

# Samachar Vaani: A Framework for providing Automated Spoken News Service

Rohit Kumar

Department of Computer Science and Engineering, Punjab Engineering College, Chandigarh, INDIA

LTRC, IIT Hyderabad, Gachibowli,  
Hyderabad, INDIA 500019  
rohit@iiit.net

Anuj Singla

CSC India Pvt. Ltd, Pardesipura  
Indore, INDIA 452010  
a.singla@sify.com

Sanjeev Sofat

Sector 12, Chandigarh,  
INDIA 160012  
chestasofat@yahoo.com

**Abstract** – Several web based news services provide news in spoken format also. The current processes involved in recording the spoken news are time consuming and costly. This has restricted several service providers from providing news in spoken form. A new framework that uses automatic speech synthesis to enable existing web based news services to provide spoken news services to the users at no recurrent costs has been presented in this paper.

## 1. INTRODUCTION

Recent advances [1] in the field of automatic speech synthesis have enabled automated synthesis of near natural sounding speech. The availability of such automated speech synthesis systems has enabled creation of applications which can replace a lot of manual work involved in recording of speech. In this paper, one of the applications of speech synthesis in the web services domain has been presented.

The web based news services provide primarily displayed news in form of text articles and pictures. The process involved in generating spoken news is time consuming, manual (hence error - prone), un-economical as well as rigid. Section II discusses the details of the costs involved in the manual process.

Automatic speech synthesis can be used for replacing the existing process of generating spoken news by a process involving much less costs and almost no human effort. A framework for automation of spoken news services and the various issues involved are discussed in Section III. A system for automated spoken news generation for a Hindi news website is described in Section IV.

It must be mentioned here that there have been previous efforts [2] [3] in the direction providing an automated interactive news service by using speech synthesis. The genuinity of our effort lies in that it is aimed at enabling existing news service providers to provide more service with the infrastructure typically available with them at no recurrent costs.

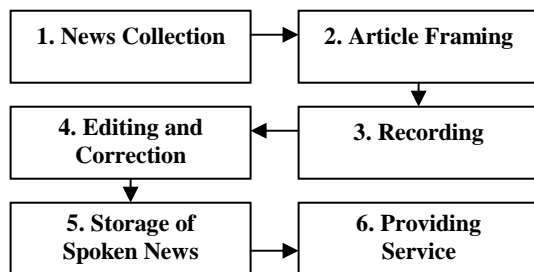


Fig.1 Manual Process of Spoken News Generation

## II. COSTS INVOLVED IN MANUAL PROCESS OF SPOKEN NEWS GENERATION

The manual process for generating the news in spoken format adds heavily to the costs involved in news production. A typical layout of the manual process is shown in fig. 1. Some of the key costs involved in providing spoken news service using this manual process are discussed.

### A. Recording

After Framing of articles, News reports are recorded manually in a studio environment. Payments to the voice talents are involved in this process besides the cost of infrastructure and its maintenance. Though the cost of infrastructure is one time cost, the payments to voice talents is recurrent. Also this manual process of recording adds in the possibility of lot of mistakes leading to re-recording which accumulates into the overall costs.

### B. Editing and Correction

Since the manual process of recording is error prone. Due to this post – recording editing and in some cases re-recordings are required. Editing requires more infrastructure and specialized manual talent. Also re-recording is both time consuming and costly.

### C. Storage of Spoken News

The recorded news needs to be stored and requires a lot of storage infrastructure which is another recurrent cost. Moreover maintenance costs are involved. Spoken news cannot be discarded as it may be needed in future and re-recording is often undesirable.

## III. SAMACHAAR VAANI: A FRAMEWORK

Samachar Vaani is the result of our efforts to come up with a framework for enabling existing news service providers to provide spoken news service with the infrastructure typically available with them. Fig. 2 shows the Samachar Vaani framework.

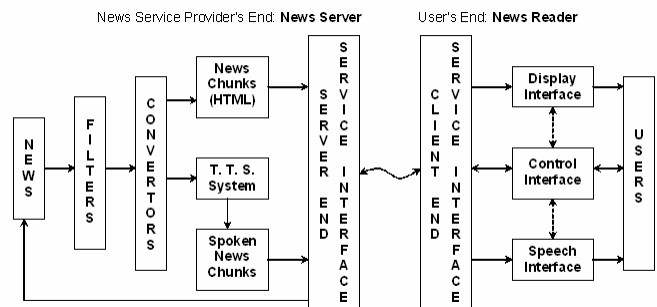


Fig.2 The Samachar Vaani Framework

Samachar Vaani framework comprises of two components namely the News Server and the News Reader. The model of deployment as depicted in fig. 2 would be that the News Server will be installed at the news service provider's end and the News Reader will be distributed to the users through which they can access the news service.

#### A. News Server

The News collected and framed into articles by the News service provider through various sources is accessed by the News Server. These news articles are passed through a set of filters and converters as shown in fig. 2.

The role of filters is to process the news to extract only the content in a suitable form depending upon the way it will be used in succeeding blocks. The news articles are expected to be annotated with markers specifying news category. The filters are basically parsers for these annotation markers in the news content. Besides this filters also add certain desirable prompts in the news content wherever appropriate.

Convertors are plugabble modules into the framework which convert the news content from filters to formats suitable for use with the TTS and the display at the client end. The problem of having so many fonts and encodings for regional languages is solved by using appropriate convertor plugins depending upon the format of the content at source.

A high quality Text to Speech (TTS) System based on concatenative synthesis by optimal unit selection from a large corpus is used to generate the spoken news content. The TTS has been placed on the server end of the framework due to issues related to plugability of new voices and size of speech databases required for high quality synthesis.

#### B. News Reader

The News Reader acts as the client to the News Server and is deployed on the user's end. The News Reader is a small, easy to use, easy to download and easy to install software. To achieve these characteristics it is further desirable to keep the Text to Speech System on the service provider's end.

The Client End Control Interface allows the user to send authentication / identification information and view / hear news content. The control interface brings about the synchrony between both the output interfaces providing the user with a multimodal interface. The control interface also stores the downloaded news content for offline access to the latest downloaded news.

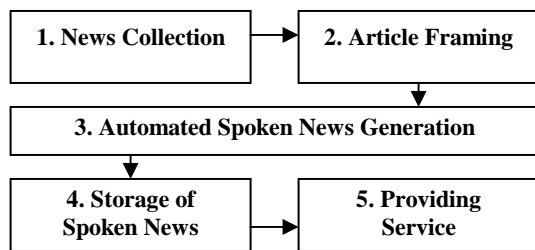


Fig.3 Process of Automated Spoken News Generation

Fig. 3 shows the revised processes and costs involved in automated spoken news generation using a system based on the Samachar Vaani framework. The manual processes of recording and editing are replaced by an automated process which besides bringing down the costs reduces the response time to generate spoken news. Also it should be mentioned that the storage requirements in step 4 of fig. 3 are lesser than those in step 5 of fig. 1 because now only the recently synthesized spoken News content needs to be stored and old news can be re-synthesized whenever required. The use of above framework for generating spoken news does not require any special infrastructure or manpower.

### IV. A SYSTEM FOR INDIAN LANGUAGE SPOKEN NEWS SERVICE

Based on the described framework described, a system for enabling automated spoken news service for the BBC Hindi website. The present system used a generic Indian language TTS System [4] which can be used with large number of Indian languages besides Hindi.

On the server end, a HTML parser parses the web pages available on the BBC web server and provides the filters with news. The news content on the BBC web server is annotated with SPAN tags to specify the type of content. The filters use the annotations to generate chunks of displayed and spoken news content. The spoken news content is converted to z-notation [4] required by the text to speech system.

### V. CONCLUSION

A framework for implementing an automated spoken news service using automatic speech synthesis systems has been proposed. Use of automated speech synthesis can lead to reduced cost and time required for providing spoken news service. The system enables the existing News service providers to provide more service to the users with the infrastructure typically available with them.

### VI. ACKNOWLEDGMENTS

We are thankful to Prof. Rajeev Sangal, Director, IIIT Hyderabad for allowing us to use the Indian language speech synthesis system developed at LTRC, IIIT Hyderabad.

### VII. REFERENCES

- [1] A.J. Hunt and A.W. Black, "Unit Selection in a concatenative speech synthesis system for a large speech database," in *Proceedings of Intl. Conf. on Acoustics, Speech and Signal Processing*, 1996, pp 373-376.
- [2] Ananova, virtual newscaster, <http://www.ananova.com/>
- [3] Ms Khabri, [http://www.corebee.com/khabari\\_details.htm](http://www.corebee.com/khabari_details.htm)
- [4] S.P. Kishore, Rohit Kumar, Rajeev Sangal, "A Data Driven Synthesis Approach for Indian Languages using Syllable as Basic Unit", in *Proceeding of Intl. Conf. on NLP*, pp. 311-316, 2002.