Journal of Theoretical and Applied Information Technology
31[st]March 2019. Vol.97. No 6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DEVELOPING BIG DATA FRAME WORKS FOR EFFICIENT INTELLIGENT TRANSPORT SYSTEM FOR SAFETY AND SECURITY

P. KIRAN KUMAR[1], A. VIJAY GOPAL[2]

[1,2]Assistant Professor, Vignan Institute of Technology and Science, Hyderabad.

ABSTRACT

Big data is an advanced technology where we can use its services for our application. Big data handles large volumes of data upon a sudden rise and request of the users. Transportation system is one where we have to handle big data. Based on the gathered data the decision has to be taken for safer measures and efficient, profitable and secured. In this paper we conducted a survey on big data in ITS (Intelligent Transportation System). Data collection, analytics and frame works can be explained. We concluded from our survey by using big data we will resolve our traditional traffic problems. We attempted to provide architecture for ITS using big data thencalculation of bayonet traffic flow, calculation ofaverage speed of a road, querying the travel path of a vehicle, checking and controlling the fake vehicles.

Keywords: ITS, big data*, vehicle tracking, Large Data Storage, Cloud Computing.*

1. INTRODUCTION

In recent days big data became a popular technology. It applications have been expanding to all domains day by day. As well as the corresponding frame works, platforms and tools on which the big data can be worked very efficiently in terms of comfort and cost. Intelligent transportation systems is one where every day we are facing a lot of issues can be resolved by using big data applications handles well large volumes of data produced time to time.Big data handles a sudden rise in the data by the users when they face uneven or beneficial things from the various companies like face book, twitter etc. Big data tells the social behaviour of a group of people based on their activities over a web.

Intelligent transportation systems (ITS) have been developed since the beginning of 1970s. It is the future direction of the transportation system. ITS incorporate advanced technologies which include electronic sensor technologies, data transmission technologies, and intelligent control technologies into the transportation systems. The purpose of ITS is to provide better services for drivers and riders in transportation systems.

In ITS, data can be obtained from diverse sources, such as smart card, GPS, sensors, video detector, social medias, and so on. Using accurate and effective data analytics of seeminglydisorganized data can provide better service for ITS. With the development of ITS, the amount of data generated in ITS is developing from Trillionbyte level to Petabyte. Givensuch amount of data, traditional data processing systems are inefficient, and cannot meet the data analytics requirement. This is because they do not foresee the rapid growth of data amount and complexity.

2. LITERURE SURVEY

At present, with the rapid development of the city, people are more and more requirement on transportation,facing the normalization problem, such as city traffic congestion, traffic safety, traffic organization and so on,the traditional way of thinking has been unable to solve these problems in [5].With the rapid developmentof science and technology such as geographic information, communication, sensor and computertechnology,The new thinking mode of intelligent transportation has gradually become a practical applicationfrom concept conception, which makes traffic management more efficient, information-based and extensive in [8].

Big data is a disruptive technological change, following the cloud computing, Internet of things,which has the characteristics of large data volume, large data type, low value density, high commercial valueand fast processing speed.In

ISSN: 1992-8645www.jatit.org E-ISSN: 1817-3195

the field of transportation,the massive data include all kinds of traffic monitoring,data of services and applications,Such as roads, waterways, passenger stations and ports and other videomonitoring data, traffic flow detection of city road and highway, meteorological data, urban public transport,taxi and passenger vehicle satellite positioning data, and highway and waterway toll data, etc., these types oftraffic data are numerous and huge.

Through market research and analysis, there are some systematic products for traffic management both at homeand abroad, but there are still some problems, such as single system function, lack of integration and backwardtechnology,and that is mainly reflected in the construction of application system distributed mass data, lack ofeffective integration of traffic data, low utilization rate, the data value cannot be brought into full play, andlimited, traffic information dissemination is difficult to timely access to traffic warning etc in [9,11,21].

With the development of information technology, traffic departments urgently need a moreadvanced intelligent data analysis method, in order to carry on the efficient, real-time analysis to the massivetransportation industry's data,and to provide real-time and accurate traffic information service for travellers, so asto provide reference for traffic management departments to deal with unexpected accidents and illegal trafficbehaviour.With the emergence of mass traffic data, the technology of large data analysis brings new opportunitiesfor the development of intelligent transportation.Large data technology storage capacity and computing powerwill be more reasonable configuration of traffic resources,more effectively support transportation planning,management, operation, service and security, and provide new ideas and means for public safety and socialmanagement.

There are a number of surveys in the literature on the application and challenges of Big data analytics in RTS context. However, most of thesestudies tend to focus on a specific aspect of RTS.

For instance, in Hodge in [23], a survey of wireless sensors network technologyfor monitoring and analyzing railway systems, structures, vehicles, and machinery was conducted. A survey of railway-relatedplanning and scheduling issues in Europe was provided inTurner [22].

As another exampleNunez and Attoh-Okine in [13,15]conducted a literature review on the application of metaheuristic optimization in railway engineering. Some other survey articles onthe application of data analytics in a specific aspect of RTS can be found inSoleimanmeigouni, in [17], Singh in [24]. Hodge in [5], Thaduri [19], Griffin [6], Summit [27], Figueres-Esteban [25] explained the same.

To the best of authorsknowledge, the literature in thisfield of study suffers from the lack of a holistic survey which takes a broad perspective of RTS as awhole and cross-maps with Big data analytics.Our survey develops a taxonomy framework in Section 2, which identifies the areas of RTS and connects them with the level ofanalytics, Big data analytics models, and techniques. The developed framework aims to provide a complete picture of where and how Big data analytics hasbeen applied in RTS. To obtain this objective the study considers four aspects namely, theareasof railway transportation in which bigdata analytics is applied, thelevelof big data analytics in rail transportation, types of big datamodelsand big datatechniquesused toapply these models.

3. RELATED WORKS

Big Data has become a hot topic in both academia and industry. It represents large and complex data sets obtained from all kinds of sources. Many of the most popular data process techniques contain Big Data techniques, including data mining, machine learning, artificial intelligence, data fusion, social networks and so on. Many people use Big Data analytics in various fields, and have achieved great success.

Big data is becoming a research focus in intelligent transportation systems (ITS), which can be seen in many projects around the world. Intelligent transportation systems will produce a large amount of data. The produced big data will have profound impacts on the design and application of intelligent transportation systems, which makes ITS safer, more efficient, and profitable. Studying big data analytics in ITS is a flourishing field.

ISSN: 1992-8645www.jatit.org E-ISSN: 1817-3195

This paper first reviews the history and characteristics of big data and intelligent transportation systems. The framework of conducting big data analytics in ITS is discussed next, where the data source and collection methods, data analytics methods and platforms, and big data analytics application categories are summarized. Several case studies of big data analytics applications in intelligent transportation systems, including road traffic accidents analysis, road traffic flow prediction, public transportation service plan, personal travel route plan, rail transportation management and control, and assets.

4. ARCHITECTURE OF I TS ON BIG DATA PLATFORM

Intelligent transportation system on big data platform is a combination of multiple systems, models, department, technology. It can be said, It is a comprehensive system of system science, management science, mathematics, economics, behavioural science, and information technology. From the architecture, the platform includes basic business layer, data analysis layer and information publishing layer. As shown in Figure 1.

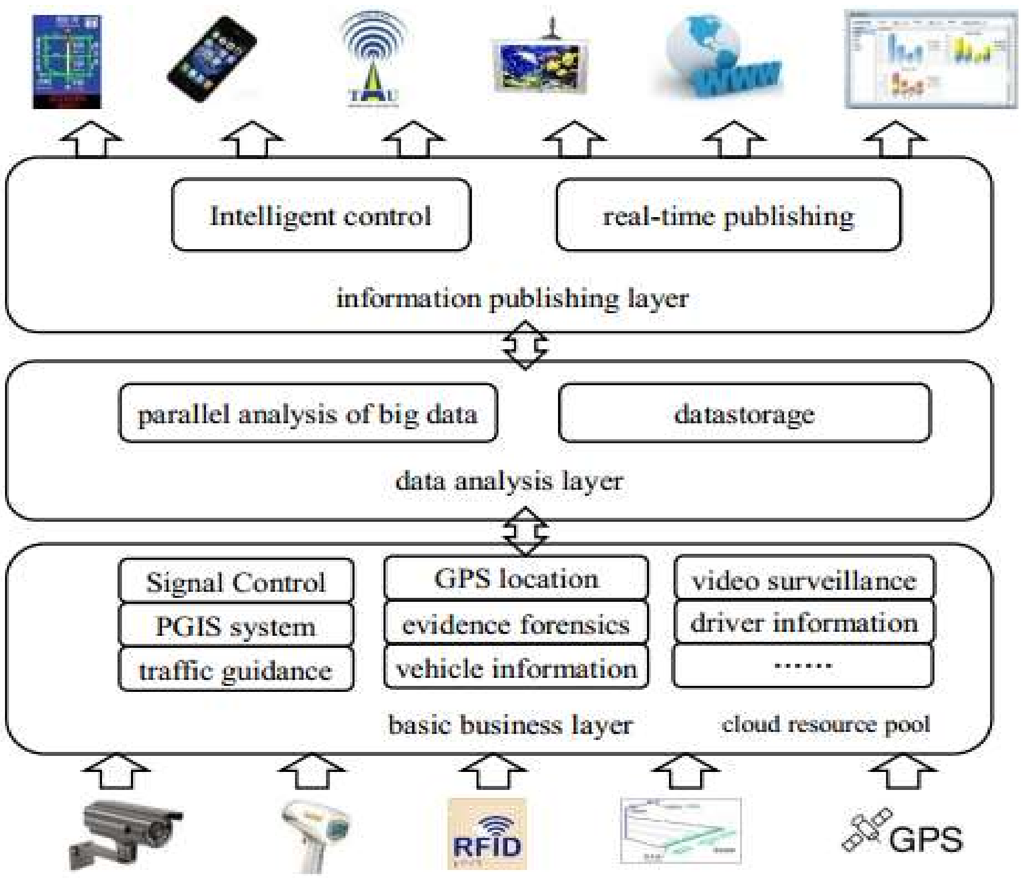

Fig.1Architecture of Intelligent Transportation on Big Data Platform

The basic business layer is the foundation of data analysis layer and information publishing layer, itsmain function is to complete the basic work of the various business units, and to produce basic business data. Itincludes traffic information collection system, signal control systems, video surveillance systems, illegalevidence forensics system, 122 alarm receive and dispose system, GPS vehicle location tracking system, trafficguidance system, vehicle information management system, driver information management system, PGIS system, and so on. the service of basic business layer is the basis for the work of the various business units,itsdata comes from data acquisition systemmentioned above, storage and handling of data is veryimportant.Therefore, cloud computing

ISSN: 1992-8645www.jatit.org E-ISSN: 1817-3195

technology can be used on the basic business layer, decentralized systemcan be

integrated into the cloud, this will ensure the security and stability of the application system, and providean efficient computing environment.

5. THE KEY TECHNOLOGYON DATA ANALYSIS LAYER

The difference between intelligent transportation systems(ITS) and the traditional traffic control system lies in its intelligent features, ITS can carry out intelligent control based on traffic condition. Hadoop ecosystem has a natural advantage in dealing with traffic big data.

Calculation of bayonet traffic flow

ITS can calculate the traffic flow of every bayonetin a certain time interval, such as 5 minutes, 10 minutes, 15 minutes, or other period of time, and push the calculated data to data publishing layer, report to the traveler, policy makers, business supervisor. Statistical analyzes were performed using Hadoop MapReduce parallel model, which is the most efficient way.The datagotfrom HBase database including bayonetID, directionID, passingtime.

The key in map() function is bayonetID and directionID, the value in map() function is passingtime. the output <key-value> pair of Map() function is <key, one>, the key include bayonetID, directionID and passingtime, the value is one. Reduce () function can calculatethe sum of one direction of traffic flow,between the start time and end time in a bayonet, the output <key-value> pair is <bayonetID_directionID_passingtime, count>.

Calculation of average speed of a road

The average speed of a road is an important indicator of the efficiency of road traffic, in general, the higher speed of traffic, the higher the traffic efficiency.The average speed is not the speed measured by radar at a place and a time point. Because it can only represent a point, but can’t represent the whole road.

$$\bar{v} = \frac{n \times s}{\sum_{i=1}^{n}(t_{end} - t_{start})}$$

where, s is the distance between adjacent bayonets, tend is the time the vehicle run out the road segment, tstart is the time the vehicle run into the road segment, tend and tstart must be the time that the same vehicle run intoand run out the road segment. The vehicle leaving or entering in the middle section of the road does not included.

6. BIG DATA ANALYSIS CLOUD PLATFORM APPLICATION ADVANTAGES

Improving the efficiency of traffic operation:

Data technology can improve traffic operation efficiency, road network capacity and facilities frequency, and control traffic demand. Traffic improvement involves a large amount of engineering, and large volume characteristics of large data help to solve this dilemma. The real time of big data makes traffic running more reasonable. When static idle data is processed and need to be used, it can be used intelligently. Large data technology has higher prediction ability; it can reduce the probability of false positives and missed reports, and provide real-time monitoring for the dynamic nature of traffic.

Improving the level of traffic safety:

The real-time and predictability of large data technology can help improve the data processing ability of traffic safety system, such as joint roadside inspection vehicle trajectory detector, big data technology rapid integration of various sensor data, construct the security model after a comprehensive analysis of the traffic safety, which can effectively reduce the possibility of traffic accident. In the field of emergency rescue, with its fast response time and comprehensive decision model, large data can provide auxiliary function for emergency decision-making command, improve emergency rescue capability, reduce casualties and property losses.

Providing environmental monitoring methods:

Large data technology plays an important role in reducing road traffic congestion and reducing the

ISSN: 1992-8645www.jatit.org E-ISSN: 1817-3195

impact ofautomobile transportation on the environment.Through the establishment of regional traffic emission monitors and forecasting model to share traffic operation and environmental data, and establish a traffic operation andenvironment data sharing test system. Large data technology can effectively analyze the impact of traffic on the environment, while analyzing historical data.The technology can also provide the basis for decision making oftraffic signal intelligent control to reduce traffic delay and reduce emissions in [11].A low emission trafficsignal control prototype system and a vehicle emission environment impact simulation system are established.

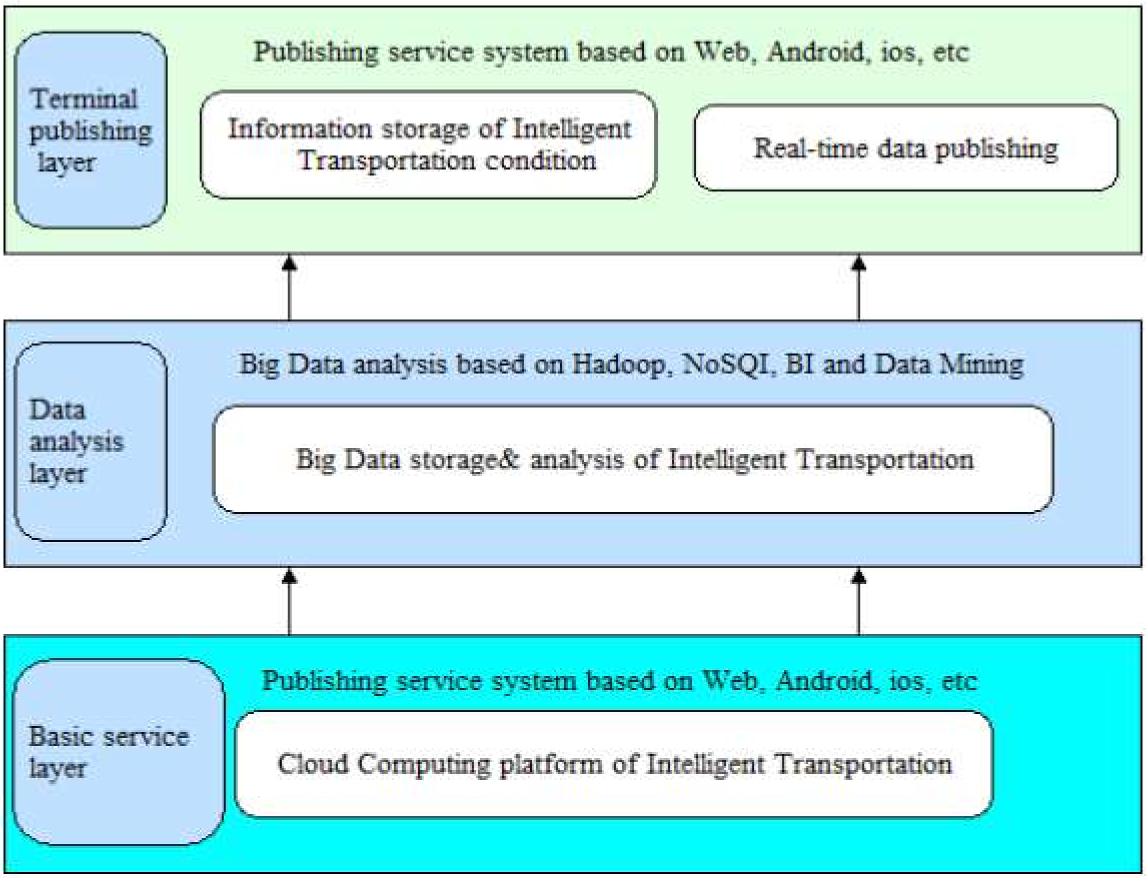

Figure 2.Intelligent Transportation Large Data Analysis Cloud Platform Overall Architecture

7. TECHNICAL REALIZATION

Platform specific technical realization idea is using a large data processing and storage technology of traffic data, combined with the big data and cloud computing technology, The use of expert mathematical model for mass traffic data for multi-dimensional analysis and mining, and through cloud publishing services, the analysis results will be communicated to all types of terminals, improve people's perception of the condition of the road bridge and traffic conditions, it enables traffic participants to complete traffic evaluation and decision-making quickly, comprehensively and accurately, and realizes intelligent traffic management.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

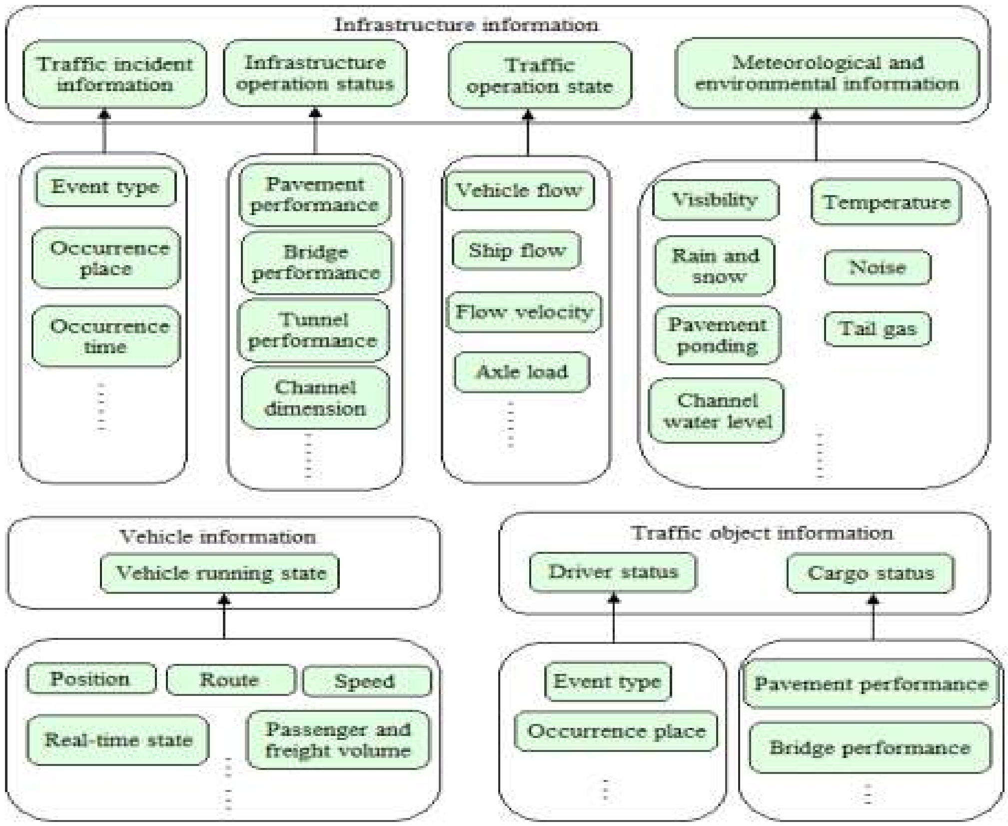

Fig.3. Virtualization

8. THE ARCHITECTURE OF CONDUCTING BIG DATA ANALYTICS IN ITS

Intelligent transportation system incorporates advanced technologies which include electronic sensor technologies, data transmission technologies, and intelligent control technologies into the transportation systems. The data collected by the intelligent transportation systems (ITS) are increasingly complex and are with Big Data features. Big companies including Gartner IBM and Microsoft put forward that that Big Data could be described by three Vs, i.e., volume, variety, and velocity.

Data collection layer: Data collection layer is the basis of the architecture, since it provides the necessary data for the upper layer. The data come from diverse sources such as induction loop detectors, microwave radars, video surveillance, remote sensing, radio frequency identification data, and GPS, etc. Details about collection of Big Data will be introduced in next sections.

Data analytics layer: Data analytics layer is the core layer of architecture. This layer is primarily to receive data from the data collection layer, and then apply various Big Data analytics approaches and the corresponding platform to complete data storage, management, mining, analysis, and sharing. Details about the Big Data analytics approaches and platform will be introduced in next sections.

Application layer: Application layer is the topmost layer in this architecture. It applies the data process results from the data analytics layer in different transportation circumstances, for example, traffic flow prediction, traffic guidance, signal control, and emergency rescue, etc.

Journal of Theoretical and Applied Information Technology
31st March 2019. Vol.97. No 6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

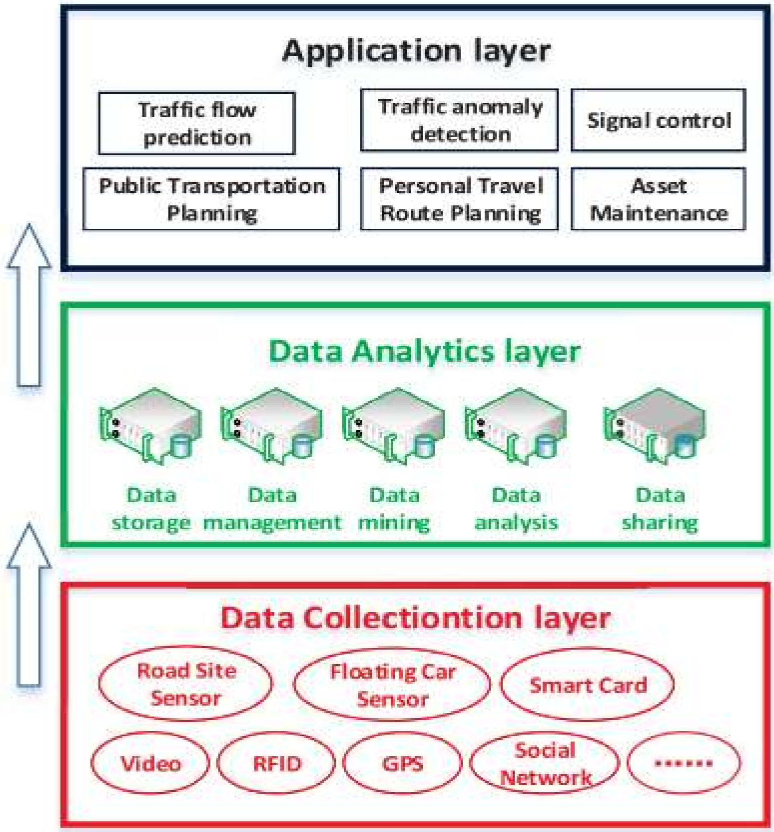

Fig.4.Data Analytics For ITS

Source	Tools	Data
Smart Grid	Smart Grid	OD flows, Travel time
GPS	GPS	Vehicle position, density, speed
Video	Video camera	Vehicle position, density, speed, classification
Road site sensor	Induction loops, road tubes, Microwave Radar, toll plazas	Vehicle position, density, speed, classification
Floating car sensor	License plate reorganization	Vehicle position, density, speed, classification
Wide area sensors	GPS, Cell phone tracking, air bone sensors	OD flows, Travel time

9. BIG DATA COLLECTION IN ITS

People unconsciously participate in the collection, transmission and application of Big Data in ITS. The technology development in ITS has led to an increase in the complexity, diversity and amount of data created and collected from vehicle, and people movements. According to different sources in ITS, Big Data in ITS can be primarily categorized into the following types, and the collected data is illustrated in Table 1 about the following of Big Data from Smart Cards, GPS, Videos, Sensors, CAV and VANET, Passive Collection and Other Sources.

10. BIG DATA ANALYTICS METHODS ITS

Machine learning is most popular modelling and analytics theory in Big Data ecosystems, which makes it easy to derive patterns and models from large amount of data. In ITS areas, machine learning theory has also be widely used to conduct data analytic. Depending on the completeness of data set that is available for learning, Machine learning models can be

ISSN: 1992-8645www.jatit.org E-ISSN: 1817-3195

categorized into supervised, unsupervised and reinforcement learning algorithms. With the recent rapid development of Artificial Intelligence, the powerful deep learning models have also been adopted to ITS recently.

Supervised Learning

Labelled training data is used in supervised learning algorithms [7]. The models use input data and the target outputs (labels) to learn the function or map between them. Combined with the learned model and the input data, the unseen outputs can be predicted. Among all the supervised learning models, linear regression, decision trees, neural networks, and supportvector machines, are the most frequently used in ITSs.

Unsupervised Learning

Unsupervised learning normally also referred as clustering focus on learning natural group from unlabeled multidimensional data [7]. K-means is the most popular unsupervisedlearning tool, and it has been widely adopted in highway transportation planning [3], and travel time prediction [7].

11. BIG DATA APPLICATIONS IN ITS

Big Data provides technical supports for the development and applications of ITS. By efficient, accurate and timely data collection, analyzing and processing in road and railtransportation system, the Big Data applications can provide the public with convenient and high efficient transportation. In order to identify problems, improving ITS efficiency, reducing costs and deriving valuable insights, Big Data applications in ITS can be divided into the following six categories.

- Road Traffic Accidents Analysis
- Road Traffic Flow Prediction
- Public Transportation Services Planning
- Personal Travel Route Planning
- Rail Transportation Management and Control
- Asset Maintenance

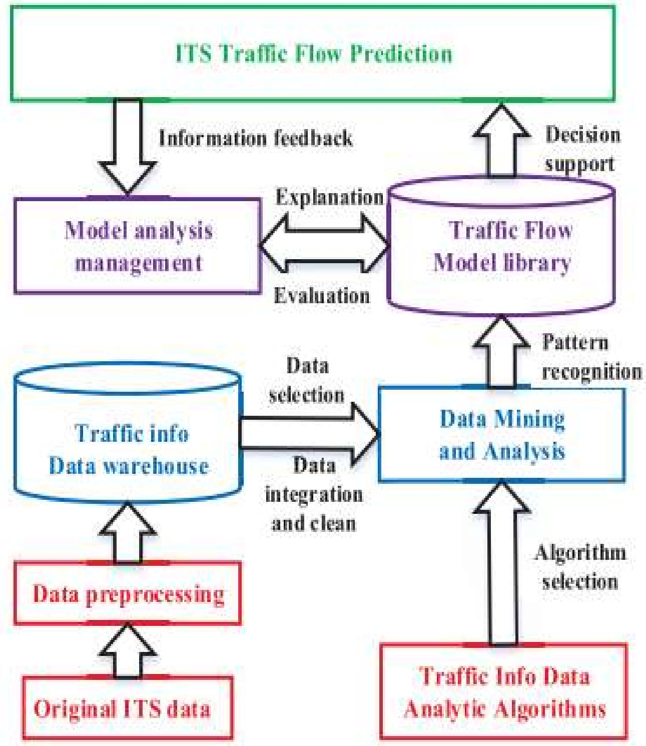

Fig.5. A typical traffic flow prediction model.

12. CONCLUSIONS

In this paper, we presented the development of Big Data andthe relevant knowledge of ITS. The framework of conductingBig Data analytics in ITS was discussed. We summarized thedata source and collection methods, data analytics methodsand platforms, and Big Data analytics application categoriesin ITS. We presented several applications of Big Data analytics in ITS, including asset maintenance, road traffic flowprediction, road traffic accidents analysis, public transportationservice planning, personal travel route planning and rail transportation management and control. Several open challengesof using Big Data analytics in ITS were discussed in thispaper, including data collection, data privacy, data storage, dataprocessing, and data opening. Big Data analytics will haveprofound impacts on the design of intelligent transportationsystem, and make it safer, more efficient and profitable.

ISSN: 1992-8645www.jatit.org E-ISSN: 1817-3195

REFERENCES

[1] Big Data Analytics in IntelligentTransportation Systems: A Survey, Li Zhu, Fei Richard Yu , Fellow, IEEE, Yige Wang, Bin Ning, Fellow, IEEE, and Tao Tang, IEEE Transactions on Intelligent Transportation Systems.

[2] Application of Big Data in Intelligent Traffic System, Gang Zeng, IOSR Journal of Computer Engineering (IOSR-JCE).

[3] Research on the application of big data in intelligent transportation system, Shengneng Hu1, Wei Liu2, Revista de la Facultad de Ingeniería U.C.V., Vol. 32, N°5, pp. 517-524, 2017.

[4] Ravindra Changala, Automated Health Care Management System Using Big Data Technology, at Journal of Network Communications and Emerging Technologies (JNCET), Volume 6, Issue 4, April (2016), 2016, pp.37-40,ISSN: 2395-5317, ©EverScience Publications.

[5] Cao W.D. (2015). Application of data mining in intelligent transportation system analysis, Computerengineering, 31(4), 91-92.

[6] Porter G., San Diego U.C., Jolla A. (2016). Decoupling Storage and Computing in Hadoop with Super DataNode. ACM SIGOPS Operating System Review, 44 (2), 15-21.

[7] Wang P. (2010). Key technologies and application examples of cloud computing, Beijing: People's Posts andTelecommunications Press.

[8] Ravindra Changala, "Retrieval of Valid Information from Clustered and Distributed Databases" in Journal of innovations in computer science and engineering (JICSE), Volume 6, Issue 1,Pages 21-25, September 2016.ISSN: 2455-3506.

[9] G. Bello-Orgaz, J. J. Jung, and D. Camacho, “Social big data: Recentachievements and new challenges,” Inf. Fusion, vol. 28, pp. 45–59,Mar. 2016.

[10] M. Chen, S. Mao, and Y. Liu, “Big data: A survey,” Mobile Netw.Appl., vol. 19, no. 2, pp. 171–209, Apr. 2014.

[11] H. Chen, R. H. Chiang, and V. C. Storey, “Business intelligence andanalytics: From big data to big impact,” MIS Quart., vol. 36, no. 4,pp. 1165–1188, 2012.

[12] T. B. Murdoch and A. S. Detsky, “The inevitable application of bigdata to health care,” JAMA, vol. 309, no. 13, pp. 1351–1352, 2013.

[13] M. Mayilvaganan and M. Sabitha, “A cloud-based architecture for bigdata analytics in smart grid: A proposal,” in Proc. IEEE Int. Conf.Comput. Intell. Comput. Res. (ICCIC), Dec. 2013, pp. 1–4.

[14] Ravindra Changala, "Data Mining Techniques for Cloud Technology" in International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE),Volume 4, Issue 8, Pages 2319-5940, ISSN: 2278-1021, August 2015.

[15] L. Qi, “Research on intelligent transportation system technologies andapplications,” in Proc. Workshop Power Electron. Intell. Transp. Syst.,2008, pp. 529–531.

[16] S.-H. An, B.-H. Lee, and D.-R. Shin, “A survey of intelligent transportation systems,” in Proc. Int. Conf. Comput. Intell., Jul. 2011,pp. 332–337.

[17] N.-E. El Faouzi, H. Leung, and A. Kurian, “Data fusion in intelligent transportation systems: Progress and challenges—A survey,” Inf.Fusion, vol. 12, no. 1, pp. 4–10, 2011.

[18] J. Zhang, F.-Y. Wang, K. Wang, W.-H. Lin, X. Xu, and C. Chen, “Datadriven intelligent transportation systems: A survey,” IEEE Trans. Intell.Transp. Syst., vol. 12, no. 4, pp. 1624–1639, Dec. 2011.

[19] Q. Shi and M. Abdel-Aty, “Big data applications in real-time trafficoperation and safety monitoring and improvement on urban expressways,” Transp. Res. C, Emerg. Technol., vol. 58, pp. 380–394,Sep. 2015.

[20] N. Mohamed and J. Al-Jaroodi, “Real-time big data analytics: Applications and challenges,” in Proc. Int. Conf. High Perform. Comput.Simulation, Jul. 2014, pp. 305–310.

[21] X. Lin, P. Wang, and B. Wu, “Log analysis in cloud computing environment with hadoop and spark,” in Proc. 5th IEEE Int. Conf. BroadbandNetw. Multimedia Technol. (IC-BNMT), Nov. 2013, pp. 273–276.

ISSN: **1992-8645** www.jatit.org E-ISSN: **1817-3195**

[22] M. Zaharia et al., “Fast and interactive analytics over Hadoop data withspark,” USENIX Login, vol. 37, no. 4, pp. 45–51, 2012.
[23] D. Corrigan, P. Zikopoulos, K. Parasuraman, T. Deutsch, D. Deroos,and J. Giles, Harness the Power of Big Data the IBM Big DataPlatform. 1st ed. New York, NY, USA: McGraw-Hill, Nov. 2012.
[24] Ravindra Changala, Data Mining Challenges With Big Data,International Journal for Research in Applied Science and Engineering Technology (IJRASET)” with Impact Factor 1.241, ISSN: 2321-9653,Volume 3 Issue VI, June 2015.
[25] L. Basche, “Says solving ‘big data’ challenge involves more than justmanaging volumes of data,” Bus. Wire, San Francisco, CA, USA,Tech. Rep., Jun. 2011.
[26] M. Bagchi and P. R. White, “The potential of public transport smartcard data,” Transp. Policy, vol. 12, no. 5, pp. 464–474, 2005.
[27] M.-P. Pelletier, M. Trépanier, and C. Morency, “Smart card data usein public transit: A literature review,” Transp. Res. C, Emerg. Technol.,vol. 19, no. 4, pp. 557–568, 2011.
[28] Y. Liu, X. Weng, J. Wan, X. Yue, and H. Song, “Exploring datavalidity in transportation systems for smart cities,” IEEE Commun.Mag., vol. 55, no. 5, pp. 26–33, 2017.