IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 45-53

Big Data with reference to IoT: Architecture, Opportunities and Challenges

Ranjita Rout[1], Rupak Kumar Mohanayak[2]
[1] *Dept. Of Computer Science & Engineering, E-mail address: ranjitarout@gietbbsr.com*
[2] *Dept. Of Computer Science & Engineering, E-mail address: rkmohanayak@gietbbsr.com*

Abstract: *Big amounts of data have been produced, as the miniaturization of Internet of things (IoT) devices increases. However, such data are not useful without analytic power. Several big data, IoT, and analytics solutions have empowered people to obtain valuable insight into large data generated by IoT devices. This paper investigates the contemporary research efforts directed toward big IoT data analytics. The relationship between big data analytics and IoT is explained. We identify the opportunities resulting from the convergence of big data, analytics, and IoT as well as discuss the role of big data analytics in IoT applications. Furthermore, big IoT data analytic types, methods, and technologies for big data mining, challenges, and opportunities are discussed. Several notable use cases are also presented. Several opportunities brought by data analytics in IoT paradigm are then discussed.*
Keywords *:IoT, big data, , analytics, data mining*

I. Introduction

IoT

The Internet of Things is the concept of everyday objects – from industrial machines to wearable devices – using built-in sensors to gather data and take action on that data across a network. So it’s a building that uses sensors to automatically adjust heating and lighting. Or production equipment alerting maintenance personnel to an impending failure. Simply put, the Internet of Things is the future of technology that can make our lives more efficient.

History

The term “Internet of Things” was coined in the late 1990s by entrepreneur Kevin Ashton. Ashton, who’s one of the founders of the Auto-ID Center at MIT, was part of a team that discovered how to link objects to the Internet through an RFID tag. He said he first used the phrase “Internet of Things” in a presentation he made in 1999 – and the term has stuck around ever since.

Importance of Internet of Things

Here are some examples of the impact the IoT has on industries:

- Intelligent transport solutions speed up traffic flows, reduce fuel consumption, prioritize vehicle repair schedules and save lives.
- Smart electric grids more efficiently connect renewable resources; improve system reliability and charge customers based on smaller usage increments.
- Machine monitoring sensors diagnose – and predict – pending maintenance issues, near-term part stockouts, and even prioritize maintenance crew schedules for repair equipment and regional needs.
- Data-driven systems are being built into the infrastructure of "smart cities," making it easier for municipalities to run waste management, law enforcement and other programs more efficiently.
- But also consider the IoT on a more personal level. Connected devices are making their way from business and industry to the mass market. Consider these possibilities:
- You’re low on milk. When you’re on your way home from work, you get an alert from your refrigerator reminding you to stop by the store.
- Your home security system, which already enables you to remotely control your locks and thermostats, can cool down your home and open your windows, based on your preferences.

Area of Use

The IoT is more than just a convenience for consumers. It offers new sources of data and business operating models that can boost productivity in a variety of industries.

Health Care
Many people have already adopted wearable devices to help monitor exercise, sleep and other health habits – and these items are only scratching the surface of how IoT impacts health care. Patient monitoring devices, electronic records and other smart accessories can help save lives.

Manufacturing
This is one of the industries that benefits from IoT the most. Data-collecting sensors embedded in factory machinery or warehouse shelves can communicate problems or track resources in real time, making it easy to work more efficiently and keep costs down.

Retail
Both consumers and stores can benefit from IoT. Stores, for example, might use IoT for inventory tracking or security purposes. Consumers may end up with personalized shopping experiences through data collected by sensors or cameras.

Telecommunications
The telecommunications industry will be significantly impacted by the IoT since it will be charged with keeping all the data the IoT uses. Smart phones and other personal devices must be able to maintain a reliable connection to the Internet for the IoT to work effectively.

Automotive
From predictive maintenance to multimodal transportation and shared mobility services, bring valuable services to market by combining analytics with IoT data. The IoT also impacts transportation on a larger scale: delivery companies can track their fleet using GPS solutions. And roadways can be monitored via sensors to keep them as safe as possible.

Utilities
Smart meters not only collect data automatically, they make it possible to apply analytics that can track and manage energy use. Likewise, sensors in devices such as windmills can track data and use predictive modeling to schedule downtime for more efficient energy use.

Big Data
The term big data existed long before IoT arrived to carry out analytics. When the information demonstrates veracity, velocity, variety and volume, then it is interpreted as big data. This equates to a large quantity of data that can be both unstructured and structured, while velocity refers to data processing speed and veracity governs its uncertainty.

The volume of data generated by sensors, devices, social media, health care applications, temperature sensors, and various other software applications and digital devices that continuously generate large amounts of structured, unstructured, or semi-structured data is strongly increasing. This massive data generation results in "big data" [17]. Traditional database systems are inefficient when storing, processing, and analyzing rapidly growing amount of data or big data [18]. The term "big data" has been used in the previous literature but is relatively new in business and IT [19]. "The Digital Universe" study [21] labels big data technologies as a new generation of technologies and architectures that aim to take out the value from a massive volume of data with various formats by enabling high-velocity capture, discovery, and analysis. This previous study also characterizes big data into three aspects: (a) data sources, (b) data analytics, and (c) the presentation of the results of the analytics. This definition uses the 3V's (volume, variety, velocity) model proposed by Beyer [2]. The model highlights an e-commerce trend in data management that faces challenges to manage volume or size of data, variety or different sources of data, and velocity or speed of data creation.

Big Data Analytics
Big data analytics involves the processes of searching a database, mining, and analyzing data dedicated to improve company performance [25].
Big data analytics is the process of examining large data sets that contain a variety of data types [4] to reveal unseen patterns, hidden correlations, market trends, customer preferences, and other useful business information [5]. The capability to analyze large amounts of data can help an organization deal with considerable information that can affect the business [6]. Therefore, the main objective of big data analytics is to assist business associations to have improved understanding of data, and thus, make efficient and well-informed

decisions. Big data analytics enables data miners and scientists to analyze a large volume of data that may not be harnessed using traditional tools [5].

Big data analytics require technologies and tools that can transform a large amount of structured, unstructured, and semi-structured data into a more understandable data and metadata format for analytical processes. The algorithms used in these analytical tools must discover patterns, trends, and correlations over a variety of time horizons in the data [26]. After analyzing the data, these tools visualize the findings in tables, graphs, and spatial charts for efficient decision making. Thus, big data analysis is a serious challenge for many applications because of data complexity and the scalability of underlying algorithms that support such processes [27].

Proper IoT data analytics should be growth, right-size infrastructure, and performance focussed. A hybrid approach is best for performance and future growth. A hybrid deployment consists of dedicated hosting, colocation, managed hosting, and the cloud. It provides a single optimal environment by combining the best features from multiple platforms. For performance, the best option would be a single customer dedicated server. When setting up proper analytics infrastructure for IoT data, companies should remember that not all IoT data is important.

II. Relationship BetweenIoT and Big Data Analytics

Big data analytics is rapidly emerging as a key IoT initiative to improve decision making. One of the most prominent features of IoT is its analysis of information about “connected things.” Big data analytics in IoT requires processing a large amount of data on the fly and storing the data in various storage technologies. Given that much of the unstructured data are gathered directly from web-enabled “things,” big data implementations will necessitate performing lightning-fast analytics with large queries to allow organizations to gain rapid insights, make quick decisions, and interact with people and other devices. The interconnection of sensing and actuating devices provide the capability to share information across platforms through a unified architecture and develop a common operating picture for enabling innovative applications.

The need to adopt big data in IoT applications is compelling. These two technologies have already been recognized in the fields of IT and business. Although, the development of big data is already lagging, these technologies are inter-dependent and should be jointly developed. In general, the deployment of IoT increases the amount of data in quantity and category; hence, offering the opportunity for the application and development of big data analytics. Moreover, the application of big data technologies in IoT accelerates the research advances and business models of IoT. The relationship between IoT and big data, can be divided into three steps to enable the management of IoT data. The first step comprises managing IoT data sources, where connected sensors devices use applications to interact with one another. For example, the interaction of devices such as CCTV cameras, smart traffic lights, and smart home devices, generates large amounts of data sources with different formats. This data can be stored in low cost commodity storage on the cloud. In the second step, the generated data are called “big data,” which are based on their volume, velocity, and variety. These huge amounts of data are stored in big data files in shared distributed fault-tolerant databases. The last step applies analytics tools such as MapReduce, Spark, Splunk, and Skytree that can analyze the stored big IoT data sets. The four levels of analytics start from training data, then move on to analytics tools, queries, and reports.

III. Bringing IoT and Big Data together

This disruptive technology needs new infrastructures, including software and hardware applications as well as an OS; enterprises must handle the influx of data that begins flowing in and examine it in real-time as it evolves by the minute.

That is where big data arrives into the picture; big data analytics tools have the capacity to handle large volumes of data generated from IoT devices that create a continuous stream of information.

But, in order to differentiate between them, IoT provides data from which big data analytics can extract information to generate insights required of it.

However, IoT conducts data on a completely different scale, so the analytics solution must accommodate its needs of processing and rapid ingestion followed by a fast and accurate extraction.

There are many solutions available that provide near real-time analytics on large-sized datasets, and necessarily change a full-rack database into a small server that processes up to 100 TB, so small amount of hardware is needed. The analytics database of next-generation leverages GPU technology, thus enabling even more downsizing of the hardware, i.e, 5 TB on a laptop or a big database in the car. This largely helps IoT organizations correlate the evolving number of data sets, which helps them adapt to changing trends and acquire real-time responses, solving the challenge regarding size and compromising on the performance.

There are many examples of big data and IoT working well together to offer analysis and insight. One such example is represented by shipping organizations. They have been utilizing big data analytics and sensor data to improve efficiency, save money and lower their environmental impact. They utilize sensors on their delivery vehicles in order to monitor engine health, number of stops, mileage, miles per gallon, and speed.

IoT and big data are creating waves in big agriculture. In this area, the field connects systems monitors to the moisture levels and transmits this data to farmers over a wireless connection. This data will enable farmers to find out when crops are reaching the optimum moisture levels.

IV. IoT Architecture for Big Data Analytics

The architectural concept of IoT has several definitions based on IoT domain abstraction and identification. It offers a reference model that defines relationships among various IoT verticals, such as, smart traffic, smart home, smart transportation, and smart health. The architecture for big data analytics offers a design for data abstraction. Furthermore, this standard provides a reference architecture that builds upon the reference model. Many IoT architectures are found in the literature [13]. For example, [13]offered an IoT architecture with cloud computing at the center and a model of end-to-end interaction among various stakeholders in a cloud-centric IoT framework for better comparison with the proposed IoT architecture. This architecture is achieved by seamless ubiquitous sensing, data analytics, and information representation with IoT as the unifying architecture. However, the current architecture focuses on IoT with regard to communications. To our knowledge, our proposed architecture, which integrates IoT and big data analytics, has not been studied in the current literature. IoT architecture and big data analytics. In this figure, the sensor layer contains all the sensor devices and the objects, which are connected through a wireless network. This wireless network communication can be RFID, WiFi, ultra-wideband, ZigBee, and Bluetooth. The IoT gateway allows communication of the Internet and various webs. The upper layer concerns big data analytics, where a large amount of data received from sensors are stored in the cloud and accessed through big data analytics applications. These applications contain API management and a dashboard to help in the interaction with the processing engine.

A novel meta-model-based approach for integrating IoT architecture objects is proposed. The concept is semi-automatically federated into a holistic digital enterprise architecture environment. The main objective is to provide an adequate decision support for complex business, architecture management with the development of assessment systems, and IT environment. Thus, architectural decisions for IoT are closely connected with code implementation to allow users to understand the integration of enterprise architecture management with IoT.

V. Impact of IoT in big data

The massive amount of revenue and data that the IoT will generate, its impact will be felt across the entire big data universe, forcing companies to upgrade current tools and processes, and technology to evolve to accommodate this additional data volume and take advantage of the insights all this new data undoubtedly will deliver.

DATA STORAGE:

- When we talk about IoT, one of the first things that comes to mind is a huge, continuous stream of data hitting the data storage.In response to this direct impact on big data storage infrastructure, many organizations are moving toward the Platform as a Service model instead of keeping their own storage infrastructure, which would require continuous expansion to handle the load of big data. PaaS is a cloud-based, managed solution that provides scalability, flexibility, compliance, and a sophisticated architecture to store valuable IoTdata.Cloud storage options include private, public, and hybrid models. If companies have sensitive data or data that is subject to regulatory compliance requirements that require heightened security, a private cloud model might be the best fit. Otherwise, a public or hybrid model can be chosen as storage for IoT data.

OPEN SOURCE:

- The IoT isn't built solely using open source software, but open source plays a key role. Linux serves as the operating system for many connected devices. Open source networking standards make it possible for devices from different vendors to communicate. Some IoT devices are even designed to be hackable by users in a way that extends the open source software concept to include open hardware. In all of these ways, the IoT plays on open source's strengths and brings open source to new frontiers.

BIG DATA:

- The IoT promises to take big data to a new level. IoT devices not only generate huge amounts of information, which can then be fed to data analytics tools. They also rely on data-based logic in order to perform many of their “smart” functions. Take your Nest thermostat, for example. It collects data from your home, then runs analytics based on the data it collected along with external information (like weather reports) to predict when to turn on your furnace.

CYBERSECURITY:

- Security and privacy aren't new concerns. But they are on the minds of consumers now more than ever, thanks to the seemingly never-ending reports of breaches at major organizations. The IoT serves both to feed and to alleviate those concerns. IoT devices raise huge new security challenges, especially when it comes to things like critical infrastructure. But they also offer ways to help keep users more secure by adding extra barriers of defense to data and persons.

VI. Fuse Cases

60° View of the Customer

Many enterprises use big data to build a dashboard application that provides a 360° view of the customer. These dashboards pull together data from a variety of internal and external sources,analyze it and present it to customer service, sales and/or marketing personnel in a way that helps them do their jobs.

For example, imagine the sort of dashboard an insurance company might create with information about its customers. Naturally, it would include demographic data, like customers' names, addresses, household income and family members, as well as sales information about which types of policies the customers hold. It could also pull information from the company's customer relationship management (CRM) solution about the customers' past interactions with the firm and even provide links to transcripts of recent calls, email messages or chat sessions. It might also show which pages of the company website a particular customer had recently visited, providing valuable clues about the reason a customer might be calling. The dashboard could also pull in external information, such as the customer's recent social media posts. Or if an auto insurance customer had agreed to have a tracking device from the company installed, it might even provide details about the customer's current location and recent speed.

All of that information would obviously help prepare company staff to interact with the customer, but the most sophisticated dashboards don't stop there. If it used advanced analytics or machine learning tools, the dashboard take a guess about the reason for a customer call. It could suggest opportunities for cross-selling or upselling customers on products, or if it detects that a customer might be in danger of defecting to a competitor, it might suggest potential discounts that could lower the customer's rate. Some tools can even analyze customers' language to detect their current emotions and suggest appropriate responses to sales or customer service agents.

This might sound far-fetched and futuristic, but many companies today already have systems like this one in place, and they are using them to improve customer satisfaction and increase revenues and margins.

2. Fraud Prevention

For credit card holders, fraud prevention is one of the most familiar use cases for big data. Even before advanced big data analytics became popular, credit card issuers were using rules-based systems to help them flag potentially fraudulent transactions. So, for example, if a credit card were used to rent a car in Hawaii, but the customer lived in Omaha, a customer service agent might call to confirm that the cardholder was on vacation and that someone hadn't stolen the card.

Thanks to big data analytics and machine learning, today's fraud prevention systems are orders of magnitude better at detecting criminal activity and preventing false positives. In the example already mentioned, for instance, a sophisticated fraud prevention system might be able to see that the customer had recently purchased airline tickets, sunscreen and a new swimsuit before the rental car purchase. Based on historical patterns, a predictive analytics or machine learning system would be able to tell that the rental car was thus less likely to be a fraudulent purchase.

But fraud prevention systems can get even more sophisticated than that. According to Experian, fraud tends to be concentrated in certain geographic regions—often near airports, which make it easy for criminals to move stolen goods. However, which zip codes are riskiest tends to change over time. Big data analytics can look at past records of fraudulent transaction and quickly identify changing trends. Credit card companies and retailers can then pay more attention to transactions in zip codes that are emerging as hotbeds for criminal activity.

Credit card issuers are understandably hesitant about disclosing all the advanced analytic techniques that they use to detect and prevent fraud. However, many credit card firms and other consultants offer technology, advice and services to other firms to help them set up systems to stop criminal transactions.

3. Security Intelligence

On the theme of criminal activity, organizations are also using big data analytics to help them thwart hackers and cyberattackers. Operating an enterprise IT department generates an enormous amount of log data. In addition, cyber threat intelligence data is available from external sources, such as law enforcement or security providers. Many organizations or now using big data solutions to help them aggregate and analyze all of this internal and external information to help them prevent, detect and mitigate attacks.

Big data security solutions vary in sophistication and they are sold under a wide variety of names. For example, vendors sell log analytics tools that can detect anomalies in network data, security information and event management (SIEM) tools that offer real-time analysis of security alerts generated by other security software, and user and entity behavior analytics (UEBA) solutions that use analytics and machine learning to detect unusual patterns in device or user activity. Other big data security solutions are labelled as security intelligence offerings or network intelligence offerings.

4. Data Warehouse Offload

One of the easiest — and potentially most cost-effective — ways for organizations to begin using big data tools is to remove some of the burden from their data warehouses. Even among the few organizations that haven't yet started experimenting with big data analytics, it is common to have a data warehouse that facilitates their business intelligence (BI) efforts.

Unfortunately, data warehouse technology tends to be very costly to purchase and run. And as business leaders have begun demanding more reports and insights from their BI teams, the data warehouse solutions haven't always been able to provide the desired performance.

To solve this problem, many enterprises use an open source big data solution like Hadoop to replace or compliment their data warehouses. Hadoop-based solutions often provide much faster performance while reducing licensing fees and other costs.

5. Price Optimization

Both business-to-consumer (B2C) and business-to-business (B2B) enterprises are also using big data analytics to optimize the prices that they charge their customers. For any company, the goal is to set prices so that they maximize their income. If the price is too high, they will sell fewer products, decreasing their net returns. But if the price is too low, they may leave money on the table.

Big data analytics allows companies to see which price points have yielded the best overall results under various historic market conditions. Businesses that are more sophisticated with their pricing analytics may also employ variable or dynamic pricing strategies. They use their big data solutions to segment their customer base and build models that show how much different types of customers will be willing to pay under different circumstances. B2C companies that have attempted this approach have met with mixed results, but it is fairly standard among B2B companies.

8. Social Media Analysis and Response

The flood of posts that flow through social media outlets like Facebook, Twitter, Instagram and others is one of the most obvious examples of big data. Today, companies are expected to monitor what people are saying about them in social media and respond appropriately — and if they do not, they quickly lose customers.
As a result, many enterprises are investing in tools to help them monitor and analyze social platforms in real-time. Sometimes these are standalone social media products, while at other times, they are part of a larger marketing intelligence or big data analytics solution.

9. Preventive Maintenance and Support

Many of the big data use cases mentioned so far relate to retail or financial companies, but businesses in manufacturing, energy, construction, agriculture, transportation and similar sectors of the economy can also benefit from big data. In these examples, some of the biggest benefit might come from using big data to improve equipment maintenance.

As the Industrial Internet of Things (IIoT) begins to become a reality, factories and other facilities that use expensive equipment are deploying sensors that can monitor that equipment and transmit relevant data over the Internet. They then use big data solutions to analyze that information — often in real time — to detect when a problem is about to occur. They can then perform preventive maintenance that may help prevent accidents or costly line shutdowns.

VII. Opportunities

IoT is currently considered one of the most profound transitions in technology. Current IoT provides several data analytics opportunities for big data analytics.

a. E-Commerce

Big IoT data analytics offers well-designed tools to process real-time big data, which produce timely results for decision making. Big IoT data exhibit heterogeneity, increasing volume, and real-time data processing features.

The convergence of big data with IoT brings new challenges and opportunities to build a smart environment. Big IoT data analytics has widespread applications in nearly every industry. However, the main success areas of analytics are in e-commerce, revenue growth, increased customer size, accuracy of sale forecast results, product optimization, risk management, and improved customer segmentation.

b. Smart Cities

Big data collected from smart cities offer new opportunities in which efficiency gains can be achieved through an appropriate analytics platform/infrastructure to analyze big IoT data. Various devices connect to the Internet in a smart environment and share information. Moreover, the cost of storing data has been reduced dramatically after the invention of cloud computing technology. Analysis capabilities have made huge leaps. Thus, the role of big data in a smart city can potentially transform every sector of the economy of a nation. Hadoop with YARN resource manager has offered recent advancement in big data technology to support and handle numerous workloads, real-time processing, and streaming data ingestion.

c. Retail and Logistics

IoT is expected to play a key role as an emerging technology in the area of retail and logistics. In logistics, RFID keeps track of containers, pallets, and crates. In addition, considerable advancements in IoT technologies can facilitate retailers by providing several benefits. However, IoT devices generate large amounts of data on a daily basis. Thus, powerful data analytics enables enterprises to gain insights from the voluminous amounts of data produced through IoT technologies. Applying data analytics to logistic data sets can improve the shipment experience of customers. Moreover, retail companies can earn additional profit by analyzing customer data, which can predict the trends and demands of goods. By looking into customer data, optimizing pricing plans and seasonal promotions can be planned efficiently to maximize profit.

d. Healthcare

Recent years have witnessed tremendous growth in smart health monitoring devices. These devices generate enormous amounts of data. Thus, applying data analytics to data collected from fetal monitors, electrocardiograms, temperature monitors, or blood glucose level monitors can help healthcare specialists efficiently assess the physical conditions of patients. Moreover, data analytics enables healthcare professionals to diagnose serious diseases in their early stages to help save lives. Furthermore, data analytics improves the clinical quality of care and ensures the safety of patients. In addition, physician profile can be reviewed by looking into the history of treatment of patients, which can improve customer satisfaction, acquisition, and retention.

VIII. Challenges And Future Directions

IoT and big data analytics have been extensively accepted by many organizations. However, these technologies are still in their early stages. Several existing research challenges have not yet been addressed. This section presents several challenges in the field of big IoT data analytics.

a. Privacy

Privacy issues arise when a system is compromised to infer or restore personal information using big data analytics tools, although data are generated from anonymous users. With the proliferation of big data analytics technologies used in big IoT data, the privacy issue has become a core problem in the data mining domain. Consequently, most people are reluctant to rely on these systems, which do not provide solid service-level agreement (SLA) conditions regarding user personal information theft or misuse. In fact, the sensitive information of users has to be secured and protected from external interference. Although temporary identification, anonymity, and encryptions provide several ways to enforce data privacy, decisions have to be made with regard to ethical factors, such as what to use, how to use, and why use generated big IoT data [7].

At present, no answer can address these challenges and manage the security and privacy of interconnected devices. However, the following guidelines can overcome these adversities. (a) First, a true open ecosystem with standard APIs is necessary to avoid interoperability and reliability problems. (b) Second, devices must be well protected while communicating with peers. (c) Third, devices should be hardcoded with the best security practices to protect against common security and privacy threats.

b. Data Mining

Data mining methods provide efficient and best-fitting predictive or descriptive solutions for big data that can also be generalized for new data [45]. The evolution of big IoT data and cloud computing platforms has brought the challenges of data exploration and information extraction.

c. Visualization

Visualization is an important entity in big data analytics, particularly when dealing with IoT systems where data are generated enormously. Furthermore, conducting data visualization is difficult because of the large size and high dimension of big data. This situation shows underlying trends and a complete picture of parsed data. Therefore, big data analytics and visualization should work seamlessly to obtain the best results from IoT applications in big data. However, visualization in the case of heterogeneous and diverse data (unstructured, structured, and semi-structured) is a challenging task. Designing visualization solution that is compatible with advanced big data indexing frameworks is a difficult task. Similarly, response time is a desirable factor in big IoT data analytics. Consequently, cloud computing architectures supported with rich GUI facilities can be deployed to obtain better insights into big IoT data trends [86].

d. Integration

Integration refers to having a uniform view of different formats. Data integration provides a single view of the data arriving from different sources and combines the view of data [95]. Data integration includes all processes involved in collecting data from different sources, as well as in storing and providing data with a unified view. For each moment, different forms of data are continuously generated by social media, IoT, and other communication and telecommunication approaches. The produced data can be categorized into three groups: (a) structured data, such as data stored in traditional database systems, including tables with rows and columns; (b) semi-structured, such as HTML, XML, and Json files; and (c) unstructured data, such as videos, audios, and images. Good data offer good information; however, this relationship is only achieved through data integration [96]. Integrating diverse data types is a complex task in merging different systems or applications. Overlapping the same data, increasing performance and scalability, and enabling real-time data access are among the challenges associated with data integration that should be addressed in the future.

IX. Conclusion

As we move towards future, all the things are being automated. The explosion of embedded and connected smart devices, systems, and technologies in our lives has created anopportunity to connect every 'thing' to the internet.The growth rate of data production has increased drastically over the past years with the proliferation of smart and sensor devices. The interaction between IoT and big data is currently at a stage where processing, transforming, and analyzing large amounts of data at a high frequency are necessary. We conducted this survey in the context of big IoT data analytics. First, we explored recent analytics solutions. The relationship between big data analytics and IoT was also discussed. Moreover, we proposed an architecture for big IoT data analytics. Furthermore, big data analytics types, methods, and technologies for big data mining were presented. Some credible use cases were also provided. In addition, we explored the domain by discussing various opportunities brought about by data analytics in the IoT paradigm.

According to Gartner, revenue generated from IoT products and services will exceed $300 billion in 2020, and that probably is just the tip of the iceberg. Given the massive amount of revenue and data that the IoT will generate, its impact will be felt across the entire big data universe, forcing companies to upgrade current tools and processes, and technology to evolve to accommodate this additional data volume and take advantage of the insights all this new data undoubtedly will deliver.

References

[1]. P. Tiainen, ''New opportunities in electrical engineering as a result of the emergence of the Internet of Things,'' Tech. Rep., AaltoDoc, Aalto Univ., 2016.

[2]. M. Beyer, ''Gartner says solving 'Big Data' challenge involves more than just managing volumes of data,'' Tech. Rep., AaltoDoc, Aalto Univ., 2011.

[3]. J. Gantz and D. Reinsel, ''Extracting value from chaos,'' IDC Iview, vol. 1142, pp. 1–12, Jun. 2011. [4] R. Mital, J. Coughlin, and M. Canaday, ''Using big data technologies and analytics to predict sensor anomalies,'' in Proc. Adv. Maui Opt. Space Surveill. Technol. Conf., Sep. 2014, p. 84.

[4]. N. Golchha, ''Big data-the information revolution,'' Int. J. Adv. Res., vol. 1, no. 12, pp. 791–794, 2015.

[5]. P. Russom, Big Data Analytics. TDWI, 4th Quart., 2011.

[6]. C.-W. Tsai, ''Big data analytics: A survey,'' J. Big Data, vol. 2, no. 1, pp. 1–32, 2015.

[7]. M. Chen, Related Technologies in Big Data. Heidelberg, Germany: Springer, 2014, pp. 11–18.

[8]. Z. Khan, A. Anjum, and S. L. Kiani, ''Cloud based big data analytics for smart future cities,'' in Proc. IEEE/ACM 6th Int. Conf. Utility Cloud Comput. (UCC), Dec. 2013, pp. 381–386.

[9]. P. Russom, Big Data Analytics. TDWI, 4th Quart., 2011, pp. 1–35.

[10]. S. LaValle, E. Lesser, R. Shockley, M. S. Hopkins, and N. Kruschwitz, ''Big data, analytics and the path from insights to value,'' MIT Sloan Manag.Rev., vol. 52, no. 2, p. 21, 2011.

[11]. E. Al Nuaimi et al., ''Applications of big data to smart cities,'' J. Internet Services Appl., vol. 6, p. 25, Dec. 2015. [13] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, ''Internet of Things (IoT): A vision, architectural elements, and future directions,'' Future Generat. Comput.Syst., vol. 29, no. 7, pp. 1645–1660, 2013.

[12]. C. A. Ciufo. (2014). Industrial Equipment Talking on the IoT? Better get a Gateway (Device). [Online]. Available: http://eecatalog.com/caciufo/2014/07/15/iot-gateway-adlink/

[13]. L. Atzori, A. Iera, and G. Morabito, ''The Internet of Things: A survey,'' Comput. Netw., vol. 54, no. 15, pp. 2787–2805, 2010.
[14]. H.-C. Hsieh and C.-H. Lai, ''Internet of Things architecture based on integrated PLC and 3G communication networks,'' in Proc. IEEE 17th Int. Conf. Parallel Distrib. Syst. (ICPADS), Dec. 2011, pp. 853–856.
[15]. K. Kambatla, ''Trends in big data analytics,'' J. Parallel Distrib.Comput., vol. 74, no. 7, pp. 2561–2573, 2014.
[16]. J. Manyika, Big Data: The Next Frontier for Innovation, Competition, and Productivity. McKinsey Global Inst. Rep., 2011.
[17]. I. A. T. Hashem et al., ''The rise of 'big data' on cloud computing: Review and open research issues,'' Inf. Syst., vol. 47, pp. 98–115, Jan. 2015. [20] W. B. Ali, ''Big data-driven smart policing: big data-based patrol car dispatching,'' J. Geotech. Transp. Eng., vol. 1, no. 2, pp. 1–16, 2016.
[18]. J. Gantz and D. Reinsel, ''The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east,'' IDC iView: IDC Analyze the Future, Tech. Rep., 2012.
[19]. V. Borkar, M. J. Carey, and C. Li, ''Inside Big Data management: Ogres, onions, or parfaits?'' in Proc. 15th Int. Conf. Extending Database Technol., 2012, pp. 3–14.
[20]. A. Gani, ''A survey on indexing techniques for big data: Taxonomy and performance evaluation,'' Knowl. Inf. Syst., vol. 46, no. 2, pp. 241–284, 2016.
[21]. A. Paul, ''Video search and indexing with reinforcement agent for interactive multimedia services,'' ACM Trans. Embed. Comput.Syst., vol. 12, no. 2, pp. 1–16, 2013.[25] O. Kwon and N. B. L. Shin, ''Data quality management, data usage experience and acquisition intention of big data analytics,'' Int. J. Inf. Manage., vol. 34, no. 3, pp. 387–394, 2014.
[22]. S. Oswal and S. Koul, ''Big data analytic and visualization on mobile devices,'' in Proc. Nat. Conf. New Horizons IT-NCNHIT, 2013, p. 223.
[23]. L. Candela and D. P. C. Pagano, ''Managing big data through hybrid data infrastructures,'' ERCIM News, vol. 89, pp. 37–38, Jun. 2012. [28] M. D. Assuncao, R. N. Calheiros, S. Bianchi, M. A. S. Netto, and R. Buyya. (2013). ''Big data computing and clouds: Challenges, solutions, and future directions.'' [Online]. Available: https://arxiv. org/abs/1312.4722
[24]. D. Singh and C. K. Reddy, ''A survey on platforms for big data analytics,'' J. Big Data, vol. 2, no. 1, p. 1, 2014.
[25]. A. Siddiqa, ''A survey of big data management: Taxonomy and state-ofthe-art,'' J. Netw. Comput. Appl., vol. 71, pp. 151–166, Aug. 2016.
[26]. A. Bifet, G. Holmes, R. Kirkby, and B. Pfahringer, ''MOA: Massive online analysis,'' J. Mach. Learn. Res., vol. 11, pp. 1601–1604, May 2010. [45] A. Mukhopadhyay, U. Maulik, S. Bandyopadhyay, and C. A. C. Coello, ''A survey of multiobjective evolutionary algorithms for data mining: Part I,'' IEEE Trans. Evol. Comput., vol. 18, no. 1, pp. 4–19, Feb. 2014.
[27]. V. Estivill-Castro, ''Why so many clustering algorithms: A position paper,'' ACM SIGKDD Explorations Newslett., vol. 4, no. 1, pp. 65–75, 2002.
[28]. C. Bielza and P. Larrañaga, ''Discrete Bayesian network classifiers: A survey,'' ACM Comput. Surveys, vol. 47, no. 1, p. 5, Jul. 2014.
[29]. F. Chen et al., ''Data mining for the Internet of Things: Literature review and challenges,'' Int. J. Distrib. Sensor Netw., vol. 12, Aug. 2015.
[30]. R. Luss and A. d'Aspremont, ''Predicting abnormal returns from news using text classification,'' Quant. Finance, vol. 15, no. 6, pp. 999–1012, 2015.
[31]. P. Melin and O. Castillo, ''A review on type-2 fuzzy logic applications in clustering, classification and pattern recognition,'' Appl. Soft comput., vol. 21, pp. 568–577, Aug. 2014.
[32]. A. Soualhi, K. Medjaher, and N. Zerhouni, ''Bearing health monitoring based on Hilbert-Huang transform, support vector machine, and regression,'' IEEE Trans. Instrum. Meas., vol. 64, no. 1, pp. 52–62, Jan. 2015.