Effective Voice Frame Pruning Method to Increase VoIP Call Capacity

Mosleh M. Abualhaj [1], Sumaya N. Al-Khatib [1], Manjur Kolhar [2], Alhamza Munther [3], Yousef Alraba'nah [1]

[1] *Al-Ahliyya Amman University, Faculty of Information Technology, Amman, Jordan*
[2] *Prince Sattam Bin Abdulaziz University, College of Arts and Science, Al-Kharj, Kingdom of Saudi Arabia*
[3] *Ministry of Higher Education, College of Applied Science / Sur Sultanate of Oman*

Abstract **– Voice over Internet Protocol (VoIP) is gradually dominating the telecommunication sector because it is free or inexpensive. However, one of the key problems of VoIP growth is inefficient bandwidth utilization. Several methods have been proposed to improve the VoIP bandwidth utilization, and they include VoIP packet aggregation and header compression. In this study, we investigate a new dimension to improve VoIP bandwidth utilization, that is, VoIP packet payload compression. The main idea of the proposed method, which is called the voice frame pruning (VFP) method, is to prune the leading/trailing zeros/ones of the VoIP packet payload on the basis of a certain mechanism. The VoIP packet payload is pruned at the sender side's wide area network (WAN) gateway and restored to its original form at the receiver side's WAN gateway. The implementation results of the proposed VFP method indicate good bandwidth savings based on the VoIP codec used. For example, the bandwidth savings of the LPC, G.723.1, and G.729 improved by up to approximately 5%, 8%, and 3.5%, respectively, thereby improving the VoIP bandwidth utilization and the capacity of VoIP calls.**

Keywords **– VoIP, VoIP Codecs, VoIP protocols, RTP, Bandwidth utilization.**

DOI: 10.18421/TEM91-08
https://dx.doi.org/10.18421/TEM91-08

Corresponding author: Mosleh M. Abualhaj,
Al-Ahliyya Amman University, Faculty of Information Technology, Amman, Jordan
Email: m.abualhaj@ammanu.edu.jo

Received: 22 November 2019.
Revised: 21 January 2020.
Accepted: 28 January 2020.
Published: 28 February 2020.

1. Introduction

Voice over Internet Protocol (VoIP) is one of the most widely used Internet technologies worldwide [1], [2]. VoIP is popular because of its various features. These features include low call cost (sometimes free), the ability to initiate a voice call through almost any smart device (e.g., mobile phone, iPad, and laptop), interactive voice recognition, and voicemail to email [3], [4]. However, VoIP has some limitations that hinder its propagation. The low quality of VoIP voice calls is a key limitation, especially when compared with the quality of a typical public switched telephone network [5], [6]. Another key limitation is its great waste of network bandwidth resulting from the addition of a big packet header (typically 40 bytes) to a small payload (typically 10 bytes to 30 bytes) [7], [8].

VoIP technology uses several protocols to produce a voice call. These protocols are categorized into signaling and media transfer protocols [9], [10]. Examples of signaling protocols are session initiation protocol and H.323 [3], [10]. Signaling protocols aim to start a call between two call ends. Examples of media transfer protocols include real-time transport protocol (RTP) [11], [12], inter-asterisk exchange [13], and Internet telephony transport protocol [10]. Media transfer protocols aim to carry voice data throughout the network between two call ends. RTP is the preponderant protocol among media transfer protocols. Figure 1. shows the 12-byte RTP protocol format. The proposed method utilizes and reemploys the 4-bit contributing source (CSRC) count (CC) of the RTP protocol. This field (CC field) indicates the number of CSRC identifiers in the RTP header, which are mostly unused in specific cases. Therefore, the value of the CC field is usually set to zero [14], [15].

0-1	2	3	4-7	8	9-15	16-31
V	P	X	CC	M	PT	Sequence Number
Timestamp						
Synchronization source (SSRC) Identifier						
Contributing Source (CSRC)						

Figure 1. RTP protocol header format

In addition to VoIP protocols, VoIP uses codec to produce a VoIP packet's payload. Codec is hardware or software that captures analogue voice signals, converts voice signals into digital signals, compresses the resulting digital data to save bandwidth, and produces a voice frame (a packet's payload). Typically, given that VoIP technology is delay-sensitive (the acceptable delay is 150ms), the codec produces a voice frame within a small period. Accordingly, the resultant voice frame size typically ranges from 10 bytes to 30 bytes depending on the codec used [16], [17]. Table 1. shows some of the renowned voice codecs.

Table 1.Renowned VoIP Codecs

Codec Name	G.723.1	G.726	LPC	G.729	G.728
Frame size (B)	20	30	14	10	10
Bit rate (kbps)	5.3	24	5.6	8	16

2. Related Works

As mentioned previously, VoIP suffers from the inefficient utilization of network bandwidth. VoIP developers have exerted massive efforts to enhance the bandwidth utilization of VoIP technology. This section discusses some of these efforts.

Roay patented a VoIP payload concatenation method to enhance the bandwidth utilization of VoIP technology. A typical VoIP packet contains 40 bytes of RTP/UDP/IP preamble. The patented method is based on the idea of concatenating many VoIP payloads in one preamble (header) instead of an independent preamble to each payload. These payloads must belong to different packets that are sharing the same route to one receiver. The resulting concatenated packet contains many small packets in one IP preamble. Each of the small packets within the concatenated packet consists of a voice frame and an RTP/UDP preamble. The VoIP packets are concatenated at the sender side's VoIP gateway while the receiver side's VoIP gateway de-concatenates the concatenated packet to restore the original packets. A concatenation field is appended to the UDP preamble of only the first small packet within the concatenated packet. This concatenation field is used to inform the receiver side's VoIP gateway about whether the received packet is a concatenated or normal packet. Thus, the patented VoIP method saves 20 bytes in IP preamble for each small packet inside the concatenated packet [18].

In addition to the VoIP payload concatenation methods, the VoIP packet preamble compression method is used to enhance the bandwidth utilization of VoIP technology. A compression method was proposed by Casner and Jacobson and then standardized by the Internet Engineering Task Force (IETF). The proposed method compresses the 40-byte VoIP packet preamble (RTP/UDP/IP) to only 2 bytes. Therefore, it greatly reduces VoIP packets' preamble size and thus enhances the bandwidth utilization of VoIP technology. The proposed method reduces the preamble size on the basis of two main principles. The first principle is based on the fact that some RTP/UDP/IP preamble fields do not change during a call's lifetime. Therefore, these fields can prune a packet's preamble and send it during call initiation. The second principle is based on the fact that some of the RTP/UDP/IP preamble fields of consecutive packets are changing in constant value. On the basis of this property, differential coding was applied by Casner and Jacobson to prune these fields [19]. The proposed compression method was refined by Sandlund et al. and standardized again by the IETF. The new standard of the compression method was designed to be general and applicable to different environments, including wireless ones [20].

Sze et al. proposed a method that takes advantage of VoIP payload concatenation and VoIP packet preamble compression methods. Similar to Roay [18], Sze et al. proposed a method that concatenates many VoIP payloads in one preamble. However, unlike that of Roay's method, the resulting concatenated packet in Sze at al.'s method contains many small packets in one UDP/IP preamble, and each of the small packets within the concatenated packet consists of a voice frame and an RTP preamble. In addition to concatenating many VoIP payloads in one preamble, the method proposed by Sze et al. compresses the RTP preamble in the small packet from 12 bytes to 2 bytes only, thereby enhancing bandwidth utilization. The proposed method consists of two main components, namely, Mux and De-Mux components. The Mux component resides at the sender side and performs payload concatenation and RTP preamble compression. The De-Mux component resides at the receiver side and performs payload de-concatenation and RTP preamble decompression [21].

We propose a novel method for enhancing the bandwidth utilization of VoIP technology. Unlike the discussed methods, which mainly focus on the VoIP packet preamble, our proposed method focuses on the VoIP packet payload (voice frame). The main idea of the proposed method is to compress the voice frame by pruning the leading or trailing zeros or ones on the basis of a certain algorithm. The proposed method is called voice frame pruning (VFP). In addition to enhancing the bandwidth utilization of VoIP technology, the VFP method can be combined with either or both of the discussed methods, namely, the VoIP payload concatenation and VoIP packet preamble compression methods. The following section discusses the proposed VFP method in detail.

3. Voice Frame Pruning (VFP) Method

The main goal of the VFP method is to prune the voice frame and thus reduce its size. This eventually enhances the bandwidth utilization of VoIP technology. The VFP method can be implemented at the client side or at the VoIP gateway that is connected to the wide area network (WAN) link. We choose to implement the VFP method at the VoIP gateway for several reasons. The first reason is the ability to use any VoIP client from any vendor with any device without worrying whether or not it supports the VFP method. The second reason is that implementing the VFP method once at the VoIP gateway is easier than implementing it at each client. The third reason is that the local area network usually has great bandwidth, whereas the WAN link bandwidth is limited and expensive. The fourth reason is that the VFP method can be implemented with other methods, such as VoIP payload concatenation or VoIP header compression, which are usually implemented at the VoIP gateway [3], [21], [22]. Figure 2. shows a network topology, including the location at which the VFP method can be implemented.

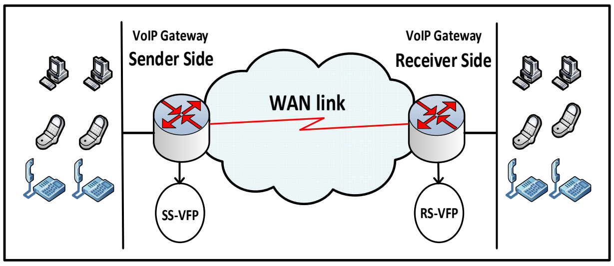

Figure 2. Location of VFP method's modules

The VFP method consists of two main modules. The first module resides at the sender side's gateway, that is, the VFP sender side (SS-VFP). The second module resides at the receiver side's gateway, that is, the VFP receiver side (RS-VFP). The SS-VFP module prunes the voice frame and produces a small voice frame (SVF), whereas the RS-VFP module restores the SVF frame to its normal size and produces an original-size voice frame (OVF). The following two subsections discuss the SS-VFP and RS-VFP modules, respectively.

3.1. SS-VFP Module Process

VFP goes through several steps at the SS-VFP module at the sender side's WAN gateway. Initially, the voice frame is separated from the VoIP packet. Next, the ones or zeros at the start and end of the voice frame (OVF) are counted. Then, the two counters, namely, the start counter (s-count) and the end counter (e-count), are compared. If the s-count is greater than or equal to the e-count, then the counted bits at the start of the OVF are removed. Otherwise, the counted bits at the end of the OVF are removed. This process produces the SVF frame. The examples in Table 2. demonstrate this part of the SS-VFP module process (frame pruning part). Notably, the values of the OVF in Table 2. are not real and are only for demonstration purposes. Subsequently, the CC field in the RTP header is set to a certain value, as discussed below. Before the final step, the SVF frame is reattached to the VoIP packet header (RTP/UDP/IP), which constitutes the VoIP packet. Finally, the VoIP packet is transmitted to the receiver's VoIP gateway. Figure 3. illustrates the SS-VFP module process.

Table 2. Process demonstration examples of SS-VFP module

OVF frame	b-count	e-count	Action	SVF frame
00000011011000110000	6	4	Remove the bits (0s) at the start of the OVF frame because of s-count greater than e-count.	11011000110000
00001111011000100000	4	5	Remove the bits (0s) at the end of the OVF frame because of e-count greater than s-count.	000011110110001
00001111011000100111	4	3	Remove the bits (0s) at the start of the OVF frame because s-count greater than e-count.	1111011000100111
11111111011000100000	8	5	Remove the bits (1s) at the start of the OVF frame because of s-count greater than e-count.	011000100000
11110100011000111111	4	6	Remove the bits (1s) at the end of the OVF frame because of e-count greater than s-count.	11110100011000

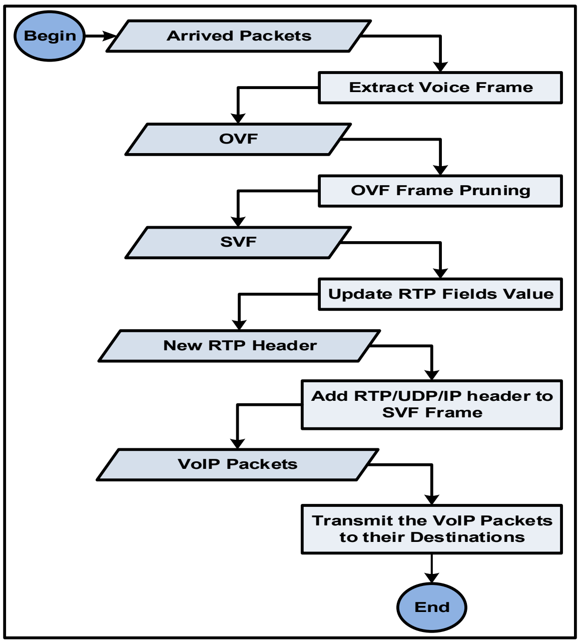

Figure 3. SS-VFP module process

As mentioned previously, the 4-bit CC fields in the RTP header are often set to zero. Therefore, the SS-VFP module at the sender side's WAN gateway uses this field to indicate the value (0 or 1) and location (start or end) of the pruned bits from the VoIP packet payload. The SS-VFP module needs two bits (4 values) for this step. Thus, only the first two bits of the CC field are used. The suggested values of the CC field are shown in Table 3.

Table 3. Suggested values for the RTP CC field

CC field value	Meaning
00	Denotes that the pruned bits are zeros from the end of the VoIP packet payload.
01	Denotes that the pruned bits are ones from the end of the VoIP packet payload.
10	Denotes that the pruned bits are zeros from the start of the VoIP packet payload.
11	Denotes that the pruned bits are ones from the start of the VoIP packet payload.

3.2. *RS-VFP Module Process*

The RS-VFP module process at the receiver side's WAN gateway goes through several steps to restore the SVF frame to its original form (OVF frame form). Initially, the voice frame (SVF frame) is separated from the VoIP packet received. Next, the length difference between the SVF and OVF frames is calculated to determine the number of pruned bits (n bits) from the OVF frame at the SS-VFP module. Then, the value of the CC field is inspected. Subsequently, the SVF frame is restored to its original form (OVF frame form) by adding the pruned bits from the OVF frame at the SS-VFP module based on Table 4. Notably, these values are the same as those in Table 3.

Table 4. Role of RTP CC field at RS-VFP module

CC field value	Action
00	Append zeros at the end of the VoIP packet payload.
01	Append ones at the end of the VoIP packet payload.
10	Append zeros at the start of the VoIP packet payload.
11	Append ones at the start of the VoIP packet payload.

The examples in Table 5. demonstrate the restoration of the SVF frame to the OVF frame using the RS-VFP module. These examples are the same examples in Table 2. Then, the CC field value is set to zeros to avoid misinterpretation at the destination of the VoIP packet. Before the final step, the OVF frame is attached to the VoIP packet header (RTP/UDP/IP), which constitutes the VoIP packet. Finally, the VoIP packet is transmitted to its destination. Figure 4. illustrates the RS-VFP module process.

Table 5.Process demonstration examples of RS-VFP module

SVF frame	CC field value	Action	OVF frame
11011000110000	10	Add n bits (0s) at beginning of the SVF frame.	000000110110001100000
000011110110001	00	Add n bits (0s) at end of the SVF frame.	000011110110001000000
1111011000100111	10	Add n bits (0s) at beginning of the SVF frame.	000011110110001001111
011000100000	11	Add n bits (1s) at beginning of the SVF frame.	111111110110001000000
11110100011000	01	Add n bits (1s) at end of the SVF frame.	111101000110001111111

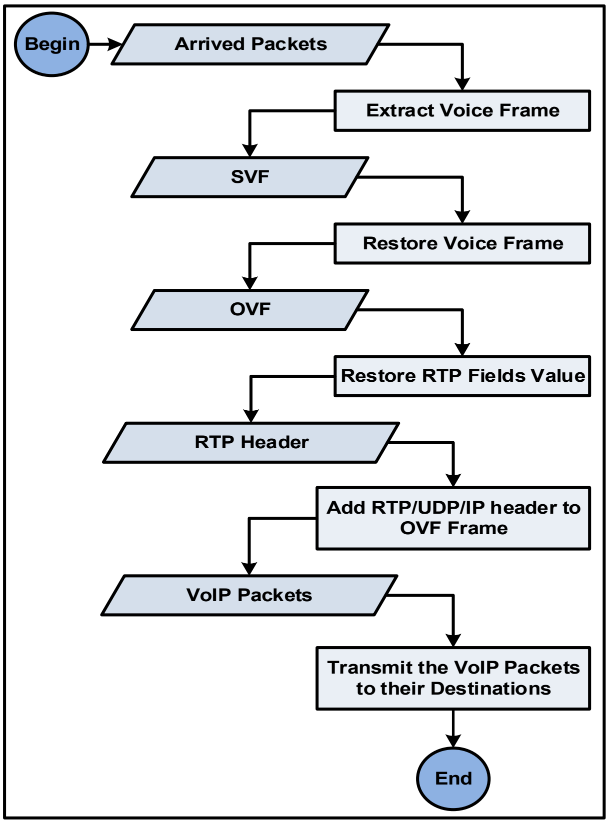

Figure 4. RS-VFP module process

4. Results and Performance Evaluation

This section discusses the VFP method's bandwidth usage. The bandwidth usage was measured on the basis of three different factors, namely, payload pruning ratio (PPR) and saved bandwidth and capacity. To perform a realistic evaluation, the result of these factors is presented with three different codecs, namely, LPC, G.723.1, and G.729.

4.1. Payload Pruning Ratio (PPR)

This section discusses the PPR of the proposed VFP method. The PPR is the percentage of pruned bits from the voice packet payload to the total size of the voice packet payload. The greater the number of bits pruned is, the better the bandwidth usage will be. The PPR was examined with 10 different calls. The average PPR of the 20 packets of each call was separately calculated. Figures 5., 6., and 7. show that the PPRs of LPC, G.723.1, and G.729 reached approximately 19%, 20%, and 16%, respectively. Therefore, considerable bandwidth savings are achieved when using the VFP method. The difference among the PPRs of the three codecs is due to the bit pattern of the voice packet payload.

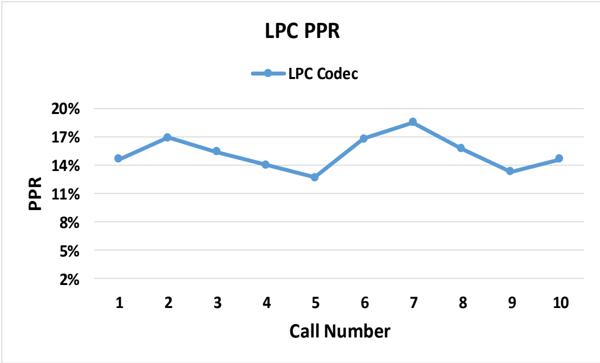

Figure 5. PPR of the LPC codec

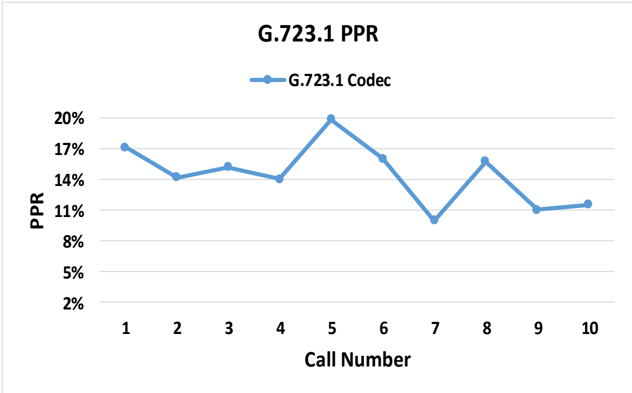

Figure 6. PPR of the G.723.1 codec

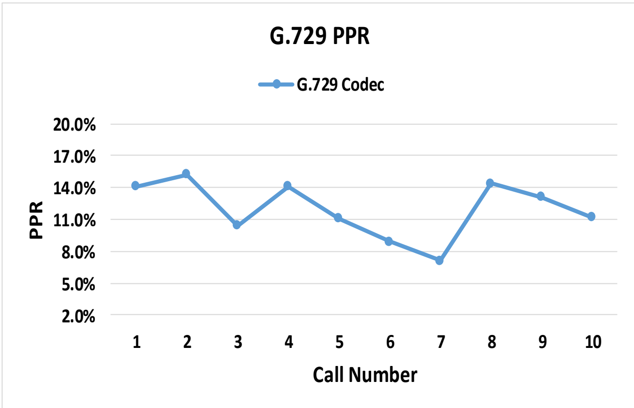

Figure 7. PPR of the G.729 codec

4.2. Saved Bandwidth and Capacity

This section discusses the saved bandwidth of the proposed VFP method and its reflection on capacity. Capacity is the number of voice calls that can run simultaneously within a specific bandwidth. Clearly, the greater the bandwidth saving is, the greater the call capacity will be. The saved bandwidth was examined with 10 different calls. The average of the saved bandwidth of 20 packets of each call was calculated. Figures 8., 9., and 10. show that the bandwidth savings of the LPC, G.723.1, and G.729 improved by up to approximately 5%, 8%, and 3.5%, respectively. Therefore, the capacity improves with similar values. The VFP method exhibits a better capacity (more calls) than the conventional method without pruning. Obviously, this feature is due to the pruning of the voice frame. Again, the differences in the bandwidth savings and capacities of the three codecs are due to the bit pattern of the voice packet payload and the payload pruning ratio.

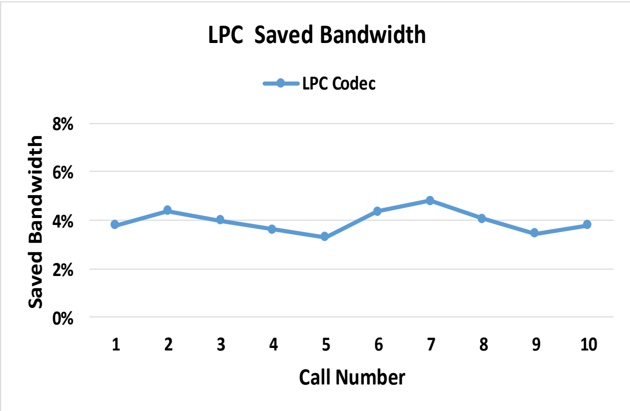

Figure 8. Saved bandwidth of the LPC codec

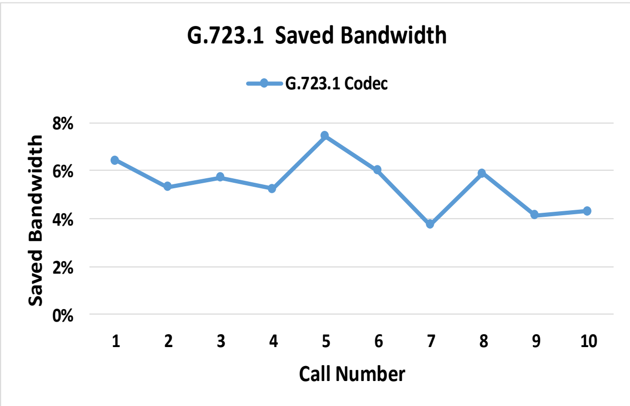

Figure 9. Saved bandwidth of the G.732.1 codec

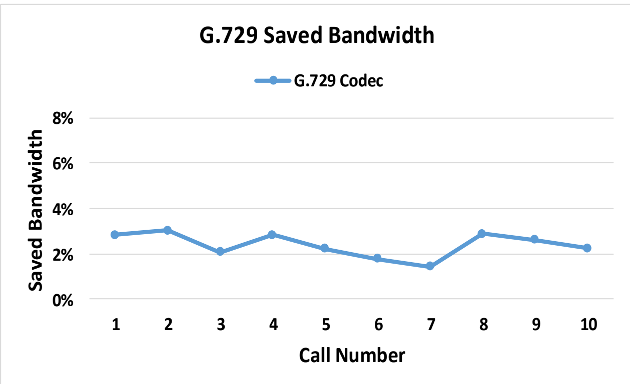

Figure 10. Saved bandwidth of the G.729 codec

5. Conclusion

VoIP is a widespread Internet technology that has replaced the traditional telecommunication network. However, several issues hinder VoIP technology propagation, and they include the inefficient bandwidth utilization of VoIP applications. VoIP packet aggregation and VoIP packet header compression are two main approaches to improving VoIP bandwidth utilization. This study investigated a new approach that works on VoIP packet payloads. VFP compresses VoIP packet payloads by pruning the leading/trailing zeros/ones of the VoIP packet payloads on the basis of a certain mechanism. The VFP method consists of two main modules, namely, the SS-VFP and RS-VFP modules. The SS-VFP module prunes the voice frame at the sender side's WAN gateway, whereas the RS-VFP module restores the SVF frame to its normal size and prunes it at the receiver side's WAN gateway. The proposed VFP method was tested with three different codecs, namely, LPC, G.723.1, and G.729. The result shows that the bandwidth savings with the three codecs improved by up to 5%, 8%, and 3.5%, thereby improving the VoIP bandwidth utilization and capacity of VoIP calls.

References

[1]. Gupta, A., & Chaudhary, A. (2018). A metaheuristic method to hide MP3 sound in JPEG image. *Neural Computing and Applications*, *30*(5), 1611-1618.

[2]. Zhang, Y., Xie, K., & Ruan, O. (2019). An improved and efficient mutual authentication scheme for session initiation protocol. *PloS one*, *14*(3).

[3]. Abualhaj, M., Qaddoum, K., & Abu-Shareha, A. A. (2016). Multiplexing VoIP packets over wireless mesh networks: a survey. *Ksii Transactions On Internet And Information Systems 10*(8), 3728-3752.

[4]. Affonso, E. T., Nunes, R. D., Rosa, R. L., Pivaro, G. F., & Rodriguez, D. Z. (2018). Speech quality assessment in wireless voip communication using deep belief network. *IEEE Access*, *6*, 77022-77032.

[5]. Olariu, C., Fitzpatrick, J., Ghamri-Doudane, Y., & Murphy, L. (2016, April). A delay-aware packet prioritisation mechanism for voice over IP in Wireless Mesh Networks. In *2016 IEEE wireless communications and networking conference* (pp. 1-7). IEEE.

[6]. Natarajan, H., Diggi, S., Kanagarathinam, M. R., Srivastava, S. K., & Bharti, C. (2019, January). *D-VoWiFi-A Guaranteed Bit Rate Scheduling for VoWiFi in non Dedicated Channel*. In 2019 16th IEEE Annual Consumer Communications & Networking Conference (CCNC) (pp. 1-6). IEEE.

[7]. Vulkan, C., Rakos, A., Vincze, Z., & Drozdy, A. (2014). U.S. Patent No. 8,824,304. Washington, DC: U.S. Patent and Trademark Office.

[8]. Fortuna, P., & Ricardo, M. (2009). Header compressed VoIP in IEEE 802.11. *IEEE Wireless communications*, *16*(3), 69-75.

[9]. Karapantazis, S., & Pavlidou, F. N. (2009). VoIP: A comprehensive survey on a promising technology. *Computer Networks*, *53*(12), 2050-2090.

[10]. Abu-Alhaj, M. M., Manjur, S. K., Sureswaran, R., Wan, T. C., Mohamad, I. J., & Manasrah, A. M. (2012). ITTP: A new transport protocol for VoIP applications. *International Journal of Innovative Computing, Information and Control (IJICIC), 8*(3), 1-10.

[11]. Saldana, J., Fernández-Navajas, J., Ruiz-Mas, J., Murillo, J., Navarro, E. V., & Aznar, J. I. (2012). Evaluating the influence of multiplexing schemes and buffer implementation on perceived VoIP conversation quality. *Computer Networks, 56*(7), 1893-1919.

[12]. Kaur, G., Kaur, J., Aggarwal, S., Singla, C., Mahajan, N., Kaushal, S., & Sangaiah, A. K. (2019). An optimized hardware calibration technique for transmission of real-time applications in VoIP network. *Multimedia Tools and Applications, 78*(5), 5537-5570.

[13]. Spencer, M., Capouch, B., Guy, E., Miller, F., & Shumard, K. (2010). *Iax: Inter-asterisk exchange version 2*. Internet Request for Comments.

[14]. Hartpence, B. (2013). *Packet Guide to Voice over IP: A system administrator's guide to VoIP technologies*. O'Reilly Media, Inc.

[15]. Schulzrinne, H., Casner, S., Frederick, R., & Jacobson, V. (2003). RTP: A Transport Protocol for Real-Time Applications, RFC 3550. *Network Working Group*, 1-89.

[16]. Gupta, N., Kumar, N., & Kumar, H. (2018, June). Comparative Analysis of Voice Codecs Over Different Environment Scenarios in VoIP. In *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 540-544). IEEE.

[17]. Ortega, M. O., Altamirano, G. C., Barros, C. L., & Abad, M. F. (2019, June). Comparison between the real and theoretical values of the technical parameters of the VoIP codecs. In *2019 IEEE Colombian Conference on Communications and Computing (COLCOM)* (pp. 1-6). IEEE.

[18]. Roay, B. (2013). U.S. Patent No. 8,553,692. Washington, DC: U.S. Patent and Trademark Office.

[19]. Casner, S., & Jacobson, V. (1999). *Compressing IP/UDP/RTP Headers for Low-Speed Serial Links*, RFC 2508.

[20]. Sandlund, K., Pelletier, G., & LE, J. (2010). *The robust header compression (rohc) framework*. RFC 5795, March.

[21]. Sze, H. P., Liew, S. C., Lee, J. Y., & Yip, D. C. (2002). A multiplexing scheme for H. 323 voice-over-IP applications. *IEEE Journal on Selected Areas in Communications*, *20*(7), 1360-1368.

[22]. Roay, B. (2013). *Generic UDP multiplexing for voice over internet protocol (VOIP)*. U.S. Patent 8,553,692.