A STUDY ON BIG DATA VISUAL ANALYTICS

K. MEENA[a1] AND P. ILANGOVAN[b]

[ab]Department of Information Technology, Periyar Maniammai University, Thanjavur, Tamilnadu, India

ABSTRACT

In every day, a vast amount of data and information rises rapidly most of it from social media, web, sales, mobile data, health data or sensors. Its quite very challenging task to analysis the bulk amount of data and extract the knowledge from them. Visual Analytics plays a major role to extract efficient knowledge from the complex data. VA tools and techniques are used for sensemaking of information. This paper presents the concepts of Big Data Visual Analytics techniques.

KEYWORDS: Visual Analytics, Knowledge, VA, Sensemaking

Today's world runs on data revolution. In this data revolution huge volume of data is being generated by every individual person. As technology keeps on updating the data generating factors is also increased as a result of these factors the size of data is increasing a lot every year. Hence capturing, processing, analyzing, visualization, searching, storage, transfer, security becomes the major challenging issues and it's difficult to process all these data by traditional method [Ilangovan, 2016]. Big data refers to huge volume, velocity, variety, value veracity of data. Figure 1 refers to the properties of Big Data

Volume : Massive Volume of Data (Expansion of Internet has created an increase in global data production)

Velocity : Speed of data to be stored, retrieved and analyzed

Value : Ability to turn our data into value

Variety : Data comes from all streams structured, unstructured and semi structured

Veracity: The data being stored and mined are meaningful to the problem being analyzed ie trustworthiness of data.

Figure 1: Properties of Big Data

Big data environment will be helpful to acquire, organize, analyze, and predict the data in different forms as per the user requirement. Big Data faces major challenges because of their rapid velocity, large volume, different variety, value and veracity. The emerging field of Visual Analytics plays a major role in big data to integrate data analytics with interactive visual interface.

VISUAL ANALYTICS

The evolving field of Visual Analytics is defined as the science of analytical reasoning supported by interactive visual interfaces and was coined by Thomas and Cook in the year 2005. Visual Analytics uses visualizations, user interaction and data analysis techniques to find insight from complex, conflicting and dynamic information. The objective of visual analytics is to develop visual and interactive tools and techniques for reasoning and decision making from large data sets [Patil and Guruprasad, 2015]. Even though we are spending huge investments in time and money which leads to loss, because we still lack the possibilities to properly interact with the large volume of data sets. Visual analytics aims at bridging this gap by employing more intelligent means in the analysis process. Visual analytics brings a new promising approach to tackle the information overload problem.

VISUAL ANALYTICS MODEL

The basic idea of visual analytics is to visually represent the information, allowing the human to directly interact with the information, to gain insight, to draw conclusions, and to ultimately make better decisions. Fig 2. shows the Visual Analytics Process [Keim et.al., 2010]. The data from various sources with different types need to be filtered, noise removed and then transformed to input data streams. The data is to be processed using mathematical, statistical and data mining algorithms and models. Visualization helps the user to perceive new aspects of the data. The user can explore the data and achieve new knowledge by incorporating interactivity in visualization.
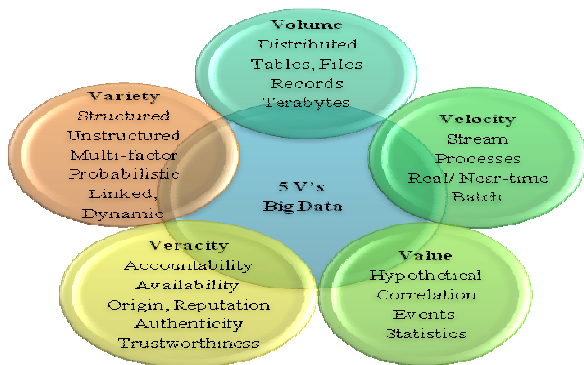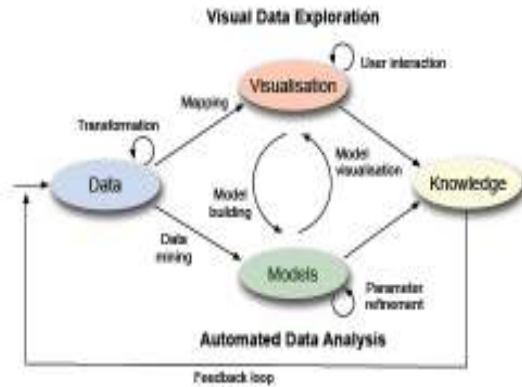
[1]Corresponding author

Figure 2: Visual Analytics Process

Fig 3. shows the structure of Visual Analytics in Big Data [Keim et.al., 2008], the rapid and wide variety of data is collected and then cleaned with proper process. Now the data is to be visualized using information visualization techniques and then transformed to data analytics with user interaction. The visualization is needed in addition to analytics to identify the structure (based on user knowledge), to bridge the semantic gap (bring in user knowledge), to help with incomplete or incorrect and changing data, to understand the problem, generate hypotheses and define the problem, to steer the analysis process in dealing with massive data (local optimization). Visual Analytics tools provides three main benefits compared with traditional method: first, the user can explore big data in a self-service 'point-and-click' fashion; second, complex ideas can be communicated with clarity and efficiency in visual graphs rather than the tabular data output; third, VA tools can display large amounts of filtered data in near real time.

The major issues in Big Data Visual Analytics are system scalability, available information Space, visual representations, interpretability, multidimensional data, network or relational data, workflow and real time analysis [Keim et.al., 2006].
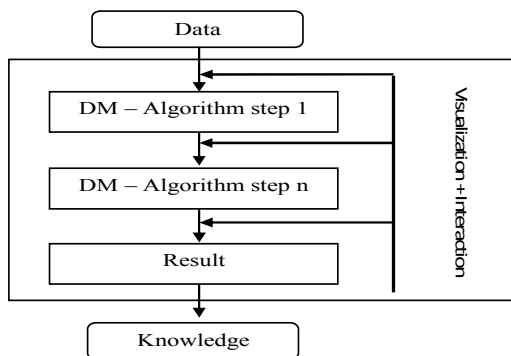

Figure 3: Structure for Visual Analytics in Big Data

The other problems are data noise, loss of information, frequent image change, perception of large image, required high performance accessories.

The important benefits of big data visual analytics is to synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data, to detect the expected and discover the unexpected, to provide timely, defensible and understandable assessments, communicate assessment effectively for action [Patil and Guruprasad, 2015]. This can be applied for wide variety of applications like education, investigation, cyber crime, credit card fraud detection, urban mobility, climate data analysis etc.

CONCLUSION

In this big data era, the importance of extracting efficient knowledge from complex data is eventually increasing, so the Visual Analytics field is also increasing parallel. This paper discusses the concepts, model and importance of Visual Analytics in big data. Furthermore it explores visual analytics, a promising method for sensemaking of big data.

REFERENCES

Aigner W., Miksch S., Müller W., Schumann H. and Tominsk C, 2008. "Visual methods for analyzing time-oriented data," Transactions on Visualization and Computer Graphics, **14**:47-60.

Card S.K., Mackinlay J.D. and Schneidermann B., 1999. "Readings in information visualization, Using Vision to Think", Academic Press Inc, pp.686, ISBN 1-55860-533-9.

Cook K., Earnshaw R. and Stasko J., 2007. "Discovering the Unexpected. IEEE Computer Graphics and Applications," **27**(5).

Keim D., Kohlhammer J., Ellis G. and Mansmann F., 2010. "Mastering the Information Age Solving Problems with Visual Analytics," Eurographics association, ISBN 978-3-905673-77-7.

Johnson C., Moorhead R., Munzner T., Pfister H.P., Rheingans P. and Yoo T.S., 2006. "NIH/NSF Visualization Research Challenges," http://vgtc.org/wpmu/ techcom/national-initiatives/nihnsf-visualization-research-challenges-report-.

Ilangovan P., 2016. "A study on Big Data and Hadoop Storage System," International Journal of Trend in Research and Development, pp.42-44, ISSN:2394-9333.

Keim D.A., Mansmann F., Oelke D. and Ziegler H., 2008. "Visual Analytics: Combining Automated Discovery with Interactive Visualizations. Discovery Science," pp.2–14, ISBN 978-3-540-88410-1. http://infovis.uni-konstanz.de/papers/2008/ComVis.pdf.

Keim D.A., Mansmann F., Schneidewind J. and Hartmut Z., 2006. "Challenges in Visual Data Analysis," Tenth International Conference on Information Visualization (IV'06). doi.ieeecomputersociety. org. ISBN 978-0-7695-2602-7.

Keim D.A., 2008. "Visual Analytics in Europe," NVAC Views. http://infovis. unikonstanz.de/papers/ 2008/NVACViews2008.pdf.

Patil S.M. and Guruprasad H.S., 2015. "A study on Visual Analytics," International Journal of Computer Science and Information Technology Research. pp. 303-308, 2015, ISSN 2348-1196.