Techniques for analysing Heavy Data stream using IOT

Amandeep Kour
Mtech scholar
Sri sai College of Engineering & Technology

Deepak Kumar
Assistant Professor
Sri sai College of Engineering & Technology

ABSTRACT

In the modern era, the road infrastructure failed to cope up with the exponential increase of road traffic. There is a thrust to find a smarter ways to deal with such transportation system. Intelligent Transport System is at the forefront edge of this, one of the points is exact and hassle free forecasts that guarantee smooth and bother free driving and authoritative experience. In such manner, ITS being looked into for quite a few years and furthermore a field of consistent growth of works and advancement after some time, there is a wealth of writing on data stream expectation. Data stream datasets generated through the application of IoT are operated upon by the existing techniques. Data stream flow analysis is conducted to tackle the issues of data stream forecasting. This paper presents a systematic analysis of previous aggregate work on data stream prediction, highlight the marked changes and presents future directions for research work.

Keywords: Data stream prediction, Data stream Dataset, IoT, Data stream flow, data stream forecasting

1. INTRODUCTION

[1]Intelligent transportation system is a technique or an application in electronic or non electronic forms for producing information through advanced sensors, computers and communication technology that improve the process of data stream forecasting. ITS is wide field providing assistance in the field of driver assistance, inter vehicle communication, air data stream control, road sign prediction, number plate detection, congestion control, dynamic routing etc. ITS caters to the multidimensional needs of data stream management overlapped with number plate detection and road data stream signal prediction.

[2]Most of the issues of data stream prediction are caused due to existing infrastructure however some of the issues are also caused by poor management of data stream flow and congestion control.

[3]ITS tackles the issue of poor management of data stream flow by the use of accurate data stream monitoring and control strategies. The distributed and shared judgment and care management has be remolded an open issue at all levels of data stream forecasting systems. For the estimation of data stream prediction it requires the information that is simple and diverse from the sensors and skills.

To work efficiently there should be a ITS software system in this environment. But this system also requires credible and timely information to ensure that software can work securely and produce results within specified time. Computer systems make the interaction between human and computational devices very natural so that users can get desired data in a transparent manner. The newly introduced gadgets like mobiles, PDAs, laptops etc. make every information available anywhere at any time.

By using ITS, interactive feedback loops and video games, we can analyze the data stream related behavior changes that may occur. ITS is associated with many applications and in long term it is viable to get feasible into larger frameworks in health care.

[4]According to researchers it is suggested that use of ITS and emergence in technology is efficient enough to aware users about the current data stream and provide preventive measures. The ITS also enable user for behavior change. Distinct elements of ITS are enhancement in decision making an objective oriented. Diverting the data stream greatly depend upon the awareness of driver which will be accomplished by the use of ITS. Routing adherence is greatly impacted by this mechanism. with the help of transportation system drivers can analyze his behavior and prepare himself for taking appropriate action.

This paper is structured as, section 2 elucidating the research methodology, section 3 addressing the research questions, section 4 providing future guidelines and finally section 5 concluding the paper by presenting the results.

2. LITERATURE SURVEY

To tackle the requirements of systematic review, background analysis is conducted. The background analysis present the existing techniques that are comprehensively used to predict on road traffic.

[5]This paper depicted our examination encounters of building a keen framework to screen and control street data stream in a Nigerian city. A half and half approach got by the intersection of the Structured Systems Analysis and Design Methodology (SSADM) and the Fuzzy-Logic based Design Methodology was conveyed to create and actualize the framework. Issues were related to the present data stream control framework at the '+' intersections and this required the plan and usage of another framework to take care of the issues. The subsequent fluffy logic based framework for activity control was recreated and tried utilizing a prominent crossing point in a Nigerian city; infamous for extreme activity logjam. The new framework dispensed with a portion of the issues distinguished in the current activity checking and control frameworks.

[6]Data stream flag controller is playing increasingly and more critical parts in present day administration and control of urban traffic. This paper introduces a shrewd data streamflag controller in light of multi-microcomputer innovation. The architecture and crucial elements of the clever data stream flag controller U initially presented in detail, at that point the human-PC interface in light of visual innovation intended for the controller is figured, and lastly an application case by and by is talked about.

[7]Propelled activity data benefit framework not just give opportune and precise data stream data for activity administration work force who can adequately adjust the data stream administration control framework to an assortment of data stream conditions and street arrange limit, yet in addition help street clients, viably staying away from roads turned parking lots, diminishing auto collisions. Notwithstanding, the existing dynamic activity data is discharged for generally group of onlookers. On the off chance that the majority of the drivers utilize the dynamic data stream data to design ongoing travel courses, at that point the in general activity framework might be bothered generally, and another road turned parking lot appear in the meantime maintaining a strategic distance from the current activity stick. In light of the GIS spatial information demonstrate and the hypothesis of multi-operator, we ponder a dynamic activity data administrations innovation in view of collective multi-specialist techniques all together to show signs of improvement travel way through upgrading the communication what's more, coordinated effort between the data suppliers and voyagers. At that point the test model framework is outlined what's more, created in view of the swarm stage and java language, and some analysis data is produced by the prototype system.

[8] Dealing with the expanding activity is a major issue everywhere throughout the world. Wise Transportation System (ITS) gives answer for these issues with the assistance of new advancements. ITS is an incorporated framework that executes an expansive scope of correspondence, control, vehicle detecting and hardware advances to take care of and deal with the data stream issues. ITS is being utilized as a part of the created nations since past two decades, however it is as yet another idea when creating nations like India,

Brazil, China, South Africa and so on is concerned. In the present examination we have considered four noteworthy parts of the ITS i.e., Advanced Traveller Data System (ATIS), Advanced Data stream Management System (ADMS), Advanced Public Transportation System (APTS), and Emergency Management System (EMS). Target of the paper is to ponder different ITS engineering and model and audit such models to get top to bottom of their design. Subsequently engineering and created models throughout the times of four noteworthy branches of ITS have been inspected here to make an examination investigation of various models that have been produced by the scientists in their examinations. It will prompt the holes in the information which can be additionally considered. The paper features the conclusions extricated from the investigations of various frameworks and furthermore gives what's to come scope in the field of ITS to make it more easy to use and open.

[9]As of late notoriety of private autos is getting urban activity more swarmed. As result data stream is getting to be plainly one of vital issues in huge urban areas in everywhere throughout the world. A portion of the activity concerns are clogs and mischance which have caused a colossal exercise in futility, property harm and ecological contamination. This exploration paper introduces a novel smart activity organization framework, in view of Internet of Things, which is included by ease, high adaptability, high similarity, simple to redesign, to supplant conventional data stream administration framework and the proposed framework can enhance street activity hugely. The Internet of Things depends on the Internet, organize remote detecting and discovery advances to understand the canny acknowledgment on the labelled activity protest, following, observing, overseeing and handled naturally. The paper proposes a design that coordinates web of things with operator innovation into a solitary stage where the specialist innovation handles successful correspondence and interfaces among countless exceptionally dispersed and decentralized gadgets inside the IoT. The design presents the utilization of a dynamic radio-recurrence distinguishing proof (RFID), remote sensor advances, question specially appointed systems administration, and Internet-based data frameworks in which labelled activity items can be consequently spoken to, followed, and questioned over a system. This examination shows a review of a structure conveyed data stream reproduction display inside NetLogo, an operator based condition, for IoT activity checking framework utilizing versatile specialist innovation.

[10]This paper incorporates the plan and usage of a clever and robotized activity control framework which takes points of interest of PC vision and picture handling systems. Alongside regular PC vision strategies; this paper presents two new techniques which has low preparing cost. One of the techniques has been developed with the assistance of equipment what's more; the other one is outlined without equipment bolster. This is a finish activity administration framework which has possessed the capacity to decrease roads turned parking lots and clog on re-enacted condition. It distinguishes the quantity of vehicles on every street and relying upon the vehicles stack on every street, this framework allots improved sum of holding up time (red flag light) and running time (green flag light). This framework is a completely robotized framework that can supplant the regular pre-decided settled time based activity framework with a progressively oversaw activity framework. It can likewise distinguish vehicle condition on street and auto-change the framework as indicated by the changing street conditions which makes the framework insightful. The composed framework can help tackling data stream issues in occupied urban communities to an awesome degree by sparing a lot of worker hours that get lost attending to stuck streets. This examination concentrates on factors, ease picture preparing and activity stack adjusting.

[11] As indicated by city open travel issue trademark, the fundamental body of a paper has been submitted and has worked out one sort of in view of the Internet of things outline intelligent transportation framework. That framework gathers information by vehicle terminal and transfers information to the server through the system and makes information obvious to the purchaser passing an algorithm in the server. One viewpoint, the customer may ask about open travel vehicle data by Web. On another viewpoint, the shopper can know open travel vehicle data by station terminal. The investigations have tried that the intelligent transportation

framework can offer open travel vehicle data to numerous shoppers with helpful way along these lines this framework can take care of the city mass travel issue.

[3] This paper concentrated on the fundamental structure of canny urban Data stream Management System Based on Cloud Figuring and Internet of Things, proposed the design of canny urban Data stream Management System Based on Distributed computing and Internet of Things. The paper made a profound research on the data observing in light of Internet of things, estimation and the shrewd displaying segments what's more, learning coordinating segment. Mass estimation was acknowledged by the utilization of the distributed computing stage. The framework generally understands the shrewd observing what's more, administration of urban data stream and understands the reason for keen dig of urban traffic.

Data stream management with the implication of sensors is complex and required accuracy. Techniques devised so far still requires further enhancements for increasing accuracy of prediction. Next section presents problem definition giving parameters which can be further enhanced.

3. GAPS IN LITERATURE

Analysis of literature indicates that dataset used is offline and is not derived with the application of IoT. sensor data utilization within data stream related application is the prime cause of interest. Accurate prediction related to data stream to drivers involved along with direction sensing is missing in existing literature. Advanced application framework construction for data stream prediction is the solution for the problem.

4. COMPARISON TABLE

The comparison of various techniques that can be used to predict data streamis listed as under:-

Title	Technique	Datasets	Parameters	Merit	Demerit
A Consumer Transceiver for Long-Range IoT Communications in Emergency Environments[12]	IEEE802.11ah Wi-Fi protocol, Time Domain Least Square(TDLS)	--------	Packet Error Rate(PER), MSE	Increased range of service	Time of execution is substantially high
The advantages of IoT and Cloud applied to Smart Cities[13]	ClouT architecture which is combination of cloud and IoT is discussed	--------	-------	Sensorisation, Actuatorisation layer along with IoT have been added in CIaaS layer to extract data out of API's	CSaaS layer is still not completely defined.
Short-term data stream flow prediction using seasonal ARIMA model with limited input data[14]	SARIMA	3-Lane roadway in Chennai, India	Flow of vehicles' accuracy through MAPE	More accurate results even with data shortage	More time for computations
Smart Disease	IoT in the	Central	Predictio	Fast	Inadequate

Surveillance Based on Internet of Things (IoT) [15]	field of health care	Health Ministry	n accuracy	prediction of patterns of disease, help to take measures on time	data managers, low budget, lack of technical advisory group
Optimising Power Consumption of Wi-Fi inbuilt IoT Devices[16]	Reduce power consumption of Wi-Fi enabled devices	--------	Power consumption of various processors	Wi-Fi is better than other technologies in terms of range and security	No parameters enhancements are suggested
Energy-Efficient Location and Activity-Aware On-Demand Mobile Distributed Sensing Platform for Sensing as a Service in IoT Clouds[17]	C-MOSDEN platform	Context, activity and location aware module(Both real world and simulated lab -based data were focused on)	Energy, Storage, Communication	Sensors energy is conserved and increases lifetime of network	No focus on privacy preservation technique.
Internet of Things: Remote Patient Monitoring Using Web Services and Cloud Computing[18]	Android app is framed which takes data from IOIO-OTG board. Binary file is uploaded on cloud and processed using MATLAB	Bio-medical data like temperature, pulse,blood pressure etc.	Portability of binary data	Uniform service to patients, feasible, inexpensive	Overhead due to authentication of users. Micro-controller of higher configuration can be used.
Data Mining for the Internet of Things: Literature Review and Challenges[19]	Review of various data mining techniques and its applications is performed	--------	3 views of data mining--> knowledge, technique, application view.	Big data, data mining are hot topics to discover deep.	Parameter optimization is not considered
Combining KNN Algorithm and Other Classifiers [20]	KNN, C4.5, SVM And Naive Bayes	20 UCI Datasets	Accuracy for classsification	Higher accuracy	Execution time not considered

	Classifier(KNC)				
Intelligent Urban Data streamManagement System Based on Cloud Computing and Internet of Things[3]	Three layers of IoT architecture were combined with SOA	--------	Accuracy, Effectiveness	Specific applications were realized such as intelligent data streamcontrol, intelligent vehicle guidance, intelligent accident monitoring etc.	No real time data is involved here
Internet of things: Vision, applications and research challenges[21]	Review of IoT along with the challenges is discussed.	--------	-------	IoT applications are described ensuring its efficient use in future work	No parameter enhancement mechanism is considered
Smartphone Based Automatic Abnormality Detection of Kidney in Ultrasound Images[22]	Viola Jones algorithm, SVM, Genetic algorithm	Ultrasound images from ultrasound scanner	Prediction accuracy	Benefits rural people, can be used for emergency	Only cyst and kidney stone is considered
Spatial and Temporal Patterns in Large-Scale Data streamSpeed Prediction[23]	Unsupervised methods(k-means, self organising maps, principal component analysis) to find out global trends	Road network from Outram park to Changi in Singapore.	Prediction accuracy MSE	Spatial and temporal trends found which was not possible through use of SVM	Need to incorporate these found patterns into route guiding algorithms
Improving Data streamPrediction with Tweet Semantics[24]	Correlation analysis between data streammeasurements and number of tweets. Later optimisation framework	Data streamand data from Twitter>> San Francisco Bay area of California	MAPE and RMSE	Prediction better in comparison to auto-correlation model	Spam data presence, no work on heterogeneous traffic.

	was used.				
Road Data streamParameters Prediction In Urban Data streamManagement Systems Using Neural Networks[25]	Neural Networks	-------	Accuracy	Only for short term prediction	Better prediction model is needed for long term predico mtion of traffic
Smart video surveillance system for vehicle Detection and data stream flow control[26]	Image Processing-->Background Subtraction using Threshhold Adjusting process	Video Database	False Rejection Rate(FRR), False Acceptance Rate(FAR), Total Success Rate(TSR)	Prediction accuracy is increased by the use of video surveillance	Cameras not for night vision, situations to suspect danger not covered.
Utilizing Real-World Transportation Data for Accurate Data stream Prediction[27]	H-ARIMA+(Hybrid model of HAM and ARIMA)	Los Angeles County Transport Network	MAPE and RMSE	Short term and Long term prediction accuracy better than ARIMA, ES, NNet	Data from each sensor is studied individually. need for spatial correlations between sensors
A Comprehensive Review on Data stream Prediction for Intelligent Transport System[28]	Review of techniques used in ITS is considered like NN, fuzzy, SVM, Bayesian etc	PeMS, TMC,MIDAS, Bing data, twitter Data.	RMSE, MAPE, MRE, VAPE, EC etc	Techniques are given that can be enhanced in future work for prediction accuracy	Lack of use of deep learning techniques, datasets excluded parameters such as humidity, holidays etc.
An Aggregation Approach to Short-Term Data streamFlow Prediction[29]	Integration of MA, ES and ARIMA using NN	National Highway 107, Guangzhou, Guangdong, China	RMSE, PAE and MAPE	Accuracy is high	Situation involving multiple detectors is missing
Data streambig data prediction and visualization using Fast Incremental Model Trees-Drift Detection (FIMT-	FIMT	Department of transport UK	Prediction accuracy through MAE, RMSE and	Accuracy is high and visualization of data streampresented for better understandin	Means square error can be further reduced.

DD)[30]			SMAPE	g	
Data streamFlow Forecasting Using a Spatio-temporal Bayesian Network Predictor[31]	Bayesian Network	--------	Accuracy through MMSE	Prediction accuracy is improved since pre-processing reduces the impact of error	No real time dataset is considered

5. PROBLEM DEFINATION

The numbers of vehicles that are moved from source to destination are known as data stream flow. Data stream flow information is utilized to dissect the data stream at specific event of time. This sort of information is required so as to predict future data stream in a zone. The log can be issue if fundamental move is not made. Utilizing data stream flow information future expectation can be made hence further assistive procedures can be utilized to decrease the blockage exhibit over the system. Data stream flow information gotten from GB DATA STREAM sites can be useful in such circumstances. Data stream in a district has shifting effect. E.g. if there should be an occurrence of gathering data stream could be high henceforth it has beneficial outcome on a district. In a course data stream has negative impact so it ought to be limited. The proposed exposition work towards on street traffic. Applications of fog computing including sensors driven information for data stream prediction is missing in existing literature.

Problem definition is listed as under

- Sensor driven information through the application of fog computing is missing
- Prediction accuracy could further be produced.
- Execution time describing time consumed in providing route related information to the user can be minimized.

6. CONCLUSION

Data stream prediction using the application of fog computing is critical that can be used to monitor time critical applications such as preventing road accidents. The relevant information is required to be transferred to the source so that user who can be a driver can take appropriate action regarding route towards the destination is the prime objective of this study. Dataset derived from sensor will be used to construct real time data stream prediction framework. Accuracy will be the key parameter that could be enhanced by the application of proposed methodology.

7. REFERENCES

[1] W. Min and L. Wynter, “Real-time road data streamprediction with spatio-temporal correlations,” *Transp. Res. Part C*, vol. 19, no. 4, pp. 606–616, 2011.

[2] S. V. Kumar and L. Vanajakshi, “Short-term data streamflow prediction using seasonal ARIMA model with limited input data,” *Eur. Transp. Res. Rev.*, vol. 7, no. 3, pp. 1–9, 2015.

[3] X. Yu, “Intelligent Urban Data streamManagement System Based on Cloud Computing and Internet of Things,” pp. 2169–2172, 2012.

[4] X. Pang, C. Wang, and G. Huang, “A Short-Term Data streamFlow Forecasting Method Based on a Three-Layer K-Nearest Neighbor Non-Parametric Regression Algorithm,” no. July, pp. 200–206, 2016.

[5] O. U. Chinyere, O. O. Francisca, and O. E. Amano, “D ESIGN AND S IMULATION OF AN I NTELLIGENT T RAFFIC,” vol. 1, no. 5, pp. 47–57, 2011.

[6] C. Xiao-feng, S. Zhong-ke, and Z. Kai, “Research on an Intelligent Data streamSignal Controller,”

pp. 884–887, 2003.

[7] L. Kdqj, “An Intelligent Data streamInformation Service System based on Agent and GIS-T,” 2010.

[8] B. Singh and A. Gupta, “Recent trends in intelligent transportation systems : a review,” vol. 9, no. 2, pp. 30–34, 2015.

[9] H. O. Al-sakran, “Intelligent Data streamInformation System Based on Integration of Internet of Things and Agent Technology,” vol. 6, no. 2, pp. 37–43, 2015.

[10] T. Osman, S. S. Psyche, J. M. S. Ferdous, and H. U. Zaman, “Intelligent Data streamManagement System for Cross Section of Roads Using Computer Vision,” 2017.

[11] Y. Wang and H. Qi, “Research of Intelligent Transportation System Based on the Internet of Things Frame,” vol. 2012, no. July, pp. 160–166, 2012.

[12] M. Kim and S. Chang, “

,” vol. 62, no. 3, pp. 226–234, 2016.

[13] C. U. Scenarios and R. Architecture, “The advantages of IoT and Cloud applied to Smart Cities,” pp. 325–332, 2015.

[14] S. Vasantha Kumar and Lelitha Vanajakshi, “Short-term data streamflow prediction using seasonal ARIMA model with limited input data,” pp. 1–9, 2016.

[15] A. Mathew, F. A. S. A, H. N. Pooja, and A. Verma, “Smart Disease Surveillance Based on Internet of Things (IoT),” vol. 4, no. 5, pp. 180–183, 2015.

[16] B. D. Thomas, R. Mcpherson, G. Paul, and J. Irvine, “Consumption of Wi-Fi for IoT Devices,” no. September, pp. 92–100, 2016.

[17] C. Perera, D. S. Talagala, C. H. Liu, S. Member, and J. C. Estrella, “Energy-Efficient Location and Activity-Aware On-Demand Mobile Distributed Sensing Platform for Sensing as a Service in IoT Clouds,” pp. 1–11, 2016.

[18] J. Mohammed, C.-H. Lung, A. Ocneanu, A. Thakral, C. Jones, and A. Adler, “Internet of Things: Remote Patient Monitoring Using Web Services and Cloud Computing,” in *2014 IEEE International Conference on Internet of Things(iThings), and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom)*, 2014, pp. 256–263.

[19] F. Chen, P. Deng, J. Wan, D. Zhang, A. V Vasilakos, and X. Rong, “Data Mining for the Internet of Things : Literature Review and Challenges,” vol. 2015, no. i, 2015.

[20] Z. Zhou, C. Du, L. Shu, G. Hancke, J. Niu, and H. Ning, “Combining KNN Algorithm and Other Classifiers ,” in *2010 IEEE International Conference on Cognitive Informatics*, 2010,pp.800-805.

[21] D. Miorandi, S. Sicari, F. De Pellegrini, and I. Chlamtac, “Ad Hoc Networks Internet of things : Vision , applications and research challenges,” *Ad Hoc Networks*, vol. 10, no. 7, pp. 1497–1516, 2012.

[22] P. Vaish, R. Bharath, P. Rajalakshmi, and U. B. Desai, “Smartphone Based Automatic Abnormality Detection of Kidney in Ultrasound Images,” 2016.

[23] M. T. Asif, J. Dauwels, C. Y. Goh, A. Oran, E. Fathi, M. Xu, M. M. Dhanya, N. Mitrovic, and P. Jaillet, “Spatial and Temporal Patterns in Large-Scale Data streamSpeed Prediction.”

[24] J. He, W. Shen, P. Divakaruni, L. Wynter, and R. Lawrence, “Improving Data streamPrediction with Tweet Semantics,” pp. 1387–1393.

[25] P. Transportu and T. Pamuła, “ROAD DATA STREAMPARAMETERS PREDICTION IN URBAN DATA STREAMMANAGEMENT SYSTEMS USING NEURAL NETWORKS,” *Transport*, vol. 6, no. 3, pp. 1–6, 2011.

[26] A. A. Shafie, M. H. Ali, F. Hafiz, and R. M. Ali, “SMART VIDEO SURVEILLANCE SYSTEM FOR VEHICLE DETECTION AND DATA STREAMFLOW CONTROL,” vol. 6, no. 4, pp. 469–480, 2011.

[27] B. Pan, U. Demiryurek, and C. Shahabi, "Utilizing Real-World Transportation Data for Accurate Data streamPrediction."

[28] S. Suhas, V. K. V, M. Katti, A. P. B. V, and C. Naveena, "A Comprehensive Review on Data streamPrediction for Intelligent Transport System," 2017.

[29] M. Tan, S. C. Wong, J. Xu, Z. Guan, and P. Zhang, "An Aggregation Approach to Short-Term Data streamFlow Prediction," vol. 10, no. 1, pp. 60–69, 2009.

[30] A. Wibisono, W. Jatmiko, H. A. Wisesa, B. Hardjono, and P. Mursanto, "Knowledge-Based Systems Data streambig data prediction and visualization using Fast Incremental Model Trees-Drift Detection (FIMT-DD)," *Knowledge-Based Syst.*, vol. 93, pp. 33–46, 2016.

[31] S. Sun, C. Zhang, and Y. Zhang, "Data streamFlow Forecasting Using a Spatio-temporal Bayesian Network Predictor," pp. 273–278, 2005.