Smruti Ranjan Pattnaik *et al*, International Journal of Computer Science and Mobile Applications,
National Conference on "The Things Services and Applications of Internet of Things",
Gandhi Institute for Education and Technology (GIET) Baniatangi, 23-24 March 2018, pg. 122-128
ISSN: 2321-8363
Impact Factor: 5.515

The Impact Internet of Things (IoT) on Big Data Implementation

Smruti Ranjan Pattnaik[1], Syamanta Garud[2], Ranjita Rout[3]

[1]*Research Scholar, Centurion University BBSR, smrutiranjan.pattanaik82@gmail.com*
[2]*Research Scalar, Centurion University, syamantagarud@rediffmail.com*
[3]*Asst. Professor, GIET, Bhubaneswar, rout.ranjita@gmail.com*

Abstract
The IoT will massively increase the amount of data available for analysis by all manner of organisations. However, there are significant barriers to overcome before the potential benefits are fully realized. The growth in the number of devices connected to the Internet of Things (IoT) and the exponential increase in data consumption only reflect how the growth of big data perfectly overlaps with that of IoT. The team of big data in a continuously expanding network gives rise to non-trivial concerns regarding data collection efficiency, data processing, analytics, and security. To address these concerns, researchers have examined the challenges associated with the successful deployment of IoT. Despite the large number of studies on big data, analytics, and IoT, the convergence of these areas creates several opportunities for flourishing big data and analytics for IoT systems. In this paper, we explore the recent advances in big data analytics for IoT systems as well as the key requirements for managing big data and for enabling analytics in an IoT environment. We identify s the role of big data analytics in IoT applications. Finally, several open challenges are presented as future research directions.
***Keywords**:* efficiency, data processing, analytics, and security, big data, IoT

I. Introduction
The Internet of Things (IoT) sounds like a consumer fantasy come true -- who wouldn't want to be able to turn off the lights at home from two towns away, or leave it to their refrigerator to make sure they know when milk, butter and other staples need to be replenished? But there's more to the IoT than lifestyle enhancement. It also includes a corporate side, enabling organizations to collect and analyze data from sensors on manufacturing equipment, pipelines, weather stations, smart meters, delivery trucks and other types of machinery.

IoT analytics applications can help companies understand the Internet of Things data at their disposal, with an eye toward reducing maintenance costs, avoiding equipment failures and improving business operations. In addition, retailers, restaurant chains and makers of consumer goods can use data from smart phones, wearable technologies and in-home devices to do targeted marketing and promotions -- the business side of the IoT's futuristic world of connected consumer gear.

Prior to the emergence of the IoT, taking a step back to analyze all of the information provided by the assortment of devices it can encompass was exceedingly difficult, if not outright impossible. As David Smith, a principal program manager at Microsoft, noted in September 2014, "Getting access to data is something data scientists strive for. But all these devices are independent, and there's no way for anybody to aggregate that data together." IoT technology offers automated mechanisms for pulling machine data into data warehouses or Hadoop clusters and other big data platforms for analysis.

Building and running the kinds of big data analytics applications typically required with IoT data isn't a simple task, though. If your organization is looking to make sense of the data it collects from the Internet of Things, check out the expert insights, user stories and other resources in this IoT analytics Essential Guide.

ISSN: 2321-8363
Impact Factor: 5.515

II. ROLE OF DATA ANALYTICS IN INTERNET OF THINGS (IoT)

Internet of Things (IoT) is an interconnection of several devices, networks, technologies and human resources to achieve a common goal. There is a variety of IoT based applications that are being used in different sectors and have succeeded in providing huge benefits to the users.

Data Analytics (DA) is defined as a process which is used to examine big and small data sets with varying data properties to extract meaningful conclusions from these data sets. These conclusions are usually in the form of trends, patterns, and statistics that aid business organizations in effective decision-making processes.

Data Analytics has a significant role to play in the growth and success of IoT applications and investments. Analytics tools will allow the business units to make effective use of their datasets as explained in the points listed below.

- **Volume:** There are huge clusters of data sets that IoT applications make use of. The business organizations need to manage these large volumes of data and need to analyze the same for extracting relevant patterns. These datasets along with real-time data can be analyzed easily and efficiently with data analytics software.
- **Structure:** IoT applications involve data sets that may have a varied structure as unstructured, semi-structured and structured data sets. There may also be a significant difference in the data formats and types. Data analytics will allow the business executive to analyze all of these varying sets of data using automated tools and software.
- **Driving Revenue:** The use of data analytics in IoT investments will allow the business units to gain an insight into customer preferences and choices. This would lead to the development of services and offers as per the customer demands and expectations. This, in turn, will improve the revenues and profits earned by the organizations.
- **Competitive Edge:** IoT is a buzzword in the current era of technology and there are numerous IoT application developers and providers present in the market. The use of data analytics in IoT investments will provide a business unit to offer better services and will, therefore, provide the ability to gain a competitive edge in the market.
- There are different types of data analytics that can be used and applied in the IoT investments to gain advantages. Some of these types have been listed and described below.
- **Streaming Analytics:** This form of data analytics is also referred as event stream processing and it analyzes huge in-motion data sets. Real-time data streams are analyzed in this process to detect urgent situations and immediate actions. IoT applications based on financial transactions, air fleet tracking, traffic analysis etc. can benefit from this method.
- **Spatial Analytics:** This is the data analytics method that is used to analyze geographic patterns to determine the spatial relationship between the physical objects. Location-based IoT applications, such as smart parking applications can benefit from this form of data analytics.
- **Time Series Analytics:** As the name suggests, this form of data analytics is based upon the time-based data which is analyzed to reveal associated trends and patterns. IoT applications, such as weather forecasting applications and health monitoring systems can benefit from this form of data analytics method.
- **Prescriptive Analysis:** This form of data analytics is the combination of descriptive and predictive analysis. It is applied to understand the best steps of action that can be taken in a particular situation. Commercial IoT applications can make use of this form of data analytics to gain better conclusions.

There have been scenarios where in IoT investments have immensely benefitted from the application and the use of data analytics. With the change and advancement in technology, there are emerging areas in which data analytics can be applied in association with IoT. For instance, actionable marketing can be carried out by applying data analytics to the product usage. IoT analytics will also allow the increased safety and surveillance abilities through video sensors and application of data analytics methods.

Healthcare is one of the prime sectors of every country and the utilization of data analytics in IoT based healthcare applications can provide breakthroughs in this area. The reduction of the healthcare costs, enhancement of telehealth monitoring, and remote health services, increased diagnosis and treatment can be achieved using the same.

The utilization of data analytics shall, therefore, be promoted in the area of IoT to gain improved revenues, competitive gain, and customer engagement.

ISSN: 2321-8363
Impact Factor: 5.515

III. IOT 'THINGS'

A huge number of 'things' *could* join the IoT, whose recent rise to prominence is the result of several trends conspiring to cause a tipping point: low-cost, low-power sensor technology; widespread wireless connectivity; huge amounts of available and affordable (largely cloud-based) storage and compute power; and plenty of internet addresses to go round, courtesy of the IPv6 protocol (2^128 addresses, versus 2^32 for IPv4).
Estimates and projections of the current and future number of internet-connected objects vary, depending on the definitions used and the optimism of whoever is doing the estimating and projecting. The best-known figures come from Cisco, which puts the current (number at around 14.8 billion and the expected number in 2020 at around 50 billion (and that's just 2.77 percent of an estimated 1.8 trillion potentially connectable 'things'):

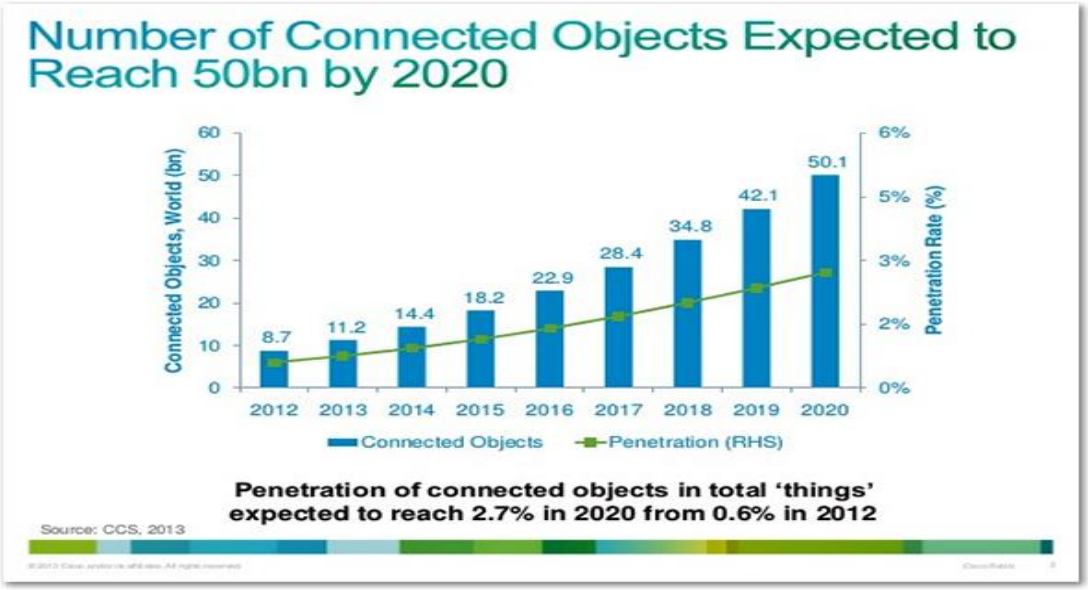

Image: Cisco

EMC and IDC are somewhat more conservative, putting the 2020 IoT population at 32 billion, while Gartner comes in with 26 billion.

BIG DATA TECHNOLOGIES

When selecting the technology stack for big data processing, the tremendous influx of data that the IoT will deliver must be kept in mind. Organizations will have to adapt technologies to map with IoT data. Network, disk, and compute power all will be impacted and should be planned to take care of this new type of data. From a technology perspective, the most important thing is to receive events from IoT connected devices. The devices can be connected to the network using Wi-Fi, Bluetooth, or another technology, but must be able to send messages to a broker using some well-defined protocol. One of the most popular and widely used protocols is Message Queue Telemetry Transport (MQTT). Mosquitto is a popular open-source MQTT broker. Once the data is received, the next consideration is the technology platform to store the IoT data. Many companies use Hadoop and Hive to store big data. But for IoT data, No SQL document databases like Apache Couch DB are more suitable because they offer high throughput and very low latency. These types of databases are schema-less, which supports the flexibility to add new event types easily. Other popular IoT tools are Apache Kafka for intermediate message brokering and Apache Storm for real-time stream processing.

Smruti Ranjan Pattnaik *et al*, International Journal of Computer Science and Mobile Applications,
National Conference on "The Things Services and Applications of Internet of Things",
Gandhi Institute for Education and Technology (GIET) Baniatangi, 23-24 March 2018, pg. 122-128
ISSN: 2321-8363
Impact Factor: 5.515

IOT AND BIG DATA

The IoT and big data are clearly growing apace, and are set to transform many areas of business and everyday life. In its 2015 Internet of Things predictions, IDC notes that: "Today, over 50% of IoT activity is centered in manufacturing, transportation, smart city, and consumer applications, but within five years all industries will have rolled out IoT initiatives".
In its 2014 Digital Universe report, EMC and IDC see the IoT creating new business opportunities in five main areas, summarized in this slide:

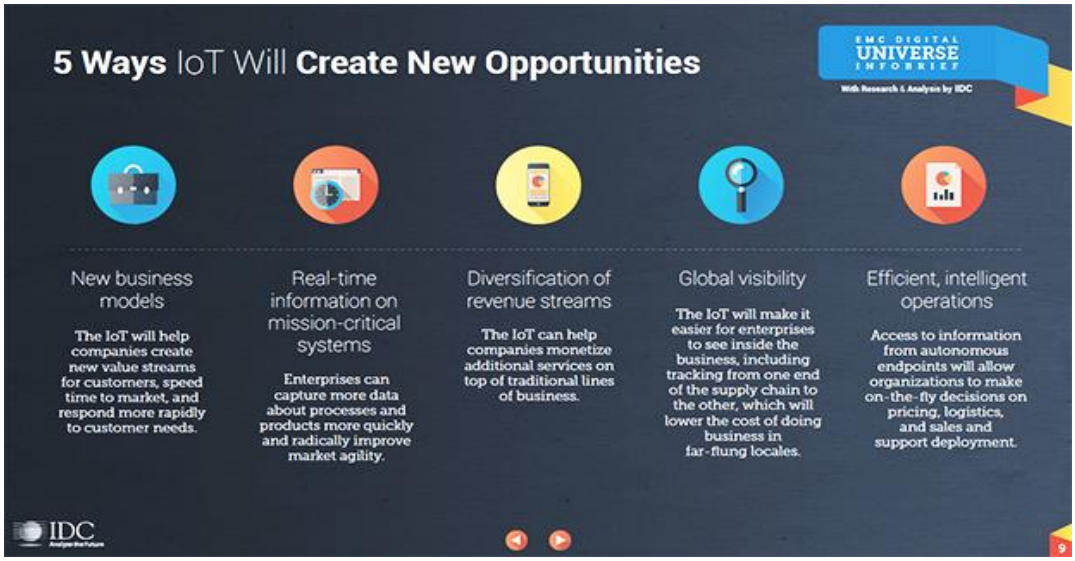

Image: EMC/IDC

To deliver on these opportunities, according to EMC's Bill Schmarzo, a new generation of IoT applications is required to address specific business needs such as: predictive maintenance; loss prevention; asset utilization; inventory tracking; disaster planning and recovery; downtime minimization; energy usage optimization; device performance effectiveness; network performance management; capacity utilization; capacity planning; demand forecasting; pricing optimization; yield management; and load balancing optimization.
If these and other nuts and bolts of the IoT/big data revolution can be put in place, there's a great deal of economic value to play for. In the private sector, Cisco expects the value to be driven in five main areas: asset utilisation ($2.5 trillion); employee productivity ($2.5 trillion); supply chain and logistics ($2.7 trillion); customer experience ($3.7 trillion); and innovation, including reducing time to market ($3.0 trillion). In the public sector the main proposed drivers are: employee productivity ($1.8 trillion); connected militarized defense ($1.5 trillion); cost reductions ($740 billion); citizen experience ($412 billion); and increased revenue ($125 billion).
As usual when it comes to the IoE/IoT, Cisco's economic value predictions err on the optimistic side; other analysts are generally more conservative -- although the numbers bandied about are still huge.

CHALLENGES TO WIDESPREAD IOT/BIG DATA VALUE DELIVERY

Before the IoT/big data nexus can deliver on its promise, there are a number of barriers to overcome. The main ones are summarized below.

ISSN: 2321-8363
Impact Factor: 5.515

Standards

For the IoT to work, there must be a framework within which devices and applications can exchange data securely over wired or wireless networks. One player in this area is One M2M, an umbrella organization including seven standards bodies, five global ICT for a hand over 200 companies (mostly from the telecoms and IT industries).

Not surprisingly, given the scope and potential value of the IoT market, there are plenty of other standards bodies vying to get their ideas adopted. These include: the All Seen Alliance, Google's The Physical Web, the Industrial Internet Consortium, the Open Interconnect Consortium and Thread.

Security & privacy

According to IDC, "Within two years, 90% of all IT networks will have an IoT-based security breach, although many will be considered 'inconveniences' .Chief Information Security Officers (CISOs) will be forced to adopt new IoT policies". Progress on data standards but there's no doubt that security and privacy is a big worry with the IoT and big data -- particularly when it comes to areas like healthcare or critical national infrastructure.

Network & data centre infrastructure

The vast amounts of data that will be generated by IoT devices will put enormous pressure on network and data centre infrastructure. IoT data flows will be primarily from sensors to applications, and will range between continuous and bursty depending on the type of application:

WAYS THE INTERNET OF THINGS WILL CHANGE BIG DATA

Before the Internet of Things (IoT) came along—billions of networked sensors and devices capable of generating enormous amounts of new, unstructured real-time data—big data was already really, *really* big. To tackle this, businesses small and large have taken to the cloud and reworked their IT architectures to create more flexible, scalable ways to manage their data.

However, for those businesses and data scientists looking to capitalize on the high-value, target-rich data the IoT will be churning out over the next decade, there will be even more to consider when it comes to data architecture.

IoT data presents a number of challenges for companies. Here are a few ways companies can meet these challenges head-on.

1. **MORE DATA MEANS COMPANIES WILL HAVE TO RETHINK THEIR IT AND DATA CENTER INFRASTRUCTURES**

For all of its potential, effective IoT data analytics will hinge on **better IT infrastructures**—data centers, server clusters, cloud-based computing, and more. Businesses that want to leverage IoT data will need to invest in long-term IT architecture planning. Because this new influx of data from sensors and devices will put more pressure on existing networks and data centers and require more power to process it. Before data experts can even begin applying analytics, data needs to be aggregated and organized—and this will be no small feat.

Whether it's a consumer company gathering data from wearable and mobile devices, or enterprise organizations processing data from industrial sensors and manufacturing equipment, upgrades will be inevitable. Services like Hadoop, with its distributed server clusters and parallel processing, will be important, as will the people who know how to set it up and work with its more tricky aspects.

Data centers themselves will most likely lean toward a more distributed approach, with tiered mini centers that pull data, then send it on to be processed further in second- and third-tier clusters. Obviously, this approach will have an impact on data storage, bandwidth, and backup.

ISSN: 2321-8363
Impact Factor: 5.515

2. WITH THE IOT, QUALITY DATA WILL BE ACTIONABLE DATA.

Finding the information that's actionable and capable of creating real, meaningful change. More isn't always more, and many companies collecting automated data from sensors will likely have more data than they know what to do with.
Complex estimations aside, the 20+ billion devices predicted to be around by 2020 are going to have an inevitable effect on the three V's of big data: volume**,** velocity**,** and variety**.** More, faster, and less structured data will be pouring in from sensored devices. IoT data is unique in that it's only really valuable to us if it's actionable, and that percentage of the massive–and totally *new* streams of data coming in–will be a bit easier to manage. Sifting through this data will be the job of business analysts who know what questions they want their data to answer, and of data scientists who know how to get those answers.

3. NOSQL DATABASES WILL MOST LIKELY OUTPACE TRADITIONAL RDBMSS.

Much of this IoT data will be unstructured, meaning it can't be easily sorted into tables like a relational database management system (RDBMS). No SQL databases like Couch base, Cassandra, and MongoDB will be able to offer IoT data scientists the flexibility they need to organize data in a way that makes the data usable.
More data means we'll need more places to aggregate the data, and more power to process it—often in real-time scenarios. Microsoft Azure, Cloud era, Amazon, and Apache's cloud-based computing platform Hadoop, with its Hive and Pig components and Spark processing engines, are all poised to take on this surge of new IoT data.

4. BEYOND COLLECTING DATA, BUSINESSES NEED TO CHOOSE A SOFTWARE STACK FOR PREPROCESSING AND ANALYZING IOT DATA.

Once this massive amount of data is collected and organized, businesses need to have the right plan and software stack in place to analyze it. Carefully choosing a stack of software and databases will ensure the system can handle the types and the scale of the data anticipated.
First, because much of this data will be raw and unstandardized, it needs to be transformed and preprocessed with tools like Hadoop's Pig component, then stored in a database. Analytics tools like Apache Storm, which is especially suited for the continuous streams of real-time data the IoT will generate, should be put in place for analytics. The overall analytics solution should be strategic specifically for IoT data, its speed, and its volume.

5. MORE SKILLED—DATA ANALYSTS TO MAKE IOT DATA VALUABLE.

Companies will need to have the right people in place to analyze and make all of this structured, unstructured, or semi-structured data into valuable business insights.
To make the most of your data, you'll need skilled business analysts who know what they're looking for from the data, what questions to ask of it, and how that data will translate into value for the company. Then, it's up to the data scientist to do the looking, answer those questions, and deliver that value, through a combination of.

ISSN: 2321-8363
Impact Factor: 5.515

IV. CONCLUSION

In day-to-day life the IoT and big data are clearly growing in the activities of manufacturing, transportation, smart city, consumer applications and health care. A brief introduction of IoT and Big data and their roles in different technologies with their wide spread of use in this generation, has been explained in this paper. The challenges to widespread big data are
Standards, security and privacy, network and data structure infrastructure. This paper also focuses on the ways to overcome the issues based on infrastructure, quality of data, choosing software stack, analyzing data and skilled manpower. If the above issues are suitable resolved, accurate information will be disseminated, which will lead to a conducive environment for organizational development and the progress of the society at large.

REFERENCES:

1. Ravi Gorli , "Interlinking OF IoT, Big data, Smart Mobile app with Smart Garbage Monitoring", International Journal of Computer Sciences and Engineering, Volume- 05, Issue-01, Page No (70-74), Jan -2017, EISSN: 2347-2693
2. J. P. Conti, " The Internet of Things" , Communications Engineer, 2006, pp.20-25.
3. https://www.edgefx.in/different-types-ofsensors-with-applications/
4. http://searchbusinessanalytics.techtarget.com/ essentialguide/IoT-analytics-guideUnderstanding-Internet-of-Things-data
5. http://www.datamation.com/applications/whybig-data-and-the-internet-of-things-are-aperfect-match.html
6. http://www.zdnet.com/article/the-internet-ofthings-and-big-data-unlocking-the-power
7. https://www.analytixlabs.co.in/blog/2016/10/0 8/iot-and-big-data-they-work-in-tandem/
8. http://data-informed.com/the-impact-ofinternet-of-things-on-big-data/
9. https://www.linkedin.com/pulse/iot-big-datascientists-bill-mccabe
10. http://www.sas.com/en_us/insights/articles/big-data/big-data-and-iot-two-sides-of-thesame-coin.html
11. http://www.zdnet.com/article/the-internet-ofthings-and-big-data-unlocking-the-power/

