The World Wide Telecom Web Browser

Sheetal Agarwal, Arun Kumar, Amit Anil Nanavati, Nitendra Rajput

IBM India Research Laboratory

4, Block C, Vasant Kunj, Institutional Area, New Delhi - 110070, INDIA.

{sheetaga, kkarun, namit, rnitendra}@in.ibm.com

ABSTRACT

As the number of telephony voice applications grow, there will be a need for a browser to surf the Web of interconnected voice applications (called as VoiceSites). These VoiceSites are accessed through a telephone over an audio channel. We present the concept and architecture of *T-Web Browser*, a World Wide Telecom Web browser that enables browsing the Web of voice applications through an ordinary phone. This browser will support rich browsing features such as history and bookmarking.

Categories and Subject Descriptors: H.4.3 [Communications Applications]: Information browsers

General Terms: Design, Human Factors

Keywords: Developing Regions, Voice Browser, WWTW, HSTP

1. INTRODUCTION

Currently, voice applications are accessed using a telephone device as shown in Figure 1. A Voice Browser is used to access multiple applications $A_1, A_2, ..., A_n$ that are hosted on the same Application Server (*App Server* in the figure). Since all the voice applications are on the same server, and are accessed by the same Voice Browser, it is possible for the Application Server to maintain a browsing history of the applications being browsed. The applications A_1 to A_n can provide a link to each other [1] .
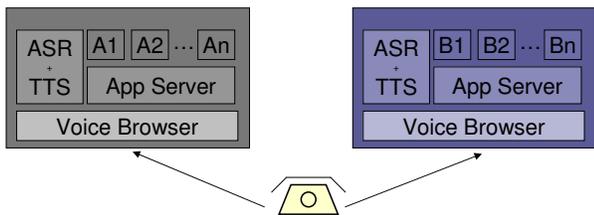

Figure 1: The current deployment and access model.

However, in this architecture, the Voice Browser at the first site will not be able to surf a link from an application in set A to any application in B. *How to enable surfing such cross-browser links is the problem we address in this paper.*

By allowing a voice application on one Voice Browser to link to another voice application deployed on another Voice Browser, the reachability of these applications can be significantly increased. This is analogous to the web applications in the World Wide Web, where any website can connect to any other website, regardless of where they are deployed.

Continuing the Web analogy further, we have shown that creating and deploying voice applications (called VoiceSites) is as easy as creating a webpage [3]. This is especially important for people in developing countries where Internet/PC availability is low, since creating VoiceSites can be done by *speaking* through an ordinary telephone. We also described the process of enabling *hyperspeech* link to another voice application. and presented a Hyperspeech Transfer Protocol to support *hyperspeech* links [2]. Such an interconnection of VoiceSites open several possibilities for telephony voice applications and can create a web parallel to the WWW, known as WWTW [4].

As an example, a fisherman can create his VoiceSite that has information and pricing of the fishes available with him. He can link his VoiceSite to a payment gateway VoiceSite to enable transactions. Villagers can call his VoiceSite and order a fish and make payment while the fisherman is busy at his lake farm. In a practical scenario, it is likely that these two VoiceSites are hosted at different Voice Browsers.

In this paper, we present the concept and architecture of a T-Web Browser. This browser will enable navigation of interlinked voice applications that are deployed across different Voice Browsers. The architecture illustrates that the Browser can be implemented as a special VoiceSite (Section 2) and can support standard browsing features such as *go-back* and *bookmark*. The T-Web browser can also be implemented on the device itself, but that would require speech recognition support on the device.

2. WWTW BROWSER ARCHITECTURE

We implement the T-Web Browser as a special VoiceSite. In addition to a dialog flow (authored as VoiceXML-jsp), the T-Web Browser VoiceSite uses a database to maintain a history of the current user session. It uses an additional database to store the caller bookmarks. The history consists of the title and the phone numbers of the VoiceSites called by the user. The bookmarks contain the phone number and a name tag for each bookmarked VoiceSite. The operational model of the T-Web Browser is shown in Figure 2.

We describe the working of T-Web Browser using the circled steps shown in the figure.

1. User calls the T-Web Browser to access a VoiceSite.
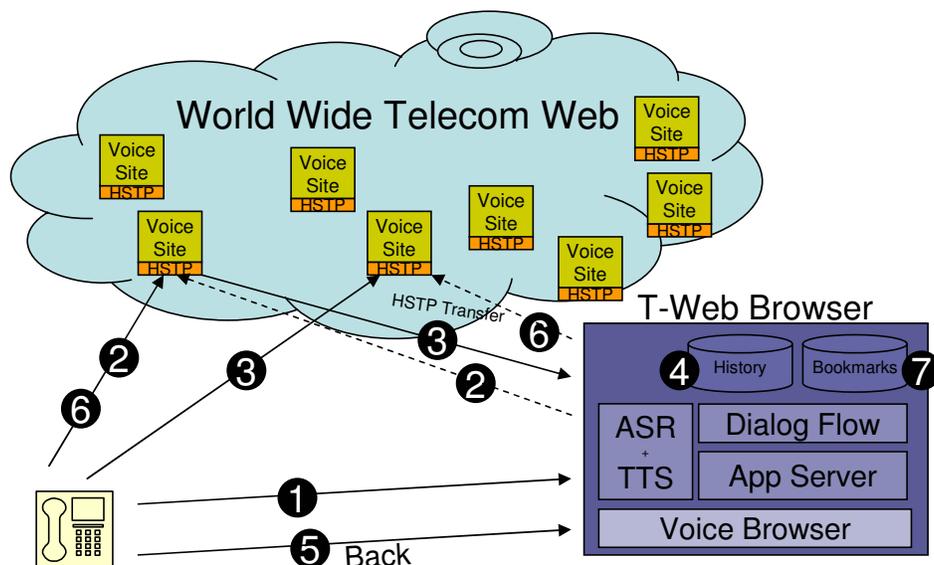2. The T-Web Browser transfers the call to the phone number of the VoiceSite through HSTP.

Figure 2: The T-Web Browser interaction with VoiceSites.

3. When a user selects a hyperspeech link to browse to the other VoiceSite, the session is transferred to the target VoiceSite and HSTP passes the call transfer information to the T-Web Browser.

4. This information is stored in the Browser history.

5. The user issues a Browser command, e.g. *go-back*.

6. The T-Web Browser instructs the HSTP layer on the current VoiceSite to initiate a transfer to the earlier VoiceSite phone number.

7. At anytime, the user can say *bookmark* to bookmark the currently browsed VoiceSite.

The HSTP layer implements the protocol for transfer of user session from one VoiceSite to another. We will have to modify the HSTP protocol and the message format for it to support browsing features. An additional field will have to be added to the HSTP message. This is the phone number of the next VoiceSite to which the user will surf. The current VoiceSite application context will be transmitted to the T-Web Browser in the application context field of the HSTP message.

The above architecture requires that a user can access the VoiceSite and the T-Web Browser simultaneously. This requires the availability of two simultaneous voice channels from the user phone. This can be realised in two ways: (a) through a three-party conference call between the user, the T-Web Browser and the VoiceSite, (b) by having both the voice channels active, and the user can put one channel on hold while talking through the other. The first approach requires that the VoiceSite and the T-Web Browser should be able to disambiguate the user utterance and identify whether it is a command to the browser or an interaction with the VoiceSite. The latter approach needs a phone (and the service provider) that has the ability to provide two simultaneously active calls. Since neither of the two requirements are unsurmountable, any of these methods can be used in the implementation of the T-Web Browser.

This architecture can support rich browsing features similar to multi-tabbed browsing on the WWW by enabling simultaneous active calls to different VoiceSites. A user can put other VoiceSites on hold while interacting with one.

2.1 Implementation

We have implemented the HSTP protocol layer that will enable the call and context transfer required by the T-Web Browser. The HSTP protocol has been implemented as a Java class library and the API can be accessed from any VoiceXML application through a Java Bean. We are in the process of developing the prototype of the T-Web Browser VoiceSite. The VoiceSite will be authored in VoiceXML-jsp. We will deploy this VoiceSite in Apache TomCat. This will be accessed through phone calls that are intercepted by the Dialogic telephony card. We use the Genesys Voice Browser to interpret VoiceXML.

3. CONCLUSION

In this paper, we briefly presented the requirement for a Browser that can browse VoiceSites that are deployed across domains. We described the concept and a high level architecture of a T-Web Browser that is implemented as a VoiceSite. We are in the process of implementing this browser, and keenly await user feedback through field studies.

4. REFERENCES

[1] 1800-555-TELL. http://www.tellme.com.

[2] S. Agarwal, D. Chakraborty, A. Kumar, A. A. Nanavati, and N. Rajput. HSTP: Hyperspeech Transfer Protocol. In *ACM Hypertext*, UK, September 2007.

[3] A. Kumar, N. Rajput, D. Chakraborty, S. Agarwal, and A. A. Nanavati. Voiserv: Creation and delivery of converged services through voice for emerging economies. In *IEEE WoWMoM*, Finland, June 2007.

[4] A. Kumar, N. Rajput, D. Chakraborty, S. Agarwal, and A. A. Nanavati. WWTW: A World Wide Telecom Web for Developing Regions. In *ACM SIGCOMM Workshop on Networked Systems For Developing Regions*, Japan, August 2007.