International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 05 | May - 2022 Impact Factor: 7.185 ISSN: 2582-3930

TO SURVEY ON TRAFFIC DATA USING DATA VISUALIZATION

[1]Mirunalini.R, [2]Sathiya.A

[1]UG Student, Data Science, Sri Krishna Arts and Science College, Coimbatore, Tamil Nadu, India.

[2]UG Student, Data Science, Sri Krishna Arts and Science College, Coimbatore, Tamil Nadu, India.

EMAIL ID: [1]mirunalinir21bds026@skasc.ac.in, [2]sathiyaa21bds043@skasc.ac.in

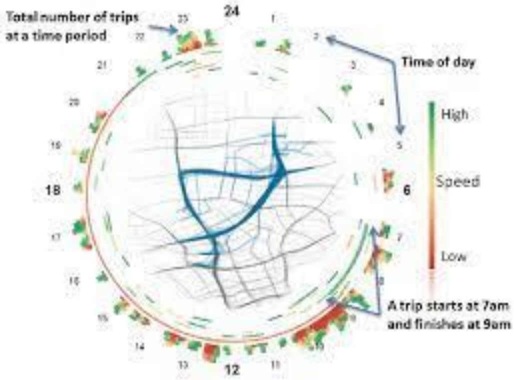
Abstract - Data visualization means it represents the distribution and structure of Datasets and also reveals the hidden patterns in the data. As there is Rapid development in States and cities, traffic congestion and complex Traffic data are being generated. Based on the traffic Data Visualization which is being generated, the users can be able to Directly interact and correlate with the data conveniently and visually. Traffic data investigates the structured survey of the state of art in the Traffic Data visualization. The data set can be generated using the Tools like data wrapper, Tableau, Google charts, etc.., Moreover we Propose a paper direction for optimizing traffic data visualization.

KEYWORDS - Data Visualization, Traffic Survey, Internet of Vehicles.

I.INTRODUCTION

Traffic is the passage of motorised vehicles, non-motorized vehicles, and pedestrians on the road. Nowadays, due to the rapid and massive development of states and cities, the transportation system is also growing rapidly. Because of the rapid development of transportation, the traffic on the roads is massive. For instance, an estimated 40% of the population spends at least one to two hours on the road every day. Transportation has a pivotal role in ensuring citizens' livability, security, safety, and health. To solve these transportation problems, many methods have been developed. To solve the problems, some of the methods are being integrated with the new technologies in their solutions, such as sensors, video images, and global positioning systems (GPS), which provide a large amount of data. To begin, in order to determine the average of transportation, a road traffic survey is conducted, which determines the number and type of vehicles passing through a specific point over a specific period. camera which captures the images of the number and type of vehicle that is passing through a certain point normally works. But this method has some disadvantages of requiring more manpower and cost, which cannot be applied where long term interpretation is required, such as traffic surveys that should be continuously monitored all the time. Among all these methods, visualisation is one such method that has attracted people Figure: 1.1 Traffic survey at a particular time period.

and increased attention. There are different traffic data visualisation methods, each of which has



different characteristics and design principles. Among many different methods, datasets are one such method that is more suitable for processing. Traffic data sets are usually multiscale or infinite-dimensional. Mainly, traffic data visualisation systems are used to visualise the information and analyse the visual technology.

The traffic data visualisation system usually consists of four data statuses and three process phases. The four data statuses include the original data, processed data, visualization, and visual symbols. The original data collection, which collects the information from two mainstream data sources: video surveillance and GPS of the vehicle, The processed data will have the information from the data that is being converted into a machine-readable form of data. The visual symbols generally show bar charts, line charts, and scatter charts. The visualisation shows the animations and 3D videos.

The three phases of data are the preprocessing stage, the visual transformation stage, and the visualisation mapping stage. The preprocessed stage is the stage in which the cleaned data is processed from the raw data. Visual transformation is the process in which the processed data is converted into visual symbols, and visualisation mapping is the process in which the user can perform interactions with the visual symbols in visual forms. In this way, massive data can be discovered easily.

II.BUSINESS DATA VISUALIZATION STYLES

A large number of visual analysis styles have been developed for large-scale and high- speed business data. Considering the current centralized response work in the field of business data visualization analysis, this section substantially introduces four representative styles, videlicet, Web Virtual Reality Geographical Information System (WebVRGIS), IoV (Internet of Vehicles) distributed armature, SMASH armature, and LDA- grounded content modelling.

2.1WebVRGIS Grounded Business Analysis and Visualization System:

X Li etal. [6] present a WebVRGIS grounded business analysis and visualization system, which uses Virtual Reality Geographical Information System (VRGIS) to collect geospatial data in real-time and grounded on this data to perform complex visual analysis, operation and expansion. The system not only has introductory visual commerce functions but also intelligent passenger inflow vaticination functions. The invention of this work includes real- time dynamic business data collection, 3D Civilians analysis, and business mecca release. Through the analysis of long- term dynamic business data, the system completely evaluates the service compass of corners, to realize the comprehensive evaluation of public transport service capacity.

2.2 IoV Distributed Architecture:

Mohamed etal. (7) present an IoV distributed armature, which aims to prize real- time business situations in each road network section. The armature follows the IoV terrain specification to collect and dissect big data business and provides effectiveness, scalability and high performance. This armature is grounded on three layers IoV subcaste, pall computing subcaste, and fog computing subcaste. The IoV subcaste is substantially used to connect vehicles, road structure and communication technology. The fog calculating subcaste is used to collect, process and store

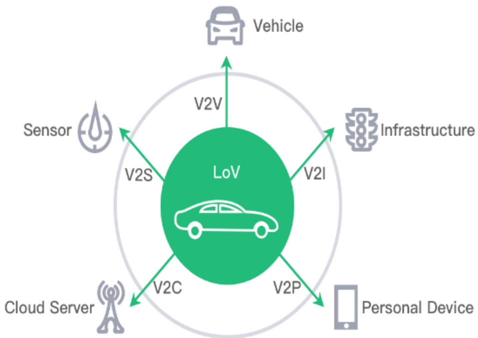


Figure2.2 Internet of Vehicles.

information generated by vehicles and structure connected by the IoV subcaste instantly. Eventually, the pall calculating subcaste carries out business data visualization, dynamic data storehouse, and fog computation results analysis. The IoV distribution armature is favored by the development of the smart metropolises including MAN networks.

2.3 SMASH Architecture:

Yikai etal. (2) proposed the SMASH armature, a general and largely scalable pall- grounded result. The SMASH armature is a software mound a set of well- defined software combinations that work together to form a complete frame able of handling multidimensional visualization tasks for developing the needed operations in a given sphere. It's the crucial function to meet the requirements of dealing with high- speed, high volume and high- perfection business data. The technologies involved in enforcing SMASH armature substantially include GeoServer, GeoMesa, Apache Spark, and Apache Hadoop. In terms of structure, the master knot and numerous slave bumps constitute the computing subcaste and data subcaste of the SMASH armature. Utmost of the factors of these two layers run in the Docker vessel. GeoServer's bumps generally run at the garçon position of the SMASH armature. The necessary plugins for the GeoMesa runtime are installed on this garçon, which also enables it to give geospatial information generated by the computing subcaste and stored in the data subcaste. It's worth mentioning that inventors are free to make the required views and publish them on GeoServer.

2.4 LDA- Grounded Content Modelling:

The system contains two main phases, videlicet, visual disquisition and data processing. During the data processing phase, a data cleaning process is first performed to cancel invalid GPS records. After the data is gutted up. LDA topic modelling has used to find semantically important areas. In addition, area characteristics are described by calculating area attribute statistics. Eventually, the precious areas are named grounded on semantic results and trait statistics. The unique point of this system is the visual design of three coupled views LDA view, business inflow view, parcels and view. It can help business directors understand the semantics of motifs implicit in line data and explore a variety of business attributes.

III.ANALYSIS OF TRAFFIC DATA

Here we introduce five representative and widely used traffic data visualisation methods. Although each method has different design principles and adopted technical support, the goal is to effectively analyse massive and complex traffic data and find valuable clues hidden in this data. Therefore, the analysis of traffic data is also crucial. The datasets collected by the electronic monitor installed along the road or the sensor installed on the vehicles are referred to as traffic data. Examples mainly include GPS data of vehicles, human displacement or GSM position data, and video/image recording of electronic intelligent monitoring equipment. In this section, we analyse the following two types of Traffic data in detail according to the working mode of the sensor:

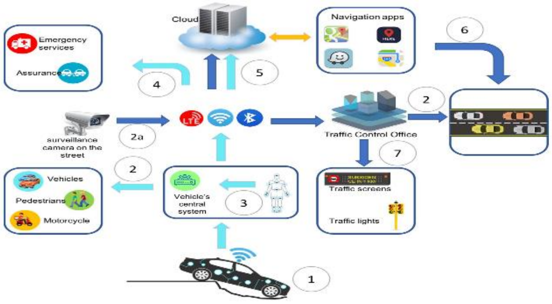
3.1 Traffic Data Collected by Location-Based Sensors:

After entering the sensor range, the location of the object is recorded. For example, SMASH architecture and IoV distributed architecture. In the SMASH architecture, the traffic volume data used is based on the Sydney Coordinated Adaptive Transport System (SCATS), which collects the traffic flow of a specific intersection in a specific period. However, the challenge with SCATS data is that they can be very storage intensive. For example, data from a single intersection may include the results of millions of vehicles passing through the intersection. In this case, the SMASH architecture provides scalable traffic data analysis function, which meets the needs of speed and accuracy in dealing with such large traffic data.

In IoV distributed architecture, the datasets used are from the road traffic database of Mohamed City, Morocco. In IoV distributed architecture, data storage occurs twice. The first relates to the storage of vehicle travel logs, and the second relates to the storage of road condition description. Firstly, the vehicle travel logs are generated in the JSON format and then published in the "vehicle" topic of the first Kafka server. Secondly, the road conditions are described by using Spark Mlib and dashboard microservices to store microservices for real-time visualization.

3.2 Traffic Data Collected by Device-Based Sensors:

The device carried by the subject actively records and sends back location information and other clues. For example, LDA-based topic modelling, TripMiner, and WebVRGIS based traffic analysis and visualization system. In LDA-based topic modelling, Hangzhou taxi GPS records and Point of Interest (POI) data are used. Taxi GPS data not only records the personal activity and travel history but also provides the whole city coverage.



Therefore, it can

Figure:3.2 Traffic collecting Sensors.

effectively reveal the potential traffic information and flows in the city. The records captured by onboard GPS equipment include vehicle number, speed, geographical location, vehicle condition (occupied or vacant), timestamp, and other fields. GPS records have the characteristics of time and space,which are preprocessed by LDA based topic modeling in the process of data processing. With GPS records, we can extract the driving track of each taxi according to the license plate number.

POI is a piece of important semantic information used to understand the relationship between human movement and activity. In LDA based topic modeling, POI is used as auxiliary semantic information to derive potential transportation preference modes in different regions. POI data is collected and obtained through the Place API provided by Baidu Maps, including the names, categories, and geographical locations of different organizations. In TripMiner, the data used are collected from vehicles travelling in a European funded project called EURO-FOT, which includes a large number of sensor readings inside and outside the vehicle, and in some cases, video recordings. The EURO-FOT database provides 170000 trips, equivalent to 100 vehicles and 206 drivers. Records provided by the database include location, vehicle speed, ambient temperature, radio value, fuel consumption, and driver ID.

In the WebVRGIS based traffic analysis and visualization system, the datasets used are multiple consecutive real-time dynamic traffic data in Shenzhen. At present, bus data is used. Through monitoring and analysis, the time and space distribution of passenger flow and hub service capacity under different modes of transportation are studied from different times and perspectives. The initial data table mainly stores data collected by the GPS recorder installed on the bus, including data collection time, license plate number, location, vehicle driving data, and no-load status.

IV.CONCLUSION

Big data brings many opportunities and challenges to the field of traffic data analysis, and visual analysis technology provides an understandable way to analyze big data, which greatly improves the accuracy of data analysis and greatly shortens the analysis time. Therefore, traffic data visualization plays a key role in solving the problems caused by large-scale, multi-mode data. This paper summarizes five representative and widely used methods of traffic data visualization, namely: WebVRGIS based traffic analysis and visualization system, TripMiner, IoV distributed architecture, SMASH architecture, and LDA-based topic modelling. Then, according to the working mode of the sensor, classify the traffic data, and analyze the reason that each traffic data visualization method adopts a specific type of traffic datasetsand the records provided by the traffic datasets. Furthermore, according to the attributes and characteristics of traffic data, this paper compares and analyzes five traffic data visualization methods in detail from seven aspects: scalability, data storage, data update, interactivity, reliability, data anomaly detection, and spatiotemporal visualization. Finally, it is concluded that SMASH architecture performs better in processing high speed and large flow traffic data.

In addition, in order to improve the robustness, performance, and processing efficiency of our traffic data visualization analysis method, we have envisaged the following points in the short term: x Further research will focus on combining multiple types of real-time datasets with SMASH architecture. x Another direction worth exploring is the visual analysis of social data. The purpose is to collect data from the network and social media for analysis, to find the potential traffic habits and routes of users, so as to further optimize the construction of intelligent transportation system.

REFERENCES

[1] A. Wibisono, W. Jatmiko, H. A. Wisesa, B. Hardjono, and P. Mursanto (ETAL), "Traffic big data prediction and visualization using Fast Incremental Model Trees-Drift Detection (FIMT-DD)," KnowledgeBased Syst., vol. 93, pp. 33–46, 2016.

[2] Y. Gong, L. Morandini, and R. O. Sinnott (ETAL), "The design and benchmarking of a Cloud-based platform for processing and visualization of traffic data," 2017 IEEE Int. Conf.

Big Data Smart Comput. BigComp 2017, pp. 13–20, 2017.

[3] U. Demiryurek, F. Banaei-Kashani, and C. Shahabi (ETAL), “TransDec:A spatiotemporal query processing framework for transportation systems,” Proc. - Int. Conf. Data Eng., pp. 1197–1200, 2010.

[4] C. Jia, H. Wang, and L. Wei (ETAL), “Research on Visualization of MultiDimensional Real-Time Traffic Data Stream Based on Cloud Computing,” Procedia Eng., vol. 137, pp. 709–718, 2016.

[5] H. H. Hochmair, “Spatiotemporal pattern analysis of taxi trips in New York city,” Transp. Res. Rec., vol. 2542, no. 2542, pp. 45–56, 2016.

[6] X. Li et al., “WebVRGIS based traffic analysis and visualization system,” Adv. Eng. Softw., vol. 93, pp. 1–8, 2016.

[7] M. Nahri, A. Boulmakoul, L. Karim, and A. Lbath, “IoV distributed architecture for real-time traffic data analytics,” Procedia Comput. Sci., vol. 130, pp. 480–487, 2018.