INTERACTIVE TELEPHONE DIRECTORY: VOICE CONTROL SERVICE

Vladimir Trnka
Department of Electronics and Multimedia Communications,
Technical University of Košice, Park Komenského 13, 041 20 Košice,
Slovak Republic.
Tel: +421-907-583793
E-mail: trnka@intrak.sk

ABSTRACT

In this report I will try to give you short description of projection and realisation voice control service – interactive telephone directory. This service on voice demand through telephone network give you a telphone number for particular employee from Faculty of Electrical Engineering and Informatics Technical university of Kosice. Report also contains description of Nuance system which we used to build the application.

1 INTRODUCTION

What happen when the user dialed a telephone number for this service. After connection is made up, service welcome user and then tell him a first prompt. Here asked him about the name of faculty. In second prompt service asked him for the name of department where is particular employee working. Then follow the prompt which asked user about a name of employee. After succesful speech recognition the system find a information about this person in database and tell them to user. Information contain titul, full name, feature and telephone number. Dialog then ask you if you would like to repeat the number.After this in next prompt you will get three options:

- continue by look for the name of employee from the same department
- continue with the name from different department
- finish the dialog

1.1 REALISATION

Realisation consist of five phases:

1.1.1 Realisaton and preparation of particular grammars

Grammar is set of terms. The main task of grammar is narrow the number of words during searching in recognition process on in advance defined words. Grammar should contains all possible user's answers which can be said in that part of dialog.

1.1.2 Create a vocabulary

For our application we used a Czech language pack which already containts dictionary. You always have to check if word you used in grammar is included in dictionary. If not you have to define pronounce for that word.

1.1.3 Test grammar

After you created a grammar and defined pronounce for every word it's importat to perform test of recognition. Quality of speech recogniton is evalueted by confidence level. Rate of confidence level is from 0 – 100. Optimal is about 40. If the word was not recognize correct you need to redefine it in dictionary.

1.1.4 Construction of dialog

For writing a dialog we used VoiceXML programming language. The structure of this language is similar to HTML. Database is realized like a text document with extension *.csv*. Connection with database is perform through ODBC. Each entry in database contains items titul, name, surname, feature and digits of telephone number. Every item in database become a track after extended by extension *.wav* .

1.1.5 Track recording

For recording and editing tracks we used a tool Xwavedit developed by Nuance. It suppot Sphere headered form of wav files.

For realisation this application were used developer tools produce by Nuance company.

1.2 Nuance system

Main reason why we choose automatic speech communication system Nuance is that this system support Czech language what alloved us to built Slovak application by used Czech phonemes. Type of Nuance architecture is client – server. System Nuance also allow you to test your voice dialogs.

Main processes which allow the fuction of the system:

1.2.1 Process Recognition Client (RecClient)

Control interaction between application and system. Treat with voice input and output. Is also responsible

for forwarding voice input to other process and inform application about user's answers.

1.2.2 Process Recognition Server (RecServer)

Perform speech recognition and interpretation. Also perform a verification of speaker if it is necessary for application. To accomplish these tasks it needs set of acoustic models and grammars.

1.2.3 Application Programmable Interface API

Set of programs (Java, C++) which are use for creating voice applications.

1.2.4 Resource Manager

It ensures that recognition and verification tasks are distributed evenly across the available recognition servers, thus reducing hardware requirements and improving quality of service.

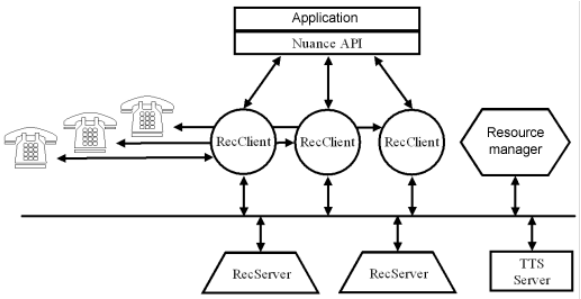

Figure. 1: Nuance archirecture

1.3 Implementation

This application can operate over IP network assuming of using signalling protocols of IP telephony H.323 and SIP. System Nuance support both of these protocols. It can also operate over public switching network. To accomplish this we had to buy telephone card. Card connect computer to telephone exchange. System Nuance support card from NMS (Natural Micro Systems), Aculab and Dialogic.

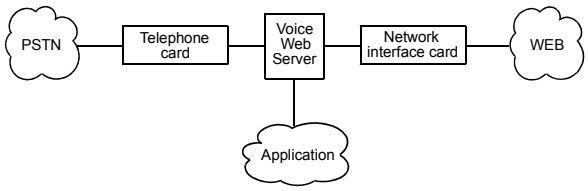

Figure. 2: Topology

Application is now running over public switching network. In our case we used telephone card Dialogic D/120 JCT – LS. Service described is available on telephone number +421-55-602-2297.

REFERENCES

[1] NUANCE COMMUNICATIONS, Inc.: Introduction to the Nuance System, Nuance Speech Recognition System 7. June 2000

[2] LIHAN, S. – JUHÁR, J. – ČIŽMÁR, A.: The use and potential of Voice XML in building speech dialogue systems for telecommunications, Conference proceedings or ICETA 2001

[3] ČIŽMÁR, A. – JUHÁR - J. LIHAN, S.: Automatizované hlasom ovládané služby v telekomunikáciách, 1.vyd. Košice: ELFA, 2002.