Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 6.017

IJCSMC, Vol. 6, Issue. 10, October 2017, pg.26 – 34

A SURVEY ON IoT PERFORMANCES IN BIG DATA

M.Praveen Kumar[1], S.P Santhoshkumar[2], T.Gowdhaman[3], S.Syed Shajahaan[4]
#1 Assistant Professor/Information Technology, Rathinam Technical Campus, Coimbatore
#2 Assistant Professor/Computer Science & Engineering, Rathinam Technical Campus, Coimbatore
#3 Assistant Professor/Computer Science & Engineering, Rathinam Technical Campus, Coimbatore
#4 Head of the Department of CSE& IT, Rathinam Technical Campus, Coimbatore
#1 santhosh.cse@rathinam.in, #2 praveen.it@rathinam.in,
#3 gowdhaman.cse@rathinam.in, #4 aortc@rathinamtechzone.com

Abstract - The Internet of things (IoT) is the inter-networking of physical devices, vehicles (also referred to as "connected devices" and "smart devices"), buildings, and other items embedded with electronics, software, sensors, actuators, and network connectivity which enable these objects to collect and exchange data. [1][2][3] The IoT allows objects to be sensed or controlled remotely across existing network infrastructure, [4] creating opportunities for more direct integration of the physical world into computer-based systems, and resulting in improved efficiency, accuracy and economic benefit in addition to reduced human intervention. The upcoming IoT will be greatly presented by the enormous quantity of heterogeneous networked embedded devices that generate intensively "Big data". Enormously a large amount of data is being collected today by many organizations and in a continuous raise. It turns out to be computationally inefficient to analyze such a massive data. The quantity of the available raw data has been expanding on an exponential scale. In a massive database, the valuable information is hidden. The new developed Big data techniques can handle many challenges that face data analysis and have the ability to extract valuable information. This survey shows the study of IoT and Big data. The survey discusses Big data on IoT and how it is created. Many IoT existing, future application and a variety of IoT technologies whether wired or wireless are viewed. Challenges and techniques that solve these issues are discussed and the architecture of IoT is observed.

Index Terms - Big Data, Internet of Things (IoT), Heterogeneous data, IoT architecture, IoT applications.

I. INTRODUCTION

The Internet of Things may be a hot topic in the industry but it's not a new concept. In the early 2000's, Kevin Ashton was laying the groundwork for what would become the Internet of Things (IoT) at MIT's AutoID lab. Ashton was one of the pioneers who conceived this notion as he searched for ways that Proctor & Gamble could improve its business by linking RFID information to the Internet. The concept was simple but powerful. If all objects in daily life were equipped with identifiers and wireless connectivity, these objects could be communicate with each other and be managed by computers. At the time, this vision required major technology improvements. After all, how would we connect everything on the planet? What type of wireless communications could be built into devices? What changes would need to be made to the existing Internet infrastructure to support billions of new devices communicating? What would power these devices? What must be developed to make the solutions cost effective? There were more questions than answers to the IoT concepts in 1999. Today, many of these obstacles have been solved. The size and cost of wireless radios has dropped tremendously. IPv6 allows us to assign a communications address to billions of devices. Electronics companies are building Wi-Fi and cellular wireless connectivity into a wide range of devices. ABI Research estimates over five billion wireless chips will ship in 2013. Mobile data coverage has improved significantly with many networks offering broadband speeds. While not perfect, battery technology has improved and solar recharging has been built into numerous devices. There will be billions of objects connecting to the network with the next several years. For example, Cisco's Internet of Things Group (IOTG) predicts there will be over 50 billion connected devices by 2020. The vast existence of varieties of things, such as sensors, actuators, and mobile phones, result in the great presence of the IoT notion. Behind all expectations, IoT enhances the living standards. The advantages related to link sensor data or networking between sensors is widely deployed in many fields. It contains environmental monitoring, disaster management, human motions, health, smart cities, and understanding social phenomenon [7]. IoT generates numerous amount of data therefore called "Big data," that provides advanced analytic techniques and offers a vision that makes machine usage easier and efficient. The Big data analysis is required to take advantage of its potential for high-level modeling and knowledge engineering. The possibility of the data flow from physical resources to future Internet facilities is what we need to an analysis by using Big data analytic techniques. The Big data challenge is how to understand the interaction between human and smart objects. The basis of the Internet was human to human interaction when the human determines the content to be used by another human, but with the IoT the objects determine the content. Therefore, the impact on our lives is an open issue that needs understanding how IoT plays an important role in a smart environment and smart world [8].The rest of this paper is organized as follows. Section II introduces creating knowledge and Big data. Section III discusses different types of applications of IoT. Section VI shows the technologies of IoT. Section V discusses techniques and methodologies. Section VI shows IoT Architecture. Section VII explains the previously related work. Finally, challenges and future Directions of IoT are discussed.

II. CONCEPTS OF BIG DATA

The notion of the Big data is related to the computer science since the earliest computing days. The data volume that goes beyond the processing capacity of the usual database and cannot be handled by traditional database techniques is called, "Big data." However, if we have a large amount of data, it requires different approaches like techniques, tools, and architecture with the aim to solve new problems or old problem in better ways [9]. A report released by Gartner [10] says entering into a connected devices world; IoT is estimated to accelerate and reach to 26 billion connected devices by 2020. It has been observed four main Big data challenges (the four Vs.):

- Expanding data volume.
- Expanding velocity of data as in/out and change of data.
- Expanding variety of data types and structures.
- Expanding data veracity.

The fifth V is suggested as the value [11], which is the contribution of Big data and able to make decisions. IoT will rapidly increase the volume, variety, and velocity of data. Therefore, enterprises begin to hold on current Big data challenges. As usual, solving the problems of data storage, integration, and IoT analytics is the responsibility of IT. In IoT, a huge amount of raw data is collected on an ongoing basis. Therefore, it is essential to develop new techniques able to transform raw data into valuable knowledge. For instance, in the medical domain, significant actions done by human- like eating, drinking, breathing, and signs can be detected by transforming meaningful sensor raw streams to it. The collected data will be intensive. It is expected to be an enormous amount of sensor data streams. These streams of data may be used in different ways for different purposes. Therefore, the resources of data and how it was processed must be known, and the privacy and security must be provided [1]. The coming IoT will be greatly presented by the enormous quantity of heterogeneous networked embedded devices that produce intensive or "Big data". The collected Big data may not have any value unless analyzing, interpretation, and understanding. The data mining techniques are the mainly recommended methods to be used in extracting knowledge from raw data [1], [6].

III. APPLICATIONS OF IoT

3.1. Smart Home:

With IoT creating the buzz, 'Smart Home' is the most searched IoT associated feature on Google. But, what is a Smart Home? Wouldn't you love if you could switch on air conditioning before reaching home or switch off lights even after you have left home? Or unlock the doors to friends for temporary access even when you are not at home. Don't be surprised with IoT taking shape companies are building products to make your life simpler and convenient. Smart Home has become the revolutionary ladder of success in the residential spaces and it is predicted Smart homes will become as common as smartphones.

The cost of owning a house is the biggest expense in a homeowner's life. Smart Home products are promised to save time, energy and money. With Smart home companies like Nest, Ecobee, Ring and August, to name a few, will become household brands and are planning to deliver a never seen before experience.

3.2. Wearables:

Wearables have experienced a explosive demand in markets all over the world. Companies like Google, Samsung have invested heavily in building such devices. But, how do they work?

Wearable devices are installed with sensors and softwares which collect data and information about the users. This data is later pre-processed to extract essential insights about user. These devices broadly cover fitness, health and entertainment requirements. The pre-requisite from internet of things technology for wearable applications is to be highly energy efficient or ultra-low power and small sized.

3.3. Smart Cities:

Smart city is another powerful application of IoT generating curiosity among world's population. Smart surveillance, automated transportation, smarter energy management systems, water distribution, urban security and environmental monitoring all are examples of internet of things applications for smart cities. IoT will solve major problems faced by the people living in cities like pollution, traffic congestion and shortage of energy supplies etc. Products like cellular communication enabled Smart Belly trash will send alerts to municipal services when a bin needs to be emptied. By installing sensors and using web applications, citizens can find free available parking slots across the city. Also, the sensors can detect meter tampering issues, general malfunctions and any installation issues in the electricity system.

3. 4. IoT in agriculture:

With the continuous increase in world's population, demand for food supply is extremely raised. Governments are helping farmers to use advanced techniques and research to increase food production. Smart farming is one of the fastest growing field in IoT. Farmers are using meaningful insights from the data to yield better return on investment. Sensing for soil moisture and nutrients, controlling water usage for plant growth and determining custom fertilizer are some simple uses of IoT.

3.5. IOT in Healthcare:

Connected healthcare yet remains the sleeping giant of the Internet of Things applications. The concept of connected healthcare system and smart medical devices bears enormous potential not just for companies, but also for the well-being of people in general. Research shows IoT in healthcare will be massive in coming years. IoT in healthcare is aimed at empowering people to live healthier life by wearing connected devices. The collected data will help in personalized analysis of an individual's health and provide tailor made strategies to combat illness.

IV. TECHNOLOGIES

The IoT involves devices to acquire technology from the physical world and transform them into data. Technologies in IoT can be divided into data acquisition and network acquisition technologies. **-** Data Acquisition Technologies Because of the fast growth in computer hardware, software, the Internet, and sensors, the mobile communications have developed to enhance services. Additionally, they have extended into new application areas, with better services and features with minor costs.

A. Two-dimensional code:

It is a barcode that represent the data that the machine can read it. The one-dimensional code can read characters and numbers only and cannot read Chinese letters and images. The two- dimensional code was evolved to solve the one-dimensional code issues. The two-dimensional code handles black and white pixels that are represented on a 2D plane to save information. In the two-dimensional code, (0) expresses white, and (1) expresses black. The advantage of the two-dimensional code is the ability to express a variety of information as sounds, images, texts, and numbers. The Two-dimensional bar code was evolved by algorithms. The usual structure of images is monochrome BMP that result in the minimum volume in bytes [18]. b) RFID Technology RFID technology can read remote source at a long distance. The identification code related to a tag so that, the resulting tag code can be sent to one or more readers. RFID involves of data communication between devices' readers and RFID tags, and it is a standard technology that can be used by many constructors. Therefore, accurate standards occur to confirm appropriate implementation. These standards are the EPCGlobal UHF v.1.2.0 and ISO 18000-6C. Many bank cards and roll tags are using passive tags [18], [19].

- Digital sensors a digital sensor that is an automated or electrochemical sensor, wherever conversion and transmission of data are completed digitally. The requirement of digital measurement and wireless transmission manage remote PC-based sensor diagnostics, tracking, and analysis. Applying the massive operating capacity, result in hard analog electronics requirements enhancement [22].
- The biosensor is an analytical device used for determining analytic that include biological components and physicochemical detector. Biosensors depend on screen printed, so it is used for a large extent construction. Despite the biosensors composed from bio-components, many challenges face biosensors as long response time, short constancy, and poor generation [23].

a) Zigbee Zigbee is a wireless network technology constructed for little sensor degree. This protocol includes the network layer, the infrastructure layer, and the application layer. These layers are defined in its concepts [18].

b) Z-Wave Z-Wave is a wireless interconnection technology that authorizes spread from a management entity to one or more entities in the network. It consists in its architecture of the network, infrastructure, and application layers[18].

c) 6LoWPAN Low-power Wireless Personal Area Networks (LoWPANs) are wireless networks that constituted of a vast quantity of low-cost devices. They are measured with similar wireless networks. LoWPAN includes challenges as small packet sizes, low bandwidth, low power, large volumes of devices, battery groove, and unreliability from radio connectivity challenges. When they are merged with the Internet Protocol (IP), the constraints of LoWPAN are suited. So, there is 6LoWPAN [18].

V. TECHNIQUES AND METHODOLOGIES

There are several techniques and tools for solving many IoT data management challenges like Big data, cloud computing, semantic sensor web, data fusion techniques, and middleware. • Big Data Analytics and Tools There are many techniques or methodologies that can solve IoT data processing and analytics issues in many concepts, fig.1 showed the Apache Hadoop ecosystem.

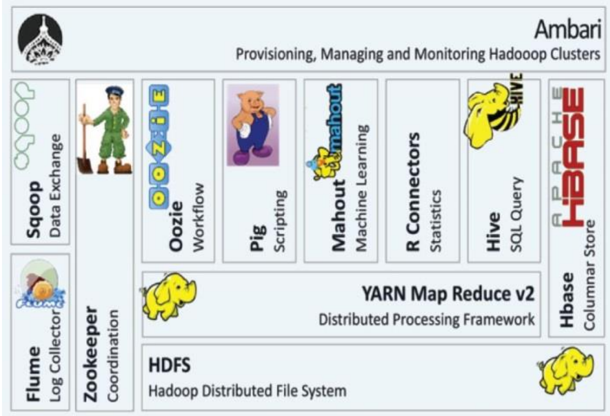


Figure. 1: Apache Hadoop ecosystem.

A. *Hadoop:*

Hadoop is an open source mission that managed by the Apache Software Foundation. Big data can be collected and handled by Hadoop. Hadoop is proposed to parallelize data processing through computing nodes to hurry computations and hide latency. There are two main components for Hadoop: Hadoop Distributed File System (HDFS) and Map Reduce engine. HDFS stores enormous data constantly set and reproduce it to the user application at high bandwidth. MapReduce is a framework that is used for processing massive data sets in a distributed fashion through numerous machines [24].

B. *Map Reduce*:

MapReduce was constructed as a broad programming paradigm. Some of the original employments offered all the key needs of parallel execution, fault tolerance, load balancing, and data manipulation. The Map Reduce named with this name because it includes two abilities from existing functional computer languages: map and reduce. The MapReduce framework gathers all sets with the common key from all records and joins them together. Therefore, it acquires forming one group for each one of the different produced keys. MapReduce is one of the new technologies, but it is just an algorithm, a technique for how to fit all the data. To acquire the best from MapReduce, we need more than just an algorithm. We need a collection of products and technologies created to manage the challenges of Big data [24], [25].

C. *HBase*: HBase:

It is a database model inside the Hadoop framework that looks like the original system of Big Table. The HBase has a column that operates as the key and is the only index that can be used to get back the rows. The data in HBase is also saved as (key, value) sets, where the subject in the non-key columns can be represented by the values [19].

D. *Hive*:

The already deployed tools for data warehousing are not able to be suitable especially in the situation wherever, data is accessible everywhere; they are costly and often privately-operated. Such as the notion like MapReduce is there, it requests for the ability to write job procedures. Map Reduce jobs are difficult to track the characteristics of reusable code as some jobs are business particular some of the time. Hive may be thought as the necessary portion of Hadoop system and views at the top that principally is the organization for the data warehouse. Hive cannot treat with applications and transactions of the real time those are achieved online. The motivation behind it is a complicated technique [26].

E. *Pig*:

The Pig implementation designed within the Hadoop framework to offer additional database as functionality. A table in Pig is a group of tuples, and every field is a value or a set of tuples. So, this framework permits for nested tables, which is a great notion. Pig also provides a scripting language called PigLatin that offers all the common concepts of SQL, such as projections, joins, sorting, and grouping. PigLatin differs from SQL as scripts are procedural and are simple for programmers to be understood. The PigLatin language offers a higher extraction level to the MapReduce framework, as a query in PigLatin may be converted into a sequence of MapReduce tasks [19].

F. *Mahout*:

Mahout is mainly built on an Apache open-source library which able to be scaled and managed for the massive volume of data. These segments rely on three significant machine learning missions that Mahout presently operates.

- Collaborative filtering
- Clustering
- Categorization/Classification [26].

G. *NoSQL*:

It is an abbreviation to Not only SQL, and the most usual notion for non-relational databases. These databases are appealed to operate better than SQL databases. Various types of NoSQL databases, which are keyvalue pair document, column-oriented, and graph databases, that permit programmers to display the data suitable to the structure of their used applications. Because of the growth of the Internet usability and the accessibility of low-cost storage, a massive quantity of structured, semi-structured and unstructured data are acquired and saved for different types of applications. This data is usually denoted to as Big data. Google, Facebook, Amazon, and several other enterprises use NoSQL databases [27], [28].

BigTable:

A Big Table development is initiated in 2004 and is now used by a much of Google applications, such as MapReduce. It is often used for producing and altering data stored in BigTable, Google Reader, Google Maps, Google Book Search, Google Earth, Blogger.com, Google Code hosting, Orkut, YouTube, and Gmail. Google's motivation for evolving its specific database contain scalability, and better control of performance features. BigTable is augmented for data read processes, by distributed data storage management model, which is based on column storage to enhance data retrieving effectiveness. The main components of BigTable are a row, column, record tablet, and timestamp. Amongst them, there cord tablet is a link to the set of row.

- Semantic Sensor Web The quantity of existing sensors will be enormous, and the gathered data will be intensive. If we have the ability to put the collections of data into a homogeneous and heterogeneous form, then the interoperability problems of understanding the data will rely on the semantic technologies to process the data. There are many aspects of semantic sensor Web as: Ontology Ontology is the core of any semantic technology as semantic sensor Web. It is a tool for knowledge allocations and usage. Semantic Ontologies can be divided into some formats as OWL and RDF [19], [27].

- OWL: OWL stands for Web Ontology Language. It defines discrete data substitution format. The great benefit of this ontology format is that there is no limitation to represent constraints as domain or range constraints.
- RDF: RDF is an abbreviation for description research framework. It is a research description language. This language determines the way that resources can interconnect with each other and perform interpretations [19].

In the subsequent subsections, some IoT middleware proposals are listed:

1. *UBIWARE:* It is an agent-based middleware that characterizes each source as a software agent. An agent is responsible for supervising the state of the source, and supporting the interoperation of the source with other elements. The core notions of UBIWARE is to permit automatic discovery, orchestration, choreography, invocation and execution of different Business Intelligence services.
2. *Hydra:* The Hydra middleware involves of a service-oriented architecture. It depends on Web services to support the resource discovery, description, and access that relies on XML and Web protocols. Hydra network uses a proxy to connect the restricted devices to it. The two principle tasks achieved by Hydra developers are
   i. Integrating non-Hydra devices and
   ii. Connecting Hydra-enabled devices to a network.
3. *Link Smart Middleware:* The Link Smart middleware deployed in the Hydra project permits the integration of heterogeneous physical devices into applications via a Web service interface for directing any physical device irrespective of its network technology, such as Bluetooth, RF, ZigBee, RFID, and Wi-Fi. Link Smart relies on a semantic model-driven architecture and permits the use of devices as services both by embedding services in devices and by proxy services for devices. The semantic description of devices relies on ontologies using OWL, OWL-s and SAWSDL.

4. Open IoT: The Open IoT project offers an open-source middleware platform. It allows the development of IoT applications rendering to a utility cloud computing delivery model. Open IoT role is to recognize the idea of on request access to IoT services obtained over clouds of Internet-connected objects, the called sensing as a service, offering a "cloud-of-things".

VI. IoT ARCHITECTURE

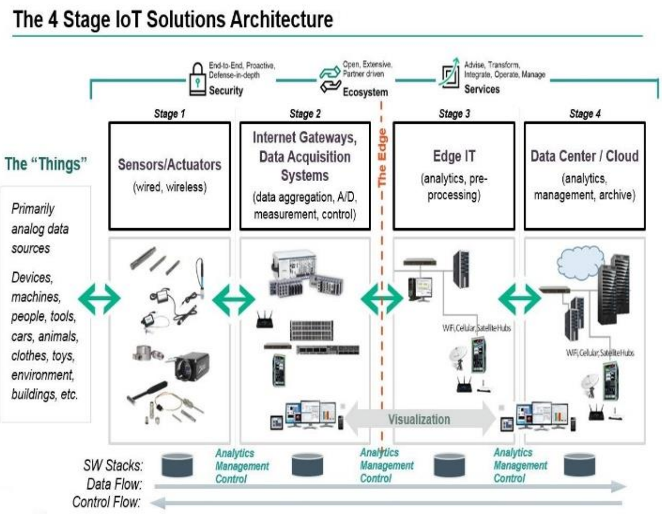
There are many ways to begin an enterprise or industrial Internet of Things (IoT) journey. What's important is not to let the perceived complexity of the IoT obscure the possibilities for implementing what should be very rewarding projects. You can get started by laying a solid framework for your IoT system.

The IoT is more than Internet-connected consumer devices. Sooner or later, your IT organization will need you to create an infrastructure to support it. Energy companies already use networked sensors to measure vibrations in turbines. They feed that data through the network to computing systems that analyze it to predict when machines will need maintenance and when they will fail. Jet engine manufacturers embed sensors that measure temperature, pressure, and other conditions to improve their products. Even a gift basket business can deploy sensors to constantly monitor the temperature of perishable products. If temperatures in storage or in transit start to rise, they can expedite deliveries. This has the dual advantage of increasing customer satisfaction while avoiding product spoilage.

The IoT opens far more opportunities than most organizations are pursuing today. And even when they do pursue them, IT leaders don’t always fully engage in IoT-related requests. Industry discussions of IoT initiatives often emphasize complexity and involve areas of the business in which IT hasn’t traditionally been involved. The proliferation of unfamiliar IoT technologies—

some of them industry-specific—and the vast scope of the global IoT itself can intimidate IT and obscure the possibilities for what should be very achievable and compelling projects. So, how can your infrastructure support IoT?

Stage 1 of an IoT architecture consists of your networked things, typically wireless sensors and actuators. Stage 2 includes sensor data aggregation systems and analog-to-digital data conversion. In Stage 3, edge IT systems perform preprocessing of the data before it moves on to the data center or cloud. Finally, in Stage 4, the data is analyzed, managed, and stored on traditional back-end data center systems. Clearly, the sensor/actuator state is the province of operations technology (OT) professionals. So is Stage 2. Stages 3 and 4 are typically controlled by IT, although the location of edge IT processing may be at a remote site or nearer to the data center. The dashed vertical line labeled "the edge" is the traditional demarcation between OT and IT responsibilities, although this is blurring. Here's a look at each in detail.



Stage 1. Sensors/actuators

Sensors collect data from the environment or object under measurement and turn it into useful data. Think of the specialized structures in your cell phone that detect the directional pull of gravity—and the phone's relative position to the “thing” we call the earth—and convert it into data that your phone can use to orient the device. Actuators can also intervene to change the physical conditions that generate the data. An actuator might, for example, shut off a power supply, adjust an air flow valve, or move a robotic gripper in an assembly process.

The sensing/actuating stage covers everything from legacy industrial devices to robotic camera systems, water-level detectors, air quality sensors, accelerometers, and heart rate monitors. And the scope of the IoT is expanding rapidly, thanks in part to low-power wireless sensor network technologies and Power over Ethernet, which enable devices on a wired LAN to operate without the need for an A/C power source.

Stage 2. The Internet gateway

The data from the sensors starts in analog form. That data needs to be aggregated and converted into digital streams for further processing downstream. Data acquisition systems (DAS) perform these data aggregation and conversion functions. The DAS connects to the sensor network, aggregates outputs, and performs the analog-to-digital conversion. The Internet gateway receives the aggregated and digitized data and routes it over Wi-Fi, wired LANs, or the Internet, to Stage 3 systems for further processing.

Gateways are still edge devices—they’re external to the data center—so geography and location matter. In the pump example, if you have 100 pump units and want to process data on-premises, you might have instant data at the pump level, aggregate the information to create a plantwide view for the facility, and pass the data on to the data center for companywide view. DAS and gateway devices may end up in a wide variety of environments, from the factory floor to mobile field stations, so these systems are usually designed to be portable, easy to deploy, and rugged enough to withstand variations in temperature, humidity, dust, and vibration.

Stage 3. Edge IT

Once IoT data has been digitized and aggregated, it's ready to cross into the realm of IT. However, the data may require further processing before it enters the data center. This is where edge IT systems, which perform more analysis, come into play. Edge IT processing systems may be located in remote offices or other edge locations, but generally these sit in the facility or location where the sensors reside closer to the sensors, such as in a wiring closet.

Stage 4. The data center and cloud

Data that needs more in-depth processing, and where feedback doesn't have to be immediate, gets forwarded to physical data center or cloud-based systems, where more powerful IT systems can analyze, manage, and securely store the data. It takes longer to get results when you wait until data reaches Stage 4, but you can execute a more in-depth analysis, as well as combine your sensor data with data from other sources for deeper insights. Stage 4 processing may take place on-premises, in the cloud, or in a hybrid cloud system, but the type of processing executed in this stage remains the same, regardless of the platform.

VII. RELATED WORK

Big data in IoT is considered as an important research topic and would be divided as follows:

A. Big Data Techniques for IoT

Many researchers that use Big data techniques to solve some IoT challenges. For example, Mesiti and Valtolina proposed a framework that able to gather data from different sources with different forms like JSON, XML, textual, and data streaming from sensors. The data collections lead the database to be unstructured and require data integration. As the world moves to develop Big data analysis techniques, they reached a solution that can be load data from heterogeneous sensors then integrate that heterogeneous sensor data using NoSQL systems. Finally, they designed a user-friendly load system for NoSQL systems by determining a plan to select appropriate NoSQL system that allow the conceptual schema to be deployed. They are continuing operations of handling schemes, and they intend to begin the implementation. In this paper, they did not observe a solution in the case of newly added data that ensure system availability and stability.

Ding et al. offered a general statistical database cluster mechanism for Big data analysis in the IoT (IoT statistic DB). One of the major problems that face IoT is how to transform sensor data into knowledge. Statistical analysis on sensor sampling data is one of the most important procedures in IoT systems. Four statistical analyzing methods were offered which, include the Euclidean-based spatial aggregation, the Network based spatial aggregation, the Euclidean-based parameter aggregation, and the Network-based parameter aggregation. The parallel processing techniques for the statistical queries are proposed, so that multiple servers can apply the statistical analysis in parallel and the performance can be enriched. They intend to discuss event detections and data mining techniques depending on the statistical analysis of IoT. It is not marked how to treat with anomaly presence.

Hayes and Capretz proposed an algorithm for anomaly detection in Big Sensor Data. In particular, an algorithm of contextual anomaly detection is suggested to progress a point anomaly detection algorithm. In this paper, they planned to create a contextual anomaly detection technique for the usage in sensor networks of streams. The technique uses a usual content anomaly detection algorithm for anomaly detection. In addition to this technique, they added a post-processing context-aware anomaly detection algorithm based on a multivariate clustering algorithm. The authors proposed a MapReduce methodology to outline the sensor profiles used in the context detector. They suggested additional modules added to this work such as, for example, a semantic anomaly detector. The anomaly detection abilities of the proposed algorithm should be enhanced.

Fazio et al. designed framework for the high heterogeneity data. They presented a new architecture able to make a dual abstraction of complex sensing infrastructures along with data they collect. There are varieties of heterogeneous smart sensors that interact with each other. Thus, integration is essential. The benefit of this work is to provide a service at a worldwide level that is scalable and flexible. The deployment of this framework depends on the Bayesian algorithm in addition to Contiki Operating System. They aim to deploy advanced applications for aggregation and filtering of data. Bayesian is a less performance algorithm, so the algorithm that has been used should be enhanced.

VIII. CHALLENGES OF IoT

The Internet of Things (IoT) is taking the world by storm. The millions of connected sensors and smart devices that are being deployed on a daily basis in homes, offices, cities and even on our persons are creating unprecedented opportunities in cutting costs, reducing energy consumption, improving efficiency and customer services, and better understanding how we interact with our environment. The IoT, hailed as one of the biggest breakthroughs in the history of the tech industry, will soon be an inherent part of every aspect of our lives, from retail shops to hotels, to cars and airplanes and practically everything we interact with. But this added utility comes with its own set of caveats and requirements, which need to be met and overcome with the proper solutions and approaches. Here are four challenges to expect in the future.

Security Challenges

IoT has already turned into a serious security concern that has drawn the attention of prominent tech firms and government agencies across the world. The hacking of baby monitors, smart fridges, Barbie dolls, drug infusion pumps, cameras and even assault rifles are portending a security nightmare being caused by the future of IoT. So many new nodes being added to networks and the internet will provide malicious actors with innumerable attack vectors and possibilities to carry out their evil deeds, especially since a considerable number of them suffer from security holes.

The more important shift in security will come from the fact that IoT will become more ingrained in our lives. Concerns will no longer be limited to the protection of sensitive information and assets. Our very lives and health can become the target of IoT hack attacks, as was shown in the hacking of pacemakers. Critical city infrastructure can also become a target, as the Ukraine power grid hack warned us last year.

There are many reasons behind the state of insecurity in IoT. Some of it has to do with the industry being in its "gold rush" state, where every vendor is hastily seeking to dish out the next innovative connected gadget before competitors do. Under such circumstances, functionality becomes the main focus and security takes a back seat.

Also, many IoT developers often come from an embedded systems programming background and are ignorant of the threats of IoT programming. They don't necessarily have the knowhow and expertise to program for the hostile connected environment of the internet, and end up dishing out code that is reliable from a functionality perspective, but can easily be exploited remotely.

Some of the data that IoT devices collect are very sensitive and are protected by legislations such as the Health Insurance Portability and Accountability Act (HIPAA) in the U.S. and are fundamentally different from our browsing and clicking habits. Yet the necessary precautions aren't taken when storing the data or sharing it with other service providers. Vendors and manufacturers must either discard this data or remove the Personally Identifiable Information (PII) to make sure that consumers aren't damaged in case of data breaches.

Another consideration to take is that while data generated about a single appliance (such as a smart toaster) might not be sensitive per-se, yet when combined with data from other devices, it can reveal information such as the consumer's life pattern, which can become very damaging if they fall into the hands of the wrong people. In many cases, criminals don't even need to pry into your encrypted communications in order to obtain the information they want. A study by LGS Innovations elaborates on this issue and presents a DIY solution to protect IoT traffic and privacy.

Connectivity Challenges

This model is sufficient for current IoT ecosystems, where tens, hundreds or even thousands of devices are involved. But when networks grow to join billions and hundreds of billions of devices, centralized brokered systems will turn into a bottleneck. Such systems will require huge investments and spending in maintaining cloud servers that can handle such large amounts of information exchange, and entire systems can go down if the server becomes unavailable.

The future of IoT will very much have to depend on decentralizing IoT networks. Part of it can become possible by moving functionality to the edge, such as using fog computing models where smart devices such as IoT hubs take charge of time-critical operations and cloud servers take on data gathering and analytical responsibilities.

Other solutions involve the use of peer-to-peer communications, where devices identify and authenticate each other directly and exchange information without the involvement of a broker. Networks will be created in meshes with no single point of failure. This model will have its own set of challenges, especially from a security perspective, but these challenges can be met with some of the emerging IoT technologies such as the Phantom protocol, or leveraging the success of other tried and tested models such as the block chain.

IX. CONCLUSION

The IoT denotes to spreading the Internet of physical objects as a room, table, or another human sensing objects as collections of features. They can be detected, determined, and accessed by devices like actuators, sensors or other smart devices. As the vast increasing of existing devices, sensors, actuators and network communications, a massive amount of data has been generated. There are many problems result in the increasing of data volume as massive, heterogeneous, noisy data, privacy, and security. Applications of IoT have been presented. Technologies have been surveyed from the perspective of data acquisition and network based. Finally, challenges and future direction have been discussed. We intend to find new techniques and tools to solve Massive–Heterogeneous issues that are found in related work.

REFERENCES

[1] Stankovic, John A, "Research Directions for the Internet of Things," IEEE Internet Things J. VOL. 1, NO.1, 2014, pp.3-9.
[2] Cecchinel, Cyril, Matthieu Jimenez, S´ ebastien Mosser, Michel Riveill, "An Architecture To Support The Collection Of Big Data In The Internet Of Things," IEEE 10Th World Congress On Services, 2014, pp. 442-449.
[3] Journal, R. (2015). That 'Internet of Things' Thing - RFID Journal. [online] Rfidjournal.com. Available at: http://www.rfidjournal.com/articles/view?4986 (Accessed on 21/4/2015).
[4] Atzori, Luigi, Antonio Iera, Giacomo Morabito, "The Internet of Things: A Survey," Computer Networks 54, 2010, pp. 2787–2805.
[5] Li, Shancang, Li Da XU, and Shanshan Zhao, "The Internet of Things: A Survey," Springer Science Business Media New York, 2014, pp.243-259.
[6] Guoru Ding, Long Wang, Qihui Wu, "Big Data Analytics In Future Internet Of Things," National Natural Science Foundation Of China, 2013, pp. 1-6.

[7] Sulayman K sowe, Takashi Kimata, Mianxiong Dong, Koji Zettsu, "'Managing Heterogeneous Sensor Data On A Big Data Platform: Iot Services For Data-Intensive Science," IEEE 38Th Annual International Computers, Software And Applications Conference Workshops, 2014, pp. 259-300.

[8] Antonio J. Jara, Dominique Genoud,Yann Bocchi,"Big Data In Smart Cities: From Poisson To Human Dynamics," 28Th International Conference On Advanced Information Networking And Applications Workshops, 2014, pp. 785-790.

[9] Stephen Kaisler, Frank Armour, J. Alberto Espinosa, William Money, "Big Data: Issues And Challenges Moving Forward', 46Th Hawaii International Conference On System Sciences, 2013, pp. 994- 1003.

[10] Gartner, 'Big Data Management and Analytics', 2015. [Online]. Available: http://www.gartner.com/technology/topics/big-data.jsp. (Accessed on 2/11/2015).

[11] InformationWeek, (2015). When Internet Of Things Meets Big Data - InformationWeek. [online] Available at: http://www.informationweek.com/big-data/big-data-analytics/when-internet-of-things-meets-big-data/a/d-id/1298137 (Accessed on 5/12/2014).

[12] Lu Tan, Neng Wang, "Future Internet: The Internet of Things," 3Rd International Conference On Advanced Computer Theory And Engineering (ICACTE), 2010, pp. 376- 380.

[13] Shashank Agrawal, Dario Vieira., "A Survey On Internet Of Things," Abakós, Belo Horizonte VOL. 1, NO. 2, 2013, pp. 78 – 95.

[14] Innova.com.tr, (2015). Innova deploys smart home and smart office services for TTNET | Innova. [online] Available at: http://www.innova.com.tr/en/news-detail.asp?haber=777C888F-C922-8822-EEE2-BBB222FF333C (Accessed on 10/1/2015).

[15] Smart Health, (2015). Walking Fit. [online] Available at: http://smarthealthusa.com/walking-fit/ (Accessed on10/1/2015).

[16] Bidnessetc.com, (2014). Internet of Things: Connecting (And Disrupting) Your Universe. [online] Available at: http://www.bidnessetc.com/business/internet-of-things-connecting-and-disrupting-your-universe/ (Accessed on 4/12/2014).

[17] Smartpharma.co.za, (2015). Welcome to Smart Pharmaceuticals. [online] Available at: http://www.smartpharma.co.za/ (Accessed on 10/1/2015).

[18] Jasper Tan, Simon G. M. Koo., "A Survey of Technologies in Internet of Things," IEEE International Conference on Distributed Computing in Sensor Systems, 2014, pp.269-274.

[19] Charu C. Aggarwal, Naveen Ashish, Amit Sheth., "The Internet Of Things: A Survey From The Data-Centric Perespective, Managing And Mining Sensor Data," Springer Science+Business Media New York, 2013, pp. 384-428.

[20] Jayavardhana Gubbi, Rajkumar Buyya, Slaven Marusic, M. Palaniswami.,"Internet Of Things (Iot): A Vision, Architectural Elements, And Future Directions," Future Generation Computer Systems 29, 2013, pp. 1645-1660.

[21] Khaled Salama, Abbas El Gamal, "Analysis Of Active Pixel Sensor Readout Circuit," IEEE Trans. Circuits Syst. I: Fundamental Theory and Applications, VOL. 50, NO. 7, 2003, pp. 941-944.

[22] Abhisek Ukil, "Towards Networked Smart Digital Sensors: A Review," IEEE, ABB Corporate Research, Switzerland, 2008, pp. 17981802.

[23] CRC Press, (2015). Biosensors in Food Processing, Safety, and Quality Control. [online] Available at: http://www.crcpress.com/product/isbn/9781439819852#googlePreviewContainer (Accessed on 21/4/2015).

[24] Kamalpreet Singh, Ravinder Kaur, "Hadoop: Addressing Challenges Of Big Data," IEEE International Advance Computing Conference (IACC), 2014, pp. 686-689.

[25] Cloudera.com, (2015). Introduction to Hadoop and MapReduce. [online]Availableat:http://www.cloudera.com/content/cloudera/en/training/courses/udacity/mapreduce.html (Accessed on 16/1/2015).

