleys detection

The peaks/valleys decision method compares each packet element with its neighboring values. If a packet element is larger than its two both neighbors, it is a peak and if a packet element is smaller than its two both neighbors, it is a valley. The positions of the peaks/valleys, give the locations as indices for all the peaks/valleys detected in this block. The inputs of the peaks/valleys elements in the regression polynomial block would be extracted from the correct packet data that precedes the lost packet. The regression polynomial therefore is performed between the amplitudes and the positions of peaks and valleys, which in point of fact, used to provide a polynomial model of the signal over the packet interval [34,35].

In order to reconstruct the lost packet in the future, should be made when the speech packet is received, we assume that the signal frame is lost The correct frame which precedes the lost frame is used to estimate of the peaks vector and the valleys vector. Position is peack if only e(n) > e(n-1) and e(n) > e(n+1) so Position is valley if only e(n) < e(n-1) and e(n) < e(n+1). The vectors of the peaks/valleys and their positions are transmitted to the regression polynomial and the Time-scaling blocks. Here, Time-scaling is the number of the consecutive frames in an erasure, increased by N samples at the positions of peaks and valleys for every 10 ms (frame) while the erasure persists. In effect, it is reset to zero at the first, correct frame that follows an erasure. In the simplest case, increasing these positions would also result in a corresponding increase or attenuation of the variation in the signal [36].

B. Regression polynomial

Regression polynomial is used for modelling non-linear relationships between peaks/valleys and their positions by fitting a polynomial model to the data (P^{i-1}_{n}, I^{I-1}_{P}) and (V^{i-1}_{n}, I^{I-1}_{V}). Unlike linear regression, which models the relationship as a straight line, regression polynomial can capture more complex and nonlinear relationships. The polynomial model gives a closer fit to the speech signal. This will lead to more accurate predictions of new values in test data. In general, polynomial model equation takes the form:

(2) $$h_\theta = \theta_0 + \theta_1 x + ... + \theta_j x^j$$

In regression polynomial, the goal is to find the best-fitting polynomial model to the data by estimating the coefficients $\theta_0, \theta_1 ... \theta_j$. Where: h_θ is the dependent variable (the variable we are trying to predict); *x* is the independent variable (the input or predictor variable); *j* is the degree of the polynomial model, which determines the complexity of the model. The steps of the regression polynomial algorithm:

- Model Fitting: Use the polynomial model equation and techniques minimize error sum of squares method. This method works by creating a best fit line through all of the available data points to estimate the coefficients that minimize the error between the predicted values and the actual values [37, 38].
- Model Evaluation: Evaluate the performance of polynomial model using metric like *MSE* (mean squared error). This metric is a measure of how well the regression fits the data. It calculates average squared difference between the observed and predicted values.
- Prediction: Once the model is trained and evaluated, it predict future outcomes based on historical data.

To find the optimal coefficients, we need an objective function to minimize, which quantifies the error between the predicted values and the actual values in the dataset, given by cost Function:

For the regression polynomial of peaks:

(3) $$MSE = J_P(\theta_j) = \frac{1}{2m_p}\sum_{k=1}^{m_p}(h_\theta(I_{Pk}{}^{i-1}) - P_k{}^{i-1})^2$$

For the regression polynomial of valleys:

(4) $$MSE = J_V(\theta_j) = \frac{1}{2m_v}\sum_{i=1}^{m_v}(h_\theta(I_{Vk}{}^{i-1}) - V_k{}^{i-1})^2$$

- Where: m_p , m_v umber of peaks and valleys in a correct frame *F(i-1)* respectively . Now, using gradient descent, we can iteratively update the coefficients to minimize the *MSE*. The gradient descent algorithm works as follows:
- Initialize the coefficients $\theta_0, \theta_1 ... \theta_j$ with some random values.
- Calculate the predicted values h_θ for each positions points using the current values coefficients with the polynomial equation.
- Calculate the gradient of the *MSE* with respect to each coefficient. The gradient is a vector that indicates the direction of the steepest increase in the *MSE*. The gradient for each coefficient can be computed as follows:

(5) $$\frac{\partial J_P(\theta_j)}{\partial \theta_j} = \frac{1}{m_p}\sum_{i=1}^{m_p}(h_\theta(I_{Pk}{}^{i-1}) - P_k{}^{i-1})^2 I_{Pk}{}^{i-1}$$

(6) $$\frac{\partial J_V(\theta_j)}{\partial \theta_j} = \frac{1}{m}\sum_{i=1}^{m_v}(h_\theta(I_{Vk}{}^{i-1}) - V_k{}^{i-1})^2 I_{Vk}{}^{i-1}$$

- Update the coefficients using the gradient and a learning rate (α). The learning rate controls the step size in each iteration. Repeat until convergence for the polynomial of peaks and valleys:

(7) $$\theta_j = \theta_j - \alpha\frac{\partial J_{P,V}\left(\theta_j\right)}{\partial \theta_j}$$

- Repeat steps until the MSE converges to a minimum or a predefined number of iterations is reached.

The learning rate (α) is an important hyperparameter for gradient descent. Even though it doesn't directly affect model like coefficients do but it can impact the performance of our model and we need to be cautious about choosing model hyperparameter. Finding an appropriate learning rate is crucial for the success of gradient descent. Gradient descent will iteratively adjust the coefficients to minimize the *MSE*, resulting in the best-fitting polynomial function that captures the relationship between the peaks/valleys and their positions.

C. Time-Scale for polynomial model

Time-Scaling is the benchmark for the number of consecutive lost frames in an erasure, increased by 80 samples at the positions of the peaks and valleys every 10 ms that lasts an erasure also called resampling, it is reset to 0 at the first correct frame following an erasure. An increase in these positions increases or decreases the variation in the signal. At the beginning of the 10 ms (80 samples) deletion, the synthetic signal is increased or attenuated, linearly, by the polynomial model therefore considered as a useful tool for representing a set of data in a linear fashion. The peaks / valleys vectors and their positions pass through scale speech for each lost frame; thus, the new positions may be expressed as.

(8) $$\hat{I}_P^{i-1} = I_P^{i-1} + N$$
$$\hat{I}_V^{i-1} = I_V^{i-1} + N$$

D. Prediction with polynomial regression

Lost packet synthesis is generated using the estimated regression model, by performing the synthesized peaks/valleys, from the following polynomial model.

(9) $$\hat{p}^i = h_\theta(\hat{I}_P^{i-1})$$
$$\hat{V}^i = h_\theta(\hat{I}_V^{i-1})$$

In packet loss *F(i)*, we insert the new synthesized peaks and valleys directly with the old ones in order to recreate the lost packet, it is formulated by.

(10) $$F(\hat{I}_P^{i-1}) = \hat{P}^i$$
$$F(\hat{I}_V^{i-1}) = \hat{V}^i$$

The others samples of the packet loss *F(i)* are approximated by adding the mean value Δ_p and Δ_p to samples of the packet correct *F(i-1)* of the history buffer, as shown in Fig 5. Where the Δ_p and Δ_p is the result of subtraction between the peaks/valleys and $\hat{P}^i$, $\hat{V}^i$. The synthesized samples is given by

(11) $$\hat{e}(n) = e(n) + \Delta_{p,v}$$

In all cases, the output is constructed by the current packets corrects of the history buffer. In this case, the peaks and valleys are ignored, because these values are already synthesized by the polynomial model as shown in equations (9 and 10)

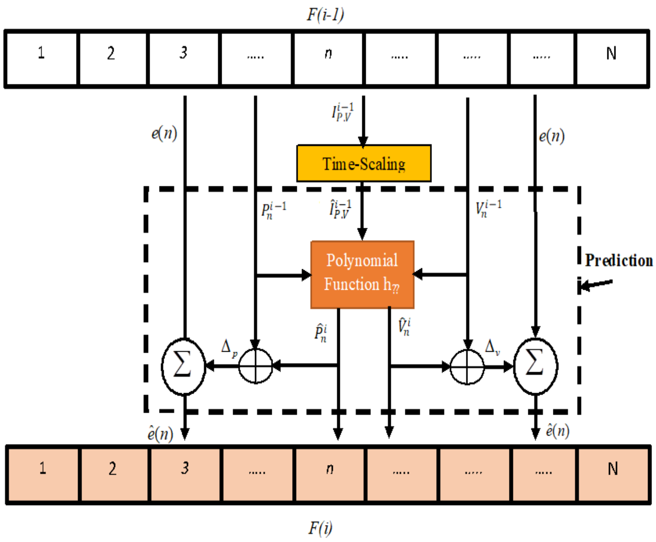

Fig. 5. Prediction with polynomial regression.

Results and discussion

To verify the performance, the proposed regression polynomial algorithm was integrated into PLC technique, we compared with and without a proposed method using different packet loss LR (loss rate). For this, we used 60 audio files and we used the speech signals taken from TIMIT [39], then obtained PESQ (perceptual evaluation of the speech quality) scores according to ITU-T Recommendation P.862 [40], and using the proposed PLC algorithm under different packet loss conditions. The performance was also compared to that with the PLC algorithm, referred to here as PLC based on WSOLA [31] and HMM [32]. In this paper, we simulated two different packet loss conditions (loss, correct), including random packet losses. During these simulations, packet loss rates of 5, and 30% were generated by two state Markov model defined in ITU-T, where Fig. 6 provides the average PESQ score sequences (−0.5 (worst) up to 4.5 (best)) under the different cases, which will be mentioned below.

As can be seen from the experiments and analysis, the quality of the voice signal is affected by various network parameters as packet loss. Initially different packet loss of first case was tested without implementing the PLC technique. From the results, it was observed that when the value of packet loss is low; i.e., LR= 5%, speech quality is degraded (PESQ going down from 4.5 to 2.11), the LR= 20%, speech quality is degraded PESQ (PESQ going down from 4.5 to 1.36). Results analysis revealed that the loss effect on voice communication increases with the value of packet loss that presents in the network.

To mitigate the effect of packet loss in VoIP, PLC technique is implemented in second case. It appears in fact that there is a difference in the PESQ scores. As expected for the first case, and after the use of the PLC technique, the improvements in the average of the speech quality are approximately raised to 0.61; i.e., LR= 10%, speech quality is improved, (PESQ going higher from 1.99 to 2.631), and LR= 30% (PESQ going higher from 0.86 to 1.30). Test results of PLC show that the optimization system can improve the perceived speech quality in IP environment and provides the best speech quality.

To develop the performance of the PLC technique for the speech of the transmission in VoIP, in third case we proposed a new PLC technique by incorporating regression polynomial. Several objective tests were performed in order to test the proposed PLC technique. From the results shown by this case, it can be observed that the PLC technique, when using the polynomial regression, offers better quality (PESQ-Average = 2.20), than the use of linear ramp (attenuation) in the PLC technique (PESQ-Average = 2.12). These results indicate a significant improvement in speech quality by the proposed algorithm, which can improve the technical performance of PLC under different packet loss conditions.

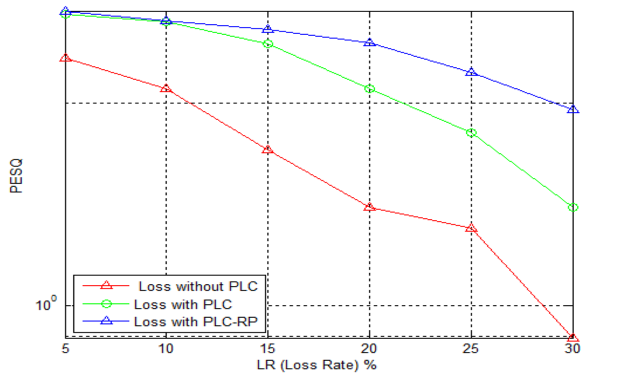

Fig. 6. PESQ results for the proposed algorithm PLC-RP and conventional PLC.

We also see through the Figure 6, the mean PESQ scores for the tested PLC algorithms at various packet loss rates LR, we notice that the PLC-RP (regression polynomial) exceeds the PLC with lower values of PESQ (0.051). In contrast to the previous situation, for when the loss rate is greater than 20%, we notice that the PLC-RP exceeds the PLC with higher values of PESQ (0.44). The results also indicate that the improvement is more pronounced for more packet loss rate LR and that the improvement is more observable when packet loss rate increases.

In the framework of the proposed method test, we compared the performance of the PLC proposed with others techniques (namely: HMM-based PLC and WOSLA-based PLC) under similar conditions. For this comparison, we simulated the packet losses and measured the quality of the output signal with PESQ through the original signal, the signal after losses and the packet reconstruction.

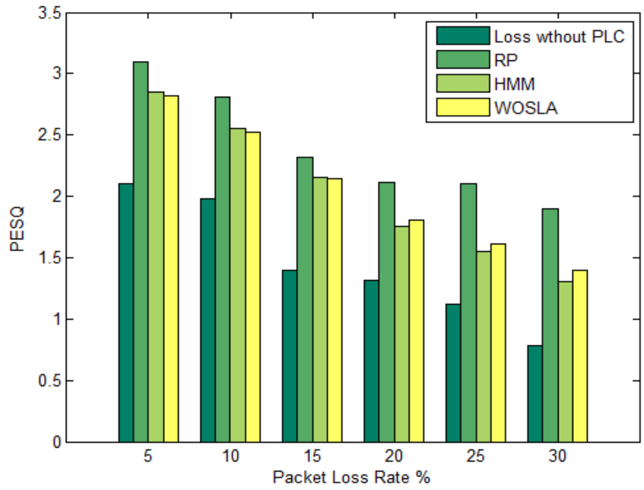

Fig. 7. PESQ scores comparison for PLC-RP and PLCs techniques.

In Fig. 7, the results obtained reveal the PESQ that are related to the three PLC techniques; the HMM-based PLC, the WOSLA-based PLC, and the RP-based PLC technique. Thus, the RP-based PLC technique offers better quality (PESQ = 2.39) than the HMM-based PLC (PESQ = 2.02) and WOSLA-based PLC (PESQ = 2.04). Thus, for our experimental data it would be suitable to use this technique since it allows to reach, in most cases, superior performances in comparison with the HMM-based PLC and the WOSLA-based PLC techniques and accordingly, we can conclude that the RP-based PLC algorithm is more efficient.

Conclusion

The strategy of PLC after the decoder consists in reconstructing the current packet lost on the basis of information already received. This technique makes it possible to replace a missing signal with a similar one. The proposed PLC algorithm combined a G.711-PLC with a method of regression polynomial. In addition, with consecutive packet losses is necessary to attenuate the signal as it progresses because the synthesized signal is more likely to diverge from the real signal. The proposed algorithm in this paper presents a solution to this deterioration by using a regression polynomial model in the components of a PLC so that the synthetic signal changes linearly. This algorithm provides a polynomial model of the peaks and valleys of the voice signal on the correct packet interval that comes before the lost packet. The main advantage compared to previous method is that the function is applied for the purpose of finding the synthetic packet approximate than the lost packet. This increases the variation linearly in the signal, which is more consistent with the real speech. In this study, we subsequently evaluated the performance of the proposed PLC algorithm under random packet loss rates of 5, and 30%, and then compared it with that of G.711-PLC. It was shown from PESQ tests, waveform comparisons, that the proposed method improves the reconstructed speech quality significantly under all the test conditions. Most importantly, the results were analyzed to evaluate, in all cases of packets losses rates, the proposed PLC algorithm that provides the best speech quality in comparison with the two well-known techniques; HMM and WOSLA. Future work can consider using different audio representation such as Time-Frequency spectrograms, or other lower representations of audio as one practical advantage of it is that it can be used to increase the model's accuracy and efficiency.

REFERENCES

[1] *M. Chandra M, and M, Ray, "Comparative study of PCM, LPC, and CELP speech coders used for VoIP applications," Intelligent computing, communication and devices, pp. 579-587, 2015.*

[2] J. Cermak, "Verbal descriptors for VoIP speech sounds," International Journal of Speech Technology, vol. 7, pp 81-91, 2004.

[3] L. Larzon, M. Degermark, et al, "The Lightweight User Datagram Protocol (UDP-Lite), "IETF. RFC 3828, 2004.

[4] H. Schulzrinne, S. Casner, R. Frederick and V. Jacobson," RFC 3550: RTP: A Transport Protocol for Real-Time Applications," Request for Comments, IETF, 2003.

[5] A. Gomez, J. Carmona, A. Peinado and V. Sanchez, "One-pulse fec coding for robust celp-coded speech transmission over erasure channels, "IEEE Transactions on Multimedia, vol. 13, pp.894-904, 2011.

[6] L. Zhongbo, Z. Shenghui and B. Stefan, "Comparison and optimization of packet loss recovery methods based on AMR-WB for VoIP, "Speech Communication. ELSEVIER, vol. 54, pp. 957-974, 2012.

[7] G. Cermak, "Subjective Quality of Speech Over Packet Networks as a Function of Packet Loss," Delay and Delay Variation. International Journal of Speech Technology, pp. 65-84, 2002.

[8] W. Jim and H. Mahbub, "The issue of useless packet transmission for multimedia over the Internet," Computer Communications. ELSEVIER, vol. 26, pp. 1240-1254, 2003.

[9] F. Neves, S. Soares and P Amado, "Optimal voice packet classification for enhanced VoIP over priority-enabled networks," Journal of Communications and Networks, vol. 20, no. 6, pp. 554-564, 2018.

[10] P. Vern, P, "End-to-End Internet Packet Dynamics, " IEEE/ACM Transactions on Networking, vol. 7, no. 3, pp. 277-292, 1999.

[11] L. Roychoudhuri and E.S. Al-Shaer, "Real-Time Packet Loss Prediction Based on End-to-end Delay Variation," IEEE etransactions on Network and Service Management, vol .2, no.1, pp.29-39, 2005.

[12] W. Jiang and H. Schulzrinne, "Modeling of packet loss and delay and their effect on real-time multimedia service quality," In Proceedings of NOSSDAV, 2000.

[13] P. Vijaya and N. Urmila, "Error concealment schemes for speech packet transmission over IP network," Systems Signals and Image Processing, IEEE Xplore, pp. 188-193, 2008.

[14] I. Akinori and N. Takeshi, "Packet loss concealment of VoIP under severe loss conditions," The 15th International Symposium on Wireless Personal Multimedia Communications 2008.

[15] A. Bakri, A. Amrouche and M. Abbas, "An improved packet loss concealment technique for speech transmission in voip," International Conference on Natural Language and Speech Processing, 2018.

[16] H. Sanneck, K. Ben Younes, and B. Girod, "A new technique for audio packet loss Concealment," IEEE Global Telecommunications Conference, pp. 48-52, 1996.

[17] N. Benamirouche, et al, "Low complexity forward error correction for CELP-type speech coding over erasure channel transmission," International Journal of Speech Technology vol. 19, p. 717-730, 2016.

[18] D. Rodríguez D Z, et al, "Impact of FEC codes on speech communication quality using WB E-model algorithm," Wireless Days (WD) Publisher IEEE, 2019.

[19] B.K. Lee and J.H Chang, "Packet loss concealment based on deep neural networks for digital speech," IEEE/ACM Transactions on Audio Speech and Language Processing, vol. 24, no. 2, pp. 378-387, 2016.

[20] F. Mehran, "Interleaver choice in serially concatenated codes for soft real-time wireless speech transmission applications," In IEEE 2011 19th Iranian Conference on Electrical Engineering, 2011.
[21] A. Bakri and A. Amrouche, "Implementing the PLC Techniques with G 729 Coded to Improving the Speech Quality for VoIP Transmission," Intentional Conference Nanoelectron Communication. Renew Energy, 2013.
[22] N. Aoki, "VoIP packet loss concealment based on two-side pitch waveform replication technique using steganography," IEEE Region 10 Conference TENCON, pp. 52-55, 2010.
[23] L. Zhibin, L. Jing and Q. Xiaojun, "An effective hybrid low delay packet loss concealment algorithm for MDCT-based audio codec," Applied Acoustics, vol. 154, pp. 170-175, 2019.
[24] W.T. Liao, et al," Adaptive Recovery Techniques for Real-Time Audio Streams", In IEEE Proceedings IEEE INFOCOM, vol. 2, pp. 815-823, 2001.
[25] S. Leandro and M. Edjair, "Survey on application-layer mechanisms for speech quality adaptation in VoIP," ACM Computing Surveys (CSUR), vol. 45, no. 3, pp. 31-36, 2013.
[26] T. Chua and D. Pheanis, "QoS evaluation of sender-based loss-recovery techniques for VoIP," IEEE Netw, vol. 20, no. 6, pp. 14-22, 2006.
[27] M. Li et al, "Packet Loss Concealment Using Enhanced Waveform Similarity Over Lap-and-Add Technique with Management of Gains," International Conference on Wireless Communications. Networking and Mobile Computing, 2009.
[28] S. Agnihotri, et al, "A new technique for improving quality of speech in voice over IP using time-scale modification," in Proc IEEE Int Conf Acoust, Speech Signal Process, pp. 2085-2088, 2002.
[29] ITU-T G.711, "A High Quality Low-Complexity Algorithm for Packet Loss Concealment with G.711, " 1999.
[30] M. Elsabrouty, M. Bouchard and T. Aboulnasr, "Receiver-based packet loss concealment for pulse code modulation (PCM G.711) coder," Signal Processing, vol. 84, no. 3, pp. 663-667, 2004.
[31] J. Qiang, B. Changchun and C. Zihao, "Packet Loss Concealment Based on Phase Correction and Deep Neural Network", Applied Siences, vol. 12, no. 19, pp.1--17, 2022.
[32] T. Gueham and F. Merazka, "Packet loss concealment method based on hidden Markov model and decision tree for AMR-WB codec", Multimedia Tools and Applications, 2023.
[33] J. Wiley, "VoIP voice and fax signal processing," General & Introductory Electrical & Electronics Engineering published simultaneously in Canada, 2008.
[34] P. Thévenaz and T Blu," Unser M. Unser Interpolation Revisited," IEEE Transactions on Medical Imaging, vol. 19, no. 7, pp. 739-758, 2000.
[35] J. Wang and J.D. Gibson, "Parameter interpolation to enhance frame erasure robustness of CELP coders in packet networks," Proc IEEE Int Conf Acoust Speech Signal Process, pp. 721-724, 2001.
[36] A. Bakri and A. Amrouche, "Packet loss concealment-based estimation of polynomial interpolation for improving speech quality in VoIP," International Journal of Intelligent Systems Technologies and Applications, vol. 19, no. 5, pp. 486-499, 2020.
[37] O. Eva, "Modelling using polynomial regression," Procedia Engineering, vol. 48: pp. 500-506, 2012.
[38] D. Christopher, et al, "Improving the Operationalization of interest congruence using polynomial regression," Journal of vocational behaviour, vol. 104, pp. 154-169, 2018.
[39] J.S. Garofolo," Getting started with the DARPA TIMIT CDROM: an acoustic phonetic continuous speech database," NIST Tech Report, 1988.
[40] ITU-T Recommendation, "Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs," 2007.