IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received September 30, 2020, accepted October 5, 2020, date of publication October 8, 2020, date of current version October 21, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3029597

Data Collection in Sensor-Cloud: A Systematic Literature Review

IHSAN ALI[1], (Senior Member, IEEE), ISMAIL AHMEDY[1], (Member, IEEE), ABDULLAH GANI[1,5], (Senior Member, IEEE), MUHAMMAD TALHA[2], (Member, IEEE), MUHAMMAD AHSAN RAZA[3], AND MOHAMMAD HOSSEIN ANISI[4], (Senior Member, IEEE)

[1]Department of Computer System and Technology, Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur 50603, Malaysia
[2]Deanship of Scientific Research, King Saud University, Riyadh 11543, Saudi Arabia
[3]Department of Information Technology, Bahauddin Zakariya University, Multan 60000, Pakistan
[4]School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, U.K.
[5]Faculty of Computing and Informatics, University Malaysia Sabah, Kota Kinabalu 88400, Malaysia

Corresponding authors: Ihsan Ali (ihsanalichd@siswa.um.edu.my), Ismail Ahmedy (ismailahmedy@um.edu.my), and Abdullah Gani (abdullahgani@ums.edu.my)

This work was supported in part by the Faculty of Computer Science and Information Technology through the Postgraduate Research, University of Malaya, under Grant PG035-2016A and Grant LR003-2019. The authors are grateful to the Deanship of Scientific Research, King Saud University for funding this Research Project through Vice Deanship of Scientific Research Chairs (DSRVCH).

ABSTRACT The integration of cloud computing and Wireless Sensor Networks (WSNs) to create Sensor-cloud helps in extending the data processing capability and storage capability of WSNs. Knowing how weak WSNs are with regards to communication ability, how to collect and upload sensory data to the cloud in limited time has become an issue in Sensor-cloud. In the last decade, with increasing interest by researchers in the domain, a considerable amount of research works have been conducted and published in the research domain. The main objective of this study is to systematically review the current research on data collection in Sensor-cloud. Hence, the study also aims at identifying, categorizing, and synthesizing important studies in the field of study. Accordingly, an evidence-based methodology is utilized in this study. By doing so, 43 relevant studies were identified and retrieved to answer the formulated research questions. The systematic methodology offers a methodical and rigorous study selection and evaluation process that is repeatable and precise. The result shows that research on data collection in Sensor-cloud is relatively consistent with stable output in the last five years. Ten proposal contributions were identified with System, Framework, and Algorithm being the most used by the selected studies. In conclusion, key research challenges and future research directions were identified and discussed for researchers to propose effective solutions to the existing challenges. Although research on data collection in Sensor-cloud is gaining some traction in recent years, the works in the domain are not sufficient and concrete proposals are needed to improve data collection.

INDEX TERMS Data collection, sensor cloud, the Internet of Things (IoT) wireless sensor networks (WSN), systematic literature review (SLR).

I. INTRODUCTION

Recently,Wireless Sensor Networks (WSNs) were mostly deployed in many applications, such as forest fire detection [1], agriculture [2], health monitoring [3], and so on. Hence, WSNs used for these applications normally generate a vast amount of data that necessitates to be collected and processed in a minimal time period with relatively low delay. However, sensors are known to have a limited battery with limited computing capability and storage capability to support huge data transmission and processing. This constraint frequently leads to a small network lifetime. As a solution, the data processing and storage abilities of WSNs can be extended using cloud computing [4]. With cloud computing, WSNs performance can be improved, such as service quality, computation latency, energy consumption, and so on. Therefore, the integration of WSNs and cloud computing is termed as Sensor-cloud. The last 10 years have seen quite a number of works on data collection in Sensor-cloud by proposing different solutions on ways to enhance the efficiency and

The associate editor coordinating the review of this manuscript and approving it for publication was Noor Zaman.

effectiveness of data collection. In recent years, many surveys and review papers were published on data collection of sensory data from sensor devices in the research domain (see Section II). In a study by Khan *et al.*, the authors presented a taxonomy of numerous data collection schemes that used sink mobility [5]. The authors identified some unresolved issues in the field of study. Waghmare and Chatur conducted a survey on energy-efficient data collection and routing algorithms in WSNs. The current issues and limitations of the algorithms studied were also discussed [6]. In another study by Yetgin *et al.*, the authors reviewed the current studies in WSNs which includes their design constraints, applications, and lifetime estimation models [7]. However, based on our knowledge, a systematic literature review (SLR) on data collection in Sensor-cloud is non-existent in this research domain. Therefore, in this study, this SLR will try to fill the research gap by the identification, categorization, and synthesisation of important works in the field of study. Therefore, an evidence-based systematic methodology is utilized in this paper to ensure that significant and important studies on data collection in Sensor-cloud in the past 10 years (2011 - 2020) are identified and retrieved. The methodology has a systematic selection and evaluation process with a detailed and repeatable studies selection process. This paper further presents results that is based on the identified selected studies overall demographics and characteristics, the contributions (with regards to data collection in Sensor-cloud) of the selected studies, the evaluation mechanisms utilised by the selected studies, and the performance measures utilised by the selected studies in the research domain. The study main contributions are as follows:

- The conduction of a broad systematic review on data collection in Sensor-cloud.
- The Analysis and synthesisation of current studies in the field of study.
- Identification of the existing research challenges in the field of study and highlighting the areas that need attention from researchers.

The remaining of this study is planned as follows. The related work is presented in Section II. Section III articulates the research method used. Section IV presents the results with respect to the defined research questions (see Section III-A). Section V outlines the general discussion of the SLR study. And lastly, the conclusion is presented in Section VI. The paper is organized as illustrated in figure 1.Abbreviations used in this paper are defined in table 11.

II. RELATED WORK

This section highlights and discussed the existing review and survey studies in the field of study. Highlighting these studies will aid in articulating and solidifying the need for this SLR to be conducted.

Francesco *et al.* conducted a survey on data collection in WSNs [8]. The survey gives a comprehensive taxonomy of WSNs architectures with the role of mobile elements. The authors also outline the data collection process and the existing issues and challenges were also highlighted. Wankhade and chavhan conducted a review on data collection methods. The authors further show the comparative study of various data collection methods and sink nodes data collection methods [9]. In another study by Nair and Jose, the authors conducted a survey on data collection methods and routing algorithms in WSNs. The authors also outline the existing issues and challenges in the field of study [10].

Another study by Khan *et al.* produced a taxonomy of distinct data collection schemes which used sing mobility [5]. The authors identified some unresolved issues in the field of study. Waghmare and Chatur conducted a survey on energy-efficient data collection and routing algorithms in WSNs. The current issues and limitations of the algorithms studied were also discussed [6]. In another study by Yetgin *et al.*, the authors reviewed the current studies on WSNs. Facets such as WSNs design constraints, applications, and lifetime estimation models [7]. A mini review was conducted by Ali *et al.* on data collection in smart communities using sensor cloud [11].

Lastly, other surveys such as [12], [13] also conducted a survey on data collection on WSNs. However, based on the review and survey papers discussed, there are no systematic studies in the field of study. Therefore, this study's objective is to fill this gap. Table 1 list all the review and survey studies with their limitations.

III. RESEARCH METHOD

In conducting an SLR, the identification, evaluation, interpretation, and reporting the research that is associated to a research domain of interest is necessary by a researcher [14]–[16]. In this study, the adoption of an evidence-based searching and study selection procedures was done with the aim of improving transparency. Consequently, to conduct an SLR, a search plan has to be followed which is transparent, fair, and also unbiased. Therefore, the search plan has to guarantee the broadness of the search for assessment [17], [18]. To this time, based on our knowledge, there is no SLR study that rigorously review and analyse the current research on data collection in Sensor-cloud (see Section II). Therefore, the aim of this study is to fill this research gap. To do so, we conduct an SLR by utilizing Kitchenham's methodology [19]. The systematic review procedures is the combination of many stages that have to be completed in a disciplined manner, these stages include the development of a review protocol, conducting a systematic review, analysis of the results, results reporting, results visualization, and finally discussion of the research findings.

A. RESEARCH QUESTIONS

The general objective of this paper is to have some insight of studies that are based on data collection in Sensor-cloud. Hence, to have a comprehensive view of this research domain, the SLR formulated four significant research questions (RQs). These RQs will help in categorizing and

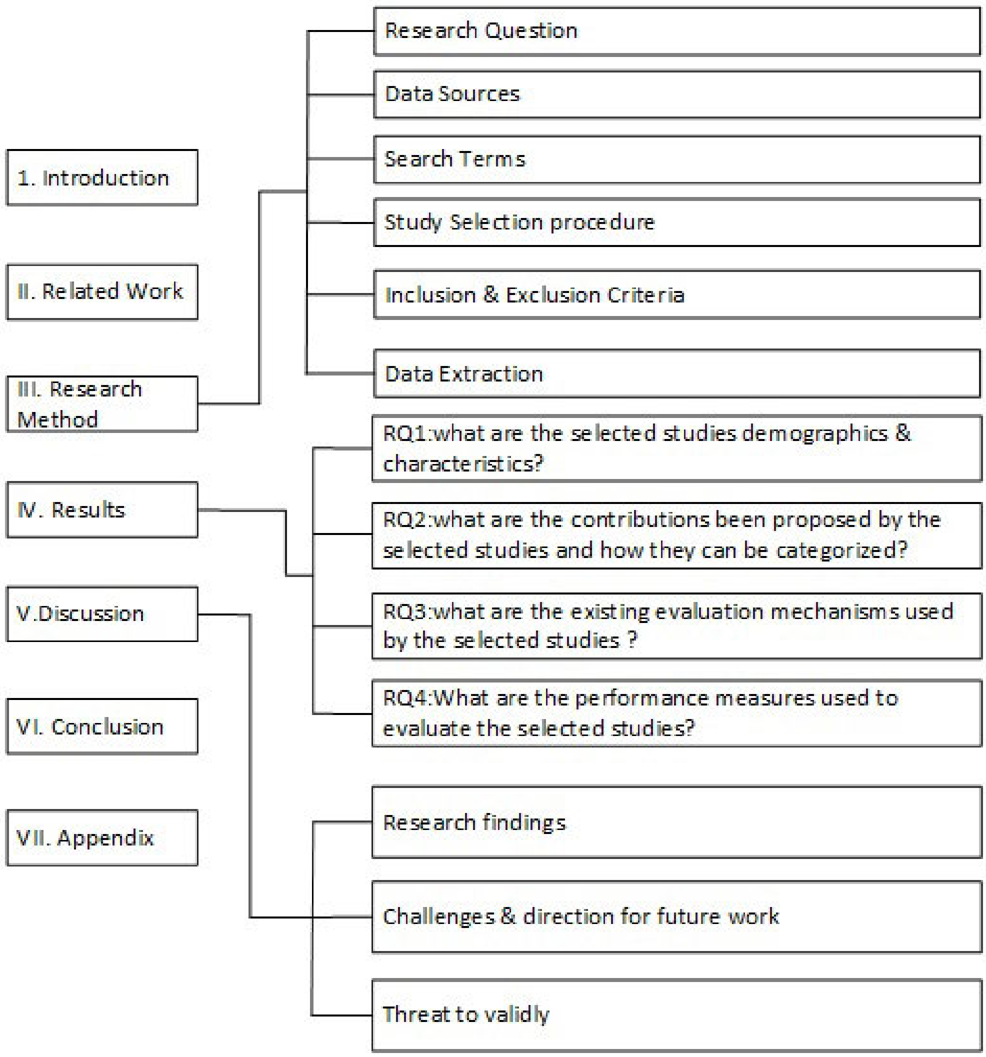

FIGURE 1. **Overview of Research Work.**

understanding the existing research in this domain and further identify the limitations and future research directions in the area of study. The four formulated RQs are presented below.

- RQ1: What are the selected studies demographics and characteristics?
- RQ2: What are the contributions been proposed by the selected studies and how they can be categorized?
- RQ3: What are the existing evaluation mechanisms used by the selected studies?
- RQ4: What are the performance measures used to evaluate the selected studies?

B. DATA SOURCES

In Table 2, the five electronic databases used in this study are highlighted. Hence, in this study, we considered these databases to be the prime data sources for retrieving any possibly relevant studies. On the other hand, Google Scholar was excluded. This is due to the issues of lack of precision of searched results with results overlapping from other data sources. Hence, all the important studies that are in Google Scholar are already retrieved by the other sources.

C. SEARCH TERMS

To successfully search for important studies, search terms are vital. In a study by Keele *et al.* [14], the author recommended Population, Intervention, Comparison, and Outcome (PICO) perspectives. These perspectives was largely utilized by many SLRs and Systematic mapping studies [20]–[22]. However, in this study, with respect to the general foundation of PICO structure, we constructed a generic Search string to sustain the stability of search on many databases. Thus, to conduct the search in the data sources (Table 2), the outline generic Search string serves as a guide.

Generic: (Sensor cloud AND Data collection)

D. STUDY SELECTION PROCEDURE

In this stage (study selection process), the main aim is to effectively identify studies that are significant to the

TABLE 1. Existing review and survey papers.

Ref.	Year	Title	Limitation
[8]	2011	Data collection in wireless sensor networks with mobile elements	• Not focused on data collection in sensor • Not a systematic literature review
[9]	2013	A review on data collection method with sink node in wireless sensor network	• Not focused on data collection in sensor cloud • Not a systematic literature review
[10]	2013	Survey on data collection methods in wireless sensor networks	• Not focused on data collection in sensor cloud • Not a systematic literature review
[5]	2014	A comprehensive study of data collection schemes using mobile sinks in wireless sensor networks	• Not focused on data collection in sensor cloud • Not a systematic literature review
[6]	2014	Energy Efficient Data Collection and Routing Algorithm in Wireless Sensor Network: A Survey	• Not focused on data collection in sensor cloud • Not a systematic literature review
[7]	2017	A survey of network lifetime maximization techniques in wireless sensor networks	• Not focused on data collection in sensor cloud • Not a systematic literature review
[12]	2017	A survey on multipath routing protocols for QoS assurances in real-time wireless multimedia sensor networks	• Not focused on data collection in sensor cloud • Not a systematic literature review
[11]	2018	Data collection in smart communities using sensor cloud: recent advances, taxonomy, and future research directions	• Not focused on data collection in sensor cloud • Not a systematic literature review
[13]	2020	A comprehensive survey on trajectory schemes for data collection using mobile elements in WSNs	• Not focused on data collection in sensor cloud • Not a systematic literature review

TABLE 2. Electronic databases.

Database Name	Link
IEEE Xplore	http://ieeexplore.ieee.org/
Science Direct	http://sciencedirect.com/
ACM	http://dl.acm.org/
Springer Link	http://link.springer.com/
Wiley	http://onlinelibrary.wiley.com/

objectives of our SLR study. In Figure 2, the study selection procedure (SSP) of this study is presented. The study selection process is in three phases, each of these stages was accomplished through an in-depth consensus meeting between the researchers to make sure that there is high confidence with least bias in the study selection process. Hence, if a particular study is in multiple sources, we only take one into consideration with respect to our search order. We initially found 3569 studies through our search. The search results of the study were integrated for different searchers (which are all the authors). The authors also carry out a preliminary screening of the 3569 study collected. This screening is with respect to studies' title, abstract, and conclusion. Hence, for each study screened, two researchers evaluated it to finally resolve if the study would be included. Consequently, for a study that was judge otherwise (the study

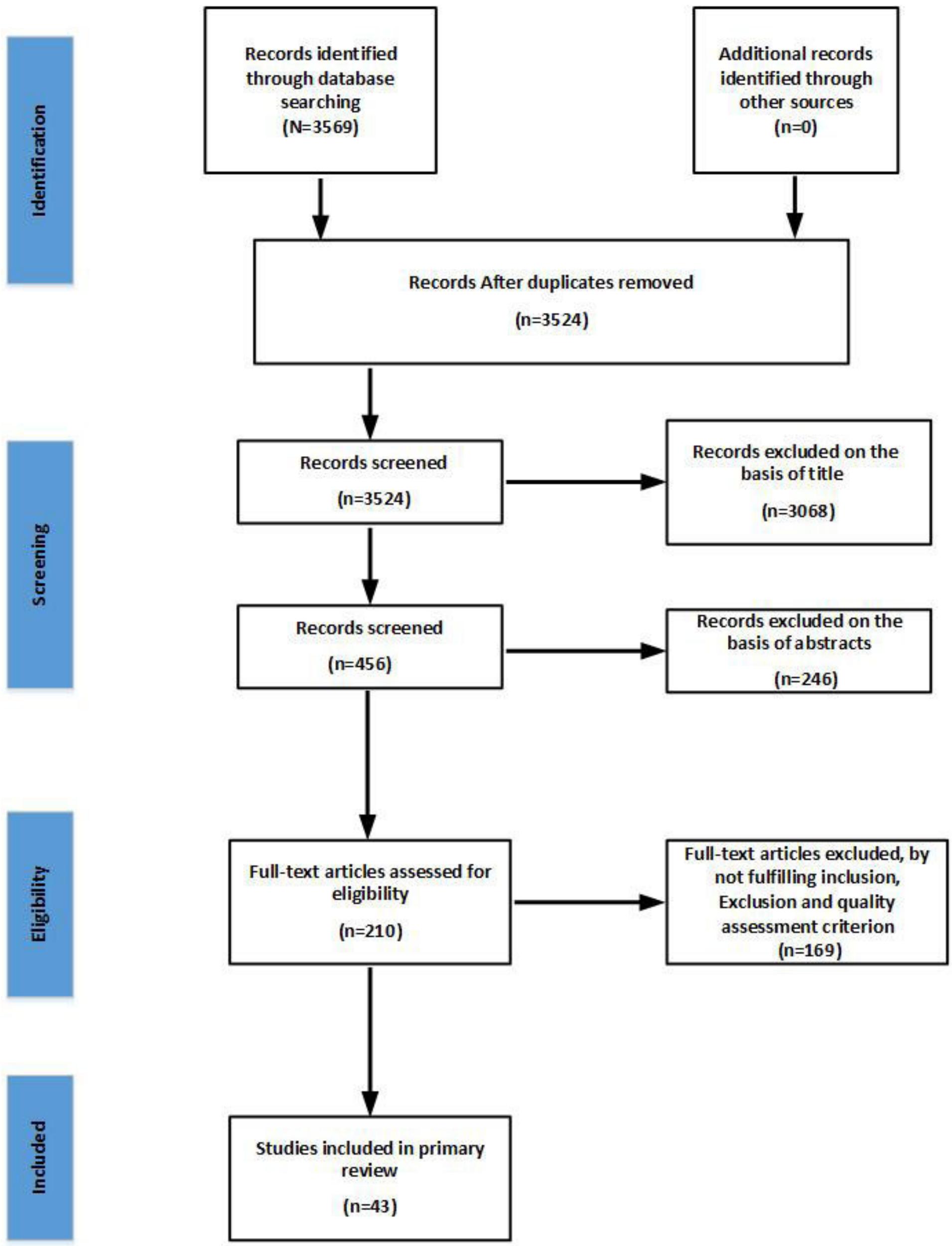

FIGURE 2. Study Selection Procedure.

should be excluded), further discussion was carried out by the two researchers who conducted the evaluation of the studies until an agreement was established. The aim of this screening was to primarily remove studies that were clearly not relevant or they are duplicate or they did not work on data collection in Sensor-cloud.

E. INCLUSION AND EXCLUSION CRITERIA

In the quest to answer the defined RQs in this SLR, we formulated and used well-articulated inclusion (IC) and exclusion (EC) criteria to help in choosing relevant studies from the data sources. The criteria were used on all the studies collected in the different stages of the SSP (see Figure 2). We further set the data collection period from January 2011 to August 2020 (10 years) for studies search, this is to make sure that only the latest studies were included. Moreover, we also include early cited studies, as long as the full study text was available. In Table 3 and 4, we outlined the IC and EC criteria used in this SLR respectively. These criteria were utilized in the second and third stages of the SSP (see Figure 2). In the second stage, the IC and EC criteria were used based on the studies' titles, abstracts, and conclusions. Thus, 210 out of 456 studies were selected in the second stage. In the third stage, to improve the confidence in studies coverage, we applied a snowballing

TABLE 3. Inclusion Criteria.

Number	Inclusion Criteria
IC1	Studies that are Peer-reviewed
IC2	Studies that focuses on data collection in Sensor-cloud
IC3	Relevant articles that are published from 2011 - 2020
IC4	Studies that have the potential of answering the research questions

procedure on 210 full-text studies examined. On the same note, a backward and forward snowballing was conducted. To conduct backwards snowballing, the researchers search through the study reference list and remove studies that do not meet the criteria of this study. For forward snowballing, the researchers analysed the studies based on the studies' citing the study being examined. With this, each study citing a particular study is examined. Therefore, in this study, we consider the inclusion and exclusion of a study based on the IC and EC criteria in Table 3 and Table 4 respectively and the quality attributes outlined in Section III-F. Hence, both criteria were used concurrently to the full-texts of all the 210 studies. Lastly, 43 studies were finally selected for this study

F. QUALITY ASSESSMENT CRITERIA

Quality assessment (QA) is critical and highly important in every SLR. QA of the studies was conducted in the third stage of the SSP. The inclusion and exclusion with the QA criteria were used to the retrieved studies in the second stage of the SSP. 210 studies were collected by the researchers in the third stage where each study was examined by the researchers to remove bias.

Consequently, to evaluate the quality of the selected articles, we designed a questionnaire. The design questionnaire was inspired by earlier systematic studies [21], [23]. As a result, a scale of 1-4 served as the final quality score for a particular article.

1) QA1: The paper gives a contribution to data collection in Sensor-cloud. The possible answers were "Yes (+1)", "Partially (+0.5)", and "No (+0)". 2) QA2: The paper provides a sufficient literature review of the research domain. The possible answer was "Yes (+1)", "Partially (+0.5)", and "No (+0)". 3) QA3: The paper clearly defined the study goals and objectives. The possible answers were "Yes (+1)", "Partially (+0.5)", and "No (+0)". 4) QA4: The contributions and limitations of the paper are clearly stated. The possible answers were "Yes (+1)" and "No (+0)".

G. DATA EXTRACTION

After the second stage of the SSP, the selected articles were then analysed by the review teams. Therefore, each article's full text was analysed by at least two researchers. As a result, vital information was extracted to a data extraction form. The form was composed of key list of items. These items are as follows.

- Title
- Publication year
- Publication venue
- Type of contribution
- Evaluation mechanism
- Performance measures used for evaluation
- Citation count of an article

IV. RESULTS

The results with respect to the RQs of this study are presented in this section.

A. RQ1:WHAT ARE THE SELECTED STUDIES DEMOGRAPHICS AND CHARACTERISTICS?

From the 210 studies that were examined based on all the defined criteria, 169 studies were removed while 43 were finally selected for this study. We intensely and critically analysed the 43 selected studies in order to answer all the RQs presented in Section III-A. In Table 5, all the selected studies are outlined in detail.

1) PUBLICATION OVER TIME

From Figure 3, we present the total number of studies that were published based on the year of publication (2011 – 2020). In the last 10 years, there is a considerable amount of attention given to the field of study by researchers at a progressive passion. We observed that 2011 was the least active year with zero studies published. In other words, there is no study published in that year. However, throughout the years, we have seen an increased interest from researchers, particularly from 2016 – 2020. This can be explained by acknowledging the build-up that occurs from 2012 to 2015 where a stable number of studies have been published, with 11 key studies published in those years. In these years (2012 - 2015), key works have been published, such as S14, S20, S21, S28, S39, and S40 that serve as the foundation for new and veteran researchers to contribute to this new and interesting research field. The reader will also observe that in the year 2017 and 2019, there are many studies published in comparison to the rest of the years with seven studies each. This could be explained by the fact that some of the most popular high ranked Journal and Conference have produced some studies this year. Journal and Conference like Transactions on Industrial Informatics and Conference on Wireless Sensor Networks. In 2020, a conclusion cannot be driven due to our search cap (Section III). Hence, the year has to end for us to know the total number of studies published. In general, despite a slow start in the early years (2011 - 2015), the research activity in the field of study continues to gain

TABLE 4. Exclusion Criteria.

Number	Exclusion Criteria
EC1	Studies that are not written in the English language
EC2	Studies that are not associated with the RQs
EC3	Gray articles; for instance, articles without key information, like publication date/type, issue numbers, and volume were excluded
EC4	Duplicate articles (latest study is included in a situation where multiple studies on the same theme are available). The rest are excluded
EC5	Studies with unclear results and findings

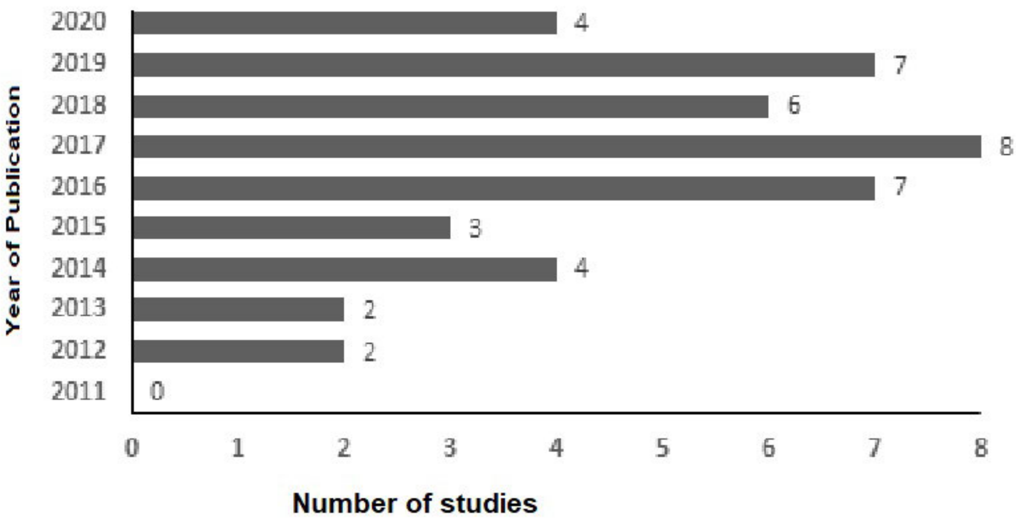

FIGURE 3. Number of articles published per year.

momentum with stable growth, mainly in the last 5 years (2016 to 2020).

2) PUBLICATION CHANNEL AND QUALITY SCORES

In Table 5, we listed the publication channels, publication year, and citation count for each study. Generally, five different publication channels were identified, which are Journal, Conference, Symposium, Workshop, and Magazine. We observed that most of the studies were published in Conferences with 19 studies (44.19%) of the selected studies, 14 studies (32.56%) published in Journals, 6 studies (13.95%) were published in Symposiums, 3 studies (6.98%) were published in Workshops, and lastly, 1 study (2.32%) was published in Magazine (further presented in Figure 4). With this, the general quality of the selected studies is relative, because only 32.56%of the selected studies were published in Journals. Even though it is not a bad number, hence, more quality Journal publications are needed to improve the quality of research in the research domain. We also examined the selected studies for quality based on our quality criteria in Section III-F. In Table 5, we presented the quality score for each study. The results of the quality analysis demonstrate that all studies score more than 1. Also, only four studies score 2 which are S8, S9, S13, and S30. Ten studies score 4 (S1, S2, S3, S4, S5, S11, S26,S34, S42 and S43) and ten studies score 3.5 (S6, S12, S23, S24, S25, S31, S33, S35, S36, and S40).

3) PUBLICATION SOURCE

With respect to the publication sources, Table 7 classifies all the studies based on their publication sources. This classification will aid in finding the publication sources that produce more studies in the field of study for the last decade. Additionally, we also present the publishers of each publication source. In total, 37 sources that published the selected studies were identified. Transactions on Industrial Informatics, Internet of Things Journal, Conference on Wireless Sensor Networks, and Global Communication Conference were the top contributors with 2 publications each, respectively. We also found that most of the studies published in the top publication sources (i.e. S5, S26, S31, S33, and S36) have a high-quality score of 3.5 and above based on the quality assessment conducted in Table 6. From Table 7, we found five publishers which are IEEE with 24 publication sources, followed by ACM (6), Springer (3), Science Direct (3), and Wiley (1). Furthermore, Figure 4 presents the publication channels. From the figure (Figure 4), one can see that majority of the selected studies were published in Conferences Journals with

TABLE 5. Overview of selected studies.

Identifier	Study reference	Publication year	Publication channel	Citation count	Contribution
S1	[21]	2020	Journal	3	Protocol
S2	[22]	2020	Journal	102	Framework
S3	[4]	2020	Journal	23	Model
S4	[23]	2017	Journal	1	Framework
S5	[2]	2019	Conference	0	System
S6	[24]	2012	Conference	57	Framework
S7	[25]	2019	Conference	0	Method
S8	[26]	2016	Conference	16	Algorithm
S9	[27]	2017	Symposium	21	Algorithm
S10	[28]	2016	Symposium	2	System
S11	[29]	2018	Conference	1	Investigation
S12	[30]	2012	Workshop	26	System
S13	[31]	2013	Symposium	5	Approach
S14	[32]	2015	Workshop	28	System
S15	[33]	2017	Workshop	15	System
S16	[34]	2016	Conference	2	Protocol
S17	[35]	2016	Conference	14	Framework
S18	[36]	2016	Symposium	11	System
S19	[37]	2013	Conference	34	Architecture
S20	[38]	2014	Conference	3	Model
S21	[39]	2015	Conference	3	Algorithm
S22	[40]	2020	Journal	0	Framework
S23	[41]	2018	Conference	4	Topology
S24	[42]	2019	Journal	14	Framework
S25	[43]	2017	Journal	56	Algorithm
S26	[44]	2018	Journal	43	Algorithm
S27	[45]	2014	Journal	16	System
S28	[46]	2014	Conference	6	Approach
S29	[47]	2019	Conference	0	System
S30	[48]	2017	Conference	6	System
S31	[49]	2018	Conference	5	Model
S32	[50]	2019	Conference	0	System
S33	[51]	2019	Journal	31	Method
S34	[52]	2016	Conference	8	Framework
S35	[53]	2017	Magazine	37	Investigation
S36	[54]	2019	Journal	26	Protocol
S37	[55]	2018	Symposium	1	System
S38	[56]	2018	Journal	30	Approach
S39	[57]	2015	Conference	10	Model
S40	[58]	2014	Symposium	5	Framework
S41	[59]	2017	Conference	10	System
S42	[60]	2017	Journal	200	Framework
S43	[61]	2016	Journal	164	Framework

19 studies, followed by Journals, Symposiums, Workshops, and Magazine with 12, 6, 3, and 1, respectively.

4) CITATION IMPACT

From Figure 5, the number of citations of all the selected studies were given. Hence, the citation count of each individual study is retrieved from Google scholar. Therefore, the citation count can change at any time. Overall, from our selected studies, we identified 3 study that has more than 100 citations. These studies are [25], [60], [61]. We further find 7 studies with or more than 30 citations. These studies are [27], [40], [46], [47], [54], [56], [59]. Generally, the overall number of citations for the selected studies is 1039, and the average citations per paper is 24.16.

5) GEOGRAPHICAL DISTRIBUTION

Figure 6 gives the top 7 most active countries in the field of study. They are top 7 because only seven countries has 2 or

TABLE 6. Quality evaluation of the selected studies.

Study	QA1	QA2	QA3	QA4	Total score
S1	1	1	1	1	4
S2	1	1	1	1	4
S3	1	1	1	1	4
S4	1	1	1	1	4
S5	1	1	1	1	4
S6	1	0.5	1	1	3.5
S7	1	0.5	1	0	2.5
S8	1	0.5	0.5	0	2
S9	1	0.5	0.5	0	2
S10	1	0.5	1	0	2.5
S11	1	1	1	1	4
S12	1	0.5	1	1	3.5
S13	1	0	1	0	2
S14	1	0.5	1	0	2.5
S15	1	1	0.5	0	2.5
S16	1	1	0.5	0	2.5
S17	1	1	0.5	0	2.5
S18	1	1	0.5	0	2.5
S19	1	0.5	0	0	1.5
S20	1	0.5	0	1	2.5
S21	1	0.5	0	0	1.5
S22	1	0.5	0	1	2.5
S23	1	1	0.5	1	3.5
S24	1	1	0.5	1	3.5
S25	1	1	0.5	1	3.5
S26	1	1	1	1	4
S27	1	1	1	0	3
S28	1	1	1	0	3
S29	1	0.5	1	0	2.5
S30	1	0	1	0	2
S31	1	0.5	1	1	3.5
S32	1	0.5	1	0	2.5
S33	1	0.5	1	1	3.5
S34	1	1	1	1	4
S35	1	1	0.5	1	3.5
S36	1	1	0.5	1	3.5
S37	1	1	0.5	0	2.5
S38	1	1	0.5	0	2.5
S39	1	0.5	1	0	2.5
S40	1	0.5	1	1	3.5
S41	1	0.5	1	0	2.5
S42	1	1	1	1	4
S43	1	1	1	1	4

more than 2 studies. Generally, we identified 19 active countries. These countries are China with 15 studies, followed by India (4), United Kingdom (UK) (3), United States (3), Greece (2), Australia (2), Italy(2), New Zealand (1), Japan (1), South Korea (1), Czech Republic (1), Singapore (1), Germany (1), Taiwan (1), Norway (1), Egypt (1), Romania (1), Malaysia (1), and Serbia (1), respectively. This result is formulated based on the first authors' country of institutions. The result based on Figure 5 shows that China is the most active country.

TABLE 7. Publication Source.

Publication	Publishers	Studies	No.	Percentage
Transactions on Industrial Informatics	IEEE	S33,S36	2	4.7
Internet of Things Journal	IEEE	S22, S26	2	4.7
Conference on Wireless Sensor Networks	Springer	S5,S7	2	4.7
Global Communication Conference	IEEE	S21,S31	2	4.7
Transactions on Sustainable Computing	IEEE	S25	1	4.7
Journal on Selected Areas in Communications	IEEE	S24	1	2.4
Journal of Parallel and Distributed Computing	Science Direct	S3	1	2.4
Future Generation Computer Systems	Science Direct	S2	1	2.4
Digital communications and Networks	Science Direct	S1	1	2.4
International Journal of Communication Systems	Wiley	S4	1	2.4
International Conference on Electronics, Computers, and Artificial Intelligence	IEEE	S23	1	2.4
Transactions on Big Data	IEEE	S38	1	2.4
2nd High Performance Computing and Cluster Technology Conference	ACM	S11	1	2.4
International Conference on I-SMAC	IEEE	S41	1	2.4
International Conference on Software Architecture	IEEE	S30	1	2.4
Consumer Electronics Magazine	IEEE	S35	1	2.4
International Symposium on Computer	IEEE	S37	1	2.4
International Conference on Advanced Technologies Systems and Services in Telecommunications	IEEE	S29	1	2.4
International Conference on Virtual Reality and Intelligent Systems	IEEE	S32	1	2.4
Cloud Computing	IEEE	S27	1	2.4
Workshop on Scientific Cloud Computing	ACM	S14	1	2.4
International Conference on Recent Trends in Information Technology	IEEE	S20	1	2.4
International Symposium on Cluster, Cloud, and Grid Computing	IEEE	S40	1	2.4
The Second Nordic Symposium on Cloud Computing and Internet Technologies	ACM	S13	1	2.4
International Conference on ICT Convergence	IEEE	S19	1	2.4
International Workshop on Data-Intensive Distributed Computing Date	ACM	S12	1	2.4
International Conference on Computer Science and Information Technology	Springer	S6	1	2.4
International Congress on Big Data	IEEE	S39	1	2.4
International Computer Software and Applications Conference	IEEE	S28	1	2.4
International Conference on High Performance Computing and Communications: International Conference on Smart City	IEEE	S34	1	2.4
Symposium on Women in Research	ACM	S10	1	2.4
Asia-Pacific Services Computing Conference	Springer	S8	1	2.4
IEEE Sarnoff Symposium	IEEE	S18	1	2.4
International Conference on Computer Engineering and Systems	IEEE	S16	1	2.4
Symposium on Cloud Computing	ACM	S9	1	2.4
International Workshop on Computer Aided Modeling and Design of Communication Links and Networks	IEEE	S15	1	2.4
International Joint Conference on Neural Networks	IEEE	S17	1	2.4
IEEE Network	IEEE	S43	1	2.4
IEEE Access	IEEE	S42	1	2.4

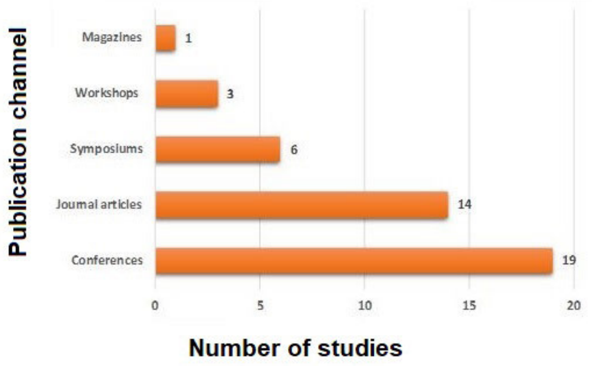

FIGURE 4. Publication channels.

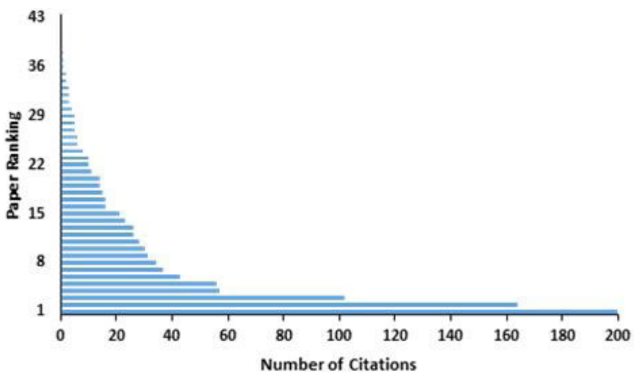

FIGURE 5. Citation Impact.

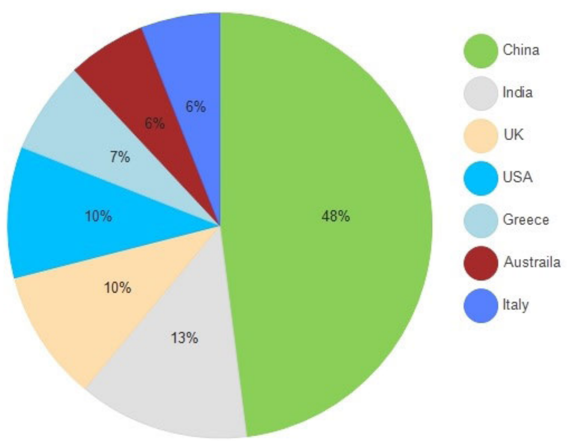

FIGURE 6. Top 7 countries with the most studies.

B. RQ2:WHAT ARE THE CONTRIBUTIONS BEEN PROPOSED BY THE SELECTED STUDIES AND HOW THEY CAN BE CATEGORIZED?

In answering this RQ, we look at the contributions that are proposed by the selected studies in this SLR. Based on our analysis, we have identified 10 contributions which are System with 27.95% of the studies, followed by Framework 23.25%, Algorithm 11.635% , Model 9.30% , Protocol 6.98%, Approach 6.98%, Investigation 4.65%, Method 4.65%, Architecture 2.33% and Topology 2.33%. In this section, the studies with their respective contributions will be discussed in detail. We observed that 12 studies have proposed Systems for data collection in Sensor-cloud. In a study by Zhang *et al.*, an agriculture irrigation system was proposed through the use of sensor-cloud technology in the agricultural sector [2]. The proposed system aids in collecting and the efficient processing of sensing data in agriculture irrigation. The result shows the performance of the proposed system in terms of energy consumption. In another study by Pansare and Bajad, a new system is proposed to help in detecting errors in a large sensor data during transmission [31]. The system shows some promise. Ward and Barker introduced a scalable distributed data collection cloud system [33]. The proposed system helps in collecting sensor data to a cloud system. The experimental result reveals some improvement. In a study by Li *et al.*, a cloud-based data streaming system named WaggleDB was proposed [35]. The system is proposed to address the challenges of data collection, data availability, efficiency, and so on, that are in cloud data infrastructure. The result shows some improvement. Charalampidis *et al.* introduced a fog-enabled IoT system utilized for sensory data collection. The experimental result shows some promise where the system reduces energy consumption [36]. In another study by Soultanopoulos *et al.*, the authors presented a system implementation and IoT service architecture for a gateway service running on smart devices [39]. The system is built to help in the processing of sensor data prior to their transfer to the cloud. The result indicates that the proposed system supports fast data collection with real-time communication. In a study by Wu *et al.*, the authors proposed a system named Concinnity. The proposed system takes sensor data from the source to the destination via a cloud data repository [48]. A case study was conducted by the authors. The result shows some progress in terms of data anomalies detection. A remote health system was introduced by Stojanovic *et al.* [50]. The proposed system used sensor fusion which allows the processing and examination of IoT data from sensor devices. The result shows that the system increased accuracy. Gesvindr *et al.* proposed a system used for collecting sensor data from smart homes by utilizing TapHome solution [51]. The result shows some promise. Min proposed a multi-network data acquisition system that is based on cloud platform with real-time data update of sensory data [53]. In a study by Wang *et al.*, a system was proposed. The proposed system integrates blockchain technology that regards each mobile database as a block [58]. The proposed system aid in data collection and analysis. The result shows some promise. Maiti *et al.* proposed a data collection system that supports the storage of sensors data to the cloud [62]. The analysis shows some promise.

Furthermore, we observed that 10 studies proposed Framework. A three-layer framework was proposed by Wang *et al.* The framework is used multiple mobile sinks with fog structure [25]. The aim of the proposed framework is to break the bottleneck of data collection from WSNs to the cloud. The framework was compared with various existing traditional solutions. The experimental result reveals that the framework can help in the improvement of throughput and the reduction of transmission delay. Mao *et al.* introduced a framework for a multi-cloud environment named parallel cloud data possession checking scheme [26]. The proposed framework uses a homomorphic verification tag that is generated by a pallier cryptosystem to support unlimited query challenges with support for error localization and data correction. The result of the evaluation demonstrates the security and efficiency of the proposed scheme. Dash *et al.* investigated the key design issues and current challenges for sensor-cloud [27]. Hence, in addressing the identified design issues, the authors introduced a framework that integrates sensor-cloud with sensor networks. In a study by Ghanavati *et al.*, a cloud-based wireless body area networks (WBANs) framework was proposed [38]. The framework is tailored toward real-time health monitoring of patients. The main objective of the proposed framework is to combine both mobile technology and cloud computing to provide services for patients. Based on a case study conducted, the result shows some promise. Liang *et al.* proposed a reliable trust computing mechanism (RTCM) [43]. The framework helps in enhancing the reliability and efficiency of data transfer to the cloud. The result shows some promise. A framework named an efficient privacy-preserving-based data collection and analysis (P2DCA) for Internet of Medical Things(IoMT) applications was proposed by Usman *et al.* [45]. Hence, the proposed framework is aimed to protect against privacy issues when collecting data to the cloud. The result demonstrates that the proposed framework is better than the current schemes. Bhuiyan *et al.* proposed a cloud-enabled remote structural health monitoring (cSHM) framework for remote structural health event detection [55]. The proposed framework helps in facilitating the secured sensor data collection on the cloud. The experiment result reveals that the proposed framework performs very well in terms of data protection. An event-driven data collection framework in sensor-cloud was proposed by Bhunia *et al.* [63]. The framework utilizes fuzzy logic to make certain of efficient data collection. The result shows some promise.

Enzo *et al.* proposed a novel paradigm coined fog of everything (FoE) paradigm. The proposed paradigm integrates fog computing and internet of everything. The result shows a good outcome [60]. In another study by Enzo *et al.*, the authors outlined the major challenges in conducting real-time energy-efficient management of resources at mobile devices and internet-connected data centres [61].

Algorithm was proposed by five studies out of the selected studies, which amount to 11.63% of the selected studies. Traub *et al.* proposed an algorithm that schedules read across a huge amount of sensors based on the data-demands [30]. The algorithm aim is to enhanced data transfer from sensor nodes to sensor-cloud. The experimental result shows that data transmission effectiveness was improved. With the issues of how to upload sense data to the cloud within a small time which turn into a bottleneck of sensor systems, Li *et al.* proposed the utilization of multiple mobile sinks which will aid in uploading data from WSNs to cloud [29]. The authors further designed a new algorithm which will schedule the multiple mobile sinks. Based on simulations conducted, the results demonstrate that the proposed algorithm performs very well with respect to a decrease in data upload latency and minimal energy consumption. Argyriou proposed an algorithm to maximized data delivery to the cloud for post-processing for each sensor in WSN [42]. Based on the simulation, the result demonstrates the algorithm performance with regards to raw sensor data collection to the cloud. With the upload of sensor data to the cloud within a small time becoming a bottleneck, Wang et al, proposed the utilization of multiple mobile sinks to help in data collection [46]. Furthermore, to reduce data delivery latency, the authors proposed a time adaptive schedule algorithm (TASA) for data collection through multiple mobile sinks. The result demonstrates that the proposed algorithm can gather data from WSNs to cloud with limited latency and minimal energy consumption. Hence, makes the sensor-cloud sustainable. Tao *et al.* proposed a secure data collection algorithm named secure data with the goal of addressing security concerns during data transfer [47]. The simulation result reveals that the proposed algorithm is useful when applied for security protection.

From the studies selected, four studies proposed a model for data collection in sensor-cloud. In a study by Wang *et al.*, the authors proposed an edge-based model for data collection. The model works in a way where the data retrieved from WSNs is processed separately by algorithms on edge servers from privacy computing [4]. As highlighted by the authors, the benefits of the proposed model is twofold. The model helps in preserving data privacy and it is implemented by different storage methods. Based on a rigorous experiment and theoretical analysis, the proposed model was validated and has shown some promise. To deal with constant and long-duration monitoring and collection of data from sensors, Grace and Sumalatha introduced a model for sensor-cloud coined senud controller [41]. The proposed model combines both a sensor gateway and a cloud gateway. The result shows that the model supports large data collection efficiently. In a study by Chen *et al.*, data collection scheme was proposed [52]. The proposed scheme protects the collected data from attackers while maintaining data correlation. The simulation result shows that the proposed scheme is very efficient for data collection to the cloud with strong privacy properties. Lawson and Ramaswamy proposed a model for monitoring tradeoff, an architecture that changes based on data quality, and customer data stream best matching cloud service. The authors concluded that their system will perform better.

With respect to Protocol proposals, three studies proposed Protocol out of the selected studies. Wang *et al.* proposed a new scheme, named energy-efficient and anonymous data collection. The proposed scheme is specifically for mobile edge networks (MENs). The aim of the scheme is to get a

TABLE 8. Proposed contributions in the field of study.

Contribution	Studies	No.	Percentage
System	S5, S10, S12, S14, S15, S18, S27, S29, S30, S32, S37, S41	12	27.95
Framework	S2, S4, S6, S17, S22, S24, S34, S40,S42,S43	10	23.25
Algorithm	S8, S9, S21, S25, S26	5	11.63
Model	S3, S20, S31, S39	4	9.30
Protocol	S1, S16, S36	3	6.98
Approach	S13, S28, S38	3	6.98
Investigation	S11, S35	2	4.65
Method	S7, S33	2	4.65
Architecture	S19	1	2.33
Topology	S23	1	2.33

balance between data privacy and energy consumption where the privacy information of sensors is concealed in the course of communication [24]. The result based on simulation shows that the proposed scheme is better than existing schemes with respect to lifetime and energy consumption. A data transfer protocol was proposed by El Mougy and El-kerdany. The protocol was built with some principles of TCP to tackle the issue of data collection from Bluetooth low energy (BLE) sensors to the cloud [37]. Based on a simulation conducted, the result shows that the proposed protocol allows for a reliable data transfer and also reduces energy consumption. In a study by Wang *et al.*, the authors proposed a bidirectional prediction-based underwater data collection protocol [64]. The proposed protocol uses mobile edge elements for data collection from end to cloud. The result shows that the cost of data collection was reduced and bandwidth utilization increases.

With respect to Approach proposals, three studies proposed it out of the selected studies. Gejibo *et al.* investigate the challenges that are related to remote mobile data collection to a central cloud and further proposed an approach that can provide solutions to data protection, sharing, and recovery [34]. The authors conclude that the underlying challenges have to be further investigated. In a study by Nakagawa *et al.*, the authors proposed an approach named m-cloud [49]. The approach aids in collecting sensor data using cloud resources for IoT data. The result indicates some progress. Wang *et al.* proposed a comprehensive trustworthy data collection approach (CTDC) for sensor-cloud systems [59]. Based on an extensive simulation, the result shows that CTDC improved performance in data collection.

Next, With respect to Investigation and Method proposals, they were proposed by two studies each out of the selected studies. A study by Yang *et al.* focuses on studying the data curation problems of IoT big-sensing-data processing on the cloud [56]. The authors highlight the current trends with future research directions for big-sensing-data processing. Abdul Rahman *et al.* conducted a chain of experiments that measure the energy consumption of two IoT sensor nodes that are transferring data to the cloud [32]. Hence, the experimental result will be useful in comparing IoT sensor nodes implementation in both wired and wireless scenarios. The result shows that the wireless connection consumes extra power in comparison to wire connection. Wang *et al.* proposed a data cleaning method that is based on a mobile edge node during data collection to sensor-cloud [54]. The experimental result demonstrates that the proposed method enhanced the efficiency of data cleaning with enhanced data integrity and reliability. Also, the proposed method further decreases the energy consumption of the industrial sensor-cloud system (SCS). In another study by Wang and Wang, to address the issue of bandwidth and real-time data collection issues of large-scale data collected from IoT devices to a central cloud, the authors introduced a new data collection method that uses deep learning technology [28]. Based on an experiment, the proposed method performs effectively.

Architecture and Topology are the least proposed with one study each. Piyare *et al.* proposed an extensive architecture that helps in integrating WSNs with the cloud [40]. Based on an experiment conducted, the result shows some promise. Mihai *et al.* proposed a three-layer topology for smart data monitoring and processing. The aim of the topology is to lessen the sending of raw data to the cloud, hence, to improve the ratio between useful information and noise [44]. The simulation result shows some improvement. Table 8 presents the list of the identified contributions with respect to the studies that proposed them((x-Axes represent number of studies and Y-Axes represent year of publication). In Figure 7, the proposed contributions are presented with respect to the year of publication.

C. RQ3:WHAT ARE THE EXISTING EVALUATION MECHANISMS USED BY THE SELECTED STUDIES?

To fully know the contributions in terms of evaluation mechanism used by the studies selected, we outline and classify the current identified evaluation mechanisms used by the selected studies and further categorized the studies based on which evaluation mechanisms they utilized. In Table 10, the studies with respect to the evaluation mechanism they utilized are presented. In totality, we identify six evaluation mechanisms.

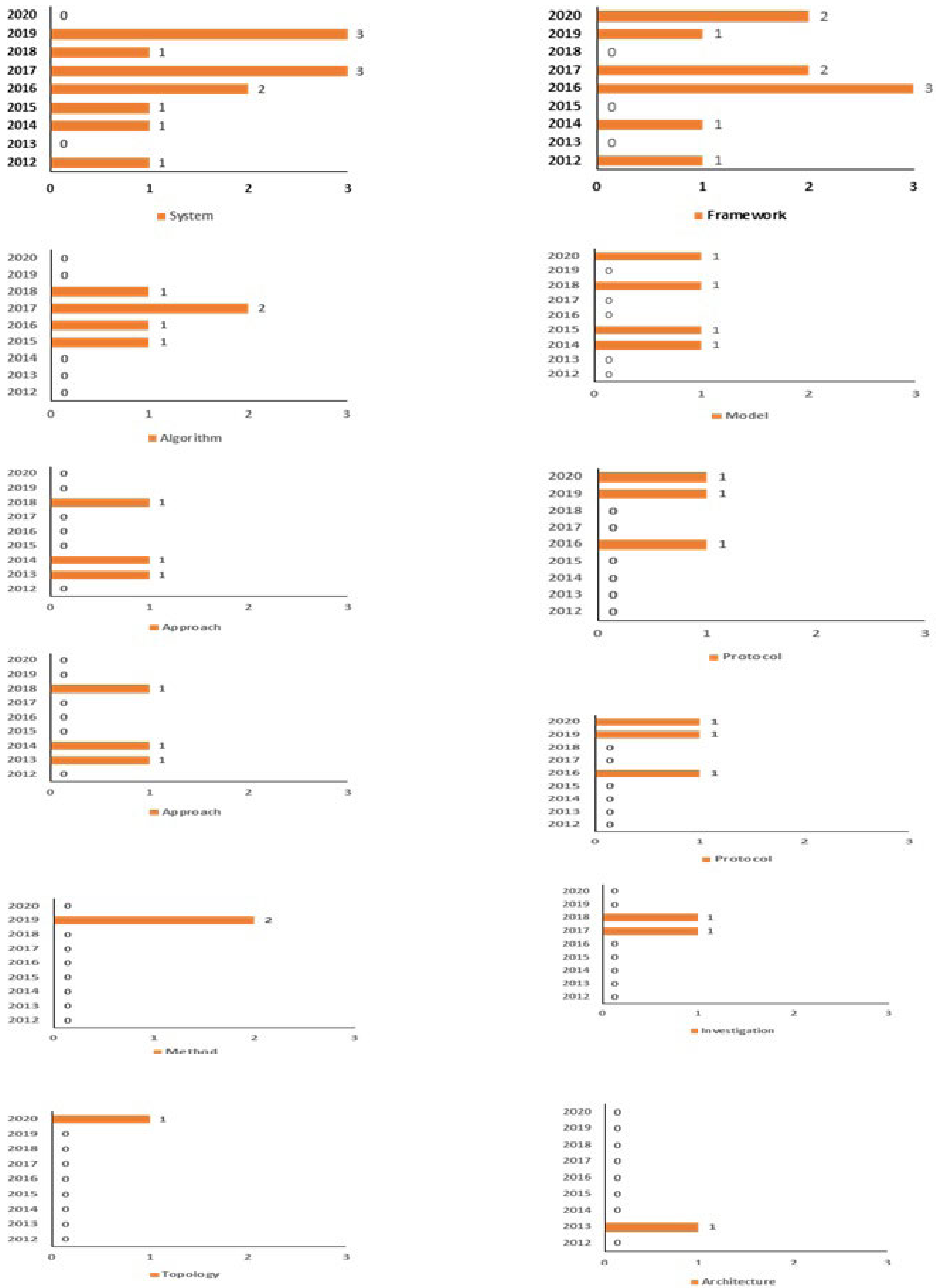

FIGURE 7. Analysis of the proposed contributions with respect to years of publication.

These mechanisms are Experiment with 16 studies, followed by Simulation (13)), Theoretical analysis (6), Hybrid (3), Case study (3)), and None (2). We created a new categorization named Hybrid. This categorization is made when a particular study utilizes more than one evaluation mechanism. For example, in a study by Wang *et al.*, both Experiment and Theoretical analysis were utilized for evaluation [4]. Furthermore, Liang *et al.* also used Theoretical analysis and Simulation combined for evaluation [43]. And lastly, Mao *et al.* also used two evaluation mechanisms for evaluating their proposals [26]. Hence, any study that did not use any evaluation mechanism or did not evaluate a given work, we categorized such study as ''None'' (as shown in Table 9).

Our result shows that Experiments are the most conducted in the field of study with 37.21% of the selected studies utilizing it (as shown in Table 9), followed by Simulation and Theoretical analysis with 30.23% and 13.95% of the studies, respectively. Furthermore, we found out that 6 out of the 16 studies that used Experiment are studies that contribute Systems for their proposals (S12, S14, S15, S18, S29, S30),

TABLE 9. Evaluation mechanisms used by the selected studies.

Evaluation mechanism	No.	Studies	Percentage
Experiment	16	S2, S7, S9, S11, S12, S14, S15, S18, S19, S20, S24, S29, S30, S33, S34, S40	37.21
Simulation	13	S1, S5, S8, S16, S21, S23, S25, S26, S31, S36, S37, S38,S42	30.23
Theoretical analysis	6	S6, S13, S28, S35, S39, S41	13.95
Hybrid	3	S3, S4, S22	6.98
Case study	2	S17, S27,S42	4.65
None	2	S10, S32	4.65

while 3 out of the 13 studies that used Simulation contributes Algorithm for their proposals (S8, S25, and S26). Hence, these are the most contributions among the highlighted most popular evaluation mechanisms.

D. RQ4.WHAT ARE THE PERFORMANCE MEASURES USED TO EVALUATE THE SELECTED STUDIES?

In answering this RQ, we identified 36 (83.72%) studies out of the selected studies that used performance measure for evaluation, while seven studies did not use any performance measure as shown in Table 10 (S6, S10, S28, S32, S35, S39, S41). From the selected studies, various and diverse performance measures where identified. Most of the studies used a combination of more than one performance measure. However, despite the performance measures been so diverse, one of the most conducted and most common is Energy consumption performance measure. We observed that 16 out of the 43 selected studies used energy consumption for their evaluation (S1, S2, S5, S8, S11, S15, S16, S17, S19, S25, S26, S33, S36, and S38). This amounts to 37.21% of the selected studies in this research domain. The dominance of measure such as energy consumption is relative due to the research directions of these studies, where the studies focus on data collection from sensor devices to sensor-cloud. When dealing with data collection in WSNs, energy is always a huge concern. Hence, the domain is dominantly solution proposal driven, where a researcher normally has to propose a new system, model, framework or algorithm, and so on (as shown in Table 8) to help in the collection of data in sensor-cloud. Table 10 highlights the performance measures used by the selected studies.

V. DISCUSSION

In this article, we conducted an SLR on data collection in sensor-cloud. Data collection in sensor-cloud has gained substantial attention from researchers in the past 10 years. Of recent, the collection of sensory data from sensor devices to the cloud has become a crucial and important issue in the research domain. In this section, the results related to the RQs are summarized and discussed through the presentation of the research findings, research challenges, and future work directions.

A. RESEARCH FINDINGS

The key objective of this SLR is to examine the current works in the area of study. To do that, 41 studies were selected based on the adopted methodology in Section III for analysis. Hence, these selected studies were deeply analysed and synthesized to help in addressing the RQs is outlined in table 2. The main findings of this SLR are presented as follows.

Based on our analysis with regards to demographics of the selected studies, we observed some stability with a consistent output of publications in the past 5 years. The result shows that 2011 was the least active year with zero studies published. In other words, there is no study published in that year. Furthermore, we have seen an increased interest from researchers, particularly from 2016–2020. This can be explained by acknowledging the build-up that occurs from 2012 to 2015 where a stable number of studies have been published, with 11 key studies published in those years. Hence, in these years, key works have been published, such as S14, S20, S21, S28, S39, and S40 that serve as the foundation for new and veteran researchers to contribute to this new and interesting research field. We found that most of the studies were published in Conferences with 44.19% of the selected studies, this makes it the highest publication channel amount all the identified publication channels. With this, the general quality of the studies selected is relative, because only 32.56% of the studies selected were published in Journals. Even though it is not a bad number, hence, more quality Journal publications are needed to improve the quality of research in the research domain. With respect to the quality of the selected studies, the result demonstrates that 23.25% of the studies have a total quality score of 4 (which is the highest score), while also 23.25% has a quality score of 3.5. This indicates that the selected studies have some quality, however, the majority of the studies 53.50% score less than 3.5 as there quality score. Hence, more quality studies are needed in the area of study. With respect to the publication source, we identified four sources that are more noticeable. These sources are Transactions on Industrial Informatics, Internet of Things Journal, Conference on Wireless Sensor Networks, and Global Communication Conference with 2 publications each, respectively.

With respect to contributions proposed in the field of study, we identified 10 key contributions. Out of the ten identified contributions by the selected studies, three were found to be more proposed by researchers, which are System, Framework, and Algorithm with 27.95%, 23.25%, and 11.63% respectively. In answering RQ3, we found out that six evaluation mechanisms were utilized by the selected studies, which are Experiment, Simulation, Theoretical analysis, Hybrid,

TABLE 10. Performance measures utilized by the selected studies.

Identifier	Publication year	Quality Score	Performance measures
S1	2020	4	Lifetime, energy consumption
S2	2020	4	Hops count, energy consumption
S3	2020	4	Encoding time, decoding time, time of transmission
S4	2017	4	Users' storage complexity, communication overhead, computation overhead
S5	2019	4	Energy consumption
S6	2012	3.5	Nil
S7	2019	2.5	Speed of data collection, network traffic, bandwidth
S8	2016	2	Latency, energy consumption
S9	2017	2	Data transmission
S10	2016	2.5	Nil
S11	2018	4	Energy consumption
S12	2012	3.5	Bandwidth usage
S13	2013	2	Integration, availability, efficiency, data protection, security
S14	2015	2.5	Latency
S15	2017	2.5	Energy consumption
S16	2016	2.5	Delay, energy consumption
S17	2016	2.5	Cost, energy consumption
S18	2016	2.5	Data transmission
S19	2013	1.5	Sensor data accessibility, alert notification time, energy consumption
S20	2014	2.5	Optimization in compression, optimization in bandwidth utilization
S21	2015	1.5	Data transmission
S22	2020	2.5	Computational efficiency, reliability
S23	2018	3.5	Performance
S24	2019	3.5	Data transmission
S25	2017	3.5	Energy consumption, data delivery latency
S26	2018	4	Energy consumption, computational time, hardware frequency rate
S27	2014	3	Data anomalies
S28	2014	3	Nil
S29	2019	2.5	Accuracy
S30	2017	2	Data ingestion scalability test, impact of the resource collocation, impact of the wallet key tactic, impact of data sharing
S31	2018	3.5	Data recovery
S32	2019	2.5	Nil
S33	2019	3.5	Delay, energy consumption
S34	2016	4	Accuracy of sensor status, mean of absolute truth error (AE), error rate
S35	2017	3.5	Nil
S36	2019	3.5	Latency, energy consumption
S37	2018	2.5	Temperature data, humidity data
S38	2018	2.5	Energy consumption, transmission distance, network throughput
S39	2015	2.5	Nil
S40	2014	3.5	Temperature, humidity
S41	2017	2.5	Nil
S42	2017	4	Energy consumption and performance
S43	2016	4	Energy consumption and performance

TABLE 11. Definitions of all acronyms mentioned in the paper.

Abbreviations	
WSNs	Wireless Sensor Networks
SC	Sensor Cloud
DC	Data Collection
IoT	Internet of Things
SLR	Systematic Literature Review
RQ	Research Question
PICO	Population,Intervention, Comparison, and Outcome Perspectives
SSP	Study Selection Procedure
QA	Quality Assisment
IC	Inclusion Criteria
EC	Exclusion Criteria
WBAN	Wireless Body Area Network
RTCM	Reliable Trust Computing Mechanism
P2DCA	Privacy-preserving-based Data Collection and Analysis
IoMT	Internet of Medical Things
cSHM	cloud-enabled remote Structural Health Monitoring
FoE	Fog of Everything
TASA	Time Adaptive Schedule Algorithm
MENs	Mobile Edge Networks
BLE	Bluetooth Low Energy
CTDC	Comprehensive Trustworthy Data Collection
SCS	Sensor Cloud System

and Case study. 37.21% of the studies selected utilized Experiments, followed by Simulation with 30.23%. These mechanisms are the most used by the studies selected whereby cumulatively they were utilized by 67.44% of the selected studies. Furthermore, we found out that 6 out of the 16 studies that used Experiment are studies that contribute Systems for their proposals (S12, S14, S15, S18, S29, S30), while 3 out of the 13 studies that used Simulation contributes Algorithm for their proposals (S8, S25, and S26). Performance measures are key when it comes to measuring and evaluating a proposal's effectiveness with respect to data collection. In this study, 36 studies were identified to used performance measures for evaluation out of the 43 selected studies. We found out that one of the most conducted and most common is the Energy consumption performance measure. We observed that 14 out of the 41 selected studies used energy consumption for their evaluation (S1, S2, S5, S8, S11, S15, S16, S17, S19, S25, S26, S33, S36,S38, S42 and S43). This amounts to 37.21% of the selected studies in this research domain. The dominance of measure such as energy consumption is relative due to the research directions of these studies, where the studies focus on data collection from sensor devices to sensor-cloud. When dealing with data collection in WSNs, energy is always a huge concern.

B. CHALLENGES AND DIRECTION FOR FUTURE WORK

A comprehensive review of the selected studies was conducted in this study. This SLR findings will allow researchers to know the existing contributions on data collection in Sensor-cloud. The study will also help researchers to know the evaluation mechanism and performance measures utilised by the studies selected in data collection. Therefore, the identified challenges with respect to the scope of this study were highlighted in this section. Also, the direction for future works is also given for further research in this research domain.

From figure 3 we have seen that the research output in this domain is relatively stable from the last 5 years. However, despite the stability, the output is quite low, where maximally, there is no publication year that has more than eight studies. This is a cause for concern looking at how important the research area is. Hence, we urge the research community to be more active. With 32.56% of the studies selected published in Journals, the general quality of the selected studies is perceived to be poor based on the research team consensus. Hence, we encourage both new and veteran researchers to publish more papers in Journal sources, because in general, Journal publications are more qualitative and have more depth. Evaluation mechanism such as Case study and the newly categorized in this study Hybrid, have received less attention from the research community. The combination of more than one evaluation mechanism (Hybrid) is very important and essential for rigorous evaluation of a given proposal. Hence, for future works, more evaluation should be conducted with a Case study and Hybrid mechanisms. Lastly, Performance measures are key when it comes to measuring and evaluating a proposal's effectiveness with respect to data collection. We observed that measures such as Energy consumption have dominated the field where most of the studies have utilized it (see Table 10). However, more diverse performance measures should be utilized and taking into consideration for a more rigorous evaluation. Hence, measures such as latency, data transmission, and delay should be used more by future works.

In addition, data collection in Sensor-cloud falls under the Fourth industrial revolution [65], where Internet of things (IoT) technologies such as WSNs are combined with cloud computing to provide real-time interface between the physical and virtual worlds. This paper research domain falls under this realm. However, due to the limitations of industry 4.0 (Fourth industrial revolution), various industries are looking ahead to industry 5.0 (Fifth industrial revolution), where sensory data can be collected autonomously with strong Artificial Intelligence (AI) presence. Hence, for future works, we encourage the research community in this domain to explore the utilization of industry 5.0 in their future research works. These explorations will further help in moving the research area forward to new heights.

C. THREAT TO VALIDITY

The limitations of this review have to be considered to have an overall analysis of the results gained from this SLR. Therefore, the key threats to the validity of this SLR are twofold, which are the incompleteness of the study search and the

TABLE 12. Author affiliation details.

Identifier	Authors	Institutions	Country of institution
S1	Wang	University of information science and technology	China
	Shen	University of information science and technology	China
	Wang	University of information science and technology	China
	Liu	University of information science and technology	China
S2	Wang	Huaqiao university, Xiamen	China
	Zeng	Huaqiao university, Xiamen	China
	Lai	Xiamen university	China
	Cai	Huaqiao university, Xiamen	China
	Liu	Huaqiao university, Xiamen	China
	Tian	Huaqiao university, Xiamen	China
	Chen	Huaqiao university, Xiamen	China
	Wang	University of information science and technology	China
S3	Wang	Huaqiao university, Xiamen	China
	Mei	Huaqiao university, Xiamen	China
	Jia	University of Macau	Macau
	Zheng	Macquire university	Australia
	Wang	Guangzhou university	China
	Xie	Zhejiang gongshang university	China
S4	Mao	Beijing university, Beijing	China
	Chen	Beijing university, Beijing	China
	Zhang	Beijing university, Beijing	China
	Xu	Beijing university, Beijing	China
	Zhou	Beijing University of Posts and Telecommunications	china
	Liu	Beijing university, Beijing	China
S5	Zhang	Jiangsu university, Zhenjiang	China
	Xiong	Jiangsu university, Zhenjiang	China
	Wang	Jiangsu university, Zhenjiang	China
S5	Dash	Biju patanaik university of technology	India
	Sahoo	Siksha "O" anusandhan university	India
	Mohapatia	Siksha "O" anusandhan university	India
	Pati	Siksha "O" anusandhan university	India
S7	Wang	North china university of technology	China
	Wang	North china university of technology	China
S8	Li	Huaqio university	China
	Wang	Huaqio university	China
	Wang	Guangzhou university	China
	Liang	Guangxi university	China
	Chen	Shijiazhuang kelin electric co, ltd.	China
S9	Traub	Technische universitat berlin	Germany
	Breb	Technische universitat berlin	Germany
	Rabl	Technische universitat berlin	Germany
	Katsitodimus	SAP innovation center	Germany
	Markl	Technische universitat berlin	Germany
S10	Pansare	Mescoe pune	India
	Bajad	Mescoe pune	India

biases on study selection. Hence, in this section, all these threats are discussed.

Firstly, with respect to the incompleteness of the study search, key studies can be missed in the process of retrieving the studies. This can affect the general completeness of the study search. Therefore, to alleviate this threat and further make sure that all significant and prospective studies are taking into consideration, a general search was done on the selected databases (see Table 2). This data sources contain a huge amount of Journals, Workshop, Conference,

TABLE 12. *(Continued.)* Author affiliation details.

Identifier	Authors	Institutions	Country of institution
S11	Rahman	University sains islam Malaysia, nilai	Malaysia
	Halim	University sains islam Malaysia, nilai	Malaysia
	Alwi	University sains islam Malaysia, nilai	Malaysia
	Seman	University sains islam Malaysia, nilai	Malaysia
	Mohammad	IoT security lab, cyber security	Malaysia
	Taufiq	IoT security lab, cyber security	Malaysia
	Mohammad	IoT security lab, cyber security	Malaysia
	Abiden	IoT security lab, cyber security	Malaysia
S12	Ward	University of st Andrews	UK
	Barker	University of st Andrews	UK
S13	Gejibo	University of Bergen	Norway
	Grasso	University of Bergen	Norway
	Mancini	University of Bergen	Norway
	Mughal	University of Bergen	Norway
S14	Li	Illinios institute of technology, Chicago	USA
	Keahey	Argonne national laboratory	USA
	Wang	Illinios institute of technology, Chicago	USA
	Zhao	Illinios institute of technology, Chicago	USA
	Raicu	Illinios institute of technology, Chicago	USA
S15	Charalampidis	Institute of computer science	Greece
	Tragos	Institute of computer science	Greece
	Fragkiadakis	Institute of computer science	Greece
S16	El mougy	German university in cairo	Egypt
	El-kerdany	German university in cairo	Egypt
S17	Chanavati	Deakin university, geelong	Australia
	Abawajy	Deakin university, geelong	Australia
	Izadi	Deakin university, geelong	Australia
S18	Soultanopoulos	Technical university of crete	Greece
	Sotiriadis	Technical university of crete	Greece
	Petrakis	Technical university of crete	Greece
	Amza	University of Toronto	Canada
S19	Piyare	Mokpo national university	South korea
	Park	Mokpo national university	South korea
	Maeng	Mokpo national university	South korea
	Park	Mokpo national university	South korea
	Oh	Mokpo national university	South korea
	Choi	Mokpo national university	South korea
	Choi	Mokpo national university	South korea
	Lee	Mokpo national university	South korea
S20	Grace	Anna university, Chennai	India
	Sumalalha	Anna university, Chennai	India
S21	Argyriou	University of Thessaly	Greece
S22	Liang	Guangxi university	China
	Zhang	Guangxi university	China
	Leung	Guangxi university	China
S23	Mihai	University politehnica of Bucharest	Romania
	Hanganu	University politehnica of Bucharest	Romania
	Stamateseu	University politehnica of Bucharest	Romania
	Popescu	University politehnica of Bucharest	Romania
S24	Usman	Swinburne university	Australia
	Jan	Abdul wali khan university mardan	Pakistan
	He	University of technology Sydney	Australia
	Chen	University of technology, Melbourne	Australia

TABLE 12. *(Continued.)* Author affiliation details.

Identifier	Authors	Institutions	Country of institution
S25	Wang	Huaqiao university, Xiamen	China
	Li	Huaqiao university, fengze	China
	Wang	Guangzhou university	China
	Cao	Hong Kong polytechnique university	China
	Bhuiyan	Fordham university	USA
	Jia	Shanhai jiaotong university	China
S26	Tao	BaoBaoji university of art and science	China
	Bhuiyan	Fordham university	USA
	Abdalla	Huaiyin institute of technology	China
	Hassan	King saud university	Saudi Arabia
	Zain	University of technology MARA	Malaysia
	Hayajneh	Fordham university	USA
S27	Wu	Imperial college, London	UK
	Birch	Imperial college, London	UK
	Silva	Imperial college, London	UK
	Lee	Imperial college, London	UK
	Tsinails	Imperial college, London	UK
	Guo	Imperial college, London	UK
S28	Nakagnwa	Osaka university	Japan
	Hiji	Tohoku university	Japan
	Kikuchi	Kochi university	Japan
	Fukumoto	Kochi university	Japan
	Shimojo	Osaka university	Japan
S29	Stojanovic	University of nis	Serbia
	Stojanovic	University of nis	Serbia
	Dordevic	University of nis	Serbia
	Stojnev	University of nis	Serbia
S30	Gesvindr	Masaryk university	Czech republic
	Michaikova	Tapltome, bratislawa	Slovaika
	Buhnovn	Masaryk university	Czech republic
S31	Chen	Nanjing university of posts and telecommunication	China
	Zhu	Nanjing university of posts and telecommunication	China
	Zhao	Nanjing university of posts and telecommunication	China
	Zhang	Nanjing university of posts and telecommunication	China
	Wang	Nanjing university of posts and telecommunication	China
	Yang	Nanjing university of posts and telecommunication	China
S32	Li	Xian eurasla university	China
S33	Wang	TCL corporation	China
	Zheng	Macquarie university	Australia
	Wang	University of California	USA
	Kumar	VIT university	India
S34	Bhuiyan	Fordham university	USA
	Wang	Guangzhou university	China
	Choo	University of texas at san Antonio	USA
S35	Yang	Stratus, unitec institution of technology	New Zealand
	Pulhai	University of technology Sydney	Australia
	Mohanty	University of north texas	USA
	Kougianos	University of north texas	USA

and Symposium in this domain that are indexed. Furthermore, the selected studies were backward and forward referenced searched to make sure that significant studies are taken. Even though we took all these actions to enhance the general completeness of the study search, this paper can still suffer from selection bias. This is due to the fact that other libraries

TABLE 12. *(Continued.)* Author affiliation details.

Identifier	Authors	Institutions	Country of institution
S36	Wang	Huaqiao university, Xiamen	China
	Zhao	Huaqiao university, Xiamen	China
	Cai	Huaqiao university, Xiamen	China
	Jia	University of macau	China
	Liu	Central south university	China
S37	Wang	National Taipei university of technology	Taiwan
	Hsu	National Taipei university of technology	Taiwan
	Hsiao	Hungkuo delin university of technology	Taiwan
S38	Wang	Huaqiao university, Xiamen	China
	Li	Huaqiao university, Xiamen	China
	Feng	Beijing jiaotong university	China
	Xu	Sichuan university	China
	Liang	Guangxi university	China
	Chen	Huaqiao university	China
	Liu	South china university	China
S39	Lawson	Georgia gwinnwtt college	USA
	Ramaswamy	University of Georgia	USA
S40	Bhunia	National university of Singapore	Singapore
	Mukherjee	Jadavpur university	India
S41	Maiti	National institute of technology rourkela	India
	Sahoo	National institute of technology rourkela	India
	Turuk	National institute of technology rourkela	India
	Satpalhy	National institute of technology rourkela	India
S42	Enzo Baccarelli	Sapienza University of Rome	Italy
	Paola G. Vinueza Naranjoi	Sapienza University of Rome	Italy
	Michele Scarpiniti	Sapienza University of Rome	Italy
	Mohammad Shojafar	Sapienza University of Rome	Italy
	Jemal h. Abawajy	Sapienza University of Rome	Australia
S43	Enzo Baccarelli	Sapienza University of Rome	Italy
	Nicola Cordeschi	Sapienza University of Rome	Italy
	Alessandro Mei	Sapienza University of Rome	Italy
	Massimo Panella	Sapienza University of Rome	Italy
	Mohammad Shojafar	Sapienza University of Rome	Italy
	Julinda Stefa	Sapienza University of Rome	Italy

like EI Compendex, Taylors & Francis, Emerald Insight, and Citeceerx were not taking into consideration.

Secondly, with respect to the study selection process, in order to decrease bias by researchers, we formulated a very clear and precise IC/EC criteria. Each researcher can have a different view of the IC/EC criteria, hence, the results of study selection of individual researchers are possibly going to differ. To alleviate this bias, we conducted a pilot selection so as to make sure that an agreement between the researchers is attained on the general meaning of the criteria. The possible mismanagement of duplicate studies is an additional threat. Hence, the threat might have marginally changed our results. 32 potential duplication were found and were thoroughly assessed to find out if they are the same study. In addition, to select a study, the final decision is taking by the two researchers that did the search process. Therefore, any disagreement that arises between the two researchers will be fixed between them. This will be done through discussion between them until a tangible agreement is established. Furthermore, the other researchers will check the final selected studies. For this study, peer-reviewed papers were solely include. Nonetheless, it's likely that we have missed some vital non-peer-reviewed studies in this domain.

VI. CONCLUSION

In this study, an SLR was conducted that presents a 10 year (2011-2020) summary of the current literature on data collection in Sensor-cloud. Out of the 3569 papers retrieved from the initial search conducted, 210 papers were selected based on rigorous analysis, of which 43 studies were selected based on the defined IC and EC criteria.

Our findings show that the research in this domain is relatively new, with a moderate and stable amount of studies published in the last 5 years. 44.19% of the studies selected where published in Conferences, followed by Journals, Symposiums, Workshops, and Magazine with 32.56%, 13.95%, 6.98%, and 2.32%, respectively. With respect to quality assessment, the result demonstrates that 23.25% of the studies have a total quality score of 4, which is the highest score that is set in this study. However, the majority of the studies, 53.50% score less than 3.5 as there quality score. With respect to the publication source, we identified four sources that are more noticeable. These sources are Transactions on Industrial Informatics, Internet of Things Journal, Conference on Wireless Sensor Networks, and Global Communication Conference with 2 publications each, respectively. Furthermore, the result of our analysis shows that there are 10 main contributions which are System with 27.95% of the studies, followed by Framework (23.25%), Algorithm (11.63%), Model (9.30%), Protocol (6.98%), Approach (6.98%), Investigation (4.65%), Method (4.65%), Architecture (2.33%), and Topology (2.33%). We observed that System and Framework are the most proposed contributions in the field of study. On evaluation mechanisms, we found out that six evaluation mechanisms were utilized by the selected studies, which are Experiment, Simulation, Theoretical analysis, Hybrid, and Case study. 37.21% of the studies selected utilized Experiments, followed by Simulation with 30.23%. These mechanisms are the most used by the studies selected whereby cumulatively they were utilized by 67.44%of the selected studies. With respect to performance measures, 36 studies were identified to used performance measures for evaluation out of the 43 selected studies. We found out that one of the most conducted and most common is the Energy consumption performance measure. We observed that 16 out of the 43 selected studies used energy consumption for their evaluation.

Finally, our research shows there is a substantial amount of interest by the researchers in the research domain considering the consistency in publication in the last 5 years. With this consistency, we expect more contributions with respect to proposals in years to come. Moreover, with the research challenges and future research directions presented in V-B, researcher must take them into consideration to help in tackling the identified challenges.

APPENDIX

Table 12 presents the selected studies information concerning the authors' names, institutions, and countries of the studies institutions.

REFERENCES

[1] D. C. Harrison, W. K. G. Seah, and R. Rayudu, ''Rare event detection and propagation in wireless sensor networks,'' *ACM Comput. Surveys*, vol. 48, no. 4, pp. 1–22, May 2016.
[2] M. Zhang, S. Xiong, and L. Wang, ''Sensor-cloud based precision sprinkler irrigation management system,'' in *Proc. China Conf. Wireless Sensor Netw.* Singapore: Springer, 2019, pp. 184–197.
[3] A. Cammarano, D. Spenza, and C. Petrioli, ''Energy-harvesting WSNs for structural health monitoring of underground train tunnels,'' in *Proc. IEEE INFOCOM*, Apr. 2013, pp. 75–76.
[4] T. Wang, Y. Mei, W. Jia, X. Zheng, G. Wang, and M. Xie, ''Edge-based differential privacy computing for sensor–cloud systems,'' *J. Parallel Distrib. Comput.*, vol. 136, pp. 75–85, Feb. 2020.
[5] A. Khan, A. Abdullah, M. Anisi, and J. Bangash, ''A comprehensive study of data collection schemes using mobile sinks in wireless sensor networks,'' *Sensors*, vol. 14, no. 2, pp. 2510–2548, Feb. 2014.
[6] M. K. A. Waghmare and P. Chatur, ''Energy efficient data collection and routing algorithm in wireless sensor network: A survey,'' Dept. Comput. Sci. Eng., Govt. College Eng., Amravati, India, Tech. Rep., 2014.
[7] H. Yetgin, K. T. K. Cheung, M. El-Hajjar, and L. Hanzo, ''A survey of network lifetime maximization techniques in wireless sensor networks,'' *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 828–854, 2nd Quart., 2017.
[8] M. Di Francesco, S. K. Das, and G. Anastasi, ''Data collection in wireless sensor networks with mobile elements: A survey,'' *ACM Trans. Sensor Netw.*, vol. 8, no. 1, pp. 1–31, Aug. 2011.
[9] S. R. Wankhade and N. A. Chavhan, ''A review on data collection method with sink node in wireless sensor network,'' *Int. J. Distrib. Parallel Syst.*, vol. 4, no. 1, pp. 67–74, Jan. 2013.
[10] N. M. Nair and A. F. A. Jose, ''Survey on data collection methods in wireless sensor networks,'' *Int. J. Eng. Res. Technol.*, vol. 2, no. 12, pp. 2220–2225, Dec. 2013.
[11] I. Ali, A. Gani, I. Ahmedy, I. Yaqoob, S. Khan, and M. H. Anisi, ''Data collection in smart communities using sensor cloud: Recent advances, taxonomy, and future research directions,'' *IEEE Commun. Mag.*, vol. 56, no. 7, pp. 192–197, Jul. 2018.
[12] M. Z. Hasan, H. Al-Rizzo, and F. Al-Turjman, ''A survey on multipath routing protocols for QoS assurances in real-time wireless multimedia sensor networks,'' *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1424–1456, 3rd Quart., 2017.
[13] S. K. Singh and P. Kumar, ''A comprehensive survey on trajectory schemes for data collection using mobile elements in WSNs,'' *J. Ambient Intell. Humanized Comput.*, vol. 11, no. 1, pp. 291–312, Jan. 2020.
[14] S. Keele, ''Guidelines for performing systematic literature reviews in software engineering,'' Ver. 2.3 EBSE, Softw. Eng. Group School Comput. Sci. Math., Keele Univ., Keele, U.K., Dept. Comput. Sci., Univ. Durham, Durham, U.K., Tech. Rep., 2007.
[15] A. Goyal, V. Gupta, and M. Kumar, ''Recent named entity recognition and classification techniques: A systematic review,'' *Comput. Sci. Rev.*, vol. 29, pp. 21–43, Aug. 2018.
[16] R. Bavaresco, D. Silveira, E. Reis, J. Barbosa, R. Righi, C. Costa, R. Antunes, M. Gomes, C. Gatti, M. Vanzin, S. C. Junior, E. Silva, and C. Moreira, ''Conversational agents in business: A systematic literature review and future research directions,'' *Comput. Sci. Rev.*, vol. 36, May 2020, Art. no. 100239.
[17] Y. Xiao and M. Watson, ''Guidance on conducting a systematic literature review,'' *J. Planning Edu. Res.*, vol. 39, no. 1, pp. 93–112, Mar. 2019.
[18] A. S. Ghiduk, M. R. Girgis, and M. H. Shehata, ''Higher order mutation testing: A systematic literature review,'' *Comput. Sci. Rev.*, vol. 25, pp. 29–48, Aug. 2017.
[19] J. Pérez, J. Díaz, J. Garcia-Martin, and B. Tabuenca, ''Systematic literature reviews in software engineering—Enhancement of the study selection process using Cohen's kappa statistic,'' *J. Syst. Softw.*, vol. 168, Oct. 2020, Art. no. 110657.
[20] K. A. Alam, R. Ahmad, A. Akhunzada, M. H. N. M. Nasir, and S. U. Khan, ''Impact analysis and change propagation in service-oriented enterprises: A systematic review,'' *Inf. Syst.*, vol. 54, pp. 43–73, Dec. 2015.
[21] S. Ouhbi, A. Idri, J. L. Fernández-Alemán, and A. Toval, ''Requirements engineering education: A systematic mapping study,'' *Requirements Eng.*, vol. 20, no. 2, pp. 119–138, Jun. 2015.
[22] A. Zakari, S. P. Lee, K. A. Alam, and R. Ahmad, ''Software fault localisation: A systematic mapping study,'' *IET Softw.*, vol. 13, no. 1, pp. 60–74, Feb. 2019.
[23] A. Zakari, S. P. Lee, R. Abreu, B. H. Ahmed, and R. A. Rasheed, ''Multiple fault localization of software programs: A systematic literature review,'' *Inf. Softw. Technol.*, vol. 124, Aug. 2020, Art. no. 106312.
[24] A. Wang, J. Shen, C. Wang, H. Yang, and D. Liu, ''Anonymous data collection scheme for cloud-aided mobile edge networks,'' *Digit. Commun. Netw.*, vol. 6, no. 2, pp. 223–228, May 2020.
[25] T. Wang, J. Zeng, Y. Lai, Y. Cai, H. Tian, Y. Chen, and B. Wang, ''Data collection from WSNs to the cloud based on mobile fog elements,'' *Future Gener. Comput. Syst.*, vol. 105, pp. 864–872, Apr. 2020.
[26] J. Mao, Y. Chen, Y. Zhang, X. Xu, Y. Zhou, and J. Liu, ''Parallel checking of content integrity in multi-cloud storage of heterogeneous sensor systems,'' *Int. J. Commun. Syst.*, vol. 30, no. 5, p. e3163, Mar. 2017.
[27] S. K. Dash, J. P. Sahoo, S. Mohapatra, and S. P. Pati, ''Sensor-cloud: Assimilation of wireless sensor network and the cloud,'' in *Proc. Int. Conf. Comput. Sci. Inf. Technol.* Berlin, Germany: Springer, 2012, pp. 455–464.
[28] Z.-H. Wang and J. Wang, ''An IoT data collection mechanism based on cloud-edge coordinated deep learning,'' in *Proc. China Conf. Wireless Sensor Netw.* Singapore: Springer, 2019, pp. 76–89.

[29] Y. Li, T. Wang, G. Wang, J. Liang, and H. Chen, "Efficient data collection in sensor-cloud system with multiple mobile sinks," in *Proc. Asia–Pacific Services Comput. Conf.* Springer, 2016, pp. 130–143.
[30] J. Traub, S. Breß, T. Rabl, A. Katsifodimos, and V. Markl, "Optimized on-demand data streaming from sensor nodes," in *Proc. Symp. Cloud Comput.*, Sep. 2017, pp. 586–597.
[31] J. R. Pansare and V. D. Bajad, "Errors detection in big sensor data on cloud using time efficient technique," in *Proc. ACM Symp. Women Res. (WIR)*, 2016, pp. 12–14.
[32] A. H. Azni, A. F. A. Rahman, N. H. M. Alwi, and K. Seman, "Measuring sensor to cloud energy consumption," in *Proc. 2nd Int. Conf. Internet Things, Data Cloud Comput. (ICC)*, 2017, pp. 43–47.
[33] J. S. Ward and A. Barker, "Semantic based data collection for large scale cloud systems," in *Proc. 5th Int. Workshop Data-Intensive Distrib. Comput. Date (DIDC)*, 2012, pp. 13–22.
[34] S. Gejibo, D. Grasso, F. Mancini, and K. A. Mughal, "Secure cloud storage for remote mobile data collection," in *Proc. 2nd Nordic Symp. Cloud Comput. Internet Technol. (NordiCloud)*, 2013, pp. 8–14.
[35] T. Li, K. Keahey, K. Wang, D. Zhao, and I. Raicu, "A dynamically scalable cloud data infrastructure for sensor networks," in *Proc. 6th Workshop Sci. Cloud Comput. (ScienceCloud)*, 2015, pp. 25–28.
[36] P. Charalampidis, E. Tragos, and A. Fragkiadakis, "A fog-enabled IoT platform for efficient management and data collection," in *Proc. IEEE 22nd Int. Workshop Comput. Aided Model. Design Commun. Links Netw. (CAMAD)*, Jun. 2017, pp. 1–6.
[37] A. El Mougy and A. El-Kerdany, "Reliable data transfer for collecting data from BLE sensors using smartphones as relays to the cloud," in *Proc. 11th Int. Conf. Comput. Eng. Syst. (ICCES)*, Dec. 2016, pp. 414–419.
[38] S. Ghanavati, J. Abawajy, and D. Izadi, "An alternative sensor cloud architecture for vital signs monitoring," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2016, pp. 2827–2833.
[39] T. Soultanopoulos, S. Sotiriadis, E. G. M. Petrakis, and C. Amza, "Data management of sensor signals for high bandwidth data streaming to the cloud," in *Proc. IEEE 37th Sarnoff Symp.*, Sep. 2016, pp. 53–58.
[40] R. Piyare, S. Park, S. Y. Maeng, S. H. Park, S. C. Oh, S. G. Choi, H. S. Choi, and S. R. Lee, "Integrating wireless sensor network into cloud services for real-time data collection," in *Proc. Int. Conf. ICT Converg. (ICTC)*, Oct. 2013, pp. 752–756.
[41] S. S. Grace and M. R. Sumalatha, "SCA—An energy efficient transmission in sensor cloud," in *Proc. Int. Conf. Recent Trends Inf. Technol.*, Apr. 2014, pp. 1–5.
[42] A. Argyriou, "Data collection from resource-limited wireless sensors for cloud-based applications," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–5.
[43] J. Liang, M. Zhang, and V. C. M. Leung, "A reliable trust computing mechanism based on multisource feedback and fog computing in social sensor cloud," *IEEE Internet Things J.*, vol. 7, no. 6, pp. 5481–5490, Jun. 2020.
[44] V. Mihai, C. E. Hanganu, G. Stamatescu, and D. Popescu, "WSN and fog computing integration for intelligent data processing," in *Proc. 10th Int. Conf. Electron., Comput. Artif. Intell. (ECAI)*, Jun. 2018, pp. 1–4.
[45] M. Usman, M. A. Jan, X. He, and J. Chen, "P2DCA: A Privacy-Preserving-Based data collection and analysis framework for IoMT applications," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1222–1230, Jun. 2019.
[46] T. Wang, Y. Li, G. Wang, J. Cao, M. Z. A. Bhuiyan, and W. Jia, "Sustainable and efficient data collection from WSNs to cloud," *IEEE Trans. Sustain. Comput.*, vol. 4, no. 2, pp. 252–262, Apr. 2019.
[47] H. Tao, M. Z. A. Bhuiyan, A. N. Abdalla, M. M. Hassan, J. M. Zain, and T. Hayajneh, "Secured data collection with hardware-based ciphers for IoT-based healthcare," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 410–420, Feb. 2019.
[48] C. Wu, D. Birch, D. Silva, C.-H. Lee, O. Tsinalis, and Y. Guo, "Concinnity: A generic platform for big sensor data applications," *IEEE Cloud Comput.*, vol. 1, no. 2, pp. 42–50, Jul. 2014.
[49] I. Nakagawa, M. Hiji, Y. Kikuchi, M. Fukumoto, and S. Shimojo, "M-cloud–distributed statistical computation using multiple cloud computers," in *Proc. IEEE 38th Int. Comput. Softw. Appl. Conf. Workshops*, Jul. 2014, pp. 301–305.
[50] D. H. Stojanovic, N. M. Stojanovic, I. Dordevic, and A. I. S. Ilic, "Sensor data fusion and big mobility data analytics for activity recognition," in *Proc. 14th Int. Conf. Adv. Technol., Syst. Services Telecommun. (TELSIKS)*, Oct. 2019, pp. 66–69.
[51] D. Gesvindr, J. Michalkova, and B. Buhnova, "System for collection and processing of smart home sensor data," in *Proc. IEEE Int. Conf. Softw. Archit. Workshops (ICSAW)*, Apr. 2017, pp. 247–250.
[52] S. Chen, X. Zhu, X. Zhao, H. Zhang, K. Wang, and G. Yang, "Fog computing assisted efficient privacy preserving data collection for big sensory data," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6.
[53] M. Li, "Design of multi-network data acquisition system based on cloud platform," in *Proc. Int. Conf. Virtual Reality Intell. Syst. (ICVRIS)*, Sep. 2019, pp. 98–100.
[54] T. Wang, H. Ke, X. Zheng, K. Wang, A. K. Sangaiah, and A. Liu, "Big data cleaning based on mobile edge computing in industrial sensor-cloud," *IEEE Trans. Ind. Informat.*, vol. 16, no. 2, pp. 1321–1329, Feb. 2020.
[55] M. Z. A. Bhuiyan, G. Wang, and K.-K.-R. Choo, "Secured data collection for a cloud-enabled structural health monitoring system," in *Proc. IEEE 18th Int. Conf. High Perform. Comput. Commun., IEEE 14th Int. Conf. Smart City, IEEE 2nd Int. Conf. Data Sci. Syst. (HPCC/SmartCity/DSS)*, Dec. 2016, pp. 1226–1231.
[56] C. Yang, D. Puthal, S. P. Mohanty, and E. Kougianos, "Big-sensing-data curation for the cloud is coming: A promise of scalable cloud-data-center mitigation for next-generation IoT and wireless sensor networks," *IEEE Consum. Electron. Mag.*, vol. 6, no. 4, pp. 48–56, Oct. 2017.
[57] V. Lawson and L. Ramaswamy, "Data quality and energy management tradeoffs in sensor service clouds," in *Proc. IEEE Int. Congr. Big Data*, Jun. 2015, pp. 749–752.
[58] S.-Y. Wang, Y.-J. Hsu, and S.-J. Hsiao, "Integrating blockchain technology for data collection and analysis in wireless sensor networks with an innovative implementation," in *Proc. Int. Symp. Comput., Consum. Control (IS C)*, Dec. 2018, pp. 149–152.
[59] T. Wang, Y. Li, W. Fang, W. Xu, J. Liang, Y. Chen, and X. Liu, "A comprehensive trustworthy data collection approach in sensor-cloud system," *IEEE Trans. Big Data*, early access, Mar. 1, 2018, doi: 10.1109/TBDATA.2018.2811501.
[60] E. Baccarelli, P. G. V. Naranjo, M. Scarpiniti, M. Shojafar, and J. H. Abawajy, "Fog of everything: Energy-efficient networked computing architectures, research challenges, and a case study," *IEEE Access*, vol. 5, pp. 9882–9910, 2017.
[61] E. Baccarelli, N. Cordeschi, A. Mei, M. Panella, M. Shojafar, and J. Stefa, "Energy-efficient dynamic traffic offloading and reconfiguration of networked data centers for big data stream mobile computing: Review, challenges, and a case study," *IEEE Netw.*, vol. 30, no. 2, pp. 54–61, Mar. 2016.
[62] P. Maiti, B. Sahoo, A. K. Turuk, and S. Satpathy, "Sensors data collection architecture in the Internet of mobile things as a service (IoMTaaS) platform," in *Proc. Int. Conf. I-SMAC (IoT Social, Mobile, Analytics Cloud) (I-SMAC)*, Feb. 2017, pp. 578–582.
[63] S. S. Bhunia, J. Pal, and N. Mukherjee, "Fuzzy assisted event driven data collection from sensor nodes in sensor-cloud infrastructure," in *Proc. 14th IEEE/ACM Int. Symp. Cluster, Cloud Grid Comput.*, May 2014, pp. 635–640.
[64] T. Wang, D. Zhao, S. Cai, W. Jia, and A. Liu, "Bidirectional prediction-based underwater data collection protocol for end-edge-cloud orchestrated system," *IEEE Trans. Ind. Informat.*, vol. 16, no. 7, pp. 4791–4799, Jul. 2020.
[65] S. Nahavandi, "Industry 5.0—A human-centric solution," *Sustainability*, vol. 11, no. 16, p. 4371, Aug. 2019.

IHSAN ALI (Senior Member, IEEE) received the M.Sc. degree from Hazara University Manshera, Pakistan, in 2005, and the M.S. degree in computer system engineering from the GIK Institute, in 2008. He is currently pursuing the Ph.D. degree with the Faculty of Computer Science and Information Technology, University of Malaya. He has more than five years teaching and research experience in different countries, including Saudi Arabia, USA, Pakistan, and Malaysia. He was also an Organizer of the Special Session on fog computing with Future 5V in 2017. He has published more than ten articles in international journals and conference proceedings. His research interests include wireless sensor networks, underwater sensor networks, sensor cloud, fog computing, and the IoT. He was a Technical Program Committee Member of IWCMC in 2017, AINIS in 2017, and Future 5V in 2017. He serves as a Reviewer for *Computers and Electrical Engineering*, *KSII Transactions on Internet and Information Systems*, *Mobile Networks and Applications*, the *International Journal of Distributed Sensor Networks*, the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, *Computer Networks*, IEEE ACCESS, and *IEEE Communications Magazine*.

ISMAIL AHMEDY (Member, IEEE) received the Ph.D. degree in computer science. He is currently a Senior Lecturer with the Department of Computer Systems and Technology, Faculty Science Computer and Information Technology, University of Malaya. His research interests include underwater acoustic sensor networks, embedded systems, and wireless sensor networks.

ABDULLAH GANI (Senior Member, IEEE) received the Diploma degree in computer science from ITM, the B.Phil. and M.Sc. degrees in information management from the University of Hull, U.K., and the Ph.D. degree in computer science from The University of Sheffield, U.K. He has a vast teaching experience in a number of educational institutions locally and abroad schools, such as the Malay Women Teaching College, Ministry of Education, the Rotterham College of Arts and Technology, Rotterham, U.K., and The University of Sheffield. He is currently a Professor with the Dean of the Faculty of Computing and Informatics, University Malaysia Sabah. He is also an Honorary Professor with the Department of Computer System and Technology, Faculty of Computer Science and Information Technology, University of Malaya, Malaysia. He has published more than 150 academic articles in respectable journals internationally with top ten% ranking. He received a number of citations in Web of Science and Scopus databases. His interest in research kicked off, in 1983, when he chosen to attend the three-month Scientific Research Course with RECSAM, Ministry of Education, Malaysia. His current research interests include self-organized systems, machine learning, reinforcement learning, wireless related networks, mobile cloud computing, big data, and the IoT. He was a Fellow of the Academy of Sciences Malaysia (FASc) for engineering and computer science discipline. He received the Teaching Certificate from the Kinta Teaching College, Ipoh.

MUHAMMAD TALHA (Member, IEEE) received the Ph.D. degree in computer science from the Faculty of Computing, University of Technology, Malaysia. He is currently with the Deanship of Scientific Research, King Saud University, Riyadh, Saudi Arabia. His research interests include image processing, medical imaging, features extraction, and classification and machine learning techniques.

MUHAMMAD AHSAN RAZA received the M.S. degree in computer science from Bahauddin Zakariya University, Pakistan, and the Ph.D. degree from Universiti Malaysia Pahang, Malaysia. He is currently a Lecturer with the Department of Information Technology, Bahauddin Zakariya University. He has published a number of articles in national and international conferences and journals. His main research interests include semantic web, information retrieval, big data, and the IoT systems.

MOHAMMAD HOSSEIN ANISI (Senior Member, IEEE) was a Senior Research Associate with the University of East Anglia, U.K., and a Senior Lecturer with the University of Malaya, Malaysia. As a Computer Scientist, he has designed and developed novel architectures and routing protocols for the Internet of Things (IoT) enabling technologies, including wireless sensor and actuator networks, vehicular networks, heterogeneous networks, and body area networks. He has a strong collaboration with industry and working with several companies in U.K., with focus on monitoring and automation systems based on the IoT concept, capable of reliable and seamless generation, transmission, processing, and demonstration of data. He is currently an Assistant Professor with the School of Computer Science and Electronic Engineering, University of Essex, and the Head of the Internet of Everything Laboratory. He has published more than 80 articles in high-quality journals and several conference papers. His research results have directly contributed to the technology industry. His research interests include real world application domains, such as energy management, transportation, healthcare, and other potential life domains. He is a Fellow of the Higher Education Academy. He is an Executive/Technical Committee Member of several conferences, a Technical Committee Member of the Finnish-Russian University Cooperation in Telecommunications (FRUCT), and a Senior Member of the Institute of Research Engineers and Doctors (IRED). He is also a member of ACM, the IEEE Council on RFID, the IEEE Sensors Council, the IEEE Systems Council, and the International Association of Engineers (IAENG). He received the Excellent Service Award for achievements from the University of Malaya. He was a recipient of two medals for his innovations from PECIPTA in 2015 and IIDEX expositions in 2016. He also received several International and national funding awards for fundamental and practical research, as a PI and a Co-I. He serves as an Associate Editor for a number of journals, including IEEE ACCESS, *Ad Hoc and Sensor Wireless Networks*, *IET Wireless Sensor Systems*, the *International Journal of Distributed Sensor Networks*, the *KSII Transactions on Internet and Information Systems* journals, and *Journal of Sensor and Actuator Networks*. He also serves as a Guest Editor for special issues of journals and a Lead Organizer for special sessions and workshops at the IEEE conferences, such as the IEEE CAMAD, the IEEE PIMRC, and the IEEE VTC.