

3.3 Use of Machine Translation and Text to Speech Technology for innovative solutions for Mobile Services

Introduction

This report is intended to present a brief overview of various innovative applications for cellular networks that can be implemented using indigenous Indian language technologies. As the country's cellular phone network grows beyond cities into semi urban and rural areas, cellular companies need to readapt their solutions to meet the needs of a completely different set of consumers, with different demographic and social profiles, along with different requirements.

One of the most powerful by products of the telecom revolution in India has been the sheer scale at which information is now available to different people. The simple SMS (Short Message service) has become a truly powerful information delivery vehicle helping people stay connected at all times, exchange critical information (through person to person SMS, or through SMS information portals), and even to build interpersonal relationships (using applications such as chat).

Despite its growing popularity, SMS continues to face certain limitations – especially when viewed in context of the great diversity of the Indian telecom market. Some of these limitations are:

- Language – Today – most of SMS is predominantly in English, or in Romanised vernacular languages. However, English is still understood by only a small minority of the Indian population, with regional languages still dominating the landscape. As a result, a large percentage of the population may not be able to use the benefits of SMS
- Technology limitations – a large number of telephone subscribers and customers may have access to telephones that do not permit the display of text. As a result, it may not be possible to deliver SMS messages to subscribers using these handsets.

This paper attempts to highlight indigenous technological developments that now allow deployment of innovative applications to surmount

the above two barriers. Along with an overview of the technology, this paper also proposes the various applications that may be deployed with suggested deployment models and issues involved in deployment. Owing to the status of Hindi as the national language and its use as the predominant Indian language of communication, this paper restricts itself to applications for Hindi (which can however be extended gradually to all Indian languages)

Overview of the Technologies

Technologies required for applications in Indian languages broadly require two different kinds of technologies:

1. Input technologies: These would be required to make the information available in Indian Languages and can take the form of:
 - a. OCR (Optical Character Recognition) for Hindi (For converting written text in Hindi into electronic format): These would largely be used by service providers providing content for the applications.
 - b. Machine Aided Translation (MAT) tools (For converting written text in English into Hindi)
 - c. Hindi input mechanisms on cell phones: These would be used by the subscribers of the services to enter text in Hindi on their handsets. These could be further divided into:
 - i. Devanagiri Keyboard lay out / mappings
 - ii. Font support in handset platforms
 - iii. Transliteration support in handsets
 - d. Automatic Speech Recognition (ASR): to allow the subscriber to enter text by speaking into the phone.
2. Output technologies: These would be required to convey the information to the recipient in Hindi and encompass:
 - a. Hindi text display technologies: Linked to (c) above, these would be required to display Hindi text on the screen of the recipient's handset. These could take the form of text messages (and hence

be further editable), or could also be conveyed to the recipient using a popular format known as a picture message.

b. Machine Aided Translation (MAT): To convert English text sent by the service provider or another subscriber in an SMS into Hindi for display on the recipient's handset

c. Text to Speech Technology: This would be required to deliver text information in Hindi in speech format to the recipient. While all the above technologies are under development, two readily available, mature technologies today are Machine Aided Translation and Text to Speech Software. Hence these two technologies have been considered as the primary technologies around which the applications would be built for purposes of this paper.

AnglaHindi – the Machine Aided Translation System

While Translation systems for translating from between International languages have been around for some time, it has been a long cherished dream of Indian linguists to evolve a translation system which can translate text with a high degree of accuracy from English (in the Roman script) to Hindi (in the Devanagiri script). AnglaHindi is the result of this vision, developed after painstaking research over the past 7 years, at the Indian Institute of Technology, Kanpur.

AnglaHindi is a Machine Aided Translation System which translates the English text into its Hindi (Devanagiri) equivalent. Developed by researchers at the Indian Institute of Technology and marketed by Prologix Software Solutions Pvt. Ltd., AnglaHindi effortlessly translates unrestricted English text into grammatically correct Hindi text.

The English language has a set of extremely complex rules which can lead to situations of the same word assuming multiple meanings based on the usage which it is subjected to.

An example of the above can be seen through two forms of the English word 'treatment'

- Maria was being treated at the city hospital

- The King always treated his guests with great warmth and humility

While 'treated' would mean '?????' in the first sentence, in the second it would mean '????????'.

As a result of this complexity and multiplicity of the meanings of a word based on its usage in the English language, any translation effort from English (human or otherwise) is always undertaken based on the usage of the word.

AnglaHindi recognizes this complexity and for situations similar to the above example, allows users to choose from multiple translations. In addition, users get the option of customising/ training their versions of the translator to most correctly render the translation at the first go itself.

Based on the nature/complexity of the input sentences and multiple meanings of a word, the system produces a number of possible translations. Out of these translations, one translation is taken as the first choice and is displayed on the screen.

An easily customisable editor interface allows the user to select and modify the output phrases up to the level of characters, if needed, with the help of an onscreen keyboard.

On an average, users have reported translation accuracies of as high as 70-75% when the translator has been untrained, and approximately 90-95% post training for the particular domain in which the translator is to be used.

AnglaHindi's architecture allows it to be plugged in easily with technologies like Text-To-Speech & Optical Character Recognition to build enterprise wide innovative applications.

Vaachak – the Indian language Text to Speech

Synthesiser

Vaachak is the first commercially available, high quality Text to Speech



Synthesiser for Indian languages. Developed over the past four years at Prologix Software Solutions Pvt. Ltd., Vaachak has the ability to transform written text into clear, audible speech. With an architecture

that can allow customisation to support all Indian languages, Vaachak is currently available for Hindi and work on developing the TTS for other Indian languages is currently in progress.

Based on the latest speech synthesis technology, Vaachak provides the following features for implementation:

- Windows XP, Windows 2000, Windows NT, Windows 98/ 95, Linux compliant
- Natural and Pleasant sounding synthesised speech
- Multiple Male/ Female voices; Frequency samples: 11 KHz (telephony), 22 KHz (Desktop); Compressed version for reduced engine size available
- Supports both Unicode and ISCII input
- Available in Client Server and Desktop editions
- Ability to control pitch/ tempo of speech through easy to use tags/ commands
- Large custom dictionary implementation to handle abbreviations, special pronunciations
- Rule based approach helps control:
 - Option to speak numbers as individual digits (phone numbers, etc.)
 - Indian Currency (Rupees, Paise), Arabic Number System (Crores, Lakh, etc.), Decimal digits

Proposed Solutions

This section attempts to capture the extent of applications which could be implemented for cellular subscribers using the above mentioned technologies and also seeks to highlight deployment models and key issues therein.

English (text) to Hindi (text) SMS

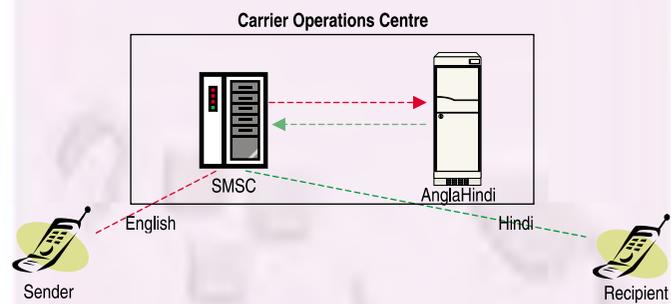
Solution Description:

This service would allow a user to send an SMS in English (Roman), which would be delivered to the recipient in Hindi (Devanagiri). The solution would encompass the range of sms terminology which would be effectively translated into Hindi and then delivered to the ultimate recipient.

Technologies Required:

- AnglaHindi – English to Hindi Translation

Deployment Models:



An AnglaHindi server running a suitably tuned version of AnglaHindi would be running in conjunction with the service provider's SMS at the operations centre. Any message from the user would be translated at the operations centre before being delivered further. The solution could be deployed in the following models:

SMS to designated translation enabled SMSC number

- The service provider designates a special number on the SMSC – for e.g. 2222 for handling English to Hindi Translation
- The subscriber sends an English sms (for e.g.: 'will reach @ 8') through the message centre number 2222
- The SMSC on receiving the message, passes it to the AnglaHindi translation server
- The AnglaHindi translation server translates the message into: '8 ??? ????? ??? ???' and sends it back to the SMSC
- The SMSC forwards the message to the desired recipient

SMS to general SMSC number with special identifier in message

- The service provider does not designate a special number on the SMSC, but programs the SMSC to parse the message for a special identifier at the beginning of the message to translate the message (for e.g. trnhin – for translate to Hindi).

- The subscriber sends an English sms (for e.g.: 'trnhin will rch @ 8') through his normal message centre.
- Based on the rules that have been set in the SMSC, and on the basis of the keyword entered by the subscriber at the beginning of the message, the SMSC passes it to the AnglaHindi translation server
- The AnglaHindi translation server translates the message into: '8 ??? ????? ??? ???' and sends it back to the SMSC
- The SMSC forwards the message to the desired recipient

Enablers Required:

Since most cell phones are able to support messaging in English (in the roman script), the following would be essential towards operationalising this service:

- Recipient's phone to have the capability to display Devanagiri text. This could be done through the following methods:
- Devanagiri text display: If the handset supports this. However this is true only for a very small percentage of handsets in the market today. It is important that handset manufacturers adopt common standards to support Hindi text under guidance from the Ministry and TDIL group and it may thus be required to involve them in this initiative. AnglaHindi currently supports Unicode output and can easily be customised to deliver the translated text in any font format.
- As a Picture Message: Most cell phones are now able to support picture messages, and to bypass the problem of Hindi display being unavailable, in the short term, the Devanagiri message could be sent as a picture message so that it displays correctly on the recipient's computer. For this, a small module would need to be written on the AnglaHindi server that converts the output text into a picture message.
- Requirement of sms corpus for adequate tuning of AnglaHindi: AnglaHindi would need to be suitably tuned to be able to comprehend the specialised sms language and hence a large enough

corpus of sms messages would be required to adequately tune the engine prior to deployment.

Key Issues:

- Hindi Font support: This seems to be the biggest impediment towards operationalising this service. Very few cell phones in the market today have the capability to support Hindi text, and the font standard/ etc. followed by them appear to be proprietary. As a long term measure, it may be important to involve some cell phone vendors in this exercise to be able to support Hindi and other Indian language scripts.

English (Text) to Hindi (Speech) SMS

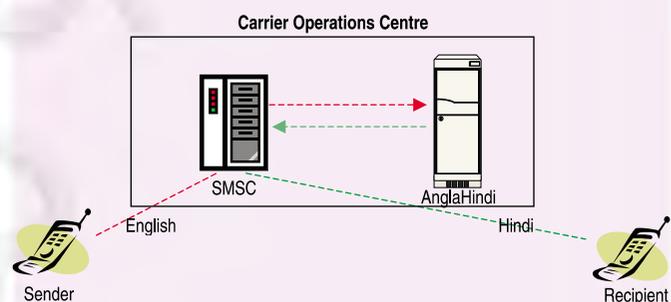
Solution Description:

This solution could enable sms to be sent in English from a subscriber's phone, translated into Hindi at the SMSC, and subsequently delivered to the subscriber as a voice message.

Technologies Required:

- AnglaHindi – English to Hindi Translation
- Vaachak – Hindi Text to Speech Synthesiser

Deployment Models:



It is proposed to deploy both the translation server as well as the text to speech server at the central operations centre to fulfil the translation and text to speech synthesis tasks. Carriers could use either of the above described methods to direct subscribers to an appropriate number from which the message would be translated and then converted into Hindi speech to be delivered to the recipient, i.e.:

- Creating a separate SMSC number for this service and asking subscribers to send messages through this SMSC in case they want it to be translated

and converted into speech

- Or, by keeping the same SMSC number and giving some appropriate keyword before the message (e.g. transpeech) to indicate to the SMSC that the message should be translated and converted into speech. Since both the above options have been discussed in detail earlier, we will restrict ourselves to understanding the processing and delivery options for this service.

Message Processing

- On identification of the message required to be processed at the SMSC, the text is extracted and sent to the AnglaHindi server
- The AnglaHindi server translates the message to Hindi and further sends it to the Vaachak – Text to Speech Synthesis server.
- This server receives Hindi input from AnglaHindi and converts the text into clear speech.
- This generated speech is sent back to the SMSC as an audio file in the appropriate format.

Message Delivery

Message Delivery could happen using either of the following two mechanisms:

Through a multimedia message

If the network and the service provider support advanced features such as Multi Media Messaging (MMS), the message could be sent on to the recipient as an MMS in the form of an audio file, which the person could open to listen to the Hindi message.

The disadvantage however with this approach is that the percentage of handsets which can support features such as MMS is still low in the country and such handsets may also be unaffordable for the target segment of consumers who would find a value in this application.

Through an Outbound call

Alternatively the cellular service provider could automatically place an outbound call to the designated recipient's number. This will help in not

just delivering an SMS to a cellular phone, but also enabling a 'Wireless to Fixed phone' sms feature with translation from English to Hindi

To the VoiceMail box of the user

In case the service provider offers a VoiceMail service and the recipient has subscribed to the service, he message could also be delivered to the VoiceMail box of the recipient as an audio message. This could be accompanied by the generation of an SMS to the user (in Hindi or English – based on user preference) informing him of the special voice message waiting for him, allowing him to pick it up at his own convenience.

Enablers Required:

None – all components required for enabling this service are currently available and ready for deployment

Key Issues:

Technology issues:

- Mode of delivery of Voice Message (VoiceMail, MMS, Outbound call) – This would need to be resolved for implementation to proceed

Hindi SMS to Speech

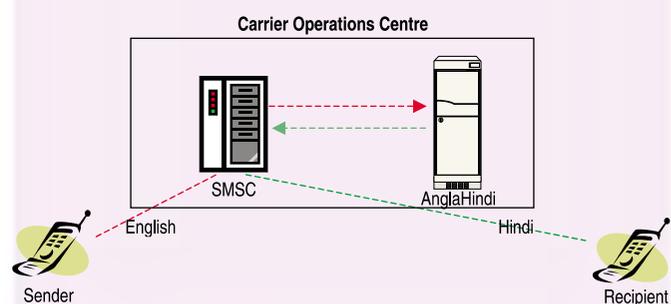
Solution Description:

This service would allow a person to send a Hindi SMS, which is converted into Hindi speech and then delivered to the recipient.

Technologies Required:

- Vaachak – Hindi Text to Speech Synthesiser

Deployment Models:



Deployment of the solution would be similar to the models discussed above:

- The sender types a Hindi message on his cellular phone and sends it to the SMSC
- The SMSC extracts the Hindi text and sends it to the Vaachak Server
- The Vaachak Server converts the Hindi text into speech and sends it back to the SMSC
- The audio file containing the synthesised text is delivered to the recipient using any of the delivery models discussed in 5.2.3.2 above.

Enablers Required:

- Hindi input mechanisms on cell phones: These would be used by the subscribers of the services to enter text in Hindi on their handsets. These could be further divided into:
 - a. Devanagiri Keyboard lay out / mappings
 - b. Font support in handset platforms
 - c. Transliteration support in handsets

Key Issues:

- Hindi font support for the range of cellular equipment – from handsets to SMSCs, with an option to support Unicode character sets is a critical component towards operationalising this service. It would be in the long term interests of all service providers that all technology/equipment vendors are able to converge onto a common standard for supporting Indian language fonts for value added applications for the future.

Hindi (text) to English (text) SMS

Solution Description:

This service would allow a user to send an SMS in Hindi (Devanagiri), which would be delivered to the recipient in English (Roman).

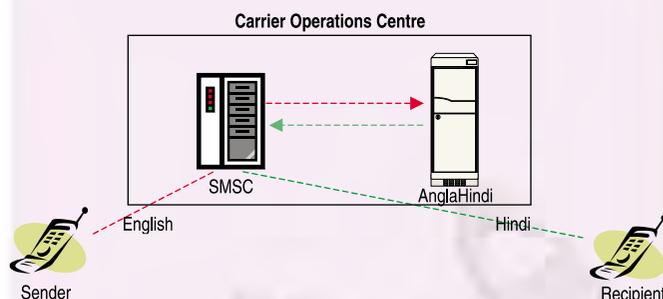
Technologies Required:

- AnglaHindi – Hindi to English Translation (to be released shortly)

Deployment Models:

The service would work similar to the English (text) to Hindi (text) service described earlier.

Enablers Required:



- Hindi input mechanisms on cell phones: These would be used by the subscribers of the services to enter text in Hindi on their handsets. These could be further divided into:
 - Devanagiri Keyboard lay out / mappings
 - Font support in handset platforms
 - Transliteration support in handsets

Key Issues:

- As discussed earlier, Hindi font support across the various building blocks of the solution is a critical component of the solution and would need to be resolved so that this service can be implemented effectively

SMS portal information in Hindi

Solution Description:

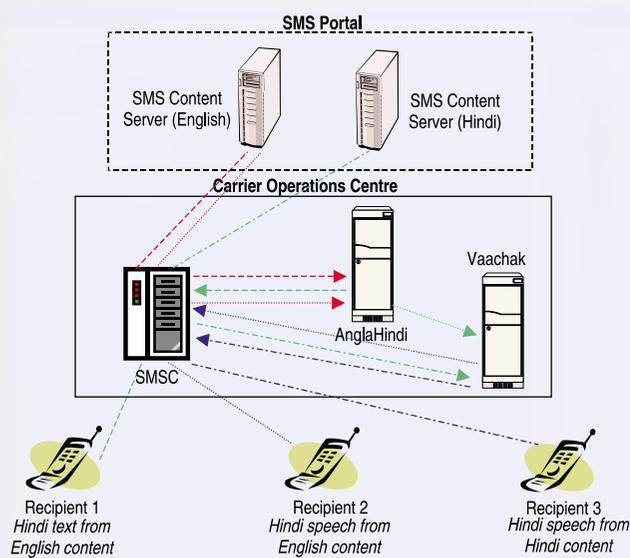
This service would allow information from SMS portals (such as Indiatimes 8888, yahoo, etc.) be made available in Indian Languages (either in text or speech format) to cellular and other phones.

Technologies Required:

- AnglaHindi – English to Hindi Translation
- Vaachak – Hindi Text to Speech Synthesiser

Deployment Models:

Deployment models could be manifold depending on the kind of service that is required (whether it is an auto SMS alert service, or a SMS info. request service, etc.). This paper restricts itself to discussing the Hindi information processing and delivery to the user of a mobile or fixed telephone.



English SMS information to Hindi

- An SMS Content Server sends the information in English to the SMSC to forward to the recipient.
- Based on the user's preferences (at the SMS portal, or as part of the request sent for the information by the user), this information is either forwarded (in English) by the SMSC to the recipient, or sent on for further processing:

a) The message content is sent on to the AnglaHindi translation server, where it is translated into Hindi and then returned to the SMSC from which it is forwarded to the user. Message handling in this case is similar to the English (text) to Hindi (text) SMS application described earlier

b) Post processing by the translation engine, the message is further forwarded to the Vaachak Text to Speech Synthesiser, where it is converted into speech and returned to the SMSC for delivery to the recipient. Message handling in this case is similar to the English (text) to Hindi (speech) SMS application described earlier

Hindi SMS information to Speech

- An SMS Content Server sends information in Hindi to the SMSC to forward to the recipient.
- Based on the user's preferences (at the SMS portal, or as part of the request send for the information by the user), this information is either forwarded

(in Hindi) by the SMSC to the recipient, or sent on for further processing:

- In case the message is to be converted into speech, it is sent by the SMSC to the Vaachak TTS server from where it is returned to the SMSC as an audio file to be forwarded to the recipient. Handling in this case is similar to the Hindi SMS to Speech application described earlier.

Enablers Required:

- Enablers required are similar to the requirements for English to Hindi SMS and SMS to voice applications as discussed earlier

Key Issues:

The services discussed so far require some key issues to be resolved on the path to deployment:

- Standardisation of Hindi font support across:
 - Handsets: for display, input and handling of text
 - SMSCs: for processing of text
 - SMS Portals: for generation of Hindi text in a standardised format so as to enable seamless integration with the SMSCs and technologies such as AnglaHindi and Vaachak
- Understanding of SMSC applications: A key element in the above applications is that of the SMSC and the way it handles the text messages (forwards them directly, sends to a particular server just for translation, or instructs the translation server to forward to TTS for conversion into speech). It will also be important to understand the SMSC interfaces in detail and use this understanding for successful integration with the translation and TTS technologies.
- For all translation requirements, collection of a sms/ text corpus for tuning the translation engine for most optimum translation will also be a critical exercise in the deployment exercise.

Hindi Voice Portals

Solution Description:

The concept of a voice portal is becoming extremely popular amongst telecom service providers

worldwide with applications such as news, weather reports, unified messaging, etc. on demand taking on the role of key value added applications. The concept of a voice portal is extremely simple:

- A subscriber calls into a published voice portal number
- On calling into the voice portal, the subscriber is directed to an Interactive Voice Response System (IVRS) which offers a menu of options, for e.g.:
 - Current News
 - Weather
 - Traffic Information
 - Sports News
 - Ring tones
 - Email access. Etc.
- The subscriber can select the appropriate option either through Speech Recognition (ASR) Technology, or by pressing the keys on the telephone (for e.g. 1 for Current News, 2 for Weather information, etc.)
- On selecting the appropriate option, the information is downloaded from an appropriate resource (in text), converted into speech in real time and played out to the user.

Some of these services already exist in some for or the other in different cellular networks, however there is currently no service available to service the demands of subscribers in Hindi and other Indian languages.

These services are different from the sms based services described earlier principally because of two reasons:

- Higher interactivity with the subscriber – in which the subscriber is an active participant in the application
- The subscriber calls into the application as opposed to text messaging, which also enables other applications to be deployed – for e.g. voice chat, conferencing, etc.

Technologies Required:

- Vaachak – Text to Speech Synthesiser
- AnglaHindi – Machine Aided Translation tool

Deployment Models:

The deployment of this service would happen similar to the services described earlier. However the key difference would be that instead of being routed through the SMSC, the information would now need to be routed through an Interactive Voice Response System, which would in turn communicate with the Information database and text to speech servers to deliver the information to the subscriber who has called in.

Enablers Required:

- Interactive Voice Response Servers (to manage the calls coming into the system and rout the relevant information)

Key Issues:

- The technology exists for operationalising and enabling this service. Ease of deployment is greater than services that work in conjunction with the SMSC – primarily because there is hardly any need to ensure commonality of font standards between different parts of the solution. Text would be stored in the database, from which it will be picked up by the IVRS and forwarded to the TTS for conversion into speech. Additionally, since the user interaction happens through speech and telephone keypad based commands, the integration is much easier.
- Voice Portals are powerful applications. However like conventional ‘web portals’ the content on voice portals needs to be well planned in order to generate a high subscriber interest and return calls. It therefore also requires constant attention, monitoring and maintenance. It should therefore also be a business decision that is well thought of by the service providers interested in implementing this solution.