CROWDSOURCING BASED DESCRIPTION USING GIS

Sheeba Edward 1st , Poonkuzhali k 2nd , K.Jaraline Kirubavathy 3rd , Dr. V. Thulasi Bai 4th
1st Final Yr,ECE Dept.,2nd Final Yr,ECE Dept.,3rd Asst.Prof, ECE Dept.,4th HOD, ECE Dept.,
KCG College of Technology,Chennai,India

ABSTRACT: Nowadays the world is connected with lots of information's. Crowdsourcing is the process of acquiring, integrating and analyzing big and heterogynous data. Recently social media feeds are rapidly emerging as a noval platform for providing and dissemination of information for geographic. Such information are fetched through twitter. The tweets are extracted. This extraction is done for the analysis of the people who are need of colleges, products, hospitals etc..So this analysis is useful in different fields. In this paper we have implemented student database analysis and necessities

Keyword*: Crowdsourcing , geographic information system*

I. INTRODUCTION

Now a days, information extraction is highly useful and valuable. Information Extraction is the automated retrieval of specific information related to selected topic. Information extraction depends on named entity recognition (NER), a sub-tool used to find targeted information to extract. NER recognizes entities first as one of several categories such as location (LOC), persons(PER) or organizations (ORG). Once the information category is recognized, an information extraction utility extracts the named entity's related information and constructs a machine-readable document from it, which algorithms can further process to extract meaning. By Michael Marchionda - Crowdsourcing is a growing, fast-paced and effective way for organizations to gather the best ideas from online communities and use them in ways that benefit both the organization and contributor. In this paper we have discussed about information extraction that is location from twitter which helps people from organization and many other backgrounds.

II. LITERATURE SURVEY:

[1]. The process of acquisition, intergeration, and analysis of big data generated by a diversity of sources in urban spaces, such as sensor, devices, vehicles.**(Building knowledge base of urban emergency events IEEE-2015).**

[2]. An emergency events is an unexceptional events that exeed the capacity of normal resources and organization **(crowdsourcing timeline description urban emergency 2016).**

[3]. Some important data and analystic tasks cannot be completely addressed by automated processes.**(crowdsourced data management survey 2016.)**

[4]. Based on an empirical study of a substantial amount of project initiated by relevant cultural heritage institutions.**(crowdsourcing cultural heritage domain 2016).**

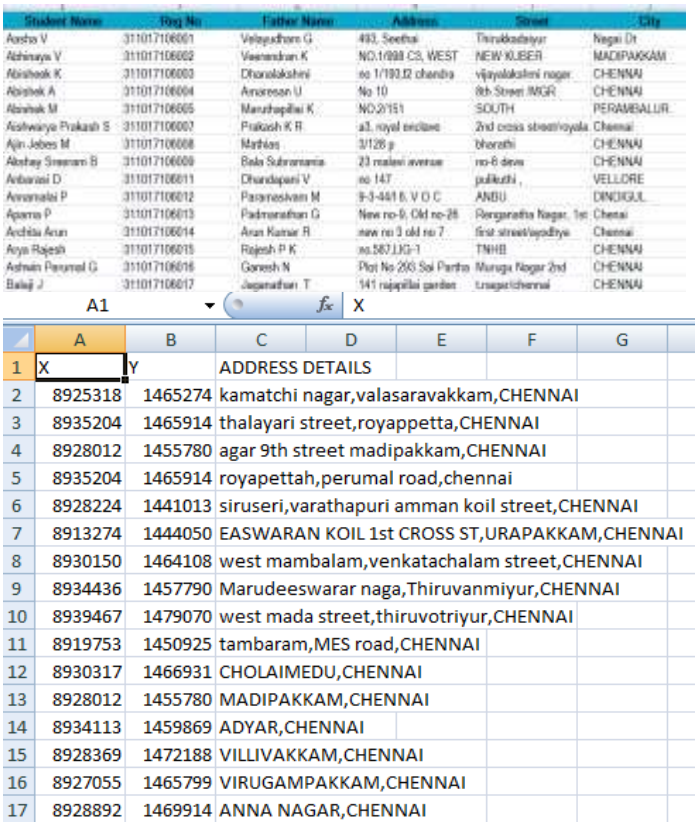
III. PROSPOSED METHOD:

Methodology

The methodology used for the exploratory data analysis comprised the following steps: (i) the acquisition/characterization of data sources ;(ii) Setting up GIS ;(iii) Integration of Excel Database in GIS;(iv) Installation and Execution of Python;(v) Information extraction and mapping in GIS;

A. Acquisition of data sources:

Source: The collection of student database is from the colleges , of branch2017, 2016 and 2015. The information given is about (i) name; (ii) register number; (iii) phone number; (iv) street; (v) area; (vi) city. But for our method we just need the city, area and street. The database is in the form of co-ordinates and then being saves. This Excel database is the starting point or a base in this project. It helps us to know where the students are coming from and where the college is known more and less. This helps to make advertisement or a camp or an awareness about the college. This makes people to know about the college and makes it well known. Excel database is stored in (.csv) file to make it easier and a known for GIS.



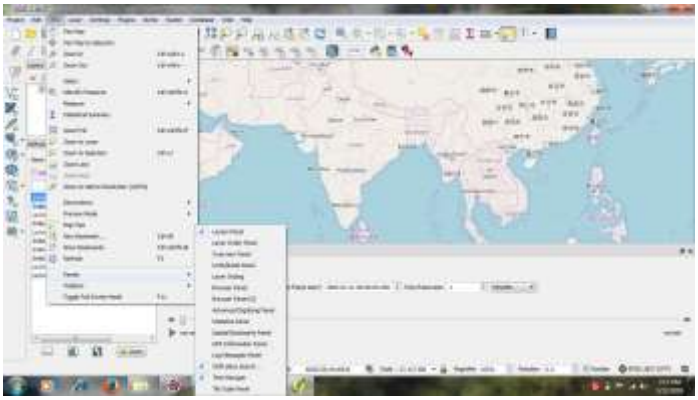
	A	B	C
1	X	Y	ADDRESS DETAILS
2	8925318	1465274	kamatchi nagar,valasaravakkam,CHENNAI
3	8935204	1465914	thalayari street,royappetta,CHENNAI
4	8928012	1455780	agar 9th street madipakkam,CHENNAI
5	8935204	1465914	royapettah,perumal road,chennai
6	8928224	1441013	siruseri,varathapuri amman koil street,CHENNAI
7	8913274	1444050	EASWARAN KOIL 1st CROSS ST,URAPAKKAM,CHENNAI
8	8930150	1464108	west mambalam,venkatachalam street,CHENNAI
9	8934436	1457790	Marudeeswarar naga,Thiruvanmiyur,CHENNAI
10	8939467	1479070	west mada street,thiruvotriyur,CHENNAI
11	8919753	1450925	tambaram,MES road,CHENNAI
12	8930317	1466931	CHOLAIMEDU,CHENNAI
13	8928012	1455780	MADIPAKKAM,CHENNAI
14	8934113	1459869	ADYAR,CHENNAI
15	8928369	1472188	VILLIVAKKAM,CHENNAI
16	8927055	1465799	VIRUGAMPAKKAM,CHENNAI
17	8928892	1469914	ANNA NAGAR,CHENNAI

B. Setting up GIS:

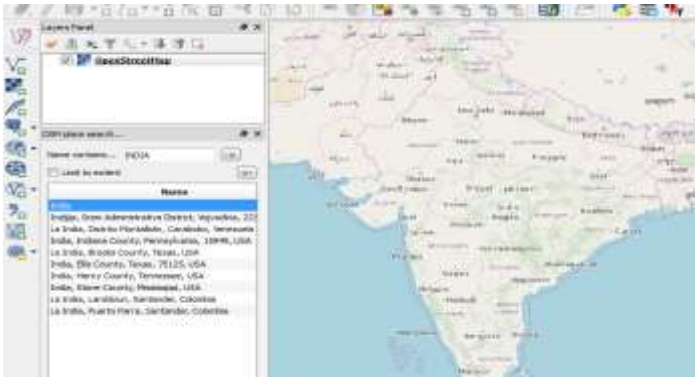
The version we use in our project for GIS is QGIS 2.18.27(Las Plasma).The next step after collecting the database is to integrate it with GIS and map the places of the collected data. We first have to do some settings and installation according to our project. The installation steps and what to be installed is detailed below:

- Open a new project.
- Install plugins (Timer manager, WKT, Open layer plugin, OSM place search).

- Click on the PANELS.



- Click on OPEN LAYER PLUGIN, a map will be appeared.
- A open street map tittle will be appeared in layer panel.
- Select India in OSM place search (select country according to your project).

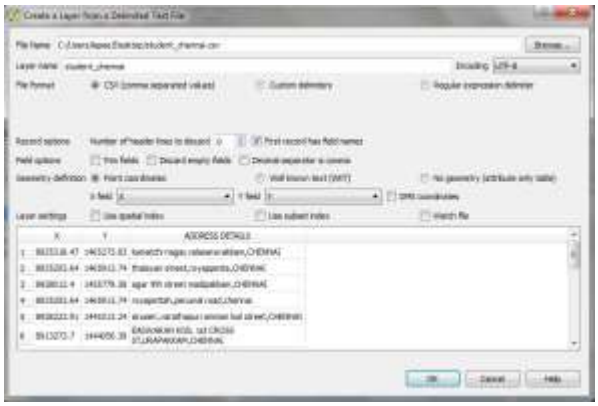


The next step is to import the database in GIS. Thus the installations are collected in this step.

C . Integrating excel database with GIS:

The excel database is being imported in this method. Here LAYERS play a major role where add layers is selected and a delimited text file is created. Browse the files and the data base is added. Then the X Co-ordinates and Y co-ordinates are plotted in the open street map. the following are the steps:

- Click on LAYERS, select ADD LAYERS.
- Next step is to add delimited text layers and browse for the files.



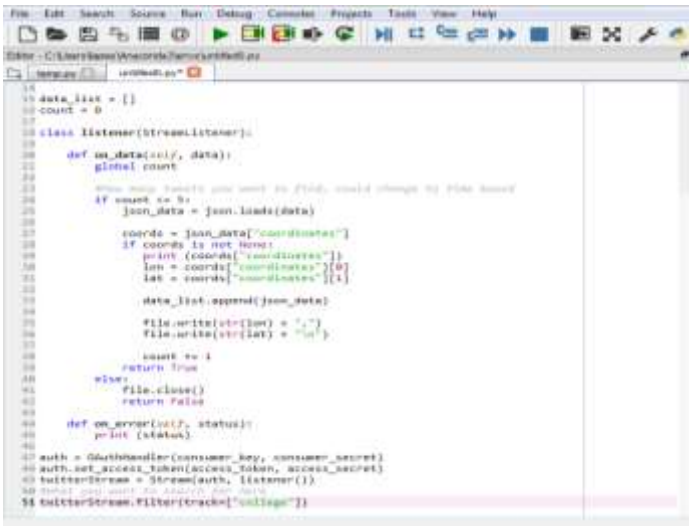
-

The file is then changed and the locations are being pinned in the map giving the exact indication of locations on the map.

D . Installation and execution of python:

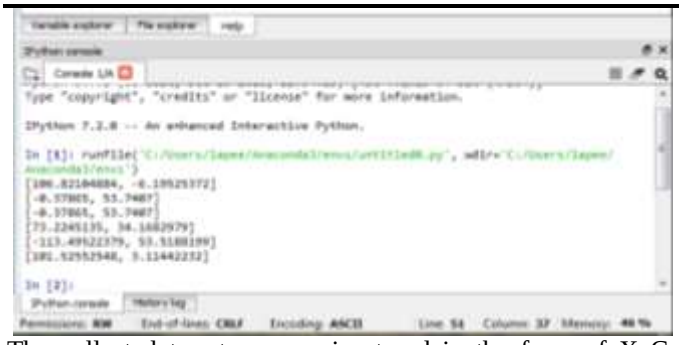
Python is a general purpose programming language. Hence, you can use the programming language for developing both desktop and web applications. Also, you can use Python for developing complex scientific and numeric applications. Python is designed with **features** to facilitate **data analysis** and visualization.

The coding used here is:

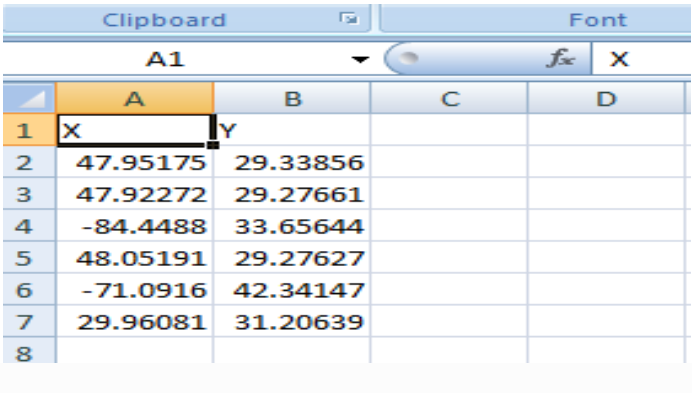


E . Fetching of tweets and mapping in GIS :

The python program is been run and the tweets are collected . The number of tweets is to be mentioned in the program according to programmers . The output gives the co-ordinates of the location from where the tweets are being tweeted.



The collected tweets are again stored in the form of X Co-ordinates and Y Co-ordinates in the excel spread sheet . This output is later given to the QGIS . This co-ordinates gives the exact location of the tweets to the users .



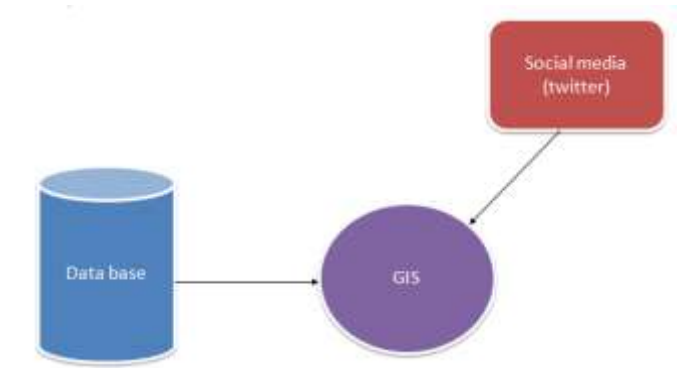
Clipboard Font

A1 X

	A	B	C	D
1	X	Y		
2	47.95175	29.33856		
3	47.92272	29.27661		
4	-84.4488	33.65644		
5	48.05191	29.27627		
6	-71.0916	42.34147		
7	29.96081	31.20639		
8				

IV. SYSTEM ARCHITECTURE

The blocks of the proposed system consists of database, geographical information system(GIS),social media(Twitter).



BLOCK DIAGRAM

SOFTWARE DETAILS:

1)Quantum Geographic Information System(QGIS GIS (Geographic Information Systems) provides an effective way to gather, store, analyze and present this real-world spatial data. GIS uses a base map – i.e., a street map or aerial photo – to allow the viewer to easily visualize the feature and location they are looking at. Using GIS helps everyone to make informed decisions. You can map any feature (e.g. building, tree, railway line, proposed shopping center, national park boundary) and record, see and analyze its size, location and how it changes over time. With GIS, you can analyze topographic, environmental, demographic and land use data to help you with your business decisions and project management. GIS is a diverse field, requiring expertise in a range of areas ranging from cartography, systems administration, relational database management, programming, and of course, spatial analysis. Even the well-versed GIS professional will come across many difficult tasks that require guidance from others in the field. Luckily, there is a strong support community readily accessible both via the Internet and in-person through focused user group meetings. media sites such as Facebook, Twitter, and Linkedin are important sites to connect with other GIS professionals.

2)Twitter: Is an American online news and social networking service on which users post and interact with messages known as "tweets". Tweets were originally restricted to 140 characters, but on November 7, 2017, this limit was doubled for all languages except Chinese, Japanese, and Korean.[14] Registered users can post, like, and retweet tweets, but unregistered users can only read them. Twitter was created in March 2006 by Jack Dorsey, Noah Glass, Biz Stone, and Evan Williams and launched in July of that year. The service rapidly gained worldwide popularity.

3)TWEEPY: Python is great language for all sorts of things. Very active developer community creates many libraries which extend the language and make it easier to use various services. One of those libraries is tweepy. Tweepy is open-sourced, hosted on GitHub and enables Python to communicate with Twitter platform and use its API. For an introduction on the library Python.

4)ANACONDA: Anaconda is the installation program used by Fedora, Red Hat Enterprise Linux and some other distributions. During installation, a target computer's hardware is identified and configured and the appropriate file systems for the system's architecture are created. Finally, anaconda allows the user to install the operating system software on the target computer. anaconda can also upgrade existing installations of earlier versions of the same distribution. After the installation is complete, you can reboot into your installed system and continue doing customization using the initial setup program. Anaconda is a fairly sophisticated installer .Installation can be scripted with kickstart to provide a fully unattended installation that can be duplicated on scores of machines. It can also be run over VNC on headless machines. Anaconda provides advanced debugging features such as remote logging, access to the python interactive debugger, and remote saving of exception dumps.

5)EXCEL DATABASE: The Excel Database functions work with an Excel Database. The Database functions perform basic operations, such as Sum, Average, Count, etc., and additionally use criteria arguments, that allow you to perform the calculation only for a specified subset of the records in your Database. A database management system is a software tool that makes it possible to organize data in a database.

FUNCTION: The ten functions in the DBMS are: data dictionary management, data storage management, data transformation and presentation, security management, multiuser access control, backup and recovery management, data integrity management, database access languages and application programming interfaces, database communication .The most simple answer **is** we need databases b/c they organize data in a manner which allows us to query data, sort data, and manipulate data in various ways. We need some way to collect, store, retrieve, sort, graph, and manipulate the data of the world.

6)PYTHON: is an interpreted, High-level language. It supports multi algorithm paradigms. Python is one of the most in-demand programming languages in the industry. It is favored by companies such as Instagram, YouTube, and Spotify. Python's diversity, adaptability and its elegant, easy-to-master basics set it apart from other programming languages, making it popular for use in web development, machine learning, data science, scientific computation, and cloud infrastructure.

Getting Data from Twitter Streaming API:

1. Create a twitter account if you do not already have one.
2. Click "Create New App"
3. Fill out the form, agree to the terms, and click "Create your Twitter application"
4. In the next page, click on "API keys" tab, and copy your "API key" and "API secret".

V.CONCLUSION

The collected database gives the location of students and gives knowledges where the more students comes from and less where less number of students is covered. The people can make advertisement etc to know more about their necessities in areas where less number are covered. Thus GIS(Geographical Information System) helps us in giving the accurate location of the database which is been imported. Crowdsourcing plays a major role which is used as a collective intellectual gathering of information that comes from the public and is then used to complete a related task. Twitter linked with python helps to get the tweets. The output co-ordinates makes it easier to plot in the QGIS software.

VI. FUTURE SCOPE

In future, our works is to get the information about the person who tweets which meant not only location but also the name of the person and at what time is the tweet is being tweeted. We can also see the persons last tweeted time. This helps the work more simpler.

VII. REFERENCES

[1] **Building knowledge base of urban emergency events based on crowdsourcing of social media** Zheng Xu1,2,*, Hui Zhang2, Chuanping Hu1, Lin Mei1, Junyu Xuan3, Kim-Kwang Raymond Choo5, Vijayan Sugumaran6 and Yiwei Zhu4(2016).

[2] **Crowdsourced Data Management: A Survey** Guoliang Li Jiannan Wang Yudian Zheng Michael J. Franklin (2015).

[3]**Crowdsourcing for Reference Correspondence Generation in Endoscopic Images.** Lena Maier-Hein1,Sven Mersmann1, Daniel Kondermann2,Speidel(2014).

[4]**Crime mapping using GIS** Zheng Xu,Yunhuai Liu,Hui Zhang,Xiangfeng LuoLin MeiChu anping Hu april(2017).

[5] **Robust crowdsourcing based Indoor location system additional and article information** Zeng Xu Yunhuai(2015)

[6]**A taxonomy of crowdsourcing based on the task complexity**. Robbie.T.nakastu, Elissa b.Grossman, charalambos L.lacorou(2014)

[7] **Crowdsourced Based on GPS** keten razirabadkar, Sidhant gadre, Rajeeev sebastian, Nikhi pwiredi (2015)

[8]**Geographic information system application and research opportunity for the information system** B.E mennecke ; M.D crossland (2013)

[9]**Remote sensing and geographic Information system(GIS) for developing coutries** M.mohamed, r.plante(2017)

[10] **The application of geographical information system to earthquake disaster** hangxia luo; Zing Zhang ; Gaoming zhu(2011)