Applying SIP VoIP to Context-Aware Telephony

Hsien-Chang Wang∗, Gu-Jun Chen∗, and Jenq-Muh Hsu**

∗Department of Information Management, Chang Jung Christian University
396 Chang Jung Rd., Sec.1, Kway Jen, Tainan 71101, Taiwan, R.O.C.
E-mail: wangbb@mail.cju.edu.tw, kgiking@gmail.com

**Department of Computer Science and Information Engineering, Chung Cheng University
168, University Rd., Min-Hsiung Chia-Yi, Taiwan, R.O.C.
E-mail: hsujm@cs.ccu.edu.tw

[Received January 20, 2007; accepted May 22, 2007]

The simple, efficient Session Initial Protocol (SIP) is widely used to integrate context-awareness computing techniques. We propose a SIP-based research framework that integrates automatic speech generation to make communication between users more convenient. Based on the proposed approach, we built and tested a SIP context server and several SIP agents hosted on PCs. If the called party is absent or busy, a context-aware message is automatically replied to the caller via synthesized natural speech. Experiments conducted by 50 college students using ADSL connected to the SIP server showed encouraging results, especially in user satisfaction in message transfer rates and service waits.

Keywords: SIP, context-aware, ubiquitously, session control, VoIP

1. Introduction

The rise of Internet, cellphone, and satellite use now makes it possible for anyone to contact anyone at anytime and anywhere [6]. The caller is usually unaware, however, of the callee’s status (context information), making further call processing difficult to decide in the absence of such context information [10]. Because nonuniform network protocols complicate communication, the need arises for simple, cheap gateways to expedite communication. A context-aware system [1] must enable (1) presentation, (2) automatic execution, and (3) tagging. Presentation provides context information to users. Automatic execution provides services automatically, and adapts to different conditions. Tagging integrates context information and automatic execution modules.

Depending on the information available, it is possible to determine the user’s status and provide required services [2]. A standard context-aware system must provide (1) context organization, (2) context utilization, and (3) adaptation [8] (Fig. 1).

The context organization module (Fig. 1) focuses on the collection, storage, abstraction, and interpretation of context information, which is centralized and conveyed to the context utilization module, which, in turn, sends context information to context-awareness applications, then goes to the next procedure based on context information.

The Service Synthesizer on the Net (STONE) project [3] at the University of Tokyo is a content-aware system begun in 1999 to explore possibilities for users in accomplishing tasks seamlessly and ubiquitously. STONE focuses on the development of context-aware services in which applications change functionality based on the dynamically changing user context. Its work is structured on three basic themes: (1) network middleware that connects objects seamlessly to provide flexible services, (2) devices that locate systems and sensors that communicate context and interaction, and (3) the STONE room – a test-bed for testing and demonstrating application prototypes.

Despite progress Internet telephony has the following problems:

- A lack of unified standards.
  Standards for developers and users are neither standardized nor uniform. Different products e.g., Skype and MSN, use different standards, snarling mutual communication and raising many inconveniences.
- A lack of connection between heterogeneous networks.
  Network architectures such as Wireless, Bluetooth, and Ethernet deliver poor service quality through poor mutual interconnection.
- User service needs.
  Users require more than a simple Internet phone service. Without value-added services, current Internet
phone services remain no better than PSTN clones. They must offer better service than PSTN to justify their use.

Despite these disadvantages, SIP is an application-layer control protocol that establishes, modifies, and terminates multimedia sessions such as Internet telephony calls. SIP features the following [4]:

- It is text-based. This enables easy implementation in object-oriented programming languages such as Java and c/c++, allows easy debugging and, most importantly, makes SIP flexible and extensible.

- It involves less signaling. SIP is designed to meet only basic call-signaling protocol requirements – create, modify, and terminate – so signaling is kept as simple as possible.

- Parallel search is possible. A stateful SIP server can split or “fork” an incoming call so that several extensions can be rung at once. The first extension to answer takes the call. This is handy for user working between two locations – a laboratory and an office, for example – or where someone is ringing two parties simultaneously – both a boss and secretary, for example.

Internet Protocol (IP) telephony currently uses two different architectures: H.323, developed by the International Telecommunication Union (ITU-T), and the Session Initiation Protocol (SIP), developed by the Internet Engineering Task Force (ETF). H.323 was the first signaling protocol standard for the Voice over Internet Protocol (VoIP). SIP is becoming increasingly popular as different vendors including it in their products without waiting for its full maturation. From the beginning, SIP designers have considered critical issues such as modularity, integration with Internet services, extensibility, and simplicity. SIP itself is not a vertically integrated communication system, but is only part of the overall IP telephony architecture, which consists of a range of packet-switched protocols.

Figure 2 shows the SIP-based IP telephony protocol stack [7]. Note that SIP lies in the application layer and must be used in conjunction with other protocols to provide real-time services. When a SIP signaling session is successfully setup, for example, a Real-Time Transport Protocol (RTP) media session is opened and the virtual session begins. Note that a digitally encoded voice stream is carried over an IP network using the RTP Protocol. The RTP Control Protocol (RTCP), a companion protocol of RTP, is mainly used for providing quality-related feedback. The RTCP functions as network remote control for on-demand delivery of real-time data. The media gateway control protocol is used when interoperating with conventional circuit-switched networks. The main signaling protocol is SIP, which plays the role of H.323 in an H.323-based VoIP network.

2. Proposed SIP Network Architecture

In this research, user context information is recorded using SIP and URL. This paper is organized as follows: Section 2 discusses the detailed architecture of the proposed SIP network and the controlling protocol, Section 3 presents preliminary results, and Section 4 draws conclusions.

Under our proposal (Fig. 3), users are connected to the network via devices with SIP agents. Messages are transmitted to SIP CA, a context-aware server, via a SIP proxy server. SIP CA uses three modules: the processing module; the context utilization module, which searches for, collects, and stores user context information; and the context organization module, which responds with command messages based on judgment of the user’s environment. If the user is busy, SIP CA replies with a busy message and terminates the user’s request.

Take message communication between users A and B, for example: user A issues an INVITE message to the