International Journal of Advanced Culture Technology Vol.1 No.1 11-14 (2013)
IJACT.2013.1.1.11

IJACT 13-1-3

Novel Speech Web Architecture Based on Information Selection Agent

Hyeong-Joon Kwon[1], Tetsuo Kinoshita[2]

[1]*School of Information and Communication Engineering, Sungkyunkwan University, South Korea*
katsyuki@skku.edu

[2]*Tohoku University Japankatsyuki@skku.edu*
kino@riec.tohoku.ac.jp

Abstract

In this paper, we propose a prototype of the SpeechWeb application using the information selection agent. We describe its design and implementation method and illustrated the processing results with the aid of some screenshots. Proposed SpeechWeb application presents the associated contents to the user by the aid of dynamic voice-anchors. These contents are presented using the apriori algorithm, which is one of data mining techniques. The application is better than the existing user-initiative structure from the viewpoint of making the user's interesting induction. Moreover, we believe that our proposed application is effective in information retrieval through wired and wireless telephone networks.

Keywords: *VXML, SpeechWeb, Voice Information System*

1. Introduction

This overcomes the problems of the existing VXML application, wherein it is not possible to include different scenarios. This structure, termed user-initiative, compares a given word with the dictionary after eliminating the stop words, e.g. an article and a prepositional word, from all the words that are detected every time from the information demanded by the user. If the detected word is the same as the keyword in the word dictionary, that word is presented to the user as the recognition candidate of the next input. However, the load of processing every user's input is significantly high. Moreover, the word dictionary database is built by application developers. Therefore, it cannot be consistently guaranteed that the user will receive the correct information. For example, this system can't process about synonym and polysemy in Korean.

In this paper, we have implemented an intelligent agent by applying the data mining algorithm. This agent searches for the content related to the information demanded by the users. The agent explores the association rules; consequently, voice-anchor is liked association information is presented to the user. This exploration into association rule among information is based on experienced information by users to use a service before, item sets of existing inputted keyword. This intelligent handling method saves the time required by the users to search for required information and to obtain additional information regarding a specific topic. Our proposed agent presents the users with positive associate information based on an appropriate threshold. Furthermore, it reduces the server's loads because the cutoff processes for eliminating the stop words, which

Manuscript Received: Mar. 28, 2013 / Revised: Apr. 10, 2013 / Accepted: Jun. 7, 2013
Corresponding Author: katsyuki@skku.edu
Tel:+82-31-290-7196, Fax: +82-31-290-7196
Dept. of and Communication Engineering, Sungkyunkwan University, South Korea

require many operations, are reduced.

The structure of this paper is as follows. In Section 2, we design and implement the proposed SpeechWeb application using the intelligent data mining agent. Then, we explain the method of implementing this application. In Section 3, we provide the results of an experiment performed using proposed system. The conclusions are provided in Section 4.

2. Proposed System

To overcome the above problem, we implement a visual-voice information system with a special agent, which explores the association rules using the user's input database transactions from the past in a visual Web environment and presents the correct voice-anchor with appropriate content to the user. Our agent generates an XML document on a daily basis using the investigation results of the association rules which is one of data mining techniques. The contents of this XML document include the association rules. Our proposed system analyzes this XML document and provides the required information to the user.

2.1 System Architecture

Our system instantaneously provides information service to the user in visual and voice environments. The system obtains input keywords from visual users, and the agent examines the association rules using the user's input transactions on a half-hourly basis. Figure 1 shows the architecture of our proposed system.

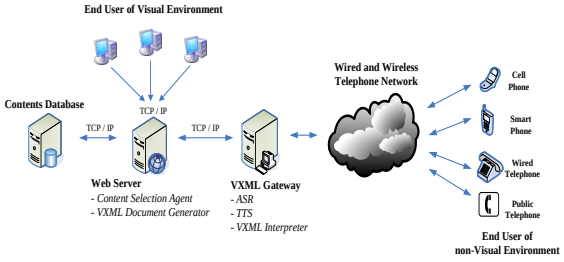


Figure 1. System architecture

2.2 Proposed Agent

Our information selection agent is implemented using the apriori algorithm, which is one of the data mining techniques. The apriori algorithm is useful for determining the association relation among the items. The association rule "If X is occurred, Y will be occurred" is represented as follows [6].

$$R: X \Rightarrow Y$$

To determine the association relation, the apriori algorithm uses three key elements—Support, Confidence, and Lift values. Support implies the total number of transactions divided by the number of transactions that include itemset X, and Confidence implies the reliability of Support. Lift estimates the association rule that is computed by Confidence and Support [6].

$$Conf(R) = \frac{Supp(X \cup Y)}{Supp(X)}$$

$$Lift(R) = \frac{Conf(R)}{Supp(X)} = \frac{Supp(X \cup Y)}{Supp(X)Supp(Y)}$$

The apriori algorithm is repetitively executed by Support and Confidence. Further information on this algorithm is available in [6]. Further, in Section 3, we provide the simulation results of the apriori algorithm.

Figure 2 shows the processing result of our agent. The result is an XML document that includes the total number of transactions, threshold, Support, Confidence, and Lift. In Figure 3, the threshold is 20%, and the number of transactions is 40.

2.3 VXML Document Generator

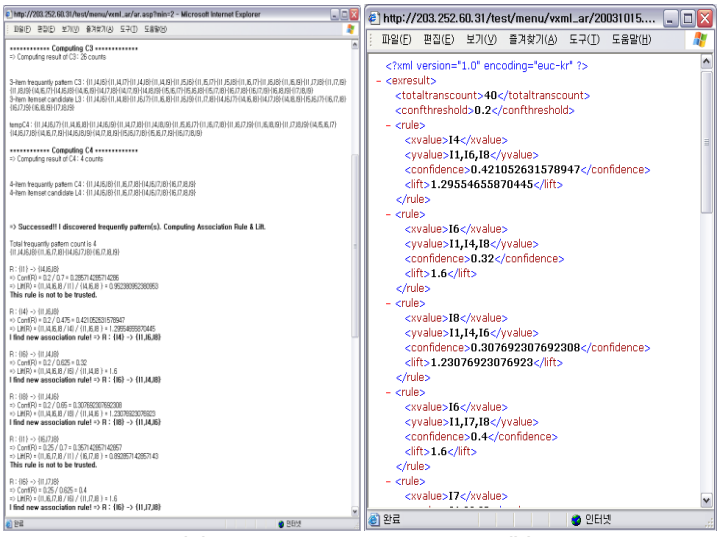
We generate the VXML document with an XML document that is generated by our information selection agent. For example, in Figure 2, if the user input is I4, according to the first rule, our system will present I1, I6, and I8 as the associated information. On the other hand, a VXML document or a scenario and recognition candidate is dynamically transformed by the user input. Figure 3 shows an example of the VXML document that is generated by our VXML document generator when the user input is I6, second rule in Figure 3.

3. Experimental Results

3.1 Experimental Conditions

The hardware and software of our system that are used for the experiments are composed of an Intel Xeon 3.0 GHz Dual CPU, 2 GB RAM, Intel Dialogic D/120 JCT-LS PCI computer and telephony integration (CTI) board [7], Microsoft Windows 2000 service pack 4 OS, and internet information service (IIS) version 5.0 as the Web container. HUVOIS version R1.1 by Korea Telecom [8] is used as the automatic speech recognition (ASR) and text-to-speech (TTS) engines. Since the total number of transactions is 40 and the total number of items is 18, the threshold is 0.2.

3.2 Processing Result of Proposed Agent



(a) (b)

Figure 2. Execution result of information selection agent

Figures 2(a) and (b) show the execution results of our agent and the XML document generated by it, respectively. For example, in Figure 2(b), if the user input is I4, our system will present I1, I6, and I8 to the user as the next information. If Lift is greater than 1, it will be selected based on the rule of the apriori algorithm [6].

3.3 VXML Document Generation

Figure 3 shows the simulation result of our VXML document generator. It generates an accurate VXML document using the XML document shown in Figure 2(b). Figure 3 shows the processing result for a user input of I6.

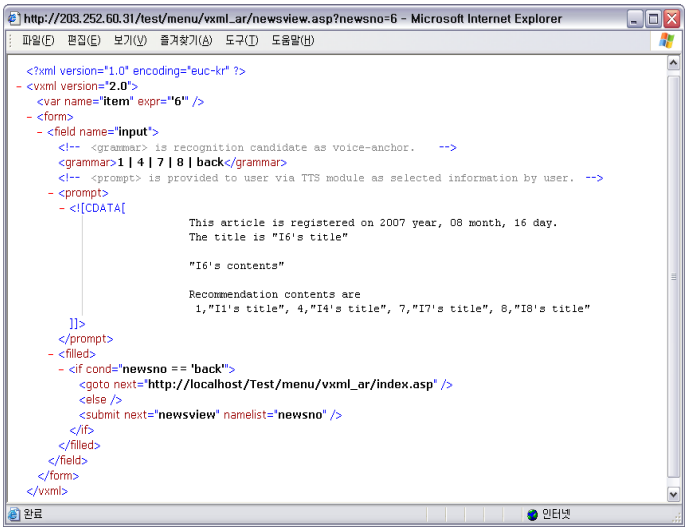


Figure 3. Simulation result of information selection agent

The system states the article's title; then, it asks for the user input. Second, as shown in Figure 3, the system states the selected content's registration date and main contents; then, it presents the recommended contents using the voice-anchor, which comprises a number. These steps are further repeated by the system.

4. Conclusion

We have proposed a prototype of the SpeechWeb application using the information selection agent. We have described its design and implementation method and illustrated the processing results with the aid of some screenshots. Our proposed SpeechWeb application presents the associated contents to the user by the aid of dynamic voice-anchors. These contents are presented using the apriori algorithm, which is one of data mining techniques. Our proposed SpeechWeb application is better than the existing user-initiative structure from the viewpoint of making the user's interesting induction. Moreover, we believe that our proposed application is effective in information retrieval through wired and wireless telephone networks.

References

[1] Eric Auchard, Google tests directory assistance for ph-ones, Routers Technology News, available at <http://www.reuters.com/article/technologyNews/idUSN0641979620070407>
[2] Scott McGlashan et al., Voice Extensible Markup Language Version 2.0 Specifications, W3C Standard document, available at <http://www.w3.org/TR/voicexml20>
[3] Hemambaradara Reddy et al., "Listener-Controlled Dy-namic Navigation of VoiceXML Documents", LNCS Vol.3118, Springer (2004)
[4] Vankayala et al., Dynamic Voice User Interface Using VoiceXML and Active Server Pages, LNCS Vol.3841, Springer (2006)
[5] Hyeong-Joon Kwon et al., "A User-Controlled Voice-XML Application Based on Dynamic Voice Anchor and Node", LNCS Vol.4743 (2007)
[6] Pang-Ning Tan et al., *Introduction to Data Mining*, A-ddison Wesley (2005)
[7] Intel Dialogic Homepage, available at <http://www.dialogic.com/manuals/sr60winpci/IssuesByProduct.pdf>
[8] KT Service Develop Laboratory, HUVOIS Solution Manual R1.1 (2003)