RETRIEVING WEB PAGES VIA SHORT MESSAGE SERVICE

Vidya .K.B[1], K. Suriyakrishnaan[2] Dr. D. Sivakumar [3], Shalini Singh[4], S. Supretha[5]

[1,4,5] *Student, Dept. of Electronics and Communication Engineering, Easwari Engineering College, Ramapuram, Chennai*

[2]*Assistant Professor M.E ,(Ph.D), Dept. of Electronics and Communication Engineering, Easwari Engineering College, Ramapuram,Chennai*

[3] *Professor M.E, Ph.D, Dept. of Electronics and Communication Engineering, Easwari Engineering College,Ramapuram,Chennai*

---***--

ABSTRACT: *In Remote areas, situations might arise that there is no internet connectivity but an availability of telecom signal, this situation will create an inconvenience in this digital era. To overcome this issue we have developed a project which can access web pages through telecom SMS. This can be done by using GSM modem. This acts as an SMS server. PC connected to the internet will act as a web server and for interfacing web server and SMS server, python is used. The query will be sent in the SMS text format. Python does the function of HTML parsing .The client is the mobile handset through which queries are sent through SMS text format. Python recognizes the URL sent by the android application to the server in SMS format. Python also collects information from the web server and HTML parsing is done. Finally, the message is sent to the mobile handset. The message is received by the Android application in the handset and is converted to HTML and the webpage is opened in a web browser.*

Keywords: GSM Modem, HTML parsing, SMS server, Web Server

1. INTRODUCTION

Text -based messages are sent to and fro from mobile telephones on GSM network using Short Message Service (SMS).The maximum length of each message is 160 characters. According to [1] there is a trend that SMS isbecoming more and more popular because it facilitates more extended modes of communications.

Also, from [1] Messaging Trends Survey in 2007, 25% of respondents sent SMS from their cell phones, including one in three (32%) teens in America (Knowledge Networks, 2007).

SMS usage in Africa is rapidly growing .It can be used for a wide variety of uses , for example, diagnoses and treatment of health workers is done using bulk message service in Mozambique, Uganda and Malawi .

Transactional alerts that are sent through SMS whenever a change occurs in a bank account are SMS based. For example, whenever a credit card is used, SMS is sent to the mobile phone that is linked with the credit card. Mobile marketing has become very popular since the beginning of 2000s in

European countries and Asia. Real time notifications are the necessity of many applications nowadays.

People who are ignorant about a lot of events are mainly in remote areas. SMS is a very simple way of communication for even the illiterates. SMS has become widely popular because it can get through even when the network is busy for hours. [2]The short message business has emerged very well due to the rapid development of internet and mobile network.The colleges , supermarkets and weather forecasting centers are the various marketing and notification providers who use an emerging technology, the SMS gateway.

SMS messaging has its great potential in car parking systems, M-learning, automated agriculture answering system, market information etc. The SMS system has become a great advantage to farmers. Even if the farmers are unaware of using the internet, they can send SMS using a mobile phone and get instant replies regarding the market price of various crops. In this way, the crops are sold at reasonable rates which is beneficiary to farmers.

The whole idea is to know where SMS can be used to help human life and to provide comfort. The SMS based information system is widened in this proposal to provide an offline internet service using SMS information service.

2. LITERATURE SURVEY:

The importance of SMS based system in mobile marketing is dealt here. As in [3] The American Marketing Association suggest sequential marketing stages as well as temporal and spatial separation of buyers and sellers. The most vital information for farmers today is market opportunities and commodity prices, training farmers on using market information and weekly prices [4].Shortcomings of this system is that the current technologies limit each message to a maximum of 160 characters.

Next is the smart car parking technology which is used to find the park- and- ride facility locations. Under [5] m-parking technology, customers initially register their mobile number, the license plate number of the car and their credit card number with the car park authorities, to create what is called

a virtual parking ticket account. This completes the initial, one-time registration. Whenever the customer wants to park his car, he should send a text message which consists of the car number, location code of the parking area and the time duration for which the car should be parked. In return, he receives an acknowledgement text from the car park message centre. A reminder would be sent 10 minutes earlier the ticket could expire along with the bill which can be payed using a credit card. [6]Bulk SMS Software is a software application or web based software application that allows sending out mass SMS text messages from computer PC via a GSM device or the Internet network connection

SMS administrative support is very useful for library applications. As in [7], libraries can reach out and serve students ubiquitously by sending and receiving SMS-based library information.SMS base services are very helpful for library applications. Alerts such as renewal date, due date can be sent through SMS alerts to the customers.

The next application is how the SMS based market system is useful to farmers. Marketing of farmer's products into the market at a reasonable profit is still a struggle for a lot of farmers, whereas the modern society has many technological aids such as web based and mobile applications. In Kenya , an information service company known as the M Farm offers SMS solutions for selling products. The farmers depend on the system to get information on price for their produce. This helps farmers to make a better sale and to compare price of products at various markets without making use of the shrewd middleman. The farmers are facilitated greatly by the growth and popularity of the SMS service. It helped the farmers with two other components that allowed farmers to come together and to buy bulk inputs. With higher volumes, famers can buy in bulk and transport costs are lower. Here, the farmers specifies the quantity of input needed and the desired location.
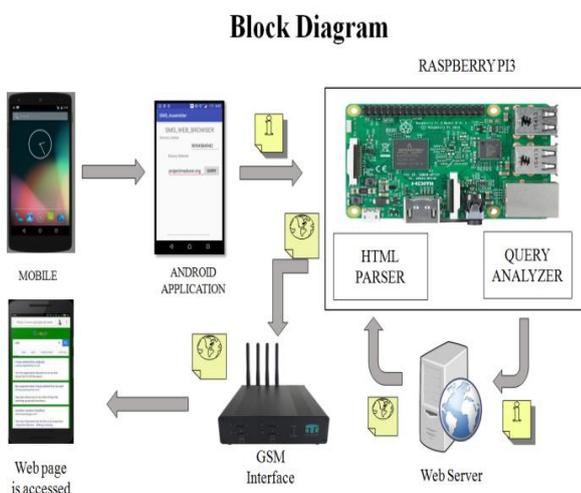
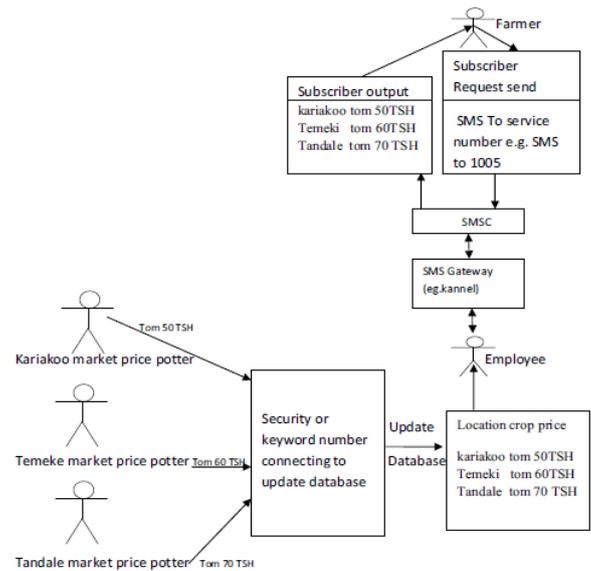## 3. SYSTEM MODEL:

Fig.3.1. System model

Figure 1 .Diagram of an SMS-based market information system

Fig .3.2. Diagram of an SMS based market information system.

Advantages of SMS based marketing system: Marketing of products is a big challenge to many small scale farmers. There is information Service Company, MFarm which offers an SMS based solution for selling their products in Kenya and it uses a module for price information service which runs on code 3535. The farmers rely on the system to get information on price for their produce. This enables the farmers to get the real time market price information, and so they don't have to deal with the shrewd middleman. Through this, farmer can compare the price of products at various markets and make up their minds on where their produce will fetch the best price. With the help of SMS services, the farmers were able to come together and buy inputs in bulk quantity which lowers the transport cost. It can be easily understood by farmers as they have to only send an SMS specifying the needed input quantity and the location. M-Farm identifies the companies that deal with the stated input. If the farmers like the offered price, they place an order. Farmers can also sell in bulk.

a. HARDWARE REQUIREMENTS:
* USB GSM Modem
* Raspberry pi

b. SOFTWARE REQUIREMENTS:
* Python IDLE
* Android Studio
* Internet Protocol

4. PROPOSED WORKING:

A successful implementation of Offline SMS Web server is developed using Python as Front end and GSM server as Backend, Android studio App design act as a middle ware which generates web queries by means of SMS text and the

SMS web queries are handled using a dedicated GSM SMS Server.

The received SMS queries will be further converted as URL query to "**beautiful soup**" python package, the paragraphs in the web URL are extracted as soup text and the split html paragraphs are again converted as SMS contents and finally organised as single SMS in Android APP.

The proposed design is fluent in parsing the webpages and immediately transferring the files into SMS texts.

Keywords: beautiful soup, android studio app.

5. MODULE DESCRIPTION:

a. RASPBERRY PI 3

Fig.5.1 Raspberry Pi 3

The Raspberry pi 3 acts as a mini computer. It has a quad-core ARM Cortex-A53 processor. The Raspberry pi 3 is enhanced with 2.4 GHz Wi-Fi 802.11n (150mbps) and Bluetooth 4.1(24 Mbps).It also has a 10/100 Ethernet port.

In general the raspberry pi can be operated with any type of generic USB computer mouse and keyboard and is also used with USB data storage capabilities.

b. PYTHON:

Fig.5.2 Python program

The python program is used for interfacing the SMS and Web Server. Beautiful soup concept is used to extract the required data from the HTML and XML files. Hence the web scraping is done using the beautiful soup concept in python.
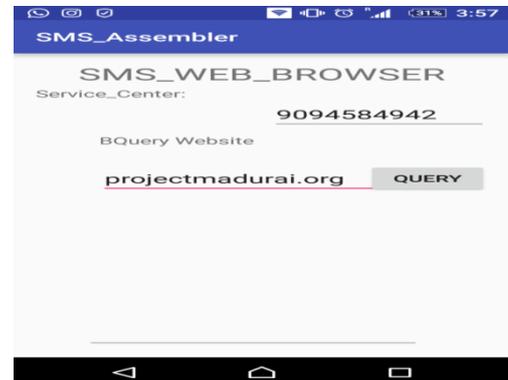
c. ANDROID APPLICATION:

Fig.5.3. Android application

Android is a Linux-based operating system for mobile devices. Android application in the mobile will act as the client and helps to send the queries through the SMS format.

Through android studio applications can be run on different devices which is powered by the application developed.

6. WEB SERVER:

INTRODUCTION

A web server is a server software or hardware that processes requests via HTTP (i.e.) the Hypertext Transfer Protocol. Through HTTP the communication between the client and the server is done. The incoming networks are processed by the web server and serves contents to the World Wide Web.

A web server usually serves content using the HTTP (Hypertext Transfer Protocol). The contents are mostly HTML images and documents but any type of file can be included.

An IP address (Internet Protocol add**ress**) is nothing but an identifier which is assigned to each computer and other device connected to a TCP/IP Network that is used to locate and identify the node in communications with other nodes on the network.

Domain Name System (or Service or Server), an Internet service that translates domain names into IP addresses. Because domain names are alphabetic, they are easier to remember. The Internet however, is really based on IP addresses. Whenever a domain name is used , a DNS server will translate the domain name into its corresponding IP address. For example, the domain name www.example.com might translate to 198.105.232.4**.**
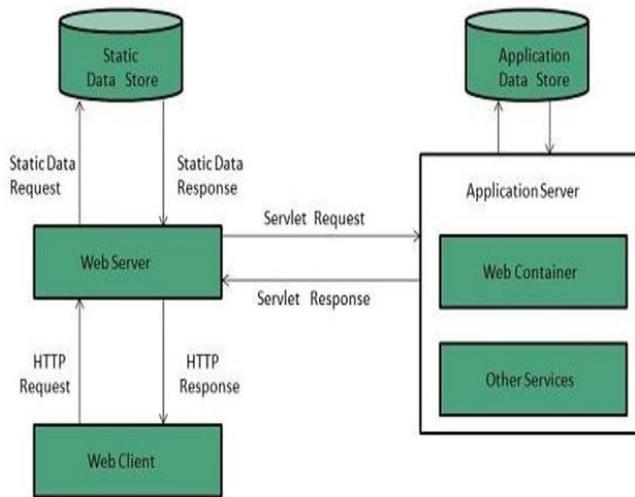
WORKING:

Fig.6.1. Working of a web server

When you enter in the address field of your browser the URL http://www.aprelium.com/doc/sample.html

First, the browser slices the URL in 3 parts:

- http://: This indicates the document to be accessed can be extracted from the web server. The standard form of communication between the browsers and the web servers is the HTTP protocol.

- www.aprelium.com: The required document can be downloaded from this host name

- /doc/sample.html: This indicates the virtual path of the document in the www.aprelium.com's web server.

The IP address of the computer will be calculated by the DNS server when the domain name www.aprelium .com is given. The ISP or the company generally runs the domain name server.

A connection channel is established by the browser with the web server on the computer in which the IP address was given by the DNS server and it requests the document on the host whose name is www.aprelium.com and which virtual path is doc/sample.html. The host name has to be specified by the browser as modern web servers has the ability to serve more than a single computer having a single IP address only. In such a case, the IP address of this computer is associated with more than one domain name.

The requests are decoded and the virtual path is mapped to a real one by the server such that it matches the existing file. The file is sent to the browser by the server with some useful information such as its last modification time and its MIME type. The browser decides how to display the receive documents using the Multipurpose Internet Mail Extension (MIME), in the considered example , it is found to be an HTML file. Thus the server sets its MIME type to text/html format and hence the browser understands that the file should be rendered in the form of text.

Sometimes a URL is entered without an explicit filename such as http://www.aprelium.com/doc. The browser sends the request to the web server as in the previous example. The server detects that the virtual path maps to a directory and not to a file. It searches then in this directory an index file. Index files are usually named index.html or index.htm. If it finds for example index.html, it acts as if the requested URL washttp://www.aprelium.com/doc/index.html. If no index file is found, the web server generates a listing of the directory contents and sends it to the browser or reports an error.

6.1. Some of the Leading WEB SERVERS

- **Apache HTTP Server**

This is the most popular web server. The Apache Software Foundation developed it. Apache web server is open source software. It can be installed on almost all OS including Linux, UNIX, Windows, FreeBSD, Mac OS X and more. About 60% of the web server machines run the Apache Web Server.

- **Internet Information Server (IIS)**

A high performance web server from Microsoft is the Internet Information Server. IIS runs on Windows NT/2000 and 2003 platforms and also on upcoming windows version. IIS comes bundled with Windows. It is easy to administer IIS as it is tightly integrated with the operating system.

- **Lighttpd**

The lighttpd is an open source and free web server that is distributed with the FreeBSD OS. It is fast, secure and consumes much less CPU power. Lighttpd also run on windows, Linux, Mac OS X, and Solaris operating systems.

7. ANDROID STUDIO:

Android Studio[8] is the official integrated development environment (IDE) for the Android platform. It is built in purpose for Android to improve your development and helps in building the highest-quality apps for all the Android devices. It offers the custom-tailored tools for Android developers which includes code editing, debugging, testing, and profiling tools.

Android is a Linux-based operating system for mobile device. The Open Handset Alliance, led by Google, and other companies developed Android.

[9] Developers need to develop only for Android and their applications should be able to run on different android

powered devices which is called as the unified approach offered by Android.

7.1. FEATURES OF ANDROID STUDIO

- Instant Run
- Intelligent code editor
- Fast and feature-rich emulator
- Robust and flexible build system
- Designed for teams
- Optimized for all Android devices

7.2. Steps to Create a Simple Android Application using Android Studio
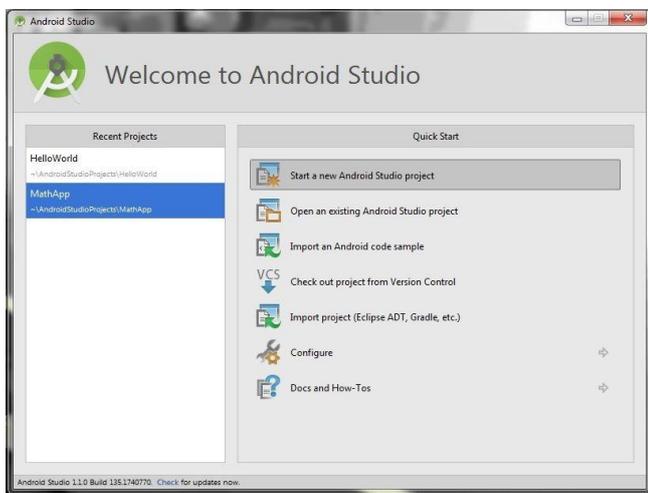
Step 1: Install Android Studio

Fig.7.2.1. Installation of Android Studio

Step 2: Open a New Project

- Open Android Studio.
- Under the "Quick Start" menu, select "Start a new Android Studio project."
- Name the project as "Hello World" on the "Create New Project" window that appears.
- Note where the project file location is and change it if desired.
- Click "Next."
- See that the "Phone and Tablet" is the only box checked.
- The minimum SDK should be below the phone's OS level if the app is to be tested on a phone. Click "Next."
- Select "Blank Activity."
- Click "Next."
- All of the Activity name fields should be left as they are.
- Click "Finish."

Step 3: Editing the Welcome Message in Main Activity

- Go to the activity_main.xml tab in case it is not already open.
- The Design tab should be open on the activity_main.xml display.
- Click and drag the "Hello, world!" from the upper left corner of the phone display to the center of the screen.
- On the left side, there will be a project file system from which values folder should be opened.
- Double-click the strings.xml file from the values folder.
- In this file, find the line "Hello world!"
- After the "Hello world!" message, add "Welcome to my app!"
- Navigate back to the activity_main.xml tab.
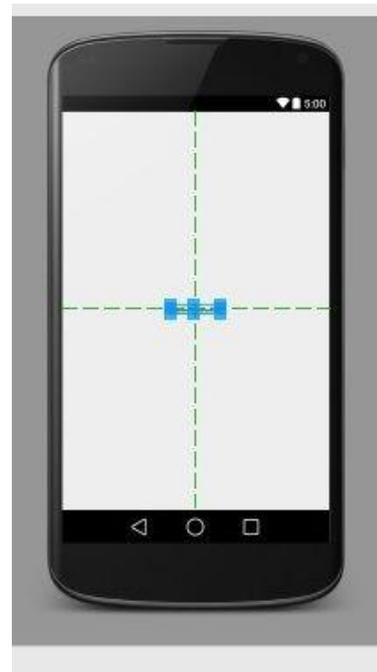- Make sure that your centered text now reads "Hello world! Welcome to my app!"

Fig.7.2.2 Editing the Welcome Message in Main Activity

Step 4: Adding a button to the Main Activity

- Go to the Design tab of the activity_main.xml display.
- Find Button , under the heading Widgets, from the Palette menu to the left of the phone display.
- The Button to be centered is clicked and dragged underneath the welcome message.
- The button should be still selected.
- Find the field for "text" in the Properties menu which is on the right side of the window.
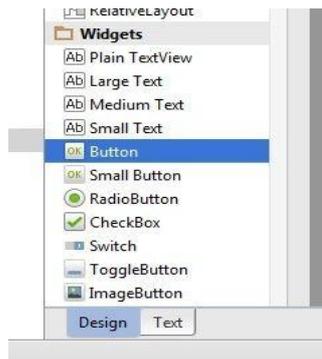- Change the text from "New Button" to "Next Page."

Fig 7.2.3 Adding Button to Main Activity

Step 5: Testing the Application

- To run the application, the green play symbol from the toolbar which is at the top of the Android Studio window is clicked.
- When the "Choose Device" dialog appears (this may take a few moments), select the "Launch emulator" option.
- Click OK.
- The application will automatically launch the app upon the virtual phone being unlocked,when the emulator opens.

RESULT:

Fig 7.2.4 Application on virtual emulator

Thus , the application has been developed on the virtual emulator and has been checked.

8. SYSTEM DESIGN

8.1. Python GSM Modem Package Introduction

PYTHON-GSMMODEM

GSM modem module for Python[10]

Python gsm modem is meant for the easy control of gsm attached to the system. The system must include a few useful command line utilities for proper interaction with the GSM modem.

Its features include:

- Methods which are flexible for sending SMS messages, checking signal level, etc.
- Starting and making calls have become convenient after the use of API in the system
- Call-back methods are used to maintain the received messages and the icoming phone calls.
- Supports SMS PDU and text mode
- Supports tracking SMS status reports
- Wraps AT command errors into Python exceptions by default
- You can easily generate your own AT commands to the modem with error checking, or read/write directly from/to the modem if preferred.
- Comprehensive test suite.

Bundled utilities:

- **GSMTerm**: For communication with an attached GSM modem, a serial terminal known as the GSMTerm is used. It contains built-in help for many AT commands ,command completion, context-aware prompt, history etc.
- **sendsms.py**: It is a command line script to send SMS message.
- **identify-modem.py**: to identify attached modem. The debug information used for development of python-gsmmodem can also be provided.

Requirements

- Python 2.6 or later
- Pyserial

8.2. WEB Beautiful Soup HTML PARSER

Beautiful Soup[11] is a Python library for getting data out of HTML, XML, and other markup languages. Example: If you've found some webpages that display data relevant to your research, such as date or address information, but that do not provide any way of downloading the data directly. It helps to pull particular contents from a webpage, remove the HTML markup, and save the required information. It is a web scraping tool that helps you to clean up and parse the documents you have pulled down from the web.

Once the pip command is installed, run the following command in the terminal to install Beautiful Soup4 or bs4:

pip install beautifulsoup4

This line has to be prefaced with "sudo" using which the permission to write to the root directories and re-entry of password is done. With sudo, the command is:
sudo pip install beautifulsoup4

Get a webpage to scrape

Get a copy of the HTML page(s) want to scrape. You can combine BeautifulSoup with urllib3 to work directly with pages on the web. The steps, mainly, focuses on using Beautiful Soup4 with local (downloaded) copies of html files.
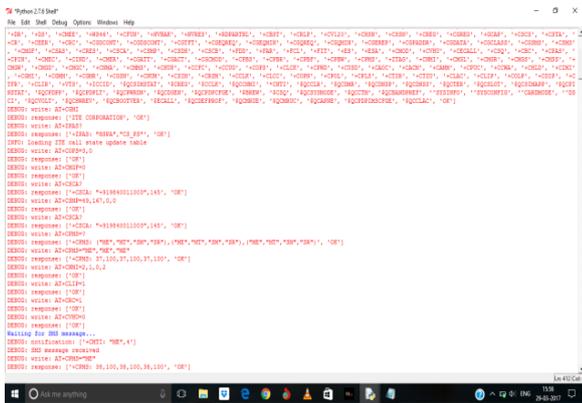

Fig 8. Waiting for SMS

The Short Message Service incorporates a number of reliability mechanisms which are, delivery acknowledgement and multiple retries[12]. Therefore, security is not compromised at all .

CONCLUSION:

A successful implementation of Offline SMS Web server has been developed by using Python as Front end and GSM server as Backend, Android studio App design act as a middle ware which generate web queries via SMS text and the SMS web queries are handled using a dedicated GSM SMS Server. The received SMS queries has been further converted as URL query to beautiful soup python package. The texts in the web URL are extracted as soup text and the splitted html paragraphs are again converted as SMS contents and finally organised as single SMS in Android APP. The proposed methodology is fluent in parsing the webpages and immediately transferring the files into SMS texts. The intangible benefits of the proposed SMS offline server are discussed and an effective protocol has been developed.

FUTURE WORK:

The SMS parsing delay is negligible but the SMS splitting and sending module is complex in nature as it requires more duty cycle. So, to overcome this, a SMS PDU concatenated mode will be initiated through which joint messages which are more length than 160 characters can be parallel sent which will further improves the timing efficiency.[13]The increasing potential the SMS offers for integration with existing messaging services and its ability to offer a successful replacement for the transmission control and Internet protocols as far as low-bandwidth-demanding applications are concerned.

REFERENCE:

[1] The Effectiveness of Short Message Service for Communication With Concerns of Privacy Protection and Conflict Avoidance by Vincent Cho Humphry Hung Department of Management and Marketing, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

[2]Short Message Service using SMS Gatewa.y, International Journal on Computer Science and Engineering by Veena K.Katankar, Dr.V.M.Thakare

[3] Mobile Marketting by Roger Storm, Martin Vendel ,Halmstad University, Spetsvinkelgatan 29, SE-302 50 Halmstad, Sweden.

[4] SMS Based Information Systems by Pankaj kumar Jaiswal, September, 2011

[5]Network flow-based strategies for identifying rail park-and-ride facility locations by Horner, M.W. and S. Groves, published in 2007.

[6] International Journal of Computer Sciences and Engineering , Research Paper Volume-02, Short Message Service: Offline Notification System through Bulk SMS for Android Application by Harshad Kale, Ganesh Rane , Sagar Shende and Swapnil Shinde

[7] Digital Libraries: Social Media and Community Networks by Walsh and Kroski, Hong Kong Institute of Education, 2012.

[8] Android Studio 2 Development Essentials by Neil Smyth published in 21st July 2014.

[9] Android   Programming in a day, book by Sam Key published in 2015.

[10] python-gsmmodem 0.9, by Francois Aucamp.

[11] Getting Started With Beautiful Soup, Book by Vineeth G.Nair, published in the year 2014.

[12] Published in INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE, Analysis of the Reliability of a Nationwide Short Message Service on 6-12 May 2007 by X. Meng, P. Zerfos, V. Samanta

[13] The global system for mobile communications short message service, IEEE Personal Communications, June 2000 by G. Peersman, S. Cvetkovic, P. Griffiths, and H. Spear.