

Research Article

IoT-Enabled Big Data Analytics Architecture for Multimedia Data Communications

Muhammad Babar ¹, **Mohammad Dahman Alshehri** ², **Muhammad Usman Tariq**,³
Fasee Ullah ⁴, **Atif Khan** ⁵, **M. Irfan Uddin** ⁶, and **Ahmed S. Almasoud**⁷

¹Department of Computer Science, Allama Iqbal Open University, Islamabad, Pakistan

²Department of Computer Science, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

³Abu Dhabi School of Management, Abu Dhabi, UAE

⁴Department of Computer Science and IT, Sarhad University of Science and Information Technology, Peshawar, Pakistan

⁵Department of Computer Science, Islamia College Peshawar, Peshawar 25120, Pakistan

⁶Institute of Computing, Kohat University of Science and Technology, Kohat, Pakistan

⁷Prince Sultan University, College of Computer and Information Sciences, Saudi Arabia

Correspondence should be addressed to Fasee Ullah; fasee.csit@suit.edu.pk

Received 27 September 2021; Accepted 24 November 2021; Published 17 December 2021

Academic Editor: Deepak Gupta

Copyright © 2021 Muhammad Babar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The present spreading out of the Internet of Things (IoT) originated the realization of millions of IoT devices connected to the Internet. With the increase of allied devices, the gigantic multimedia big data (MMBD) vision is also gaining eminence and has been broadly acknowledged. MMBD management offers computation, exploration, storage, and control to resolve the QoS issues for multimedia data communications. However, it becomes challenging for multimedia systems to tackle the diverse multimedia-enabled IoT settings including healthcare, traffic videos, automation, society parking images, and surveillance that produce a massive amount of big multimedia data to be processed and analyzed efficiently. There are several challenges in the existing structural design of the IoT-enabled data management systems to handle MMBD including high-volume storage and processing of data, data heterogeneity due to various multimedia sources, and intelligent decision-making. In this article, an architecture is proposed to process and store MMBD efficiently in an IoT-enabled environment. The proposed architecture is a layered architecture integrated with a parallel and distributed module to accomplish big data analytics for multimedia data. A preprocessing module is also integrated with the proposed architecture to prepare the MMBD and speed up the processing mechanism. The proposed system is realized and experimentally tested using real-time multimedia big data sets from authentic sources that discloses the effectiveness of the proposed architecture.

1. Introduction

Internet of Things (IoT) is one of the latest concepts in the current age. The future of this globe is IoT which will be going to alter today's world objects into intelligent and smart objects [1]. The term IoT was introduced in the late 1990s, yet any other components like semiconductors and wireless networks exist for quite a long time [2]. The Internet of Things is comprised of hardware and software tools. The hardware consists of the associated devices with sensors hav-

ing a network among them, and a software component comprises data storage and analytics programs that help in presenting information for users. The IoT involves communication between different objects in an intelligent fashion. The IoT comprises a network of sensors connected with various devices, which gives information that can be evaluated to initiate different actions. IoT is mainly used to define smart devices equipped for sending data back remotely to a specific application or a computer server to provide some assistance in making people make smarter decisions. IoT

keeps monitoring machines and sensors, even when they are placed in immensely remote locations or places with extremely tough climate conditions [3]. The latest advancement in technologies, computing powers, storage sizes, and energy sources provides better ingredients for the IoT world [4]. IoT is aimed at connecting the physical structure, the IT requirements, the business, and the social requirements to influence the mutual intelligence of the city [5]. With the growth of IoT data growing at a frightening hop, the impending of IoT is also dignified [6].

The multimedia big data (MMBD) vision is also gaining prominence with the increase of IoT. The IoT devices generate gigantic multimedia data. The massive data is termed big data that plays a great part in the information management of intelligent applications. The multimedia data refers to the various media types including videos and animations along with text and audio. Multimedia communications change the lifestyle of people, thus changing the way we use devices present in a technological environment. Multimedia data communications have been evolved naturally, and it will be mixed between user experience and the adoption of the user by technology which unites to take advantage of living in a smart environment and government and corporate putting employs mechanisms to make multimedia communications in living easier and more livable [7]. Multimedia communications involve big data when there are more vital problems. Traditional data analytics of multimedia data face a lot of bottlenecks to process the heterogeneous data of the IoT environment. Therefore, the incorporation of IoT with big data plays a vital part in the direction of multimedia data computation. The multimedia big data management in the IoT setting is serving to solve the challenges associated with people and society including lighting automation, controlling traffic, and automation of building.

Big data management systems provide computing, analysis, storage, and control to resolve the issues of sustainability. Multimedia data generated in the IoT environment is the big data. The interactions among all the components involved in the IoT setting create a new type of big data collection of various types of applications and services. The IoT is regarded as the fourth basic need for human beings in the near future. Everything surrounding human being gets connected to the network every second. The IoT paradigm allowed us to connect through these different types of IoT devices in multiple ways that generate different types of data. Big data analytics is combined with IoT application development to truthful processing and computation of the generated multimedia data. Parallel and distributed processing platforms are used for processing big data followed by intelligent decision-making. Big data analytics of multimedia data has released a sphere of possibilities and opportunities in the industry including retail, energy, healthcare, transportation financial services, and manufacturing. In addition, some researches have been carried out to understand the conviction of community that covers health management, waste management, water controlling, traffic controlling, parking management, and so forth [8]. The MMBD has fascinated attention by providing an image of the worldwide infrastructure of multimedia communication.

Consequently, processing the immense data has turned out to be a need for smart community development. However, there are challenges featured by big data and IoT that are interoperability issues, heterogeneity problems, data valuing challenges, data format issues, normalization of data, incompleteness data, data filtration, scaling the data, etc. [9, 10]. A scalable infrastructure will be required to handle the massive influx of devices for multimedia communications. This article proposes a generic parallel and distributed framework to process the huge data efficiently and overcome the processing issue. The proposed scheme is a layered framework with a parallel and distributed module using big data analytics. A preprocessing module is also integrated with the proposed architecture to speed up the processing mechanism in the IoT-enabled environment. Specific datasets are utilized to realize the proposed multimedia big data management architecture to optimize the processing method of data.

2. Literature Review

The multimedia big data generation is growing rapidly day by day as industries and societies move towards the IoT applications that generate various types of data. This increase demands an efficient computational and analytics of multimedia data. An efficient architecture for MMBD management is a key pillar for the quality of IoT-enabled multimedia systems. It is bound to enhance the services including healthcare, traffic, parking, smart home, and surveillance. IoT-enabled multimedia information deals with sensor technologies that generate different types of multimedia data. In the most recent decade, academia and researchers both have extensively studied the IoT-enabled systems [11, 12]. Therefore, the concept is considerably raised in the entire world for linking various diverse devices, objects, or “things” with the Internet. The IoT-based applications generated MMBD that require to be computed to have resourceful decision management. As relocation to built-up endures better occupation chances, a more efficient system is required in handling challenges in the major areas.

A massive data is created by sensors and other digital devices resulting in big data. Traditional databases are ineffective for storing, processing, and evaluating purposes; that is why big data terminology has been introduced in the IT field [13]. The traditional approaches lack efficient cluster management and processing. Numerous sensor data are processed in a household and office environment that collects run-time data in an appropriate and inappropriate format [14]. The collected data was not preprocessed properly to remove the anomalies and aggregate in a uniform format. Many organizations use different analytical techniques, i.e., genetic algorithms, neural networks, and sentiment analysis, to learn the behavior of different types of data that helps in process discovery and analysis [15]. Big data brings altogether not only a large data volume but several data types that never would have been considered together previously.

The impression of objects connected through the internet is expanded with the implausible expansion of multimedia information that brought up IoT as an important gesture

[16]. Furthermore, the IoT environment allows both a static and mobile object to link to any object anywhere and anytime, which eventually produces any type of data [17]. Hence, the vital objective of such environments is to build up a model that could be used for the processing of heterogeneous information. The “things” are linked with the web employing various technologies such as ZigBee and Bluetooth that is another factor of having the distributed, heterogeneous, and diverse multimedia data. Although offline processing can assist in designing the MMBD processing framework, it lacks real-time decision-making. The hurried expansion of automation takes away the concentration of scientists in the way of effective designs. The customary processing design could present advantages to the researchers and industry.

The majority of investigative research depends on the new technologies of multimedia, and now, the societies are completely dependent on these technologies [18, 19]. To design a generic and efficient smart community schema, it is essential to vigilantly examine the big data collected from the community [20]. Several methods are designed to analyze the data based on the Hadoop processing mechanism for the betterment of the societies. For instance, the CiDAP (City Data and Analytics Platform) architecture is proposed that is used for data processing [21]. This architecture consists of an IoT broker and IoT agent that achieve higher throughput. Similarly, various other proposals are proposed to tackle this issue [22, 23]. These proposals help in different perspectives and provide methods to process data; however, they are not explicitly accessible for further use of research and lack an efficient processing mechanism. MMBD is very significant to regulate technology relevant to IoT because the IoT is a new-fangled technology that generates big data with a variety of formats [24]. Rapidly, this technology has been taken up by several groups, organizations, associations, and firms for better expansion of IT. There are a couple of IoT-based architectures that have been designed based on big data analytics [25, 26].

Unluckily, no model portrays a general configuration that could be utilized by all and sundry. In the same way, an inclusive architecture is required to grip data of diverse things and compute accordingly. The literature renders some notable challenges that require to be considered, that is, proper collection of the data from various IoT devices, noise removal, data analytics, and decision-making in the smart community development. To improve the storage and processing issues of MMBD generated in the IoT-enabled environment, we propose a specific architecture that overcomes the challenges faced in the IoT-enabled infrastructure. The proposed architecture is based on distributed and parallel processing that would be performing efficient processing using a multimedia big data analytics mechanism.

3. Proposed Framework

The proposed parallel and distributed architecture are connected with multimedia big data sources. The data sources comprised of multimedia big data such as weather, water,

traffic, health, and parking. The workflow of the proposed parallel and the distributed scheme is provided in Figure 1. Data gathering is done by the respective units collected from various devices. To devise effective parallel and distributed architecture, the data must be sensibly scrutinized before computation. The data is generated by different devices such as environmental monitoring sensors, security monitoring sensors, power monitoring sensors, facility monitoring sensors, traffic, and transportation monitoring sensors. The data is properly collected by the various departments such as the smart health monitoring department, water management department, traffic controlling authorities, and weather monitoring department. This process is known as edge caching. The predefined data is given to the proposed parallel and distributed architecture to process using proposed modules. This overall data collection is a part of a distributed system known as caching. It involves the overall data management including the aggregation, collection, and storage for the diverse multimedia data.

The data is also preprocessed before injecting into the proposed scheme to remove noise and make uniform format and anomalies for speeding up the processing activities. Afterward, the data is divided into different chunks for parallel processing. The distributed storage mechanism is also taken into consideration to assist the parallel processing. The preferred technique for storage is the Hadoop distributed file system (HDFS). The map-reduce parallel processing paradigm for big data processing is preferred that requires distributed storage. Hence, we preferred the HDFS storage technique. Premediated algorithms (e.g., capacity algorithms, DP algorithms) are applied for data processing in the cluster. The processed data are sent for the decision-making to the corresponding services providers, which are finally provided to the users. Afterward, the Hadoop processing unit is used to process the data which is stored in the distributed storage. Lastly, the analyzed data is operated for multimedia system planning. The data is collected from the smart community multimedia systems, and the decisions are sent back to the same systems.

The proposed architecture is connected with several multimedia sources that are smart home, smart water management system, electricity management system, smart environment, smart surveillance system, and so forth. The objective is to realize a scheme to perform efficient processing of multimedia big data. The multimedia systems are the data sources for the proposed system and a mediator between system and user. Architecturally, the anticipated solution consists of 3 modules which are data management, processing, and service management that are shown in Figure 2.

3.1. Multimedia Big Data (MMBD) Management. The MMBD management module involves the overall data organization including big data aggregation, acquisition and collection, and big data storage. The data is distributed across various devices for computation to get the load from the central server or cloud. Intelligent applications are supported by acquiring data via the Internet from various devices. Various devices including sensors, cameras, and

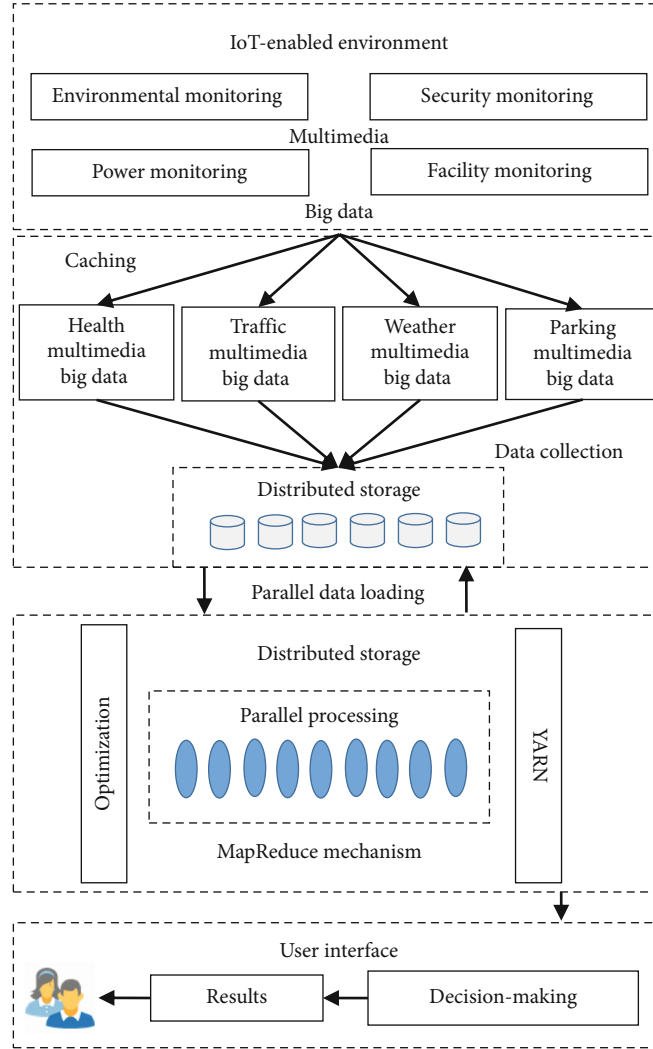


FIGURE 1: Workflow of the proposed framework.

object-mounted devices record the information of the environment in the different domains. This MMBD is later utilized for analysis to get insights and produce intelligent decisions. This layer is responsible for MMBD collection from various sources that are used to manage the multimedia systems and services. A practical community does not only hold a huge quantity of data but also includes versatile and wide-ranging processing areas. The multimedia systems implementation is dependent on all types of big data processing with heterogeneous environment. Data collection is used to transform signals that are assessed in practical circumstances and convert outcomes to the digital form for processing. The MMBD collection is done by a special system that pull out the data using various environmental objects to collect real-time MMBD. The data collection layer further includes the data aggregation, where the data are grouped based on the identification of the connected devices. This aggregation is applied because the data is gigantic and need to be grouped for efficient processing.

3.2. Multimedia Big Data (MMBD) Processing. The multimedia big data processing involves the distribution of MMBD into various splits for the parallel processing using multiple nodes in the cluster. The storage and processing of the MMBD are done using the Apache Spark distributed and parallel big data analytics platform. The process of computation from the distributed environment is based on the specific algorithms available in the distributed environment. On the other hand, cluster management is the particular activity in the system architecture. We utilized the specific utility called Yet Another Resource Negotiator (YARN) for cluster management. This layer is the central processing unit and is accountable for processing including training and inferencing. Initially, raw multimedia data that may include the irrational data combination, missing values, and values beyond the range are managed before processing the data. If the data is not inspected for such problems, there could be misleading results during decision-making. Therefore, the conversion is also performed to scale the data to a

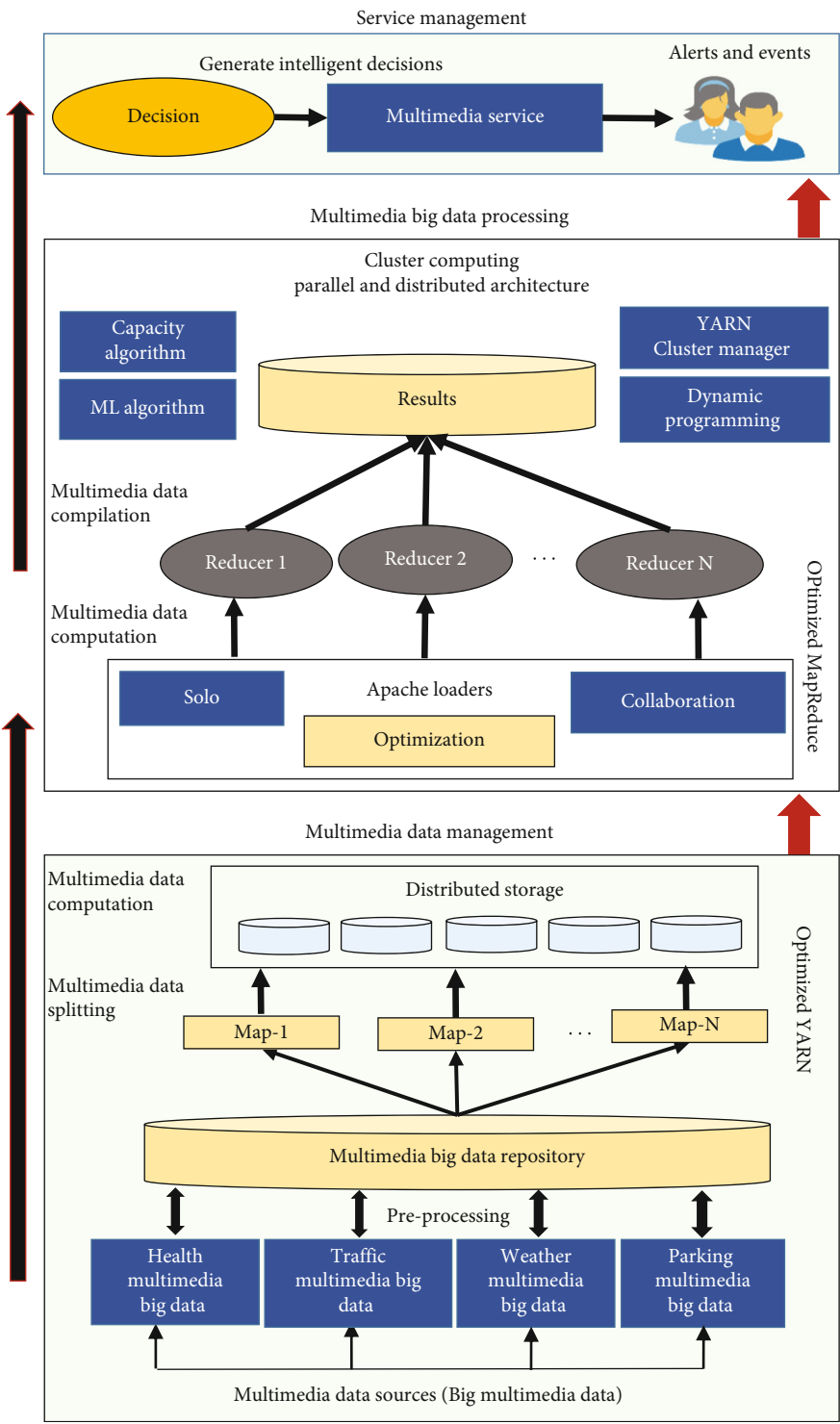


FIGURE 2: Proposed architecture.

particular specified scale. Subsequently, data is taken by a parallel and distributed processing unit that is the backbone of the proposed architecture.

The parallel and distributed processing unit processes huge data distributed and parallel form. It uses MapReduce programming that carries work in mapping and reducing processes. The storage obligation is supported using the distributed filing system. The proposed architecture is based on a parallel and distributed computing paradigm that is utilized for processing and computation. An optimized map-reduce model is introduced to implement the MMBD analytics. The proposed optimized model processes huge datasets in parallel. It executes the analytics processes in a distributed manner with high availability. The proposed architecture also overcomes the machine failures, machine's performance issues, and effective communications. The job distribution in the computing cluster is performed using the YARN cluster management framework. The YARN model is equipped with dynamic programming for job distribution and resources management in the cluster. The earlier versions of the parallel and distributed platforms (e.g., traditional Hadoop) utilized MapReduce for both processing and cluster management, which created the communication overhead and decreases the performance. On the other hand, the Yet Another Resource Negotiator (YARN) is favored as it performs cluster management separately. A detailed explanation is provided in the revised manuscript.

In the context of YARN-based cluster management, the job is a YARN application. The Application Master implementation in the YARN-based solution is provided by MRAppMaster. There could be many tasks in the phase of mapping or reducing stage for every split of MMBD. In addition, the map and reduce phases can be interleaved; therefore, the reduce phase may start before the end of the map phase. When an application is submitted to YARN considering the edge information, some additional information is given to YARN infrastructure particularly a configuration, a JAR file, and input/output information. The configuration might be limited as some parameters might not be specified in that case; for job execution, the default values are utilized. If the data file size is too large (larger than the block size of HDFS (256 MB in proposed architecture)), then there would be two or more two map splits related to the same input file. For stream processing, running map-reduce is not a good idea. Additional utilities (e.g., Apache Spark) can be added with optimized MapReduce to process the data in memory (RAM) rather than put it on a disk, etc.

The predefined method called `getSplits()` of the predefined class called `FileInputFormat` is customized with information for implementing map-reduce jobs using dynamic programming algorithms. The MRAppMaster requests for containers to the resource manager required executing the job. After the setup of the system, for every $\langle k, v \rangle$ tuple enclosed in a split of the map, the `map()` method is called. Therefore, `map()` accepts a specific key, a specific value, and a specific mapper context. We utilized the parallel and distributed platform. We preferred the use of the capacity scheduling algorithm for parallel processing. Similarly, the DP (dynamic programming) is utilized for recursion as the

huge data is divided into chunks repeatedly for parallel processing. The DP is one of the best algorithms for resolving a problem by recursively breaking it down into simpler sub-problems and producing the optimal solution. The Apache loader is the utility that loads the Apache Hadoop files and libraries for map-reduce and other operations.

The output of the optimized mapper is stored to a particular buffer using the context. When the overall optimized-map split has been processed and executed, the clean method is called by a run. There is no action performed by default but the user might decide to override it. MapReduce is the batch-processing mechanism. The concept behind MapReduce is that data of a particular edge is first grouped into small portions. After that, these edge-map splits or portions are processed in a distributed manner to create desired results. MapReduce is the extension of the traditional MapReduce model which is easy and generally useful for various diverse applications such as bioinformatics, web mining, and machine learning technologies. On the other hand, stream processing is meant for processing and executing the data record-wise as the data is pulled in and updated incrementally. The result is reorganized with each new record of the data. Stream processing queries run continuously and are never ending.

3.3. Service Management. This layer exists on top of the proposed architecture and is responsible for making decisions and communicating the corresponding decision to the corresponding departments of the smart community. It is the end-user or application layer. The intelligent decisions and results are sent to the management centers. Afterward, the events are classified and the corresponding users are informed accordingly. First of all, the sophisticated actions are kept at community development departmental glassy and then dispatched to the users. The application interfaces are responsible for the subservice selection and the classification of the events. Moreover, the decision-making process is carried out using a properly maintained rule engine, where several rules based on particular thresholds are defined. IoT-enabled MMBD applications could be education that involves people in dynamic learning environments to adjust to the rapid changes of society, traffic lights that improve the transportation systems and overall traffic patterns, smart grid that improves the productivity and supply of electric power, healthcare that monitors and analyzes health issues daily or on-demand basis, smart energy that helps in management related to the providing the intensity of power with the definite claim of the citizens, and smart environment that provides weather info to improve the country's agriculture and other possible harmful circumstances.

4. Results and Discussion

We utilized the open-source platform Apache Hadoop version 3.0 (parallel and distributed platform) for the validation of the proposed framework. Additionally, Apache Spark is configured with the Hadoop version 3.0 for real-time stream processing and computation. The discussion about the results is provided in this section. Results are produced by

utilizing various authentic multimedia big data sets to assess the proposed MMBD framework using parallel and distributed paradigms and premeditated algorithms. The pre-processing mechanism is carried out before the core processing to remove the anomalies and noise from multimedia big data sets. Thus, notable efficiency is achieved in the processing time and throughput. The proposed architecture is implemented using Hadoop and Apache Spark parallel and distributed framework along with optimized premeditated algorithms. The datasets include transportation and vehicular data, pollution data, and water data [27–29]. These datasets are preferred due to the utilization of this dataset in the literature. We deliberately executed almost the similar kind of queries for the processing time and throughput comparison of the proposed IoT-enabled MMBD system using optimized map-reduce and YARN for parallel and distributed processing.

The utilization of the water is assessed to accomplish sustainable management and control of the water consumption in the region. The unpredictable depletion of water can be a catastrophe for society. The dataset used in the proposed research includes the data of the city of Surrey, Canada. The data is collected for various sources that contain the intake of water by houses. The information is processed using our suggested processes. The outcomes are revealed in Figure 3. It demonstrates the families spent water more than 83500 liters a month. The well-defined threshold is 83500 according to the defined rule. The water consumption higher than the threshold or TLV is specifically emphasized that could cause scary circumstances for the consultants. It is perceived that almost 55% of the users utilized more water than the TLV. The users above TLV limit usages have utilized water between 115000 and 120500 liters that is pretty frightening. Current production approaches can be utilized by the industry to govern the challenges of the users.

Furthermore, the traffic dataset is also experimentally tested with the proposed optimized MMBD architecture. The average speed is exposed in Figure 4. The average speed is observed in the pretty similar complete day, excluding from 12:50 to 17:50 for a few automobiles.

Similarly, the pollution dataset is also experimentally tested with the proposed optimized MMBD framework using customized map-reduce and YARN for parallel and distributed processing. The dataset of the pollution is thoroughly investigated. The pollution of a specific day is also shown in Figure 5.

The selection criteria of the parameters for the analysis of the results were the parameters of the base papers. We preferred the same set of queries utilized by base papers. We also preferred the same datasets for comparison. Moreover, a similar configuration (storage, CPU, cluster, and nodes) is preferred for comparison. Equally, a comparison to related architectures is also provided concerning system processing time in Figure 6. The proposed architecture processing time is better than the existing solutions provided in the literature for smart city planning. The performance improvement is due to the customization in the YARN cluster management and utilizing MapReduce. The YARN and MapReduce are the extensions of the traditional YARN

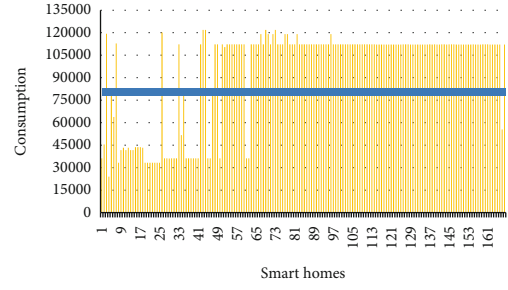


FIGURE 3: Consumption of water in smart homes.

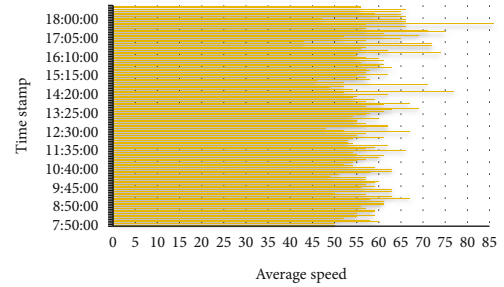


FIGURE 4: Average speed.

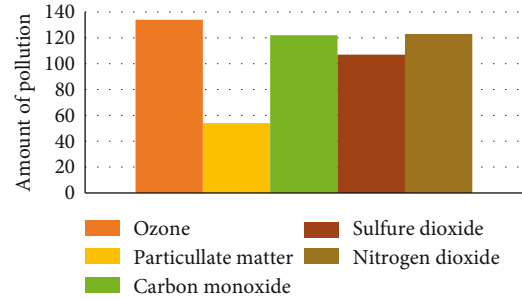


FIGURE 5: Pollution.

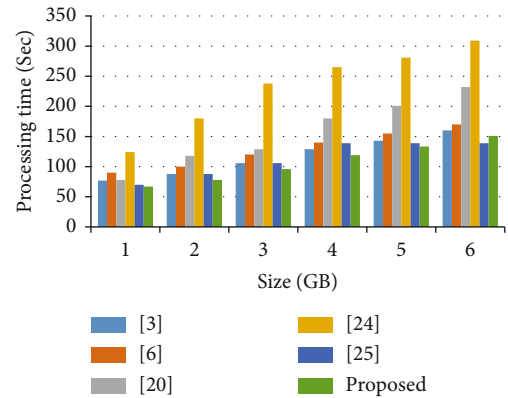


FIGURE 6: Processing time comparison with existing solutions.

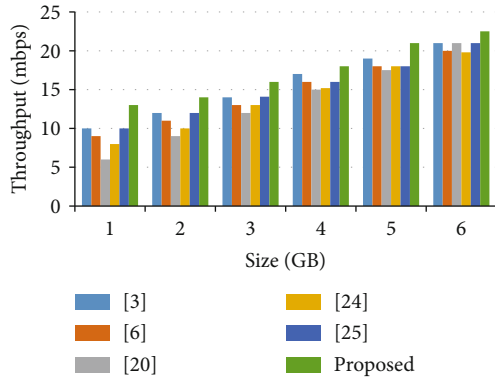


FIGURE 7: Throughput comparison with existing solutions.

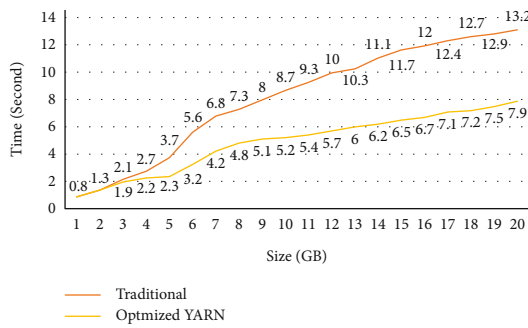


FIGURE 8: Traditional vs. optimized cluster management.

and MapReduce for edge computing. The results show that the proposed architecture is better than existing architectures in the context of processing time.

Finally, a comparison to related architectures is also provided about throughput in Figure 7. The proposed architecture throughput is much better than the existing solutions provided in the literature for planning due to the customization of the traditional distributed framework. Figure 8 highlights the effectiveness of optimized YARN. The throughput increases due to the decrease in the processing time. The processing time decreases because the parallel processing platform is optimized with different block sizes, replica numbers, and the use of the configured utility.

5. Conclusion

This paper proposes a generic scheme to process data in the parallel and distributed mechanism to overcome the processing issues. Parallel processing is utilized to perform big data processing. This research intends the explicit apprehension of multimedia communication to make possible efficient data processing and decision-making. Multimedia systems provide computing, storage, and analysis, to solve the challenges. However, it becomes challenging to tackle with the diverse IoT settings. The proposed architecture is a layered framework with a parallel and distributed architecture using big multimedia data analytics. A preprocessing module is also integrated with the proposed architecture to speed up the processing mechanism of big data produced

by IoT devices in the IoT-enabled environment. Specific datasets are utilized to realize the proposed architecture to optimize the processing of data. The proposed system is realized using real-time datasets from various sources. The proposed architecture is experimentally tested with authentic datasets that reveal the effectiveness of the proposed architecture.

Data Availability

The data used to support the findings of the study are included within the article.

Conflicts of Interest

The authors declare that they do not have conflict of interest.

Acknowledgments

This study was supported by Taif University Researchers Supporting Project (number TURSP-2020/126), Taif University, Taif, Saudi Arabia.

References

- [1] Z. Liao, S. Nazir, H. U. Khan, and M. Shafiq, "Assessing security of software components for Internet of Things: a systematic review and future directions," *Networks*, vol. 2021, article 6677867, 22 pages, 2021.
- [2] M. Ahmad, T. Younis, M. A. Habib, R. Ashraf, and S. H. Ahmed, *Recent Trends and Advances in Wireless and IoT-enabled Networks*, Springer Nature Switzerland AG, 2019.
- [3] M. Babar, F. Arif, M. A. Jan, Z. Tan, and F. Khan, "Urban data management system: towards Big Data analytics for Internet of Things based smart urban environment using customized Hadoop," *Future Generation Computer Systems*, vol. 96, pp. 398–409, 2019.
- [4] M. Elkhodr, B. Alsinglawi, and M. Alshehri, "Data provenance in the internet of things," in *2018 32nd International Conference on Advanced Information Networking and Applications Workshops (WAINA)*, pp. 727–731, Krakow, Poland, 2018.
- [5] L. Ru, B. Zhang, J. Duan et al., "A detailed research on human health monitoring system based on Internet of Things," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 5592454, 9 pages, 2021.
- [6] M. Babar and F. Arif, "Real-time data processing scheme using big data analytics in internet of things based smart transportation environment," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 10, pp. 4167–4177, 2019.
- [7] H. Gao, Y. Duan, L. Shao, and X. Sun, "Transformation-based processing of typed resources for multimedia sources in the IoT environment," *Wireless Networks*, vol. 27, no. 5, pp. 3377–3393, 2021.
- [8] W. Shang, J. Chen, H. Bi, Y. Sui, Y. Chen, and H. Yu, "Impacts of COVID-19 pandemic on user behaviors and environmental benefits of bike sharing: a big-data analysis," *Applied Energy*, vol. 285, 2021.
- [9] Z.-N. Li, M. S. Drew, and J. Liu, "Network services and protocols for multimedia communications," in *Fundamentals of Multimedia*, pp. 535–582, Springer, 2021.

- [10] L. Tan, K. Yu, F. Ming, X. Chen, and G. Srivastava, "Secure and resilient Artificial Intelligence of Things: a HoneyNet approach for threat detection and situational awareness," *IEEE Consumer Electronics Magazine*, vol. 2021, p. 1, 2021.
- [11] K. Yu, Z. Guo, Y. Shen, W. Wang, J. C. Lin, and T. Sato, "Secure Artificial Intelligence of Things for implicit group recommendations," *IEEE Internet of Things Journal*, vol. 2021, 2021.
- [12] Z. Guo, K. Yu, Y. Li, G. Srivastava, and J. C.-W. Lin, "Deep learning-embedded social internet of things for ambiguity-aware social recommendations," *IEEE Transactions on Network Science and Engineering*, vol. 2021, 2021.
- [13] M. Khan, J. Iqbal, M. Talha, M. Arshad, M. Diyan, and K. Han, "Big data processing using internet of software defined things in smart cities," *International Journal of Parallel Programming*, vol. 48, no. 2, pp. 178–191, 2020.
- [14] B. Anthony Jnr, S. Abbas Petersen, D. Ahlers, and J. Krogstie, "Big data driven multi-tier architecture for electric mobility as a service in smart cities," *International Journal of Energy Sector Management*, vol. 14, no. 5, pp. 1023–1047, 2020.
- [15] E. Carter, P. Adam, D. Tsakis, S. Shaw, R. Watson, and P. Ryan, "Enhancing pedestrian mobility in smart cities using big data," *Journal of Management Analytics*, vol. 7, no. 2, pp. 173–188, 2020.
- [16] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: a survey on enabling technologies, protocols, and applications," *Communications Surveys and Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015.
- [17] M. A. Razzaque, M. Milojevic-Jevric, A. Palade, and S. Clarke, "Middleware for internet of things: a survey," *IEEE Internet of Things Journal*, vol. 3, no. 1, pp. 70–95, 2016.
- [18] Y. Gong, L. Zhang, R. Liu, K. Yu, and G. Srivastava, "Nonlinear MIMO for industrial Internet of Things in cyber-physical systems," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 8, pp. 5533–5541, 2021.
- [19] S. Uppoor, O. Trullols-Cruces, M. Fiore, and J. M. Barcelo-Ordinas, "Generation and analysis of a large-scale urban vehicular mobility dataset," *IEEE Transactions on Mobile Computing*, vol. 13, no. 5, pp. 1061–1075, 2014.
- [20] M. Khan, B. N. Silva, and K. Han, "Internet of things based energy aware smart home control system," *IEEE Access*, vol. 4, pp. 7556–7566, 2016.
- [21] B. Cheng, S. Longo, F. Cirillo, M. Bauer, and E. Kovacs, "Building a big data platform for smart cities: experience and lessons from Santander," in *2015 IEEE International Congress on Big Data*, pp. 592–599, New York, NY, USA, 2015.
- [22] "FIWARE open source platform," <https://www.fiware.org/>.
- [23] "SCOPE a smart-city cloud-based open platform and ecosystem," <https://www.bu.edu/hic/research/highlighted-sponsored-projects/scope/>.
- [24] M. Khan, B. N. Silva, and K. Han, "A Web of Things-based emerging sensor network architecture for smart control systems," *Sensors*, vol. 17, no. 2, p. 332, 2017.
- [25] M. Babar, A. Rahman, F. Arif, and G. Jeon, "Energy-harvesting based on internet of things and big data analytics for smart health monitoring," *Sustainable Computing: Informatics and Systems*, vol. 20, pp. 155–164, 2018.
- [26] A. Ahmad, A. Paul, S. Din, M. M. Rathore, G. S. Choi, and G. Jeon, "Multilevel data processing using parallel algorithms for analyzing big data in high-performance computing," *International Journal of Parallel Programming*, vol. 46, no. 3, pp. 508–527, 2018.
- [27] "T. Dataset, Dataset collection," <http://iot.ee.surrey.ac.uk:8080/datasets.html#traffic>.
- [28] "P. Dataset, Dataset collection," <http://iot.ee.surrey.ac.uk:8080/datasets.html#pollution>.
- [29] "Dataset water meters," <http://data.surrey.ca/dataset/water-meters>.