Edge computing: A systematic mapping study

Jalal Sakhdari | Behrooz Zolfaghari | Shaghayegh Izadpanah | Samaneh Mahdizadeh Zargar | Mahla Rahati Quchani | Mahsa Shadi | Saeid Abrishami | Abbas Rasoolzadegan

Department of Computer Engineering, Ferdowsi University of Mashhad, Mashhad, Iran

Correspondence
Saeid Abrishami, Department of Computer Engineering, Ferdowsi University of Mashhad, Mashhad, Iran.
Email: s-abrishami@um.ac.ir

Summary
Edge computing is a new way of computing that uses resources at the edge of a network to solve the problem of communication delays in applications that require immediate responses. This field has received a lot of attention from the research community over the past few decades, leading to a significant increase in publications. To better understand the field, a systematic mapping study (SMS) was conducted using a three-tier search method that involved defining quality criteria to extract relevant search spaces and studies. This resulted in the selection of 112 search spaces out of 805 and 1440 studies out of 8725. The SMS addressed 8 research questions to identify the main topics, architectures, techniques, and other important aspects of edge computing.

KEYWORDS
edge computing, fog computing, systematic mapping study, systematic review

1 | INTRODUCTION

As the number of internet-based applications and the amount of data they handle continues to grow, the remote computational resources offered by cloud computing data centers have become an effective solution for data storage and processing needs. However, for applications that require mobility support, location awareness, and low latency, this model is not suitable. This is because the heavy computation and data of such applications need to be transferred to cloud data centers that are geographically far from the users running them. The resulting long communication latency makes it impractical for such applications.

Edge computing has emerged as a potential solution to the aforementioned problem. The core idea of edge computing is to place computing and storage resources in close proximity to end-users and at the network’s edge. This approach significantly reduces communication latency. Researchers have extensively studied edge computing for approximately two decades. The concept was initially introduced by Akamai, who created cache servers at the network’s edge.1 This approach has also been used in peer-to-peer (P2P) systems to distribute tasks and workloads among peers. Over time, the concept of edge computing has evolved, and various architectures have been developed to support it.

The literature on edge computing includes various definitions.2-4 In this SMS, we focus on research that involves sending processing or storage tasks to devices located near the end-user, which falls under the category of edge computing. While edge computing shows promise for meeting the quality requirements of new applications, there are still many operational challenges. However, due to the increasing use of IoT applications and the significant rise in mobile and wearable devices, research on edge computing has experienced substantial growth. Given the large number of papers published in this field, conducting a literature survey can help researchers identify the primary areas of focus in edge computing. In essence, we need a guide to assist researchers in conducting more effective searches on each aspect of edge computing. Given the vastness of research on edge computing, we require a sophisticated research methodology that can comprehensively cover relevant studies and be impartial, thorough, and auditable (traceable). To the best of our knowledge, while many research studies survey existing work in the field, none of them utilize a standardized and systematic approach to searching and reviewing papers.

To perform an advanced and comprehensive literature review, evidence-based software engineering recommends two well-known research techniques: systematic literature review (SLR) and systematic mapping study (SMS).5,6 Although these two systematic reviews employ the same search and data extraction methodology, they differ in their goals, which determine their usage.7 In an SMS, researchers seek to obtain general
information on a specific research field. Specifically, they aim to classify topics and identify existing trends in that field without delving into the details of each paper. In contrast, an SLR investigates a primary subset of research studies and extracts more specific data from them to analyze a topic (sub-topic) in-depth and identify the advantages and shortcomings of different proposals. This difference stems from the research questions (RQs) in each of these two techniques. The research questions for an SMS are high level, such as which sub-topics have been addressed, and provide an overview of the literature in specific topic areas. Therefore, an SMS should be used when identifying scopes and categorizing a large number of studies, while an SLR should be employed when specific data from a limited number of studies is required. It is worth noting that an SMS can be performed as a pre-SLR review.

In order to carry out an SMS on edge computing, a structured approach is necessary to identify relevant research studies in this area. This involves extracting search spaces and publications related to the topic, and then selecting a subset of these studies that meet appropriate quality criteria. To achieve this, we used a well-established methodology as described in previous studies. The key aspect of this methodology is to use an appropriate search method to identify a large number of studies related to edge computing. This was done by conducting searches at three different levels: manual searches, backward snowballing, and database searches. We also defined quality criteria to select studies that are relevant and of high quality for further analysis. These criteria were used to evaluate and select the most suitable studies from the extracted items. The process of extracting studies was evaluated separately to ensure accuracy and completeness of the data. A supplementary file containing comprehensive and complete information is available at https://github.com/jalalsakhdari/SMS. In conducting this SMS, we drew upon our past experiences and feedback from previous reviews to ensure a high-quality and comprehensive review of the literature on edge computing.

In this work, we use an SMS to investigate the topic of edge computing. We have classified the various topics related to edge computing and also categorized relevant research studies within each topic. Our ultimate objective is to help researchers interested in this field to easily identify research trends by specifying topics and sub-topics, similar to what has been done in previous SMS studies.

In order to achieve this, we have formulated 8 research questions which we endeavored to answer and scrutinize throughout the systematic review procedure. These questions were formulated with the aim of accomplishing various objectives such as identifying crucial topics in the realm of edge computing, recognizing the architectures that are utilized, identifying the techniques that are commonly employed, examining the applications of edge computing in other domains, identifying researchers who are currently active in this field, and determining the current search spaces in the domain of edge computing. The results of this investigation will serve as a valuable resource for researchers and developers who are interested in the subject of edge computing.

The rest of this article is organized as follows: Section 2 outlines the steps involved in an SMS process, followed by a description of the research methodology employed in this article, and a discussion of potential threats to the validity of this review. Section 3 presents and analyzes the results obtained from our SMS, while Section 4 compares our SMS with other secondary studies. In Section 5, we present some implications of our SMS results for researchers, practitioners, and educators interested in the field under study. Finally, in Section 6, we conclude our review.

To enhance the clarity of our findings, we have also included a comprehensive supplementary document which provides detailed information (accessible from https://github.com/jalalsakhdari/SMS). Throughout this article, we will refer to specific sections of this document which may be of interest to readers seeking more information on the method and data extracted. These documents are labeled with the prefix SupFile, followed by a postfix indicating the relevant table. For example, if we reference [SupFile]_(E1,T1), it refers to Table T1 of the Excel file E1. Additionally, to avoid confusion, all abbreviations used in this article are summarized in Table 1.

### 2 RESEARCH METHODOLOGY

This article employs a methodology that relies on the updated SMS guidelines described in Reference. The SMS process of this article is composed of three main phases: planning, evaluation, and conducting. Due to space limitations, we provide a brief summary of each step below. A more comprehensive explanation of these steps can be found in Appendix A (available in the Appendix A file at https://github.com/jalalsakhdari/SMS).

1. Specifying the Scope and RQs (each RQ along with its rationality is summarized in Table 2),
2. Planning of the search strategy which involves specifying the search strategy, search spaces, and search strings,
3. Specifying the search spaces,
4. Specifying the search strings,
5. Planning the study selection process,
6. Specifying the search and study selection evaluation strategy,
7. Planning the data extraction and classification process.

After completing the previous steps of identifying the included studies, it is important to investigate and analyze the data to respond to the RQs of the SMS accurately. The primary objective of this SMS is to identify the primary topics and sub-topics in the field of edge computing, and the research tree obtained during this process is shown in Figure 1. Appendix A provides a detailed description of how to extract the research tree.
TABLE 1 Summary of abbreviations.

| Abbr. | Term                  | Abbre. | Term                      |
|-------|-----------------------|--------|---------------------------|
| EC    | Edge Computing        | D2D    | Device-to-Device          |
| FC    | Fog Computing         | M2M    | Machine-to-Machine        |
| CO    | Computation Offloading | BS     | Base Station              |
| Ac.   | Architecture          | SDN    | Software-Defined Network  |
| DM    | Data Management       | NFV    | Network Function Virtualization |
| RM    | Resource Management   | TDMA   | Time Division Multiple Access |
| NM    | Network Management    | FDMA   | Frequency Division Multiple Access |
| Ec.   | Economic              | CDMA   | Code Division Multiple Access |
| S&P   | Security & Privacy    | P2P    | Peer to Peer              |
| UE    | User Equipment        | RAN    | Radio Access Network      |
| MD    | Mobile Device         | BBU    | Base Band Unit            |
| VM    | Virtual Machine       | RRH    | Radio Remote Head         |
| RQ    | Research Question     | ANET   | Vehicular Ad-hoc Network  |
| MEC   | Mobile Edge Computing | QoS    | Quality of Service        |
| MCC   | Mobile Cloud Computing| QoE    | Quality of Experience     |
| IoT   | Internet of Things    | -      | -                         |

TABLE 2 Defined research questions (RQs).

| No. | Research question                                                                 | Rationale                                                                                                                                                                                                 |
|-----|----------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1   | What is the level of activity in the field of edge computing, and how are studies distributed across different types and publication years? | By examining the current volume of research and general trends, we can gain a better understanding of the appeal of this field. Comparing the amount of research conducted in different publication years can also provide insight into the maturity of edge computing. |
| 2   | Who are the most active researchers and research venues in the field of edge computing, and how are they distributed geographically? | Understanding the demographics of edge computing research is a useful starting point for researchers who want to explore this field, as it can identify the most active scholars, venues, and countries involved in this research area. |
| 3   | What are the primary research topics being studied in the field of edge computing? | In order to gain insight into the current state of research, it is necessary to identify and categorize these topics, evaluate and analyze their distribution, and identify potential trends in the field. |
| 4   | How are applications distributed among the various research topics?               | To gain a better understanding of the importance and practicality of each research topic, it is necessary to identify the percentage of studies published for each application within a given research topic relative to the total number of studies. |
| 5   | How are various architectures distributed across different research topics?       | Architecture can be Fog Computing, Cloudlet, MEC and so on.                                                                                                                                              |
| 6   | Which techniques are commonly utilized in the field?                              | To gain a better understanding of the various research topics, it is necessary to identify the primary techniques used within each topic and analyze their relationship with other aspects. These techniques may include game theory, heuristics, and others. |
| 7   | What forms of empirical evaluation have been employed in the field?               | By empirical evaluation, we refer to whether the environment being studied is real, simulated, or a testbed.                                                                                               |
| 8   | Which qualitative requirements (QoS) have primarily been considered in the move towards edge computing? | Answering this research question can provide insight into when researchers should use edge computing. QoS considerations may include factors such as time, cost, energy consumption, and more. |
During the search process (including the planning and conducting phases), various factors may affect the validity of the SMS. Therefore, it is crucial to discuss potential threats to validity as a rubric to ensure the credibility of the study. The primary purpose of the validation process is to provide evidence to address any threats that the systematic review process may encounter. Appendix A presents some salient evidence related to these threats.

3 | RESULTS OF THE STUDY

The previous sections have outlined the search process and its particulars. In this section, the results of the review of the edge computing field are discussed in accordance with the research questions (RQs) presented in Section 2, as listed in Table 2. Additionally, the questions are addressed and analyzed according to the levels of the research tree.

3.1 | What is the level of activity in the field of edge computing, and how are studies distributed across different types and publication years?

This section presents a statistical analysis of the number of publications in various years, as well as the percentage of studies conducted on different topics. Figure 2 depicts the number of papers published in the field of edge computing over the years, where the horizontal axis represents the years and the vertical axis represents the number of published papers. The data utilized to generate the diagrams in this section are derived from the [SupFile]_(E2,T1) and [SupFile]_(E1,T5) files.

Figure 2 illustrates that prior to 2014, only a handful of papers had been published in the field of edge computing. In previous years, studies had been conducted on performing computations in proximity to the user, such as cloudlets,10 cyber foraging,11 and nomadic computing.12 However, the emergence of formal and standardized edge computing can be traced back to the period of 2013–2015. In 2013, IBM and Nokia introduced...
the Joint Radio Applications Cloud Server (RACS) project, which was a platform for edge computing over 4G/LTE networks. Following this, efforts began to standardize edge computing under the guidance of the European Telecommunications Standards Institute (ETSI). The first Mobile Edge Computing congress took place in London in 2015, followed by the first IEEE/ACM Symposium on Edge Computing a year later. These events marked the beginning of edge computing as a new avenue of research, providing numerous research opportunities for academics and industry professionals alike.

Figure 2 also highlights a noticeable point, which is the sharp increase in the number of publications in recent years. This surge in research activity can be attributed to various factors that justify the increasing popularity of this field. Some of these factors include:

**Maturity of IoT and emergence of new applications:** With the widespread use of smartphones and the growing prevalence of IoT-based technologies, a variety of applications such as healthcare, smart communities, social networks, and VANETs have become increasingly popular. Consequently, a large number of studies have been conducted to adapt these applications to the edge computing paradigm, leveraging edge and fog computing to improve their quality.

**Existence of several operational challenges:** As an emerging area of research, edge computing faces numerous challenges in areas such as resource management, networking, and QoS assurance. These challenges have caught the attention of many researchers in the field, driving increased research activity in recent years.

**Advantages of edge over remote cloud:** The proliferation of billions of internet-connected devices has made remote cloud computing impractical for new applications. It can result in heavy traffic on the backbone network, leading to communication delays. Consequently, a significant amount of research has been conducted to migrate and adapt remote cloud services and applications to distributed edge resources.

**Close relationship to other research area and technologies:** Edge computing is a rapidly developing research field that is closely related to various other areas of research, including cloud computing, networking, data analysis and processing, security, artificial intelligence, medicine, and more. As a result, many researchers from these different fields have published papers in edge computing, leading to interdisciplinary research and the discovery of new insights.

Figure 3 illustrates the number of published papers on edge computing over the years, segmented by topic. The horizontal axis represents the years, while the vertical axis indicates the number of published papers. As the graph depicts, there has been an upward trend in research publications across all topics. However, the resource management and computation offloading topics have experienced the highest acceleration compared to other topics. Resource management is a critical issue in edge computing due to the limited and heterogeneous resources with variable loads. Hence, the number of researches in resource management has grown more sharply than other topics. Computation offloading is also an active research area with various challenges such as mobility management, partitioning, and decision making, making it a subject of great interest. Following resource management and computation offloading topics, the data management topic has experienced the highest growth in the number of publications. This growth can be attributed to the typical applications in this field, as most edge computing applications such as healthcare, crowdsourcing, streaming, and social networks involve user data. Therefore, there is a strong need to provide methods for storing, processing, and analyzing data in a distributed edge environment, making data management one of the fastest-growing topics in this field.

### 3.2 Who are the most active researchers and research venues in the field of edge computing, and how are they distributed geographically?

This section provides an introduction to active journal researchers and journals focused on edge computing. The information contained in this section can serve as a useful starting point for new researchers. To identify active researchers, the number of studies published by each author is tallied. The top ten authors with the highest number of publications are selected as active researchers, and the number of their publications in each
The information about active researchers is presented in the form of a bubble diagram, as shown in Figure 4. The horizontal axis in the diagram represents different topics while the vertical axis shows the top ten authors who have published the highest number of journal articles. The size of the bubbles corresponds to the number of papers published on each topic. According to the statistics shown in this diagram, Dr. F. Richard Yu and Dr. Victor C. M. Leung have the highest number of published papers among all the other authors. The majority of their published papers are related to resource management and offloading, which are currently hot research areas in the field of edge computing. On the other hand, some other authors, such as Dr. Chunlin Li, Dr. Mohsen Guizani, and Dr. Tian Wang, have focused more specifically on security and privacy topics.

To identify active journals in the field of edge computing, similar to the process for active researchers, the number of papers published in each journal was tallied. The top ten journals with the highest number of relevant publications were then selected as active journals. The journal column in [SuppFile]_(E2T2) was used for this purpose. Additionally, information on the number of studies published by other journals per topic is available in [SupFile]_(E4T5). Figure 5 illustrates the active journals in edge computing along with the number of papers published in each one. "IEEE Access," "IEEE Internet of Things Journal," and "Future Generation Computer Systems" are the journals with the largest number of published papers. There are several reasons for the popularity of these journals.

"IEEE Access Journal" is a multidisciplinary journal that covers all IEEE fields, with a review and publication time of 4 to 6 weeks. This feature can be of great interest as papers on edge computing are being published at a high speed. After "IEEE Access," "IEEE Internet of Things Journal" has the largest number of papers. This journal’s main area covers all aspects related to the Internet of Things, including IoT system architecture, IoT communication, networking protocols, and applications such as smart cities, smart environments, and smart homes, which are closely related to edge computing. Since no specific journal has been established to cover edge computing, this journal can be considered one of the most specialized journals in the field of edge computing. The third journal is "Future Generation Computer Systems," which is one of the oldest and most popular journals in computing, covering a wide range of computer systems.

Another analysis presented in this section is the proportion of papers published in journals per topic. As there is a broad thematic diversity in edge computing research, analyzing this information can aid in the search for specific sub-topics. Figure 6 illustrates the percentage of papers published in various topics for each of the leading journals. In this figure, the horizontal axis represents the journals, and the vertical columns represent the percentage of publications in each topic for the corresponding journal. Since some papers may belong to multiple topics, the sum of the percentages in certain journals may exceed 100. Based on this chart, most of the papers published in the "Future Generation Computer Systems" and "Sensors" journals relate to security and privacy topics. The network management topic has a higher percentage in journals such as "Journal of Network and Computer Applications," "IEEE Journal on Selected Areas in Communications" and "IEEE Transactions on Wireless Communications,"
which focus more on computer networks. The resource management topic, which is widely regarded as the most challenging topic in this field, also accounts for the highest percentage of published papers in most of the journals.

The next analysis of this section belongs to the geographical distribution of research in the field of edge computing. The purpose of this analysis is to help researchers identify the leading countries in this area. To produce the statistics, data from the [SupFile]_(E1,T1) was used. Table 4 displays some of the contents of this file. It’s worth noting that the geographical location of authors is determined based on the countries mentioned in their affiliations, as different authors from various countries may have participated in each paper. Figure 7 presents the percentage of participation of different countries in the reviewed studies. The top ten countries with the highest number of papers are shown in this figure, while the countries with low frequency are grouped under the ‘Other’ category. The [SuppFile]_(E4,T7) file provides information on the number of papers published from other countries.

The chart in Figure 7 demonstrates that China is responsible for the majority of published papers on edge computing, accounting for approximately one-third of all papers. There are a number of reasons why this is the case. One reason is the existence of densely populated cities in China and the widespread use of smartphones among the Chinese population. As of the end of 2018, China had the highest number of mobile users in the world, with 850 million smart mobile users, which is more than double the number in the second country, India.\textsuperscript{15} This fact has increased the demand for new technologies such as 5G and ultradense networks in order to ensure access to the data network for users. In addition, there is a
growing need for edge computing to improve user Quality of Service (QoS). On the other hand, the Chinese are leaders in using technologies such as the Internet of Things (IoT), smart homes, and mobile applications based on artificial intelligence (AI). Various technologies, including semantic analysis, speech and image recognition, have been rapidly implemented in the Chinese smartphone market. The ownership of various smart home related products in China is also significantly higher than the global average. Since the operation of these technologies and applications requires fast processing with low communication delays on edge resources, edge computing has received much attention from Chinese researchers. Following China, the United States, Canada, Australia, South Korea, the United Kingdom, Italy, India, Spain, and Japan have the highest number of papers on edge computing, and these countries are often ranked as the most technologically advanced in the SCImago country rank.

The final analysis in this section pertains to the keywords and common terms that have been gathered from the reviewed papers. Familiarity with these keywords and common terms gives researchers a better understanding of the subject matter and assists them in conducting more effective
FIGURE 8 Main keywords in the field of edge computing.

searches. To conduct this analysis, the observed keywords in the keywords section of all papers have been collected. For some papers that lack a keyword section, an expert has proposed some keywords related to the content of the paper. Figure 8 displays a cloud diagram of the keywords with the highest frequency, representing the most common terms in this field based on the reviewed papers. The size of each word reflects the number of times it has been repeated.

3.3 What are the primary research topics being studied in the field of edge computing?

In this section, we will discuss the existing research scopes in the field of edge computing from the perspective of reviewed papers. Figure 1 displays these scopes in a research tree, which organizes them into topics and sub-topics. If certain studies cover multiple subjects, they are categorized into more than one topic. Additionally, there is a sub-topic named "miscellaneous," which includes fields that do not fit into any of the other sub-topics. To ensure that we properly covered all of the topics in the research tree, some other review papers have been examined. However, we limited our review to journal articles, and we discuss the effects and threats of this limitation in our article in Section 3.1.4 of Appendix A. Even though these review papers covered conference papers, our obtained research tree indicates that the topics were covered appropriately. Furthermore, to accurately extract the topics and sub-topics, we selected the covered topics for each included article by analyzing the title, abstract, and full text. The complete information about these topics is reported in [Sup-File] [E2T4]. In the following sections, we will explain each topic and its sub-topics in detail.

3.3.1 Computation offloading

This research topic focuses on papers that introduce innovative methods in the field of offloading. Among the reviewed papers, this topic has received the second-highest attention from the research community, after resource management. Mobile Devices (MDs) are often constrained by limitations in CPU processing power, memory, and battery life. Offloading is a critical aspect of edge computing, which enables users to execute resource-intensive applications on MDs through edge resources. The primary objective of offloading is to enhance task execution speed and power efficiency, either individually or together. Offloading can be directed towards remote cloud resources, edge resources, or a combination of both (known as collaborative offloading). Network communication is necessary when using cloud and edge resources. In cases where network instability occurs, collaborative offloading is a beneficial alternative. However, it is important to note that papers that solely utilize cloud resources as a destination, such as, are outside the scope of this research. Such papers generally belong to the field of mobile cloud computing (MCC). Sub-topics related to offloading research are discussed in the following sections.

Decision making

This particular sub-topic encompasses studies that explore situations where users are faced with various choices when it comes to offloading and are thus required to make a decision. In these studies, the problem of decision making plays a crucial role in achieving a balance between the user’s benefits and reducing offloading overhead. Various approaches have been proposed to address this issue. In some studies, the decision making
problem involves choosing between local execution, limited resource execution on cloudlets or edge devices, or remote execution on unlimited cloud resources with high access latency.\textsuperscript{20,21} Additionally, if it is possible to connect to other users’ devices through a Mobile Adhoc Network (MANET), this can also be considered as another offloading option.\textsuperscript{22,23} Another decision making scenario involves selecting from multiple resources of the same provider that are available near the user. This type of decision making is more common in ultra-dense networks where users may have access to multiple macro or small cell base stations simultaneously.\textsuperscript{24,25}

Mobility management

Mobile edge computing is constructed upon cellular networks and Wi-Fi access points, making mobility an inherent feature and an essential factor in distributed cloud computing and edge computing. User mobility can lead to failures in offloaded requests, making it necessary to develop an appropriate mechanism for mobility management to enhance QoS in edge computing applications. The reviewed papers propose various solutions to address the mobility issue. The most significant of these solutions involve predicting mobility and user paths using methods such as,\textsuperscript{26,27} modeling user mobility with Markov chains,\textsuperscript{28} and analyzing movement patterns.\textsuperscript{29} User mobility, coupled with limited coverage of edge servers, may result in reduced performance, service continuity, and consequently, QoS violations. In such cases, service migration is considered a suitable solution, which will be further discussed in the resource management topic.

Partitioning

Partitioning is another area of research within the computation offloading topic. Partitioning involves dividing computation tasks or data to execute on different devices separately. This allows for parallelization, resulting in more efficient application performance. A crucial issue in partitioning is the type of computational tasks. In general, researches consider two common models: binary tasks and partial tasks.\textsuperscript{30} Binary tasks are highly integrated and cannot be decomposed; hence they must be executed entirely on a single resource. On the other hand, partial tasks can be broken down into smaller tasks and executed in parallel. Partitioning a task can be performed by breaking the code and executing independent processes and components on different resources,\textsuperscript{31} or dividing input data into different sections.\textsuperscript{32} Partitioning can be done based on various criteria such as energy\textsuperscript{33,34} or resource efficiency.\textsuperscript{31}

Computation migration

Computation migration involves dynamically transferring computing tasks (in whole or in part) from User Equipments (UEs) to nearby or remote sources. There are several methods for achieving this, such as the system VM method used in cloudlets\textsuperscript{35} and ThinkAir,\textsuperscript{36} application-level virtualization in MAUI,\textsuperscript{37} and Instruction Set Architecture (ISA) command emulation in Reference 38. In the system VM method proposed in Reference 35, the user’s application is migrated as a VM from the UE to an edge resource. This method involves migrating the entire VM with the program, resulting in higher overhead than other methods.\textsuperscript{38} In MAUI, application-level virtualization is used to reduce overhead. In these methods, specific parts of the code, such as functions or classes that need to be offloaded, are identified through annotation and offloaded using application-level virtualization, such as Dalvik VM in Android.\textsuperscript{39} In the ISA emulation method (process migration), the amount of data to be offloaded is much less compared to VM and application-level virtualization. However, due to the difference between the architecture of mobile devices (usually ARM) and the architecture of servers (generally Intel), ISA emulation is required while offloading the process state between heterogeneous processors.\textsuperscript{38}

3.3.2 Resource management

Resource management is one of the most critical issues in edge computing and has therefore received a lot of attention from the research community, as observed in the reviewed papers. Resources in edge computing can be classified into three categories: computational resources, storage resources, and network resources. These resources are available to users in a heterogeneous, limited, and distributed manner, unlike cloud resources. Moreover, they are more dynamic and, in some cases, distributed competitively among users.\textsuperscript{40} Resource management in edge computing involves a set of control processes for allocating and reclaiming resources to user tasks or requests based on various criteria such as latency, cost, and energy. The main research subfields in this area include resource allocation and scheduling, migration and placement, load balancing, resource estimating and utilization, and resource sharing, which are discussed below.

Resource allocation and scheduling

Resource allocation and scheduling refer to the process of assigning resources to specific tasks or requests based on predefined criteria and constraints. In edge computing, improper resource allocation can result in reduced resource efficiency, energy constraint violations, or missed user deadlines. Various scheduling methods have been proposed in the reviewed papers, including task execution on VMs\textsuperscript{41} or containers on the edge,\textsuperscript{42} assigning user sessions to application samples,\textsuperscript{43} running services on edge resources,\textsuperscript{44} and network resource allocation.\textsuperscript{45}
Resourcemigration and placement
The process of selecting an appropriate location for a service, VM, or computing resource is handled by a placement algorithm. Additionally, live migration refers to transferring a VM from one host to another without disrupting the VM service. These processes can be performed either proactively (offline) or reactively (online). In proactive methods, like the one described in Reference, multiple versions of the service are pre-placed on various hosts. Reactive methods, such as the one outlined in Reference, rely on user mobility and select the most suitable edge node for migration or placement based on this mobility. The complexities associated with user mobility, path uncertainty, and the dynamic and heterogeneous nature of resources at the edge of the network make edge resource placement and migration more challenging than cloud computing.

Load balancing
Load balancing is a crucial mechanism that ensures the even distribution of workload among nodes, preventing overload on a single or a few nodes. This directly impacts the reduction of users’ response time. In edge computing, where computational nodes are distributed geographically, load balancing becomes a primary concern. In the reviewed papers, the load balancing problem is predominantly addressed by creating and distributing multiple versions of a program or service on fog nodes or different cloudlets, and then routing data and requests to these nodes based on the level of load.

Resource sharing
Resource sharing is a crucial solution to overcome the limitations of resources at the edge of the network, particularly the limited resources of UEs. Generally, resource sharing addresses three main issues: (1) insufficient specific types of resources in UEs to execute tasks, (2) insufficient overall resources in UEs, and (3) the use of additional resources to accelerate task completion. Resource sharing can be accomplished between UEs through technologies like D2D introduced by 5G, Mobile ad-hoc, Wi-Fi, and Bluetooth, or by creating clusters between edge resources. In the case of resource sharing between UEs, computing, network, or storage resources are shared for better, more secure, and faster execution. For instance, to enhance network reliability, edge devices can use not only their communication resources but also the heterogeneous network resources of other UEs to meet QoS requirements. Leveraging resource sharing techniques at the edge can not only improve resource efficiency but also enable the execution of resource-intensive applications.

Resource estimation
This sub-topic focuses on reviewing papers related to resource estimation, which is one of the initial steps in resource management. Resource estimation involves predicting the number of resources needed to complete a task or handle a computational workload. It is also essential for managing fluctuations in resource demands and ensuring QoS. Effective resource estimation can improve efficiency and fairness in resource management.

Resource utilization
Resource utilization involves maximizing the use of available resources while taking into account important factors such as time, energy, and cost. Typically, multiple criteria are considered simultaneously in resource optimization studies. To achieve this goal, resource provisioning and consumption processes are often modeled using various optimization methods. Linear programming, PSO, Lyapunov, game theory, heuristic, and meta-heuristic are among the most commonly used methods for optimizing resource usage.

3.3.3 Architecture
Some of the papers that were reviewed have proposed an architecture for conducting distributed computations on edge resources, or have enhanced existing edge architectures with additional capabilities. This has led to the recognition of architecture as a distinct area of research. The papers in this field aim to identify the different components of an architecture and explore their interactions, with the goal of enabling new capabilities or simplifying the implementation of specific applications. The research in this area can be classified into two sub-topics: Application-specific architecture and General-purpose architecture, which will be described in further detail below.

Application-specific
This sub-topic focuses on research that propose an architecture for running a specific application in an edge computing environment. Since each application has unique features and requirements in terms of resource allocation and network traffic, some studies have modified the infrastructure or topology of the elements in the edge architecture to meet these specific needs. Examples of these modifications include the presentation of architectures designed for healthcare applications, crowdsensing applications, and automated monitoring systems.
General-purpose
This research sub-topic is focused on the design of a general-purpose architecture that can serve as a baseline configuration for distributed edge computing. The aim of these studies is to create a new or modified architecture that adheres to new criteria or incorporates new capabilities to previously established architectures. Some examples of these studies include proposing an architecture for achieving agreement and consensus between nodes within a specific community, developing a trust evaluation architecture, reducing energy consumption in edge computing systems, and presenting a blockchain-based architecture for ensuring security.

3.3.4 Network management
One of the key foundations identified in the reviewed papers is network management. This research area is dedicated to optimizing parameters related to network infrastructure and applying new generation cellular network technologies to the edge computing paradigm. This topic also encompasses all aspects of modifying current standard network architectures to accommodate the capabilities of edge computing. The reviewed papers showcase several research studies in the field of network management, which have been classified into several sub-topics based on standard network technologies. The following paragraphs provide a detailed explanation of each sub-topic.

Access control
The field of communication encompasses two prominent concepts—network access architecture and multiple access schema. The access network, also referred to as the Radio Access Network (RAN), is responsible for connecting a mobile device to its core network (CN) by residing between them. On the other hand, multiple access schema, also known as channel access method, facilitates sharing a communication channel or physical communication medium between numerous users using multiplexing. The reviewed papers focused on both these concepts and are discussed below.

Access network architectures: The Radio Access Network (RAN) is a well-known architecture in mobile and cellular networks that has developed over the years alongside advancements in mobile communications. Its main function is to connect devices such as M2Ms or computers to the CN by acting as a middleman. The RAN is composed of a collection of base stations (BS) that each cover a specific area depending on their transmission power. Each BS is further divided into a Remote Radio Unit (RRU) for handling radio functions and a Base Bound Unit (BBU) for processing the baseband and channels. The data transmission between the RRU and BBU is carried out through the Common Public Radio Interface (CPRI). Cloud-RAN (C-RAN) is a variation of the RAN that has the potential to handle as many BS as required by using virtualization technology. In C-RAN, the BBUs are virtualized and shared among operators in a centralized BBU pool. The Fog RAN (F-RAN) architecture takes the central computing concept of C-RAN and transfers it to the network’s edge. In this architecture, a fog node can cache content and provide computation capabilities. Cisco first introduced this concept to take advantage of local signal processing, computing, collaborative resource management, and distributed storage and caching at the edge of the network. Since then, research in the field of edge computing has focused on combining the Multi-access Edge Computing (MEC) with this technology. Research in this area has often focused on designing new F-RAN-based architectures for resource management, adding mobility management mechanisms, or providing methods to increase resource efficiency.

Multiple access schemas: There are two well-known Multiple Access (MA) schemes in wireless communication systems: Orthogonal Multiple Access (OMA) and Non-Orthogonal Multiple Access (NOMA). The main difference between them is that in OMA, the bandwidth is divided among UEs based on frequency (such as FMDA), time (TDMA), or code (CDMA), while in NOMA, it is divided fairly among all UEs. For example, connecting thousands of IoT devices, such as vehicles in vehicular adhoc networks for intelligent transportation, with OMA would require thousands of bandwidth channels, but with NOMA, they can be served using a single channel. However, in NOMA networks, some users with poor channel conditions may experience low data rates. Despite this, NOMA is considered one of the most promising radio access techniques for next-generation wireless communications, including 5G. Cognitive radio (CR) is another form of wireless communication where the transceiver can intelligently detect which communication channels are in use and move into vacant ones while avoiding occupied ones. In fact, NOMA is a special case of CR.

Network abstraction and orchestration
This sub-topic discusses the research related to utilizing Software-Defined Networking (SDN) technology in edge computing environments. SDN is a centralized network configuration architecture that allows centralized programming and control of network traffic. In this configuration method, the control plane is separated from the data plane and is managed by a central module with a global view of the traffic and network’s overall state.

Network Function Virtualization (NFV) is a technology that is typically used in conjunction with SDN. NFV involves virtualizing the network equipment functions, allowing networking functions to be performed on computing resources like VMs or containers located in the cloud or at the network’s edge. The use of these two technologies has rapidly transformed the development of network functions and the evolution of network-based architectures. They offer significant benefits, including cost reduction, increased network flexibility and scalability, and reduced time-to-market for new applications and services.
The decentralized nature of the edge computing environment is a fundamental issue that results in several problems related to computing, storage, security, and traffic control. To address this issue, researchers have explored the use of centralized control, decision making capabilities, and programmability offered by SDN. SDN has become increasingly popular in edge computing research, as it offers several solutions such as resource management, cloudlet management, traffic control, and mobility management.

Traffic management and engineering
This research sub-topic is concerned with traffic modeling optimization and engineering in edge computing, which is a crucial aspect of network operations. One of the most significant characteristics of edge computing is the heavy traffic load at the network’s edge caused by a large number of user requests for various services. As a result, many researchers in the field of edge computing have focused on traffic management and engineering. Traffic engineering is a process that optimizes network resources by allocating bandwidth and selecting traffic routes to meet service requirements. In the reviewed literature, video streaming applications were the most frequently studied applications in the field of traffic management and engineering as video streaming traffic accounts for the largest portion of network traffic.

Slicing and overlaying
Network slicing is a technology that has been proposed to address the diverse applications and business models in 5G, and it has received significant attention in both academia and industry. The concept behind network slicing is to divide physical network infrastructure entities into isolated logical network components with specific functions that can meet different application requirements. Each of these network components is designed for a particular purpose or service.

As the requirements of edge-based computing services are highly diverse, network slicing provides a flexible and promising solution to meet the needs of various services. In the papers reviewed, network slicing has been utilized for different purposes, including separating MEC services from traditional services, dividing computational resources based on energy criteria for various user-requested services, and providing application-based QoS.

3.3.5 Security and privacy

Similar to any other infrastructure, edge computing is not immune to hostile and aggressive agents. Edge computing is based on a distributed and mostly unreliable platform. Therefore, security is a major concern in edge computing, and has been discussed at various levels in the literature. The most critical security issues in this field include the design of access control mechanisms, privacy preservation, trust management, and attack prevention, which are explained in more detail below.

Access control
Access control is a crucial security feature, particularly for applications that involve user data, such as storage applications. The aim of access control is to design a mechanism to monitor and ensure that data is accessible only to authorized individuals who have permission to access it. Attribute-based encryption is a well-known technique for providing access control, but it comes with a significant computation overhead for encryption/decryption. Some researchers focus on reducing this overhead by outsourcing the encryption/decryption computation to edge servers. In addition, blockchain technology, which has recently gained the attention of researchers in various industries, is also being considered in this area.

Attack detection
In the realm of edge computing, there are numerous possible attacks that can target different levels of the network or application, and can cause damage to varying degrees. The main goal of this research sub-topic is to prevent or detect these attacks. Some examples of common attacks in this area include the "Man in the Middle Attack," "Selective Forwarding Attack," and "Data Injection Attack."

Trust evaluation
Measuring trust values between components in a community is an important security feature used to detect attacks and secure systems. In edge computing, determining the degree of trust between pairs is particularly important due to the presence of low-reliable heterogeneous edge nodes. Consequently, several studies have been conducted on the topic of trust in edge computing. Trust in social networks has been one of the applications considered in this field. In social networks, trust can be verified by a user, a service provider, or by observing the connection between two entities.
Privacy
Most applications of edge computing involve uploading content, making the privacy of users vulnerable to threats. This issue has gained more attention in personal data-driven applications like crowdsensing and healthcare applications. In healthcare applications, the health-related data measured by wearable sensors on users is sent to nearby edge servers for analysis. In crowdsensing applications, data sensed by volunteer user devices are collected for use in a specific application. In both scenarios, private user data may be accessible to application providers. Another type of private user data that can be compromised through edge computing applications is the user’s location, which is used for navigation or map-based applications. Preserving user location privacy has also received more attention in research.

3.3.6 Data management
Due to the rise of IoT and the significant increase in smart devices generating data, data management has become a crucial issue in the field of edge computing. As a result, efficient methods for data management have received considerable attention in recent years. These methods include collecting, storing, caching, and processing data. Previously, these operations were performed in remote clouds with satisfactory results. However, due to the surge in data volume and the bandwidth limitations of the backbone network, this is no longer feasible. Edge computing offers a solution to this problem by sending data to the nearest edge equipment instead of a remote cloud. After reviewing papers in this field, several research sub-topics were identified, which we will explain below.

Caching
This research sub-topic pertains to network edge caching. Caching is a widely used technique to improve data access speed, decrease the load on backbone networks, and enhance user experience on the edge. This technique involves caching frequently used data such as videos, news, weather updates, and other popular content at the network’s edge. Consequently, the content access delay and the load on the backbone network are reduced for the users. Moreover, the implementation of caching can significantly enhance performance. Predicting the popularity of the content that should be cached is one of the primary issues in caching. Various methods have been utilized in the reviewed papers to address this problem, such as learning techniques, location utilization, and social network-based methods. Furthermore, caching can improve the performance of computation offloading in wireless networks. The objective is to cache the results of computation and avoid processing the same task again.

Data analysis
In the field of IoT, various services and applications, including smart cities, smart transportation, and healthcare, rely on data. To analyze and extract knowledge from data collected from UEs or sensors, it is necessary to send the data to data centers. Edge computing is an effective solution that allows performing some or all of the data analysis operations at the network’s edge. This approach reduces the amount of data sent to data centers and enhances the users’ request response time. Data analysis at the edge can include various operations, from simple tasks such as data aggregation, preprocessing, matchmaking, and filtering to more complex ones such as data mining applications. Researchers have also explored the use of big data techniques, such as stream processing, and their adaptation to the edge infrastructure in this field.

Data distribution
With the rise of new applications, cloud-based data storage has become practically inefficient due to high traffic between the user and the cloud, as well as high latency and cost. To solve this problem, the distribution of data between partner devices at the edge can be used, despite the limited storage resources available at the edge compared to the cloud. However, significant challenges must be considered in this field, including node mobility and the determination of the geographical location of data. For example, in Reference, an architecture is presented for distributing large volumes of data across automotive networks for content sharing. Content placement is also a challenge in most CDN applications, including which involves distributing content appropriately between resources and devices at the edge.

Data dissemination
Data dissemination refers to the process of determining the optimal path for data transfer, which can occur between UEs, nodes at the edge of the network, and between edge nodes and remote cloud. In one study, a k-means clustering algorithm was used to transmit data efficiently over longer distances with less energy consumption in a Low-Power Wide-Area Network (LPWAN). Devices were clustered based on traffic priorities to optimize data transfer. Another study proposed a combined data propagation framework using SDN and Delay-Tolerant Networking (DTN) to reduce data costs and traffic between fog and cloud nodes. The control plane is located in the cloud, while the data plane resides at the fog nodes.

Data replication
Data replication is a widely used technique in cloud computing, where a particular set of data is stored in multiple data centers with different geographic locations to enhance its availability and reliability. This method is also employed in edge computing to improve access speed, QoS,
and bandwidth consumption. However, the replication mechanisms designed for cloud computing may not be appropriate for edge servers due to limited resources and diverse traffic in edge RANs. Therefore, new replication methods need to be developed for the edge computing paradigm.

### 3.3.7 Economics

This topic pertains to research that considers economic factors alongside other criteria. The number of papers published in this field is comparatively low, indicating significant research opportunities in this area. In the reviewed papers, economic theories are utilized to enhance resource productivity, maximize profits, and create a fair environment between providers and customers.

The primary economic research problem examined in the reviewed papers is pricing, which deals with designing mechanisms to determine the price of resources and establishing penalties for violating the guaranteed QoS. Pricing is a strategic issue between the supplier and the buyer, with the supplier aiming to increase resource productivity and profit, while the buyer seeks to minimize costs while adhering to time and energy constraints. To address this, matching and game theory techniques, such as the Stackelberg game, coalition game, and market game, have been widely employed.

Auction-based pricing is a well-known and widely used technique in the edge computing environment, owing to the limitation of edge server resources and the fluctuation of user requests, which may create a competitive environment, especially during peak times. An auction is one of the most common ways to allocate resources efficiently and fairly in such cases. In an auction, there are typically two main agents: the seller and the buyers. In edge computing, buyer agents are often mobile users, and seller agents are service providers or cloudlets. However, in some cases, the mobile users themselves can also play the role of the seller. When a buyer agent needs a resource, he/she sends a request with a bid price. Then, the sellers select the buyers based on the offered bid prices and provide resources to the winning buyers. In edge computing, auctioned items can be either processing resources alone or processing and communication resources together.

### 3.4 How are applications distributed among the various research topics?

In this section, we will discuss the common applications that are suitable for running on the edge platform and investigate the distribution of these applications in the field of edge computing. These applications have been extracted from various papers and are listed in Table 5. It is important to note that papers that introduce their innovation in a general way without limiting it to a specific area are categorized as General, and papers with very few repetitions that cannot be classified into any existing categories are categorized as Other. In the following comparison between applications, we will focus on the most commonly used applications and neglect the Other and General categories to provide a more detailed analysis.

**Table 5** The distribution of applications per topic

| Application      | Co | RM | DM | Ar. | NM | S&P | Ec. |
|------------------|----|----|----|-----|----|-----|-----|
| General          | 185| 311| 114| 99  | 140| 92  | 28  |
| IoT              | 76 | 132| 60 | 75  | 55 | 75  | 9   |
| Smart Community  | 9  | 30 | 22 | 36  | 13 | 20  | 1   |
| Vehicular        | 28 | 50 | 32 | 38  | 39 | 24  | 3   |
| Healthcare       | 8  | 14 | 30 | 28  | 8  | 19  | 0   |
| Streaming        | 10 | 18 | 19 | 7   | 8  | 3   | 0   |
| Industria        | 14 | 12 | 9  | 12  | 6  | 4   | 0   |
| Crowdsensing     | 2  | 5  | 3  | 8   | 2  | 10  | 0   |
| AR/VR            | 2  | 9  | 3  | 4   | 3  | 0   | 0   |
| Other            | 10 | 15 | 25 | 16  | 11 | 12  | 1   |

*Smart community includes smart city, smart home, smart grid, etc.*
3.4.1 IoT applications

IoT refers to the process of connecting a large number of physical devices to the internet to sense, collect and share data. In recent years, IoT has gained significant attention and has been utilized in various applications. However, the high volume of data collected by IoT devices, which usually have limited memory and processing power, has made data processing challenging. Initially, cloud computing was suggested as a solution to this issue, but it had some limitations. The most critical barriers to using cloud computing in IoT applications are bandwidth and latency problems in sending and receiving data. Therefore, edge computing has emerged as a promising solution. Edge computing can overcome network-related issues and reduce latency by transmitting data to computing resources located at the edge of the network. As shown in Table 5, IoT papers have the highest number of publications, indicating the strong compatibility of IoT applications with edge computing. These papers are mainly categorized into resource management, security, architecture, and computation offloading topics.

3.4.2 Vehicular applications

Vehicular applications are designed to enhance the driving experience, and they are among the most frequently discussed applications in edge computing research papers. Researchers have come up with various useful applications in fields such as traffic management and parking reservation to amaze vehicle drivers. However, many of these applications, including autonomous driving and traffic management, require specific processing, storage, and communication capabilities. Edge computing offers an excellent solution to meet the requirements of vehicular applications with its proximity to users, low latency, high mobility support, real-time communication, context-awareness support, and low development costs. Nevertheless, network stability issues due to vehicles' mobility and the fast response requirement of these applications are among the most important challenges. As a result, most papers on this application fall into the resource management, network management, and architecture topics, which deal with these issues.

3.4.3 Smart community

This category encompasses all applications that have utilized artificial intelligence techniques to enhance their capabilities in the smart community. Most papers related to these applications belong to the architecture topic. Some examples of such papers are managing smart city applications, connecting IoT devices to a smart hospital, and presenting an architecture for managing big data in real-time for smart transportation applications. All of these architectures are based on edge/fog computing.

3.4.4 Healthcare

Healthcare applications involve monitoring systems that collect patient data either directly from the patient or through wearable sensors. These systems analyze the patient’s daily activities, eating habits, and sleeping patterns to provide helpful recommendations for a healthy lifestyle or to detect medical emergencies. The large amount of data generated by continuous sensing and the need for rapid response in emergency situations have made healthcare applications a popular topic in edge computing research. Due to the personal nature of health data, privacy is a crucial concern. Therefore, data management, architecture, security, and privacy are important topics in this field.

3.4.5 Streaming

This category encompasses applications that require continuous and real-time processing of data streams. One of the most common examples of such applications is multimedia streaming. These applications demand real-time processing, and high latency during execution can significantly degrade the user experience. Edge computing can be leveraged to enhance the quality of multimedia services by bringing processing and storage resources closer to user devices. Resource management and bandwidth allocation to control latency and jitter, as well as managing cached data, are crucial issues in this category of applications. Therefore, articles related to these applications are mostly classified in the resource management and data management topics.

3.4.6 Industrial

Industry applications are another category that researchers of edge computing have targeted. With the transition to the fourth revolution of manufacturing, also known as Industry 4.0, issues such as communication between industrial equipment and data analysis have been raised in order to
increase automation in decision-making and reduce human involvement. However, these applications require the rapid processing and analysis of large volumes of generated data, which can be conveniently performed through edge computing.

3.4.7 Crowdsensing

Crowdsensing is an application category that involves distributing large-scale tasks among a group of individuals who voluntarily participate through their mobile devices. These tasks often involve sensing data through mobile sensors, reporting on a situation, or completing a small part of a larger job. Companies use crowdsensing to acquire large amounts of low-cost data for their online services. However, due to the widespread adoption of this application and the increase in sensed data traffic, edge computing has become a suitable option for implementing crowdsensing. Nonetheless, this type of application faces various threats due to its voluntary nature, including sabotage, false data submissions, and user privacy concerns. Therefore, privacy and security are major challenges in this field, and many articles related to this application focus on privacy and security topics. To address these challenges, several reputation management and privacy-preserving techniques have been presented in the literature on edge computing.

3.4.8 AR/VR

Augmented and virtual reality applications have gained significant attention in recent years. Augmented reality tags and identifies objects, or adds extra elements to live videos captured on user smartphones, while virtual reality allows users to experience virtual interactions through VR headsets. These applications require fast data communication and analysis to provide users with immersive, real-time interactions, making them a popular research topic in edge computing. Similar to streaming applications, augmented and virtual reality are highly delay-sensitive, and reducing latency is crucial to improve the quality of service. Proper resource allocation, task placement, and decision-making policies can achieve this. Therefore, the majority of articles related to these applications are classified under the resource management topic.

3.5 How are various architectures distributed across different research topics?

In this section, we present a statistical analysis of edge architectures used in different topics among the reviewed papers. Figure 9 displays the frequency of utilization of different architectures by researchers. The horizontal axis represents the topics, the vertical axis represents the architectures, and the volume of the bubbles represents the number of articles utilizing the related architecture in the relevant topics. The primary edge computing architectures include fog, MEC, and cloudlet. However, in some edge computing studies, none of these architectures are explicitly mentioned, and only the term “edge computing” is used. In this study, we have categorized these articles as general architecture and investigated them under the same heading as edge computing.

Among the reviewed papers, various architectures have been proposed to exploit nearby resources. The most widely used and well-known architectures include fog, cloudlet, mobile edge computing (also referred to as multi-access edge computing), and mist computing. Although these concepts have similarities and are sometimes used interchangeably in the literature, there are some differences between them.

Fog computing is an architecture that utilizes devices located near the user, including servers, computers, routers, and switches on the network edge, to process or pre-process data before transmitting it to the cloud. This architecture is hierarchical, where all available processing capabilities...
along the user’s path to the remote cloud are utilized. On the other hand, the cloudlet architecture employs high-power computers or clusters of computers that are situated close to the user. These cloudlets are often virtualized, and their resources can be used by surrounding mobile devices to boost their processing power. Cloudlet operators can be cloud service providers who aim to offer their services in proximity to the user.

MEC is an evolution of mobile computing that incorporates edge computing, where IT and cloud computing capabilities are delivered through the RAN (Radio Access Network) in 5G and 4G. This architecture provides edge computing services such as processing, storage, and network through the telecommunication network infrastructure. It was subsequently expanded to encompass a broader range of applications beyond the specific functions of mobile phones and is now known as Multi-access Edge Computing. Lastly, Mist computing is a computational model that enables scattered computing to be carried out at the most end nodes, that is, the IoT devices themselves. Since this architecture is managed by the devices themselves, it presents more complexities compared to other architectures.

The diagram in Figure 9 reveals that the cloudlet architecture has received relatively less attention compared to the other three architectures. While the cloudlet architecture was an important early model for improving the speed and quality of mobile devices by bringing cloud services closer to users, newer and more versatile architectures have emerged, leading the research community to explore alternative options. Fog computing, for instance, offers a more flexible solution by enabling resources to be located anywhere along the edge of the network, from edge devices to the remote cloud. In addition, recent advancements in mobile networks have prompted operators to focus on MEC as the main edge server technology, replacing cloudlets. Furthermore, there are significant variations in the number of papers on different architectures in different topics, which may be attributed to differences in their structures, features, and applications. We will discuss some of these features in the following sections.

There have been numerous research studies conducted on the MEC architecture in relation to offloading, as shown by a significant number of publications on this topic. Offloading is a technique that is commonly used to reduce energy consumption and increase the speed of execution for mobile applications that are constrained by limited memory, processing power, and energy. A considerable proportion of the publications on offloading are concerned with mobility management, decision making, and partitioning for mobile applications, particularly those that operate on cellular networks and the MEC architecture.

In the resource management topic, the number of publications in three main architectures has a high number as it is the most challenging topic in the edge computing. In the data management topic, the studies conducted on the MEC architecture is also less than the other two architectures. According to this study, MEC applications are mostly related to mobile applications and less used for data-intensive applications. One reason for this is the constraints on the location of the servers in the MEC architecture. The MEC architecture is placed on a RAN, in which the servers must be adjacent to the cellular network base stations, but in the other architectures, there is no such limitation. In MEC architecture, it is not possible to bring servers closer to the data source. Another reason for the fog preference to MEC architecture in data management topic is that the multi-layered architecture of fog and the presence of cloud support so that, fog nodes can be used to pre-process or filter data before sending it to the cloud for storage or analysis.

In the architecture topic, MEC had a lower number of publications compared to fog and edge architectures. Research in this category aims to increase the efficiency of a specific application or improve a general parameter by modifying standard architectures. MEC architecture is more structured and less flexible than other architectures because it is designed to exist within a regulated telecommunications infrastructure. This means that MEC architecture has limited flexibility for structural changes. On the other hand, other architectures provide more flexibility to incorporate new architectural ideas and make changes.

Finally, in the security and privacy topic, fog architecture has the highest number of publications. Unlike MEC architecture, fog nodes are not owned by a provider and are provided by independent individuals. Since fog nodes are often deployed in areas with relatively weak protection, they may encounter various malicious attacks. In addition, devices in the fog are usually deployed without strict monitoring and protection. Therefore, security issues in this architecture have received more attention than in other architectures. Another notable point in Figure 9 is that the number of publications on the economic topic is lower than in other areas, but the two paradigms of MEC and edge are more favored than others. This indicates the particular need for designing pricing and auction models in the edge computing environment.

The papers also mentioned paradigms such as osmotic computing, mist computing, and application-centric computing, which did not have significant statistical representation. For instance, Sharma et al. used osmotic computing to propose a trust management framework for social network applications. Rahman et al. proposed integrating a five-tier cloud, fog, and mist computing environment for IoT applications in healthcare and next-generation e-healthcare systems. Their method aims to handle and route offline/batch data in real-time with high QoS and low end-to-end latency.

3.6 Which techniques are commonly utilized in the field?

The emergence of edge computing has brought about new challenges in distributed resource development and management. Moving storage and computation to the edge of the network improves Quality of Service (QoS) and Quality of Experience (QOE), particularly in applications where low latency is critical. However, efficiently utilizing heterogeneous resources to meet user needs is a challenging task. To address these challenges, various techniques have been proposed in the literature, which are summarized in Figures 10 and Table 6. Figure 10 and Table 6 show the
percentage of techniques used in different topics. As some articles may be categorized into more than one topic, the sum of the percentages may exceed one hundred. Figure 11 shows the number of publications per topic. In the following section, we will explain these techniques and their wide usage in edge computing related issues.

3.6.1 Heuristic techniques

Heuristic techniques are a useful solution for finding near-optimal values in situations where full optimization is difficult and time-consuming. The term “heuristic” refers to a method that, based on experience or judgment, is likely to produce a near-optimal solution to a problem. These techniques start with a well-defined mathematical representation of a problem. In edge computing applications, where time is a critical QoS criterion, heuristic methods that can find near-optimal solutions quickly have been widely used by researchers in this field. The key features of heuristic methods are their ease of implementation, ability to show improvement in each iteration, fast production of results, and robustness. Heuristic techniques are most commonly used in resource management, where studies have attempted to optimize energy, latency, cost, and computational complexity. Heuristic techniques are generally applicable methods for solving scientific problems, and therefore, this technique is commonly used in almost all topics, particularly in resource management.
3.6.2 Linear and non-linear programming

Linear programming is a mathematical approach used for solving complex optimization problems by modeling objective functions and constraints with linear functions. However, if any of the functions in the problem statement involve non-linear variables, it becomes a non-linear optimization problem. Linear programming falls under the category of convex optimization problems, which are formulated as a set of objectives and constraints. Some non-linear problems, such as MINLP, are non-convex in nature and can be challenging to solve. The difficulty in solving these problems lies in the fact that local optimal solutions may not be equivalent to global optimal solutions.

Convex and non-convex optimizations are centralized methods, meaning that they require a centralized agent to gather all necessary information and make decisions on behalf of all involved. As a result, they are not scalable and are only suitable for small scenarios. Additionally, due to their static nature, they are not well-suited for dynamic applications. However, as previously mentioned, they can achieve global or near-global optimization.

As illustrated in Figure 10 and Table 6, this technique has garnered significant attention in the resource management field compared to other topics because it is a common method for solving a wide range of optimization problems in resource management. For instance, some of these optimization problems include balancing energy consumption and delay at the edge, efficiently and effectively transporting, storing, placing, and processing vast amounts of IoT data, maximizing storage utilization while simultaneously reducing service latency and improving energy savings, and optimizing task offloading in mobile edge computing environments.

3.6.3 Game theory

In IoT-based computing environments, such as edge and fog computing, game theory techniques are used to address the competition among different entities. In such networks, there are distributed resources that need to be allocated fairly among IoT devices and users while also taking into account the profit of service providers. To tackle this challenge, researchers have turned to game theory, which has become widely adopted in papers after heuristic and learning techniques. Game theory techniques are decentralized, meaning that each user makes their own decisions locally. As a result, these methods are scalable and can handle large and dynamic scenarios. However, a limitation of game theory techniques is that they only lead to the Nash equilibrium, which may not always be optimal.

An important observation from Table 6 is the relatively high proportion of economic articles that utilize this technique compared to other techniques. Within the economic topic, most articles pertained to pricing, a strategic issue in which each agent aims to maximize their benefits. Game theory techniques, therefore, offer an effective solution for such problems. Stakelberg game, coalitional game, and matching games are among the most commonly used techniques in edge computing research papers. They are well-suited to addressing common edge problems such as resource allocation, sharing, and pricing.

3.6.4 Meta-heuristic techniques

Meta-heuristic techniques are a well-known group of optimization techniques that are widely used in scientific research to find good solutions for NP-hard problems. These techniques are often inspired by nature, and many problems can be effectively modeled using them. The fast convergence
feature of meta-heuristic algorithms makes them a great solution for time-constrained edge problems such as resource management and task allocation.

As can be seen in Table 6, similar to heuristic techniques, meta-heuristic techniques are predominantly used in resource management and offloading topics for the same reasons. Some widely used meta-heuristic algorithms in edge computing research include genetic algorithms, ant colony optimization, simulated annealing, and particle swarm optimization.

3.6.5 Markov decision process

The Markov Decision Process (MDP) is a mathematical technique used to model sequential decision-making problems under uncertainty. It involves the interaction of an agent with the environment and consists of a sequence of decision problems. MDPs can be used to model the state evolution dynamics of a stochastic system controlled by an agent that selects actions at each time step. The process of selecting such actions is called an action policy, which can be history-dependent or only consider the current state.

One advantage of MDP is that it can be used in both distributed and centralized scenarios, making it suitable for modeling problems in edge computing. However, it faces two main challenges. First, to reach the optimal state, decision makers need knowledge of the system state to make informed decisions, which requires addressing the data collection problem. Second, MDPs face the problem of curse of dimensionality, meaning the state space of the corresponding dynamic problem is intractably large, making it not scalable. As a result, classic MDP algorithms such as value iteration policy and policy iteration policy are not suitable for addressing the MDP problem. In reviewed studies, methods such as reinforcement learning have been used to solve these problems.

Table 6 shows that the percentage of articles about offloading is significantly higher than the other topics, which is likely because this technique is a useful option for modeling a wide range of edge computing decision-making problems that are raised in offloading topics. For example, some of these problems include task offloading decision-making, selecting the next hop in a software-defined vehicular network, and random access processes in M2M communications.

3.6.6 Lyapunov optimization

The Lyapunov optimization technique is a useful method for optimizing performance objectives subject to queue stability, either real or virtual. This technique is particularly effective in solving problems that deal with the tradeoff between utility maximization and delay or any time-average constraints. Edge computing problems, such as resource management and offloading, can be highly complex. Using this optimization method, the computational complexity of edge computing problems can be significantly reduced. For this reason, the use of Lyapunov optimization in modeling edge computing problems is widespread. Although Lyapunov optimization is a centralized modeling method, due to its low computational complexity, it can be used in large-scale scenarios. However, Lyapunov optimization only achieves local optimization because it is a single step optimization method and does not repeat the process.

The Lyapunov optimization technique is widely employed in solving optimization problems in edge computing, with various criteria. For instance, some works utilize this technique to minimize power consumption while adhering to quality of service constraints for users, while others minimize communication resource consumption subject to delay constraints. Additionally, some works aim to minimize energy consumption for task execution while preventing free-riding behavior of users by imposing incentive constraints.

Table 6 reveals that this technique has been predominantly used in resource management, offloading, and network management topics, where stabilizing and balancing the system along with other criteria are common issues. Therefore, using the Lyapunov optimization technique is a suitable option for optimizing such problems.

3.6.7 Graph theory

Graph theory is a mathematical field concerned with studying problems that involve graphs, and it finds applications in a variety of scientific and engineering domains. Graph-based techniques offer an effective way to solve problems that can be represented by objects and their relationships. In edge computing, graph-based techniques are among the most commonly employed, with many articles modeling the edge computing platform as a graph, where nodes represent computing servers and edges represent the links between them. Additionally, some applications in this field are represented by graphs where nodes signify tasks or functions, and edges denote the dependencies between them.

According to the statistics presented in Table 6, the use of graph theory techniques in resource management is significantly higher than in other topics. This is likely due to the fact that resources in edge computing are distributed and can often be modeled as graphs. Among the reviewed papers, the most commonly used graph theory techniques are graph coloring, graph matching, and partitioning.
3.6.8 Learning techniques

Learning techniques refer to the process of equipping computers with the ability to learn from data without requiring explicit programming.\textsuperscript{218} These techniques aim to improve computer systems automatically through experience.\textsuperscript{219} Machine learning techniques are particularly effective in handling complex problems, such as those involving high dimensions, where they can overcome the curse of dimensionality in MDP. Moreover, online learning techniques, which are especially suitable for dynamic environments, can be applied to both centralized and distributed applications. Since the edge computing environment is highly dynamic due to user mobility and workload changes, researchers have explored the use of learning methods to predict the state of the environment and make intelligent decisions. Additionally, learning methods are widely used to extract valuable information from raw data. This is particularly relevant in the context of edge computing and IoT devices, which generate large volumes of raw data that must be analyzed to derive meaningful insights. Therefore, the widespread adoption of learning techniques in research papers is motivated, in part, by the need to address this issue.

According to Table 6, data management articles have made greater use of this method, as these techniques are closely tied to data, both for discovering new information and improving accuracy. Reinforcement learning,\textsuperscript{220,221} neural networks,\textsuperscript{222,223} and clustering\textsuperscript{110,224} are among the most commonly used learning techniques in the field of edge computing literature.

3.7 What forms of empirical evaluation have been employed in the field?

Figure 12 presents an overview of the different evaluation methods employed in the reviewed articles. Among the most common methods found in research papers were simulation, real-system implementation, implementation on a testbed, and proof of concept methods. Notably, the majority of proposed methods were tested and evaluated using simulation, likely due to the lack of real-world edge infrastructure or the high costs and complexity associated with real-world evaluations. Examples of simulation tools used in edge and cloud computing research include FogNet-Sim,\textsuperscript{225} iFogSim,\textsuperscript{226} EdgeCloudSim,\textsuperscript{227} and IoTsim.\textsuperscript{228} Real-system implementation, on the other hand, encompasses a range of approaches, including the development of prototypes (21%) and scenarios (4%), designing a case study (12%), utilizing a real dataset (5%), using benchmarks (5%), use cases (5%) and platforms (1%).

3.8 Which qualitative requirements have primarily been considered in the move towards edge computing?

Figure 13 illustrates the different QoS categories frequently reported in the reviewed papers. This section presents statistics on the primary QoS category based on the SMI\textsuperscript{229} and ISO 9126\textsuperscript{230} quality models as they relate to the edge environment. Efficiency emerged as the most commonly cited category among the studies. Efficiency is defined as “a set of attributes that bear on the relationship between the level of performance of the software and the amount of resources used, under stated conditions.”\textsuperscript{230} Time (49%), energy consumption (26%), utilization (22%), and throughput (3%) had the most frequency among the efficiency criteria. As previously mentioned, applications in the edge computing environment are latency-sensitive, and reducing this delay has been the impetus for their migration from the cloud to the network’s edge. Therefore, it is natural that the time criterion, specifically delay and latency, should be regarded as the primary consideration when evaluating the QoS of such programs.

![Figure 12](image-url) Evaluation methods in edge computing.
The second category, following the efficiency category, is the financial category, which focuses on the amount of money spent by the client on the service.\(^{229}\) In this category, cost is the only criterion considered. The performance category, on the other hand, concentrates on the "attributes of the features and functions of the provided service."\(^{229}\) It is worth noting that 52% of studies only reported performance as their quality criterion without any mention to its sub-criteria. Some studies reported accuracy (48%) as a performance criterion in their evaluation.

The next category relates to assurance, which is concerned with the "attributes that indicate how likely it is for the service to be available as specified."\(^{229}\) Reliability (50%), Availability (26%), Failure (1%), Stability (19%) and Robustness (4%) are the most frequent criterion in this category. Agility is defined as "attributes including the impact of a service upon a client’s ability to change direction, strategy, or tactics quickly and with minimal disruption."\(^{229}\) Agility parameters such as flexibility (29%), elasticity (2%) is reported frequently in studies.

### 4 COMPARISON WITH SIMILAR REVIEW RESEARCHES

As previously mentioned, at the start of the search process, the expert randomly selected several secondary studies (surveys and reviews) as a starting point. As the search process progressed, additional review studies were discovered. A list of these studies is provided in [SuppFil]_E2,T2). Among the review studies, those that were specific to a particular sub-topic were excluded. For the remaining studies, we extracted the covered topics in the field of edge computing, and the details of this information are provided in [SuppFile]_E2,T6). Table 7 presents a comparison between our SMS and other review studies. Note that PID column represents Paper ID which can be found in [SuppFil]_E2,T2). Unlike our SMS, none of these studies used a systematic approach to cover all the related journal studies in the field of edge computing.

However, we selected seven papers, namely References 1,2,40,231-233, and 234 for comparison as they are relatively more comprehensive and cover a wider range of aspects in the field of edge computing. Table 8 illustrates the differences between our SMS and the other studies being compared. For each of the studies, we extracted the necessary information about the number of covered venues and studies. As shown in Table 7, our SMS, indicated in the last row, reviewed and covered a significantly larger collection of related studies compared to the other reviews. For instance, in the review paper,\(^{2}\) only 120 papers out of 450 cited were used in the taxonomy. Furthermore, these studies in the field of edge computing were not subjected to a systematic process and addressed fewer topics and sub-topics, as detailed in [Supp-File]_E2,T3). In contrast, our SMS employed a comprehensive and complete process to find related studies in a collection of high-quality journals.

### 5 IMPLICATIONS OF FINDINGS

This paper presents the results of a systematic mapping study (SMS) aimed at creating a guide for researchers, practitioners, and professors working in the field of edge computing. The findings of this study can be valuable to a diverse range of audiences. The paper is expected to be of interest to three primary categories of readers, namely researchers, industry practitioners, and university professors involved in this field. As such, this section provides specific tips and guidelines tailored to the needs of each of these three main groups of audiences.
### Table 7: Comparison between this SMS and other related reviews.

| PID*  | Topic | Scope | Ac. | NM | DM | RM | Ec. | CO | S&P |
|-------|-------|-------|-----|----|----|----|-----|----|-----|
| 68    | FOG   | ✓     | x   | x  | ✓  | x  | ✓   | x  | x   |
| 1610  | MEC   | ✓     | x   | x  | ✓  | x  | ✓   | x  | x   |
| 1611  | EC    | ✓     | x   | x  | ✓  | x  | x   | x  | ✓   |
| 1618  | MEC   | ✓     | x   | x  | ✓  | x  | ✓   | x  | x   |
| 5439  | FOG   | ✓     | ✓   | x  | ✓  | x  | x   | x  | x   |
| 6838  | EC    | ✓     | x   | x  | x  | x   | x  | x   |
| 5715  | EC    | ✓     | x   | ✓  | ✓  | ✓  | x   | ✓  | x   |
| 7906  | FOG   | ✓     | x   | x  | ✓  | ✓  | x   | ✓  | ✓   |
| 7907  | FOG   | ✓     | x   | ✓  | ✓  | ✓  | x   | x  | ✓   |
| 8044  | EC    | ✓     | ✓   | x  | ✓  | x   | x  | x   |
| 8242  | EC    | ✓     | ✓   | x  | ✓  | ✓  | x   | x  | ✓   |
| 8587  | EC    | ✓     | x   | x  | ✓  | ✓  | x   | ✓  | x   |
| 8775  | EC    | ✓     | x   | ✓  | ✓  | ✓  | x   | x  | ✓   |
| This SMS | EC      | ✓     | ✓   | ✓  | ✓  | ✓  | ✓   | ✓  | ✓   |

*PID means Paper ID which can be found in [SuppFile] (E2T6).

### Table 8: Statistical comparison between this SMS and other selected reviews.

| Reference | Type | Studies# | Venue# | Searching method | QC | Eval | Period |
|-----------|------|----------|--------|------------------|----|------|--------|
| 231       | Survey | 22/169   | 41     | N/A              | x  | x    | 2013–2017 |
| 2         | Survey | 120/450  | 40     | Manual           | x  | x    | 2013–2017 |
| 1         | Survey | 37/137   | 29     | N/A              | x  | x    | 2004–2017 |
| 232       | Survey | 27/99    | 14     | N/A              | x  | x    | 2002–2019 |
| 233       | Survey | 97/296   | 55     | N/A              | x  | x    | 2002–2019 |
| 40        | Survey | 70/200   | 43     | N/A              | x  | x    | 2009–2019 |
| 234       | Survey | 36/101   | 19     | N/A              | x  | x    | 2009–2019 |
| This SMS  | SMS   | 1440/1440 | 112   | Manual-Snowballing-Database | ✓  | ✓    | 2001-2019 |

Abbreviations: Eval, Search Evaluation; QC, Quality Criteria.

### 5.1 Implications for researchers

In recent years, the field of edge computing has gained immense popularity, and a substantial amount of research has been carried out in this area. In this SMS, we identified active research topics in the field of edge computing by reviewing a large number of articles. This study can serve as an excellent starting point for new researchers who are interested in edge computing (RQ3). Our analysis also revealed significant differences in the number of studies conducted across different countries (RQ2). China, the United States, and Canada are the countries with the highest percentage of published articles, accounting for 31%, 11%, and 7% respectively. These countries have advanced technology in their industries compared to other nations, and there is a strong connection between their industries and academia. Furthermore, the study identifies the leading researchers in each field (RQ2), which is valuable information for researchers seeking international collaborations.

Over the past decade, the field of edge computing has made significant progress, and numerous architectures for utilizing edge resources have been proposed (RQ5). However, some researchers may misunderstand these architectures or use them interchangeably in the literature. Thus, it is essential for researchers in this field to have a complete understanding of standard edge architectures and to pay attention to the differences and
characteristics of each of them. This knowledge helps researchers to gain a deep understanding of the fundamental concepts and propose effective and accurate solutions.

Most of the reviewed papers in the field of edge computing use simulation-based evaluation methods (RQ7). This is because the complete infrastructure for edge computing is not yet available in many parts of the world, and simulation provides a simple and inexpensive way to evaluate research in this area. As a result, various simulators such as FogNetSim++ and iFogSim have been proposed to facilitate research in this field. However, simulations often involve simplifications and hypotheses that may not reflect reality accurately. To address this, real testbeds and small prototypes can be implemented at the laboratory level to more accurately assess and identify operational challenges and problems. For instance, the Raspberry Pi can be used to build a realistic platform for researchers to create desired scenarios.

Another goal of this study was to identify and introduce common applications and techniques used in edge computing research. Optimization methods are widely used for load balancing and increasing efficiency, and knowledge of machine learning, optimization, meta-heuristics, or game theory can be practical for researchers in this field. A separate section of the study explored common applications in edge computing papers, which researchers can use to become familiar with the techniques and applications in this field (RQ6). Familiarity with these applications and techniques can aid researchers in reading and understanding articles in this field.

5.2 | Implications for practitioners

After years of research, the field of edge computing is approaching maturity. In academia, extensive research has been conducted to leverage resources in proximity to users and to address various challenges in this field (RQ1). As a result, a vast source of scientific solutions, ideas, and applications has been developed. To apply this knowledge to real-world problems and enhance daily life quality, industrial organizations must collaborate with scientific and academic departments. Such collaboration can be achieved by participating in conferences related to edge computing, facilitating knowledge transfer from researchers to industry professionals and vice versa. This connection can expedite the realization of edge computing solutions in the real world.

The emergence of the Internet of Things (IoT) has introduced new applications, some of which cannot be implemented without utilizing edge resources (RQ4). Since these applications are expected to become an integral part of people's daily lives in the future, developing and advancing edge computing platforms will be a necessity. Application and service developers should consider utilizing edge resources and distributed platforms in their designs. Additionally, data network operators and network equipment manufacturers should consider integrating the edge computing model in the design of future network infrastructure or communication equipment.

5.3 | Implications for teachers

Edge computing has emerged as a promising solution in the past decade to enhance service quality and decrease response time by bringing resources closer to users' devices. The field has witnessed substantial growth in recent years, making it a vast, independent, and popular research domain among scholars (RQ1). Therefore, it is crucial to introduce students and novices to the concepts of edge computing, to facilitate knowledge transfer and disseminate research accomplishments of recent years. Professors who specialize in cloud computing and computer networks can incorporate edge computing concepts into courses such as “cloud computing” and “distributed systems,” or refer to them indirectly in related courses such as “computer networks” and “internet engineering.” Professors can also utilize the concepts presented in section (RQ3) and the statistical results derived from the reviewed papers in sections (RQ2, RQ3) to familiarize themselves with edge computing concepts and develop educational content.

6 | CONCLUSION

As the trend towards edge computing continues to rise and the number of studies in this field increases dramatically, there is a need for a systematic review of the literature. In this paper, we conducted a SMS in the field of edge computing from 2001 to 2020. To achieve this aim, we used well-established methodologies for systematic reviews, as well as our previous experience in publishing such studies. The SMS methodology we used primarily focused on providing a robust search method. We conducted the search at three levels, including manual search, backward snowballing search, and database search. We defined criteria for selecting studies based on their quality and relevance to the field of edge computing. Out of the 8725 studies obtained, we selected 1440 studies that met these criteria. We also evaluated the search methodology to ensure that it provided adequate coverage of the studies.

Based on the methodology used, we developed eight research questions, which we answered during the SMS. The first three questions focused on reviewing the number of studies done (RQ1), introducing the most important researchers and pioneers in the field of edge computing (RQ2), and extracting the most important topics and subtopics in this field (RQ3). The next three questions (RQ4, RQ5, and RQ6) reviewed and analyzed
the most crucial applications of edge computing in each topic, the most important architectures used in each topic, and the techniques used in each topic, respectively. Questions RQ7 and RQ8 focused on reviewing the most important evaluation methods and quality criteria used in research to advance the field of edge computing. Additionally, we compared our SMS with several reviews in the field of edge computing to examine the depth of the methodology used and the various aspects considered.

Our research shows that edge computing is an active and growing research field in various geographical areas. Although the number of studies in this field has significantly increased, there are still challenges, and the development of methods and paradigms is necessary to address these challenges. Systematic review studies such as our SMS can provide more detailed reviews of each aspect using systematic literature reviews (SLRs). For future work, each of the high-level topics extracted in this paper can be explored in-depth to answer more specific research questions.

CONFLICT OF INTEREST STATEMENT
The authors have no relevant financial or non-financial interests to disclose.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are openly available in Edge Computing SMS at https://github.com/jalalsakhdari/SMS.

ORCID
Saeid Abrishami https://orcid.org/0000-0003-1421-5179

REFERENCES
1. Li C, Xue Y, Wang J, Zhang W, Li T. Edge-oriented computing paradigms: a survey on architecture design and system management. ACM Comput Surveys. 2018;51(2):1-34.
2. Yousefpour A, Fung C, Nguyen T, et al. All one needs to know about fog computing and related edge computing paradigms: a complete survey. J Syst Archit. 2019;98:289-330.
3. Baktir AC, Ozgavde A, Ersoy C. How can edge computing benefit from software-defined networking: a survey, use cases, and future directions. IEEE Commun Surveys Tutorials. 2017;19(4):2359-2391.
4. Varghese B, Buyya R. Next generation cloud computing: new trends and research directions. Futur Gener Comput Syst. 2018;79:849-861.
5. Kitchenham BA, Budgen D, Brereton OP. Using mapping studies as the basis for further research—a participant-observer case study. Inf Softw Technol. 2011;53(6):638-651.
6. Petersen K, Vakalakana S, Kuzniarz L. Guidelines for conducting systematic mapping studies in software engineering: an update. Inf Softw Technol. 2015;64:1-18.
7. Ramaki AA, Rasoolzadegan A, Bafghi AG. A systematic mapping study on intrusion alert analysis in intrusion detection systems. ACM Comput Surveys. 2018;51(3):1-41.
8. Mayvan BB, Rasoolzadegan A, Yazdi ZG. The state of the art on design patterns: a systematic mapping of the literature. J Syst Softw. 2017;125:93-118.
9. Jafari AJ, Rasoolzadegan A. Security patterns: a systematic mapping study. J Comput Lang. 2020;56:100938.
10. Verbelen T, Simoens P, De Turck F, Dhoept D. Cloudlets: Bringing the Cloud to the Mobile User. Proceedings of the Third ACM Workshop on Mobile Cloud Computing and Services; 2012:29-36.
11. Balan R, Flinn J, Satyanarayanan M, Sinnamohideen S, Yang HI. The Case for Cyber Foraging. Proceedings of the 10th ACM SIGOPS European Workshop. Association for Computing Machinery; 2002:87-92.
12. Kindberg T, Barton J. A web-based nomadic computing system. Comput Netw. 2001;35(4):443-456.
13. Satyanarayanan M. The emergence of edge computing. Computer. 2017;50(1):30-39.
14. Ai Y, Peng Y, Zhang K. Edge computing technologies for internet of things: a primer. Digital Commun Networks. 2018;4(2):77-86.
15. Number of smartphone users by leading countries. https://www.statista.com/statistics/748053/worldwide-top-countries-smartphone-users/
16. Chou W, Lam T, Chung R. Chinese consumers at the forefront of digital technologies china mobile consumer survey 2018. Accessed 28 September 2020 www.deloitte.com/content/dam/Deloitte/cn/Documents/technology-media-telecommunications/deloitte-cn-2018-mobile-consumer-surveyen-190121.pdf/
17. scimago. https://www.scimagojr.com/
18. Zhou H, Wang H, Li X, Leung VC. A survey on mobile data offloading technologies. IEEE Access. 2018;6:5101-5111.
19. Panneerselvam J, Hardy J, Liu L, Yuan B, Antonopoulos N. Mobloulbs: an energy efficient mcc collaborative framework with extended mobile participation for next generation networks. IEEE Access. 2016;4:9129-9144.
20. Ma X, Lin C, Zhang H, Liu J. Energy-aware computation offloading of iot sensors in cloudlet-based mobile edge computing. Sensors. 2018;18(6):1945.
21. Wu H, Sun Y, Wolter K. Energy-efficient decision making for mobile cloud offloading. IEEE Trans Cloud Comput. 2018;8(2):570-584.
22. Zhou B, Dastjerdi AV, Calheiros RN, Srimana SN, Buyya R. Mcloud: a context-aware offloading framework for heterogeneous mobile cloud. IEEE Trans Serv Comput. 2015;10(5):797-810.
23. Min M, Wan X, Xiao L, et al. Learning-based privacy-aware offloading for healthcare IoT with energy harvesting. IEEE Internet Things J. 2018;6(3):4307-4316.
24. Guo H, Liu J, Zhang J, Sun W, Kato N. Mobile-edge computation offloading for ultradense iot networks. IEEE Internet Things J. 2018;5(6):4977-4988.
25. Guo H, Zhang J, Liu J, Zhang H. Energy aware computation offloading and transmit power allocation in ultradense iot networks. IEEE Internet Things J. 2018;6(3):4317-4329.
26. Wang Z, Zhao Z, Min G, Huang X, Ni Q, Wang R. User mobility aware task assignment for mobile edge computing. Futur Gener Comput Syst. 2018;85:1-8.
27. Yu F, Chen H, Xu J. DMPO: dynamic mobility-aware partial offloading in mobile edge computing. Futur Gener Comput Syst. 2018;89:722-735.
28. Liao Y, Han Y, Yu Q, Ai Q, Liu Q, Leeson MS. Wireless body area network mobility-aware task offloading scheme. IEEE Access. 2018;6:61366-61376.
29. Chandrasekaran G, Wang N, Hassanpour M, Xu M, Tafazolli R. Mobility as a Service (MaaaS): a 2d-based information centric network architecture for edge-controlled content distribution. IEEE Access. 2017;6:2110-2129.
30. Cao X, Wang F, Xu J, Zhang R, Cui S. Joint computation and communication cooperation for energy-efficient mobile edge computing. IEEE Internet Things J. 2018;6(3):4188-4200.
31. Carvalho GH, Woungang I, Anpalagan A, Jaseemuddin M. Analysis of joint parallelism in wireless and cloud domains on mobile edge computing over 5g systems. J Commun Netw. 2018;20(6):565-577.
32. Silva FA, Maciel P, Matos R. Smartrank: a smart scheduling tool for mobile cloud computing. J Supercomput. 2015;71(8):2985-3008.
33. Kulatunga C, Bhargava K, Vimalajeewa D, Ivanov S. Cooperative in-network computation in energy harvesting device clouds. Sustainable Comput Inform Syst. 2017;16:106-116.
34. You C, Huang K. Exploiting non-causal cpu-state information for energy-efficient mobile cooperative computing. IEEE Trans Wirel Commun. 2018;17(6):4104-4117.
35. Satyanarayanan M, Bahl P, Caceres R, Davies N. The case for vm-based cloudlets in mobile computing. IEEE Pervasive Comput. 2009;8(4):14-23.
36. Kosta S, Aucinas A, Hui P, Mortier R, Zhang X. 2012 Proceedings IEEE Infocom. IEEE; 2012:945-953.
37. Cuervo E, Balasubramanian A, Cho D, et al. Maui: Making Smartphones Last Longer with Code Offload. Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services. Association for Computing Machinery; 2010:49-62.
38. Shuja J, Gani A, Naveed A, Ahmed E, Hsu CH. Case of arm emulation optimization for offloading mechanisms in mobile cloud computing. Futur Gener Comput Syst. 2017;76:407-417.
39. Zhou B, Buyya R. Augmentation techniques for mobile cloud computing: a taxonomy, survey, and future directions. ACM Comput Surveys. 2018;51(1):1-38.
40. Hong CH, Varghese B. Resource management in fog/edge computing: a survey on architectures, infrastructure, and algorithms. ACM Comput Surveys. 2019;52(5):1-37.
41. Mitya S, Roy A, Majumder K, Phadikar S. Multi-objective optimization technique for resource allocation and task scheduling in vehicular cloud architecture: a hybrid adaptive nature inspired approach. J Netw Comput Appl. 2018;103:58-84.
42. Zhao M, Wang W, Wang Y, Zhang Z. Load scheduling for distributed edge computing: a communication-computation tradeoff. Peer Peer Network Appl. 2019;12(5):1418-1432.
43. Van Do T, Do NH, Nguyen HT, Rotter C, Hegyi A, Hegyi P. Comparison of scheduling algorithms for multiple mobile computing edge clouds. Simul Model Pract Theory. 2019;93:104-118.
44. Mishra SK, Puthal D, Rodrigues JJ, Sahoo B, Dutkiewicz E. Sustainableserviceallocationusingametaheuristictechniqueinafogserverforindustri applications. IEEE Trans Industr Inform. 2018;14(10):4497-4506.
45. Wang X, Wang K, Wu S, et al. Dynamic resource scheduling in mobile edge cloud with cloud radio access network. IEEE Trans Parallel Distribut Syst. 2018;29(11):2429-2445.
46. Sun J, Sun S, Li K, Liao D, Sangaiah AK, Chang V. Efficient algorithm for traffic engineering in cloud-of-things and edge computing. Comput Electr Eng. 2018;69:610-627.
47. Arkian HR, Diyanat A, Pourkhalili A. Mist: fog-based data analytics scheme with cost-efficient resource provisioning for iot crowdsensing applications. J Netw Comput Appl. 2017;82:152-165.
48. Yao H, Bai C, Xiong M, Zeng D, Fu Z. Heterogeneous cloudlet deployment and user-cloudlet association toward cost effective fog computing. Concurr Comput Pract Theory. 2019;31(11):e3975.
49. Mapetu JPB, Chen Z, Kong L. Heuristic cloudlet allocation approach based on optimal completion time and earliest finish time. J Supercomput. 2015;73(4):1672-1690.
50. Roy DG, De D, Mukherjee A, Buyuya R. Application-aware cloudlet selection for computation offloading in multi-cloudlet environment. J Supercomput. 2020;75:645-665.
51. Zeng D, Gu L, Yao H. Towards energy efficient service composition in green energy powered cyber-physical fog systems. Futur Gener Comput Syst. 2020;105:757-765.
52. Roy DG, De D, Mukherjee A, Buyuya R. Application-aware cloudlet selection for computation offloading in multi-cloudlet environment. J Supercomput. 2017;73(4):1672-1690.
53. Mapetu JPB, Chen Z, Kong L. Heuristic cloudlet allocation approach based on optimal completion time and earliest finish time. IEEE Access. 2018;6:61714-61727.
54. Toczé K, Nadjm-Tehrani S. A taxonomy for management and optimization of multiple resources in edge computing. Wirel Commun Mob Comput. 2018;2018:7476201.
55. Tang M, Gao L, Huang J. Enabling edge cooperation in tactile internet via 3c resource sharing. IEEE J Selected Areas Commun. 2018;36(11):2444-2454.
56. Liu Y, Jiang J, Liu Y, Zhang Y, Wu Q. Network resource optimization configuration in edge computing environment. Int J Comput Appl. 2019;45(1):88-95.
57. Azam M, Huh EN. 2015 IEEE 29th International Conference on Advanced Information Networking and Applications. IEEE; 2015:687-694.
58. Rahman GS, Peng M, Zhang K, Chen S. Radio resource allocation for achieving ultra-low latency in fog radio access networks. IEEE Access. 2018;6:17442-17454.
59. Lin YH, Hu YC. Residential consumer-centric demand-side management based on energy disaggregation-piloting constrained swarm intelligence: towards edge computing. Sensors. 2018;18(5):1365.
64. Yang X, Chen Z, Li K, et al. Communication-constrained mobile edge computing systems for wireless virtual reality: scheduling and tradeoff. IEEE Access. 2018;6:16665-16677.

65. Song Z, Duan Y, Wan S, et al. Processing optimization of typed resources with synchronized storage and computation adaptation in fog computing. Wirel Commun Mob Comput. 2018;2018.

66. Zhang F, Ge J, Li Z, et al. A load-aware resource allocation and task scheduling for the emerging cloudlet system. Futur Gener Comput Syst. 2018;87:438-456.

67. Pace P, Aloj G, Gravina R, Caliciuri G, Fortino G, Liotta A. An edge-based architecture to support efficient applications for healthcare industry 4.0. IEEE Trans Industr Inform. 2018;15(1):481-489.

68. Marjanović M, Antonic A, Žarko IP. Edge computing architecture for mobile crowdsensing. IEEE Access. 2018;6:10662-10674.

69. Park HC, Min OG, Lee YJ. Scalable architecture for an automated surveillance system using edge computing. J Supercomput. 2017;73(3):926-939.

70. Wang SC, Tseng SC, Yan KQ, Tsai YT. Reaching agreement in an integrated fog cloud iot. IEEE Access. 2018;6:64515-64524.

71. Wang T, Zhang G, Liu A, Bhuian MZA, Jin Q. A secure IoT service architecture with an efficient balance dynamics based on cloud and edge computing. IEEE Internet Things J. 2018;6(3):4831-4843.

72. Magarawagal CMS, Yang K, Hu L, Zhang J. Energy-efficient and network-aware offloading algorithm for mobile cloud computing. Comput Netw. 2014;74:22-33.

73. Sharma PK, Chen MY, Park JH. A software defined fog node based distributed blockchain cloud architecture for IoT. IEEE Access. 2017;6:115-124.

74. Park Y, Adachi F. Enhanced Radio Access Technologies for Next Generation Mobile Communication. Springer; 2007.

75. Salman T. Cloud RAN: Basics, Advances and Challenges A Survey of CRAN Basics, Virtualization, Resource Allocation, and Challenges; 2016:1-16.

76. Bonomi F, Milito R, Zhu J, Addepalli S. Fog Computing and its Role in the Internet of Things. Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing; 2012:13-16.

77. Shantharama P, Thayagaturu AS, Karakoc N, Ferrari L, Reisslein M, Scaglione A. Layback: SDN management of multi-access edge computing (MEC) for network access services and radio resource sharing. IEEE Access. 2018;6:57545-57561.

78. Mei H, Wang K, Yang K. Multi-layer cloud-ran with cooperative resource allocations for low-latency computing and communication services. IEEE Access. 2017;5:19023-19032.

79. Zhang H, Qiu Y, Chu X, Long K, Leung VC. Fog radio access networks: mobility management, interference mitigation, and resource optimization. IEEE Wirel Commun. 2017;24(6):120-127.

80. Chen X, Zhang H, Wu C, Mao S, Ji Y, Bennis M. Optimized computation offloading performance in virtual edge computing systems via deep reinforcement learning. IEEE Internet Things J. 2018;6(3):4005-4018.

81. Garcia-Saavedra A, Iosifidis G, Costa-Perez X, Leith DJ. Joint optimization of edge computing architectures and radio access networks. IEEE J Selected Areas Commun. 2018;36(11):2433-2443.

82. Zhang Z, Sun H, Lei X. Non-orthogonal multiple access. Encyclopedia of Wireless Networks. Springer; 2018. doi:10.1007/978-3-319-32903-1_57-1

83. Wu Y, Ni K, Zhang C, Qian LP, Tsang DH. Noma-assisted multi-access mobile edge computing: a joint optimization of computation offloading and time allocation. IEEE Trans Veh Technol. 2018;67(12):12244-12258.

84. Liu B, Li W, Ma Y, Wang J, Lu G. Wireless powered cognitive-based mobile edge computing with imperfect spectrum sensing. IEEE Access. 2019;7:80431-80442.

85. Nguyen VG, Brunstrom A, Grinnemo KJ, Taheri J. SDN/NFV-based mobile packet core network architectures: a survey. IEEE Commun Surveys Tutorials. 2017;19(3):1567-1602.

86. Zhao L, Sun W, Shi Y, Liu J. Optimal placement of cloudlets for access delay minimization in sdn-based internet of things networks. IEEE Internet Things J. 2018;5(2):1334-1344.

87. Schiller E, Nikaein N, Kalofigiotou E, Gasparyan M, Braun T. CDS-MEC: NFV/SDN-based application management for MEC in 5g systems. Comput Netw. 2018;135:96-107.

88. Liao L, Qiu M, Leung V. Software defined mobile cloudlet. Mobile Networks Appl. 2015;20(3):337-347.

89. Cominardi L, Giust F, Bernardos CJ, De La Oliva A. Distributed mobility management solutions for next mobile network architectures. Comput Netw. 2017;121:124-136.

90. Liang C, He Y, Yu FR, Zhao N. Enhancing video rate adaptation with mobile edge computing and caching in software-defined mobile networks. IEEE Trans Wirel Commun. 2018;17(10):7013-7026.

91. Mehrabi A, Siekkinen M, Ylä-Jaaski A. Qoe-traffic optimization through collaborative edge caching in adaptive mobile video streaming. IEEE Access. 2018;6:52261-52276.

92. Tran TX, Hajisami A, Pandey P, Pompili D. Collaborative mobile edge computing in 5g networks: new paradigms, scenarios, and challenges. IEEE Commun Mag. 2017;55(4):54-61.

93. Xiang H, Zhou W, Daneshmand M, Peng M. Network slicing in fog radio access networks: issues and challenges. IEEE Commun Mag. 2017;55(12):110-116.

94. Alliance N. Definition of the Testing Framework for the NGMN 5g Pre-Commercial Networks Trials. Frankfurt, Germany; 2019.

95. Zhao P, Tian H, Fan S, Paulraj A. Information prediction and dynamic programming-based ran slicing for mobile edge computing. IEEE Wireless Commun Lett. 2018;7(4):614-617.

96. Xiao Y, Kunz M. Dynamic network slicing for scalable fog computing systems with energy harvesting. IEEE J Selected Areas Commun. 2018;36(12):2640-2654.

97. Han K, Li S, Tang S, et al. Application-driven end-to-end slicing: when wireless network virtualization orchestrates with NFV-based mobile edge computing. IEEE Access. 2018;6:26567-26577.

98. Zhang P, Chen Z, Liu JK, Liang K, Liu H. An efficient access control scheme with outsourcing capability and attribute update for fog computing. Futur Gener Comput Syst. 2018;78:753-762.

99. Cui H, Yi X, Nepal S. Achieving scalable access control over encrypted data for edge computing networks. IEEE Access. 2018;6:30049-30059.

100. Xu Q, Tan C, Fan Z, Zhu W, Xiao Y, Cheng F. Secure data access control for fog computing based on multi-authority attribute-based signcryption with computation outsourcing and attribute revocation. Sensors. 2018;18(5):1609.
101. Li Y, Dong Z, Sha K, Jiang C, Wan J, Wang Y. TMO: time domain outsourcing attribute-based encryption scheme for data acquisition in edge computing. IEEE Access. 2019;7:40240-40257.

102. Li C, Qin Z, Novak E, Li Q. Securing sdn infrastructure of IoT–fog networks from MITM attacks. IEEE Internet Things J. 2017;4(5):1156-1164.

103. Yaseen Q, Albala F, Jararwah Y, Al-Yayoub M. Leveraging fog computing and software defined systems for selective forwarding attacks detection in mobile wireless sensor networks. Trans Emerg Telecommun Technol. 2018;29(4):e3183.

104. Zhang Y, Zhao J, Zheng D, et al. Privacy-preserving data aggregation against false data injection attacks in fog computing. Sensors. 2018;18(8):2659.

105. Sharma V, You I, Jakayody DNK, Atiquzzaman M. Cooperative trust relaying and privacy preservation via edge-crowdsourcing in social internet of things. Futur Gener Comput Syst. 2019;92:758-776.

106. Li J, Li X, Yuan J, Zhang R, Fang B. Fog computing-assisted trustworthy forwarding scheme in mobile internet of things. IEEE Internet Things J. 2018;6(2):2778-2796.

107. Xu Q, Su Z, Wang Y, Dai M. A trustworthy content caching and bandwidth allocation scheme with edge computing for smart campus. IEEE Access. 2018;6:63868-63879.

108. Sharma V, You I, Kumar R, Kim P. Computational offloading for efficient trust management in pervasive online social networks using osmotic computing. IEEE Access. 2017;5:5084-5103.

109. Basudan S, Lin X, Sankaranarayanan K. A privacy-preserving vehicular crowdsensing-based road surface condition monitoring system using fog computing. IEEE Internet Things J. 2017;4(3):772-782.

110. Elmisery AM, Rho S, Botvich D. A fog based middleware for automated compliance with oecd privacy principles in internet of healthcare things. IEEE Access. 2016;4:8418-8441.

111. Yang X, Yin F, Tang X. A fine-grained and privacy-preserving query scheme for fog computing-enhanced location-based service. Sensors. 2017;17(7):1611.

112. Yang R, Xu Q, Au MH, Yu Z, Wang H, Zhou L. Position based cryptography with location privacy: a step for fog computing. Futur Gener Comput Syst. 2018;78:799-806.

113. Mosysiadis V, Sarigiannidis P, Moscholios I. Towards distributed data management in fog computing. WIREs Commun Mob Comput. 2018;2018:7597686.

114. Erol-Kantarci M, Sukhmani S. Caching and computing at the edge for mobile augmented reality and virtual reality (AR/VR) in 5G. Ad Hoc Netw. 2018;223:169-177.

115. Liu D, Chen B, Yang C, Molisch AF. Caching at the wireless edge: design aspects, challenges, and future directions. IEEE Commun Mag. 2016;54(9):22-28.

116. Hou T, Feng G, Qin S, Jiang W. Proactive content caching by exploiting transfer learning for mobile edge computing. Int J Commun Syst. 2018;31(11):e3706.

117. Tang Y, Guo K, Ma J, Shen Y, Chi T. A smart caching mechanism for mobile multimedia in information centric networking with edge computing. Futur Gener Comput Syst. 2019;91:590-600.

118. Zhu K, Chen Z, Zhang L, Zhang Y, Kim SW. Geo-cascading and community-cascading in social networks: comparative analysis and its implications to edge caching. Inf Sci. 2018;436:1-12.

119. Wang X, Leng S, Yang K. Social-aware edge caching in fog radio access networks. IEEE Access. 2017;5:8492-8501.

120. Di Pietro N, Strinati EC. An optimal low-complexity policy for cache-aided computation offloading. IEEE Access. 2019;7:182499-182514.

121. Wu J, Dong M, Ota K, Li J, Guan Z. FCSS: fog-computing-based content-aware filtering for security services in information-centric social networks. IEEE Trans Emerg Top Comput. 2017;7(4):553-564.

122. Huang JY, Hong WC, Tai PS, Liao IE. A model for aggregation and filtering on encrypted XML streams in fog computing. Int J Distribut Sensor Networks. 2013(5):1550147717704.

123. Fan K, Yin J, Zhang K, Li H, Yang Y. EARS-DM: efficient auto correction retrieval scheme for data management in edge computing. Sensors. 2018;18(11):3616.

124. Rahmani AM, Gia TN, Negash B, et al. Exploiting smart e-health gateways at the edge of healthcare internet-of-things: a fog computing approach. Futur Gener Comput Syst. 2018;78:641-658.

125. Dautov R, Distefano S, Bruneo D, Longo F, Merlino G, Puliafito A. Data processing in cyber-physical-social systems through edge computing. IEEE Access. 2018;6:29822-29835.

126. Dautov R, Distefano S, Bruneo D, Longo F, Merlino G, Puliafito A. Data agility through clustered edge computing and stream processing. Concurr Comput Pract Exp. 2021;33(7):1.

127. Evaluating the benefits of combined and continuous fog-to-cloud architectures. Comput Commun. 2017;113:43-52.

128. Luo G, Yuan Q, Zhou H, et al. Cooperative vehicular content distribution in edge computing assisted 5G-vanet. China Commun. 2018;15(7):1-17.

129. Moreno BM, Salvador CP, Domingo ME, Pena IA, Extremera VR. On content delivery network implementation. Comput Commun. 2006;29(12):2396-2412.

130. Molina B, Palau CE, Esteve M. Modeling content delivery networks and their performance. Comput Commun. 2004;27(15):1401-1411.

131. Kim DY, Kim S. Data transmission using k-means clustering in low power wide area networks with mobile edge cloud. Wirel Pers Commun. 2019;105(2):567-581.

132. Gao L, Luan TH, Yu S, Zhou W, Liu B. Fogroute: DTN-based data dissemination model in fog computing. IEEE Internet Things J. 2016;4(1):225-235.

133. Wu X. Combination replicas placements strategy for data sets from cost-effective view in the cloud. Int J Comput Intell Syst. 2017;10(1):521.

134. Bhalaji N. Efficient and secure data utilization in mobile edge computing by data replication. JSMAC. 2020;2(01):1-12.

135. Chang WC, Wang PC. Adaptive replication for mobile edge computing. IEEE J Selected Areas Commun. 2018;36(11):2422-2432.

136. Xinong Z, Feng S, Wang W, Niyato D, Wang P, Han Z. Cloud/fog computing resource management and pricing for blockchain networks. IEEE Internet Things J. 2018;6(3):4585-4600.

137. Liu M, Liu Y. Price-based distributed offloading for mobile-edge computing with computation capacity constraints. IEEE Wireless Commun Lett. 2017;6(3):420-423.

138. Zhang T. Data offloading in mobile edge computing: a coalition and pricing based approach. IEEE Access. 2017;6:2760-2767.

139. Yu R, Ding J, Maharjan S, Gjessing S, Zhang Y, Tsang DH. Decentralized and optimal resource cooperation in geo-distributed mobile cloud computing. IEEE Trans Emerg Top Comput. 2015;6(1):72-84.
140. Xie K, Wang X, Xie G, et al. Distributed multi-dimensional pricing for efficient application offloading in mobile cloud computing. IEEE Trans Serv Comput. 2016;12(6):925-940.

141. Nguyen DT, Le LB, Bhargava V. Price-based resource allocation for edge computing: a market equilibrium approach. IEEE Trans Cloud Comput. 2018;9(1):302-317.

142. Wang X, Sui Y, Wang J, Yuen C, Wu W. A distributed truthful auction mechanism for task allocation in mobile cloud computing. IEEE Trans Serv Comput. 2018;14(3):628-638.

143. Sun W, Liu J, Yue Y, Zhang H. Double auction-based resource allocation for mobile edge computing in industrial internet of things. IEEE Trans Industr Inform. 2018;14(10):4692-4701.

144. Zhou G, Wu J, Chen L, Jiang G, Lam SK. Efficient three-stage auction schemes for cloudlets deployment in wireless access network. Wirel Netw. 2019;25(6):3335-3349.

145. Zhang H, Guo F, Ji H, Zhu C. Combinational auction-based service provider selection in mobile edge computing networks. IEEE Access. 2017;5:13455-13464.

146. Kiani A, Ansari N. Toward hierarchical mobile edge computing: an auction-based profit maximization approach. IEEE Internet Things J. 2017;4(6):2082-2091.

147. Kato S, Shinkuma R. Priority control in communication networks for accuracy-freshness tradeoff in real-time road-traffic information delivery. IEEE Access. 2017;5:25226-25235.

148. Tang C, Wei X, Zhu C, Chen W, Rodrigues JJ. Towards smart parking based on fog computing. IEEE Access. 2018;6:70172-70185.

149. Raza S, Wang S, Ahmed M, Anwar MR. A survey on vehicular edge computing: architecture, applications, technical issues, and future directions. Wirel Commun Mob Comput. 2019:2019:3159762.

150. Cicirelli F, Guerrieri A, Spezzano G, Vinci A. An edge-based platform for dynamic smart city applications. Futur Gener Comput Syst. 2017;76:106-118.

151. Jiang H, Li J, Wen B, Xin Y, Liu J. Connecting intelligent things in smart hospitals using nb-iot. IEEE Internet Things J. 2018;5(3):1550-1560.

152. Darwish T, Bakar A, Big KBFI. Data analytics in the internet of vehicles environment: motivations, architecture, challenges, and critical issues. IEEE Access. 2018;6:15679-15701.

153. Ahmad M, Amin MB, Hussain S, Kang BH, Cheong T, Lee S. Health fog: a novel framework for health and wellness applications. J Supercomput. 2016;72(10):3677-3695.

154. Guo B, Zhang X, Wang Y, Yang H. Deep-q-network-based multimedia multi-service qos optimization for mobile edge computing systems. IEEE Access. 2019;7:160961-160972.

155. Liu R, Yuan X, Xu J, et al. A novel server selection approach for mobile cloud streaming service. Simul Model Pract Theory. 2015;50:72-82.

156. Jiang W, Feng G, Qin S, Liu Y. Multi-agent reinforcement learning based cooperative content caching for mobile edge networks. IEEE Access. 2019;7:61856-61867.

157. Vigneri L, Spyropoulos T, Barakat C. Low cost video streaming through mobile edge caching: modelling and optimization. IEEE Trans Mob Comput. 2018;18(6):1302-1315.

158. Verbeke N, Chao KM, Lewandowski J, Shah N, James A, Tian F. Modeling industry 4.0 based fog computing environments for application analysis and deployment. Futur Gener Comput Syst. 2019;91:48-60.

159. Chen L, Zhou P, Gao L, Xu J. Adaptive fog configuration for the industrial internet of things. IEEE Trans Industr Inform. 2018;14(10):4656-4664.

160. Li L, Ota K, Dong M. Human in the loop: distributed deep model for mobile crowdsensing. IEEE Internet Things J. 2018;5(6):4957-4964.

161. Ma L, Liu X, Pei Q, Xiang Y. Privacy-preserving reputation management for edge computing enhanced mobile crowdsensing. IEEE Trans Serv Comput. 2018;12(5):786-799.

162. Zhao K, Tang S, Zhao B, Wu Y. Dynamic and privacy-preserving reputation management for blockchain-based mobile crowdsensing. IEEE Access. 2019;7:74694-74710.

163. Caudell TP. Telemanipulator and Telepresence Technologies. Vol 2351. International Society for Optics and Photonics; 1995:272-281.

164. Park J, Popovski P, Simeone O. Minimizing latency to support VR social interactions over wireless cellular systems. IEEE Wireless Commun Lett. 2018;7(5):776-779.

165. Zhang Y, Jiao L, Yan J, Lin X. Dynamic service placement for virtual reality group gaming on mobile edge cloudlets. IEEE Access. 2018;7:1881-1897.

166. Chiu TC, Pang AC, Chung WH, Zhang J. Latency-driven fog cooperation approach in fog radio access networks. IEEE Trans Serv Comput. 2018;12(5):698-711.

167. Verbelen T, Simoens P, De Turck F, Dhoedt B. Adaptive deployment and configuration for mobile augmented reality in the cloudlet. J Netw Comput Appl. 2014;41:206-216.

168. Hu YC, Patel M, Sabella D, Sprecher N, Young V. Mobile edge computing—a key technology towards 5g. ETSI White Paper. 2015;11(11):1-16.

169. Preden JS, Tammemäe K, Jantsch A, et al. Toward hierarchical mobile edge computing: an auction-based profit maximization approach. IEEE Internet Things J. 2018;5(10):4656-4664.

170. Li L, Ota K, Dong M. Human in the loop: distributed deep model for mobile crowdsensing. IEEE Internet Things J. 2018;5(6):4957-4964.

171. Ma L, Liu X, Pei Q, Xiang Y. Privacy-preserving reputation management for edge computing enhanced mobile crowdsensing. IEEE Trans Serv Comput. 2018;12(5):786-799.

172. Zhao K, Tang S, Zhao B, Wu Y. Dynamic and privacy-preserving reputation management for blockchain-based mobile crowdsensing. IEEE Access. 2019;7:74694-74710.

173. Caudell TP. Telemanipulator and Telepresence Technologies. Vol 2351. International Society for Optics and Photonics; 1995:272-281.

174. Park J, Popovski P, Simeone O. Minimizing latency to support VR social interactions over wireless cellular systems via bandwidth allocation. IEEE Wireless Commun Lett. 2018;7(5):776-779.

175. Zhang Y, Jiao L, Yan J, Lin X. Dynamic service placement for virtual reality group gaming on mobile edge cloudlets. IEEE Access. 2018;7:1881-1897.

176. Chiu TC, Pang AC, Chung WH, Zhang J. Latency-driven fog cooperation approach in fog radio access networks. IEEE Trans Serv Comput. 2018;12(5):698-711.

177. Verbelen T, Simoens P, De Turck F, Dhoedt B. Adaptive deployment and configuration for mobile augmented reality in the cloudlet. J Netw Comput Appl. 2014;41:206-216.

178. Hu YC, Patel M, Sabella D, Sprecher N, Young V. Mobile edge computing—a key technology towards 5g. ETSI White Paper. 2015;11(11):1-16.

179. Preden JS, Tammemäe K, Jantsch A, Leier M, Riid A, Calis E. The benefits of self-awareness and attention in wireless cellular systems via bandwidth allocation. IEEE Wireless Commun Lett. 2018;7(5):776-779.

180. Zhang Y, Jiao L, Yan J, Lin X. Dynamic service placement for virtual reality group gaming on mobile edge cloudlets. IEEE Access. 2018;7:1881-1897.

181. Chiu TC, Pang AC, Chung WH, Zhang J. Latency-driven fog cooperation approach in fog radio access networks. IEEE Trans Serv Comput. 2018;12(5):698-711.
179. Abdelwahab S, Hamdaoui B, Guizani M, Znati T. Replisom: disciplined tiny memory replication for massive IoT devices in LTE edge cloud. IEEE Internet Things J. 2015;3(3):327-338.

180. Fan W, Liu Y, Tang B, Wu F, Wang Z. Computation offloading based on cooperations of mobile edge computing-enabled base stations. IEEE Access. 2017;6:22622-22633.

181. Gill SS, Garraghan P, Buyya R. Router: fog enabled cloud based intelligent resource management approach for smart home IoT devices. J Syst Softw. 2019;154:125-138.

182. Hao P, Hu L, Jiang J, Hu J, Che X. Mobile edge provision with flexible deployment. IEEE Trans Serv Comput. 2018;12(5):750-761.

183. Liu M, Mao Y, Leng S, Mao S. Full-duplex aided user virtualization for mobile edge computing in 5G networks. IEEE Access. 2017;6:2996-3007.

184. Matousek J, Gärtner B. Understanding and Using Linear Programming. Springer Science & Business Media; 2006.

185. Awr tel M. Nonlinear Programming: Analysis and Methods. Courier Corporation; 2003.

186. Li G, Yan J, Chen L, Wu J, Lin Q, Zhang Y. Energy consumption optimization with a delay threshold in cloud-fog cooperation computing. IEEE Access. 2019;7:159688-159697.

187. Karatas F, Korpeoglu I. Fog-based data distribution service (F-DAD) for internet of things (IoT) applications. Futur Gener Comput Syst. 2019;93:156-169.

188. Ren D, Gui X, Zhang K, Wu J. Hybrid collaborative caching in mobile edge networks: an analytical approach. Comput Netw. 2019;158:1-16.

189. Garcia F, Rachelson E. Markov decision processes. Wiley-IEEE Press; 2010:1-38.

190. Mazouz H, Boussita K, Achir N. Maximizing mobiles energy saving through tasks optimal offloading placement in two-tier cloud: a theoretical and an experimental study. Comput Commun. 2019;144:132-148.

191. Guo X, Liu L, Zhang Z, Ristami T. Data offloading and task allocation for cloudlet-assisted ad hoc mobile clouds. Wirel Netw. 2018;24(1):79-88.

192. Meng S, Wang Y, Mao Z, Sun K. Joint optimization of wireless bandwidth and computing resource in cloudlet-based mobile cloud computing environment. Peer Peer Network Appl. 2018;11(3):462-472.

193. Yu S, Langar R, Fu X, Wang L, Han Z. Computation offloading with data caching enhancement for mobile edge computing. IEEE Trans Veh Technol. 2018;67(11):11098-11112.

194. Gu Y, Chang Z, Pan M, Song L, Han Z. Joint radio and computational resource allocation in IoT fog computing. IEEE Trans Veh Technol. 2018;67(8):7475-7484.

195. Chen M, Hao Y. Task offloading for mobile edge computing in software defined ultra-dense network. IEEE J Selected Areas Commun. 2018;36(3):587-597.

196. Wang D, Liu Z, Wang X, Lan Y. Mobility-aware task offloading and migration schemes in fog computing networks. IEEE Access. 2019;7:43356-43368.

197. Farshin A, Sharifian S. A modified knowledge-based ant colony algorithm for virtual machine placement and simultaneous routing of NFV in distributed cloud architecture. J Supercomput. 2019;75(8):5520-5550.

198. Rashidi S, Sharifian S. A hybrid heuristic queue based algorithm for task assignment in mobile cloud. Futur Gener Comput Syst. 2017;68:331-345.

199. Zhang C, Cho HH, Chen CY. Emergency-level-based healthcare information offloading over fog network. Peer Peer Network Appl. 2020;13(1):16-26.

200. Rhee I, Ilkhanii E, Nikooghadam M. Secara: a security and cost-aware resource allocation method for mobile cloudlet systems. Ad Hoc Netw. 2019;86:103-118.

201. Wan J, Chen B, Wang S, Xia M, Li D, Liu C. Fog computing for energy-aware load balancing and scheduling in smart factory. IEEE Trans Industr Inform. 2018;14(10):4548-4556.

202. Murphy KP. A survey of pomdp solution techniques. Environment. 2000;2:2X3.

203. Garcia F, Rachelson E. Markov decision processes. Markov Decision Processes in Artificial Intelligence. Wiley-IEEE Press; 2010:1-38.

204. Van Le D, Tham CK. Quality of service aware computation offloading in an ad-hoc mobile cloud. IEEE Trans Veh Technol. 2018;67(9):8890-8904.

205. Li M, Si P, Zhang Y. Delay-tolerant data traffic to software-defined vehicular networks with mobile edge computing in smart city. IEEE Trans Veh Technol. 2018;67(10):9073-9086.

206. Li M, Yu FR, Si P, Zhang Y. Energy-efficient machine-to-machine (M2M) communications in virtualized cellular networks with mobile edge computing (MEC). IEEE Trans Mob Comput. 2018;18(7):1541-1555.

207. Zhang G, Chen Y, Shen Z, Wang L. Distributed energy management for multiuser mobile-edge computing systems with energy harvesting devices and qos constraints. IEEE Internet Things J. 2018;6(3):4035-4048.

208. Nan Y, Li W, Bao W, et al. Adaptive energy-aware computation offloading for cloud of things systems. IEEE Access. 2017;5:23947-23957.

209. Yang Y, Zhao S, Zhang W, Chen Y, Luo X, Wang J. Debts: delay energy balanced task scheduling in homogeneous fog networks. IEEE Internet Things J. 2018;5(3):2094-2106.

210. Pu L, Chen X, Xu J, Fu X. D2D fogging: an energy-efficient and incentive-aware task offloading framework via network-assisted D2D collaboration. IEEE J Selected Areas Commun. 2016;34(12):3887-3901.

211. Rahman MS. Basic Graph Theory. Vol 9. Springer; 2017.

212. Shi T, Yang M, Li X, Lei Q, Jiang Y. An energy-efficient scheduling scheme for time-constrained tasks in local mobile clouds. Pervasive Mobile Comput. 2016;27:90-105.

213. Jin J, Li Y, Luo J. Cooperative storage by exploiting graph-based data placement algorithm for edge computing environment. Concurr Comput Practice Exp. 2018;30(20):e4914.

214. Peng X, Ren J, She L, Zhang D, Li J, Zhang Y. Boat: a block-streaming app execution scheme for lightweight IoT devices. IEEE Internet Things J. 2018;5(3):1816-1829.

215. Wang C, Yu FR, Liang C, Chen Q, Tang L. Joint computation offloading and interference management in wireless cellular networks with mobile edge computing. IEEE Trans Veh Technol. 2017;66(8):7432-7445.

216. Ren C, Luo X, Ni W, Tian H, Liu RP. Distributed online learning of fog computing under nonuniform device cardinality. IEEE Internet Things J. 2018;5(1):1147-1159.

217. Wang T, Wei X, Liang T, Fan J. Dynamic tasks scheduling based on weighted bi-graph in mobile cloud computing. Sustainable Comput Informat Syst. 2018;19:214-222.

218. Ningning S, Chao G, Xingshuo A, Qiang Z. Fog computing dynamic load balancing mechanism based on graph repartitioning. China Commun. 2016;13(3):156-164.

219. Mahesh B. Machine learning algorithms-a review. Int J Sci Res. 2020;9:381-386.

220. Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. Science. 2015;349(6245):255-260.
220. Zhang Q, Lin M, Yang LT, Chen Z, Khan SU, Li P. A double deep q-learning model for energy-efficient edge scheduling. IEEE Trans Serv Comput. 2018;12(5):739-749.
221. Tang L, Chen X, He S. When social network meets mobile cloud: a social group utility approach for optimizing computation offloading in cloudlet. IEEE Access. 2016;4:5868-5879.
222. Pech M, Vrchota J, Bednář J. Predictive maintenance and intelligent sensors in smart factory. Sensors. 2021;21(4):1470.
223. Zhang X, Lin J, Chen Z, Sun F, Zhu X, Fang G. An efficient neural-network-based microseismic monitoring platform for hydraulic fracture on an edge computing architecture. Sensors. 2018;18(6):1828.
224. Corno F, De Russis L, Marcelli A, Montanaro T. An unsupervised and noninvasive model for predicting network resource demands. IEEE Internet Things J. 2018;5(6):4342-4350.
225. Qayyum T, Malik AW, Khalid O, Khan SU. Fognetsim++: a toolkit for modeling and simulation of distributed fog environment. IEEE Access. 2018;6:63570-63583.
226. Gupta H, Vahid Dastjerdi A, Ghosh SK, Buysa R. Ifogsim: a toolkit for modeling and simulation of resource management techniques in the internet of things, edge and fog computing environments. Software Practice Exp. 2017;47(9):1275-1296.
227. Sonmez C, Ozgude C, Ersoy C. Edgecloudsim: an environment for performance evaluation of edge computing systems. Trans Emerg Telecommun Technol. 2018;29(11):e3493.
228. Zeng X, Garg SK, Stradins P, Jayaraman PP, Georgakopoulos D, Ranjan R. Iotsim: a simulator for analysing IoT applications. J Syst Archit. 2017;72:93-107.
229. Siegel J, Perdue J. 2012 Annual SRII Global Conference. IEEE; 2012:411-415.
230. ISO 9126. ISO 9126 stand. ref.ISO/IEC. 1991.
231. Mouradian C, Naboulsi D, Yangui S, Glitho RH, Morrow MJ, Polakos PA. A comprehensive survey on fog computing: state-of-the-art and research challenges. IEEE Commun Surveys Tutorials. 2017;20(1):416-464.
232. Zhao Y, Wang W, Li Y, Meixner CC, Tornatore M, Zhang J. Edge computing and networking: a survey on infrastructures and applications. IEEE Access. 2019;7:101213-101230.
233. Gedeon J, Brandherr F, Egert R, Grube T, Mühlehäuser M. What the fog? Edge computing revisited: promises, applications and future challenges. IEEE Access. 2019;7:152847-152878.
234. Khan WZ, Ahmed E, Hakak S, Yaqoob I, Ahmed A. Edge computing: a survey. Futur Gener Comput Syst. 2019;97:219-235.
235. Zhao Z, Guardalben L, Karimzadeh M, Silva J, Braun T, Sargento S. Mobility prediction-assisted over-the-top edge prefetching for hierarchical vanets. IEEE J Selected Areas Commun. 2018;36(8):1786-1801.

SUPPORTING INFORMATION
Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Sakhdari J, Zolfaghari B, Izadpanah S, et al. Edge computing: A systematic mapping study. Concurrency Computat Pract Exper. 2023;35(22):e7741. doi: 10.1002/cpe.7741