

XML-Based Multimodal Interaction Framework for Contact Center Applications

Nikolay Anisimov, Brian Galvin, Herbert Ristock
 Genesys Telecommunication Laboratories (an Alcatel-Lucent Company)
 2001 Junipero Serra Dr.,
 Daly City, CA 94014, USA
 Tel.: +1 650 466-1347

{anisimov,bgalvin,herbertr}@genesyslab.com

ABSTRACT

In this paper, we consider a way to represent contact center applications as a set of multiple XML documents written in different markups including VoiceXML and CCXML. Applications can comprise a dialog with IVR, call routing and agent scripting functionalities. We also consider ways how such applications can be executed in run-time contact center environment.

Categories and Subject Descriptors

H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems - *Audio input/output, hypertext navigation and maps*. H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces - *Web-based Interaction*.

General Terms

Documentation, Standardization, Languages.

Keywords

Call center, contact center application, VoiceXML, Call Control XML, call routing, agent scripting.

1. INTRODUCTION

Contact centers (CC) play a very important role in contemporary business. According to some estimations [3] 70% of all business interactions are handled in contact centers. In the U.S., the number of all contact center workers is about 4 million or 2.5-3% of the U.S. workforce.

Creating business applications in contemporary Contact Centers is a very complex task. Indeed, typical CC applications comprise Interactive Voice Response (IVR) scripts, routing strategies, call control, agent scripting, reporting, etc. Each of these functions has its dedicated tools and scripting languages and a CC application designer is required to be proficient in all of them. The heterogeneous structure of CC applications is a challenge also because many of the applications, such as routing strategies, are also strongly platform dependent. Since most of the leading contact center applications remain proprietary, it is quite common that applications developed for a specific contact center product cannot be easily transferred to another one.

A proven way of achieving application uniformity, platform independence, and simplification of the task of creating business applications is to employ XML-based standards and related

technologies. XML is increasingly used as a basis for building applications in different vertical businesses. Good examples of XML-based standards for voice processing are the VoiceXML [5] and Call Control XML [2] protocols developed within W3C. They enable representation of any voice application as an XML document, and using VoiceXML and CCXML it is already possible to build simple CC applications involving only IVR (including automatic speech recognition capabilities) processing and simple call control and to represent them as a single XML document. The main advantages are obvious: uniformity, platform independence, and leveraging web technologies.

However, VoiceXML and CCXML do not address other important aspects of CC applications such as interaction workflow/service chain management (the process management task specialized in customer interaction management), interaction routing, scripting agent activities, reporting on agent (sometimes called customer service representatives – CSR) performance and traffic management, using customer profiles, conducting outbound campaigns, and interactions that are conducted in media other than voice.

In [1] we proposed some ways of extending the VoiceXML and CCXML approach in order to provide coverage for additional important contact center functionality. We proposed a methodology that is open to incremental extensions and that presents basic interaction management concepts such as platform and application, multi-script and multi-browsing, and interaction data processing without attempting a comprehensive top-down standard.

In this position paper we consider a contact center application within W3C Multimodal Interaction framework [4]. According to this approach, CC application can be represented as a set of XML documents with different namespaces. We also consider how it can be executed in typical CC environment. In this paper we focus on main concepts and principles rather than specific XML languages.

2. MULTIMODAL FRAMEWORK

Agent involvement in a contact with a customer can be considered from web perspective, see Figure 1.

One could think of it as the CSR playing the role of a browser “rendering” agent script dialog instructions written in HTML. Similar to VoiceXML an agent script specifies a dialog with a customer but in different terms. Moreover, CSRs usually use additional knowledge acquired during training process and sometime referred to as skills.

We can consider such an environment as another modality or more strictly as another implementation of voice modality. The main difference here is that a CSR-browser should be found before starting the browsing session. Moreover, the CSR should have appropriate skills and be available (not busy). This searching logic can be expressed in an XML-based form as routing strategy, see previous section.

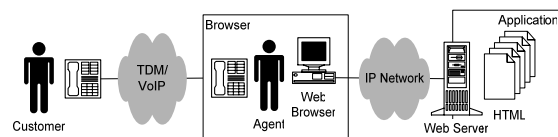


Figure 1: Agent as a voice browser

The CSR environment can be considered as a special case of W3C Multimodal architecture [4]. In this architecture, VoiceXML and Agent scripts play the role of markup languages for modality components. CCXML and XML strategy are markup languages for controller and interaction management.

3. APPLICATION STRUCTURE

We consider structure of CC application using typical application with IVR and agent involvement.

3.1 Design-Time View

The application can be designed as a set of four XML documents, see Figure 2. The root document is written in CCXML which plays the role of interaction manager markup.

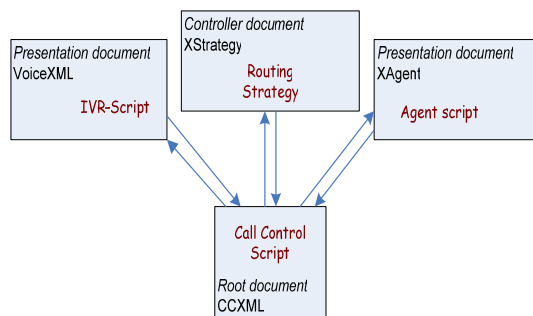


Figure 2: Application Structure

This document contains logic of call control. It is activated when a call arrives into CC. After that it invokes a presentation document with IVR script that is written in presentation markup VoiceXML. This document controls a spoken dialog with a customer that may collect needed information. After the end of the dialog, this information is returned to the call control script. Based on this information the application starts searching for the most appropriate CSR invoking routing strategy script. The script is written in a markup called XStrategy [1]. It returns the address of the most appropriate available CSR. Then the call is transferred to this CSR workplace. A corresponding agent application written in XAgent markup is then activated which helps CSR to talk to the customer.

3.2 Run-Time View

The run-time view of the CC application is depicted in Figure 3. The contact center environment comprises several application

servers, each being responsible for a particular function of contact center operation. The call control part of application is executed by CTI-Server that connects telephony and computer domains.

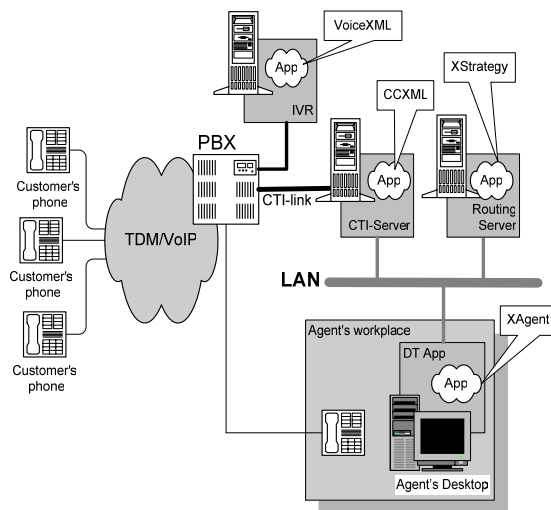


Figure 3: Run-time view

All application servers and workstations are connected via LAN and synchronized by event exchange.

4. CONCLUSION

In this paper we introduced main concepts that we believe will be important for a comprehensive and consistent scripting of all contact center functions. In particular, we considered W3C Multimodal Interaction Framework as a suitable approach for CC application design and execution. Our future plans include the incorporation of applicable existing XML specifications and the development of XML languages for specific areas of contact centers.

5. REFERENCES

- [1] Anisimov N., Galvin B., Ristock H. XML-based Framework for Contact Center Applications. In: Filipe J. *et al* (Eds). Proc. of 3rd Int. Conf. on Web Information Systems and Technologies (WEBIST 2007), Barcelona, Spain, 3-6 March, 2007. Vol. 1, 443-450.
- [2] CCXML. Voice Browser Call Control: Version 1.0. W3C Working Draft, June 29, 2005. <http://www.w3.org/voice/>
- [3] Gans N., Koole G., Mandelbaum A. Telephone Call Centers: Tutorial, Review and Research Prospects, Manufacturing and Service Operations Management, 2003, vol.5, no.2, 79-141
- [4] Multimodal Architecture and Interfaces. W3C Working Draft, December 11, 2006. <http://www.w3.org/TR/2006/WD-mmi-arch-20061211/>
- [5] VoiceXML. Voice Extensible Markup Language. Version 2.0. W3C Recommendation, March 16, 2004. <http://www.w3.org/voice>