

PRELIMINARY SCIENTIFIC RESEARCH AND INFORMATIVE REPRESENTATION OF SALES DATA

Dr. P. Harini¹

**¹Professor & HOD, Department of Computer Science and Engineering,
St. Ann's College of Engineering and Technology, Chirala, Andhra Pradesh, India.**

Email: drharinicse@gmail.com

ABSTRACT

Data are produced quite quickly because in the increase of daily knowledge. A huge volume of data from different entities is collected and is challenging to process and manipulate. Examples of large data are the data generated by an increasing array of environmental sensors, including traffic cameras and satellites, social networking website Internet operations, healthcare, the government database, sales data, etc. It is a task to process, analyze and communicate these results. Every day, online shopping pages are overwhelmed with huge sales data. It is a challenging challenge to analyze and visualize this data for the recovery of knowledge. This requires a framework that analyses and visualizes data effectively. The main objective of this paper is a framework that visualizes market data that allow users to apply market analytics, generate income and make decisions, manage business activities and monitor the progress of tasks.

Key words: Sales data, Analyses, Visualization, Report generation.

1. INTRODUCTION

Data visualization is a method aimed at efficient and direct communication of data to the consumer through graphical representation. The main aspect of the exploration method is effective and efficient data visualization. It is the intermediate between the subjective experience and the abstract context of the evidence; it is therefore an important element in the analytical journey between data and information. It is a versatile emerging technology that will allow researchers and businesses to make income decision [1]. it has a tremendous potential. Diverse challenges such as limitations on data store formats, a lack of prior expert knowledge of real-world databases, the difficulties of visualizing information using complex databases, which are more unexpected than usual, make extracting relevant information and useful knowledge from large mixed mode databases difficult even for domain masters. Data masters. Analysis of data is required in order to help visualize outcomes. One of the important phases in the business intelligence method, which extracts data from different sources and correlates them. Most organizations have many knowledge repositories throughout today's globalized sector. Human resources, sales, client management and marketing are also subject to software structures. Most of these divisions also has many datasets and

programmes. More and more data in various cloud offers, along with certain databases, is held with the introduction of SAAS lately. Data collection is a sequence of steps in the process of information exploration involving the use of specific pattern algorithms, as demanded by the real world. Huge data volumes do not matter for their quantity, but for the consistency of the knowledge they extract. For a comparatively complicated real problem with a huge data area, it would be evident, often even without support, that both information generation and data mining techniques would be inefficient. Even the domain owner will find it impossible to achieve useful outcomes in a wider dynamic database with more unexpected differences than regular ones. Data processing is essential to help visualize the findings.

One of the important phases in the business intelligence method, which extracts data from different sources and correlates them. Most organizations have many knowledge repositories throughout today's globalized sector. Human resources, sales, client management and marketing are also subject to software structures. Each department also has many databases and applications and more and more data in various cloud offers, along with some databases, are held in premise with recent use of SAAS.

2. LITERATURE REVIEW

The term visualization is an evolving study area, where many researchers have contributed from the last few decades. Various authors have proposed different techniques and technologies to support data visualization. This section elaborates about how the flow of research has been carried out by the authors and researchers from reputed journals and conferences.

Vipul Gupta, ArshanPorsohi and Poornaprajna Udupi[4] the author has proposed a Sensor: Network based approach for storing, sharing, visualizing and analyzing data from multiple devices and to interact with each other and with the end user through an open REST- based API. The author has visualized the geographical location of the data stream which when clicked pops up a tabbed window containing different associated information.

CiroDonalek, S.G. Djorgovski [5] the author has proposed a virtual reality platform for scientific data visualization, a tool for multi-dimensional data visualization using which scientist can interact with the data and their colleagues in the same space. The author has mapped data parameters in different data points, shapes, size, colors, XYZ axis and many more. The author has used iViz a visualization tool which can be run as a standalone application or in a web browser. The author has discussed about a framework of financial time series delivery and visualization which can be used in viewing the historical price movement of a stock [6]. Specialized binary tree (SB- tree) is used for representing the financial time series. Time series data server, SB-tree server and web service

contains is the three major components which are distributed on different machines. The system can reduce data volume and can capture the critical points.

Eric Martin and Vincenzo Di Bernardo [7] the author has proposed a dashboard for displaying data used for communicating and finding trends in laboratory operation. System is based on .NET scripts, SQL repository. The author depicts that data is collected from the multiple sources like admin, internet and user portal and are stored in database using XML layer, Adobe flash, Action Script, etc. Data is being visualized which is used for laboratory and staff management.

Victor Pascual-cid[8] the author has used a concept of visual web mining for analyzing the web data. A tool named WET is being used for visualization which provides a set of visual metaphor that represents the structure of the websites. The websites exploration tool is used for exploring the websites and for giving the feedback to the website owner for the betterment of the website.

Liang Zhu, Bing-Fang [9] the author has used a concept for analyzing data for examining the trend and evaluating the eco-environment impact of three gorges project. VC.NET and ArcIMs is the development platform for information system. ArcSDE and oracle 10g are used for management and use of spatial data. The author introduces method and processing and storing the data generated from cross-region, cross-department. Visualization helps in enhancing the data analysis and data mining.

Thorben sander [10] the author has discussed the problem in compliance management which becomes an obstacle for decision making for effective and efficient monitoring. The person should be provided with compliance software which will help in getting high level information about overall compliance status and low level problem regarding possible problems. The author has designed a dashboard for watching the compliance which avoids the obstacle and decision can be made effectively.

Javier Perez [11] the author has introduced a tool named SECONDA which is used for analyzing both individual and grouped evolution of projects and develops belonging to a software ecosystem, Visualization is implemented in java using JFREECHART libraries. The author has used GNOME ecosystem for studying, under SECONDA. It offers a dashboard for fast visual analysis of local and global matrixes that can be extracted from information stored in the repositories.

CheolJeong and Joseph Finkelstein [12] the author has proposed a system for monitoring the user exercising progress and presenting exercise parameters in relation to prescribed targets. This system can be used for monitoring the intensity of the levels recommended by the patients care provider. It uses a miniature wireless 3-axis acceleration tied on the wrist of the patient that transmits acceleration data. The dashboard allows graphical visualization of exercise progress in real time.

FleniStrouliaThe author introduces a system where the huge amount of data generated from the collaborative software development tool during the lifecycle of a project can be used to analyze the

performance of the individual member, or a team or manager. [13] They can analyze from different perspectives across different dimensions and visualized in different ways.

Pavan Kumar, Rashid Ahmad [14] the author has proposed a dashboard which is an integration, validation and visualizing tool for natural language processing. The system helps the system integration team to integrate and validate the system; developers to profile each module and researchers to evaluate and compare the module with the earlier versions. It also supports execution of modules on heterogeneous platform with an easy to use graphical interface developed using eclipse RCP.

3. METHODOLOGY

The primary objective in data visualization is the analysis of the data and its presentation to the end user. The primary objective of display is to simply and efficiently connect details by means of graphics. We propose a framework for the analysis and visualization of sales results. For different viewpoints, the details are diagramed into various parameters. In order to identify trends for potential forecasts, data mining would be implemented. For review and simulation, the data collection of one US store shall be taken. The dataset contains different attributes including order ID, order dates, priority of ordering, sales, client name, country, product name, product category etc. The device transformation scheme as seen in Figure 1 above where transformations are introduced between end users and the system and database. The operation from the registered user is seen before the user logs out. Dataset is stored using such parameters by following functions as shown in Figure 2 before loading the data into the database.

A. Data Parser

Data collection contains some entries that may or may not be applicable to the customer. In Java, the parsing of the attributes contained in a data set is performed using the java.util.Iterator class.

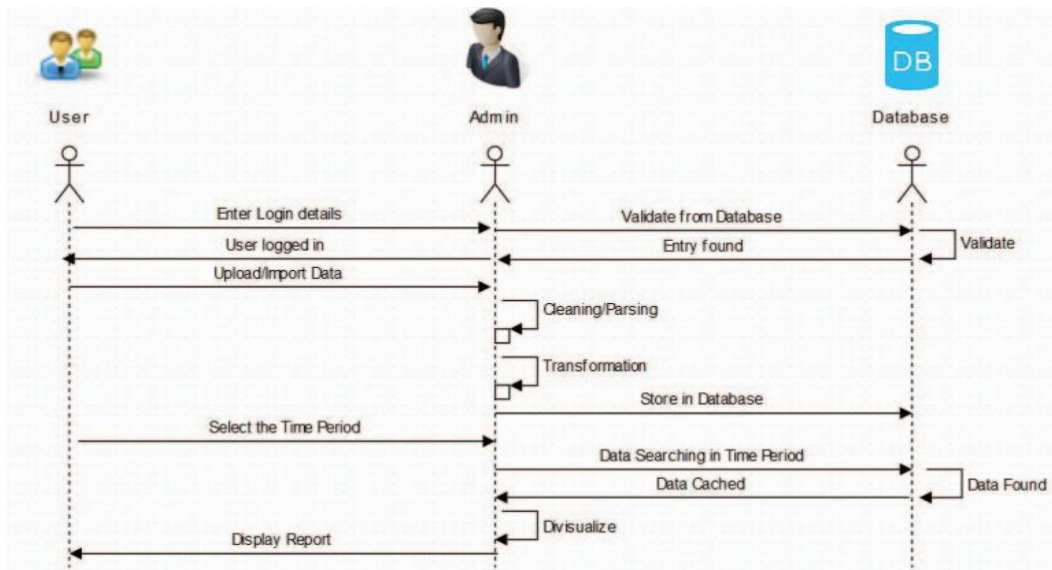


Figure1: Transition diagram for the system

B. Data Cleaner

The data set can provide details that is not valuable to the customer. This data is removed and washed from the dataset, in order to process further just the related data, which reduces the complexity of time and time.

C. Data Transfer

SFHS The working book is used to store the user's File Input Stream to convert the name of the attribute in the data collection. Attribute names can not be correctly defined. For instance, Order Id may be written to cause uncertainty. Order Id.

D. Database

Details can be downloaded into the database during analysis. The database contains valid user-specific data in the correct format.

E. Cache

The cache extracts the data, instead of a database which reduces the time taken, when the request for accessing the same data is placed.

F. Visualization

The details are visualized according to the length of the period the end user has. It is possible to imagine top clients, regional revenues, top brands and no of the customers visited.

This view allows the end customer to decide on the introduction of new products; decisions would be taken to generate income.

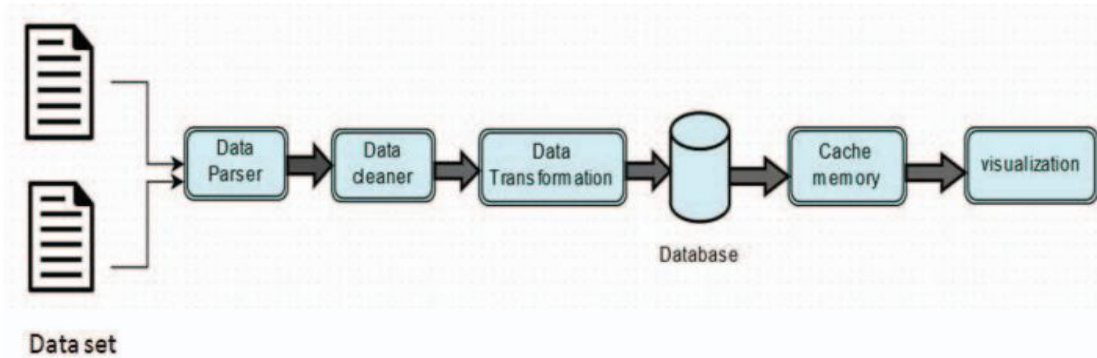


Figure2: Flow diagram for data processing carried out by the system

4. RESULTS & EVALUATION

The device data set includes many properties, which may be of little importance to the end consumer as seen in Figure3. Therefore, only the appropriate attributes from the data collection must be cleaned and extracted prior to saving it in the database. In such tasks, data is stored such as parsing, washing, and processing. Following processing by the functions that are stored in the database, Figure4 shows the results. Data must be viewed after data collection, so that the end consumer may take decisions.

Row ID	Order ID	Order Date	Order Priority	Order Quantity	Sales	Discount	Ship Mode	Profit	Unit Price	Shipping Cost	Customer Name	City
1	3	10/13/2010	Low	6	261.54	0.04	Regular Air	-213.25	38.94	35	Muhammed Meclitini	Highland Park
2	6	2/20/2012	Not Specified	2	6.93	0.01	Regular Air	-4.64	2.08	2	Ruben Darrt	Edmonds
3	32	7/15/2011	High	26	2808.08	0.07	Regular Air	1054.82	107.53	5	Liz Palletier	Elk Plain
4	32	7/15/2011	High	24	1761.4	0.09	Delivery Truck	-1748.56	70.89	89	Liz Palletier	Elk Plain
5	32	7/15/2011	High	23	160.2335	0.04	Regular Air	-85.13	7.99	5	Liz Palletier	Elk Plain
6	32	7/15/2011	High	15	140.56	0.04	Regular Air	-128.38	8.46	8	Liz Palletier	High Point
7	35	10/22/2011	Not Specified	30	288.56	0.03	Regular Air	60.72	9.11	2	Julie Creighton	Ames
8	35	10/22/2011	Not Specified	14	1892.848	0.01	Regular Air	48.99	155.99	8	Julie Creighton	Ames
9	36	11/2/2011	Critical	46	2484.7455	0.1	Regular Air	657.48	65.99	4	Sample Company A	Albany
10	65	3/17/2011	Critical	32	3812.73	0.02	Regular Air	1470.30	115.79	1	Tamara Dahlan	Plugerville
11	66	1/14/2009	Low	41	108.15	0.09	Regular Air	7.57	7.88	0	Arthur Gainer	Santa Fe
12	69	6/3/2009	Not Specified	42	1186.86	0.09	Regular Air	511.69	30.93	3	Jonathan Doherty	Garnet
13	69	6/3/2009	Not Specified	28	51.53	0.03	Express Air	0.35	1.68	0	Jonathan Doherty	Garnet
14	70	12/17/2010	Low	48	90.05	0.03	Regular Air	-107.00	1.86	2	Helen Wasserman	East Meadow
15	70	12/17/2010	Low	46	7804.53	0.05	Regular Air	2057.17	205.99	5	Helen Wasserman	East Meadow
16	96	4/16/2009	High	37	4158.1235	0.01	Regular Air	1228.89	125.99	8	Keith Dawkins	Little Rock
17	97	1/28/2010	Medium	26	75.57	0.03	Regular Air	28.24	2.89	0	Craig Yedwab	Prescott Valley
18	129	11/18/2012	Low	4	32.72	0.09	Regular Air	-22.59	6.48	8	Pauline Chand	Moore
19	130	5/7/2012	High	3	461.89	0.05	Express Air	-309.82	150.98	13	Roy Collins	Cleveland
20	130	5/7/2012	High	29	575.11	0.02	Regular Air	71.75	18.97	9	Roy Collins	Cleveland
21	130	5/7/2012	High	23	236.46	0.05	Regular Air	-134.31	9.71	9	Roy Collins	Cleveland
22	132	6/10/2010	Medium	27	192.814	0.03	Regular Air	-86.20	7.99	5	Emily Phan	Avenel
23	132	6/10/2010	Medium	30	4011.65	0.05	Delivery Truck	-603.80	130.98	5	Emily Phan	Lawston
24	134	4/30/2012	Not Specified	11	1132.6	0.01	Regular Air	-310.21	95.99	3	Michael Dominguez	Brookfield
25	135	10/20/2011	Not Specified	25	125.85	0.09	Regular Air	-89.25	4.98	4	Anne Pryor	Troy
26	166	9/11/2011	High	10	567.936	0.02	Express Air	-126.09	65.99	8	Valerie Takehito	Tamarac
27	193	8/7/2010	Critical	14	174.89	0.06	Regular Air	-37.04	12.44	6	Justin Hirsh	Bethpage
28	194	4/4/2012	Medium	49	329.03	0.1	Regular Air	-197.25	7.28	7	Maria Zeltner	West Linn
29	194	4/4/2012	Medium	6	20.19	0.04	Regular Air	-13.44	3.14	1	Maria Zeltner	West Linn

Figure3: Dataset containing multiple attributes which are not relevant to the user

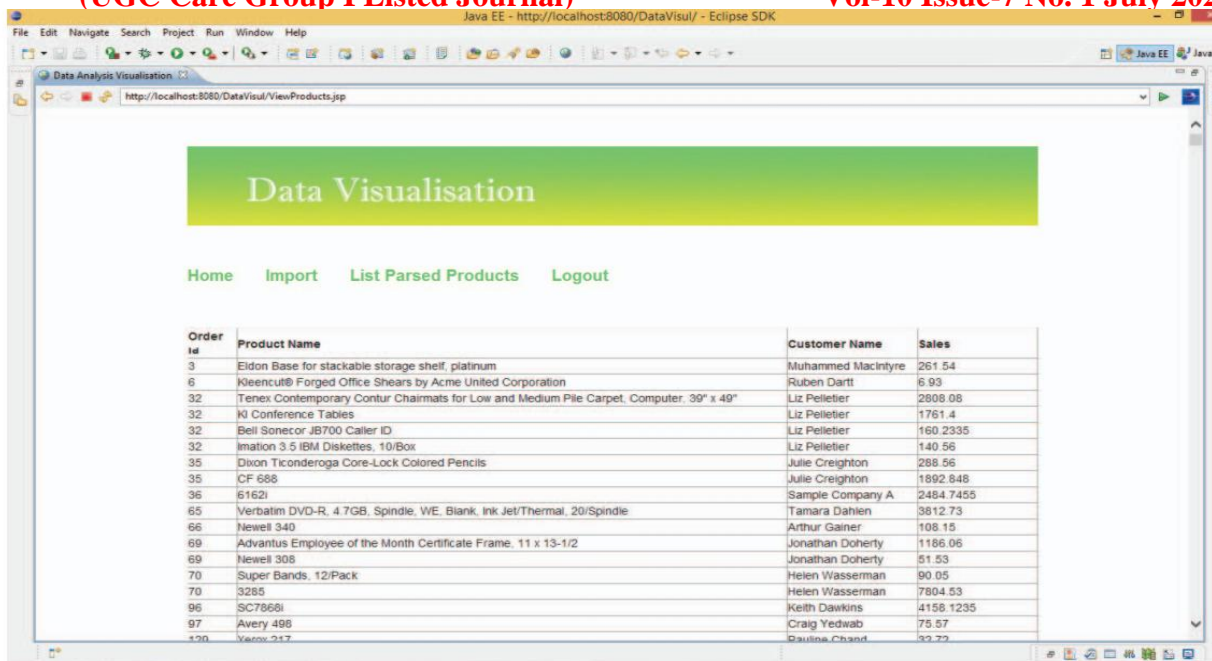


Figure4: Data stored in database after data processing.

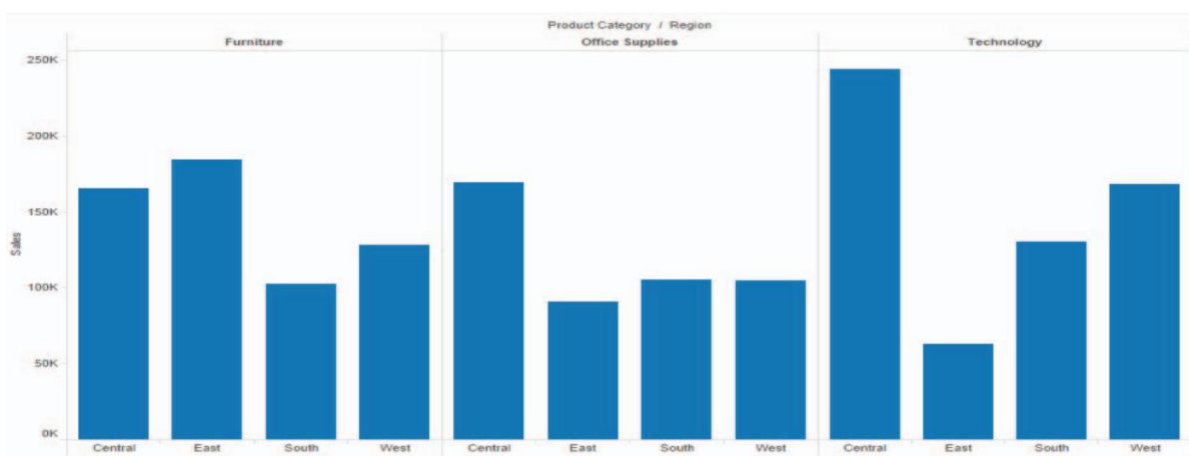


Figure5: Sales of product in particular region.

That sales of the product by segment in a given area are presented in Figure5. With this map, the maximum revenue in the area can be easily determined.

5. CONCLUSION

It is also analyzed various approaches and strategies with some weaknesses in this article. The review of number of papers from which a wide understanding about the framework needed to analyze and visualize sales data in today's environment, which allows customers and organizational owners to make proper choices and produce revenue. After compilation, I have proposed a method for importing and saving data in the data base. The data are shown with various measurements and parameters. The end consumer will decide on potential orders, quantify regional sales and enhance supply demand dependencies.

6. REFERENCES

- [1] Jiawei Han, MichelineKamber and Jian Pei, “Data Mining Concepts and Techniques”, Third edition, MK Publications, 2009.
- [2] M. Tennekes and E. de Jonge, “Top-down Data Analysis with Treemaps,” in Proceedings of the International Conference on Information Visualization Theory and Applications (IVAPP'11), pp. 236–241, March 2011.
- [3] P. Hoek, “Parallel Arc Diagrams: Visualizing Temporal Interactions,” Journal of Social Structure, vol. 12, 2011.
- [4] Vipul Gupta, ArshanPorsohi and Poornaprajna Udipi, “Sensor Network: An Open Data Exchange for the Web of Things,” in Proceedings of 8th IEEE Conference on Pervasive Computing and Communication Workshop (PERCOM), 2010.
- [5] CiroDonalek, S.G. Djorgovski, Alex cioc, Anwell Wang, Jerry Zhang, Elizabeth Lawler, Stacy Yeh, Ashish Mahbal, Matthew Graham, Andrew, Drake Scott Davidoff and Jeffrey S. Norris, “Immersive and Collaboration Data visualization using Virtual Reality Platforms,” in Proceedings of IEEE International Conference on Big Data, 2014.
- [6] Tak-chung Fu, Fu-lai Chung and Chun-fai Lam, “Adaptive Data Delivery Framework for Financial Time Series Visualization,” in Proceedings of IEEE International Conference on Mobile Business (ICMB), pp. 267-273, 2005.
- [7] Eric Martin and Vincenzo Di Bernardo, “Enterprise Dashboard Tools for Management of Share-use University Laboratory,” in Proceedings of University Conveinment Industry Micro (UCIM), 2008.
- [8] Victor Pascual-cid, “An Information Visualization System for the Understanding of Web Data,” in Proceedings of IEEE Symposium on Information Visualization (INFOVIS), 2008.
- [9] Liang Zhu, Bing-Fang, Yue-min Zhou, Xin-hui ma and Leidong yang, “Researches on Eco-Environment Data Visualization for Three Gorgeous Project,” in Proceedings of IEEE International Symposium on Information Engineering Electronic Commerce (IEEC), 2009.
- [10] Thorben sander, Malthiaskehlenbeeck and Michael H. Breithner, “Visualization of Automated Compliance Monitoring and Repairing,” in Proceedings of IEEE Database and Expert System Applications (DEXA), 2012.
- [11] Javier Perez, RomualdDeshayes, MethieuGoeminne and Tom Mens, “SECONDA: Software Ecosystem Analysis Dashboard,” in International Conference on Software Maintenance and Reengineering (CSMR), 2010.

- [12] CheolJeong and Joseph Finkelstein, “Computer-Assisted Upper Extremity Training Using Interactive Biking Exercise (iBike) Platform,” in Proceedings of IEEE Conference on Engineering in Medicine and Biological Society (EMBC), 2012.
- [13] FleniStroulia, Isaac Matichuk, Fabio Rocha, Ken Bayer, “Interactive Exploration of collaboration software development data,” in IEEE International Conference on Software Maintenance (ICSM), 2013. [14] Pavan Kumar, Rashid Ahmad, B.D. Chaudary and Mukul K. Sinha, “Enriched dashboard: - An Integration and Visualization Tool for Distributed NLP System on Heterogeneous Platform,” in International Conference on Computer Science and its Applications (ICCSA), 2013.
- [15] Ma Xin-hui, Wu Bing-fang, Zhu Liang, etc. “Study on Methods of Hydro-environment Data Visualization for Three Gorges Project” [J], Remote sensing informatics, 2007, pp. 28-31.
- [16] M. Lungu, M. Lanza, T. G^ˆOrba, and R. Robbes, “The small project observatory: Visualizing software ecosystems,” Science of Computer Programming, vol. 75, no. 4, pp. 264 – 275, 2010.
- [17] R. Chang, M. Ghoniem, R. Kosara, W. Ribarsky, J. Yang, E. Suma, C. Ziemkiewicz, D. Kern, and A. Sudjianto, Wire-vis: Visualization of categorical, time-varying data from financial transactions. In IEEE Symposium on Visual Analytics Science and Technology, pages 155–162, 2007.
- [18] “The 37 Best Tools for Visualization,” <http://www.creativebloq.com/design-tools/data-visualization712402>. Harjeet Kaur, VarshaSahni, and Dr. ManjuBala, “A survey of reactive, proactive, and hybrid routing protocols in MANET: a review”, International Journal of Computer Science and Information Technologies, vol. 4, no. 3, pp. 498-500, 2013.