Implementation of '*ASR4CRM*': An Automated Speech–enabled Customer Care Service System

Aderemi A. Atayero, Charles K. Ayo, Nicholas Ikhu-Omoregbe and Ambrose Azeta, *Members IEEE*

Abstract — **The main disadvantage of human presence in the Call centers of GSM service providers is poor response time. The preference of IVR services by Nigerian subscribers can be attributed to this fact. A system has been developed on the VoiceXML platform to serve as a *panacea* for this problem. The developed system '*ASR4CRM*' obviates human-to-human interaction in the complaint lodging and solution provision process, by replacing it with human-to-system interactivity. ASR4CRM has a 3-tier architecture. The telephone system constitutes the first tier, the VoiceXML gateway and the web server constitute the middleware, while the database constitutes the third tier. The system was tested with the top twenty-four FAQs from a leading Nigerian GSM carrier (MTN) and successfully deployed on Voxeo voice server. It has the attendant benefit of improved CRM.**

Keywords — **Automatic Speech Recognition, Customer Care Service (CCS), Customer Relationship Management (CRM), FAQ, Interactive Voice Response (IVR), Speech-enabled, Speech-to-Text (STT), Text-to-Speech (TTS), VXML.**

I. INTRODUCTION

The deregulation of the telecoms industry in Nigeria in 2001 boosted the rate of diffusion of telephone lines. It is solely responsible for the phenomenal growth in the sector from about 450,000 lines in 1999 to 57 million by December 2007, boosting the teledensity growth from 0.4% to 41% [1].

The huge number of telephone users all over the world, with an estimate of two billion fixed and mobile phones users have prompted the World Wide Web Consortium (W3C) speech interface framework to create an avenue for users to interact with web-based services via key pads, spoken commands, Interactive Voice Response, synthetic speech, and music. The W3C speech interface framework incorporates Voice eXtensible Markup Language (VoiceXML or VXML), speech synthesis markup language (SSML), speech recognition grammar specification (SRGS), voice browser call control XML (CCXML) and

semantic interpretation for speech recognition (SISR). VoiceXML controls how applications interact with a user through interactive voice response; the SRGS offers support for speech recognition used by developers to describe end users responses to token prompts; the CCXML provides telephony call control support and other dialog systems, while the SISR defines how speech grammars bind to application semantics [2].

The W3C speech interface framework is presented in [3]. Automatic speech recognizer (ASR), which is a component part of the VoiceXML Gateway, accepts speech from the user and produces text. It uses statistical grammars generated from large *corpora* of speech data based on the speech grammar makeup language (SGML). The dual tone multi-frequency (DTMF) or touch-tone, accepts touch-tones produced by a telephone whenever a user presses any key on the keypad. The language understanding component is responsible for extracting semantics from a text string using a pre-specified grammar; and the context interpreter enhances the semantics from the language understanding module by obtaining context information from a dialog history. The interface framework incorporates a dialog manager that prompts users for input, determines the requests and acts in accordance to the dialog script specified in VoiceXML. Similarly, there is a media planner is responsible for determining the format of output to be presented to the user, either as synthetic speech or pre-recorded audio.

Similarly, the recorded audio player module replays prerecorded audio files to users; and the language generators accepts text input from the media planner and presents it as audio output to user via the text-to-speech (TTS) synthesizer. Generally, VoiceXML enables the integration of voice services with data services based on the client-server paradigm [4]. A voice service is viewed as a sequence of interactive dialogs between a user and an implementation platform. Suffice it to note that a high quality of speech user interface can only be achieved through iterative design and evaluation [5].

Recent advances in ICT, particularly the world-wide web, has expanded the reach and mode of business transactions as well as incorporated customer interaction data, multi-channel communications and one-to-one interactions to deliver seamless, reliable, efficient and transparent services to customers [6], [7].

- A. A. Atayero is with the Department of Electrical and Information Engineering, Covenant University, Ota, Nigeria (phone: +234.805.256.7491; e-mail: atayero@ieee.org).
- C. K. Ayo, N. Ikhu-Omoregbe and A. Azeta are with the Department of Computer and Information Systems, Covenant University, Ota, Nigeria.

The rest of the paper is organized as follows: section 2 presents the objectives of the research, while section 3 presents the research methodology. Sections 4 and 5 present the software deployment architecture and the system implementation respectively. Section 6 presents the discussion and results while the conclusion to the work is presented in section 7.

II. RESEARCH OBJECTIVES

Presently, the Telecoms industry in Nigeria is beclouded with dismally poor quality of service (QoS), which is of great concern to the government, the monitoring agency - The Nigerian Communications Commission (NCC), and the subscribers. Amidst the problems is the inability of the service providers to respond to users' requests appropriately and in a timely manner thus leading to poor CRM. Therefore, the primary objectives of this paper include:

- To design and develop a complaint-lodging system with a view to improving the customer relationship management, customer satisfaction and subsequently customer retention rate.
- To offer a cost effective and efficient system for managing and tracking customers complaints with minimum human intervention.
- To deploy the developed system on a mobile platform for ease and ubiquity of access.

III. RESEARCH METHODOLOGY

The application was developed using tools such as:
- VoiceXML – as the telephony interface, through which a subscriber interacts with the system.
- PHP and Apache – as the front end and middle-ware respectively.
- MySQL – as the database for storing captured user information for personalization of interactive sessions.

The choice for PHP, Apache and MySQL as tools for developing the system is premised on the fact that they are open-source. This renders the developed system readily available for use without concerns related to proprietary issues.

IV. SYSTEM DEVELOPMENT ARCHITECTURE

Figure 1 shows the detailed software architecture of ASR4CRM with the location of each of the service components. The architecture consists of the presentation layer, business logic layer and data layer.

A. The presentation layer

The business logic layer provides access to ASR4CRM through the presentation layer. The services provided in this layer include the authentication and enquiry menu.

B. Business Logic Layer

The business logic layer contains all the functional service components that are performed by the system. They are organized into request and response modules with each having a corresponding sub-module.

C. Data Layer

This layer contains the services that provide data storage and retrieval. It includes a relational database and some customized tables that guarantee application dynamism.
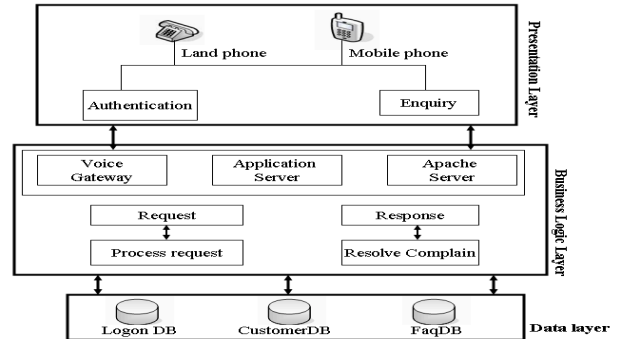

Figure 1. Software Architecture

The deployment architecture is presented in Figure 2, modeled after suggestions given in [8]. The system consists of a Voice Interface, a web server, and a VoiceXML gateway. The voice interface runs on a mobile phone and interacts with the web server and the VoiceXML gateway to make solutions available to the user. The VoiceXML gateway is responsible for communicating with the user over the telephone and performs telephony tasks. The voice server fetches VoiceXML codes, grammar and other files from the application server over the Internet. The main components of a VoiceXML gateway are:
- Voice Browser (Voice User Interface)
- Speech Recognition Engine, also called Automatic Speech Recognizer (ASR), and
- Text to Speech converter, also called (TTS).

The system was deployed and fully tested on Voxeo Community Voice Center [9]. Other platforms that the system can run on are [10] – [13]. The Voxeo™ platform allows many telephone numbers to be generated for testing an application, thereby making it simultaneously accessible to numerous users.
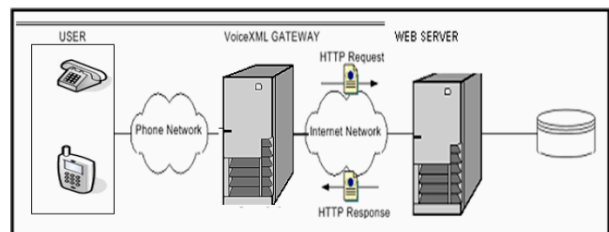

Figure 2. 'ARS4CRM Deployment Architecture

V. System Implementation

The application has been deployed on a Voxeo™ voice server and can be accessed from any mobile phone using the format:

<source country international access code>
+ <destination country code>
+ <destination area code>
+ <voxeo generated 7-digit number>.

The access code is a concatenate of digits making up the four components given above void of spaces or special characters. For example, to access *ASR4CRM* from Nigeria, a user simply dials: *009 1 412 5281595 (or 009 1 206 6078044)* and is promptly connected to the *ASR4CRM* system resident on the Voxeo™ voice server. Automated execution of the application then ensues.

The system prompts the user for a username and password and the interactive session of automated speech-enabled self-help service commences. The default username and password is *<admin>*.

Figures 3a and 3b show a live deployment of the customer care service application on a Nokia 6301 mobile phone.
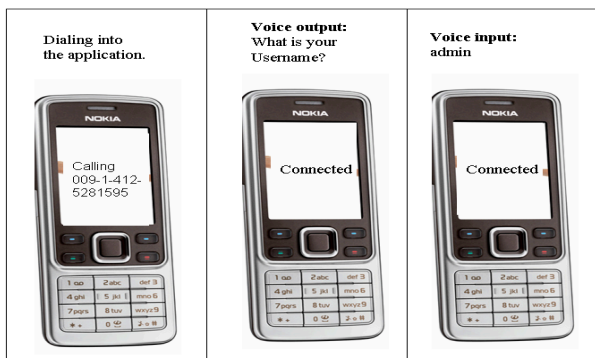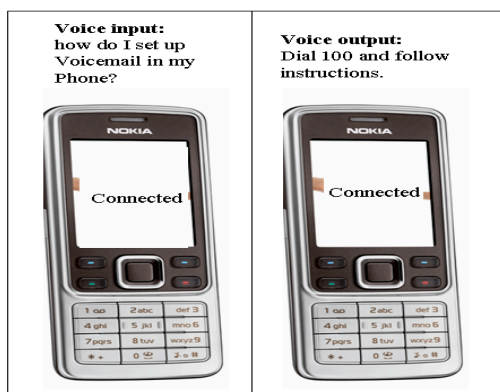

Figure 3a: Authentication session

Figure 3b: Enquiries and response session

VI. Discussion and Result

The system presented in this paper has a call center using speech-enabled Customer Relationship Management (CRM), which connects a user by telephone, 24/7/365 to the customer care information they seek. When a customer in need of a care service dials into the system She/he will have access to front desk information with options stating the problems that are handled by the system. Where a user's complaint is not currently handled by the system, the frequently asked questions (FAQ) database will be updated with the new complaint and directed to the Service Providers' Diagnostic Expert (SPDE) to proffer a solution, which is mailed to the caller at a later time.

The system database was modeled after the MTN top FAQs [14]. The system incorporates an Automated Speech-Enabled Customer Care Service (ASR CCS) proposed in [15] to minimize Human-to-Human (H2H) interaction being replaced with Human-to-System (H2S) model to reduce response time.

The system was tested with the 24 top FAQs that are available at the MTN website. The FAQs are text-based, but the *ASR4CRM* system offered interactive responses to 20 of them perfectly well, while it was observed that the remaining FAQs were found to be composed of long sentences and special characters as a result of which they were rejected. It is therefore recommended that for efficient and accurate response from the system, the FAQs should be as short as possible, preferably, about 40 words and should be void of special characters.

VII. Conclusion

We have described the implementation of *ASR4CRM* - an automated customer care service system that obviates the need for a human operator, reduces the budget allocation of corporate bodies for CCS and most importantly, improves the business to customer (B2C) relationship, which is often damaged by inevitable flaws in the human character.

However, whenever requests are made, and they are not available in the FAQ database, the system directs such requests to the SPDE (the human expert) and solutions are proffered at the shortest possible time and the database updated for future use.

Acknowledgment

VIII. References

[1] C. Akwaja, "Nigeria leads South Africa in Telecoms Growth" *Financial Standards*, vol. 9, no 116, p. 44, 2007.

[2] J. Daly, M. Forgue, Hirakawa, World Wide Web Consortium Issues VoiceXML 2.0 and Speech Recognition Grammar as W3C Recommendations, available online at: http://www.w3.org/2004/03/voicexml2-pressrelease (accessed 26th March, 2008).

[3] J. A. Larson, "Introduction and overview of W3C speech interface framework, W3C working draft 4 December 2000", available online at: http://www.w3.org/TR/2000/WD-voice-intro-20001204/ (accessed 21st February, 2008).

[4] B. M. Raghuram "Design and implementation of V-HELP system – a voice-enabled web application for the visually impaired", available online at http://netlab.unl.edu/netgroup/alumnidocs/mukesh.MS.pdf (accessed 2nd February, 2008).

[5] S. R. Klemmer, et al, "SUEDE: A wizard of oz prototyping tool for speech user interfaces", Available online at: http://citeseer.ist.psu.edu/landay02informal.html (accessed 10th February, 2008).

[6] L. S. Shari, Worthington and W. Boyes, "e-Business in manufacturing - putting the internet to work in the industrial enterprise", *ISA- The Instrumentation, Systems, and Automation Society*, pp. 41-60, 2002.

[7] J. Daly, M. Forgue, Hirakawa, World Wide Web Consortium Issues VoiceXML 2.0 and Speech Recognition Grammar as W3C Recommendations, available online at: http://www.w3.org/2004/03/voicexml2-pressrelease (accessed 16th March, 2008).

[8] B. Marquette, "Voice-Enabled applications deployed using the component server architecture", *SandCherry, Inc*. 2004.

[9] Voxeo Community, available online at: http://community.voxeo.com, (accessed 27th February, 2008).

[10] BeVocal, available online at: http://www.bevocal.com (accessed 15th February, 2008).

[11] Hey Anita FreeSpeech, available online at: http://heyanita.com, (accessed 5th January, 2008).

[12] VoiceGenie Developer Workshop, available online at: http://developer.voicegenie.com(accessed 7th March, 2008).

[13] Tellme Studio, available at: http://studio.tellme.com(accessed 14th February, 2008).

[14] MTN Nigeria FAQs, available online at: http://www.mtnonline.com/support/topfaqs.asp (accessed 11th March, 2008).

[15] A. A. Atayero, O. O. Olugbara, C. K. Ayo, and N. A. Ikhu-Omoregbe, "Design, development, and deployment of an dutomated speech-controlled customer care service system", in *Proc. GSPx 2004: The International Embedded Solutions Event*, Santa Clara, 2004.