International Journal on Future Revolution in Computer Science & Communication Engineering
Volume: 4 Issue: 3
ISSN: 2454-4248
441 – 448

Implicit Study of Techniques and Tools for Data Analysis of Complex Sensory Data

Mr. Vishwas S
Assistant Professor
Department of Computer Science & Engineering
KVG College of Engineering
Sullia, Karnataka, India
vshcool@gmail.com

Dr. Hareesh K
Associate Professor
Department of Computer Science & Engineering
Government Engineering College Ramanagara
Karnataka, India
hareeshk.gec@gmail.com

Abstract— The utility as well as contribution of applications in Wireless Sensor Network (WSN) has been experienced by the users from more than a decade. However, with the evolution of time, it has been found that there is a massive growth of data generation even in WSN. The smaller size of sensor with limited battery life and minimal computational capability cannot handle processive such a massive stream of complex data efficiently. Although, there are various types of mining techniques being practiced today, but such tools and techniques cannot be efficiently used for analyzing such complex and massively growing data. This paper therefore discusses about the generation of large data and issues of the existing research techniques by reviewing the literatures and frequently used tools. The study finally briefs about the significant research gap that calls for need of data analytical tools in extracting knowledge from complex sensory data.

Keywords- *Cloud Computing ;Data Analysis ; Sensor Data Analysis; Tools of Sensor Data Analysis.*

*****

I. INTRODUCTION

The advancement in Wireless Sensor Networks (WSN) application has become a trend in various environments, commercial, and industrial fields. The sensor network consists of many sensor nodes to act upon the region monitoring and data fetching for surroundings. The WSN has spatially distributed and self-regulated sensors that favor environmental conditions monitoring like temperature, motion, vibration, pressure, pollution and so on [1, 2].In the military, many applications are used to track the unusual activity or events. A military application also includes seismometers, audio or visual cameras for tracking unusual activities or events. It may also include the seismometers, audio or visual cameras to track the movement of larger objects. The sensors are used in the context of several environmental applications like pollution level tracking in water, climate trend, and weather detection [3]. The sensor data processing requires real-time and efficient processing of uncertain data of massive volume. Thus, sensor data will bring many challenges in it with a context of processing, storage and collection of data. Some of the challenges were stated as follows: i)bigger challenges of processing data owing to incompleteness and natural errors in data aggregation, ii) limited or less battery life of sensors that results in inability to transmit or collect the data for long periods of time. The errors in data might cause uncertainty in data representation. Therefore, a precise methods is needed to be processed when the uncertainty is present in the data. Although, sensors are also needed to be designed for processing in real-time applications, but they are already shrouded with various issues and limitations in processing data. The algorithms meant to address these issues are needed to execute data in one pass, as it is difficult to store the entire set of data because of other constraints and storage. The massive volume data storage may lead big challenge in data storage and processing [4]. Another bigger upcoming challenges face by WSN is to handle and process a massive set of data stream, which is unstructured and highly voluminous. Since 2008, it is seen that the devices connected to the internet have exceeded more than the humans present on the planet. Thus, it is considered that the data generated by the human are less than thedata generated by machines, and this gap is likely to be increasing in future days. These are widely known as big data problems in analytical applications or steam processing problems by information overload. In several cases, performing processing in-network is critical, within the network data is processed than at centralized service. In real time, within the network mining algorithms and other queries are processed, whereas the processing algorithms are distributed for effective design needs. The next level information technology is sensor data analytics[5]. Many experts predicted that sensor data volume will surpass social media for next few decades; many enterprises were waiting to know the benefit of sensor data analytics in the business point of view for upcoming years. Many organizations were failed to get this point. The legacy integration technologies are not keeping the pace for emerging many new numbers of sources for sensor data. The existing infrastructures cannot handle the massive data sets required for sensor data analytics. The sensors have become simpler and cheaper to use [4].

The prime objective of this paper is to investigate the significant issues in sensor data analytics for exploring the better possibilities of existing data mining techniques. The paper discusses about the existing techniques from the review of literature accompanied by discussion of frequently used tool in analyzing sensory data. Finally, the study discusses about the research gap from teh existing literatures. Section 1 discusses about the brief introduction to the topic under study along with background discussion and problem discussion. Section 2 discusses about the prominent issues in sensor data mining followed by highlights some some frequently used tools in Section 3. Finally Section 4 discusses about the research gap followed by summarization of study in Section 5 as conclusion.

 ISSN: 2454-4248
441 – 448

A. Background

The section includes previous studies that focused on related issues of a sensor on the cloud, data security, and privacy or efficient platforms of provision for processing and storage of sensor data.

Wang et al. [6] have presented the definition for big data and anomaly detection. The researchers have used the theory of compressive sensing in the algorithm for anomaly detection. The illustration for human detection behind gypsum and brick wall is carried out based upon radar signal of ultra-wideband. The outcome of the study gives the effective algorithm for anomaly detection for human detection through uncompressed data and compressed signals. Tiwari et al. [7] have discussed the data management on cloud and storage aspects for cloud computing. The authors used the used data stores like NoSQL and NewSQL. The domain comprehends and existences of diversity problems were discussed. The study outcome gives challenge identification in technology inconsistency and immense diversity and gives best needs for the needs for solutions. Tambe et al. [8] have illustrated a system based on real time to analyze and distribute sensor data with high velocity from the soccer game. The study implemented an OMG Data Distribution (DDS) System for data circulation and combined the algorithm to present required real-time analytics. The results are examined by comparing graphs of filtering outliers and player distance, and by per player data refresh value. The study result presents perfect solution for reliability, timeliness and distribution, scalability to store a huge number of data and in-network database. Melchor et al. [9] have presented the data channel design for integration of sensor cloud. The authors focused on a problem of establishing a channel series elastic data of sensor to users (Mobile) through jobs running on data analyzing in the cloud. The results are concluded through graphs of forwarding graphic performance and sensor data performance. The illustration outcome provides a well-designed connector toolkit of the data channel for implementation in the integration of sensor-cloud. Selia et al. [10] have discussed challenges and applications of big data analytics. The authors have discussed deep learning aspects in complex data extraction from the huge volume of data, data tagging, semantic indexing, discriminative task simplification, etc. The outcomes the discussion gives sampling criteria for data, modeling for domain adoption, data streaming, and data scalability. Rong et al. [11] have illustrated data analytics for privacy and security preserving in sensor data without compromising with presented data utility. The illustration outcome presents a new approach to getting data privacy and security in the complete lifecycle of data, data collection/generation, storage, sharing and processing and transfer. Ride et al. [12] have presented the design concept to link sensor data to cloud system. The authors have concentrated on communication and ubiquity smartphone features to unify process collection, analysis and storing for long term sensor data for the multi-user purpose. The outcome of study gives a health based implementation in the storage device. Goyal et al. [13] have discussed the utilization of efficient energy in a cloud. The authors have illustrated several methods and techniques to solve the problem of huge power consumption in a cloud system. The study outcome presents many energy efficient Particle Swarm Organization (PSO) and Firefly algorithm (FA) methods to reduce consumption of energy in cloud computation. Healy et al. [14] have presented sensor data management on the cloud. The cloud computing issues are focused on presenting platforms for data gathering and monitoring of sensor data. The outcome of the discuss gives efficient platforms for cloud-based sensors. Alamri et al. [15] have illustrated infrastructure of cloud-sensor including its architecture, approaches, and applications. The authors have concentrated on issues related to energy, memory, communication, computation and scalability for WSN. The cloud sensor infrastructure innovation capability outcome with implementation technology to face the real world situations. Ebner et al. [16] have discussed the big data analytics applications on the cloud. The problem of data collection is analyzed. The research outcome gives many efficient solutions for data collection.

Veen et al. [17] have presented cloud analysis using sensor data. The author has focused on data demand for storage and transport. The results were evaluated by comparing platform of stream processing and to existing batch and simulation results of different schedules as single, module driven and recurring. The outcome of the research gives simple cloud analysis programming interfaces. Arshi et al. [18] have given issues of security in cloud computing. The information security, data piracy, network security, computer security issues are discussed. The discussion outcomes with security approaches and security of cloud environment for business. Abadi [19] has discussed the cloud data management opportunities and limitations. The author has speculated analysis of large-scale data tasks, specific data applications and system support decisions. The discussion outcomes with a proposal of well-designed security for computing cloud environment. Lim et al. [20] have presented infrastructure of sensor-cloud and technical issues in it. The author has presented an addressed issues like subscribe/publish and sensor description middleware. Skov et al. [21] have illustrated data modeling for sensor based instrument i.e. electronic nose, the more experience of sensor data analyzing is needed for multi-way approaches. The problem of more data extraction of data for the need of proper classification and qualification are taken into account for module designing. The research outcomes with an efficient model for pre-processing, processing, post processing of data correctly have found easy, the fast potential for data analytics. Bhunia et al. [22] have discussed the framework of sensor-cloud service health-care. The authors also discussed the related issues of the health-care remote system and how the health records can be extended to include them in sensor data. The paper outcomes with a low-cost remote system for health-care on the basis of sensor-cloud.

B. Problem Description

The area of wireless sensor network possess the issues right from beginning of data collection, knowledge discovery, and mining, cleaning and data management. Even more, issues will arise in the field of research in some application domain processing of in-network. The application domains will arise in the issues in pipeline parts as well. The below section explains key issues in sensor data analysis [5].

- **Data Collection and Cleaning Issues:** Number of issues will arise in sensor data collection in context. Sensor data's are inherently uncertain and full of noise and readings of reluctant will be depending upon application domain. In case of Radio Frequency Identification (RFID) data, 30% of readings were dropped and sensors will track

 ISSN: 2454-4248
441 – 448

similar RFID objects. In case of sensors driven by a battery, many losses will take place while transmitting data and it may also incompleteness in significance due to limited or less battery life[23].

- **Data Management Issues:** The collected data will face huge challenges from largely collected data (of collected data). In many cases, the data volume will be huge so that it may drop or compress the important parts of data. The uncertainty and errors of sensor data will cause/motivate the algorithm development for database management in uncertainty[23].
- **Sensor Data Processing and Mining:** It is quite a difficult task to process a massive stream of data encapsulated by the sensors which have low computational capability. These are considered as stream data mining algorithms. The further more advantage of this algorithm that in time processing will be performed, before sending the processing results to a higher storage level, it performs partial data processing within the network[23].
- **Specific- Application Issues:** The sensor data will be used in many domains like astronomy, RFID, sensor networks for military, mobile data and in the environment. The different kinds of domains will give different issues in point of processing and storage. In Radio Frequency Identification, data may have a higher level of uncertainty and redundancy, and the applications in mobile data mining may need the mining techniques like Spatio-temporal. Recently in the aspects of database management were designed sensor data in these conditions[23].
- **Query and Data Issues:** The issues regarding the query handling and data processing are discussed here in case of sensor data. In sensor data query processing and indexing are also more important. In sensor data query processing, the sensor data will face a huge number of challenges from query and indexing prospective processing, due to receiving a massive data volume at a time. In some cases query processing of event detection, the queries are causing continuously in sensor data to detect underlying events. Moreover, it is quite difficult to understand the underlying semantics of the raw data if the size of the data is quite massive in size. Moreover, origination of defective query process will always results in false alarm in the detection process by the sensor nodes [23].

II. THE SENSOR DATA MINING ISSUES

At present, there are many methods of classification, clustering, outlier detection and frequent mining pattern extraction process in the literatures [24]. The data are needed to be filtered and compressed for even more effective analysis and mining. The big problem is that the algorithms for the conventional mining are not designed for real-time data processing. Thus, new algorithms of sensor data for stream processing are needed to analytic performs in real-time in a single pass. Additionally, the scenario of a sensor are required for processes in-network, and the processed for representation of further high-level data processing, this will reduce the data overload and transmission cost from the perspective of storage. The stream mining and compression were together integrated tightly from perspective efficiency. The modelling and compression of hidden variables will give the summary of representation and will be used in applications like outlier analysis and forecasting [23].

The biggest problems or challenges in sensor-cloud operation are establishing the data channels series from sensors for any mobile users by data analyzing processes in the cloud. The network protocol variation, user mobility and dynamism of resources in cloud computing will result in many problems in design. Analysis of data must possess the following characteristics e.g. i) deals with protocols of cloud communication, ii) include operations for data retrieval and sensor manipulation (like checking of web portal and packet filtering), iii) handle performance of job migration and availability of resources, for recovering broken (data) connection. The data analyzing designers and users were not skilled in much network protocol [25].

In case of stream processing, the data from the input is analyzed first, and the output will be available immediately, which means one will act after as another data arrives. A give skewed results will arise as the analysis is not finished actually. In retailer case, partial results comparison from one country next to other is more difficult due to different business hours and time zones. For quick response to anomalous situations, current sensor analysis is carried out as the data arrives. The process of stream approach, i.e., calculation of an average moving value of 20 latest sensors will be the best for these analyzes. Platform for a sensor with multi-purpose data analysis is carried out by using stream and batch processing combination i.e. current detection in the anomalous situation and historical analysis respectively [26].

There are many issues exists in a spectrum, regarding software and hardware components heterogeneity. In lower levels, sensor plethora exists, which functions in data gathering of different formats. The generic interface design of these kinds of data requires study on types of existing data there aggregation or conversion or adoption to a standard format. In higher level data processing, the challenge lies in control mechanism in data access and security. The gateway server authentication before forwarding the sensor data to the cloud will not be enough. Many of the customers worried about their data privacy and how the data is used by the cloud providers on a cloud. For addressing these kinds of problems of different security level assurances will be required in securing and controlling the data usage on a cloud [27].

The layered Sensor-cloud structure is shown in Fig.1. The structure consists of three main layers:

1. Layer-1 pertaining to application and user layers: This layer process user and user relevant applications. Many of the users were interested in valuable sensor data access from other Operating system platforms like windows OS, Mac OS, or Mobile phone OS for many applications. The different user platforms were allowed in structure to process and utilize sensor data without any problem occurrences due to high storage and cloud infrastructure availability.

2. Layer-2 pertaining to virtualization and sensor-cloud layers: The layer aims with physical sensors virtualization and cloud resource. The virtualization will enable the provision of cloud-sensor services and remotely the many IT resources for end-users without worry of exact sensor located. The sensors with virtualization were automatically created by the use of service templates. The templates of services were created by

 ISSN: 2454-4248
441 – 448

providers like service catalog, which automatically enables service creation to make accessible for multiple users [28].

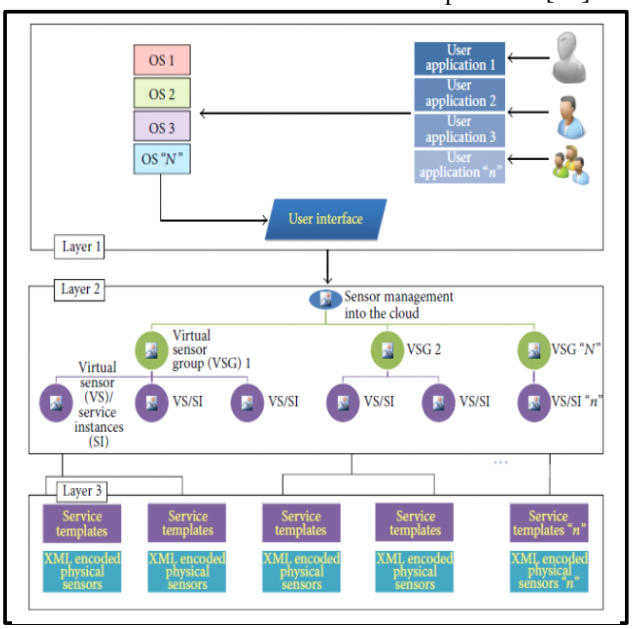

Figure 1: Layered structure for Sensor-Cloud

3. Layer-3 pertaining to tangible sensor and template creation layers: The layer deals with template creation for service and for catalog layer definition in catalog menu forming. The physical layers were retrieved and located in this layer. As every physical sensor poses own data collection and control mechanism, the standard mechanisms were defined, used to sensor access with irrespective of differences between many physical sensors. User can access the virtual sensors with well-defined standard function. The layer 2 translates virtual sensors standard functions into many specific functions to variant physical sensors (in layer 3). Physical sensors were coded with XML which enables provided services through these kinds of sensors to utilize many platforms with irrespective of other platforms. The sensor-cloud offers an aggregation platform based on web for data of sensor to develop user based applications [28].

III. TOOLS OF SENSOR DATA ANALYTICS

The section will deals with the platforms and tools for gathering, processing and monitoring sensor data. This will give cases of practical use and their features of characteristics.

A. IBM Mote Runners

IBM's Mote runner is the WSNs infrastructure [29]. It was based on targeted virtual machine to resource constrained environment for hardware and it has contains two parts as like a) runtime hardware of mote class (Ex: MEMSIC Iris or Libelium Wasp mote), b) Environment development of WSNs applications. In its core, design of mote is designed to run over very small, embedded controllers, with processor of 8-bit of low power, standard, and then reduce in maintenance cost both post deployment as well as initial investment. It shows less energy consumption [30]. It provides language friendly, resource efficient, virtual machine with high performance, which will shields applications (portable) by hardware specifics. With this Mote runner will allows programmers for implementation in programming languages like java and eclipse (environment based development) to develop WSN applications which may distributed dynamically. The features includes as follows [31]:low power,Supports in solar power harvesting as energy source,For developers it will provide software development Kit (SDK),Will support C# and Java,Will run on 8-bit micro controller with 4KB RAM and 32KB flash memory,Supports in environment developments like Visual studio, Eclipse and in MonoDevelop.This tool will focus on Mote actuators and applications related with WSN. There is no or little information is available about its wired support and sensor types. The details for data interchange are publicly not available.

B. TempoDB Platform

It is an database, aims in analyze and storage data of time series from sensor, servers, automotive telematics and smart meters. This tool favors following features [32]:Simple API (REST) for retrieval and data storage, allows SSL encryption for every data transfer, allows all data storage at no down sampling (i.e. full resolution), provide support to Internet of things, three times the data replication and provides data availability guarantee, availability of API clients in many programming languages as .net, ruby, python and Java.

As per primary aim of this platform is time management of data sets in format of IS08601 [33]. In data querying from storage of TempoDB by client API, the data returned will be formatted only by JavaScript Object Notation. Other interchange data formats will not be available. The platform has expanded, renamed as TempoIQ [34].Now it will offers alternating mechanism and data monitoring of flexible sensor, also offers tools of analytic in support of user applications.

C. Sensor-Cloud Platform

MicroStrains of sensor-cloud will offer data storage for sensor, remote management and visualization platform which leverages technologies of cloud computing to provide rapid visualization and data scalability [35]. Initially it was designed to provide support for deployment (long-term) of wireless (Micro Strain) sensors. Now, it will support any devices of third party like web-connected, sensor networks or sensors through Application Programming interface (API). The sensor-cloud platform aims to give unlimited storage to sensor data [36]. The features includes it will give storage for unlimited data, time series graphing tool and visualization, facilitates in quick development of user applications by using their cloud data, provides alerting capabilities like E-mail, SMS and so on, and provides API for data transport.It only supports formats for data interchange of Comma Interchange value (CSV) and external data representation (XDR).

Some scenarios for usage includes assets of high value monitoring based on condition and monitoring of structural health, in which data tools available are not capable of data scalability, performance, programmability or accessibility. Every communication with sensor-cloud will be performed on secure Hyper Text transfer protocol (HTTPS), will be secure. However, it will restrict low level communication by forcing HTTP use. HTTP is middleware to platform and device connection. CSV and XDR only format types of data interchange supports currently.

 ISSN: 2454-4248
441 – 448

D. Ostia Portus Platform

It was designed to take part between technologies of multiple vendors, where every vendor has individual programming language, database and platforms. This will achieve by having sets of isolated data form individual separate system [37]. This kind of platform connects many devices including networks, platforms and sensors.The platform features are Relational database support, Uses SSL i.e. Secure Socket Layer over HTTP for security. It supports in Java Message service, Internet Protocol (IP) or Transmission Control Protocol (TCP), RabbitMQ, transport protocols of Message query (MQ).It is easy to installation and use. Portus was built using open exploits and technologies to open standards over organizational data accessing and giving them by defined views of business. The core components are Web services providing by front-end, Server will be written and hosted by using C++ and C and apache respectively, and Administration control centers are written using Java and were built for run using Eclipse framework.

E.Bluwired Sensor-Cloud Platforms

It provides a platform for exploration of sensor data, analysis and interaction [38]. This facilitates in management of device data and sensor data of web enabled place in the world and hence to deploy application analysis and processing of data which gathers cloud data. It has features like Data management from web enabled place in the world of sensor data, Supports in deployment and development of data for analytical application and processing by using cloud data, A platform for interaction, exploration of sensor data and analysis, Allows visual management of data interface, Favors data storage in real-time and retrieval, and Provides mechanism for an abnormal event.

The Bluwired Sensor Data platform favors in sensor data analysis, tracking and reliable storage, which is resulted from control solutions in many applications like process control, agriculture, Irrigation, Factory automation, monitoring system for patients, gas and oil. It is a tool which does not allow an API open source for platform accessing.

F.FreshTemp Temperature Monitor

This is a cloud based monitoring system used on perishable goods [39]. This automates collection of temperature during the production, storage and transportation of perishable products by offering integrating capability with monitoring different kinds of temperature sensors of such products. It provides online dash boards, configurable alerts and data logs of real-time. This commercial tool aims to offer solutions for transportation, healthcare, industrial usage and food services. Its core has features such as Wireless monitoring for temperature, Data logs of real-time, Bluetooth data loggers, Bluetooth food probes, and Mechanism of alerting by Email, phone or SMS.

This tool favors solely for temperature sensors that limits usability for controlling and managing other types of sensor devices. Moreover, it will not provide any API programming for development of application or accessing any stored data by development.

G. SensarTrack Monitor

This is an independent monitoring multi-perform service which deals with Machine to Machine sensor monitoring. This was developed by Ms. Cannon Water Technology Inc. and was released on 2012 [40]. The Monitoring software offers sensor data visualization to user from web supported device. This was designed to enterprise with many distribution assets as tanks of bulk storage,

silos, bins and storage facilities of chemical. The tool involves features as Quick scanning ability of multiple locations, Servers for secure data, Gateways for wireless data, Finds problems early and trend tracks, and Installation is easy for non-technical personnel.

Sensatrack uses equipment of wireless communication by development of Digi Corporation on protocols of zigbee. This supports networks of wireless and wired networks. This will not support any development of application and also does not offer external API access.

H. Nimbits Platform

The platform as a service in development hardware and software solutions which connects each other and cloud. This shares and records sensor data over cloud [41]. In Nimbits, the sensor data will be stored using textual as formats of data points, eXtensible Markup language or JSON. This offers web services of REST for retrieving and logging time and some data of geo stamped (as temperature sensor reading). These servers of Nimbits will run on powerful platforms like raspberry pi (smallest) and app engine of google. The interface of graphic user has tree structure with child and parent structures [42]. The platform has features such as Ability in data sharing and recording, Easy connections of analytic tool and data, Data stored is data points, Ability in generation of events or threshold based alerts, and Visualization graphic interface of user. Nimbits is a platform of open source for Internet of things. It available easily and provides APIs, libraries and documentations in other programming languages.

I.Xively Platform

The platform which harnesses IoT power to easily and quickly transform the connected vision of product to market reality. It is called as “pachube", later "COSM". Currently, it is a LogMeln Inc. division and offers business solution via IoT. It gives a platform with applications to connect the products and devices for control of real-time, storage and management. The tool features as Secure messaging in real-time, Data storage in time series, selected data sharing, Offers easy connections to provide services for external cloud such as Facebook and Twitter, Transport Layer Security Encryption with SSL, and Message bus in real-time based on MQTT, i.e. Message Queue Telemetry Transport.The project 'Park-A-Lot' is the example of Xively platform use that has been designed for support over management system of automated parking.

J. Intelligent service of Microsoft Azure

This aims in secure connect, capture and manage the data generated by machine from assets of Line of Business sensors, industrial devices across platforms of operating systems. The service represents the Microsoft effort in addressing IoT challenges to utilize its business potentials. The tool offers to extend cloud of Microsoft azure across devices and sensors connected for virtual data analyzing, capturing them by known

 ISSN: 2454-4248
441 – 448

Microsoft tools. The tool features [43] includes Supports accelerated implementation and heterogeneous environments control in real time, Provides promised extended platform for specific requirements of industries, Supports efficient store, analyze, share data, capture and visualize, and Secure management and connection of data and devices.

The service is strongly a commercial software like Microsoft services, but its production is not ready yet (released preview version on April 2014).

K.Paho Platform

This aims in offering implementation of scalable open source for protocols of standard messaging to facilitate emerging, existing, new applications for enabling usage scenarios of Internet-of-Things and Machine-to-Machine of things. This is an Eclipse foundation part and features have MQTT based, Provides client implementation of MQTT in Python, Java, C and C++, Supports embedded platforms, Provide open API, client implementation and libraries, Enables wide integration range of messaging models, middleware and programming languages, and Enables decoupling of levels between applications and devices.

This tool strives for devices with embedded platforms and constrained networks.

L.ThingSpeak Platform

The open source platform of Internet of Things which offers API to retrieve and store data things from HTTP by using over LAN or via internet [41]. It facilitates sensor creation for logging application, applications of location tracking or social networking things for status updating. The tool was organized in a channel where an application user can store data and retrieve the data. The every channel will support data entries till eight data fields. The channel supports XML, CSV and JSON formats of data in integration application. The tool features [44] Open access API to developers, Data collection in real time, Data gathering of geolocation, Analytic tools and data processing, Status messages of device and plugins integration, Supports languages like JavaScript, Ruby, .net and Java, and Visualization of data on mobile devices and web.

ThingSpeak offers hosted service which is not a version of open source. The Open Sen se [45] is another tool of Internet of Things like features and characteristics of ThingSpeak.

IV. Research Gap

From the prior section, it can be seen that there are various sorts of tools and techniques available at present in order to process the data aggregated from the wireless sensor network. The biggest tradeoff seen in majority of the existing studies are that conventional mining techniques are not found compatible to handle the massive data generated from sensors using TDMA scheduling. The problems of data velocity, data heterogeneity, data volume, and data veracity are still not found to be implemented in the existing data analytical process included in wireless sensor network. The existing tools discussed in this paper are frequently used in the industry as well as in the research studies, but still the research papers are not found with potential benchmarking outcomes as well as evolution of any evidence that proves algorithm to be computationally potential. Hence, in a nutshell, it can be just said that studies towards complex data analysis is still in infancy stage and will require more indepth adoption. The existing data mining techniques also should be further investigated for exploring the optimal feasibility of handling massive data.

The cloud computing has becomes a computational paradigm, which is used in growing of data storage and for processing of todays as well as emerging applications. Many researches were carried out in recent past on problems of data storage and processing. The challenge of establishment of channels of elastic data series from sensors and the jobs running on cloud of data analyzing. The analysis of data must deal with communicational protocols of cloud environment and major sensors. The data process must have data manipulation of sensor and in cloud environment it needs the broken data recovery in connections.

The existing sensors on cloud will give a scalable, extensible, easy to use and interoperable with network reconstruction of sensors. There are many tool sets were developed in order to give data channel uniformly. Various methods were discussed which will take importance in energy consumption on cloud computing. The sensor-cloud architecture is established which enables the data of sensor to be categorized, processed and stored in such a manner which will

become cost effective, easily accessible, timely available. The cloud analysis capacity will be scaled down or up in demand by the use of computing nodes.

V. Conclusion

The paper discusses sensor data over cloud applications and the tools adopted in processing of data. Many issues with sensor data like processing, storage and collection of data and security has been discussed in the paper. The sensor data analytics tools such as IBM mote runner, TempoDB platform, sensor cloud platform, Ostia Portus platform, Nimbits platforms etc. are discussed. These tools are more energy efficient, secure in data management, allows visual management in data interface. Nimbits is a platform of open source for Internet of things. It available easily and provides APIs, libraries and documentations in other programming languages. SensaTrackMonitor is an independent monitoring multi-perform service which deals with Machine to Machine sensor monitoring. IBM Mote Runner was designed to take part between technologies of multiple vendors, where every vendor has individual programming language, database and platforms. It provides language friendly, resource efficient, virtual machine with high performance, which will shields applications (portable) by hardware specifics. Ostia Portus platform was designed to take part between technologies of multiple vendors, where every vendor has individual programming language, database and platforms. Bluwired Sensor-Cloud platform provides an exploration of sensor data, analysis and interaction. FreshTemp Temperature monitor is a cloud based monitoring system used on perishable goods. This automates collection of temperature during the production, storage and transportation of perishable products by offering integrating capability with monitoring different kinds of temperature sensors of such products. TempoDB platform is a database, aims in analyze and storage data of time series from

sensor, servers, automotive telematics and smart meters. MicroStrains of sensor-cloud will offer data storage for sensor, remote management and visualization platform which leverages technologies of cloud computing to provide rapid visualization and data scalability. Xively Platform is the platform which harnesses IoT power to easily and quickly transform the connected vision of product to market reality. Microsoft Azure is aims in secure connect, capture and manage the data generated by machine from assets of Line of Business sensors, industrial devices across platforms of operating systems. ThingSpeak offers hosted service which is not a version of open source. The Open Sen Se is another tool of Internet of Things like features and characteristics of ThingSpeak. Paho Platform: This tool strives for devices with embedded platforms and constrained networks. Thus the paper discusses many advanced tools for sensor data analytics on cloud and its advantages.

Future work needs to focus on addressing the problems encountered in sensor data, thereby contributing towards in-depth study as well as sensor data Analytics research quantity. Extended work will also include validating of the implemented prototype in order to expand the capability of both data collection as well as analysis aspects of the system. Further we implement the intended algorithm using cloud analyst toolkit and java programming. The algorithm will also be compared with PSO algorithm to validate the results with the assumption made by us. Present analysis, real retrieval and storage of sensor data is performed by the model, which implies that every module will communicate with an external storage system prior and after it performs the computation on the sensor data. Hence it is also possible multiple modules can extract the similar data from the storage system. In future version of analysis cloud we aim to avoid the cost of redundant retrieval. Future endeavor will be developing a sensor cloud framework and implementing a remote health-care service on top of the sensor-cloud framework.

REFERENCES

[1] K. Romer and F. Mattern, “The design space of wireless sensor networks,” *IEEE Wireless Communications*, vol. 11, no. 6, pp. 54– 61, 2004.

[2] T. Haenselmann, “Sensor networks,” *GFDL Wireless Sensor Network textbook*, 2006.

[3] G. Simon, G. Balogh, G. Pap et al., “Sensor network-based Counter sniper system,” *Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems (SenSys ’04), Baltimore, Md,USA,* pp. 1–12, 2004.

[4] S. K. Dash, J. P. Sahoo, S. Mohapatra, and S. P. Pati, “Sensorcloud: assimilation of wireless sensor network and the cloud,” *Advances in Computer Science and Information Technology. Networks and Communications*, vol. 84, pp. 455 - 464, Springer- Link, 2012.

[5] C. C. Aggarwal, “Managing and Mining Uncertain Data”, *Springer*,2009.

[6] W.Wang, D. Lu, X. Zhou, B. Zhang and J. Mu, “Statistical wavelet-based anomaly detection in big data with compressive sensing”, *EURASIP Journal on Wireless Communications and Networking*, pp.269, 2013.

[7] K. Grolinger, W.A Higashino, A.Tiwari and M.A.M Capretz, “Data management in cloud environments: NoSQL and NewSQL data stores”, *Springer Journal of Cloud Computing: Advances, Systems and Applications*, vol2, pp.22, 2013.

[8] K. An, S. Tambe, A. Sorbini, S. Mukherjee, J. P-Molina, M. Walker, N. Vermani, A. Gokhale, P. Pazandak, “Real-time Sensor Data Analysis Processing of a Soccer Game Using OMG DDS Publish/Subscribe Middleware”, *Technical report Institute for Software Integrated Systems Vanderbilt University Nashville*, pp. 13-102, June, 2013.

[9] J. Melchor, M. Fukuda, “A Design of Flexible Data Channels for Sensor-Cloud Integration", *UW Provost International Grants, Faculty-led Program, Computing & Software Systems University of Washington Bothell*, 2011

[10] M.M Najafabadi, F.Villanustre, T. M. Khoshgoftaar, N.Seliya, R. Wald and E. Muharemagi, “Deep learning applications and challenges in big data analytics”, *SpringerJournal of Big Data*, 2015.

[11] A.Chakravorty, T. Wlodarczyk, C. Rong, "Privacy Preserving Data Analytics for Smart Homes", *IEEE Security and Privacy rkshops*, 2013.

[12] D. D. Rowlands, J. R. Ride, M. W. McCarthy, L. Laakso, D.A. James, “Linking Sensor Data to a Cloud-based Storage Server Using a martphone Bridge", *SENSORCOMM: The Sixth International Conference on Sensor Technologies and Applications*, 2012.

[13] A.Goyal and N.S. Chahal, “A Proposed Approach for Efficient Energy Utilization in Cloud Data Center", *International Journal of Computer Applications*, vol.115, pp. 11, April 2015.

[14] V. C. Emeakaroha, K. Fatema, P. Healy and J.P. Morrison, "Towards a Generic Cloud-based Sensor Data Management Platform: A Survey and Conceptual Architecture", *SENSORCOMM: The Eighth International Conference on Sensor Technologies and Applications*, 2015.

[15] A. Alamri, W.S. Ansari, M. M. Hassan, M. S. Hossain, A. Alelaiwi, and M. A. Hossain, "A Survey on Sensor-Cloud: Architecture, Applications, and Approaches", *Hindawi Publishing Corporation International Journal of Distributed Sensor Networks*, 2013.

[16] K. Ebner, T. Bühnen, N. Urbach, " Think Big with Big Data: Identifying Suitable Big Data Strategies in Corporate Environments", *47th Hawaii International Conference on System Science*, 2014.

[17] J.S.V.D. Veen, B.V.D.Waaij, M. Vonder, M.D. Jonge, E.Lazovik and R.J. Meijer, "Analysis Cloud Running Sensor Data Analysis Programs on a Cloud Computing Infrastructure", *University of Amsterdam, Amsterdam, The Netherlands*, 2015.

[18] V. N. Inukollu , S. Arsi and S.R.Ravuri, "Security Issues Associated with Big Data in Cloud Computing", *International Journal of Network Security & Its Applications (IJNSA)*, vol.6, pp. No.3, May 2014.

[19] D. J. Abadi, "Data Management in the Cloud: Limitations and Opportunities”, *IEEE Computer Society Technical Committee on Data Engineering*, 2009.

[20] Y. Lim and J. Park, "Sensor Resource Sharing Approaches in Sensor-Cloud Infrastructure", *Hindawi Publishing Corporation International Journal of Distributed Sensor Networks*, 2014.

[21] T. Skov and R. Bro, "A new approach for modelling sensor based data", *Elsevier, Department of Food Science (IFV), The Royal Veterinary and Agricultural University (KVL), Rolighedsvej*, 2004.

[22] N. Mukherjee, S. S. Bhunia and P. Sil, "A Sensor-Cloud Framework for Provisioning Remote Health-Care Services", *Department of Computer Science and Engg. Jadavpur University*,2014

[23] C. C. Aggarwal, “Managing and Mining Sensor Data”, *springer*, 2013.

[24] A. Mahmood, K. Shi, S. Khatoon, and M. Xiao, "Data mining techniques for wireless sensor networks: A survey", International Journal of Distributed Sensor Networks, 2013

[25] J.Melchor and M. Fukuda," A Design of Flexible Data Channels for Sensor-Cloud Integration", *IEEE,* 2011.

[26] J. S. V.D. Veen, B.V.D. Waaij, M. Vonder, M.D. Jonge, E. Lazovik and R. J. Meijer, "Analysis CloudRunning Sensor Data Analysis Programs on a Cloud Computing Infrastructure", 2013

[27] V.C. Emeakaroha, KanizFatema, Philip Healy, John P. Morrison, "Contemporary Analysis and Architecture for a Generic Cloud-based Sensor Data Management Platform", Sensors & Transducers, Vol. 185, Issue 2, pp. 100-112, February 2015.

[28] A.Alamri, W. S. Ansari, M. M. Hassan, M. S. Hossain, A. Alelaiwi, and M. Anwar Hossain, "A Survey on Sensor-Cloud: Architecture, *Applications, and Approaches", Hindawi Publishing Corporation International Journal of Distributed Sensor Networks*, 2013.

[29] IBM, "Ibm mote runner for wireless sensor platform," http://www.zurich.ibm.com/moterunner/ [retrieved: 10-08-2015].

[30] A. Caracas, "Energy-efficiency through micro-managing communication and optimizing sleep," *Sensor, Mesh and Ad Hoc Communications and Networks (SECON), 2011 8th Annual IEEE Communications Society Conference* , pp. 55–63, June 2011.

[31] A. Caracas, T. Kramp, M. Baentsch, M. Oestreicher, T. Eirich, and I. Romanov, "Mote runner: A multi-language virtual machine for small embedded devices," Sensor *Technologies and Applications, SENSORCOMM '09. Third International Conference*, pp. 117–125, June 2009.

[32] TempoDB, "The Time Series Database Service," https://tempo-db.com/ [retrieved: 10-08-2015].

[33] Zhang, Jia et al., "Sensor data as a service–a federated platform for mobile data-centric service development and sharing," in Services Computing (SCC), 2013 IEEE International Conference,, pp. 446–453, 2013.

[34] TempoIQ, "Sensor Analytics for the Measured World," https://www. tempoiq.com/?r=1/[retrieved: 10-08-2015].

[35] MicroStrain, "Microstrainsensorcloud platform," http://www.sensorcloud.com/ [retrieved: 10-08-2015].

[36] V. K. Sehgal, A. Patrick, and L. Rajpoot, "A comparative study of cyber physical cloud, cloud of sensors and internet of things: Their ideology, similarities and differences," *Advance Computing Conference (IACC), IEEE International*, pp. 708–716, 2014.

[37]Ostia,"PortusPlatform,"http://www.ostiasolutions.com/index.php/product/platform-overview/ [retrieved: 10-08-2015].

[38] Bluwird, "Bluwired S-CLOUD Platform," http://bluwired.com/Pages/Discover/BluwiredCloud.aspx [retrieved: 10-08-2015].

[39] FreshTemp, "Cloud-based temperature monitoring tool," https://freshtemp.com/ [retrieved: 10-08-2015].

[40] SensaTrack, "Online sensor monitoring platform," http://www.sensatrack.com/index.html/ [retrieved: 10-08-2015].

[41] C. Doukas and I. Maglogiannis, "Bringing IoT and cloud computing towards pervasive healthcare," *in Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS)*, pp. 922–926, 2012

[42] A. A. Chandra, Y. Lee, B. M. Kim, S. Y. Maeng, S. H. Park, and S. R. Lee, "Review on sensor cloud and its integration with arduino based sensor network," *in IT Convergence and Security (ICITCS), 2013 International Conference*, pp. 1–4, 2013.

[43] Microsoft, "Microsoft azure intelligent system service," http://www.microsoft.com/windowsembedded/en-us/intelligent-systems-service. aspx [retrieved: 10-08-2015].

[44] ThingSpeak, "ThingSpeak Platform," https://thingspeak.com/ [retrieved: 10-08-2015].

[45] R. Piyare, "Integrating wireless sensor network into cloud services for real-time data collection," *ICT Convergence (ICTC), International Conference*, pp. 752–756, 2013.