THE INNOVATION DEVELOPMENT OF DATA TRANSPORTATION MODEL TO ANALYZE THE COMPLEX BIG DATA ENVIRONMENT IN PUBLIC SECTOR COMPANIES

S. Nawin

Department of Master of Business Administration, Dr. N.G.P Institute of Technology, India

Abstract

The data transportation brings together developers of solutions to work with digital technologies. The analysis of big data helps in achieving the set goals. The innovation center has the ability to work with data from companies under the transport sector, carry out its own research, test hypotheses and create segments for target communication companies. We collect a wide range of non-personal data about passengers and, based on analysis, can provide important information to citizens in a targeted way. In this paper an innovation model was proposed to analyze the complex big data environment in public sector companies. There are frequent changes in routes, repairs, blocking of traffic due to certain events, activities. By analyzing the data, travelers who frequent these routes can be given personalized information about changes.

Keywords:

Data, Transportation, Digital, Big Data, Communication, Innovation

1. INTRODUCTION

There is always a gap in any business, in any science, between the professional technologists and those who understand why all this can happen, and this is certainly a problem [1]. People who know the side of technology and experiment with new solutions rarely do really useful things, and even those who know how to use these developments alone cannot create a quality product [2]. Therefore, the only way to create when working with Big Data is to find new ways of communication between experts. Big Data is currently one of the main drivers of IT development [3]. It was in the era of information technology, especially after the boom in social networking sites, which a significant amount of information began to accumulate for each Internet user, which eventually led to the direction of Big Data [4]. The term "Big Data" causes a lot of controversy, many people believe that it is only the amount of accumulated information, but do not forget about the technical side, this area includes storage technologies, computing and services [5-6]. It should be noted that this area involves processing a large amount of information that is difficult to process using traditional methods [7]. The scope of Big Data technologies is broad. So, with the help of Big Data, you can learn about customer preferences, effectiveness of marketing campaigns or risk analysis [8].

It should also be noted that Big Data is one of the fastest growing areas of information technology, and according to statistics, the total amount of data received and stored doubles every 1.2 years [9]. Between 2012 and 2014, the amount of data sent monthly by mobile networks increased by 81%. Cisco estimates that the volume of mobile traffic in 2014 was 2.5 Hexa bytes per month (a unit of measurement of information equal to 10^18 standard bytes) and in 2019 it will be equal to 24.3 Hexa bytes [10]. Therefore, Big Data is already an established technology field, despite its relatively young age [11]; it has become widespread in many areas of business and plays an important role in the development of companies [12].

2. LITERATURE REVIEW

Big data makes it possible to analyze a borrower's creditworthiness and is also useful for credit scoring and underwriting. The introduction of Big Data technologies will reduce the time taken to process loan applications. With the help of Big Data, it is possible to analyze the activities of a particular customer and offer tailored banking services to him [13]. In the telecom industry, Big Data is widely used by mobile operators. Operators, together with cellular telecommunications companies, have one of the largest databases, which allows for the most in-depth analysis of the accumulated information [14]. The main goal of data analysis is to retain existing customers and attract new ones. To do this, companies segment customers, analyze their traffic, and determine the subscriber's social connection. Apart from using big data for marketing purposes, the technology is being used to prevent fraudulent financial transactions [15]. Big data is used to extract minerals, both their processing and marketing. Based on the information obtained, companies can make decisions about the effectiveness of field development, monitor replacement schedules and equipment condition, and forecast demand for products and prices [16]. Big data is very prevalent in telecommunications, engineering, IT, financial and government organizations. According to these survey results, Big Data is less popular in the education and healthcare sectors [17].

3. PROPOSED MODEL

The proposed Services include database system architecture, infrastructure development and optimization, and data storage security. Software, hardware and services combine to form end-to-end platforms for data storage and analysis. Companies like Microsoft, HP, EMC provide services for development, deployment and management of Big Data solutions. Big data has become pervasive in many business sectors. They are used in healthcare, telecommunications, commerce, logistics, financial institutions and public administration. Databases of retail outlets accumulate a lot of information about customers, inventory management system, and distribution of marketable goods. This information is useful in all areas of store operations. Therefore, with the help of the collected information, you can manage the distribution of goods, their storage and sales. Based on the information gathered, the demand and supply of goods can be predicted.

SQL is a structured query language that allows you to work with databases. Using SQL, you can create and modify data, and

the data sequence is managed by an appropriate database management system.

NoSQL refers to not only SQL (not only SQL). It includes several approaches aimed at implementing the database, which differ from the models used in traditional, relational DBMS. They are convenient for use with constantly changing data structures. For example, collecting and storing information on social networks.

Graph Reduce is used to calculation distribution model. Parallel computing is used for very large datasets (petabytes or more). In a programming interface, data is not transferred to the program for processing, but the program is transferred to the data. A separate program has thus been requested. The principle of operation is to process data sequentially through two map and reduce methods. The graph selects preliminary data and minimizes them.

Hadoop is used to implement search and contextual algorithms for high traffic sites like Facebook, eBay, Amazon. A unique feature is that since each block has at least one copy of the data, the system is protected against failure on any of the cluster nodes. At the other end.

SAP HANA is a high-performance NewSQL platform for data storage and processing. Provides high-speed request processing. Another difference is that SAP HANA simplifies the computing landscape by reducing the cost of supporting analytics systems.

In developed countries, big data has become widespread in the fields of healthcare, insurance, metallurgy, Internet companies and manufacturing companies, and most likely in the future, Russian companies from these areas will also appreciate the effect of big data implementation and adapt them.

4. RESULTS AND DISCUSSION

The proposed data transportation model (DTM) was compared with the existing Traffic Flow Forecasting of Transportation Networks (TFTN), Logic-Driven Traffic Big Data Analytics (LTBA), Deep Learning of Spatiotemporal Patterns (DLSP) and vehicular routing protocol (VRP)

Data Volume - Volume, aggregated database is a large amount of information that is difficult to process and store in traditional ways, they need a new approach and improved tools. While most companies use Big Data in the field of customer service, the second most popular direction is operational efficiency, while Big Data is currently less prevalent in the field of risk management.

The Fig.2 shows the comparison of data volume. The technologies helped reduce reporting time by 43.5 times (from 14.5 hours to 20 minutes) and improved the accuracy of cost allocation by 40 times. Also, Big Data is introduced in planning and payment regulation processes. In total, companies use more than 300 systems based on SAP solutions, 4 data centers are involved, and the number of users is 220,000.

4.1 DATA SPEED

Speed, this feature refers to both the increasing speed of data accumulation (90% of information collected in the last 2 years) and the speed of data processing. Recently, real-time data processing technologies have become increasingly in demand.

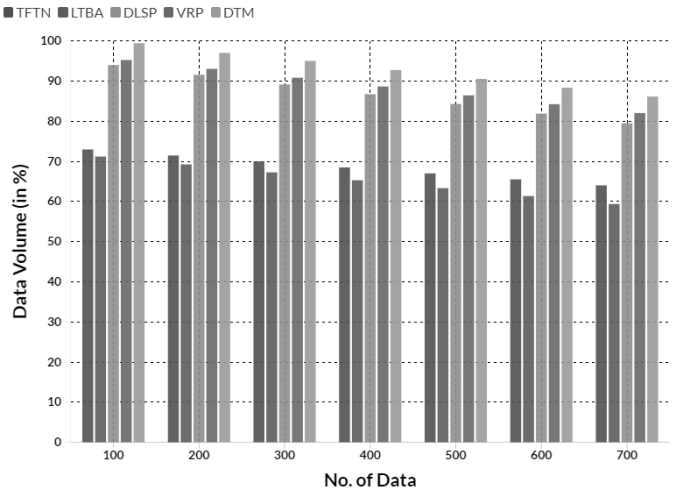

Fig.2. Comparison of data volume

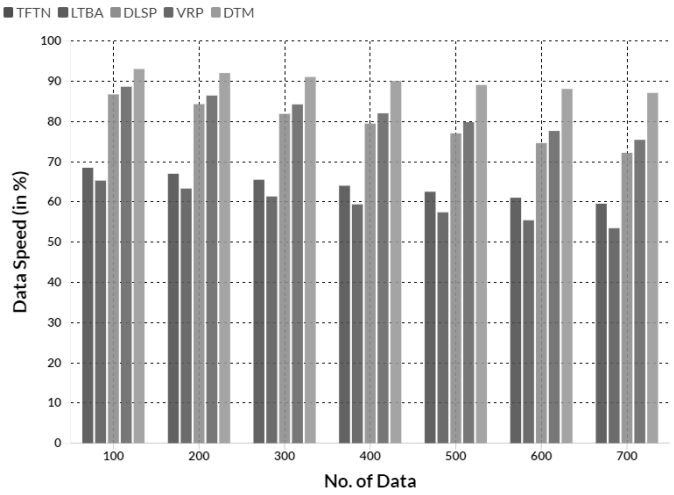

Fig.3: Comparison of data speed

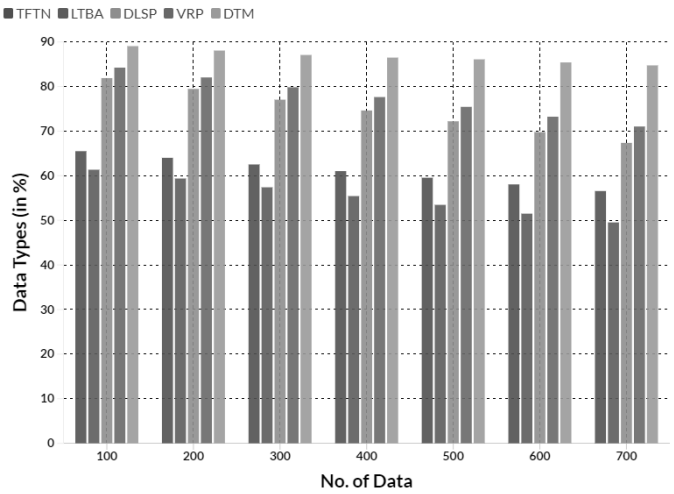

Fig.4. Comparison of data types

The Fig.3 demonstrates the comparison of data speed with the help of big data technologies, it is possible to monitor the fleet of cars, take into account fuel costs and track customer requests.

4.2 DATA TYPES VARIETY

The possibility of simultaneous processing of structured and unstructured information of different forms. The main difference between structured information is that it can be classified.

The Fig.4 demonstrates the comparison of data type. The information about customer transactions. Unstructured information includes video, audio files, free text, and information from social networks. To date, 80% of the information is included in the unstructured group. This information requires complex analysis to be useful for further processing.

4.3 DATA AUTHENTICITY

Reliability of data users have started giving importance to the reliability of available data. Therefore, Internet companies have trouble separating the actions carried out by a robot and a person on the company website, which ultimately leads to difficulty in data analysis.

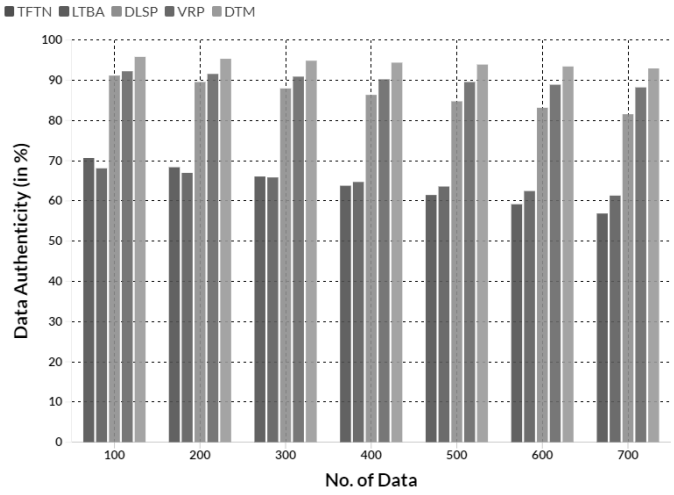

Fig.5. Comparison of Data Authenticity

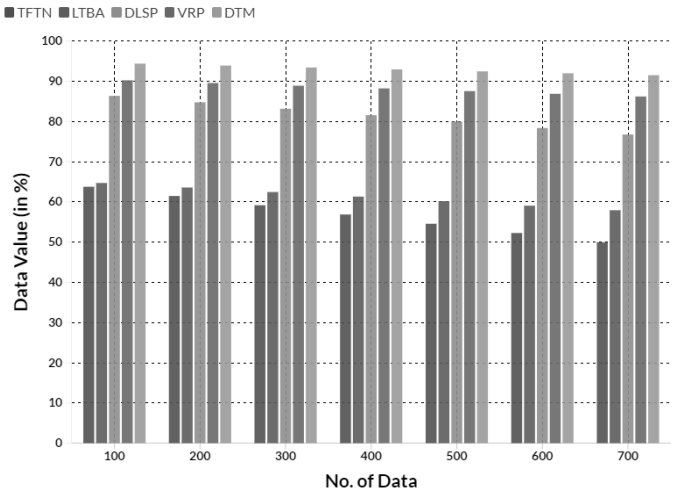

Fig.6. Comparison of Data value

The Fig.5 shows the Comparison of Data Authenticity. Big data technologies can be used to extract minerals from the gut. With their help, you can analyze the mining process itself and the most useful ways to extract it, monitor the drilling process, analyze the quality of raw materials, and process and market the final product.

4.4 DATA VALUE

The value of the information gathered. Big data should be useful to the organization and bring some value to it. For example, help improve business processes, reporting or cost optimization.

Big data technologies are used to solve national problems. These technologies enable public authorities to deliver services to people more effectively, providing targeted social support.

5. CONCLUSION

In the future there will be a trend towards the development of data visualization, analysis of media files and the Internet of Things. Despite the general stagnation of the economy, in the coming years, analysts predict further growth in the Big Data market, primarily because the use of Big Data technologies gives its users a competitive edge in increasing its operational efficiency. Business, attracting additional flow of customers, reducing risks and implementing data predictive technologies. Therefore, we can conclude that the Big Data sector in the formative stage, but the demand for these technologies is increasing every year. Based on the results of the analysis, we can conclude that the Big Data market is still in the early stages of its development, and we will observe its development and the expansion of the capabilities of these technologies in the future.

REFERENCES

[1] B. Yang and S. An, “How to Improve Urban Transportation Planning in Big Data Era? A Practice in the study of Traffic Analysis Zone Delineation”, *Transport Policy*, Vol. 127, pp. 1-14, 2022.

[2] R. Gurusamy and S.R. Seenivasan, “DGSLSTM: Deep Gated Stacked Long Short-Term Memory Neural Network for Traffic Flow Forecasting of Transportation Networks on Big Data Environment”, *Big Data*, Vol. 45, No. 2, pp. 1-15, 2022.

[3] S. Zhong and D.J. Sun, “Logic-Driven Traffic Big Data Analytics: An Introduction”, *Proceedings of International Conference on Logic-Driven Traffic Big Data Analytics*, pp. 1-32, 2022.

[4] H. Wang, Q. Li and S.B. Tsai, “Network Design Algorithm Implementation for Resilient Transportation System under Continuous Risk Perturbation with Big Data Analysis”, *Mathematical Problems in Engineering*, Vol. 2022, pp. 1-14, 2022.

[5] P. Lorenzo-Eiroa, “Space-Environment Commons: From Big Data Survey to AI, to a Post-Capitalist Blockchain Zoning Platform”, *Proceedings of International Conference on Resilient Communities and the Peccioli*, pp. 161-172, 2022.

[6] S. Zhong and D.J. Sun, “Taxi Hailing Choice Behavior and Economic Benefit Analysis of Emission Reduction based on Multi-Mode Travel Big Data”, *Proceedings of International*

Conference on Logic-Driven Traffic Big Data Analytics, pp. 227-254, 2022.

[7] Y. Wang, "Deep Learning of Spatiotemporal Patterns for Urban Mobility Prediction Using Big Data", *Proceedings of International Conference on Logic-Driven Traffic Big Data Analytics*, pp. 891-899, 2022.

[8] N.O. Aljehane and R.F. Mansour, "Big Data Analytics with Oppositional Moth Flame Optimization based Vehicular Routing Protocol for Future Smart Cities", *Expert Systems*, Vol. 39, No. 5, pp. 1-13, 2022.

[9] J. Mahmoudi and W. Luo, "Modeling the Frequency of Pedestrian and Bicyclist Crashes at Intersections: Big Data-driven Evidence From Maryland", *Transportation Research Record*, Vol. 7, No. 3, pp. 1-7, 2022.

[10] A. Song and L. Pan, "Optimization Analysis of the Emergency Logistics Identification Method Based on the Deep Learning Model under the Background of Big Data", *Wireless Communications and Mobile Computing*, Vol. 2022, pp. 1-8, 2022.

[11] S. Sun and M. Ertz, "Can shared Micromobility Programs Reduce Greenhouse Gas Emissions: Evidence from Urban Transportation Big Data", *Sustainable Cities and Society*, Vol. 85, pp. 1-9, 2022.

[12] C. Lee and S. Lee, "Exploring the Contributions by Transportation Features to Urban Economy: An Experiment of a Scalable Tree-Boosting Algorithm with Big Data", *Big Data*, Vol. 11, No. 4, pp. 577-589, 2022.

[13] M. Sutharasan and J. Logeshwaran, "Design Intelligence Data Gathering and Incident Response Model for Data Security using Honey Pot System", *International Journal for Research and Development in Technology*, Vol. 5, No. 5, pp. 310-314, 2016.

[14] H.P. Nguyen, "Applications of Big Data Analytics in Traffic Management in Intelligent Transportation Systems", *JOIV: International Journal on Informatics Visualization*, Vol. 6, No. 1-2, pp. 177-187, 2021.

[15] Y. Fang, "Research on Big Data Traffic Operation Based on Intelligent Holographic 3D Image Vision", *Scientific Programming*, Vol. 2022, pp. 1-7, 2022.

[16] J. Logeshwaran, "AICSA - An Artificial Intelligence Cyber Security Algorithm for Cooperative P2P File Sharing in Social Networks", *ICTACT Journal on Data Science and Machine Learning*, Vol. 3, No. 1, pp. 251-253, 2021.

[17] C. Wei and Y. Xiong, "The Research Development of Hedonic Price Model-Based Real Estate Appraisal in The Era of Big Data", *Big Data*, Vol. 11, No. 3, pp. 334-345, 2022.