PannoMulloKathan: Voice enabled Mobile App for Agricultural Commodity Price Dissemination in Bengali Language

Madhab Pal, Rajib Roy, Soma Khan, Milton S. Bepari, Joyanta Basu

Centre for Development of Advanced Computing, Kolkata, India

{madhab.pal, rajib.roy,soma.khan, milton.bepari,joyanta.basu}@cdac.in

Abstract

In this work we present a voice based mobile application for dissemination of agricultural commodity procurement and consumer prices. Disbursed information is crawled at daily basis from government authorized websites of agricultural marketing departments. The app incorporates mix media multiple access means in form of touch-type-see, touch-type-listen, speak-see and speak-listen modalities and also includes a robust Automatic Speech Recognition (ASR) engine in Bengali language to support real time voice queries. Colorful interactive app based user interface and ASR incorporated core client-server architecture altogether provides an efficient framework for serving registered users of different educational and economical background including people having little or no computer knowledge, semi-literate or illiterate rural people.

Index Terms: agricultural information system, voice enabled mobile app, speech recognition

1. Introduction

Adoption of Information Communication Technology (ICT) enabled information support systems for sustainable development in agricultural extension is quite important for developing countries like India, where agriculture is a major livelihood [1]. Agriculture continues to be the most important sector of Indian economy also. In view of the country wide spread of telephone and mobile network with easy availability of smart mobile devices and affordable mobile data services, both the government and private sector are now coming up with initiatives [2] to develop and support ICT based smart agriculture related solutions. Prime focus areas under these initiatives include facilitating farming and associated activities with information support, empowering farming community, promoting ago-produce marketing etc.

With a similar background, we present here our work on developing a speech based Mobile App for marketwise agricultural commodity price dissemination in Bengali language. The developed app extracts real time data from already existing authentic websites of Agri-marketing departments of govt. of West Bengal and govt. of India (AGMARKNET), manages procurement as well as consumer prices separately on daily basis and disburses the same on request from registered end users. On using the app, request can be placed any time through voice query, image selection or text typing; whereas information is provided in both speech and text forms. Voice queries in form of spoken commodity names in Bengali language are automatically recognized by a robust Automatic Speech Recognition (ASR) engine at runtime, to generate text outputs for commodity variety selection, commodity searching etc. The developed application serves end users from both rural and urban population, specifically who are not computer literate or otherwise find is difficult to access the same data during busy working hours. In fact, rural end users of this application mostly include semi-literate non-tech savvy farmers and agro-produce sellers who mainly earn their livelihood by selling farm grown products in local markets. Easy access to standardized (govt. authorized) commodity procurement price information duly on time saves these users a prior trip to the local market or asking a middleman and selling agro products below market prices. Price information for consumers, on the other hand, benefits smart consumers to have fair deals while buying agricultural commodities. This speech based mobile application is thus unique in its own way to facilitate farmers as well as consumers through a single application.

2. Key issues on Application Design

Major issues that are being thought of and decided during design of the developed application are as follows:

- *Connecting stakeholders:* simple, interactive interface with efficiently designed core system enables easy usage by all the stakeholders (govt. dept., farmer, consumer) and thus connect to each of them alike
- *Access modality:* various access modalities like touch-type-see, touch-type-listen, speak-see and speak-listen are included in the design to attract literate as well as semi-literate end users
- *User preferences:* user preferences for language, input output modality, service type etc. are recorded while user registration and maintained until next change.
- *Information coverage:* other than commodity prices, additional information on nearby markets, collection centres and outlets location and contact are also included
- *Content quality:* design should ensure completeness, reliability and timeliness (important for agriculture domain) of the information to be provided
- *Core technology:* app based colorful user interface, underlying client-server architecture and latest ASR technology in the application core to add flexibility and ease of access.

3. App overview and architecture

Overall architecture of the mobile app is dependent on three basic components; android based app user interface, a web service to which the app communicates and an authentic online information sources. Figure 1 shows the entire app architecture with data flow within the core modules.

- **Android App:** This is the client side interface through which user requests to the web service for latest available commodity price information.

- **Web Service:** Web service available online takes request from the app and responds to the same. At backend, this web service requires some important server side scripts, images, sound files, ASR engine, database and daily information crawler.
- **Information Source:** this is the authentic information source server from where relevant data is crawled daily and stored within in-house database.
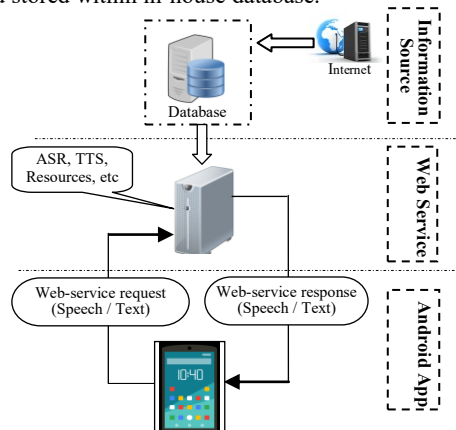

Figure 1: General block diagram of the App

4. Bengali ASR: the core component

The voice enabling technology being used in the app back end is connected word ASR. Baseline acoustic models are built using indigenously collected real world telephony speech data (of around 98 hrs) from all geographical regions of West Bengal to cover major dialectal variations of live Bengali speech. A dynamic language model is designed that automatically builds itself for daily reported agricultural commodities. Besides data collection, sustained efforts have been given in transcription, core ASR engine development, evaluation and field trial analysis. From field trial feedback analysis of the core ASR engine, distinctive error recovery methods like Signal Analysis and Decision [3], Confidence Measure and Polling [4], Complementary Information, Runtime model generation etc. are also incorporated to confirm satisfied performance in field conditions. Comparative recognition performance analysis is performed on field collected live test audio of around 5 hrs using Sphinx and Kaldi toolkits to finalize a robust backend ASR system. This reveals our best performing Kaldi ASR system having 7.9 % WER using SGMM [5] with LDA, MLLT and SAT [6] training on extracted MFCC, delta & double delta features.

5. Speech based Mobile App

After user registration, the fully designed mobile app shows a list of available agricultural commodities with specification and respective images individually for easy understanding and selection by end users. As per price availability on the authorized information source websites, the list is updated automatically in daily basis. Based on the selected language preference, user can listen to the textual price information either in Bengali or English languages, though ASR engine expects only Bengali language speech. Besides selecting, user can also search for any commodity, by typing few letters of the commodity name or speaking the commodity name. In addition to commodity prices, the app also provides other relevant information like locations of all static and mobile outlets, collection centres, contact numbers etc. Figure 2

shows some of the screenshots of the designed mobile app highlighting the main features.
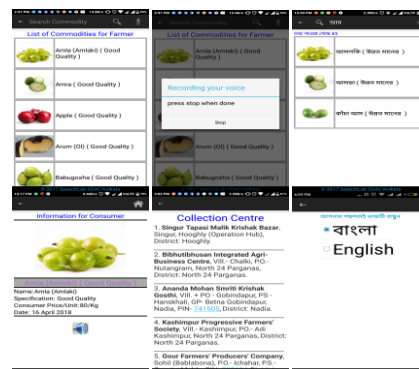

Figure 2: Screenshots of PannoMulloKathan Android App

6. Conclusions

This work focuses on overall architecture, design process and features of a voice based mobile application for agricultural commodity price dissemination in Bengali language. The ASR incorporated app is providing excellent value-addition to the already existing websites of agriculture marketing department of govt. of West Bengal and govt. of India; as no computers, internet, or even reading/writing skills are required for using this application. Presently the app is being field trialed under real world environment and updates will be included after a detail feedback analysis. We are working on extending the ASR part of the work for processing continuous word voice queries; so that, agriculture related live voice chat applications (on topics like, soil preparation, crop advisory etc.) can also be developed in similar line.

7. Acknowledgements

8. References

[1] J. Aker, "Dial "A" for agriculture: using information and communication technologies for agricultural extension in developing countries." Conf. on Agriculture for Development-Revisited, University of California at Berkeley, 2010.

[2] Saravanan, R. and S. Bhattacharjee, "Mobile phone applications for agricultural extension in India", Worldwide mAgri innovations and promise for future. New India Publishing Agency. pp 1-75, 2013.

[3] J. Basu, M. S. Bepari, R. Roy and S. Khan, "Real Time Challenges to Handle the Telephonic Speech Recognition System", proceedings of ICSIP 2012, (pub. at Lecture Notes in Electrical Engineering, Vol. 222, pp. 395-408, Dec. 2012

[4] H. Jiang, "Confidence measures for speech recognition: A survey", Speech communication, Volume 45, pp 455-470, 2003

[5] D. Povey et al., "The subspace Gaussian mixture model: A structured model for speech recognition," Comput. Speech Lang., vol. 25, no. 2, pp. 404–439, Apr. 2011.

[6] S. P. Rath, D. Povey, K. Vesel´y and Jan H. Cernock´y, "Improved feature processing for Deep Neural Networks," INTERSPEECH 2013