[1]M. K.Kishore

[2]D. Lavanya

[3]K. Dennypaul

[4]G. Bhavani

[5]S. Manikanteswarao

# A Comprehensive Review of Voip Technologies and Performance Metrics


***Abstract:*** - VoIP has emerged as a core part of today's communication, having left considerable paradigm shift with flexibilities and cost-effectiveness. These have been elaborated upon in this study regarding the major components of VoIP performance, including Quality of Service (QOS), Quality of Experience (QOE), and security. A few of the most important QoS metrics include latency, jitter, and packet loss, ensuring technical performance and robustness of VoIP systems and assuring efficient service. QOE metrics measure the satisfaction of the user and the subjective perception of call quality, bringing tremendous value as sources of information about understanding what end-users experience. The other point of the paper is associated with the growing relevance of security, where vulnerabilities like eavesdropping, DOS attacks and the need for proper encryption and authentication mechanisms have been inducted. Existing approaches measuring these metrics are analyzed, along with recent improvements in performance optimization and the challenges induced by changeable network conditions. The survey suggests better VoIP service quality and security by strongly underlining the fact that future VoIP systems should be integrated with QOS, QOE, and security measures.

***Keywords:*** VoIP, QoS, QoE, Latency, Jitter, Packet loss, Security, DOS, Encryption, Authentication, Network Conditions

## I. INTRODUCTION

VoIP is gaining importance and rapidly becoming one of the important components of modern communication. VoIP supports voice, video, and multimedia communications over IP networks. Therefore, VoIP describes flexibility, scalability, and efficiency that has popularized mass adoption in personal communications and enterprise environments. The challenge lies in constant variations in network conditions and emerging security threats that may disrupt the flow and compromise the end-user experience for a VoIP-based service. It is a survey paper, with in-depth analysis of the key determinants influencing VoIP performance. QoS, security, and QoE do count. QoS metrics in terms of latency, jitter, and packet loss play important roles both in personal and in-context implementations, while security has emerged also as an important factor: eavesdropping and denial-of-service attacks threaten the integrity and confidentiality of communications. This paper shall provide the current overview of methodologies currently in use to assess such metrics, recent advancements to optimize performance, and new challenges dynamic network conditions present. It further identifies ways to improve the design of VoIP systems with respect to QOS, QOE, and security, towards the provision of safer and better experiences.

## II. VOIP PROTOCOLS AND STANDARDS

There are three Protocols widely used in the implementation of VoIP

### *A. Session Initiation Protocol*

The basic structure of a SIP-based network is very commonly employed in VoIP systems. Actually, it is the SIP that works as a signaling protocol used for the establishment, overseeing, and termination of the different sorts of multimedia sessions including voice and video calls over IP networks. Three major components of the SIP network are: User Agent Clients, which are computers or phones that initiate and receive a session; Proxy Servers, which route SIP messages to their destinations while taking care of authentication management; Registrar Servers, which store user location by mapping the SIP address to the IP; Redirect Servers, which aid in forwarding requests to the proper server. It also envisions a gateway that relates the SIP network to the PSTN in order to allow VoIP users to reach PSTN users and vice versa. The architecture does much speaking on how relevant SIP is when it comes to

[1]Assistant Professor, [2][3][4][5]Students, Department of Electronics and Communication, Usha Rama College of Engineering and Technology, Telaprolu, India

communicating across different devices as well as networks but underlines its role in modern multimedia communications[7].

SIP networks typically feature a gateway that connects Voice over Internet Protocol (VoIP) systems to the Public Switched Telephone Network (PSTN). This connection allows VoIP callers to communicate with users on traditional landlines or mobile phones. The way a SIP network is configured plays a crucial role in its compatibility with various devices, platforms, and networks. This widespread configuration highlights its significance in contemporary multimedia communications, offering a communication system that is easy to expand, flexible, and capable of real-time interactions, which are essential in today's interconnected land.

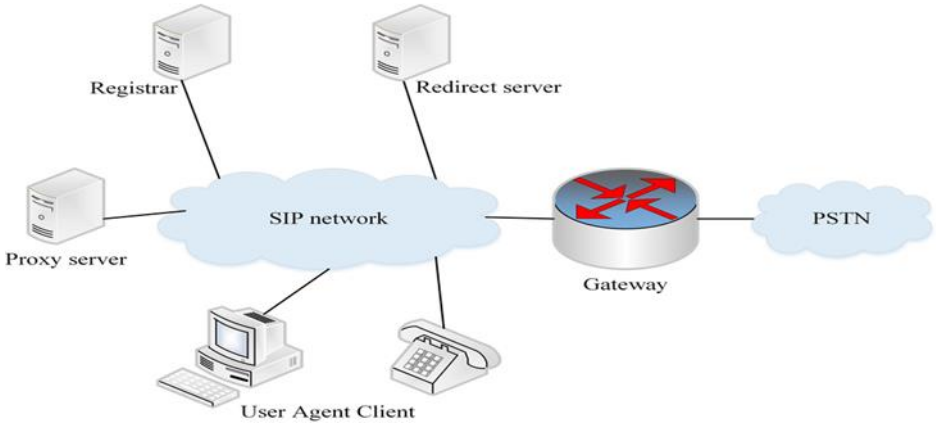


Fig.1 Components of SIP Architecture

*B. Real-Time Transport Protocol*

It is one of the elementary protocols for audio and video streams to be carried over IP networks so that real- time applications like VoIP can be facilitated. In this system, the central Call Manager will be responsible for directing the signaling for the call to control both the flow of signaling between the devices and ensure that phones from both private and public networks interact over a series of signaling messages. A private network would use the NAT router to translate the IP addresses so that an inside device can contact an outside network. Thus, the RTP Proxy plays a critical role in enabling actual data exchange, manages RTP streams between the caller and receiver, and further solves NAT traversal issues. In addition, the proxy mediates media paths to ensure RTP traffic is possible for the free flow between the private and the public networks. Such a setup clearly distinguishes the signaling function from media transport, thereby showing how control mechanisms like SIP complement RTP to ensure smooth communication of multimedia over different network infrastructures. Because in RTP, QoS is an important issue for applications that rely on real-time communication, it becomes crucial that jitter buffers and congestion control support the management of network variability since RTP implies no guarantee about the delivery of packets in any particular order, or even within a prescribed time interval. This diagram shows a solid RTP architecture that manages and routes traffic in such a way as to minimize delay and packet loss, which is critical to maintaining the quality of voice or video streams in VoIP systems. Many components, such as NAT routers and RTP proxies, or call managers work well together in enterprise VoIP or in the interconnection of public and private networks to ensure consistent, secure, and scalable performance of the communication infrastructure. In addition, RTP has Secure RTP (SRTP), which implements encryption, message authentication, and integrity to provide confidential and secure voice and video communication-a critical requirement for enterprise and public systems. In addition, RTP carries metadata that indicates the type of media being transmitted. therefore, this allows audio, video, and text type of media to be handed over appropriately by the receiver.

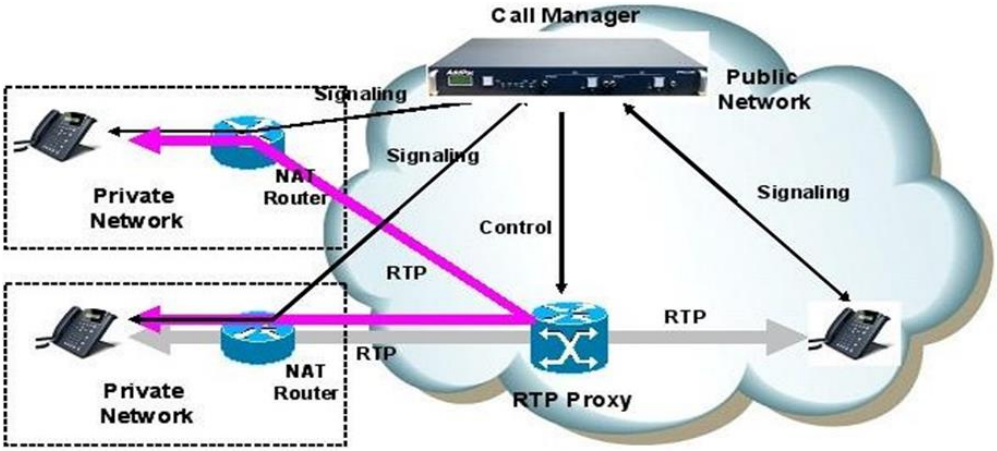


Fig.2: An Overview of Real-Time Transport Protocol

*C. Media Gateway Control Protocols*

The Media Gateway Control Protocol is a signaling and control protocol that is primarily used in VoIP systems in order to manage media gateways, which help to translate media streams between different transmission formats: traditional telephony and IP-based communications. MGCP is a client-server protocol that allows a centralized call control application, referred to as the call agent, to exercise control over the media gateways acting as endpoints in these communication systems. It also allows session control management in order to establish, modify, and terminate media sessions while ensuring efficient allocation and deallocation of resources for such sessions. It will support a number of control messages about which device could be initiated or terminated, information about playing announcements, and recording conversations, all of which are pretty important for applications like IVR systems.

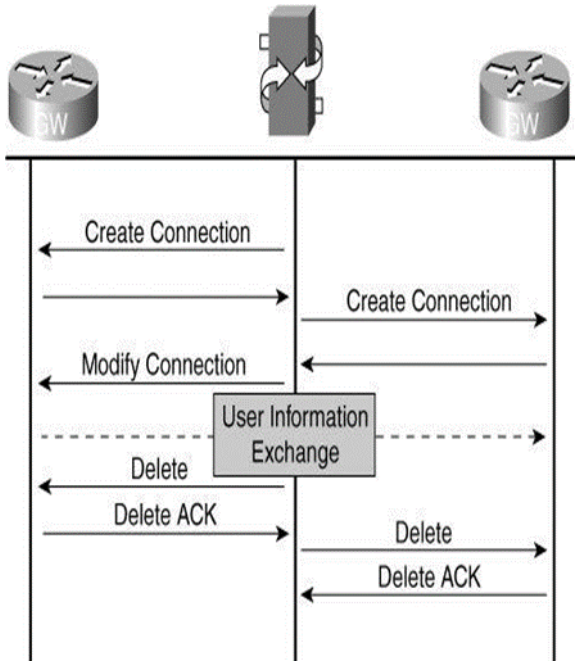


Fig.3: MGCP Calls and Connections

## III. VOIP PERFORMANCE METRICS

VoIP performance metrics are criteria used for measurements for voice communication quality and efficiency across the IP network. It will be important in both user experience and system performance while assessing real-time communications. Major VoIP performance metrics

### *A. Latency(End-to-End Delay)*

Latency means how long it takes for a voice packet to travel from sender to receiver. High latency really affects the flow of natural conversation since it is perceivable, more so in real-time communications like VoIP[7].

### *B. Jitter*

Jitter is variation in packet arrival times, typically caused by congestion within the network or inconsistent timing with regard to the transmission of packets. Receiving packets that are out of sequence or significantly delayed degrades the quality of voice. There is commonly a jitter buffer that smooths this variation and keeps the stream stable[7].

### *C. Packet Loss*

The packets of voice data do not even reach there; such a phenomenon is known as packet loss. Voice data arrives in packets, and these packets cause gaps within the audio. Even less percentages of packet loss could degrade VoIP call quality considerably. For example, a 1-3% packet loss rate creates audible problems in voice quality; serious losses[5][7].

### *D. Throughput*

Throughput measures the amount of information transferred by a network in a given time, and for VoIP, good throughput has to be maintained to ensure that voice packets can be delivered with delay or loss[7].

### *E. Mean Opinion Score(MOS)*

MOS is the level of subjectivity of the voice quality itself; in this scenario, it measures the quality of a call using a rating that ranges from 1 (bad) to 5 (excellent). Determined based on user reports, it is affected by other influences that include latency, jitter, and packet loss. In VoIP, scores of 4 and above are normally adequate.

### *F. Echo*

In echo, a caller hears an echo of his own voice when calling. The reason for this can be due to either signal reflections in the network or at the user's end. Even though some level of echo is expected, significant echo can really disrupt conversation and result in reduced user satisfaction.

### *G. Bandwidth*

Bandwidth is the amount of data transferred over the network. VoIP needs a good amount of bandwidth in order to carry the audio data streams. More bandwidth would result in congestion, packet loss, and poor call quality. VoIP codecs-for instance G.711, G.729-do voice data compression to reduce the bandwidth requirement.

### *H. Signal-to-Noise Ratio(SNR)*

This measures the ratio of the level of the desired signal to the level of background noise. Generally, the higher the signal-to-noise ratio, the clearer the calls and their quality is good. Low SNR usually brings noisy calls.

### *I. RTT: Round Trip Time*

RTT is the time taken by a packet to reach its destination and come back to the sender. High values of RTT cause communication delays and degrade real-time nature of VoIP communications.

## IV. VOIP SECURITY

VoIP Attacks/Threats Attackers usually target some popular and well-known systems and applications. VoIP has become one of such application. VoIP like any another system or applications have its weakness, thus protocol designers need to address it before successfully installing VoIP on a universal scale. This section, presents a study of attacks on the VoIP[1].

### *A. DoS(Denial of Service)*

DoS (Denial of Service) attacks which reduces the number of available IP addresses, bandwidth and other router functions. A DoS attack usually blocks the service of the server. A VoIP based DoS attack bombards a call processing application with large amounts of concurrent requests that it cannot process, causing the shutting down

of the application, thereby denying service to authorized or intended users. DoS attacks can be directed toward any network element to disrupt the system's functionality[1][2].

### *B. Network Sniffing*

Network Sniffing attacks occur when an individual or attacker is observing the network traffic patterns. Typically, any system (user/attacker) on a network that is sharing a transmission medium has the ability to view other system's traffic[1].

### *C. Eavesdropping*

Eavesdropping is an endeavor towards collecting sensitive information to prepare for a cyber-attack or to gain intelligence. In VoIP, eavesdropping is a scenario where the attacker is able to monitor signal or media contents that are exchanged between users in order to examine communications to prepare for other future attacks[1].

## V. CLASSIFICATION OF VOIP METHODOLOGIES

### *A. Packet Compression and Packet Aggregation Algorithm*

#### *1) Packet Compression*

In this method, the payload of individual data packets is compressed right before it leaves the client node into the network. This minimizes the amount of data being transferred and, therefore, reduces bandwidth usage. Packet compression makes this possible; it thus improves efficiency in transmission over bandwidth-constrained environments. Compression also reduces the probability of congestion, which leads to more efficient network performance and potentially lower packet delay as well as jitter. However, whereas compression has the effect of reducing the data size, it increases some overhead processing since packets have to be compressed before sending and decompressed on reception[3].

#### *2) Packet Aggregation*

Aggregation of multiple small packets in this context, which is divided into one large packet before they transmit. Then, it would result in every single packet having lower overhead, such as headers, thus leading to a fewer packet transmission over the network. This is the reason why packet aggregation minimizes congestion possibilities while improving throughput. This may reduce queuing and transmission time significantly, and consequently packet delay and jitter. Packet aggregation can be effective for improving network performance; however, in certain situations, especially wireless mesh networks environments, packet aggregation could easily increase packet size at the peril of fragmentation or additional delay at critical times[3].

### *B. Runt Payload VoIP (RPV) Packet Methodology*

The most important aspect of the RPV methodology is minimizing packet length data of VoIP to as low as zero octets. This, therefore improves the BW exploitation of VoIP technology, particularly for unicast Voice over IP conversations. The RPV method can be used at the VoIP client side or at the VoIP concentrator that is assigned to the WAN interface. The RPV methodology is prescribed to be used is sent at the VoIP concentrator for several reasons. First, the ability to use the RPV methodology with any VoIP client from any company with any apparatus without stressing whether it implements RPV Methodology. The local network usually has much more BW at no additional cost while WAN connection BW is limited and costly. Third, the RPV method is summable with other methods, such as Coalescence methods that are commonly used at a VoIP concentrator in VOIP. This RPV methodology involves two major constituents. One is called the payload diminishing (PD) entity, and it is located at the transmitter side's concentrator. The second one is named at the receiver side's concentrator. The VoIP packet data is it is situated at the payload reversion (PR) Produce the standard VoIP packet and recover the speech data to its original length. The PD and PR entities describe these in detail. Figure 4 presents the topology of a network and the position where the method of RPV will be applied[4].

### *C. Fuzzy Logic Approach*

In the case of VoIP quality maintenance, Fuzzy Logic provides an adaptive dynamic approach to the management of network parameters that have a direct influence on voice quality-a packet loss, delay, and jitter. Real-time communication with minimal disruptions qualifies as the state of the ideal situation for any application; however,

the traditional methods of traffic management have long been found wanting when dealing with the dynamism in network conditions. In effect, a fuzzy logic technique actually addresses this problem by interjecting a rule-based system that can handle imprecise or uncertain input data, hence making possible more elastic decisions.

In the Fuzzy Logic approach, bandwidth rate and buffer size would be identified as two parameters deemed an important network parameter used within mechanisms such as the Token Bucket Algorithm, and a set of fuzzy rules implemented as an attempt at dynamically adjusting those parameters.

For example, if the network congestion is increased, fuzzy logic may reduce the rate at which the token is supposed to be generated so that more opportunities may be there for fewer chances of packet loss and unfruitful delays. Such adaptability allows voice data to be managed in an efficient way under fluctuating conditions, leading to better QoS outcomes. This integrates fuzzy logic into management systems for traffic that assure VoIP quality to be maintained at optimal levels even in dynamic conditions within networks. This approach helps smoothen spikes within the networks to ensure consistent and minimum degradation delivery of voice packets. Therefore, users enjoy clearer voice communications without probable interruption even in sub optimal network environments. Simulation studies, such as those done in Opnet, with fuzzy logic used in combination with conventional QoS mechanisms across multiple network scenarios, depict considerable improvement for maintaining VoIP quality[6].

*D. Payload Shrinking over Internet Telephony Transport Protocol (ITTP-PS) method*

The Payload Shrinking over Internet Telephony Transport Protocol (ITTP-PS) method is a new approach to bandwidth utilization in Voice over Internet Protocol (VoIP) systems. In fact, the usual limitation of traditional VoIP systems is that more calls could be made at the same time. The challenge that ITTP-PS solves is by introducing a payload compression mechanism that reduces the size of VoIP packets during transmission, thus optimizing bandwidth usage. This ITTP consists of two components: Sender ITTP-PS and Receiver ITTP-PS. The S-ITTP-PS compresses the payload in the packet before sending it; hence, the data volume to be transmitted is reduced. When the receiver receives the compressed packet, the payload is restored to its original size by the R-ITTP-PS so that the integrity of VoIP data is guaranteed.

Another benefit of the ITTP-PS method is that it uses the flag bits of the IP protocol header for the process of payload compression and restoration. The process ensures that the information being sent is effectively conveyed concerning decompression without highly raising the overhead of the payload. It also performs communication with smother flows and resource management with no compromising quality in VoIP calls.

The ITTP-PS method revealed that bandwidth utilization was greatly enhanced in performance evaluations. The payload shrinking ratio on VoIP packet payload reached as high as 20%, allowing for much efficient data transmission. Isochronous calls capacity improvement ratio went up about 9.5%. This means that the network can carry more simultaneous VoIP calls without degrading call quality. These results thus indicate that ITTP-PS is an effective solution for the exploitation of bandwidth in VoIP systems and also a very scalable and efficient way to improve performance under high-demand scenarios.

*E. Packet Multiplexing and Carrier Fields (PMCF) method*

The Packet Multiplexing and Carrier Fields PMCF is a novel technique for solving the inefficiency problem in bandwidth usage BWU while running VoIP over IPv6. IPv6 has a large header, which brings a significant overhead especially in VoIP due to the normal transmission of very small speech frames. Therefore, such poor adaptation between large header size and the payload leads to wasted bandwidth and lower call capacity. This problem is solved using the PMCF method that integrates two of the most vital techniques used: packet multiplexing and carrier fields. Packet multiplexing is a technique in which many VoIP packets are aggregated into a single header of IPv6, thereby reducing the overhead of sending individual headers for each packet. This helps increase bandwidth efficiency because more data can travel on fewer headers. Another advantage of PMCF is that it makes use of some unused or redundant fields within the IPv6 header for parts of the speech frame. The method further optimizes the available bandwidth due to the fact that it embeds some part of the voice data in some of those fields. This double optimization will drastically reduce the amount of bandwidth that would be required for VoIP calls and improve both capacity call and general network performance.

In terms of performance evaluation, it has been determined that the PMCF technique has brought about an outstanding improvement in BWU by depending upon metrics such as call capacity, header size reduction, bandwidth saving, and shortening speech frames. Further, call capacity was increased up to 269% as compared to traditional IPv6 VoIP methods. This makes the PMCF method a promising solution for an improvement in VoIP performance in IPv6 environments under conditions of high demand for virtual communication[9].

Table 1: Numerical Comparison of Network Approaches in Terms of Performance Metrics and Security

| Approach | Throughput (Mbps) | Latency (ms) | Packet Loss (%) | Security Level (1-5) |
|---|---|---|---|---|
| PCA | 90 | 30 | 2 | 3 |
| RPV Packet | 75 | 40 | 5 | 4 |
| Fuzzy Logic | 80 | 35 | 3 | 3 |
| RPT-ITP | 95 | 20 | 1 | 5 |
| PMCF | 70 | 50 | 4 | 4 |

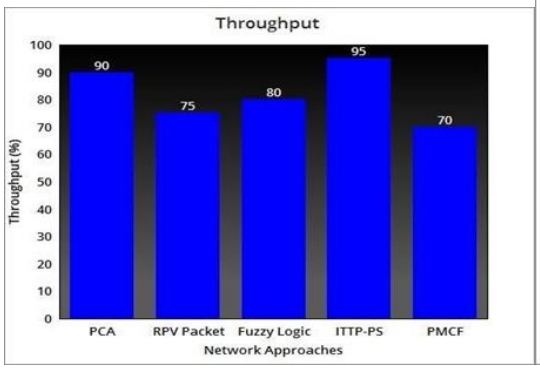


Fig.4 Throughput Comparison of Network Approaches

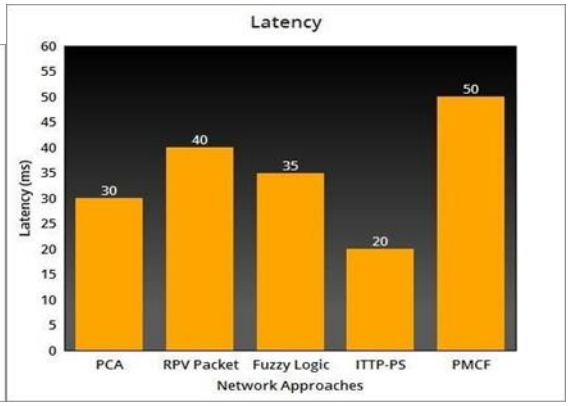


Fig.5: Latency Comparison of Network Approaches

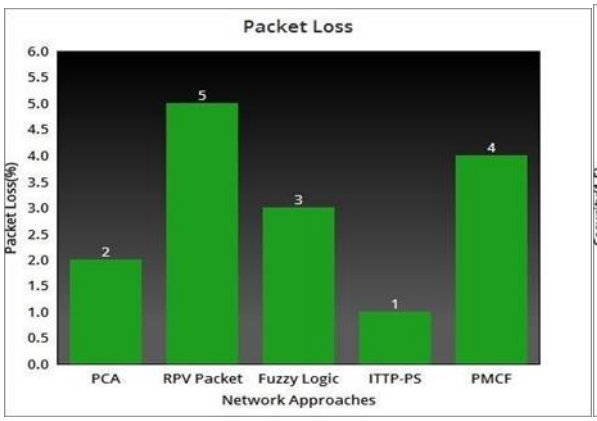


Fig.6 PacketLoss Comparison of Network Approaches

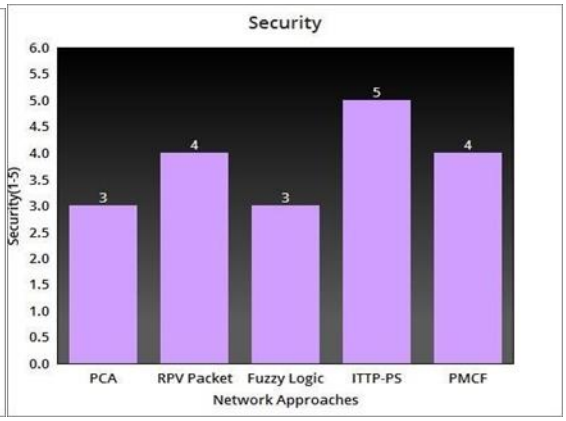


Fig.7 Security Comparison of Network Approaches

## VI. Challenges and Future Directions of VoIP

VoIP technology changed communication by passing voice over the internet, but it brings with itself challenges that must be addressed in order to secure its further development. Among them, the most critical challenge is QoS and QoE issues, especially in mobile environments where congesting the network, packet loss, jitter, and latency can activate call quality degradation. However, security risks including eavesdropping, DoS attacks, and data breaches also threaten because VoIP uses open networks and internet protocols. Besides, interoperation with other existing standards and legacy systems like the traditional PSTN is another challenge since

communication cannot rightly flow across different platforms. Scalability is another challenge, given the boom of IoT devices and 5G, which will further augment stress on VoIP infrastructures. Machine learning and artificial intelligence will remain dominant forces in improving VoIP performance; the applications include predictive optimization of real- time traffic, QoS/QoE, and network management automation. Advancements in compression and aggregation are available to increase bandwidth efficiency, while blockchain technology is capable of improvement in the voice over IP system's security and privacy.

The future of communications would rather be very highly integrated with cloud-based VoIP services. Allowing VoIP providers, VoIP providers would be able to offer scalable and cost-effective solutions that have on- demand resources. This should make expansion easier and increase the flexibility of communication capabilities for businesses. Hybrid cloud models may provide increased reliability and resilience through the spreading of VoIP workloads across several environments. Virtualization technologies, for example, NFV and containers, should make it easier to deploy services such as VoIP and reduce latency and accelerate services rollout. As the technologies mature, they shall be able to provide customized service offerings dependent on individual user requirements, which further cements the position of VoIP as a very amenable communication platform.

## VII. Conclusion

In Conclusion, this survey has explored various methods to optimizing VoIP systems in terms of performance and security are discussed and their strengths and challenges emphasized. Packet Compression & Aggregation and Runt Payload VoIP Packet methods are excellent as regards bandwidth efficiency and throughput, thus are efficient for the enhancement of performance. Again, both methods are exposed to security risks, which require robust encryption protocols such as TLS and SRTP to mitigate vulnerabilities. The Fuzzy Logic Approach is a balanced solution because, for example, it optimizes network conditions, reduces latency and jitter, though decision logic that informs these processes pose risks. This makes the ITTP-PS approach the most reliable for VoIP, enabling it to have high performance with low latencies and jitter alongside complete encryption for safe communication. It is well suited for critical real-time applications. On the other hand, PMCF, despite having a very good range of bandwidth management efficiency, faces security-related issues with increased risks from multiplexing, so it needs very strong security implementations like IPsec and TLS.

Ultimately, the future of VoIP depends on achieving an acceptable balance between performance optimization and security enhancement. Although several techniques promise improvement by many orders of magnitude higher than those introduced in this chapter, robust security will be required for further deployments of VoIP systems, particularly as communication technologies advance with IoT, 5G, and beyond

## VIII. References


[1] U. Shaw and B. Sharma, “A survey paper on Voice over Internet Protocol (VoIP),” International Journal of Computer Applications, vol. 139, no. 2, pp. 1–6, Apr. 2016.

[2] S. Jalendry and S. Verma, "A detailed review on voice over internet protocol (VoIP)," *International Journal of Engineering Trends and Technology (IJETT)*, vol. 23, no. 4, p. 161, May 2015. [Online].

[3] O. Olorunnisola, T. E. Mathonsi, and D. Du Plessis, "An algorithm to optimize concurrent VoIP calls across wireless mesh networks," *Journal of Advances in Information Technology*, vol. 14, no. 5, pp. 1–10, 2023.

[4] M. M. Abualhaj, A. A. Abu-Shareha, and S. N. Al-Khatib, "Utilizing VoIP packet header’s fields to save the bandwidth," *Transport and Telecommunication*, vol. 24, no. 1, pp. 33–42, 2023, doi: 10.2478/ttj-2023-0004.

[5] O. P. Roy and V. Kumar, "A survey on voice over Internet Protocol (VoIP) reliability research," *IOP Conference Series: Materials Science and Engineering*, vol. 1020, p. 012015, 2021, doi: 10.1088/1757-899X/1020/1/012015.

[6] M. E. A. Ebrahim and H. A. Hefny, "Fuzzy logic-based approach for VoIP quality maintaining," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 1, pp. 1–7, 2018, doi: 10.14569/IJACSA.2018.090101.

[7] S. Kumar and K. Sharma, "A review paper on Voice over Internet Protocol," in *V-IMPACT - 2016 Conference Proceedings*, 2016, pp. 1–5.

[8] Q. Shambour, S. N. Alkhatib, M. M. Abualhaj, and Y. Alraba’nah, "Effective voice frame shrinking method to enhance VoIP bandwidth exploitation," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 7, pp. 141–146, 2020, doi: 10.14569/IJACSA.2020.0110718.

[9] M. M. Abualhaj and S. N. Al-Khatib, "A new method to boost VoIP performance over IPv6 networks," *Transport and Telecommunication*, vol. 23, no. 1, pp. 62–72, 2022, doi: 10.2478/ttj-2022-0006.