

Wikipedia on Lower End Mobile

Rashmi.M¹, Chaithra M²

^{1,2}GSSS Institute of Engineering and Technology for Women, K R S Road, Metagalli, Mysore, India

Abstract: *Wikipedia on lower end mobile is a tool mainly designed for visually disadvantaged people and people having lower end mobiles. Wikipedia on lower end mobile is nothing but a mechanism to listen to Wikipedia contents through receiving SMS. The advantage of this service is that we can get instant information, which does not require reading. We can use our travelling time to know about some useful stuff. In addition, listening to information reduces strain on the eyes.*

Keywords: Wikipedia, text-to-speech, SMS Gateway, Server

1. Introduction

The aim of this paper is to bring the Wikipedia information on mobile through voice. No connection to the mobile internet is required. Mobile users are able to request Wikipedia article via SMS. An audio representation of the requested article will be generated in real time by the help of computer synthesized speech (text-to-speech) technologies. The end output will be like this, if we send "Engineering" from our mobile as SMS to a particular number assigned for this service, we will get a call from the server and we will be able to listen the Wikipedia entry for "Engineering" as voice.

2. Problem statement

Wapedia is a WAP (Wireless Application Protocol) site created by Florian Amrhein and operated by Taptu, which is a new kind of search engine for mobile phones that allows us to search and find really useful contents. Wapedia brings the contents of Wikipedia to mobile devices like mobile phones and PDA's (Personal Digital Assistance) in text form. The main disadvantage of the WAP solution is the large amount of data which will be generated by each query. Large text documents cannot be displayed on the limited screen size available in the mobile phones. For mobile Internet users without a flat fee, the download may be very costly. The downloaded file has to be stored locally on the device. Some mobile phones have very limited resources to store files. GPRS or UMTS mobile Internet access is required and these facilities are available only in smart phones and PDA's. In addition, this service cannot be used by visually disadvantaged people.

3. Solution

As an alternative to display large text documents on very small displays, audio based Wikipedia accessing can be a solution for hand held devices. To avoid those communication costs and the limitation to users of the mobile internet a pure SMS based service is created. No GPRS or UMTS mobile Internet access is required. The service works with all mobile phones today without the need of additional configuration or expert knowledge. The costs are transparent for the users, only communications fees for an SMS request and a land line call have to be paid. Audio representation of Wikipedia contents can also be used by visually disadvantaged people without modification. Especially to generate an audio representation of a text, dynamically text-to-speech conversion is the only solution.

4. Advantage

- The service works with all mobile phones without need of additional configuration or expert knowledge.
- Costs are transparent for the users, only communication fees for an SMS request and landline call charge has to be paid.

4. Working cycle

The system works like this, if a user sends "Bill_Clinton" as SMS to a service number after an approximation of 2 to 3 min he/she will get a call through which he can listen to the information about "Bill Clinton".

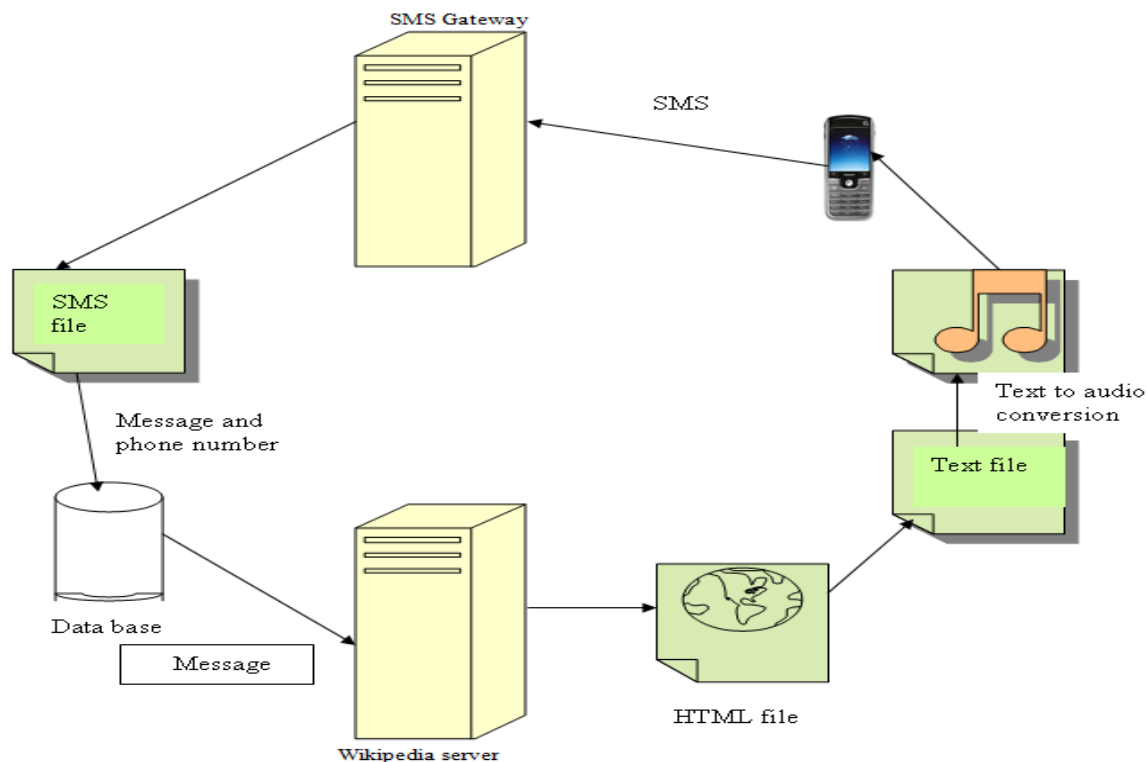


Figure 1: Transfer of information from server to mobile

In the server the following process takes place as soon as the SMS is received from the user:

- The received message is retrieved and stored in the database
- Using this message, the particular information about the requested article is obtained and stored in a file.
- Then by using the computer synthesized voice the article is read out to the user who sent the message through making a phone call to him.

This system is divided into four modules.

Module 1: Retrieving the message from the user and inserting it into the database.

We use the reliance SMS gateway to receive the SMS from the user. The SMS sent by the user is stored in the SMS.DTC file created by the SMS gateway and it also stores other information like, their phone number, date and time of sent message etc. The new message received by the user will be appended in the same file. The phone number and message are stored after a sequence of special characters (3f J), the hexadecimal value of the special characters are used for retrieving the phone number and the message. The retrieved phone number and messages along with their status are stored in the database.

Code for retrieving the message.

```
import java.io.*;
import java.lang.*;
import java.sql.*;
import java.lang.String;
```

```
public class Retrieve_msgf
public static void main1()throws IOException
```

```
{try{BufferedInputStream a=new BufferedInputStream(new
FileInputStream("c:\\ProgramFiles\\Mobile
Partner\\userdata\\SMS.DTC"));
byte b;
int x,as=0,as1=0,as2=0;
String store=new String();
String sms=new String();
String phno=new String();
String temp=" ",msg="",n="9",hex="",hex1="",hex2="";
int count=0,c;
Class.forName("sun.jdbc.odbc.JdbcOdbcDriver");
Connection co=DriverManager.getConnection
("jdbc:odbc:sms", "system", "rashmi");
PreparedStatement s1=co.prepareStatement("insert into
newsms values(?,?,?)");
PreparedStatement s2=co.prepareStatement("select count(*)
from newsms");ResultSet r =s2.executeQuery();r.next();
c= r.getInt(1);
System.out.println("c"+c);
b=(byte)a.read();
if(b!=-1)
{for(int i=0;i<356;i++)
{b=(byte)a.read();}
for(int i=0;i<11;i++)
{b=(byte)a.read();
phno=phno+(char)b;}
for(int i=0;i<490;i++)
{b=(byte)a.read();}
for(int i=0;i<128;i++)
{b=(byte)a.read();
msg=msg+(char)b;}
count++;
System.out.print("count"+count+"c"+c);
System.out.print("phno"+phno);
System.out.println("msg"+msg);
if(count > c)
```

```

{s1.setString(1,phno);
s1.setString(2,msg);
s1.setString(3,"b");
s1.executeUpdate();
}
msg="";phno="";
}while(b!=-1)
{int j=0;b=(byte)a.read();as=(int)b;
hex=Integer.toHexString(as);
b=(byte)a.read();as1=(int)b;
hex1=Integer.toHexString(as1);
hex=hex+hex1;b=(byte)a.read();
as2=(int)b;hex2=Integer.toHexString(as2);
hex1=hex1+hex2;if(hex!="00")

if(hex.equals("4a0")||hex1.equals("4a0")||
hex.equals("4b0")||hex1.equals("4b0"))
{for(int i=0;i<12;i++)
{b=(byte)a.read();
phno=phno+(char)b;}
for(int i=0;i<490;i++)
{b=(byte)a.read();}
for(int i=0;i<128;i++)
{b=(byte)a.read();
msg=msg+(char)b;}
count++;if(count>c){if(phno.charAt(0)=='9'||phno.charAt(0)
)=='0'){s1.setString(1,phno);s1.setString(2,msg);
s1.setString(3,"b");s1.executeUpdate();
}}
System.out.print("count"+count+"c"+c);
System.out.println("phno"+phno+"msg"+msg);
phno="";msg="";hex=" ";hex1=" ";hex2=" ";}
else{hex=" ";hex2=" ";hex1=" ";}}
catch(Exception e)
{System.out.println("error"+e);}
}
}

```

Module 2: Accessing the wikipedia and storing its contents into a text file.

To get the wikipedia content for the requested message received from the user, the message whose status is marked as 'n' (new message) is appended with the URL <http://en.wikipedia.org/wiki>. Then the connection is established with the URL. The wikipedia content is stored in html file.

As HTML file along with plain text contains HTML tags. These tags are removed using HTML parser and the plain text is stored in the .txt file. Once this process is completed, the status of the message in the database is changed as 's' (sent).

Module 3: Audio generation

In this module audio is generated from the text file using Java Speech API (JSAPI) which defines packages for speech synthesis [1]. JSAPI enables developers to write application that do not depend on the proprietary features of one platform or one speech engine. The javax.speech package defines classes and interfaces that define the basic functionality of an engine. The javax.speech.synthesis package and javax.speech.recognition package extend and

augment the basic functionality to define the specific capabilities of speech synthesizers and speech recognizers. The basic processes for using a speech engine in an application are as follows.

- Identify the application's functional requirements for an engine (e.g, language or dictation capability).
- Locate and create an engine that meets those functional requirements.
- Allocate the resources for the engine.
- Set up the engine.
- Begin operation of the engine - technically, resume it.
- Use the engine
- De-allocate the resources of the engine.

Module 4:

Sending audio file to mobile with the help of voicent gateway, hostname and port values are assigned. Then automatically server will make a call to the number specified and plays the audio message.

5. Database Design

This Table 1 is populated for all 'Retrieve message' and 'Accessing Wikipedia' modules the SQL operation to be verified is insert.

Table 1: Structure of newsms table

Table Name	Field name	Data type	Remarks
NEW SMS	PHNO	VARCHAR2	Retrieved from SMS.DTC file
	MSG	VARCHAR2	Retrieved from SMS.DTC file
	STATUS	VARCHAR2	The status of the new message will be initially 'n' after processing it will modified to 's'.

5. System Design

Figure 2 shows flow of the process. When SMS is sent by the user to the server, SMS content is extracted and searches for the information in the Wikipedia. Downloaded Wikipedia content will be converted in to audio file which will be played after calling back the user.

6. Conclusion

The service itself is quite usable for short Wikipedia articles. It has been reported by some users that listening to a synthetic voice for a long time is inconvenient. The Wikipedia on lower end mobile service is more usable with short articles just to refresh knowledge or to recover facts and issues. Especially the mobile phone version can be useful to act like a 'Hitchhiker's Guide to the Galaxy' for mobile users in their private everyday life. The control systems engineering group of the University uses the similar technique to generate audio teaching material (and a Podcast) out of text based material.

Audio representation of the articles make the wikipedia content available to those, who can understand English but can not read. Visually impaired people are able to use this service.

7.Future Scope

Future work on ‘Wikipedia on lower end mobile’ can cover the implementation of new regional languages like Hindi, Kannada, etc and can improve the mobile phone based user interface (the voice menu). As computer synthesized voices are not pleasant to listen for long time we can build databases of our own voices.



Chaithra M received the B.E. degree in Information Science from Visvesvaraya Technological University in 2007 and MTech in Computer Science and Engineering from PES college of Engineering, Mandya. Her current research interests are Wireless Sensor Network Security and Cloud Computing. She is currently working as Assistant Professor at the Department of Computer Science and Engineering, GSSSIETW, Mysore.

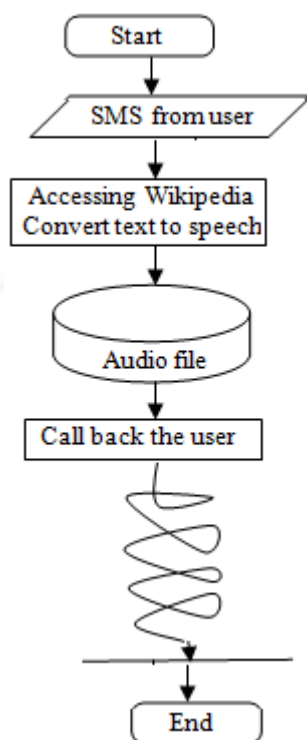


Figure 2: System Flow Chart

References

- [1] Y.Sagisaka, "Speech synthesis from text", IEEE communication magazine, vol.28,no.1,pp.35-41,1990.
- [2] O. Bohl, S. Manouchehri, and U. Winand. "Mobile information systems for the private everyday life. Mobile information systems", 3(3-4):135 – 152, 2007.
- [3] Rui Ren, Zhenjiang Miao "Emotional speech synthesis and its application to pervasive E-learning", Ubi-Media Computing, 2008 First IEEE International Conference , 431 – 435, 978-1-4244-1866-4, 2008.

Author Profile



Rashmi M received the B.E. degree in Information Science from Visvesvaraya Technological University in 2009 and MTech in Information Systems and Management from University of Mysore in the year of 2011. Her current research interests are Wireless Sensor Network Security and Cloud Computing. She is currently working as Assistant Professor at the Department of Computer Science and Engineering, GSSSIETW, Mysore.