RESEARCH ARTICLE OPEN ACCESS

Effectively Error Detection on Cloud by using Time Efficient Technique

Vishakha Dattatraya Bajad
M.E (Computer)
Savitaribai Phule University, MESCOE
Pune –India

ABSTRACT

Big sensor data is huge amount of data in both industry and scientific research application in which it generate high quantity of data. Cloud computing provides a special platform to support this challenge as it provides a flexible massive data, storage, and different software services in a scalable manner at low cost. Different technique has been developing in recent years for processing sensor data on cloud, such as sensor-cloud. But, these techniques do not provide efficient support on fast detection and locating of errors in big sensor data. For faster error detection in big sensor data sets, in this system, it develop a novel data error detection approach which gives the full feature of cloud platform and the network feature of Wireless sensor network(WSN). Firstly, it classifies a set of sensor data error types and then defined it. Specifically, in proposed system, the error detection is based on the scale-free network topology and most of detection operations can be conducted in clustered form not a whole big data set.

Keywords :- Big data, data abnormality, time efficiency, sensor networks, complex network systems.

I. INTRODUCTION

Big data is a collection of data sets so large and complex that it becomes difficult to process with on hand database management systems. One of important source for scientific big data is the data sets collected by wireless sensor networks (WSN). For WSN application to get an appropriate result, it is necessary that the data received is clean, accurate, and lossless. Effective detection and cleaning of sensor big data errors is a challenging issue. Develop a novel error detection approach by exploiting the massive storage, scalability and computation power of cloud to detect errors in big data sets from sensor networks. Time efficient approach provides fast detection and locating of errors in big sensor data with error correction, big data cleaning and recovery technique. It represents the progress of the human cognitive processes, usually it includes data sets with sizes beyond the ability of current technology, method and theory to capture, manage, and process the data within a tolerable elapsed time Big data has typical characteristics of five 'V's, volume, variety, velocity, veracity and value. Using the different time efficient technique it i.e. easy to find out and locate the errors on big data set for this reason the system uses a scale free network topology for better efficiency and faster process.

The reminder of this paper is organized as follows. In Section II, it gives the Literature review of big data. In Section III, related work and analysis provided for differentiating error types in big data sets of complex network systems, such as WSN on cloud. In Section IV, based on the proposed schema of given system which mainly include different approaches for the error detection in big sensor data with cloud Computing. Section V, it conclude the research contributions of this paper.

II. LITERATURE REVIEW

Errors classification is mainly based on different algorithm which were using in error detection process. In the given system different error classification algorithms are used which gives the textual as well as numeric error [4]. Textual error is nothing but missing values. In which commonly occurring error scenario analysis is done. Different error pattern are categorized in it. Using the different Classification methods it is easy to find out the exact type of error and according to that its location [1][2][3]. After finding the error and its location data cleaning process is carried out using this corrupted files are recover easily.

There are different challenges in wireless sensor network but the main challenge is it the source to the high volume of big data set but the incoming data must be clear[1][2]. Wireless sensor network contain the low cost low-energy operating constraints of such applications. A wireless sensor network is inherently vulnerable to different sources of transient failures. A Novel data error detection technique is also developed for data processing optimization [3][5].The experiment

results demonstrate that significant performance gains in term of spatiotemporal compression and scheduling which gives compress data size and data fidelity loss properties of sensor data network to remove any overhead on the sensor nodes, at the expense of primary buffer requirements at the data aggregator nodes, which are much less cost/energy constrained.

Processing big graph data on Cloud it is costly. Managing big graph data gives complex and multiple iterations that introduce challenges such as parallel memory bottlenecks, deadlocks, and challenges, the system propose a [3]Novel technique for effectively managing big graph data on Cloud. Mainly, the big data will be compressed with its spatiotemporal features on Cloud. By exploring spatial data correlation, the system introduces partition a graph data sets into clusters. In a cluster,[4] the workload can be partition by the inference based on time series similarity. By introducing temporal correlation, in each time series or a single graph edge, temporal data compression is performing.

III. RELATED WORK

To address various challenges of big data, research works can be found intensively from the various papers [1], [2], [3]. However, the problem can be also discussed from the perspective of parallel systems and cloud [6], [8]. In this section, related work for big data set processing on cloud platform, and data error detection and location analysis for complex network systems will be reviewed and compared. There are different types of work done on the big data set using different types of classification algorithm.In big data set a complex network is represented on which error detection algorithm is applied. The error detection algorithm firstly analysis whole big data set then it filters out it. Clustering process is done on given big data set. Clustering can significantly reduce the time and cost for detecting errors and finds the exact location of error on given data set. Efficiently locate the error on big data set and then after recover the corrupted files. This is the overall process using these errors detection strategies are done and find the errors and locate it using different time efficient technique.

This system aims to develop a novel error detection approach for the big data storage, scalability and computation power of cloud to detect and locate the errors in big data sets from sensor networks.

IV. PROPOSED SYSTEM

In proposed approach, error detection is conducted in limited time period on big data set. Clustering is done over the big data. It does not consider the whole data set instead it uses clustering. Therefore the error detection and location process can be done fast. Furthermore the system uses the User define cloud known as U-cloud. Proposed system can easily minimize the time for error detection and location in big data sets generated by large scale sensor network systems with properly defined error detection accuracy.

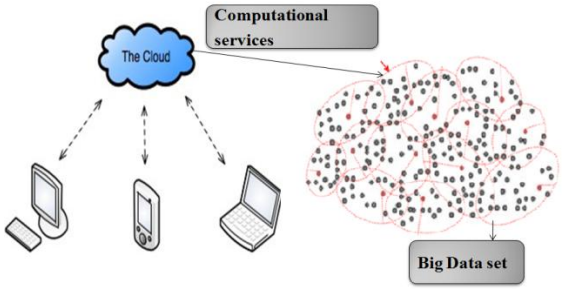

Fig. Clustering on Big Data set

In order to effectively deploy proposed algorithm on cloud, it is necessary to partition the data set before feeding to the algorithm on cloud. There are two points should be mentioned when carrying out partitioning. Firstly, the partition process could not bring new data errors into a data set or change and influence the original errors in a data set.

V. FLOW OF THE SYSTEM

That is different to the previous partition algorithm which normally divides data set according certain application preference or clustering principles. Secondly, due to the scale-free network systems being a special topology, the partition has to form the data clusters according to

the real world situation of scale free network or Cluster-head based WSN.

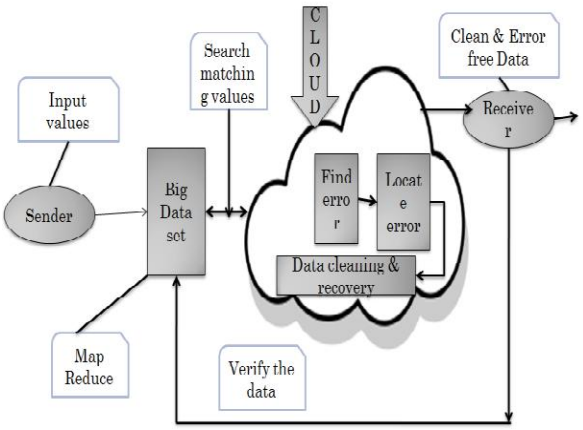

Fig. Architecture diagram

- In fig. there is a complex network and cloud platform for running error detecting algorithms.
- The error detection algorithm needs to filter the whole big data set from the network.
- Clustering is done on given big data set. It significantly reduce the time and cost.
- Efficiently locate the error and corrupted files are recovered.

For this Map reduce function is used to map the error and give its exact location on cloud

V. CONCLUSION

In order to detect and find the location of error in big data set mainly uses a sensor network systems, a novel approach is developed with cloud computing. Firstly the classification of error in big data sets is presented. Secondly, the correlation and comparison between sensor network systems and the scale-free complex networks are introduced. According to each define error type and the features from scale-free networks, the system proposed different error detection strategies for detecting and locating errors in big data sets on cloud. All the process of error detection is conducted over the User defined cloud i.e. U-cloud. The significance for the system is: 1It gives the fast error detection and locating using time efficient approach and 2) after error detection process done effectively recover the corrupted file to avoid loss of data.

REFERENCES

[1] D.J. Wang, X. Shi, D.A. Mcfarland, and J. Leskovec, "Measuremen Error in Network Data: A Re-Classification," *Social Networks*, vol. 34, no. 4, pp. 396-409, Oct. 2012.

[2] S. Slijepcevic, S. Megerian, and M. Potkonjak, "Charaterization of Lacation Error in Wireless Sensor Networks: Analysis and Application," *Proc. the Second Int'l Conf. Information Processing in Sensor Networks* (*IPSN'03*), pp. 593-608, 2003.

[3] A. Sheth, C. Hartung, and Richard Han, "A Decentralized Fault Diagnosis System for Wireless Sensor Netowrks," *Proc. IEESecond Conf. Mobile Ad-hoc and Sensor Systems (MASS '05),* Nov. 2005.

[4] M.M.H. Khan, H.H.K. Le, H. Ahmadi, T.F. Abdelzaher, and J. Han, "Dustminer: Troubleshooting Interactive Complexity Bugs in Sensor Networks," *Proc. ACM Sixth Conf. Embedded Network Sensor Systems (SenSys '08),* pp. 99-112, 2008.

[5] X. Zhang, T. Yang, C. Liu, and J. Chen, "A Scalable Two-Phase Top-Down Specialization Approach for Data Anonymization Using Systems, in MapReduc on Cloud," *IEEE Trans. Parallel and Distributed*, vol. 25, no. 2, pp. 363-373, Feb. 2014

[6] N. Laptev, K. Zeng, and C. Zaniolo, "Very Fast Estimation for Result and Accuracy of Big Data Analytics: The EARL System," *Proc. IEEE 29th Int'l Conf. Data Eng. (ICDE)*, pp. 1296-1299, 2013.

[7] J. R. Pansare and V. D. Bajad , "Errors Detection in Big Sensor Data on Cloud using Time Efficient Technique." http://dl.acm.org/citation.cfm?id=2909071&dl=ACM&coll=DL&CFID=664730537&CFTOKEN=123717