ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

Visualization of Earthquake Data using Data Mining

Sneha Patil[1], Shikha Shahu[2], Mihir Roy[3], Shruti Patil[4]

B.E Student, Dept. of Information Technology, YCCE, Wanadongri, Nagpur, India[1,2,3,4]

ABSTRACT: Earthquake is one of the natural calamities that occur due to release of energy from the earth crust. This calamity results into the loss of human life, property .The data of previous earthquakes may serve as a tool for predicting various parameters such as, highly prone regions, intensity of earthquake or for further research work. This data is usually available in the form of 'raw data'. Visualization of data using conventional excel sheet consume lot of time. This paper discuss about the project that serves as a time efficient and user friendly solution for study, analysis and representation of available earthquake data using data mining technique..

KEYWORDS: data mining, visualization, earthquake data.

I. INTRODUCTION

Earthquake - A *natural disaster* is the effect of a natural hazard (e.g., Flood, tornado earthquake, heat wave, or landslide). Earthquakes, landslides, tsunamis and volcanoes are complex physical phenomenon that leads to , environmental or human losses. Prediction of such geological disasters is the need of the day. Also, prediction of these disasters is a very complex process that depends on many physical and environmental parameters. Many approaches exits based on analysis for analyzing earthquake data. Data mining techniques can also be used for prediction of these natural hazards. The processing of post earthquake raw data is generally done using Excel sheet which is very much time consuming. This project overcomes this drawback; in which data will be uploaded just by clicking on a button .It involves application of data mining technique- visualization, which may serve as time effective solution for visualizing earthquake data. There are two waves that cause an earthquake that are P-wave and S-wave The P-waves (Primary or Pressure wave) is a pulse of energy that travels through the earth and through the liquid. It forces the ground to move back and forth as it is compressed and expanded. The S-wave (Secondary or Sheer wave) follows more slowly with swaying, rolling motion that causes the sudden moment on the ground back and forth perpendicular to the direction of the wave.

II. TECHNOLOGY USED

Data Mining:

Data Mining is a computer-aided process that digs and analyses large amount of data and then extract the knowledge or information out of it. Data mining can be used for variety of applications such as weather forecasting, electric load prediction, product designing etc.

The various data mining techniques are:

1. Preprocessing: Preprocessing is done to improve the quality of data, improve efficiency and ease of mining.
2. Association: In association, a pattern is developed based on a relationship of a particular item on other items in the same transaction.
3. Classification: It classifies data sets into classes and groups.
4. Extraction: Extraction useful data from the large data set.
5. Clustering: Clustering means forming groups of objects which are similar.
6. Cleaning: Cleaning includes removing of noisy, inconsistent, incomplete data.
7. Prediction: Prediction discovers relationship between independent variables and relationship between independent and dependent variables.
8. Visualization: It involves conversion of textual or numeric data into meaningful images

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

III. LITERATURE SURVEY

The authors G. V. Otari, Dr. R. V. Kulkarni has discussed in their paper "A Review of Application of Data Mining in Earthquake Prediction", it is a prediction of geological disaster. The main data mining techniques used for earthquake prediction are logistic models, neural networks, and decision trees, all of which provide primary solutions to the problems inherent in the prediction of earthquakes, tsunamis, and other micro seismic activities. In their paper they also aim to encourage additional research on topics, and conclude with several suggestions for further research.

The authors Gunther H. Weber, Marco Schneider, Daniel W. Wilson, Hans Hagen, Bernd Hamann and Bruce L. Kutter has discussed in their paper "Visualization of Experimental Earthquake Data", a visualization tool that starts by reading the data which describes experiment set-up and displays this data along with icons for the sensors used during data acquisition. Different sensor types (measuring acceleration, displacement and strain) were indicated by different icons. One general experiment set-up was used in a sequence of simulated earthquake events. Once a user has selected a particular event, measured data can be viewed as a two-dimensional graph or plot by clicking the corresponding sensors. Multiple sensors can be animated to obtain a three-dimensional visualization of measured data.

The authors C. Willmes , J. Weskamm, U. Baaser , K.-G. Hinzen , G. Bareth has discussed in their paper "seismogis: a tool for the visualization of earthquake data" , SeismoGIS a GIS toolset developed for the Earthquake. This system uses seismometer station network that are used to measure the current and previous earthquake data which supports to visualize and analyze the earthquake data. The Toolset consists of three parameter that are(i) an extension for ESRI's ArcGIS, (ii) an online GIS using the UMN MapServer software and (iii) a small PHP web application for dynamic KML data generation. The software parses dynamically the measured earthquake events from an ASCII file, which is supplied by the SEISAN / SEISNET software.

IV. PROPOSED WORK

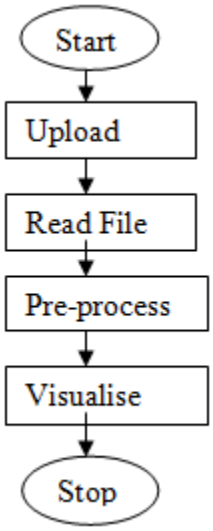

Fig.1 Flowchart

The flow of the system is depicted in fig.1 First the user uploads the raw data file. Once it gets successfully uploaded, user read the file. After that the raw data is preprocessed that is errors are removed. After preprocessing this data is visualize in the form of graphs.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

V. IMPLEMENTATION

Fig.2.Login Page

The software provides a login page as shown in fig.2.If the user has already registered then the user may directly login with his or her credentials. To login user can use their registered email-id or contact number and password

Fig.3.Registration Form

New user has to register by filling up the form as shown in fig.3.Then submit it.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

Fig.4.Upload File

The file is provided as input by selecting the 'browse' and then 'upload file' option. On reading the file using 'read file' the output obtained is as shown in fig.4.The data is available in raw format which is unreadable by user. It may also contain various unnecessary symbols, whitespaces.

Fig.5.Read File

The read file is then processed to gain tokens. Tokens are obtained by selecting 'get token' option. The output obtained is as shown in fig.5.It provides total number of tokens, word count and number count.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

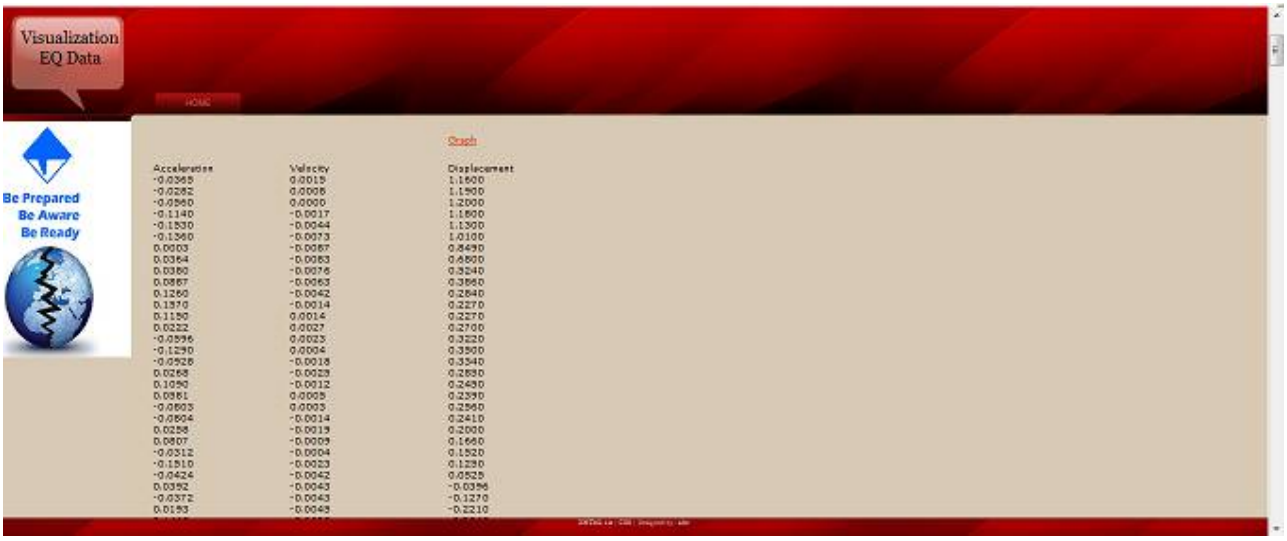

Fig.6.Get values

The tokenized information obtained in fig.5 is categorized into various groups such as velocity, acceleration and displacement as in fig.6. This is the most systematic and human readable representation of the input raw file. This information is then processed to gain graphs. User may obtain graphs between different parameters depending on his requirement.

VI. RESULT

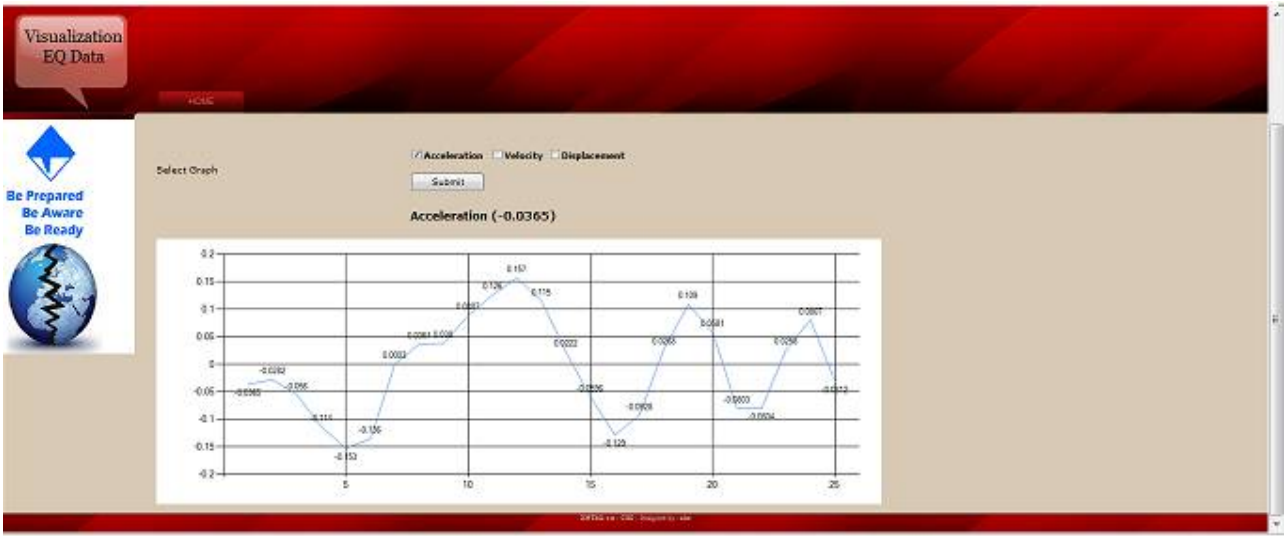

Fig.7.Time versus Acceleration

The above figure shows the graph plotted for only one parameter which is acceleration.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

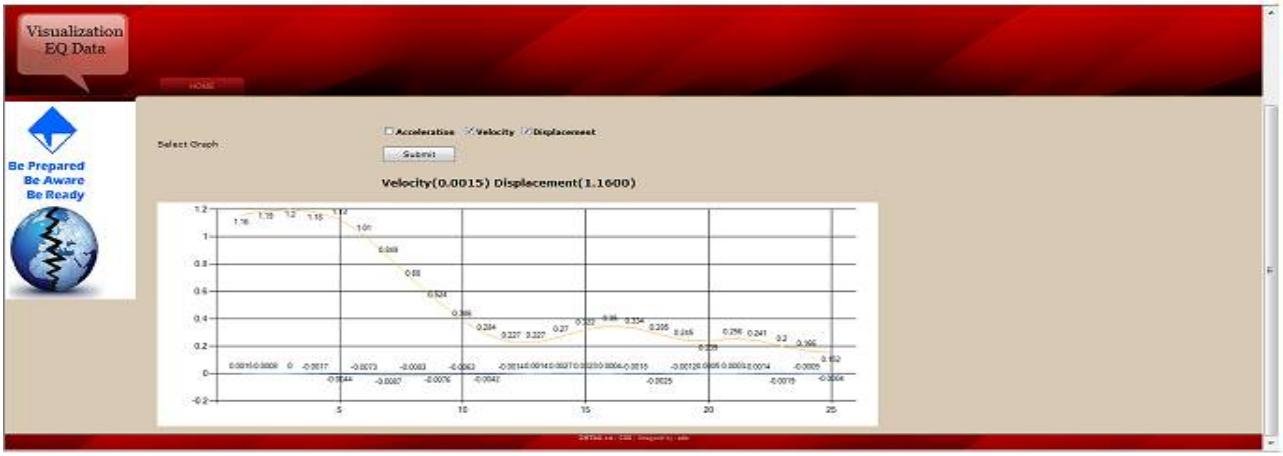

Fig.8.Time versus Velocity, Displacement

The above figure shows the graph for two parameters. User may select any two parameters according to his or her requirement .Fig 8 depicts velocity versus displacement.

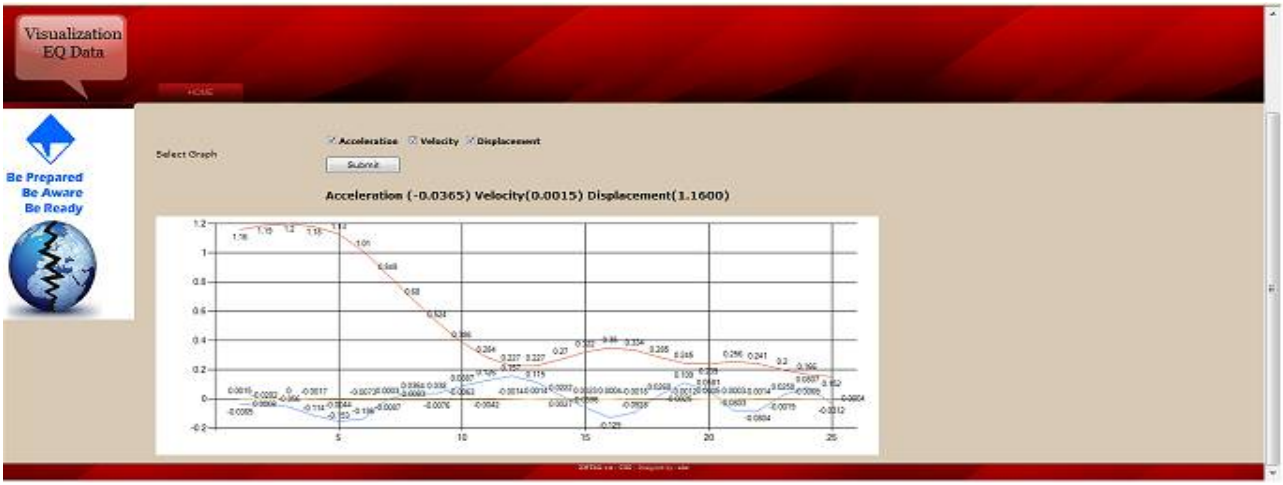

Fig.9.Time versus acceleration, velocity, displacement

The above figure shows the graph for all three parameters that is acceleration, velocity and displacement.

V. CONCLUSION

The software supports the user of the earthquake observatory to analyse the earthquake data .Another aspect of the Software is the visualization of the data, which is accomplished and archived by the earthquake observatory

REFERENCES

1. G. V. Otari, Dr. R. V. Kulkarni, "A Review of Application of Data Mining in Earthquake Prediction", International Journal of Computer Science and Information Technologies, Vol. 3 (2), 2012, 3570-3574
2. Dave A.Yuen1, Benjamin J Kadlec1, Evan F Bollig1, Witold Dzwinel2, Zachary A.Garbow1, Cesar da Silva3, "Clustering and Visualization of Earthquake data in a Grid Environment".
3. Kyle Springer, Atsushi Inoue, "Novel Visualization Scheme for Reasoning With data mining", the 28th North American Fuzzy Information ,Processing Society Annual Conference(NAFIPS2009) Cincinnati, Ohio, USA - June - 2009