ISSN: 2583-9292 (ONLINE)
DOI: 10.21917/ijdsml.2026.0192

EVALUATING APACHE HADOOP IN TELECOMMUNICATION BIG DATA: SCOPING REVIEW AND SCIENTOMETRICS ANALYSIS

Arti Sawale and Paramjeet Kaur Walia
Department of Library and Information Science, University of Delhi, India

Abstract

Liberalization has enhanced the global access to telecommunication and networking services. This could imply the need for advanced technologies to handle the growing telecom data demands. The exponential growth of mobile internet in telecommunication is adopting the big data platform for the data acquisition, storage and analysis to provide the value-added services. The main objective of this research paper is to demonstrate the use of Apache Hadoop ecosystem in Telecommunication Big data Analytics. Hadoop is an open-source and well known for distributed large dataset storage and processing mechanism with core modules HDFS, YARN, Hadoop Common and MapReduce. The study explores the use cases for Apache Hadoop ecosystem in the telecommunication sector and demonstrates the versatility of Hadoop in addressing different aspects of telecommunication data such as analysis of telecom network traffic logs, location-based services, call detail records (CDR), telecom user churn prediction, and anomaly detection. Finally, 77 research papers which implemented the Apache Hadoop in telecommunication big data processing efficiently retrieved from IEEE Xplore are reviewed and analysed. Scoping review and scientometrics analysis methodology have been used. The results of the study interpret the implications of Apache Hadoop in telecommunication big data analytics, offering insights into their use across various decision-making domains to enhance the potential impact of big data analytics in improving telecommunication services.

Keywords:

Telecommunication, Big data, Hadoop, Scientometrics, IEEE

1. INTRODUCTION

It is evident that big data is relevantly impressing all the possible sectors of technologies. During the last decade, the impact of big data on the performance of telecommunication (telecom) and networking has evolved drastically. The evolution in telecommunication technologies started with the perspective of mobility in 1980s to the advanced quality integration of networking in the 21st century. This revolution happened from 1G to 5G, i.e., wired to wireless telecommunication technologies with the high-tech mobility. Today technological advances came up with Industry 4.0, where big data and data analytics are the main components of this fourth industrial revolution not only just about data compilation but being able to create a controlled system with a high-speed and flexible type of production with different information sources [1]. Now a days humans are relying on the smart devices and every second a new device is getting connecting and disconnecting to the internet. Along with smart devices, industrial devices and communication i.e., Industrial Internet of Things (IIoT) are actually contributing to generate the large data over internet. The administration of devices and users anywhere and anytime are getting accelerate with leveraging Internet of things (IoT) in the quality of telecommunication services. However, the growing wireless environment is a large source of virtual big data, which is creating a new challenge in telecommunication sector. These challenges are unprecedented discussed by Zahid et al. [2] in the systematic review of literature of big data in telecommunication domain. Telecommunications sector continues to expand, the volume, velocity, and variety of data generated have increased exponentially. The several data types of telecom data are neither effectively analysed by Relational database- dataset and nor by traditional data analysis tools. The most important reason behind this is big data set which get updated with real time and such continues growth in amount of data get managed by big data analytics tools. Big data analytics (BDA) provides a unified solution for the unstructured and huge telecom data analytics mainly for analysis of network traffic logs, location-based services, call detail records (CDR) also for telecom user churn prediction and anomaly detection.

This scoping review aims to provide the implementation of big data analytics in telecommunication and focus on the Hadoop ecosystem. Most of the telecom big data case studies have been performed on an open-source big data analytics platform Hadoop which is subject of this literature review study. The main reason behind BDA in telecom is the variety and volume of data generated by telecom value added services, web logs, real time network traffic, signal processing and many more. The literature is reviewed with reference to the telecom big data theoretical analysis and its practical implications on BDA platform. Many studies show that, variety of data types keeps on mounting at large velocity and resurge to the huge volume of telecom big data which is complex to be managed by traditional data processing software. Therefore, it utilizes cloud computing and BDA tools to store and process large pool of structured, semi-structure and unstructured telecom data.

1.1 BRIEFING THE HADOOP ECOSYSTEM

Hadoop is an open-source framework for distributed storage and processing of large datasets, played a pivotal role in the evolution of big data. Hadoop was initially developed at Yahoo. It was inspired by Google's MapReduce and Google File System (GFS) papers. The project aimed to provide a scalable and distributed computing framework for processing large datasets. It allowed to store and process vast amounts of unstructured and semi-structured data, including social media data, sensor data, and many more. The ecosystem around Hadoop expanded with the development of additional projects such as Apache Hive, Apache HBase, Apache Pig, and more. The concept of distributed computing and the MapReduce programming model of Hadoop is shown in Fig.1. Hadoop ecosystem tools architecture has also been shown in Fig.1. with respect to storage of larger size (Volume) of unstructured and semi structured data types (variety) collected from telecom data sources storage in HBase, which is NoSQL Hadoop database for Hadoop Distributed File System (HDFS). Big data rush streaming out with real time data

processing speed (Velocity) dealt by MapReduce which enhance the scalability in Hadoop cluster.

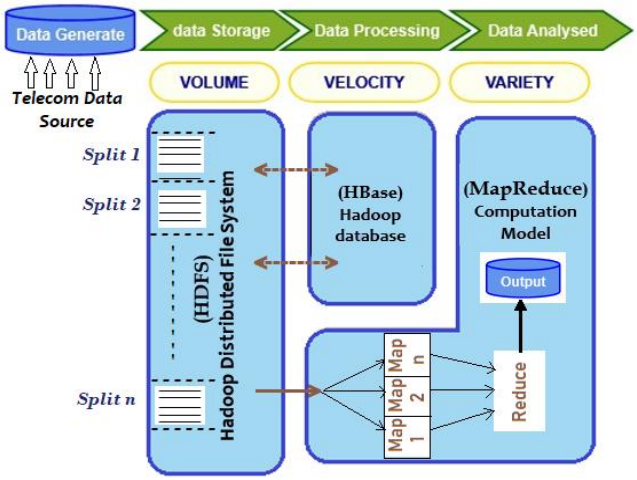

Fig.1. Hadoop Ecosystem tools for components of big data

The mention of MapReduce and HDFS specifically highlights their central role in the Hadoop ecosystem. MapReduce is the programming model for processing and generating large datasets, and HDFS is the distributed file system that enables reliable data storage across nodes in a Hadoop cluster and splits the massive data in key-value pairs and reduce the map input for further data processing. However; HIVE offers an SQL-like query engine which simplifies the MapReduce, but in many cases, MapReduce rises latency where HBase a NoSQL database on Hadoop lower the latency with faster and real time data processing.

2. REVIEW OF LITERATURE

This study, the primary focused on telecom big data analytics and referred the studies which have been conducted on the performance evaluation of telecommunications networks. Several case studies use big data analytics in telecommunication applications of cellular networks, software defined networks (SDN), optical networks, network security, CDR logs, anomaly tele-traffic etc. Telecom big data stores valuable information to support the telecom operations, based on this, Zhu et al. [3] has designed scheme to precisely analyse the reliability of wireless environment to assist telecom operator to locate the problematic base stations. Digging into the challenges of data acquisition and analysis for value-added services, Wenliang Huang et al. [4] adopted the big data platform to introduce the telecom operator centric ecosystem based on mobile internet big data. Zeydan et al. [5] proposed architecture for 5G wireless network traffic analyses essentially based on Hadoop. Also, the comprehensive literature analyses on 5G integration with smart grid presented by Rituraj et al. [6].

3. METHODOLOGY

This study adopted the Scientometrics analysis and scoping review method. The research studies published during 2011 to 2022 retrieved from IEEE Xplore database which is prominent source of comprehensive telecom literature. Various studies are available which broadly explained the need of scoping review over a systematic review study. Arksey and O'Malley [7] proposed methodological framework on the basis of the main reasons to conduct scoping studies to perform scoping review, which are examining the nature of research studies, determining the feasibility of systematic literature studies, summarising the research finding and identifying the research gap in present research studies. Many authors have applied the scoping review methodology for their research papers, Munn [8] has provided a fundamental example for the scenario of deciding the type of review to the authors for their research studies. Further, Peters et al. [9] has provided broad discussion on scoping review which he also referred as 'mapping review' i.e., mapping the topic base literature findings to determine the quantitative and qualitative approach on the topic and to represent this evidence by mapping or charting the data. Tricco et al. [10] [11] presented PRISMA-ScR (Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews) as an extension to scoping review with intend to help and develop a greater understanding of relevant terminology, core concepts of topic, and key items to report for scoping reviews which is also got citated in many research papers.

4. RESEARCH QUESTIONS

The primary research question is to study how telecom sector has evolved over big data. The review of previous studies highlights telecom big data has evolved gradually. Many research publications answer the present and the future study of big data in telecommunication domain. Exploring the big data in telecommunication fulfils the gap of expanding technologies and attempted to provide roadmap of exploring the technologies used for telecom BDA. Scoping review on application of Hadoop tools in telecom big data exposed the previous research. The existing literature also proposed the concepts for comprehensive future studies.

- RQ1. What are the study measures Hadoop ecosystem tool in telecommunication big data analytics?
- RQ2. Providing the scoping review data charting and Scientometrics analysis of annual scientific publication of Hadoop tools use for data storage, analysis and processing in telecom big data from year 2011 to 2022.

5. SCOPE OF THE STUDY

Scientific research review accelerates scientific approach to understand the previous and present research trends. An in-depth information about any research is important to know the current trends through existing literature [8]. Telecommunication is a vast field of study and IEEE Xplore provides the comprehensive access for the research publication in the field of telecommunication networking.

5.1 DATA COLLECTION CRITERIA (FINDING RELEVANT RESEARCH PUBLICATIONS)

The next stage is finding the relevant research publications with searching keyword which covers area of big data in telecommunication with below Inclusive criteria:

- **Period**: 2012-2022; Language Covered: English;

- **Document Identifier**: IEEE Journals; Conference and magazine
- **Search**: Command search and Advanced Search;
- **Data fields**: Abstract, Author keywords, IEEE and INSPECT terms
- **Inclusive keywords**: big data; Telecommunication; Telecom, Networking

A simple searching strategy considered for data collection to get maximum number of publications from the data fields in full text and metadata. According to Scientific Journal Rankings (SJR) of the year 2021, most of the journals listed in top ranks for Computer Networks and Communication were IEEE Journals, so this pilot study has carried with the intention of analysing the IEEE journals for telecommunication domain. As telecommunication is also referred as Telecom and Telco, therefore; searching with all the possible keyword the query formulated with the filters 'IEEE Conference', 'IEEE Journals' and 'IEEE Magazines' for the year 2011 to 2022. The searching strategy used is 'advanced search'.

Query formulated: ("Abstract": telecom*) AND ("Abstract": big data); ("Abstract": telecom*) AND ("Abstract": hadoop). Similar advance search performed for Author Keywords; INSPEC Controlled and Non-Controlled Terms; Index terms; IEEE terms.

Due to the limitation in exporting the results up to 2000 in CSV file with 30 fields, the priority has been given to advanced search as the output was collectively easy. Secondly to double check the command search was also performed for the comprehensive data collection of telecom big data. The included keywords are telecom, telecommunication, mobile, cellular, VOIP (Voice over Internet Protocol), smartphone, smart devices and 4G, 5G, 6G with the combine keyword search for 'big data' OR/AND 'hadoop' OR/AND 'telecom*' (* provide all the matches e.g., telecom, telecommunication etc.)

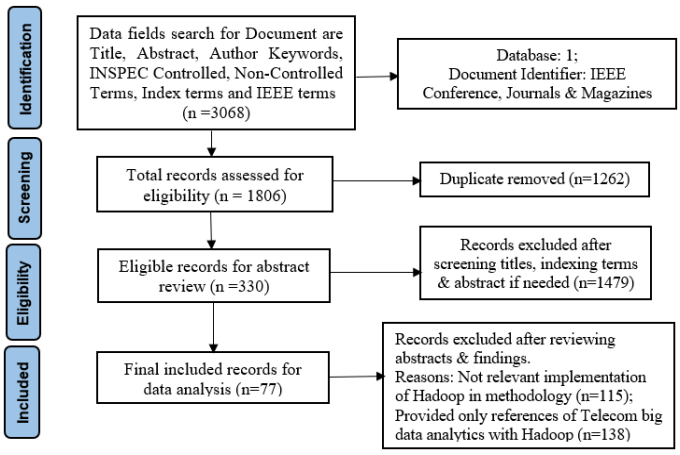

Fig.2. Documenting the literature search as per guidelines http://www.prisma-statement.org/ provides for PRISMA flowchart

As shown in Fig.2, the tentatively eligible records of 330 publications are reviewed with respect to abstracts and findings. Further some of the records have been excluded due to the reasons like; not implementing Hadoop in data analytics, provided only reference of telecom big data, only comparative studies of Hadoop with other BDA tools and some of the research titles are also excluded due to only theoretical discussion of usage of Hadoop in telecommunication sector. The total numbers of research titles included for the final scoping review data charting are 77 in number.

6. DATA ANALYSIS AND FINDINGS

6.1 ANNUAL SCIENTIFIC PRODUCTION WITH RESPECT TO ARTICLE CITATION AND FULL TEXT VIEW

Article citations measure the scope and strengthen of the study carried out. In this paper, the annual total citation counts of article cited in IEEE publication and other publications have been provided in Fig.3, which highlights yearly increasing pattern for number of articles published and cited from 2011 to 2016. The Fig.4 shows the gradual increasing pattern in annual full text views count. However; the old years articles published gets the more citations.

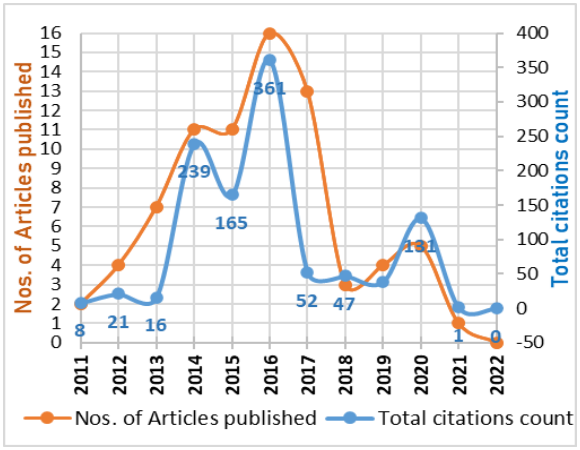

Fig.3. Annual Publications and IEEE Total Citation Count Per Year

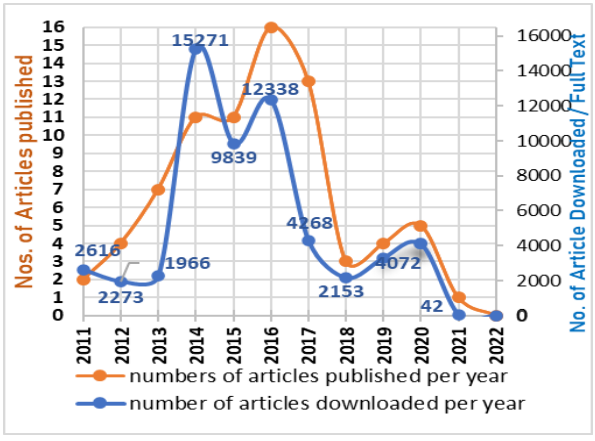

Fig.4. Annual Scientific Production and IEEE Full Text Views/ Downloads

The scoping review on Hadoop implementations for telecom big data analytics revealed that there has been continuous increase in research output, number of citations per year as well as full text views from the year 2011 to 2016. It is observed the highest

number of articles were published in 2016 and 2017 in comparison to 2018-2022. The latest years articles cannot be ignored on citations basis as citation per year has increase by number of years passed from the year of publications and also the measure of research output.

6.2 ANNUAL SCIENTIFIC PRODUCTION IN IEEE PUBLICATIONS

The data analysis pertaining to articles published in IEEE conference, journals and magazines are shown in Fig.5 and it has been observed that maximum articles are published in IEEE conference than IEEE journals and magazines. 'IEEE Network and Communications magazines' has included 4 publications and the titles listed in journals are shown in Table 1. As Hadoop is a very important and more user-friendly tools for big data analytics also it has a vast scope of studies in ongoing researches and therefore conference often received the vide range research studies. The constructive feedback and suggestions received over the research studies on big data analytics more quickly hence, researches give priority to the conference publications than journals and magazines.

The list of the conferences where more than one research papers are published out of the 77 research titles included in this study are 'IEEE International Conference on Big Data' and 'Asia-Pacific Network Operations and Management Symposium' with three publications. Also, with two publications each, the conferences included are 'IEEE International Conference on Mobile Data Management', 'International Conference on Network-Based Information Systems', 'International Symposium on Networks, Computers and Communications', International Symposium on Wireless Personal Multimedia Communications and 'IEEE International Conf on Parallel and Distributed Processing with Applications, BDCloud, SustainCom, SocialCom'.

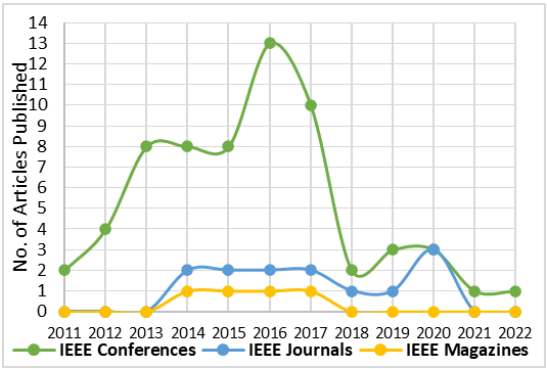

Fig.5. Annual Publication of IEEE Conference, Journals and Magazines

Table.1. List of Journals and their Scientometrics parameters

Journal titles	N(total Citations)	Journal h-index	SJR 2022	Cite score 2022
IEEE Access	2(8)	204	0.93	9
IEEE Transactions on Knowledge & Data Engineering	1(60)	190	2.35	12.4
IEEE Transactions on Industrial Informatics	1(120)	170	4	22.4
IEEE Transactions on Mobile Computing	1(14)	151	2.54	13.1
IEEE Transactions on Consumer Electronics	1(8)	108	1.46	9.9
IEEE Transactions on Services Computing	1(69(	84	2.11	12.5
IEEE Systems Journal	1(42)	98	1.44	9.1
IEEE Transactions on Cloud Computing	2(21)	61	1.43	10
Tsinghua Science and Technology	2(27)	52	1.64	10.2
Journal of Communication & Infor Networks	1(3)	10	1.58	4

The articles published in journals are provided in Table.1, where *N* is the number of articles published related to Hadoop application in Telecom Big Data has total 77 articles, which are considered in this study. The metrics parameter SJR indicator is SCImago Journal Ranking which is a scientific indicator based on Scopus database. CiteScore is metrics of yearly citations count with reference to the citations count of the previous three years, also h-index is another important metrics based on the Scopus indexed publications, where journals h-index provided the journals (h) number of articles which received at least (h) citation, which helps to calculate journals scientific productivity.

6.3 FUNDING PUBLICATIONS

In the last decade 2010-20, big data research projects have been encouraged by many funding agencies and institutes with the purpose of carrying out the advance research in big data domain. Out of 77 studies included in this paper, 8 research titles have been funded by institutes and universities as shown in Table.2.

Table.2. List of funding bodies

Funding Institutes	No. of Articles/Year
University of Cyprus	1(2015)
Deanship of Scientific Research at King Saud University	1(2016)
US National Science Foundation	1(2018)
Naval Postgraduate School; US National Science Foundation	1(2019)
The National Polytechnic School Research & Social Projection Management Unit Quito Ecuador	1(2020)
National Natural Science Foundation of China; Sichuan Science and Technology Program; Key Projects of the Sichuan Provincial Education	1(2020)
Wuhan University	1(2021)

6.4 AUTHORS AND AFFILIATIONS

The parameters listed in table 3 showing the Fig.of authors and affiliations contributions. The total number of authors for the 77 research titles included in this paper are 319 from the 135 different institutes (Affiliations) across all countries. It has been overseed that China has the maximum number of publications (122) follows by India (24), south Korea (18), USA and Ireland (12) respectively for the Hadoop studies in telecom big data. The Scientific collaboration of different affiliations across the globe from year 2011-22 is carried out and co-authorship pattern by countries is shown in Fig.6, it has been observed that authorship collaborations mostly taken places with the five countries i.e. USA, China, United Kingdom, Brazil and Ireland.

Table.3. Authors collaboration

Parameters	Counts
Total number of authors	268
Number of papers with single author	6
Number of papers with co-authors	71
Average number of authors per paper	4
Number of affiliations contributed	135
Countries of authors affiliation	31

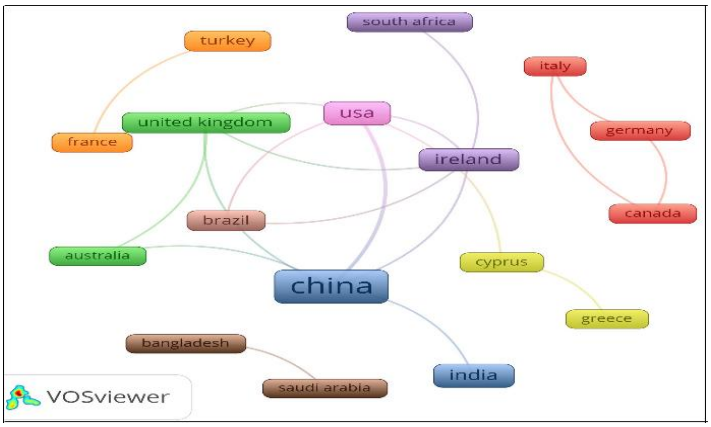

Fig.6. Co-Authorship by Countries

6.5 AUTHOR KEYWORDS CO-OCCURRENCE NETWORK VISUALIZATION AND IEEE TERMS CO-OCCURRENCE DENSITY VISUALIZATION

Note(s): Each node in a network represents an entity (e.g., Author keywords keyword mentions in telecommunication big data analytics implemented Hadoop papers), wherein: (1) the size of the node indicates the occurrence of the keyword (i.e., the number of times that the keyword occurs), (2) the link between the nodes represents the co-occurrence between keywords (occur together), (3) the bigger the node, the greater the occurrence of the keyword. Each colour represents a thematic cluster and coverage of topics (nodes) and the relationships (links) between the topics manifesting under that topic cluster.

The keywords provided by authors in their respective papers are used to highlight the scope of the research studies and findings from the content of literature. The visual analysis of authors keyword co-occurrence mentioned in 77 articles are analysed by VOS viewer and shown in Fig.7. The co-occurrence of all author's keyword around 150, keyword 'big data' occurred in 23 documents with total link strength 103, 'Hadoop' in 25 documents with total link strength. In Fig.8 the density visualization of IEEE terms i.e., the indexing terms used by IEEE is shown where the topics with yellow colour shows largest total link strength (TLS) also indicate the maximum interconnectivity with other topics covered in studies. Further the yellow coverage goes lighter shows the less connected topics or terms that means less discussed topics.

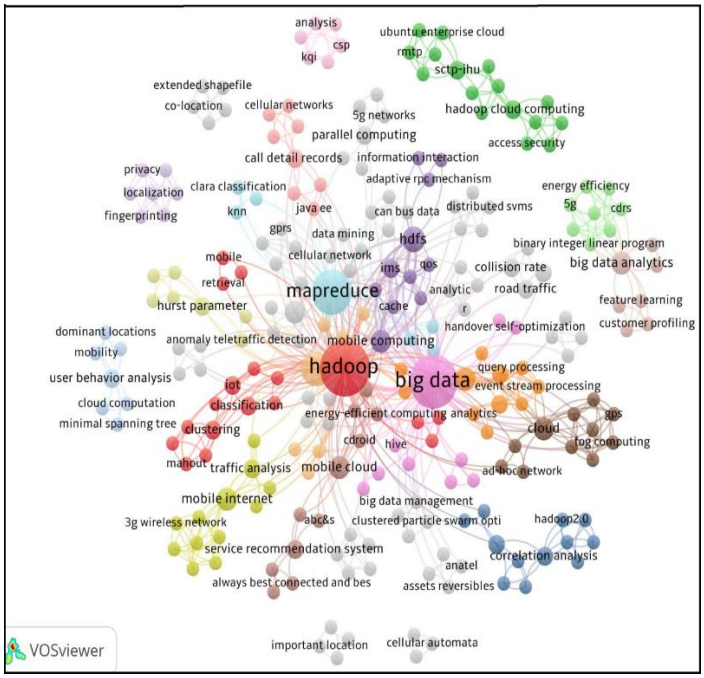

Fig.7. Example of Author keywords co-occurrence network visualization using VOS viewer.

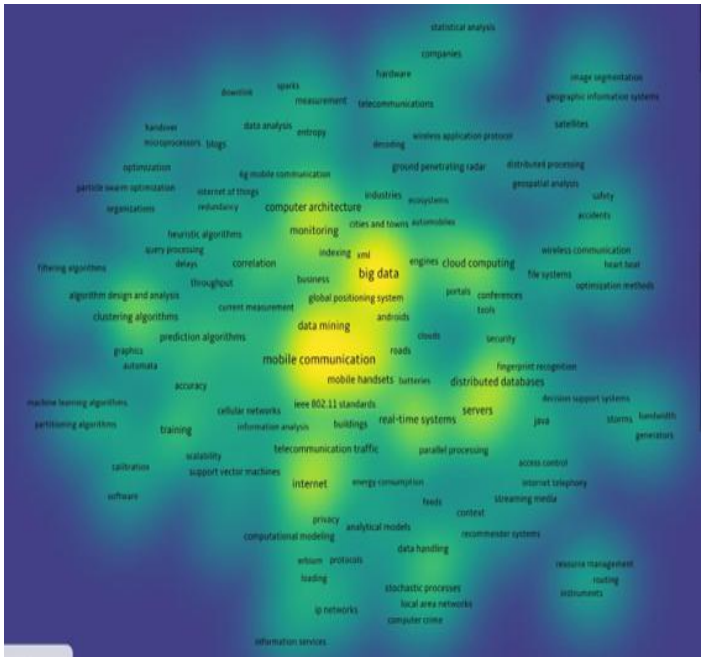

Fig.8. Example of IEEE controlled Terms co-occurrence density visualization using VOSviewer

6.6 LIST OF TOP 25 TOPICS KEYWORDS INDEXED IN IEEE INSPEC CONTROLLED TERMS AND PROVIDED BY AN AUTHOR

The Table.4 having a list of top 25 author keywords and IEEE INSPEC controlled terms (INSPEC is a major indexing database for scientific and technical subjects) which has the highest total link strength (TLS). These keyword co-occurrence statistics is an output of VOSviewer where links are the count of co-occurrence of two keywords and the total link strength of each keyword or terms shows the co-occurrence links with all other keywords in the total articles; more TLS means more frequently keywords being used in the article. In table 4, it is observed that, three keywords are common and have maximum TLS listed in INSPECT and author keywords that are 'big data', 'mobile computing' and 'cloud computing'.

7. CONCLUSION

This paper provided a summary offers a comprehensive exploration of the use of Apache Hadoop in telecommunication big data processing. Leveraging big data analytics in the telecommunications industry is leading to various benefits and improvements in telecom operations, services, and decision-making. BDA not only helps into the challenging task of data analytics for network optimization but also improvise the quality of telecom services. This study has explored the cases studies of Hadoop deploying in telecommunication big data processing efficiently. Hadoop Ecosystem has ability to integrate wide range of tools that cater to different needs in big data analytics. The data charting given in table 5 (Appendix A) is the scope of studies of 77 research paper which highlight the trending Hadoop tools application in telecommunication big data. the This adaptability ensures that Hadoop remains relevant despite evolving big data technologies. Mobile cloud computing is trending topic is discussed in most of the research studies for managing and storage of the centralized and decentralized telecom big data resources. The Scientometrics study evaluates the authors keywords, article citations, annual scientific production on telecom big data analysis with Hadoop tools. It has been observed that the system designed for real time wireless data analysis has major studies carried out in analysing the CDR pattern, cellular network behaviour, GPS logs, mobile internet traffic (4G/5G) as well as detecting of anomaly tele-traffic and propose cellular quality services. Data processing with Hadoop MapReduce implemented in maximum studies also in cluster with Apache Mahout and spark interface. Similarly, adoption of Hadoop HDFS for data storage has advantage over Apache HBase in terms of data read / write process. Some of the studies have implemented Apache pig for data manipulation operation and Hive, for simplified data analysis with SQL interface in Hadoop. Hadoop is a popular open-source framework for distributed storage and processing of large datasets. This paper provides the extensive coverage for Hadoop tools practice for telecommunication big data. Although; some studies have concluded the limitation of Hadoop ecosystem, despite the acknowledged limitations of the Hadoop ecosystem tools, they remain a priority for big data analytics. Many of the studies emphasised on improving the usage of Hadoop clusters effectively indicates a recognition that simply deploying Hadoop is not enough; optimizing its usage is crucial. This could involve tuning configurations, implementing best practices, and integrating additional tools or frameworks to complement Hadoop's capabilities. The future research can be carried out by considering publication on telecom big data in many other major databases.

Table.4. List of top most used INSPEC indexing terms and Author Keywords with their respective weightage of Links, total link Strength and the number of Articles published in IEEE database from year 2011 to 2022

INSPEC Controlled Terms	Links	TLS	Articles	Author Keywords	Links	TLS	Articles
Big Data	94	193	35	Hadoop	82	102	25
Parallel Processing	91	170	29	Big data	80	103	23
Mobile Computing	83	165	30	Mapreduce	46	65	15
Cloud Computing	79	139	24	Cloud Computing	26	33	8
Data Handling	65	105	17	Cloud	18	19	3
Data Analysis	60	102	17	Call detail records	17	17	3
Telecommunication Traffic	49	90	18	HDFS	15	21	4
Public domain software	40	53	6	Mobile Computing	14	18	3
Distributed processing	39	43	7	Big data Analytics	14	14	3
Cellular radio	37	60	11	Mobile internet	13	14	3
Telecommunication computing	36	58	9	Mobile cloud	12	12	3
5g mobile communication	27	37	5	Distributed Computing	12	13	3
Data mining	26	35	6	Hadoop cloud computing	10	10	2
Statistical analysis	25	32	4	SCTP-IHU	10	10	2
Pattern clustering	25	33	5	Hurst parameter	9	9	2
Distributed databases	24	33	5	Telecommunication	9	10	2
Mobile communication	24	28	4	Correlation analysis	9	9	2

QoS	23	27	4	IMS	9	13	2
4g Mobile Communication	21	29	5	Data processing	8	9	2
Customer services	23	27	4	5g	8	8	1
Mobility management (mobile radio)	21	34	5	Encryption and decryption	8	8	1
Query processing	19	27	6	Fog computing	8	8	1
Learning (AI)	19	28	6	GPS	8	8	1
Smart phones	18	20	3	Hadoop database	8	8	1
Recommender system	18	24	5	smartphone	8	8	1

Appendix (A)

Table.5. Data Charting (Scoping reviewed 77 research studies)

Research Title	**Scope of the Study**	**Hadoop Tools Practiced**
RT1. [5]	A case study introduced proactive caching architecture for 5G wireless networks	Cloudera Distributed Hadoop (CHD4)
RT2. [12]	Analysed CDR pattern analysis, cellular network behaviour analysis, traffic flow extraction, data prehending	Hive, HBase
RT3. [13]	designed algorithms framework for network traffic analysis of cellular network	MapReduce, HDFS, HBase
RT4. [14]	Designed an algorithm for heterogeneous large data analysis from various sources in the mobile environment	MapReduce
RT5. [15]	An analytical study of real prototype system consisting of a big-data back-end & smartphone front-end also proposed a framework for a user to localize	HDFS, MapReduce, HBase
RT6. [16]	proposed a cloud-based mobile multimedia recommendation system to reduce network overhead and speedup the recommendation process	MapReduce
RT7. [17]	Proposed a BDA method for deployment of an IoST via a LoRa wireless platform for smart cities	Hadoop YARN, Spark, GraphX
RT8. [18]	Introduced a system to detect several types of anomalies, hardware misconfiguration, infrastructure failures or user driven anomalies in large-scale telecom.	HDFS, Kafka, Spark
RT9. [4]	Introduced a big data platform for a telecom operator centric ecosystem based on Mobile Internet Big Data	HDFS, HBase
RT10. [19]	Presented a system that reduces the skewed communication patter of underlying network	Hadoop MapReduce, Apache Nutch
RT11. [20]	Proposed telecom big data exploration stack to minimize the storage needed to retain data over time & to lower the response time for spatiotemporal data exploration tasks	HDFS, Hive
RT12. [21]	Presented a novel approach to estimate in near real-time to improve user service and to reduce churn rates from set of mobile phone activity logs of customer calls data	HDFS, MapReduce, Big R
RT13. [22]	Analysis of spatio-temporal regions using call records collected from cellular network operators	HDFS, HBase
RT14. [23]	Proposed Anomaly-based Tele-traffic Intrusion Detection System (AT-IDS) design for detecting of anomaly tele-traffic	MapReduce
RT15. [24]	Propose a system that can improve the classification of large-scale remote sensing image and benefit the power of spatial big data concept	MapReduce
RT16. [25]	Presented an architecture for monitoring platform mainly to collect, analyse logs and network flow data related to Android applications lifecycle activities	HBase, HDFS
RT17. [26]	Design improved parallel KNN algorithm to classify text classification of network public opinion data	MapReduce
RT18. [27]	Proposed a cloud-enhanced system architecture for logistics tracking services equipped with 3G/GPS	MapReduce, HBase
RT19. [28]	Proposed framework for Mobile Distributed File System to analysis heterogeneous mobile cluster with varying input workloads in a real.	MapReduce
RT20. [29]	designed a distributed Support Vector Machines architecture for classifying network traffic & to analyse large scale datasets	MapReduce

RT21. [30]	Analysed the overall bandwidth consumption between data centre nodes	Continuous Hive (CHive)
RT22. [31]	Design handover optimization to improve the service quality of the cellular network	HDFS, MapReduce
RT23. [32]	Study to detect the abnormal SIMs from network running IoT	Hadoop with Apache Spark interface
RT24. [33]	Introduce a Hadoop cloud computing connected to mobile or wired networks and employed a way to link low-capacity Linux embedded platform to Hadoop via Ethernet, Wi-Fi or 3G	MapReduce
RT25. [34]	presented an integrated framework with R and Pig for scalable deep analysis and use cases for telecom data	Apache Pig/Hadoop and R
RT26. [35]	design a mobile service recommendation system in the emerging ubiquitous consumer wireless world	Hadoop, Storm, and Kafka
RT27. [36]	proposed platform for AT-IDS, anomaly tele traffic intrusion detection on Hadoop	HDFS, HBase
RT28. [37]	Proposed a hybrid recommendation algorithm improves the efficiency and accuracy based on collaborative filtering and Word2Vec for large dataset	MapReduce, Hive
RT29. [38]	Analysed and implement mobile app for Controller Area Network (CAN) bus data	MapReduce, Hive, SAS, HiveQL
RT30. [39]	characterization of the mobility models from the 6 TB 4G data traffic mobile big data processing	Hadoop
RT31. [40]	Analysed and Summarization of CDR data for criminal investigation	Hive
RT32. [41]	proposed an architecture for traffic prediction by collecting real-time traffic log data of the cellular subscribers from mobile terminals	MapReduce
RT33. [42]	designed a network traffic prediction system to process the real mobile network traffic data	MapReduce
RT34. [43]	Provided geo-processing software for processing and visualization the complex spatiotemporal data analysis	MapReduce, HDFS
RT35. [44]	Proposed a controllable Hadoop-based data centre for telecoms service providers mobile computing through the IMS network	HDFS, YARN
RT36. [45]	proposed an architecture for big data distributed processing on the road traffic data	HBase
RT37. [46]	Proposed an architecture to store and manage GPS log files in a cloud-based platform using a real time dataset	Hive, HDFS MapReduce
RT38. [47]	Architecture of the mobile network traffic analysing system for user mobility and online user distribution	MapReduce
RT39. [48]	analysed the mobile network traffic, examine the basic traffic characteristics and user behaviour characteristics	MapReduce
RT40. [49]	proposed a classifications method on large-scale mobile data utilizes Map Reduce-based distributed computing framework to accomplish Parallelization Naive Bayes	HDFS, MapReduce
RT41. [50]	proposed a 4G LTE system framework for the distributed storage of IP Multimedia Subsystem (IMS) service resources to improvise mobile network operators' quality of service	HDFS, MapReduce
RT42. [51]	Analysis of Performance Management Data from different Network Elements based on the Big Data stack	
RT43. [52]	develop a parallel genetic algorithm for the coverage problem of IoTs with massive nodes in 5G networks	HDFS
RT44. [53]	Presented the cloud infrastructure for mobile conventional and location data	HDFS with Cassandra
RT45. [54]	Analysed the real mobile Internet traffic data based on Traffic Analysis System	MapReduce
RT46. [55]	described methodology detecting anomalous traffic and normal traffic by using the self-similarity analysis in the Anomaly Tele traffic detection Measurement analysis Simulator, ATMSim environment	Hadoop Yarn HDFS MapReduce
RT47. [56]	Discussed a Hadoop based framework for ad-hoc mobile cloud computing	MapReduce
RT48. [57]	Key Performance Indicator (KPIs)/ Key Quality Indicators (KQIs) analysis and verification of distinct test case scenarios to progress real time data of mobile network	HBase
RT49. [58]	proposed solution for controlling the vehicle-to-vehicle collision prevention by using GPS location and traffic info.	HBase

RT50. [59]	Proposed a Big data platform framework for telecommunications enterprise development	HDFS, Sqoop Apache (SQL-to-Hadoop)
RT51. [60]	analysed and store the massive online logs generated by users' mobile terminals	Hive
RT52. [61]	Developed a car cloud sensor network platform that support the multidimensional information service with 3G/4G mobile	HDFS
RT53. [62]	Analysed two datasets from a commercial trial mobile network	Apache Hadoop, Mahout
RT54. [63]	proposes a model to characterize of cellular service in network design, service generation and resource allocation analysing big traffic data.	HDFS, MapReduce
RT55. [64]	implemented a real-time video/voice over IP (VVoIP) applications in Hadoop cloud computing system	HDFS, MapReduce
RT56. [65]	Proposed Flow Logs Analysis System which storage, process and compute the massive Mobile Internet flow logs efficiently.	Hadoop cloud storage
RT57. [66]	proposed an analysis system based on the Hadoop cloud computing platform in telecoms companies	HDFS, MapReduce
RT58. [67]	to classify massive mobile web log data used in generating Customer web behaviour	Hadoop with Apache Spark interface
RT59. [68]	develop a distributed end-to-end internet speed test and analysis system for mobile networks	MapReduce
RT60. [69]	Proposed a prediction model based on decision tree, using it to predict the telecommunications equipment failures with the main cause of customer complaints	Hadoop 2.0
RT61. [70]	Analysis of mobile network nodes data for predictive maintenance and troubleshooting purposes live Hadoop cluster using real-world data	hive
RT62. [71]	Analysed user behaviour in 4G mobile networks	
RT63. [72]	Framework to big data processing on the road traffic data with especially weather information	HBase
RT64. [73]	design a precision marketing system on business platform of telecom network package	Hdfs
RT65. [74]	Implemented a real-time video/voice over IP (VVoIP) applications in wireless mobile network	Hadoop cloud computing
RT66 [75]	Analysed the large-scale mobile logs to help mobile service providers	Hdfs, MapReduce
RT67 [76]	Study to evaluate the execution efficiency of a query from call data records	Hive, HDFS, MapReduce
RT68 [77]	Proposed a system to cellular Data Network flow upload analyse components of CDR	HDFS, MapReduce
RT69. [78]	Developed the detection and response mechanism to combat web malware in smartphones	MapReduce/Hadoop HDFS, Apache Mahout
RT70. [79]	Design a performance system for Call Detail Record (CDR) batch processing	MapReduce Hadoop Clus ter
RT71. [80]	developed a mobile trajectory retrieval and prototype to manipulate the real mobile trajectory data set	HBase
RT72. [81]	Proposed a system architecture to reduce storage of huge data-centric applications and perform large scale data processing	MapReduce, HDFS
RT73 [82]	conducted a correlation analysis of mobile data useful for achieving the optimization of resource allocation and communication network service provided by operators	Hdfs
RT74. [83]	Developed and analysis model of mobile user's behaviour pattern	Hdfs
RT75. [84]	presented an extensive method which dynamically adjusts the Remote Procedure Call in HDFS.	Hdfs
RT76. [85]	analysis of mobile phone user behaviour prediction to discover correlation between mobile phone user behaviour and other attributes	Hdfs
RT77. [86]	introduced an integrated scalable data analytics framework for extracting the correlative structure between large input data stream	MapReduce, Hadoop Cluster

REFERENCES

[1] J. Hamilton Ortiz, "*Industry 4.0 - Current Status and Future Trends*", Springer, 2020.

[2] H. Zahid, T. Mahmood, A. Morshed and T. Sellis, "Big Data Analytics in telecommunications: Literature Review and Architecture Recommendations", *Journal of Automatica Sinica*, Vol. 7, No. 1, pp. 18-38, 2020.

[3] C. Zhu, "A Novel Base Station Analysis Scheme Based on Telecom Big Data", *Proceedings of International Conference on High Performance Computing and Communications*, pp. 1076-1081, 2018.

[4] Wenliang Huang, Zhen Chen, Wenyu Dong, Hang Li, Bin Cao and Junwei Cao, "Mobile Internet Big Data Platform in China Unicom", *Tsinghua Science and Technology*, Vol. 19, No. 1, pp. 95-101, 2014.

[5] E. Zeydan, "Big Data Caching for Networking: Moving from Cloud to Edge", *IEEE Communications Magazine*, Vol. 54, No. 9, pp. 36-42, 2016.

[6] R. Rituraj, D.T. Varkonyi, A. Mosavi and A.V. Koczy, "5G for Smart Grids: Review, Taxonomy, Bibliometrics, Applications and Future Trends", *Proceedings of International Conference on Intelligent Engineering Systems*, pp. 275-284, 2023.

[7] H. Arksey and L. O'Malley, "Scoping Studies: Towards a Methodological Framework", *International Journal of Social Research Methodology*, Vol. 8, No. 1, pp. 19-32, 2005.

[8] Z. Munn, M.D.J. Peters, C. Stern, C. Tufanaru, A. McArthur and E. Aromataris, "Systematic Review or Scoping Review? Guidance for Authors When Choosing between a Systematic or Scoping Review Approach", *BMC Medical Research Methodology*, Vol. 18, No. 1, pp. 1-7, 2018.

[9] M.D.J. Peters, C.M. Godfrey, H. Khalil, P. McInerney, D. Parker and C.B. Soares, "Guidance for Conducting Systematic Scoping Reviews", *International Journal of Evidence-based Healthcare*, Vol. 13, No. 3, pp. 141-146, 2015.

[10] A.C. Tricco, "PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation", *Annals of Internal Medicine*, Vol. 169, No. 7, pp. 467-473, 2018.

[11] S. Anderson, P. Allen, S. Peckham and N. Goodwin, "Asking the Right Questions: Scoping Studies in the Commissioning of Research on the Organisation and Delivery of Health Services", *Health Research Policy and Systems*, Vol. 6, No. 1, pp. 1-7, 2008.

[12] D. Jiang, Y. Wang, Z. Lv, S. Qi and S. Singh, "Big Data Analysis based Network Behavior Insight of Cellular Networks for Industry 4.0 Applications", *IEEE Transactions on Industrial Informatics*, Vol. 16, No. 2, pp. 1310-1320, 2020.

[13] J. Liu, F. Liu and N. Ansari, "Monitoring and Analyzing Big Traffic Data of a Large-Scale Cellular Network with Hadoop", *IEEE Network*, Vol. 28, No. 4, pp. 32-39, 2014.

[14] M.S. Hossain, M. Moniruzzaman, G. Muhammad, A. Ghoneim and A. Alamri, "Big Data-Driven Service Composition using Parallel Clustered Particle Swarm Optimization in Mobile Environment", *IEEE Transactions on Services Computing*, Vol. 9, No. 5, pp. 806-817, 2016.

[15] A. Konstantinidis, G. Chatzimilioudis, D. Zeinalipour-Yazti, P. Mpeis, N. Pelekis and Y. Theodoridis, "Privacy-Preserving Indoor Localization on Smartphones", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 27, No. 11, pp. 3042-3055, 2015.

[16] Y. Mo, J. Chen, X. Xie, C. Luo and L.T. Yang, "Cloud-based Mobile Multimedia Recommendation System with User Behavior Information", *IEEE Systems Journal*, Vol. 8, No. 1, pp. 184-193, 2014.

[17] M. Gohar, S.H. Ahmed, M. Khan, N. Guizani, A. Ahmed and A. Ur Rahman, "A Big Data Analytics Architecture for the Internet of Small Things", *IEEE Communications Magazine*, Vol. 56, No. 2, pp. 128-133, 2018.

[18] L. Rettig, M. Khayati, P. Cudre-Mauroux and M. Piorkowski, "Online Anomaly Detection over Big Data Streams", *Proceedings of International Conference on Big Data*, pp. 1113-1122, 2015.

[19] M.V. Neves, C.A.F. De Rose, K. Katrinis and H. Franke, "Pythia: Faster Big Data in Motion through Predictive Software-Defined Network Optimization at Runtime", *Proceedings of International Parallel and Distributed Processing Symposium*, pp. 82-90, 2014.

[20] C. Costa, G. Chatzimilioudis, D. Zeinalipour-Yazti and M.F. Mokbel, "Efficient Exploration of Telco Big Data with Compression and Decaying", *Proceedings of International Conference on Data Engineering*, pp. 1332-1343, 2017.

[21] E. Diaz-Aviles, "Towards Real-Time Customer Experience Prediction for Telecommunication Operators", *Proceedings of International Conference on Big Data*, pp. 1063-1072, 2015.

[22] V. Kolar, "People in Motion: Spatio-Temporal Analytics on Call Detail Records", *Proceedings of International Conference on Communication Systems and Networks*, pp. 1-4. 2014.

[23] H.D.J. Jeong, W. Hyun, J. Lim and I. You, "Anomaly Teletraffic Intrusion Detection Systems on Hadoop-based Platforms: A Survey of Some Problems and Solutions", *Proceedings of International Conference on Network-based Information Systems*, pp. 766-770, 2012.

[24] I. Chebbi, W. Boulila and I.R. Farah, "Improvement of Satellite Image Classification: Approach based on Hadoop/MapReduce", *Proceedings of International Conference on Advanced Technologies for Signal and Image Processing*, pp. 31-34, 2016.

[25] A. Lahmadi, F. Beck, E. Finickel and O. Festor, "A Platform for the Analysis and Visualization of Network Flow Data of Android Environments", *Proceedings of International Symposium on Integrated Network Management*, pp. 1129-1130, 2015.

[26] S. Du and J. Li, "Parallel Processing of Improved KNN Text Classification Algorithm based on Hadoop", *Proceedings of International Conference on Information, Communication and Networks*, pp. 167-170, 2019.

[27] A. Ben Ayed, M. Ben Halima and A.M. Alimi, "Big Data Analytics for Logistics and Transportation", *Proceedings of International Conference on Advanced Logistics and Transport*, pp. 311-316, 2015.

[28] I. George, C.A. Chen, R. Stoleru and G.G. Xie, "Hadoop MapReduce for Mobile Clouds", *IEEE Transactions on Cloud Computing*, Vol. 7, No. 1, pp. 224-236, 2019.

[29] D. Le Quoc, V. D'Alessandro, B. Park, L. Romano and C. Fetzer, "Scalable Network Traffic Classification using Distributed Support Vector Machines", *Proceedings of International Conference on Cloud Computing*, pp. 1008-1012, 2015.

[30] B. Theeten and N. Janssens, "CHive: Bandwidth Optimized Continuous Querying in Distributed Clouds", *IEEE Transactions on Cloud Computing*, Vol. 3, No. 2, pp. 219-232, 2015.

[31] C.L. Lee, W.S. Su, K.A. Tang and W.I. Chao, "Design of Handover Self-Optimization using Big Data Analytics", *Proceedings of International Symposium on Network Operations and Management Symposium*, pp. 1-5, 2015.

[32] T. Zhang, H. Li, L. Xu, J. Gao, J. Guan and X. Cheng, "Comprehensive IoT SIM Card Anomaly Detection Algorithm based on Big Data", *Proceedings of International Conferences on Ubiquitous Computing and Communications (IUCC) and Data Science and Computational Intelligence (DSCI) and Smart Computing, Networking and Services*, pp. 602-606, 2019.

[33] B.R. Chang, H.F. Tsai, Z.Y. Lin and C.M. Chen, "Access Security on Cloud Computing Implemented in Hadoop System", *Proceedings of International Conference on Genetic and Evolutionary Computing*, pp. 77-80, 2011.

[34] M. Wang, S.B. Handurukande and M. Nassar, "RPig: A Scalable Framework for Machine Learning and Advanced Statistical Functionalities", *Proceedings of International Conference on Cloud Computing Technology and Science*, pp. 293-300, 2012.

[35] I. Ganchev, Z. Ji, M. O'Droma and L. Zhao, "Smart Recommendation of Mobile Services to Consumers", *IEEE Transactions on Consumer Electronics*, Vol. 63, No. 4, pp. 499-508, 2017.

[36] I.R. Lee, S.K. Ye and H.D.J. Jeong, "Detecting Anomaly Teletraffic using Stochastic Self-Similarity based on Hadoop", *Proceedings of International Conference on Network-based Information Systems*, pp. 282-287, 2013.

[37] Y. Xiao and Q. Shi, "Research and Implementation of Hybrid Recommendation Algorithm based on Collaborative Filtering and Word2Vec", *Proceedings of International Symposium on Computational Intelligence and Design*, pp. 172-175, 2015.

[38] I. Nkenyereye and J.W. Jang, "Integration of Big Data for Querying Can Bus Data from Connected Car", *Proceedings of International Conference on Ubiquitous and Future Networks*, pp. 946-950, 2017.

[39] W. Sun, D. Miao, X. Qin and G. Wei, "Characterizing User Mobility from the View of 4G Cellular Network", *Proceedings of International Conference on Mobile Data Management*, pp. 34-39, 2016.

[40] S. Khan, H. Azmi, F. Ansari and S. Dhalvelkar, "Simple Implementation of Criminal Investigation using Call Data Records (CDRs) through Big Data Technology", *Proceedings of International Conference on Smart City and Emerging Technology*, pp. 1-5. 2018.

[41] W. Jiewu, F. Wentao, H. Chunjing and Z. Xing, "User Traffic Collection and Prediction in Cellular Networks: Architecture, Platform and Case Study", *Proceedings of International Conference on Network Infrastructure and Digital Content*, pp. 414-419, 2014.

[42] H. Cui, Y. Yao, K. Zhang, F. Sun and Y. Liu, "Network Traffic Prediction based on Hadoop", *Proceedings of International Symposium on Wireless Personal Multimedia Communications*, pp. 29-33, 2014.

[43] I. Abdul, M. Alkathiri and M.B. Potdar, "Geospatial Hadoop (GS-Hadoop) an Efficient Mapreduce based Engine for Distributed Processing of Shapefiles", *Proceedings of International Conference on Advances in Computing, Communication and Automation*, pp. 1-7, 2016.

[44] Y. Seraoui, M. Bellafkih and B. Raouyane, "A High-Performance and Scalable Distributed Storage and Computing System for IMS Services", *Proceedings of International Conference on Cloud Computing Technologies and Applications*, pp. 335-342, 2016.

[45] I.J. Lee, "Big Data Processing Framework of Road Traffic Collision using Distributed CEP", *Proceedings of International Symposium on Network Operations and Management*, pp. 1-4, 2014.

[46] Y. He, F. Zhang, Y. Li, Jun Huang, Ling Yin and Chengzhong Xu, "Multiple Routes Recommendation System on Massive Taxi Trajectories", *Tsinghua Science and Technology*, Vol. 21, No. 5, pp. 510-520, 2016.

[47] Chao Dong, Shuo Zhang, Zhenming Lei, Jie Yang and Gang Cheng, "Analyzing GPRS Mobile Network Traffic with Map Reduce", *Proceedings of International Conference on Wireless Communications and Networking*, pp. 129-134, 2013.

[48] D. Peng, Yuanyuan Qiao and J. Yang, "Analyzing Traffic Characteristics between Backbone Networks based on Hadoop", *Proceedings of International Conference on Cloud Computing and Intelligence Systems*, pp. 149-153, 2014.

[49] X. Wang, B. Sheng, L. Xue and Z. Xiao, "Classification of Customer Requirements on Map Reduce-based Naive Bayes", *Proceedings of International Conference on Big Data Analysis*, pp. 1-4, 2016.

[50] Y. Seraoui, B. Raouyane and M. Bellafkih, "An Extended IMS Framework with a High-Performance and Scalable Distributed Storage and Computing System", *Proceedings of International Symposium on Networks, Computers and Communications*, pp. 1-6, 2017.

[51] D. Martinez-Mosquera, R. Navarrete and S. Lujan-Mora, "Development and Evaluation of a Big Data Framework for Performance Management in Mobile Networks", *IEEE Access,* Vol. 8, pp. 226380-226396, 2020.

[52] Y. Zhang, W. Yu, X. Chen and J. Jiang, "Parallel Genetic Algorithm to Extend the Lifespan of Internet of Things in 5G Networks", *IEEE Access,* Vol. 8, pp. 149630-149642, 2020.

[53] A.F. Da Conceicao, J. Sanchez, T. Barabasz, A. Mamani-Aliaga, B.G. Dos Santos and M.F. Mendonca, "Open Architecture for Mobile Data Collection using Cloud Computing", *Proceedings of International Conference on Mobile Data Management*, pp. 160-165, 2013.

[54] I. Yang, H. He and Y. Qiao, "Network Traffic Analysis based on Hadoop", *Proceedings of International Conference on Wireless Communications, Vehicular Technology, Information Theory and Aerospace and Electronics Systems*, pp. 1-5, 2014.

[55] H.D.J. Jeong, "Analysis and Detection of Anomalous Network Traffic", *Proceedings of International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, pp. 403-408. 2016.

[56] C. Barca, "A Virtual Cloud Computing Provider for Mobile Devices", *Proceedings of International Conference on Electronics, Computers and Artificial Intelligence*, pp. 1-4, 2016.

[57] Chennabassamma and R. Sumathi, "Analysis and Verification of Key Performance Parameters of Cellular Network on CEMoD Portal", *Proceedings of International Conference on Recent Trends in Electronics, Information and Communication Technology*, pp. 2095-2100, 2017.

[58] N. Sathishkumar and K. Rajakumar, "A Study on Vehicle-to-Vehicle Collision Prevention using Fog, Cloud, Big Data and Elliptic Curve Security based on Threshold Energy Efficient Protocol in Wireless Sensor Network", *Proceedings of International Conference on Recent Trends and Challenges in Computational Models*, pp. 275-280, 2017.

[59] Z. Wang, G. Wei, Y. Zhan and Y. Sun, "Big Data in Telecommunication Operators: Data, Platform and Practices", *Journal of Communications and Information Networks*, Vol. 2, No. 3, pp. 78-91, 2017.

[60] W. Gan, "Design of Network Precision Marketing based on Big Data Analysis Technology", *Proceedings of International Conference on E-Commerce and Internet Technology*, pp. 77-81, 2020.

[61] C.P. Hwang, M.S. Chen and H.F. Wang, "Development of Car Cloud Sensor Network Application Platform", *Proceedings of International Conference on Advanced Information Networking and Applications Workshops*, pp. 975-978, 2016.

[62] I. Si, C.H. Lung, S. Ajila and W. Ding, "An Empirical Investigation of Mobile Network Traffic Data for Resource Management", *Proceedings of International Congress on Big Data*, pp. 291-298, 2016.

[63] W. Sun, X. Qin and G. Wei, "Modelling and Mining the Temporal Patterns of Service in Cellular Network", *China Communications*, Vol. 12, No. 9, pp. 11-21, 2015.

[64] B.R. Chang, H.F. Tsai, Z.Y. Lin, C.M. Chen and C.F. Huang, "Intelligent VVoIP Implementation in UEC Cloud Computing", *Proceedings of International Conference on Machine Learning and Cybernetics*, pp. 519-524, 2012.

[65] Y. Qiao, Z. Lei, J. Yang and G. Cheng, "FLAS: Traffic Analysis of Emerging Applications on Mobile Internet using Cloud Computing Tools", *Proceedings of International Symposium on Wireless Personal Multimedia Communications*, pp. 1-6, 2013.

[66] F. Tongke, "Hadoop-based Data Analysis System Architecture for Telecom Enterprises", *Proceedings of International Conference on Computational and Information Sciences*, pp. 1277-1279, 2013.

[67] R. Kanagasabai, A. Veeramani, H. Shangfeng, K. Sangaralingam and G. Manai, "Classification of Massive Mobile Web Log URLs for Customer Profiling and Analytics", *Proceedings of International Conference on Big Data*, pp. 1609-1614. 2016.

[68] I. Uzun and O. Abul, "End-to-End Internet Speed Analysis of Mobile Networks with MapReduce", *Proceedings of International Symposium on Networks, Computers and Communications*, pp. 1-6. 2016.

[69] Q. Yang, G. Ji and W. Zhou, "The Correlation Analysis and Prediction between Mobile Phone Users Complaints and Telecom Equipment Failures Under Big Data Environments", *Proceedings of International Conference on Advanced Robotics and Mechatronics*, pp. 201-206. 2017.

[70] I. Skracic and I. Bodrusic, "A Big Data Solution for Troubleshooting Mobile Network Performance Problems", *Proceedings of International Convention on Information and Communication Technology, Electronics and Microelectronics*, pp. 472-477, 2017.

[71] S. Jiang, B. Wei, T. Wang, Z. Zhao and X. Zhang, "Big Data Enabled User behavior Characteristics in Mobile Internet", *Proceedings of International Conference on Wireless Communications and Signal Processing*, pp. 1-5. 2017.

[72] I.J. Lee, "Big Data Processing Framework of Learning Weather Information and Road Traffic Collision using Distributed CEP from CCTV Video: Cognitive Image Processing", *Proceedings of International Conference on Cognitive Informatics and Cognitive Computing*, pp. 400-406, 2017.

[73] Q. Liu, "Design and Implementation of Precision Marketing System under Cloud Computing Environment", *Proceedings of International Conference on Smart Grid and Electrical Automation*, pp. 231-234, 2017.

[74] B.R. Chang, H.F. Tsai, C.F. Huang, Z.Y. Lin and C.M. Chen, "Implementation of Mobile Video/Voice Over IP and Access Control on Cloud Computing", *Proceedings of International Conference on Granular Computing*, pp. 89-94, 2011.

[75] I. Koca, I. Ari, U. Kocak, O. Calikus and C. Sezgin, "Parallel and Pipelined Processing of Large-Scale Mobile Communication Data using Hadoop Open-Source Framework", *Proceedings of International Conference on Signal Processing and Communications Applications*, pp. 1-4. 2012.

[76] J.C. Tseng, "A Successful Application of Big Data Storage Techniques Implemented to Criminal Investigation for Telecom", *Proceedings of International Symposium on Network Operations and Management*, pp. 1-3, 2013.

[77] Zhengxiang Ke, Jun Liu, Dazhong He and Qing Yan, "Toward Scalable and High-Performance Data Processing for Cellular Data Network", *Proceedings of International Conference on Communication Technology*, pp. 393-397, 2013.

[78] H. Adas, S. Shetty and W. Tayib, "Scalable Detection of Web Malware on Smartphones", *Proceedings of International Conference on Information and Communication Technology Research*, pp. 198-201, 2015.

[79] I. Agung and A.I. Kistijantoro, "High Performance CDR Processing with MapReduce", *Proceedings of International Conference on Telecommunication Systems Services and Applications*, pp. 1-6, 2015.

[80] I. Lv, X. Wang, F. Huang, J. Yang, T. Wu and Q. Yan, "TREST: A Hadoop based Distributed Mobile Trajectory Retrieval System", *Proceedings of International Conference on Data Science in Cyberspace*, pp. 341-346. 2016.

[81] S. Koley, S. Nandy, P. Dutta, S. Dhar and T. Sur, "Big Data Architecture with Mobile Cloud in CDroid Operating

System for Storing Huge Data", *Proceedings of International Conference on Computing, Analytics and Security Trends*, pp. 12-17, 2016.

[82] B. Liu, "A Study on Correlation Analysis and Application of Communication Network Service", *Proceedings of International Conference on Universal Village*, pp. 1-6, 2018.

[83] R. Hong, "Research on Mobile User Behavior Mining Model based on Big Data", *Proceedings of International Conference on Mechanical, Control and Computer Engineering*, pp. 1-13, 2019.

[84] I. Wang, Y. Yang, J. Zhang and L. Wang, "Remote Procedure Call Optimization of Big Data Systems based on Data Awareness", *Proceedings of International Conference on Parallel and Distributed Processing with Applications, Big Data and Cloud Computing, Sustainable Computing and Communications, Social Computing and Networking*, pp. 441-447, 2020.

[85] G. Hong and G. Nan, "Research and Application of a Multidimensional Association Rules Mining Algorithm based on Hadoop", *Proceedings of International Conference on Parallel and Distributed Processing with Applications, Big Data and Cloud Computing, Sustainable Computing and Communications, Social Computing and Networking*, pp. 636-643, 2021.

[86] S. Robitzsch, F. Zaman, S. Van Der Meer, J. Keeney and G.M. Muntean, "Magnet: Real-Time Trace Stream Analytics Framework for 5G Operations Support Systems", *IEEE Network*, Vol. 31, No. 5, pp. 6-13, 2017.