Multichannel Interactive Speech Self-Service Systems

Nina Petukhova, Mais Farkhadov, Alexey Troshchenko
Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

Abstract

This paper describes building of voice applications using speech technologies and remote sources of information. Examples of Multichannel Interactive Speech Self-Service Systems (MISSS) are shown. The application Aeroflot-Info enables telephone subscribers to obtain flight schedule and flight status information from the remote public web-service of "Aeroflot-Russian Airlines" company. Presented Multichannel Interactive Speech Self-Service System is build as a scalable, standard based system using state-of-the-art technologies such as SIP for telephone infrastructure, VXML for dialog control, ASP.NET for dynamic content generation, web-services SOAP for information access, speaker-independent speech recognition technology for understanding user input. The application Financial Services has distributed architecture and enables telephone subscribers to obtain the official exchange rates and other information from web-service of Central Bank of the Russian Federation.

I. Introduction

In modern Contact Centers and Voice Self-Service Systems we can see new approaches for their system architecture [1]. They are built using SIP, VXML/CCXML, MRCP standards. They are scalable and distributed and their system components are interchangeable. Now not only it is possible to use corporate applications and database access, but with VXML standard it is also possible to access information from the web-services. Using information from the public remote web-services extends Contact Centers functions and raises their attractiveness for clients.

We assembled MISSS system based on these technologies and in this paper we describe a voice application using this system.

Paper structure is as follows: Section 2 describes the system architecture. Section 3 is dedicated to the Aeroflot-Info voice application and dialogue strategies classification. Section 4 describes MISSS with distributed architecture. Section 5 presents conclusions and future work.

II. MISSS Architecture overview

Typical Dialogue System consists of several main logical components (see Fig.1): *VXML interpreter,* which requests VXML documents with instructions from *Document Server* and processes them. While processing VXML Interpreter interacts with *Implementation Platform*. The Implementation Platform in turn interacts with telephone infrastructure including Voice Gateways, SIP terminals etc, ASR and TTS resources and Media Servers, containing voice prompts.