EasyChair Preprint

№ 10403

Solution for Extending the Cloud-Based Call Services with Voice Recognition

Azzaya Battogtokh, Uyanga Sambuu and Enkhtuul Bukhsuren

June 15, 2023

Solution for extending the cloud-based call services with voice recognition

Battogtokh Azzaya[1]. Sambuu Uyanga[2*], Bukhsuren Enkhtuul[3]
[123]*Department of information and Computer Sciences, School of Engineering and Applied Sciences, Ulaanbaatar, Mongolia*
{18m1num0368[1], uyanga[2], enkhtuul[3]}@seas.num.edu.mn*

Abstract

The Information and Communication Technology (ICT) for improving the organizational culture and services, enhancing the transparency of the organizations' activities, improving the quality of services, and satisfying the needs and requirements of modern societies and markets. This article proposes the technological solution for extending the Cloud-based Call Services with voice recognition. This solution can be fully implemented in the organization's information system in a short time and at a low cost.

Keywords: *ICT, cloud computing, call service, information system, voice recognition*

1. Introduction

Telephone communication has long been intertwined with the dominant consumer inclination, primarily due to its versatile spectrum of channels dedicated to the provision of organizational care. As a result, the establishment of a customer care center or call center has become an integral constituent within the operational structure of every support-centric organization. Enhancements in customer call center operations have led to increasingly sophisticated approaches in retrieving schedules, locations, and other pertinent information. These advancements are driven by customer interactions encompassing order placement, information acquisition, and even personalized information customization. Conversely, while the organization provides users with the option to access the system via the Internet, it remains imperative to accommodate individuals who lack Internet accessibility, ensuring their uninterrupted access to services through telephonic means and enabling them to retrieve their pertinent information. Hence, prioritizing customer identification verification becomes paramount for organizations as it allows them to not only authenticate customers during phone interactions but also offer tailored information and leverage system functionalities to foster customer retention and support. Consequently, customer service centers encounter the imperative of fulfilling requests that necessitate customer identification through telephonic engagements.

In contemporary times, both public and private entities operating in the market, irrespective of their nature, magnitude, and scope of operations, have extensively adopted call services as a means to disseminate information to the public and engage with consumers. Driven by the expanding market demand and technological advancements, service providers have embraced cutting-edge technologies in their call services, particularly leveraging cloud technology in recent years. Cloud computing constitutes not only an information technology paradigm but also a transformative force that has revolutionized the business landscape, ushering in substantial changes and disruptive innovations. It signifies a paradigm shift in IT that is swiftly reshaping how organizations administer and dispense IT services via the Internet [1]. It encompasses the convergence of networked software and virtualized hardware, encompassing their comprehensive management and provision of services.As a result, through the process of uploading and utilizing essential software and data on Internet-based servers, users are relieved of the burden of ensuring the completeness, reliability, information security, and technical services of their operations. Instead, the onus lies with the service provider to assume responsibility for these aspects.Conversely, the adoption of cloud computing empowers users to select products and services at equitable market prices and engage in pay-as-you-go models, affording them comprehensive control, management, and strategic planning of their usage. In contemporary times, amidst the advancement of cloud computing, information management based on cloud computing has emerged as one of the novel concepts that has been introduced. Utilizing cloud computing for data management has

become a prominent competitive strategy employed by organizations. As the significance of information management in generating organizational value and facilitating business processes expands, facilitating cloud-based productivity becomes imperative for organizational success within an evolving business landscape [2]. In their study, Herhalt and Cochrane highlighted that users have the ability to leverage cloud computing limitlessly as per their requirements, rather than constructing their own physical infrastructure. This preference for Internet-based cloud computing services enables users to pay only for the services they utilize [3]. Moreover, the act of capturing and storing substantial volumes of data presents numerous opportunities and advantages, including enhanced service quality, innovation in service offerings, and heightened competitive advantage for organizations.

In Mongolia, it has become customary for consumers to avail themselves of services such as card deactivation, retrieval of recent transaction details, and other related inquiries by contacting the bank's designated telephone number. To facilitate the provision of the service, the call center operator obtains the user's registration number and verifies their identity through a series of predetermined questions. Upon successful verification, the service is attributed to the user's identity. Nevertheless, there exists a potential risk wherein an individual possessing knowledge of the customer's registration number and the corresponding answers to registered questions may exploit this information to gain unauthorized access to the service through phone calls.

Hence, to ensure the precision and security of the user identification and verification process, it is imperative to promptly devise a solution that addresses the challenges inherent in user identification and verification.

2. Related work

A considerable amount of research has been conducted on the topic of user authentication.For instance, in their research paper titled 'Specifics of Receiving and Processing Phase Information in Voice Authentication Systems,' Mykola Pastushenko, Viacheslav Pastushenko, Oleksandr Pastushenko, et al. (2019) assert that while user static biometric features like facial images and fingerprints are extensively employed for identification and authentication, they are susceptible to falsification, rendering the system unable to fulfill the requirements of both developers and users.In recent times, there has been a notable proliferation in the utilization of dynamic (behavioral) biometric features of users, exemplified by voice recognition systems. This development has shifted the paradigm of user voice signal processing, augmenting amplitude-frequency spectrum analysis with investigations into phase data that were previously disregarded during the authentication process. A study titled 'Remote User Authentication Using a Voice Authentication System' was conducted by Shilpa Kaman, K. Swetha, Syed Akram, and G. Varaprasad in 2013. The research work highlights the simultaneous rise in the usage of applications involving sensitive information, such as e-commerce, e-learning, e-banking, and e-health services, and the corresponding surge in information-accessible terminals, underscoring the genuine need for robust and user-friendly authentication methods.Conventional personal identification methods, such as passwords and PIN codes, possess inherent vulnerabilities that impede their ability to fulfill the demands of data security. Biometrics pertains to the automated recognition of an individual by utilizing their physiological and behavioral attributes. In their research paper titled 'VoiceMask: Anonymize and Sanitize Voice Input on Mobile Devices,' Jianwei Qian, Haohua Du, Jiahui Hou, Linlin Chen, Taeho Jung, Xiang-Yang Li, Yu Wang, and Yanbo Deng (2017) proposed an approach to enhance the accuracy of speech recognition by leveraging cloud technology. They introduced a tool called Voice Mask, which acts as an intermediary between users and cloud technology, sanitizing and anonymizing users' voice data before transmitting it to the cloud. The research paper further examined potential privacy risks and explored methods to safeguard users against the exposure of their identity and password information within the cloud environment.

3. Research material methodology

3.1 Cloud-based call services

The effective adoption of cloud technologies yields distinct performance and advantages that vary according to the selected delivery model for organizations. Presently, numerous cloud-based services exist, each implementing organizational cloud service strategies in unique ways [5]. The delivery of services through cloud technology relies on fundamental models: public cloud, private cloud, and hybrid cloud.

A cloud-based call service solution offers a range of benefits, encompassing the following:

- It offers cost-effectiveness and economic efficiency;
- It enables call reception across various devices, such as desktop or laptop computers, IP phones, mobile phones, or smartphones, irrespective of the device being used;

● In scenarios of high call volume where calls cannot be immediately answered, the system facilitates the reception of voicemail messages from callers and the delivery of call information via email;
● It offers the capability to function as a Call Center, ensuring equitable distribution of calls among operators;
● It presents the opportunity for cost and time savings by eliminating the need to acquire primary equipment.
● It provides the capability to send and receive faxes in PDF format via email, eliminating the need for additional faxing equipment;
● It incurs minimal engineering costs;
● It provides continuous updates and the opportunity to leverage the latest advancements in technology periodically.
● It offers the potential for seamless integration;
● It allows for simultaneous usage across multiple locations;
● It enables time-saving and promptness;

When examining cloud computing from the standpoint of hardware provisioning and pricing, the following three aspects of innovation can be observed:

1. The availability of unlimited on-demand computing resources capable of accommodating load growth has eliminated the need for cloud computing users to pre-plan resources.
2. The user will have the flexibility to increase consumption only when necessary.
3. Computational costs can be paid for short durations, such as hourly processor usage or daily storage, and released as needed [6].

3.2 User identification and authentication

Authentication encompasses the intricate procedure of authorizing and granting users privileged access to an information system, meticulously validating their credentials, and establishing their identity to ensure secure and authorized interactions within the system. There are three prominent authentication mechanisms widely employed in information systems: password entry, smart card authentication, and biometric authentication. These mechanisms serve as key pillars in verifying the identity and granting access to authorized users, each leveraging unique factors such as knowledge-based credentials, physical tokens, or distinctive biological traits for robust and reliable authentication processes. Each authentication mechanism operates in distinct ways, exhibiting inherent strengths and limitations. Consequently, understanding the unique characteristics of each mechanism is crucial in selecting the most suitable approach for a given context. By comprehending the advantages and disadvantages associated with each authentication mechanism, organizations can make informed decisions that align with their security requirements and user experience considerations.

3.3 Solution for extending the cloud-based call services with voice recognition

Cloud-based call services can be efficiently deployed utilizing Internet networks and softphones, empowering organizational management to access the system through online call center software. This enables them to seamlessly review call recordings, assess customer satisfaction levels, monitor call center operators performance, and access comprehensive reports, all while reaping numerous benefits and circumventing excessive expenditures. By leveraging the advantages offered by cloud-based call services, organizations can achieve enhanced operational efficiency and cost-effectiveness, thereby optimizing their call center operations. Figure 1 depicts the overarching architecture of the system, wherein the user's identification is facilitated through the call center system. Subsequently, the user gains access to pertinent information from the organization's core system and is empowered to execute various feasible actions. This illustration provides a visual representation of the interconnected components and their roles within the system, showcasing the flow of information and interactions between the user, call center system, and the organization's core system.

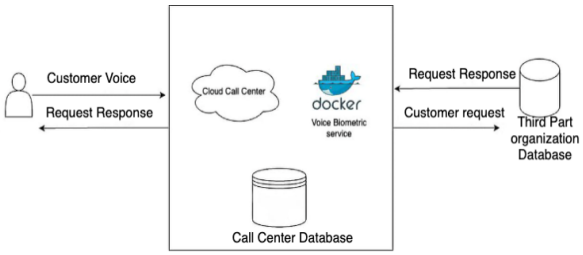

Figure 1. Solution architecture

User: To engage directly with the organization's phone calls and access the pertinent information handled within the call system, it is imperative to undergo a process of verification and identification. This ensures that authorized individuals can securely interact with the call system and retrieve the relevant data they require. By successfully completing the verification and identification steps, users are granted appropriate access privileges, enabling them to effectively collaborate with the call system and retrieve the necessary information.

Call center system: In addition to the calling system, the user's voice can be saved as a model. When the user calls again, a new model is created, which is then compared with the previously stored models in the database to identify the user. Once the user is identified, the relevant data or user request is transmitted to the organization through the API. Subsequently, the user's voice is utilized to retrieve the corresponding response, which can be listened to by the user. This process enables efficient and accurate identification of users based on their voice patterns, facilitating seamless communication and interaction with the organization's services.

Organization: The system ensures the smooth operation of API requests originating from the call center system and effectively responds to customer inquiries. It manages and coordinates these requests to ensure timely and accurate handling. By efficiently handling API requests, the system ensures seamless communication between the call center and the organization, facilitating effective customer service and support.To utilize the voice recognition system as a web service, the software development kit (SDK) of the system needs to be installed and integrated into the web server. The customer organization's systems, such as CRM, ERP, and ticketing systems, can communicate with the voice recognition system by exchanging data through web services. This integration enables the seamless transfer of customer information and facilitates the utilization of voice recognition capabilities within the organization's existing infrastructure. By leveraging web services, the voice recognition system can be seamlessly incorporated into the customer organization's systems, enhancing their functionality and enabling efficient data exchange.

3.4 Process definition

Based on an analysis of the overall system architecture and system requirements, the process of integrating voice recognition to extend the call service can be outlined as follows:

3.4.1 The process of registering the user: The process of registering the user's voice in the voice recognition system and creating a template to be associated with the user's information can be implemented through three approaches. These approaches include recording the conversation stored in the call center system through the application used by the company's customers, as well as creating a voice model when the user initiates a call to the company's did number. The process of creating a voice model through a user call:

- Customer call to the service provider organization
- Choose create a voice template IVR
- Call center notify start audio recording
- Server save audio record to file
- Check Antispoof
- Response antispoof result
- If successfully checked revoke create voice template API
- Response api result by audio
- Hang up call

3.4.2. The process of verifying the real-time voice recording is as follows:

- Customer call to the service provider organization
- Choose create a voice template IVR
- Call center notify start audio recording
- Server save audio record to file
- Revoke anti spoof API
- Response anti spoof API result
- Play result of matching API response
- Customer listen response and hang up the call

3.4.3 The process of identifying a user by comparing their voice pattern involves the following steps:

- Customer call to the service provider organization
- Choose create a voice template IVR
- Call center notify start audio recording
- Server save audio record to file
- Revoke anti spoof API
- Response anti spoof result
- If successfully checked revoke create voice template API
- Response api result
- If successfully created voice template revoke match customer 2 voice templates API
- Response api response
- Play result of matching api response
- Customer listen response and hang up the call

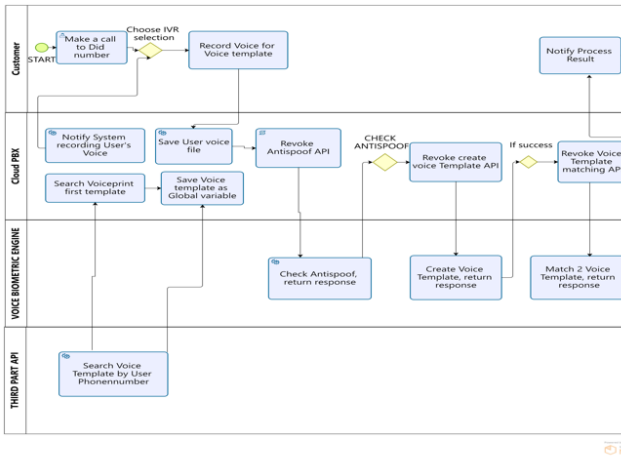

Figure 2. User authentication by voice process

4. Recommended solution

In the implementation of the customer identification solution, the design of the customer identification system was formulated by integrating the call center system with the user organization system, incorporating advanced technologies, and considering the overall requirements and objectives of the solution.

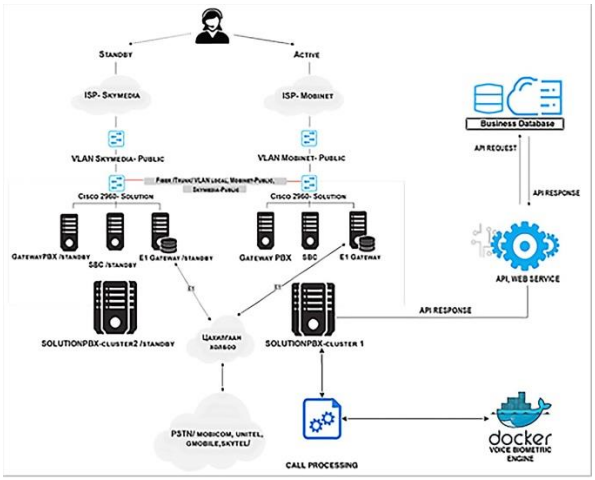

Figure 3. Solution Detailed Architecture

Call center system: Call center systems are designed to operate in both active and passive modes, ensuring uninterrupted functionality. The active version serves as the core component of the solution, while the standby version acts as a backup in the event of a failure in the main system, ensuring continuous system operation. In both the active and standby scenarios, the call center system comprises the following components:

- Client – User;
- ISP Mobinet or Skymedia - Internet network;
- VLAN Mobinet or Skymedia - public - network switch;
- CISCO 2960 Solution - a switch to connect the solution to mobinet or skymedia network;
- Gateway PBX – Mobicom, Unitel, Skytel, G-Mobile mobile numbers are registered and peered into the solution;
- SBC (Session Border Controller) performs the role of secure transmission of calls or conversations;
- E1 gateway - is responsible for providing speech transmission line or multilevel speech transmission;
- Telecommunications: Provide a landline telephone line;
- Solution PBX main server or server where conversations take place. When the server receives a call, it captures the user's actions, invokes the relevant web service from the third part system, executes a script to convert the received response into speech, and delivers the returned value in JSON format from the web service to the user as speech.

In the case of a user organization system, it is not recommended to send data directly to the system database from external or separate systems, as it poses security risks. Instead, it is advisable to transfer data through an API web service that is protected by an API key or token.For this purpose, it is recommended to utilize the AWS cloud service. AWS offers a wide range of services with various options that can meet the requirements of the system. Behind the AWS API Gateway, there is the AWS Lambda function, which represents a serverless computing service. The Lambda function processes incoming parameter values through the API gateway, performs searches in the database, and generates a response to the request. This serverless architecture enables efficient and scalable processing of API requests without the need for managing traditional server infrastructure.The call service we provide, based on cloud computing, falls under the category of Platform as a Service (PaaS). The web service, which processes data from the user's enterprise system and returns the response, belongs to the Infrastructure as a Service (IaaS) type, thanks to the utilization of the AWS cloud. Consequently, our solution leverages a combination of PaaS and IaaS to deliver its functionality.

API, Web service: The implementation of the AWS web service utilizes several AWS services, including AWS EC2, AWS API Gateway, and AWS Lambda. These services are used to achieve different functionalities within the web service. Specifically, the web service includes a search functionality based on the

user's phone number, a functionality to create a voice template for user voice authentication and store it in the database, a functionality to store the user's voice authentication result in the database, and a functionality for logging user activity. Each of these functionalities is implemented through dedicated web services integrated into the overall solution.

Voice biometric engine: The IDRND technology enables voice recognition for user identification. This technology offers the flexibility to be installed on a company's own server using the SDK version. One notable feature is that the technology does not store the user's voice used to create their voice model on the system. In a study conducted in 2022, IDRND ranked first in terms of its ability to detect voice interference and forgery. Due to its advanced capabilities, this technology has been proposed as a viable option for implementing a voice recognition system in call center systems.

5. Conclusion

In this research work, a comprehensive solution was developed to extend an organization's information system using cloud-based services. The solution was specifically modeled and demonstrated in the context of a call service. The aim was to provide a practical framework that can be applied to other service domains within an organization.

Through the integration of the call center system, users can access the organization's system using voice authentication. This enables them to not only access their information but also make necessary updates or modifications to their data within the system. The voice authentication process serves as an additional layer of security and convenience for users interacting with the system.

Our solution is designed to cater to customer service organizations, specifically through call centers. By implementing our solution, these organizations can enhance their customer service capabilities and improve the overall experience for their customers. The integration of voice authentication and other cloud-based services can streamline operations, enable secure access to customer information, and facilitate efficient communication between agents and customers.

Business organizations that already have CRM, ERP, and other record-keeping information systems can greatly benefit from adopting cloud-based call services. By integrating these services into their existing systems, organizations can expand their information systems, support electronic migration, and provide enhanced services to their customers. Cloud-based call services offer advantages such as improved accessibility, reduced costs, and streamlined operations. Customers can save time by accessing services through various devices, and organizations can optimize their human resources by leveraging the efficiency and scalability of cloud technologies. Overall, implementing cloud-based call services enables businesses to enhance their information systems, improve customer satisfaction, and achieve greater operational efficiency.

In the future, there is a potential for extending the voice recognition solution to include text-based verification as well. This expansion would allow for the verification and authentication of users not only through their voice but also through text-based inputs such as passwords, PIN codes, or other textual identifiers. By incorporating text-based verification alongside voice recognition, organizations can offer users more flexibility in how they authenticate their identity. This expansion could enhance security measures and provide additional options for user authentication, depending on the specific needs and preferences of the users and the organization.

6. References

[1] Shawish, A., & Salama, M. *Cloud computing: Paradigms and technologies. In F. Xhafa, & N Bessis (Eds.), Inter-cooperative collective intelligence: Techniques and applications, studies in computational intelligence 495*, Springer-Verlag Berlin Heidelberg, 10.1007/978-3-642-35016-0 2, 2014, pp. 39–67

[2] Abhishek et al, “A Strategic Value Appropriation Path for Cloud Computing”, *Journal of Management Information Systems* 35(3):740-775, Sep 2018

[3] Herhalt, J., Cochrane, K. (2012). Exploring the Cloud: A Global Study of Governments‘ Adoption of Cloud. https://images.forbes.com/forbesinsights/StudyPDFs/exploring-cloud.pdf, Mongolia, Oxford Policy Management

[4] Sajjad A. B., Awais A, *Challenges in education management information system indeveloping countries, Information and Emerging Technologies* (ICIET)-2010 International Conference. Karachi, Pakistan, 2010

[5] Gary Garrison, Robin L, Wakefield, Sanghyun Kim, “The effects of IT capabilities and delivery model on cloud computing success and firm performance for cloud supported processes and operations”, *International Journal of Information Management* 35 (2015) 377–393, 2018

[6] Mrs. Ashwini Sheth, Mr. Sachin Bhosale, Mr. Harshad Kadam, *Research paper on cloud computing*, (ISSN 2231-2137), 2021