Voice Driven AI-based Automated Voip PBX for SMEs

Ryan Fernando
Faculty of Science and Engineering
University of Plymouth
Plymouth, United Kingdom
ryanmario.fdo@gmail.com

Chamara Disanayake
Department of Software Engineering and Computer Security
Faculty of Computing, NSBM Green University
Homagama, Sri Lanka
chamara.d@nsbm.ac.lk

Chamindra Attanayake
Department of Software Engineering and Computer Security
Faculty of Computing, NSBM Green University
Homagama, Sri Lanka
chamindra.a@nsbm.ac.lk

***Abstract*— In the current telecommunications industry, conventional keypad-operated Interactive Voice Response (IVR) systems present usability issues, such as complex menu configurations, prolonged handling durations, and poor user experiences. This paper introduces a voice-driven Artificial Intelligence-based automated VoIP PBX communication platform that transforms traditional VoIP PBX systems. The system is developed with Asterisk as the primary telephony engine and combined with Google Cloud's Speech to Text (STT) API, allowing users to engage through natural speech commands instead of navigating through complex menu hierarchies. The implementation includes a custom dial plan that records voice commands, transcribes them into text, recognizes keywords, directs calls based on established logic for immediate decision making, and performs AI-based automated tasks. Testing conducted in virtual machine environments demonstrated improved responsiveness and decreased call handling times, indicating enhanced user interaction despite issues associated with network dependency and voice recognition precision in noisy environments. This solution establishes a foundation for further enhancements, including customer database integration, multilingual support, and advanced speech recognition modules, ultimately creating more efficient and natural communication in telephony environments. Several performance metrics and comparative outcomes like Average Response Time, Intent Recognition Accuracy, etc of this system and traditional system will be evaluated in future work to assess the effectiveness of the proposed PBX system against the traditional PBX system.**

Keywords*— *Interactive Voice Response (IVR), Voice over Internet Protocol (VoIP), Private Branch Exchange (PBX), Artificial Intelligence, Natural Language Processing (NLP), Asterisk

I. INTRODUCTION

The use of Artificial Intelligence (AI) is transforming sectors into the fast-paced digital era we live in today. The integration of Artificial Intelligence has transformed operations, consumer interaction, and decision making in a variety of industries, including healthcare, financial services and banking. However, the adaptation of AI technology has not been fully utilized in the telecommunications sector, especially in the era of voice communication systems. In VoIP Private Branch Exchange (PBX) systems, traditional customer engagement techniques rely on static, menu-driven Interactive Voice Response (IVR) menus, which frustrate customers.

To modernize client engagement in telecommunications, this report presents the development of a voice-driven AI-based automated VoIP PBX. The report investigates how customers can communicate with a PBX system using straightforward, natural voice commands instead of navigating intricate IVR menu trees by utilizing cutting-edge technologies like speech recognition and natural language processing.

The limits of the existing IVR systems and the potential for AI to improve communication have served as an inspiration for this project. As part of the project, a system that can comprehend vocal requests intelligently, interpret them, route calls appropriately, and guide users with AI-based automated tasks. With this user-centered approach, inflexible menu systems are giving way to more adaptable, human-like interactions.

The problem statement, determination, and main project goals are described in this report. The implementation approach, challenges faced, testing methods, and results are presented after a technical review of the frameworks and tools utilized. The paper also addresses the importance of incorporating AI-based automation into PBX systems and looks at various other possibilities for further study and development of this system.

The goal of this research is to show that significant changes are achievable even in sectors that are reluctant to embrace AI. This approach improves user experience, speeds up call processing, and opens the door for more sophisticated, voice-driven telecommunication systems by incorporating Artificial Intelligence into a conventional PBX environment.

A. Problem Identification

Traditional PBX systems provide significant challenges for both customers and service providers due to their reliance on sophisticated Interactive Voice Response (IVR) menus and the need for human operators or customer agents. For clients, scrolling through several IVR menus is excessively time-consuming, frequently leading to unnecessary delays in acquiring services or information. This is especially problematic in emergencies where immediate prompt response is essential. The extended time of these encounters frequently results in increased call charges, which creates additional financial strain for clients without providing proportional value or satisfaction.

The technical complexity of traditional IVR systems in current PBX systems presents a significant obstacle to efficient communication. Many clients fail to understand the

complex vocabulary used in the prompts, resulting in confusion and a sense of inadequate service. This challenge is amplified by the naturally irritating experience of sequentially selecting from various menu alternatives, especially when the desired service shows only after navigating through numerous preliminary choices. Even after investing significant amounts of time browsing through menus, users often discover that the available selection fails to address their specific requirements or concerns.

The limitations of traditional methods frequently result in recurrent calling patterns. Dissatisfied with first outcomes, clients typically redial to test alternative menu selections, hoping to achieve more satisfactory results. This cyclical pattern contributes to a continuous impression of limited engagement and interaction, ultimately leading to customer boredom, irritation, and potentially service provider abandonment. When the customer expectations remain consistently unmet, many opt to transition to alternative service providers offering more efficient communication options.

The procedure of making appointments or reporting complaints creates more issues in the current PBX systems. Customers often must browse through many IVR menus before eventually connecting with a human operator to fulfil these crucial tasks. This human-dependent procedure presents multiple inefficiencies and areas of potential failure. Operators may be unavailable during high call periods, resulting in longer wait times that annoy customers and postpone issue resolution. When connections are established, confusion between customers and operators usually leads to erroneous data entry, appointment scheduling mistakes, or inadequate complaint filing. The manual nature of these encounters also produces variability in service quality, as client experiences vary greatly based on individual operator knowledge, communication skills, etc. Furthermore, the limited availability of human operators during peak business hours limits customers' ability to make appointments or file complaints at their convenience, establishing artificial constraints that contrast with modern expectations for 24/7 service availability.

From the service providers' perspective, traditional PBX systems generate major operational issues. Servers experience traffic overload due to prolonged or repeated calls from unsatisfied customers, affecting overall system performance and potentially degrading service quality for all users. The financial consequences are also significant, where big investments in conventional system development and maintenance offer fewer benefits when customer satisfaction levels are poor. Additionally, the training requirements for customer service agents and human operators contribute tremendous expenses and logistical complexity to organizational operations.

Most importantly, the credibility and dependability of service providers are seriously threatened by current traditional PBX systems. Systems that frequently irritate and frustrate customers have a direct negative impact on customer loyalty and brand reputation.

B. Objectives

The primary goal of this research project is to develop an intelligent system that enhances customer service by automating critical tasks traditionally performed by human operators in current PBX systems. This allows us to improve customer communications and interactions by offering an AI-based automated experience that directly addresses customer requirements. The objectives in this system are defined as follows:

- Implemented Speech to Text integration using Google Cloud's Speech to Text (STT) API, enabling real-time translation of spoken commands into text for further analysis and decision-making processes for the system.
- Developed an intelligent call routing system that analyses customer requests and determines their intentions through Natural Language Processing (NLP) capabilities, directing calls to the most appropriate department or personnel.
- Implemented AI-based automated support functionalities that manage routine customer service tasks without human agent intervention, providing instructions and responding to frequently asked questions (FAQs).
- Implemented an automated complaint registration system capable of recording and processing verbal customer complaints, converting speech to text and storing this complaint information on the service provider's complaint handling platform (webpage).
- Developed an AI-based automated appointment booking system featuring an interactive voice-based module that facilitates structured dialogue between customers and the system, capturing spoken responses regarding preferred date, time, service type, and personal information, etc. and finally storing this information on the service provider's booking handling platform.

The integration of these objectives creates a comprehensive solution that addresses the common limitations of traditional PBX systems while leveraging current AI technologies to create a more responsive, efficient, and user-friendly communication platform.

II. LITERATURE REVIEW

In recent years, VoIP PBX systems have drastically transformed organizational communication infrastructures. According to R. Bryant and Akanksha, Asterisk, due to its open-source adaptability [1-2], have achieved widespread adoption as a leading PBX solution [3]. Despite technological improvements in numerous areas [4], many organizations continue to rely on old PBX systems for call routing and basic functionality, which fundamentally lack intelligent interaction capabilities. The integration of Artificial Intelligence and Natural Language Processing in telephony is still in its early stages, which offers significant potential for creating more innovative, voice-driven user experiences.

Several research projects have studied the fundamental elements of VoIP PBX systems. Studies such as "Wireless Enabled Voice over Internet Protocol (VoIP) Network Application Using Asterisk PBX [5]" provide core knowledge on developing VoIP PBX systems using Asterisk. However, these experiments mostly focus on creating basic call handling and networking features rather than applying AI for enhanced interactions. [6]. This research gap underlines the need for more complex approaches that utilize modern computing capabilities.

More recent experiments have begun studying AI assistant integration inside PBX systems. The research "Design and implementation of a VoIP PBX integrated Vietnamese virtual assistant: a case study [7]" describes one such technique through the development of a RASA chatbot to support users in understanding and routing queries. This research introduced NLP-based call handling utilizing a conversational assistant architecture. Although there are basic similarities to the current project, their implementation primarily depended on external chatbot platforms. [8], which introduced added complexity and potential latency difficulties.

In contrast, this project extends the typical Asterisk PBX design by embedding AI-based capabilities directly inside the Dial-plan, avoiding dependency on third-party chatbot solutions. The system leverages Google Cloud's Speech to Text API primarily for transforming spoken client input into textual data, while integrating NLP functions such as keyword filtering and intent matching directly using the REGEX tool within the Asterisk Dial-plan. This architecture approach supports rapid understanding of user intentions for call routing and other activities while simultaneously decreasing system complexity, processing time, and resource needs by removing additional external AI platform dependencies.

Furthermore, both traditional systems and previous research have neglected the importance of sustained conversational interaction in telephony contexts. This project addresses this limitation by introducing two additional AI-based automated features that enable users to independently book appointments and file complaints through natural voice interactions. These capabilities represent significant advancements over existing implementations by providing comprehensive voice-based alternatives to traditional human operator-dependent processes.

The literature review demonstrates a development from simple PBX deployments toward increasingly intelligent systems. However, most existing solutions either lack AI integration totally or implement it through external dependencies that impose extra complexity. This research contributes to the field by proving that successful AI integration may be performed directly within Asterisk's Dial-plan, enabling a more responsive and efficient system while keeping architectural simplicity.

III. METHODOLOGY

A. Methodology

This section covers the methods adopted for designing the voice driven AI based automated VoIP PBX system, aimed at transforming traditional telecommunication experiences through natural language processing and speech recognition technology. The development process followed Agile methodology principles to ensure adaptability and responsive execution throughout the project lifecycle.

The Agile methodology was selected for this project due to its inherent flexibility, support for incremental modifications, and capacity to react to new technological issues. This technique proved particularly beneficial given the complicated integration of voice processing technology with traditional PBX infrastructure, allowing for continual modification based on testing outcomes and growing requirements.

Development continued through four planned sprints, each targeting various parts of the system functionality and integration. The initial sprint, covering three weeks, focused on developing the foundational environment and architectural layout. This step entailed building a virtualized development environment, setting up important system service dependencies, and developing thorough architectural diagrams to illustrate the communication flow between user voice inputs and system responses. Additionally, basic planning was undertaken for integrating telephony features, user interface processes, and speech processing modules, forming the project's technical foundation.

The second sprint focused on core functionality development during a two-week timeframe. This phase prioritized implementing key call handling capabilities and basic interactive elements within the Asterisk environment. Preliminary call routing techniques were built and tested to ensure essential system functionality before advancing to more advanced capabilities. This stepwise approach allows for early identification and resolution of any design restrictions.

Speech processing integration was the third and most intensive sprint, involving six weeks of focused development. During this vital step, the Asterisk server was connected to Google Cloud's Speech to Text API, offering real-time voice transcription capabilities. A dynamic call routing Dial-plan was built based on keyword and phrase recognition using REGEX, boosting the system's capacity of identifying user intentions. This sprint also built sophisticated logging methods and fault detection protocols. [9] to facilitate system diagnostics and troubleshooting. Automatic service initiations were programmed to prevent errors following system reboots, while fault tolerance measures were developed to address probable network outages and service interruptions.

The final sprint, covering four weeks, focused on system optimization and refinement through rigorous testing and performance advancements. Performance testing was conducted across numerous end devices using realistic call simulations to validate system responsiveness and correctness. Data extraction techniques from Asterisk to web interfaces were evaluated to ensure consistent information flow, while web interface designs were upgraded to better data presentation and user interaction. Automation scripts underwent modification to optimize system reliability and performance, with final adjustments performed based on testing outcomes and identified possibilities for improvement.

Throughout the development process, continual testing and assessment guaranteed that each component functioned correctly both individually and as part of the integrated

system. The methodology emphasizes scheduled sprints with specific deliverables, providing a framework for systematic progress while keeping the flexibility essential to manage technical issues as they develop during the implementation phase.

B. *High-level Architectural Diagram*

The image below (Fig 1) is the high-level architectural representation of this system. The customer interacts with the system through a mobile phone or SIP phone, which connects to the Asterisk PBX deployed on a virtual machine in the cloud. Voice commands from the customer are processed through the Asterisk Dial-plan, which serves as the core control module for the system. Speech input is sent to the Speech-to-Text API for conversion into a textual format for analysis. Based on the interpreted intent, the system can route calls to specific individuals or direct them to appropriate departments, or activate specialized service modules including automated FAQs, complaint filing, and appointment booking. This architecture eliminates traditional IVR menu navigation by creating direct pathways between natural language inputs and system responses, enhancing user experience while maintaining capabilities for human intervention when necessary. The system does not support inbound or outbound calls since a SIP trunk was not obtained from an internet service provider. Nevertheless, if SIP trucking were to be enabled, the system would continue to operate in accordance with procedures and functionalities that are already described in this system.

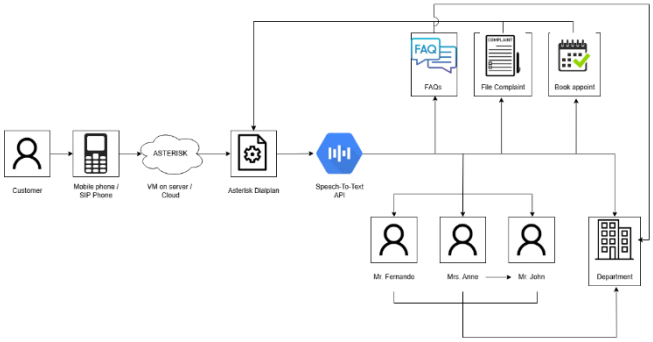

Fig. 15. High-level Architectural Diagram

C. *User Features*

Service providers' system administrators have comprehensive access to the environment and configuration infrastructure through Asterisk's Dial-plan (extensions.conf), enabling customization according to organizational requirements. This centralized configuration approach allows for efficient adaptation of call routing rules, speech recognition parameters, and automated response behaviours without requiring extensive code modification.

The implementation gives the administrator access to Google Cloud's dashboard for monitoring Speech-To-Text API performance indicators, usage statistics, and managing associated payment arrangements. This connection permits data-driven optimization of speech recognition parameters while retaining transparency in use and operating expenses.

All the customer-generated data, including verbal complaints and appointment bookings handled through the interface, are consistently logged in the service provider's web portal, where only authorized individuals can access and delete if required.

D. *Implementation*

For the development of the system, the structure involves the implementation of several key tools and technologies. The main goal of this system proposal is to implement an interactive PBX system, where a user gets a natural human-like interaction, by preventing the need for DTMF (Dual-Tone Multi-Frequency, scrolling through IVR menus and human operators.

Asterisk is the core of the system, which is an open-source telephony software which controls SIP configurations. The dial plan in Asterisk, which is extensions.conf is conFigd accordingly to handle calls, and for the integration of AI functionalities for AI-based automated tasks. It also includes verbose logging to provide logs to observers for failures or errors, and also where fallback mechanisms were implemented. AEAP (Asterisk External Application Protocol) module and Node.js service help Asterisk to communicate with external services integrated into the system to achieve the proposed functionalities. Basic configurations for assigning users, authentications and IP phones in the network are done in Asterisk's sip.conf configuration file.

To handle speech processing, a simple lightweight Node.js server was developed, which listens to port 9099. For the connection between Asterisk and the Google Cloud service, Node.js acts as the middleware, forwarding the audio to the speech-to-text transcription module.

The tool used for converting users' speech into text in order for the system to process users' requirements and enabling voice interactivity is Google Cloud's Speech to Text API. The transcription process provides real-time transcription processing with an accuracy of 90% to provide seamless interactivity.

The webhosting module, for the purpose of customer complaints and customer appointment bookings monitoring, is done by the Apache web server, which is used to host complaints and book webpages locally. The relevant data are extracted from Asterisk and are shown in the mentioned webpages through Asterisk Gateway Interface (AGI), PHP automation scripts.

IV. Observations And Discussions

In the testing phase for the system, the process was planned and conducted to evaluate the overall performance of the system and its dependability in managing its main functions. The Asterisk Dial-plan was validated, which helps to provide a foundational PBX system, also by making sure that the system could function smoothly, understanding users' requirements and enabling the user to have interactive experience in a variety of ways possible. To ensure that the system operated as planned and successfully, specifications were established throughout the development and design stages.

Testing was accomplished by running a range of testing scenarios that mimic real-world use cases, also by guaranteeing a thorough evaluation of the functionalities of the system. The testing scenarios were designed which cover all the vital functionalities. The testing procedures go beyond confirming the system's technical functionalities but also provide insightful insights on the responsiveness and usability of the system for end user experience, mimicking real-life scenarios.

The end-to-end process from the user initiating the call to the system, the capacity to gather, analyse and save data was the main priority of the testing process. The system's transcription tool's capabilities were tested by providing special considerations, but it is the main tool for comprehending and analysing the user's input. The effectiveness of the system's functionalities was heavily dependent on the ability to extract useful and actionable information from voice interactions.

In addition, an analysis of the connection between web-based dashboards for complaints and bookings was thoroughly conducted. The dashboards allow system administrators to view and handle clients' contacts by visualizing the data that was extracted during the calls. At this testing stage, a verification process was carried out to ensure that the recorded data was correctly documented and displayed, to ensure seamless connectivity flow between the data from the dashboard and the PBX system.

Everything taken into consideration, the system's testing phase was essential to make sure that the system is able to provide a seamless and less complex experience for the user [10]. Identification of faulty errors and error fixing were done during the testing phase to ensure a rigid and dependable solution.

A. Automated Call Handling Test

A test was conducted on the system's ability to understand users' requirements by using implemented Natural Language Processing (NLP) functionalities for handling calls efficiently and for conducting automated tasks. A test was done by calling the extension (550), where the user was required to call an individual called "Ryan Mario", and with the backend processing, the system was able to understand and directly forward the call to that relevant individual's assigned extension. The image below shows the test which was done using two IP softphones [11] to test automated call handling using NLP functionalities.

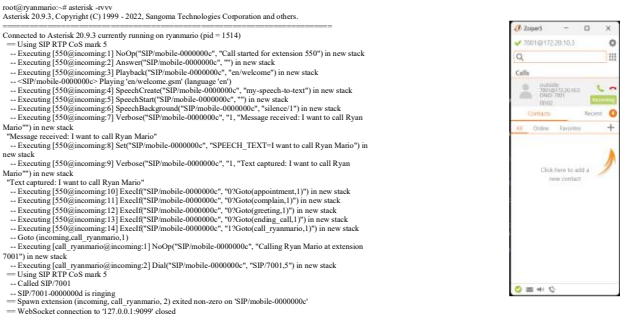

Fig. 16. Automated Call Handling Test

B. Complaint Logging Test

A test was conducted to simulate the process of filing a complaint through the system. Firstly, the user dialed the system's extension (550), and after that, the system was able to understand that the user wanted to file a complaint. The system then continued to follow the automated process that was designed in order to file a complaint. The user was requested to explain the complaint, and from there, the system did its part by storing and displaying the complaint on the service provider's complaints webpage.

The image below shows the overview of the complaint-filing process testing conducted in the Asterisk CLI [12].

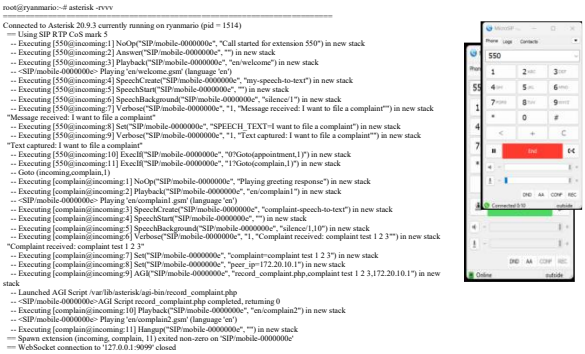

Fig. 17. Complaint Logging (Asterisk CLI Test)

The image below shows the complaint that was filed and is now shown on the Complaints Dashboard webpage.

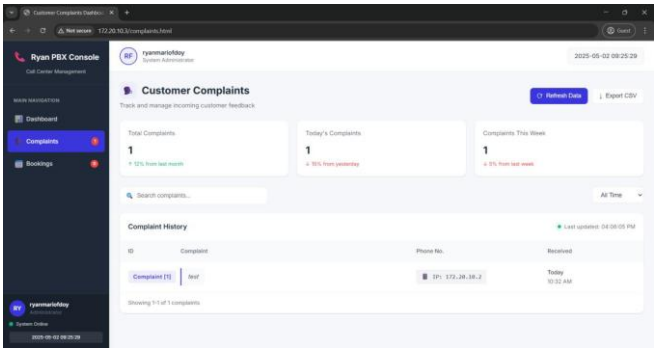

Fig. 4. Complaint Logging (Webpage Test)

C. Appointment Booking Test

The second testing scenario was booking an appointment through the system. Once the user initiated the call and gave the requirement to book an appointment, the system follows the designed automated process for booking an appointment, which involves a back to back questionnaire where the system is designed to ask several relevant questions and for each response from the user for each question is stored, analysed and finally shown in the appointment bookings dashboard.

The image below shows the overview of the bookings process test conducted in the Asterisk CLI.



Fig. 5. Appointment Booking (Asterisk CLI Test)

The image shown below shows that the appointment that was booked is now shown in the appointment bookings webpage.

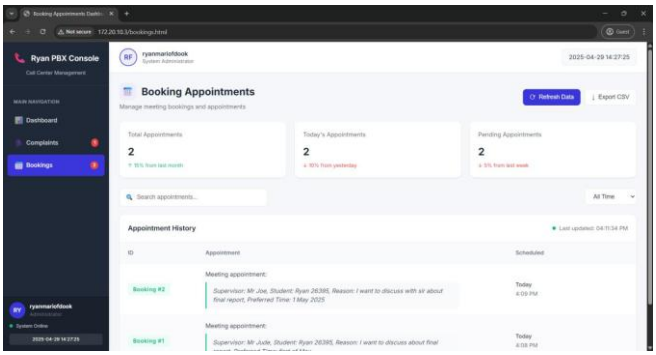

Fig. 6. Appointment Booking (Webpage Test)

V. CONCLUSION

A. Conclusion

The conversion of a conventional PBX into a voice-driven, Artificial Intelligence-based enhanced telephony infrastructure has shown notable improvements in both operational efficiency and user experience. Conversational voice commands replaced traditional menu-based navigation (IVRs), enabling callers to express demands without the need for manual keyboard input, such as making appointments, filing complaints, or contacting particular departments. This change led to shorter call handling times, less user annoyance, and an interface that was easier to use for people with different backgrounds of technical proficiency.

In conclusion, the study effectively illustrated that integrating AI-based speech interfaces into an open-source phone platform is feasible. In addition to streamlining user interactions, a conversational approach to call routing creates a scalable, expandable basis for upcoming advancements in intelligent communications.

B. Future Works

Several areas for future work have been identified to further increase the system's capabilities. The use of advanced AI technologies, including conversational AI and machine learning based requirement identification, could facilitate enhanced automation and tailored client interactions. Integrating the customer database will allow effortless access to user histories and preferences, facilitating personalised responses and enhanced service management.

Multilingual support is a vital aspect, as it expands the system's applicability worldwide and improves accessibility. The implementation of advanced speech recognition modules which is capable of handling different accents and noisy environments would further improve the accuracy and reliability of voice-based interactions. Such enhancements will improve robustness of the system, reliability, intelligence and a universally accessible communication platform.

Further, another important aspect to be considered in running a system of this nature is implementing a strong security mechanism, which can withstand any security related threats in any vulnerable environment. A security threat that could happen is a dial plan injection, where an attacker can misconFig the dial plan (extensions.conf) for malicious call routing. According to Pelayo Nuño and Carla Su´arez [13], an automated diagnosis engine can be implemented that parses the Asterisk dial plan for syntax or structural errors and vulnerabilities like use of unsafe Asterisk functions like (SHELL, RECEIVEFAX, CONFBRIDGE) and dial plan injection patters using _X, or _X! If the dial plan contains unsafe patterns, it will then automatically rewrite into a secure version or if an unsafe function is found, it will log a warning to the administrator. This diagnosis engine not just detects problems; it also fixes them automatically without the need for manual human intervention.

REFERENCES

[1] R. Bryant, L. Madsen, and J. Van Meggelen, "AsteriskTM: The Definitive Guide." [Online]. Available: www.it-ebooks.info

[2] Akanksha, Y. Surya, R. Kumar, and V. Kumar, "Cloud-Hosted IP-PBX System with Asterisk and Oracle Cloud: A Scalable and Cost-Effective Solution for Business Communications," in *2024 2nd International Conference on Disruptive Technologies, ICDT 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 837–841. doi: 10.1109/ICDT61202.2024.10489415.

[3] Paul. Mahler, *VoIP telephony with Asterisk*. Signate, L.L.C., 2004.

[4] S. A. Grushko, A. P. Pshenichnikov, E. E. Malikova, and A. Y. Malikov, "Virtual Asterisk IP-PBX Operation Studying and Exploring at the University," in *2022 Systems of Signals Generating and Processing in the Field of on Board Communications, SOSG 2022 - Conference Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/IEEECONF53456.2022.9744336.

[5] E. Odjidja, S. Kabanda, W. Agangiba, and R. Annan, "WIRELESS VOIP IMPLEMENTATION USING ASTERISK PBX AND OPEN SOURCE SOFTPHONE," European Alliance for Innovation n.o., Sep. 2018. doi: 10.4108/eai.21-3-2018.2275638.

[6] P. Nuño, F. G. Bulnes, S. Pérez-González, and J. C. Granda, "Asterisk as a Tool to Aid in Learning to

Program," *Electronics (Switzerland)*, vol. 12, no. 5, Mar. 2023, doi: 10.3390/electronics12051160.

[7] H. S. Hoang, A. K. Tran, T. P. Doan, H. K. Tran, N. M. D. Dang, and H. N. Nguyen, "Design and implementation of a VoIP PBX integrated Vietnamese virtual assistant: a case study," *Journal of Information and Telecommunication*, vol. 7, no. 2, pp. 201–226, 2023, doi: 10.1080/24751839.2023.2183631.

[8] R. Kumar Sharma and M. Joshi, "An Analytical Study and Review of open Source Chatbot framework, RASA." [Online]. Available: www.ijert.org

[9] P. Montoro and E. Casilari, "A comparative study of VoIP standards with asterisk," in *Proceedings - 2009 4th International Conference on Digital Telecommunications, ICDT 2009*, 2009, pp. 1–6. doi: 10.1109/ICDT.2009.8.

[10] P. Nuno, C. Suarez, E. Suarez, F. G. Bulnes, F. J. Delacalle, and J. C. Granda, "A Diagnosis and Hardening Platform for an Asterisk VoIP PBX," *Security and Communication Networks*, vol. 2020, 2020, doi: 10.1155/2020/8853625.

[11] Jim. Van Meggelen, Jared. Smith, and Leif. Madsen, *Asterisk: The Future of Telephony, 2nd Edition Van Meggelen, Jim*. O'Reilly Media, Inc., 2007.

[12] M. S. Mack and A. C. Rhodes, "The Asterisk Handbook Version 2," 2003. [Online]. Available: http://www.openoffice.org.

[13] P. Nuno, C. Suarez, E. Suarez, F. G. Bulnes, F. J. Delacalle, and J. C. Granda, "A Diagnosis and Hardening Platform for an Asterisk VoIP PBX," *Security and Communication Networks*, vol. 2020, 2020, doi: 10.1155/2020/8853625.