Towards Crowdsourcing Internet of Things (Crowd-IoT): Architectures, Security and Applications

Kenneth Li Minn Ang 1,*, Jasmine Kah Phooi Seng 2,3 and Ericmoore Ngharamike 4

1 School of Science and Engineering, University of the Sunshine Coast, Petrie, QLD 4502, Australia
2 School of AI and Advanced Computing, Xian Jiaotong Liverpool University, Suzhou 215123, China; Jasmine.Seng@xjlu.edu.cn
3 School of Computer Science, Queensland University of Technology, Brisbane, QLD 4000, Australia
4 Department of Computer Science, Federal University Oye-Ekiti, Oye-Ekiti 371104, Nigeria; ericmoore.ngharamike@fuoye.edu.ng
* Correspondence: lang@usc.edu.au

Abstract: Crowdsourcing can play an important role in the Internet of Things (IoT) applications for information sensing and gathering where the participants are equipped with geolocated devices. Mobile crowdsourcing can be seen as a new paradigm contributing to the development of the IoT. They can be merged to form a new and essential platform in crowdsourcing IoT paradigm for data collection from different sources and communication mediums. This paper presents a comprehensive survey for this new Crowdsourcing IoT paradigm from four different perspectives: (1) Architectures for Crowd-IoT; (2) Trustworthy, Privacy and Security for Crowd-IoT; (3) Resources, Sharing, Storage and Energy Considerations for Crowd-IoT; and (4) Applications for Crowd-IoT. This survey paper aims to increase awareness and encourage continuing developments and innovations from the research community and industry towards the Crowdsourcing IoT paradigm.

Keywords: crowdsourcing; Internet of Thing (IoT); crowdsourcing IoT; mobile sensing; social sensing; collaborative sensing

1. Introduction

Howe [1] coined the concept of crowdsourcing in 2006 and described it as a new web-based business strategy that leverages the creative solutions of a distributed network of individuals through an open call for proposals. Since then, researchers in many fields have adopted the concept in developing techniques for many research problems. Over the past decade, crowdsourcing approaches have been developed to facilitate information processing and communication systems. The critical concept of crowdsourcing is the action of outsourcing and distributing tasks performed by workers to a large group of people termed as the crowd. In this approach, a group of socially connected workers with diverse expertise and experience are tasked to work together and collaboratively to solve problems which can be collecting, submitting, or sharing their data and opinion via the Internet. The task allocation and outcome collection are performed using an architecture involving the crowdsourcing platform and users.

With modern society coming into an unprecedented information age and the rapid development of smartphones and mobile technologies, crowdsourcing offers vast computing resources. This leads to a new paradigm called mobile crowdsourcing. Mobile crowdsourcing provides a sensing and data collection paradigm by utilizing intelligent devices to acquire sensing data. Workers can use both the sensing and computing advantage of intelligent devices.

There are some review papers on which have focused on crowdsourcing or mobile crowdsourcing in general [2,3] or for specific applications or industries (e.g., the authors...
in [4] give a survey on crowdsourcing for agriculture applications and the authors in [5] gives a survey on crowdsourcing for disaster applications. However, there is currently no comprehensive survey or review paper on crowdsourcing IoT and edge crowdsourcing to our best knowledge. This paper aims to fill the gap in this survey area and give valuable insights into the research area. This paper first discusses the background knowledge on the traditional and mobile crowdsourcing paradigm and highlights representative crowdsourcing platforms. Next, the paper focuses on IoT and reviews the existing crowdsourcing architecture literature related to IoT, including crowdsourcing for the Social IoT. The centralized architectures and layered structures for crowdsourcing IoT are discussed. A new crowdsourcing Internet of Things (Crowd-IoT) architecture is proposed. The proposed architecture has seven layers. These layers are Identification Layer, Crowd Object Layer, Communication/Network Layer, Middleware Layer, Crowd Computation Layer, Application/User Layer, Security, Privacy, and Trust Layer. The Crowd-IoT preserves the advantages from the advancement of IoT and mobile technologies while offering specific processing and computation of the crowd data.

In centralized architectures, the server interacts with participants after receiving the tasks from requesters. For a large number of connected users and multiple interactions, the large volume of sensing data creates significant computational overheads for the conventional crowdsourcing platform. Furthermore, conventional crowdsourcing techniques cannot meet the requirements for time-sensitive IoT tasks due to unforeseen delays and variations in response time. An approach termed edge computing attempts to use the current and surrounding infrastructures, components, or devices for information processing and return it to the cloud. Edge and Fog computing approaches provide an alternative to improving conventional crowdsourcing platforms. Edge and Fog are emerging technologies for IoT applications that can meet the requirements for computational, storage, and distributed access amongst the users and the Internet. On the one hand, many edge or fog nodes, e.g., base stations, routers, and so on, are flexible to be deployed close to users in IoT. On the other hand, the sensed or collected data can be utilized by adversaries to perform hostile attacks and inferences.

Despite the importance of crowdsourcing, the privacy issue is an issue for large-scale deployment of crowdsourcing applications. One of the reasons is that task requestors or owners and users may fully trust the crowdsourcing server. For the participants, the tasks and sensing data may contain private information such as point of interest, location, and so on. On the other hand, the task requestors may inadvertently reveal private information such as their daily routine when they release the tasks. Thus, it is essential to keep the sensing tasks unknown to the server. Furthermore, task requestors may consider the outcomes or results as private assets and may not want to disclose them. In the crowdsourcing IoT, it is critical to preserve the privacy of both the requestors and the participants.

Besides privacy, trust is also another critical factor in establishing confidence and privacy between the entities. The current and existing crowdsourcing approaches which resolve privacy and trust issues do not focus on privacy and trust during data exchanges in IoT or crowdsourcing IoT. This paper also reviews trust and privacy by integrating crowdsourcing and IoT at a later stage. Furthermore, the existing solutions utilize a trust scoring system and a centralized server. These scoring systems can be exploited and manipulated by eavesdroppers. Therefore, the specific research works of trust and privacy for edge-based crowdsourcing are also reviewed. Privacy is also dependent on the security of data and infrastructure. Security threats usually come from the external environment. The external attacks may involve eavesdropping on communication channels to intercept the encrypted sensing data. These attacks may be disguised as legitimate and authorized participants to submit false data to the fog nodes in the fog architecture.

Hence the latter part of the paper also reviews some security aspects in crowdsourcing IoT and edge-based crowdsourcing. Finally, the paper presents the research works on diverse applications of crowdsourcing IoT. The applications include transport and smart
cities, health, localization, monitoring and guidance, 3D, video and live streaming, semantics and data mining, disaster/emergency and risk management, ocean/marine and environment, recruitment and workers, finance, budget and planning, and so on. To the best of the authors’ knowledge, there are no comprehensive reviews on specific research works from different aspects for crowdsourcing IoT and edge-based crowdsourcing. The remaining paper is organized as follows. Section 2 presents the research methodology employed in this paper. Section 3 presents architectures for crowdsourcing and a proposed crowdsourcing IoT architecture. This is followed by Sections 4 and 5, which discuss trustworthiness, privacy, and security in Crowd-IoT and resources, sharing, storage, and energy considerations in Crowd-IoT. Various applications for the Crowd-IoT are discussed in Section 6. Section 7 concludes the paper.

2. Research Method

One of the major objectives and contributions of this paper is to present a comprehensive survey for the new Crowdsourcing IoT paradigm from four different perspectives: (1) Architectures for Crowd-IoT; (2) Trustworthy, Privacy and Security for Crowd-IoT, (3) Resources, Sharing, Storage and Energy Considerations for Crowd-IoT, and (4) Applications for Crowd-IoT. The literature review approach has been recognized as a valid approach and a crucial step in structuring a research field, and so it is an essential component of a research project. As a result, this study employs a four-step research process for gathering and analyzing the literature which includes—(1) Defining unit of analysis, (2) Selecting the classification context, (3) Collecting publications and delineating the field, (4) Analyzing or evaluating the materials. Thus, within the constraints of this purpose, this study delivers the literature encompassing over 625 journal research articles and over 130 cited references at the end.

The literature has been structured under four main categories, namely: (1) Architectures for Crowd-IoT, (2) Trustworthy, Privacy and Security for Crowd-IoT, (3) Resources, Sharing, Storage, and Energy Considerations for Crowd-IoT, and (4) Applications for Crowd-IoT. Google scholar, IEEE Explore, and Scopus databases were used to search for the relevant papers from 2006 onwards, while others have been obtained via cross-referencing. However, as illustrated in Figure 1, a wide range of papers were found between 2017 and 2020, which aided the authors in refining their search to uncover any missing papers, both traditional and new, linked to the keywords. In finding relevant publications, the authors have also considered a single research paper as the unit of analysis, and have employed a set of keywords to ensure the collection of a large number of studies. Table 1 shows an overview of the paper structure and the main classification descriptors.
Figure 1. Statistics of distribution of the papers by the publication year.

Table 1. Overview of the paper structure and main classification descriptors.

<table>
<thead>
<tr>
<th>Classification Descriptors</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Architectures for Crowd-IoT</td>
<td>[6–23]</td>
</tr>
<tr>
<td>Centralized architectures for Crowd-IoT</td>
<td></td>
</tr>
<tr>
<td>Decentralized architectures for Crowd-IoT</td>
<td>[24–38]</td>
</tr>
<tr>
<td>Trustworthy, Privacy and Security for Crowd-IoT</td>
<td></td>
</tr>
<tr>
<td>Privacy in Crowd-IoT</td>
<td>[39–45]</td>
</tr>
<tr>
<td>Trustworthy in Crowd-IoT</td>
<td>[46–51]</td>
</tr>
<tr>
<td>Security in Crowd-IoT</td>
<td>[52–58]</td>
</tr>
<tr>
<td>Resources, Sharing, Storage, and Energy Considerations for Crowd-IoT</td>
<td></td>
</tr>
<tr>
<td>Considerations for Crowd-IoT</td>
<td>[59–63]</td>
</tr>
<tr>
<td>Applications for Crowd-IoT</td>
<td></td>
</tr>
<tr>
<td>Localization</td>
<td>[64–70]</td>
</tr>
<tr>
<td>Monitoring and guidance</td>
<td>[71–77]</td>
</tr>
<tr>
<td>Video and live streaming</td>
<td>[78–85]</td>
</tr>
<tr>
<td>Semantics and data mining</td>
<td>[86–89]</td>
</tr>
<tr>
<td>Disaster and risk management</td>
<td>[90–99]</td>
</tr>
<tr>
<td>Transportation and smart cities</td>
<td>[100–106]</td>
</tr>
<tr>
<td>Ocean and environmental</td>
<td>[107–111]</td>
</tr>
<tr>
<td>Recruitment and workers</td>
<td>[112–117]</td>
</tr>
<tr>
<td>Finance and pricing</td>
<td>[118–123]</td>
</tr>
<tr>
<td>Health and medical</td>
<td>[124–128]</td>
</tr>
</tbody>
</table>
3. Architectures of Crowdsourcing and Proposed Crowdsourcing Internet of Things

In this section, types and research of crowdsourcing architectures will be discussed. The section first discusses the general structure and primary components of crowdsourcing. Typically, crowdsourcing systems implemented using the Internet and client-server architectures can utilize common devices such as smartphones and general-purpose computers. An intensive review of the research works will be provided. The section also presents the centralized and decentralized architectures of crowdsourcing. A centralized approach for crowdsourcing typically utilizes a client-server architecture or model. The computational requirements of these services are performed at the central point or server where the computing resources are located at the main or centralized data center. On the other hand, the decentralized architecture for crowdsourcing is a distributed approach. The layered structure enables peer-to-peer processing and in-situ processing to avoid all components connecting to a central and reduce the traffic congestion to/at the center. After reviewing the centralized and decentralized architectures for crowdsourcing, it is noted that many architectures have been proposed for crowdsourcing, but they are not specifically designed for crowdsourcing IoT. Thus, this section will propose a new architecture for crowdsourcing IoT called Crowd-IoT.

3.1. Centralized Architectures of Crowdsourcing

Most studies on general architectures for crowdsourcing focused on the functions such as data storage, capture, data processing, task management, security, scalability, and privacy. Table 2 shows a summary of the research works on centralized architectures of crowdsourcing. The non-functional aspects (e.g., trust, security, and privacy) will be discussed in the next section (Section 3). Several architectures of applications focused on task management and distributing tasks to participants. Most of these works are based on the general or centralized architecture of crowdsourcing. There are some studies or reviews on crowdsourcing, e.g., its applications, techniques, and technologies. To the best of our knowledge, there is no comprehensive review of the architectures of crowdsourcing IoT or IoT architectures for crowdsourcing. In the review paper by Phuttharak [129], the authors provide a survey of the literature on mobile crowdsourcing research. From the current mobile architectures and applications, the authors classified the architectures of crowdsourcing based on typical roles, components, and functionalities. The review provides a good understanding of typical functionalities, design aspects, and concerns during the development of mobile crowdsourcing systems.

Table 2. Summary of the research works on centralized architectures of crowdsourcing.

<table>
<thead>
<tr>
<th>Authors &amp; Ref</th>
<th>Year</th>
<th>Focus Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cornelius et al. [6]</td>
<td>2008</td>
<td>The authors presented AnonySense, a general architecture focusing on privacy-aware mobile crowdsourcing. Sensing tasks are distributed across anonymous participating mobile devices, verified, anonymized, in this participatory sensing mode.</td>
</tr>
<tr>
<td>Das et al. [7]</td>
<td>2010</td>
<td>The authors presented a Platform for Remote Sensing using Smartphones (PRISM) which is general architecture for crowdsourcing applications. It allows developers to package their applications as executable binaries, which offers efficiency and also the flexibility of reusing existing code modules.</td>
</tr>
<tr>
<td>Ra et al. [9]</td>
<td>2012</td>
<td>The authors designed and implemented Medusa, which is a novel programming framework to provide high-level abstractions for specifying the steps for a crowd-sensing task. Medusa</td>
</tr>
<tr>
<td>Authors</td>
<td>Year</td>
<td>Description</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Hupfer et al. [10]</td>
<td>2012</td>
<td>The authors presented MoCoMapps (Mobile Collaborative Map-based Applications) to provide a mobile service for users to use and create mobile map-based applications (e.g., data entry, data look-up, and collaboration).</td>
</tr>
<tr>
<td>Hu et al. [11]</td>
<td>2013</td>
<td>A mobile cyber-physical system called Vita was presented to allow users to perform mobile crowdsensing tasks. Vita is a universal architecture that integrates the service-oriented architecture with a resource optimization mechanism for crowdsensing.</td>
</tr>
<tr>
<td>Khorashadi et al. [12]</td>
<td>2013</td>
<td>The work is about general architectural components, including the collection, aggregation, and analysis of data.</td>
</tr>
<tr>
<td>Hetmank et al. [13]</td>
<td>2013</td>
<td>Crowdsourcing architecture/system with a focus on campaign management. Contains four components: (1) Task management; (2) User management; (3) Contribution management; and (4) Workflow management.</td>
</tr>
<tr>
<td>Difallah et al. [14]</td>
<td>2013</td>
<td>The authors focused on crowd tasks and proposed an underlying categorization structure, which extracted from the task descriptions and deployed the worker profiles from social network information. Then the most suitable workers and tasks are automatically matched.</td>
</tr>
<tr>
<td>Fuchs-Kittowski &amp; Faust [15]</td>
<td>2014</td>
<td>Architecture for mobile crowdsourcing applications. The architecture is basic and only has client/participant and server. On the server-side, recruitment and interaction with participants are performed. On the client-side, functions, e.g., user interface and data capture are provided. The data include geospatial data captured and shared using mobiles or smartphones.</td>
</tr>
<tr>
<td>Zhao and Zhu [16]</td>
<td>2014</td>
<td>The authors focused on the functional components for the transformation process. These components are: (i) Assigners—initiate and manage the task, (ii) Providers—respond to task and submit solutions; and (iii) Intermediation platform—connects assigners and providers, functions as crowdsourcing enabler.</td>
</tr>
<tr>
<td>Ponciano et al. [17]</td>
<td>2014</td>
<td>Crowdsourcing architecture/platform with a focus on distributed applications. It contains three contributions: (1) QoS requirements for requesters; (2) Strategies for platform management and execution; and (3) Human aspects and worker characteristics.</td>
</tr>
<tr>
<td>Luz et al. [18]</td>
<td>2014</td>
<td>A task-oriented crowdsourcing system was proposed to connect the worker, the requester. Authors focused on task management in crowdsourcing and complex tasks. Crowd tasks are classified into subtasks (e.g., partition, aggregation, qualification, and grading).</td>
</tr>
<tr>
<td>Ren et al. [19]</td>
<td>2015</td>
<td>The authors described two models for mobile crowdsourcing: (1) Internet-based and; (2) Local-based. In the Internet-based model, mobile users can be utilized as service providers. In the local-based model, users within the vicinity can provide cloud services in local-based crowdsourcing.</td>
</tr>
<tr>
<td>Giang et al. [20]</td>
<td>2015</td>
<td>Authors used crowdsourcing to exploit the sensors and other equipment for setting up IoT networks and focused on a more</td>
</tr>
<tr>
<td>Authors</td>
<td>Year</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------</td>
<td>-------------</td>
</tr>
<tr>
<td>Zhang and Marchiori</td>
<td>2017</td>
<td>Authors applied crowdsourcing to simplify and focused on reducing the dependency of IoT networks on the service gateways to reduce the cost of the system. Their work showed the potential of applying crowdsourcing in futuristic IoT.</td>
</tr>
</tbody>
</table>

The authors in [129] proposed a generalized centralized mobile crowdsourcing architecture for IoT. Figure 2 shows the centralized mobile crowdsourcing architecture for IoT in which the information processing is performed at the cloud server. This architecture has four layers: (1) Mobile sensing/gathering layer; (2) Connectivity/network layer; (3) Crowd processing layer; and (4) End-user layer. The mobile sensing layer has sensors or sensing devices, including user devices such as mobile/smartphones, wearable devices, smart appliances, smart vehicles, and so on.
The major task of the connectivity layer is to provide network connectivity to mobile crowdsourcing. These communication networks can be cellular networks (3G/4G/5G), wireless sensor networks (WSN), Wi-Fi, Bluetooth, and vehicular ad hoc networks (VANETs). These communication technologies transfer the data collected by the sensors,
including mobiles, to the cloud or server. In the crowd processing layer, the data from sensors/mobiles and crowd tasks are stored, processed, analyzed, and visualized. Cloud computing which is a centralized paradigm, plays a vital role in the above purposes.

The Social Internet of Things (SIoT) [22] is another approach to spreading tasks and expanding traditional crowdsourcing networks. Utilizing crowdsourcing in SIoT has advantages such as the advancement of large social networks leading to a larger group of workers. SIoT is usually organized according to mutual friendship and thus can be obtained without incentive issues. The authors in [23] proposed an architecture for crowdsourcing in SIoT. The authors proposed a trustworthy crowdsourcing model in SIoT. Their system model of crowdsourcing in SIoT is shown in Figure 3. The architecture contains three components called sensing entities, end-users, and social cloud. It is a 3-layered structure that includes the base layer, component layer, and application layer. The function of the base layer is to provide the infrastructure for computation and storage to support the upper-layer services. The component layer provides tools for basic component implementation such as owner control, ID management, object profiling, and so on. The function of the application layer is to serve as the interface for the sensing entities, human participants, and end-users.

3.2. Decentralized Architectures of Crowdsourcing

Fog computing for a decentralized computing structure has become more important and can offer benefits such as wide geographical distribution, low response delay, and location awareness. In the fog-based mobile crowdsourcing architecture, caching/forwarding nodes are utilized between the platform and users. The fog layer can be added to give low latency service for IoT. The addition of this layer not only offers fog computing advantages. For tasks with spatial-temporal requirements, the distribution of fog node locations can provide a precondition for task forwarding and improve task allocation efficiency. It also helps to prevent the crowdsourcing platforms from eavesdropping on users’ privacy. The previous subsection presented the generalized or centralized...
crowdsourcing architectures, whereas this subsection focuses on the decentralized crowdsourcing architectures, including fog architecture. Table 3 summarizes the research works for decentralized architectures, including fog architectures.

**Table 3. Summary of the research works on decentralized architectures of crowdsourcing.**

<table>
<thead>
<tr>
<th>Authors &amp; Ref</th>
<th>Year</th>
<th>Focus Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhang et al., [24]</td>
<td>2013</td>
<td>Proposed a distributed crowdsourcing system in Delay Tolerant Network (DTN) in a collaborative way to achieve minimal makespan.</td>
</tr>
<tr>
<td>Tang and Yuan, [25]</td>
<td>2013</td>
<td>The monitoring quality function was defined and applied to find the spatial-temporal correlation between monitoring nodes.</td>
</tr>
<tr>
<td>Min and Kim, [26]</td>
<td>2014</td>
<td>Proposed topology repair algorithm based on redundant node selection for the problem of shortening network life due to excessive energy consumption of individual nodes.</td>
</tr>
<tr>
<td>Cheung et al., [27]</td>
<td>2015</td>
<td>Proposed an asynchronous and distributed task selection in mobile crowdsensing.</td>
</tr>
<tr>
<td>Najimi et al., [28]</td>
<td>2015</td>
<td>Based on effective energy, the authors presented a node selection technique to choose nodes that satisfy the energy consumption requirements and the average monitoring rate of monitoring nodes.</td>
</tr>
<tr>
<td>Jiang and Huang, [29]</td>
<td>2015</td>
<td>The authors presented an intrusion tolerance method against malicious nodes to solve the problems of network vulnerability by deploying redundant nodes.</td>
</tr>
<tr>
<td>Yang et al., [31]</td>
<td>2016</td>
<td>Proposed a task allocation scheme by utilizing social relationships in crowdsourcing systems and loading balancing in a distributed model.</td>
</tr>
<tr>
<td>Ni et al., [33]</td>
<td>2017</td>
<td>Fog-based architecture for spatial crowdsourcing to select fog nodes that meet the spatial attribute requirements for task assignment and the collection of outcomes or results.</td>
</tr>
<tr>
<td>Ni et al., [34]</td>
<td>2017</td>
<td>Fog-based vehicle crowdsourcing architecture and how the fog nodes interact with the sensing vehicle when performing the task, to improve the node selection algorithm. Their studies revealed that the fog nodes were selected based on the spatial attribute of tasks. The temporal attribute of tasks was not considered, thus causing inaccurate task allocation.</td>
</tr>
<tr>
<td>Rahman et al., [35]</td>
<td>2017</td>
<td>Mobile-edge computing architecture for position monitoring and considering node energy consumption and correlation.</td>
</tr>
<tr>
<td>Chowdhury and Benslimane, [36]</td>
<td>2018</td>
<td>A redundant node identification and relocation strategy was presented to enhance energy efficiency and coverage of the network.</td>
</tr>
<tr>
<td>Balevi and Gitlin, 2018 [37]</td>
<td>2018</td>
<td>The authors presented the optimization of the number of selected nodes to minimize transmission delay and maximize the average data rate. Their work aimed to use appropriate methods to determine redundancy and improve the monitoring quality and robustness of the network.</td>
</tr>
</tbody>
</table>

The authors in [22] proposed a generic architecture of decentralized architecture of mobile crowdsourcing applications. The architecture is illustrated in Figure 4. The architecture is similar to a mobile peer-to-peer network where each peer can communicate with other peers in the local vicinity and contribute to identifying mobility patterns. In the crowdsourcing process, mobile devices explicitly or implicitly capture sensor data in the background. The decentralized architecture performs the information processing and storage in the mobile device and local database.
The authors in [38] proposed a fog-based spatial-temporal crowdsourcing architecture shown in Figure 5. The proposed architecture has three layers: (1) Crowdssourcing platform layer; (2) Fog layer; and (3) User layer. The fog nodes in the fog layer are network components closer to the user end, for example, base stations (BSs), roadside units (RSUs), and edge routers. The functions of fog nodes include computation and storage space. Each node can decide its interest attributes based on the behavior of the mobile users within its vicinity to achieve task matching.
Figure 6. Elements and architecture for proposed Crowd-IoT. (a) Crowd-IoT elements; (b) Crowd-IoT architecture.

The Identification Layer gives a distinctive identity to each object within the Crowd-IoT architecture. For mobile crowdsourcing, mobile numbers are assigned to a specific
subscriber station in a phone network. The mobile numbers could be either global or local. A global number begins with a country code defined by ITU-T. Local numbers can be extended by providing global uniqueness when calling outside the local area. The Crowd Object Layer gathers crowdsensing data from the IoT objects and sensing elements and sends/forwards the data onto a central server/cloud storage for information processing. In the traditional IoT, these data and object sources would be captured from sensing devices such as RFID tags, embedded, mobile and wearable sensors. The role of the communication and network layer is to provide network connectivity to the various devices and components within the mobile crowdsourcing system. There are various communication networks in mobile crowdsourcing that can be utilized, such as Wi-Fi, WSNs, cellular technologies, and vehicle networks (e.g., VANETs).

One of the challenges of mobile crowdsourcing IoT is the requirement to handle multiple devices that can produce heterogeneous and different data types. For the challenge of heterogeneous data sources, one of the possible solutions is to have a middleware platform in the crowdsourcing architecture for information exchange and collaboration among the devices. Within the architecture of Crowd-IoT, the middleware layer provides a software layer between the network communications, operating system, and applications. As a result, our Crowd-IoT architecture will enable the efficient use of mobile crowdsourcing IoT in several applications in a smart city environment with numerous heterogeneous data sources. Such applications includes smart transportation, smart health, smart homes, and so on. The middleware service components can be classified as functional or non-functional. Examples of functional middleware components include resource discovery, data, and code management; and examples of non-functional requirements include components for scalability, reliability, and security. The Crowd Computation Layer serves as the processing layer in the Crowd-IoT architecture and contains the hardware and software components. Figure 7 shows the architecture and components of the crowd computation and processing layer. The layer consists of five components: (1) Crowd Context Database, (2) Big Data and Processing, (3) Task manager, (4) Contribution manager, (5) Workflow Management.
The Application and User Layer in the Crowd-IoT provides services to the service providers and end-users for crowdsourcing applications. Examples of applications for service providers include allowing requesters to initiate the tasks and retrieve results through web-based or mobile technologies. Examples of applications for end-users include allowing entities (e.g., individuals, vehicles) to send requests to the server/platform and receive the results.

From our study, it is observed that the choice of crowdsourcing structure, components, and type of architecture depends on the type of application. Applications with a focus on data storage, task management and distribution, security, and privacy adopted the centralized architecture. On the other hand, applications with high concerns for energy consumption, network connectivity and coverage, high processing speed, load balancing, security, and privacy utilized decentralized architecture.

Security is a major challenge for Crowd IoT implementations due to the lack of common standard and architecture for IoT security and the requirement to handle data in heterogeneous networks. Security, privacy, and trust components provide services that meet these requirements in the Crowd-IoT. The trustworthiness, privacy, and security for crowdsourcing in IoT will be further elaborated in the next section (Section 3).

4. Trustworthy, Security and Privacy in Crowdsourcing Internet of Things

In crowdsourcing, current research for privacy protection can be grouped based on two different phases found in crowdsourcing: (1) Task allocation; and (2) Data aggregation. Task allocation is an essential component and indispensable service for the response time and ensuring task quality. To have efficient task allocation and recruitments, the center or server needs to utilize data-related participant and worker profiles. This poses a
threat to privacy leaks. There are privacy protection techniques for the requestors/owners or participants/workers during the phase of task allocation. There are few research works for privacy-preserving task allocating mechanisms to preserve the privacy of the locations for the end-users. The authors in [130] proposed a helpful task assignment to protect task and worker privacy protection based on cryptosystems. Although there are many approaches to preserving privacy in crowdsourcing, the focus of this section is on studies that protect privacy in crowdsourcing IoT and Fog based crowdsourcing.

For privacy-aware data aggregation, the current research can be grouped into two approaches: (1) Collection-oriented; and (2) Statistic-oriented. The former aims to collect high-quality sensed data, while the latter targets to perform different statistical analysis on collected or sensed data. Privacy concerns can be mitigated by increasing the perceived trustworthiness and information security in crowd systems. Trust and reputation management systems are commonly used widely in crowdsensing to assess participants’ trustworthiness. This section presents some of these key works by classifying them into three main categories: (1) Trustworthy in crowdsourcing IoT, (2) Privacy in crowdsourcing IoT, and (3) Security in crowdsourcing IoT. Some authors and works may consider more than one aspect in their research.

4.1. Privacy in Crowd-IoT

For specific works of privacy in IoT, privacy maintaining framework for IoT was developed by Kung [39]. In this work, the authors focused on data controllers, data processors, and integrators for privacy preservation. The authors in [40] proposed a privacy-preserving IoT architecture to maintain privacy in large-scale IoT networks. The authors in [41] proposed an approach termed Lightweight Privacy-preserving Data Aggregation (LPDA) for Fog computing-enhanced IoT. The proposed LPDA combines homomorphic encryption, Chinese Remainder Theorem, and hash chaining to enable fog device aggregation at the network edge. Although there are many works on privacy issues in crowdsourcing, there are few works on privacy preservation in specialized crowdsourcing IoT. This subsection reviews recent and specific research on centric crowdsourcing IoT and edge-based crowdsourcing IoT.

Yang [42] proposed a crowdsourcing privacy framework under fog computing architecture. Figure 8 shows the framework, which has four components: (1) Requester who collects data and sends the tasks to the crowdsourcing server, (2) Server to assign tasks to participants, provide incentive mechanism and manage the data collection, (3) Fog nodes for tasks assignment and data collection in the selected region, and (4) Workers who select the task (fog node) and report their location information. The authors proposed a novel approach to protect participants’ location information. The key idea is to utilize the fog nodes to collect statistics of different paths. The proposed architecture contains three main components (iterative statistics, adaptive sampling, and secure reporting). Theoretical analysis and experiments were performed to show that the approach could provide a strict privacy guarantee and enhance the performance.
Figure 8. Crowdsourcing privacy framework under fog computing architecture [42].

Zhang [43] aimed to address the issue of privacy-preserving crowdsourcing. They proposed an approach to managing the workers’ computational requirements by allocating parts of the computation to the fog node. The proposed fog-assisted crowdsourcing system for task allocation is shown in Figure 9. There are five key components in this system: (1) Crowdsourcing server; (2) Workers; (3) Customers; (4) Fog node; and (5) Key generation center (KGC). The first four components are general in any crowdsourcing IoT or fog-based crowdsourcing. The KGC is a trusted entity responsible for initializing the system parameters (e.g., public key, master secret key) and allocating secret keys to customers and workers. The fog nodes are deployed at the network edge. These nodes share the partial computation from the end-users or works. The privacy-preserving task allocation scheme performed the privacy-preserving task-and-keywords matching. The workers who satisfy the task requirements can decrypt the cipher-text of task content and can be appointed. The workers produce the cipher-text of interest keywords and send the data to the crowdsourcing server. Their approach achieves privacy protection of task content and interest keywords and resists the attacks of the workers.

Figure 9. Fog-assisted crowdsourcing system for task allocation [43].

Spatial crowdsourcing is a specific type of crowdsourcing and has drawn much attention in mining. It allows multiple task owners or requestors to outsource their tasks to the central server for recruiting participants. The task release and data aggregation may have privacy threats for both parties. The central server does not fully know the entire geographical user distribution. Wu [44] proposed to address the above problem by presenting a fog-assisted spatial crowdsourcing architecture. Figure 10 shows their fog-assisted spatial crowdsourcing system. The architecture contains the following components: (1) Spatial crowdsourcing server which provides aggregation services by assigning corresponding tasks of requestors or owners to fog nodes, (2) Tasks owners who generate spatial aggregation tasks and collects the data from assigned locations, (3) Fog nodes which are deployed at the network edge to manage the task distributions in a geographical area, (4) Mobile participants who carry the mobile sensing devices and perform the sensing tasks, and (5) Authority center which has the responsibility for managing the system including registration of system entities, and key distribution.

In their architecture termed as PTAA (Privacy Aware Task Allocation and Data Aggregation), fog nodes are deployed in various locations to distribute the tasks and perform the data aggregation in a privacy-aware manner. The PTAA approaches consists of sev-
eral phases: (1) System initialization; (2) Task generation/allocation; (3) Data collection/aggregation; (4) Data decryption; and (5) Data update. In their proposed scheme, bilinear pairing and homomorphic encryption were applied. There are two steps of task allocation in PTAA. This allows the server to allocate the task to fog nodes without requiring knowledge of the task content. To protect the user or participant data privacy, sensing data are encrypted before being sent to the fog nodes. Fog nodes can also detect invalid data, such as modified data from adversaries. In the aggregation phase, secure aggregation protocols enable the fog nodes and the server to compute the encrypted aggregated result.

![Diagram](image1.png)

**Figure 10.** PTAA fog-assisted spatial crowdsourcing architecture [44].

The authors in [45] proposed a privacy-preserving incentive mechanism involving multi-hop routing to preserve task owners or requester’s privacy for social IoT crowdsourcing. Figure 11 shows the privacy-preserving incentive mechanism for the social crowdsourcing service. The authors incorporated privacy and budget feasibility constraints into a problem formulation of utility maximization. The objective of the mechanism is to improve the overall service performance by dividing it into subtasks and routing it to workers. Guidelines on the subtasks division from an entire task and the tradeoff between task accuracy and privacy are provided. In their proposed work for social IoT, people with the direct connection to task requester are considered as 1-hop friends, while 2-hop friends are those who do not have a direct connection to a requester but have a direct connection with 1-hop friends. There are two stages in their multi-hop routing incentive policy: (1) One-hop myopic routing; and (2) Multi-hop payment allocation, to have an equilibrium to motivate workers to forward subtasks to friends with the appropriate expertise. They showed that the cost for the task requester would increase with the length of the task forwarding chain in multi-hop routing. Simulation results verified the theoretical analysis, and the proposed work could achieve high information gain with a privacy guarantee.

![Diagram](image2.png)
Figure 11. Privacy-preserving incentive mechanism for social crowdsourcing [45].

4.2. Trustworthy in Crowd-IoT

Trust and privacy in the previous subsection are usually synonymous. A solution that targets protecting trust relationships should also be protecting user privacy and device information. Many trust-based approaches target to balance the connectivity between service or task requestor/owner and workers. This section discusses recent and specific works related to crowdsourcing IoT for centric or edge computing. In crowdsourcing IoT, a network with trust is important for sustaining applications. However, trust also needs strict privacy-preserving mechanisms that can assure protected environments for all users and prevent the leak of an individual’s personal data and damage of social reputation.

An earlier work on trustworthy crowdsourcing in cloud-centric IoT can be found in [46]. The authors proposed mobility-aware crowdsourcing (MACS) architecture in IoT for smart city management. The MACS applied an incentive mechanism that chooses the users based on an auction and makes payments to the chosen users based on their bids. The MACS also predicts the future locations of participating users in the auction. The authors also proposed Mobility-Aware Trustworthy Crowdsourcing (MATCS) for crowdsourcing IoT. The MATCS first confirms the presence of the participants and updates the trustworthiness of users. The bids, marginal values, and trustworthiness are used to determine the payments to the auction winners. The authors performed simulations to evaluate the performance of MATCS and showed the improvement of smart city management authority (up to 70%) and the decrease in disinformation probability (up to 55%).

The authors in [47] proposed a crowdsourcing assignment model based on social relationship cognition and community detection. Figure 12 shows the proposed architecture of the crowdsourcing model and a use case. In the scenario, based on the mobility of users, a trusted route is established from the service requester (SR) to the service provider (SP) for the assignment of crowdsourcing tasks to realize efficient distribution and delivery of requests. The assignment of crowdsourcing uses decision-making factors, including the service quality factor (SQF), link reliability factor (LRF), and region heat factor (RHF). The procedure has the following steps: (1) Detection of crowdsourcing service candidates in the target location; (2) Computation of crowdsourcing decision-making factors of SQF, LRF, and RHF; and (3) Assignment of crowdsourcing sensing task using the crowdsourcing algorithm. For performance evaluation, simulations were conducted on the mobile crowdsourcing platform developed by the authors. The experimental results showed the correctness and robustness of the approach.
Future Internet 2022, 14, 49

Figure 12. Crowdsourcing assignment model based on social relationship cognition and community detection [47].

Bahutair [48] proposed just-in-time memoryless trust for crowdsourced IoT services. The provider’s trustworthiness is evaluated without requiring historical records (memoryless). Figure 13 shows their framework, which utilizes the service session characteristics and its surrounding environment to infer the service’s trustworthiness. Their approach has three stages: (1) IoT service initiation, (2) IoT service monitoring; and (3) Trustworthiness assessment. Another challenge of assessing trust in dynamic nature is its dependability on how the service is used. To overcome this challenge, the same author Bahutair [49], proposed an approach termed Adaptive Trust framework for crowdsourced IoT services. The framework of usage-based trust assigns trust values to IoT services based on the consumers’ usage.

Figure 13. Just-in-time memoryless trust for crowdsourced IoT services [49].

Xiuwen [50] investigated the impact of trustworthy relationships between Social IoT nodes on social team crowdsourcing mechanism sensing in social IoT. The authors focused on two aspects: (1) Trust-aware data quality estimation; and (2) Task valuation optimization with team connectivity. For the first aspect, they first proposed a trust-aware quality estimation for crowdsourcing in social IoT (termed as TAQ-Crowd). Figure 14 shows the TAQ-Crowd model. There are two major components in TAQ-Crowd. The first component is the TAQ model, an unsupervised probabilistic graphical model to estimate sensing quality. The model utilizes trustworthy relationships to increase the quality estimation. The second component is a computationally efficient mechanism that approximates the maximal overall task valuation and ensures communication connectivity for collaborated task execution. The authors designed a task assignment algorithm termed as CS-Selection. Concerning cost constraints from route planning, a variation of the Traveling Salesman Problem (TSP) algorithm is used to produce a routing network for team
communications. The proposed TAQ-Crowd framework is evaluated through simulations. The results showed the proposed TAQ-Crowd model can achieve good performance and gives near-optimality results for the task valuation.

Sharma [51] considered both trust and privacy-preserving in their solution. They proposed a novel architecture of edge-crowd integration for cooperative trust and privacy preservation in social IoT. In their approach, crowdsources such as mini-edge servers are used to form an edge-crowdsourcing network. The network model, which consists of Stations (BSs) and the Access Points (APs), is shown in Figure 15. Their model contains multiple query manager systems (Main Query Manager (MQM), Crowd Query Manager (CQM), and Individual Query Manager (IQM)). For privacy, data flow, device confidence and authorization, and device authentication are included. For trust, their approach uses crowd-resources to maintain trust-based on-demand trust relaying system. The authors presented a case study for preventing fake news in S-IoT by using the proposed computational paradigm for trust and privacy preservation through edge-crowdsourcing.

Figure 14. Crowdsourcing in social IoT (TAQ-Crowd) [50].

Figure 15. The architecture of edge-crowd integration for cooperative trust and privacy preservation in social IoT [51].
4.3. Security in Crowd-IoT

Privacy and security are related to crowdsourcing. As mentioned in Section 3.1 on privacy in crowdsourcing of IoT, some existing approaches resolve users’ privacy by data encryption. However, some encryption algorithms, e.g., public-key encryption methods, are not suitable for crowdsourcing because they require the parties to issue the keys for decryption to be made online. Implementing crowdsourcing with a large number of users and tasks is complex. On the other hand, conventional symmetric encryption methods are not applicable to crowdsourcing because the data can be compromised after the keys have been revealed. This subsection reviews specific research works on security in crowdsourcing IoT, followed by security in edge-based crowdsourcing.

The authors in [52] proposed a ciphertext-policy attribute-based encryption (CP-ABE) mechanism for crowdsourcing. The proposed crowdsourcing system with CSMA-ABE is shown in Figure 16. The system components contain requesters, workers, a semi-trusted third party, and attribute authorities. In the CSMA-ABE scheme, the requestor applies the CP-ABE encryption to encrypt the task data after specifying an access policy based on the task requirement. After task matching, the partially decrypted task is sent back to workers. The workers decrypt the task to obtain the task data and the symmetric key. Each worker encrypts the task result using the symmetric key and sends the encrypted result back to the platform. Their work focused on key management and distribution in multiauthority scenarios while ensuring forward/backward security of tasks.

![Crowdsourcing system with CSMA-ABE](image)

**Figure 16.** Crowdsourcing system with CSMA-ABE [52].

In crowdsourcing IoT, signcryption is a cryptographic primitive for ensuring the authenticity and confidentiality of crowdsourced data. A novel identity-based signcryption (IBSC) approach using bilinear pairing was proposed by Karati [53] for crowdsourcing industrial IoT. The architecture for crowdsourced industrial IoT and the network model of the proposed signcryption scheme are shown in Figure 17. There are four major units: (1) Cloud Server for data processing and storage; (2) PKG for generation of system and secret keys; (3) IIoT Data Owner, which applies the receiver and PKG public keys to signcrypt data; and (4) Receiver which allows users to obtain a secret key from the PKG and perform unsigncryption over signcrypted data. The performance comparison shows that the IBSC scheme is efficient in terms of communication and computation costs with signcryptext size and is suitable for crowdsourced industrial IoT applications.
Remote user authentication is important for crowdsourcing IoT. The authors in [54] proposed an approach based on a chaotic map for a user authentication scheme that is secure for crowdsourcing IoT. Their scheme considered three factors: smart cards, passwords, and personal biometrics. These factors can be authenticated by the Remote server, which participates in the key agreement process. The proposed approach has some important features. First, it does not store or transmit any user-sensitive information to protect the user’s privacy. That means there is no revealing of sensitive information during the mutual authentication and key agreement phase. For authentication, the authors applied zero-knowledge proof, fuzzy extractor, and Chebyshev polynomial chaotic maps. The chaotic map zero-knowledge proof allows the user to prove the ownership of a secret key without revealing sensitive information. A fuzzy extractor can then extract the random string and the auxiliary string from a biometric input. There are five phases in the overall process: (1) Registration phase; (2) Login phase; (3) Authentication and key establishment phase; (4) Password and biometric change phase; and (5) Smart card or device revocation phase. Their approach has the advantages of high security and low communication and computational costs. These approaches can benefit resource-constrained and battery-limited devices in IoT and the next generation 5G communication.

A group of researchers studied the potential to combine crowdsourcing, IoT, and 5G. Considering 5G IoT is the ecosystem formed by 5G together with IoT, the authors in [55] analyzed the idea of crowdsourcing analysis against cyberattacks in 5G IoT. They started to analyze some use cases in crowdsourcing focusing on security. The potential areas of a 5G ecosystem where crowdsourcing can stop or mitigate the effect of local and remote attacks and discourage cybercriminal activities and behavior is highlighted. They also studied how these can be applied at different levels involving crowds which can be service providers, infrastructure, and users. Trusting 5G to solve IoT security issues or that IoT devices are more secure with full deployment of 5G may not be correct. Cooperation mechanisms are needed so that the IoT device data can be used for the detection and mitigation of threats.

There are few researchers working on the security aspects of fog-based crowdsourcing. The authors [33] proposed a fog-based spatial crowdsourcing (Fo-DSC) framework for task assignment and secure data deduplication. Figure 18 shows the system model of Fo-DSC, which contains three layers: (1) Service layer, (2) Fog layer, and (3) Mobile user layer. The spatial crowdsourcing (SC) server assigns the tasks to the fog nodes in the sensing vicinity. The fog nodes then distribute the task to users based on the task requirements from SC-server. There are six phases in the Fo-DSC framework: Setup, T-Releasing, T-Allocation, DCollection, D-Ddeduplication, and D-Reading. For secure data deduplication, Fo-DSC framework allows fog nodes for data deduplication without disclosing any information about the reports. Message-lock encryption is used for encryption in the D-
Collection phase to achieve data confidentiality. A fog node checks the identification of sensing data in two crowdsensing reports by comparing the ciphertexts. Consistency is checked after decryption to prevent duplicate-faking attacks. To record the contribution of the mobile users with deduplicated data, key homomorphic signature is utilized to sign the sensing data, followed by aggregation of signatures of mobile users by fog nodes. The proxy re-encryption scheme is applied to prevent spatial task data from being disclosed to external attackers. The scheme also enables the SC-server to re-encrypt for generating the ciphertexts that fog nodes can decrypt.

Some researchers used blockchains [56] to improve the security of decentralized or fog-based crowdsourcing. To prevent malicious users’ participation, user privacy leakage, and so on, Yu [57] proposed a fog-blockchain architecture for crowdsourcing reputation management. Figure 19 shows the proposed framework. There are four elements: crowdsourcing platform, the task requesters, fog-blockchain layer, and mobile workers layer. A cross-layer privacy protection scheme was designed to separate the tasks and user’s identity and tasks. There are three phases: system initialization, crowdsourcing services, and reputation management. Each user is registered on the crowdsourcing platform and provides their information in the system initialization phase. In the second phase crowdsourcing service phase, a task is submitted by the task requestor to the platform, which assigns the encrypted task to the fog nodes which match the task requirements. In the last phase of reputation management, the platform then decrypts, analyzes, and evaluates the task results. The Workers’ reputation values were then calculated, and the reputation records were feedback to the fog nodes. At last, the workers’ reputation with the real user ID was updated by the fog nodes update in the maintained blockchain. Theoretical security analysis was evaluated, and experiments were conducted to evaluate the proposed CrowdR-FBC framework.
In machine learning, federated learning, also known as collaborative learning, can train an algorithm across multiple decentralized edge devices or servers. This is different from traditional centralized machine learning techniques, which either require local datasets to be uploaded to one server or often assume that local data samples are identically distributed in traditional decentralized approaches. Zhao [58] proposed a blockchain-based crowdsourcing federated learning system for IoT devices. Several technologies such as edge crowdsourcing and computing, mobile technologies, blockchain, distributed system, and storage, and federated learning, are used to build up the system. Figure 20 shows an overview of their system. There are three major components, manufacturers, customers, and blockchain in the system. For a crowdsourcing federated learning task, manufacturers send a request. Then customers who are interested sign on their trained models to the blockchain, which serves as a centralized server to collect these models. In the crowdsourcing tasks, authors also designed an incentive mechanism to attract participation. The blockchain will monitor updates during the federated training. Their system can guarantee the model accountability to prevent malicious attacks.

The study in this section reveals that approaches employing public-key encryption methods are unsuitable for crowdsourcing since they need the persons issuing the keys for decryption to do so online. This will make the implementation of crowdsourcing with a large number of users and tasks very complex. Conventional symmetric encryption
methods on the other hand are inapplicable to crowdsourcing since the data can be compromised once the keys are revealed. Using blockchain techniques and advanced secure encryption mechanisms hold a great promise for crowdsourcing IoT systems.

5. Resources, Sharing, Storage and Energy in Crowd-IoT

As mentioned in previous sections, mobile edge computing is one of the technologies for envisioning IoT and implies computational speed and processing at the edge of the network. The fundamental idea of mobile edge computing is to utilize the idle computing and storage resources at the mobile edge via the appropriate service or data caching and task offloading. It is different from the conventional centralized cloud computing approaches. Mobile edge caching is a promising technology for next-generation mobile networks to offer cloud services and storage capabilities at the edge of networks. This section will provide insights on resources, storage, and energy requirements for crowdsourcing IoT or edge-based crowdsourcing. These topics include mobile edge caching and sharing, multi-dimensional resources, energy services, and requirements.

One of the problems with mobile edge computing is how to offload tasks to the right locations at the edge. Mobile edge computing servers need to be deployed in advance to offload tasks to the edge cloud. However, it is not a trivial task to deploy and maintain a server for both economic and technical reasons. The authors in [59] investigated how mobile devices could share the heterogeneous resources at the edge. They proposed a multi-dimensional resource crowdsourcing architecture for sharing in mobile edge computing. The aim is to offload tasks optimally to mobile devices while maximizing the overall task completion and minimizing the total energy cost. Figure 21 shows the proposed sharing architecture with four mobile devices. Their simulation results showed their approach could reduce the task failure probability by 10% to 35% and the total cost by 25% to 85%.

![Figure 21. Architecture for sharing of devices in mobile edge [59].](image)

Crowdsourcing energy as a service or crowdsourced energy service has the potential to create a green computing environment. The authors in Lakhdari [60] focused on designing crowdsourced energy services to meet the IoT users’ energy requirements in confined areas. Figure 22 shows the crowdsourced energy services architecture. The authors proposed a new composability model considering (i) the energy usage behavior and (ii) the spatio-temporal features of services/queries of the IoT devices. The crowdsourcing energy services are formulated as a multi-objective optimization problem to meet the user’s energy requirements in the earliest and shortest time intervals. Their experimental results revealed the scalability and effectiveness of their approach in those composition scenarios.
Jiang [61] studied and conducted the economic analysis and strategic interactions between content providers (CP) and multiple edge devices (EDs). The interaction between the CP and EDs can be: (1) The CP can determine to share a certain revenue with EDs as the incentive of caching and sharing contents; (2) The EDs can determine whether to cache and share contents for the CP. This interaction is formulated as a 2-stage Stackelberg game. In the first stage, the CP maximizes its profit by determining the incentive mechanism offer and the ratio of revenue shared with EDs. On the other hand, in the second stage, EDs maximize their payoffs by determining whether to cache contents based on the incentive scheme and the caching cost. Figure 23 illustrates the Crowd-MECS architecture for edge caching and sharing. Their experimental results showed that the proposed Crowd-MECS could significantly improve the CP profit and ED total welfare.

Figure 23. Crowd-MECS architecture for edge caching and sharing [61].

Chen [62] presented a device-to-device (D2D) crowd task offloading scheme for 5G mobile edge computing that enables a huge crowd of devices at the network edge to exploit network assisted D2D collaborative sharing of computation and communication resources. Figure 24 shows the D2D crowd framework. The scheme aimed to create collaborative task execution in an energy-efficient manner for mobile users at the network edge. To achieve this aim, the authors first formulated an energy-efficient D2D crowd task assignment problem such that a task can be executed either locally on its primary mobile device or offloaded to another nearby device. They further presented a graph-matching-based optimal task assignment policy to discover the minimal weight matching solution for the D2D crowd task assignment problem using Edmonds’ Blossom algorithm. Performance evaluation of the proposed scheme via numerical analysis demonstrated an improved performance of over 50% energy consumption reduction in comparison to the scenario of local task execution.
Tang [63] designed a general 3C framework that allows mobile users to exchange all three categories of resources (communication, computation and caching) via D2D connection. Figure 25 illustrates the general 3C framework. The authors focused on an energy minimization problem to quantify the benefit of the designed framework and show that the 3C framework always achieve a lower overall energy usage when compared to previous 1C/2C models. In addition, they demonstrated that energy savings are greatest when the probability of user connection and content caching ratio are neither too high nor too small. Their numerical results revealed that overall 3C framework may cut total energy consumption by 82.98% when compared to the 1C/2C models.

IoT networks are usually composed of heterogeneous and resource constrained devices. Numerous novel applications and services that require hybrid types of resources can be supported by pooling and sharing diverse processing and communication capabilities among mobile devices. Because of the heterogeneity of transmission technologies used in devices and the time-dependent feature of wireless transmission, the cellular connection’s quality among devices might differ even when they are in the same location. Therefore, more effective data offloading techniques are required such that devices with weak cellular connection can offload their data to devices in close proximity with strong cellular connections to enhance energy efficiency. With modern mobile and IoT devices equipped with powerful sensors, capable of gathering and communicating large amounts of data, several applications can leverage on the crowd of collaborative devices’ computation resources at the network edge for efficient in-situ data processing across diverse

Figure 24. D2D crowd framework [62].

Figure 25. General 3C framework [63].
6. Applications for Crowd-IoT

The Internet of Things (IoT) has changed the way individuals interact with physical world things. IoT enables physical items to interface and share information by connecting smart objects in a network structure, and enabling new types of interaction between objects and people. The extraordinary growth of the IoT enables smart objects to communicate with one another, resulting in a plethora of ubiquitous applications. Crowdsourcing is viewed as a viable approach in the IoT paradigm for offering effective sensing, computing, as well as processing services to a specific task supplied by customers, effectively combining the abilities of the crowd. Crowdsourcing is a unique sourcing approach employing collaborative brainstorming by the community. It enables users, also known as crowdsourcers, to benefit from a big and open group of individuals, generally known as the crowd, cooperating to produce cumulative outcomes using a cloud crowdsourcing platform [131]. Crowdsourcing can play an important role in Internet of Things (IoT) applications for information sensing and gathering where the participants are equipped with geo-located devices [132]. Thus, mobile crowdsourcing can be seen as a new paradigm that is contributing to the development of the IoT. They can be merged to form a new and essential platform in crowdsourcing IoT paradigm for data collection from different sources and communication mediums for various applications in environment, transportation, health, surveillance, and so on. This new paradigm can be used for the practical deployment for complex distributed problem-solving and providing reasonable solutions for communication and cooperation among intelligent devices. It is frequently utilized in a number of applications that affect many elements of our life. These application areas are classified as generic and specific areas.

6.1. General Applications

Crowdsourcing IoT has been generally and extensively used in research in a number of fields. This subsection reviews research in these fields, which includes localization, monitoring and guidance, 3D, video and live streaming, and semantic and data mining.

6.1.1. Localization

In many emerging fields, location consciousness has become progressively necessary. In the context of environment, location information is studied and classified as outdoor and indoor localization. Several tools and techniques have been developed to estimate accurate location information. For outdoor localization, the global positioning system (GPS) has been widely and successfully used to accurately estimate location but fails to provide such accuracy in indoor location estimation due to its costly energy tags and weak penetration of its signal via a complex indoor environment [64,65]. However, indoor location-based services are realized using various systems and techniques such as Wi-Fi, ultra-wide band (UWB), RFID, Bluetooth, light and magnetic field, FM radio, acoustic analysis, sensor network etc. Amidst the afore-named systems, Wi-Fi signal has been broadly utilized compared to other techniques due to its low cost, vast coverage features, and its ability to successfully allow signal accumulation and aggregation operation in indoor locations with no need for extra investment in infrastructure [65–67]. Intelligent localization that can scale with huge data without human perception is becoming increasingly important in next-generation Internet of Things (IoT) systems. Localization based on crowdsourcing has been shown to be beneficial for commercial location-based IoT applications.

Li [68] proposed an enhanced localization technique based on crowdsourcing which avoids the intervention of users and tuning of parameters. Figure 26 shows the diagram of proposed crowdsourced-based localization method The proposed technique integrated

devices. Incentive mechanisms also need to be incorporated to encourage device collaboration.
three sensor (wireless, inertial and magnetic sensors) and utilized crowdsourced sensor data to update simultaneously both the magnetic and wireless databases throughout the localization process. Since in multi-sensor integration, it is complicated to set and tune the variability of fingerprinting solutions, the authors presented the fingerprinting accuracy indicator (FAI) to predict the accuracy of both wireless and magnetic fingerprinting from signal, geometry and database level. They further proposed an FAI-enhanced extended Kalman filter (EKF) to effectively increase the accuracy and reliability of localization. The investigation of the advantages and drawbacks of each FAI factor on magnetic and wireless fingerprinting using proposed crowdsourcing localization technique showed that the weighted DSF based FAI successfully predicted the location error and outliers. On the other hand, the geometry-based FAI proved effective in predicting short-term errors as well as outliers but was deficient in detecting long-term error. The signal strength FAI, on the contrary, could comparatively predict long-term location error effectively.

![Diagram of proposed crowdsourced-based localization method](image)

Figure 26. Diagram of proposed crowdsourced-based localization method [68].

He and Chan [69] proposed a survey reduction and signal map construction method called Surecose based on implicit crowdsourcing, and cooperation of low-cost collaborative sensors. Figure 27 shows the system flow of the proposed survey-reduction framework. First, the authors used a light-weight sparse collection of signal map to instantiate Surecose. After the background application is set up, naïve users which act as implicit surveyors unintentionally upload crowdsourced signals. It is possible for phones to recognize other phones (mobile sensors) or beacons (fixed sensors). To label the location of the sensors, the proposed method calculates the proximity of the interacting IoT devices and can clearly locate the users through collaborative local labelling as well as sensor calibration in spite of the device used. Since estimating user locations (using radio signals) is usually sparse and irregular in the site, preprocessing is therefore necessary for complete scaling of the signal map for actual use. To address this, the authors modified the Gaussian Process (GP) to enable the construction of signal map anywhere. Using the crowdsourced signals that have been labeled, the proposed method adaptively predicts signals at the structured grids of reference points (RPs) and developed a comprehensive signal database far off the sparse one. The modified GP takes into consideration the directions of the received signal in regard to APs, which in turn adjust to the complex and segregated indoor environment. The proposed method was experimented in different environments (campus hall, international airport, and shopping mall), and results showed that Surecose can
adjust to different environments and construct a signal map with over 70% reduction in survey cost.

Utilizing crowdsourced Wi-Fi fingerprinting together with Micro-Electro-Mechanical System (MEMS) sensors, Yu [67] introduced an autonomous algorithm for 3D indoor localization. First, the authors designed an enhanced complementary filter (ACF) and used it to integrate data from multi-sensors to produce precise attitude information, constraining cumulative errors that emanate from the gyroscope update. The ACF variables could be adaptively adjusted via quasi-static magnetic field (QSMF) and external acceleration detection. In addition, gradient decent algorithm was designed to enhance the forward pedestrian dead reckoning (PDR) crowdsourced trajectories to further reduce the cumulative error and enhance the robust generation of MEMS sensors database. Furthermore, a quality evaluation (QE) criterion was introduced to approximate the correctness of the crowdsourced trajectories and obtain a weighted fusion so as to achieve autonomous generation and updating of a highly accurate Wi-Fi fingerprinting database. Finally, the authors introduced an adaptive particle filter (APF) to obtain autonomous and accurate performance of crowdsourcing-based 3D indoor localization, using information supplied by the MEMS sensors and Wi-Fi fingerprinting. Experiments in real-world complex indoor environments proved that the proposed method was very effective.

Zhao [70] Leveraged on smartphone crowdsourced Wi-Fi data and inertial measurement unit (IMU) data to construct a map-free and calibration-free indoor positioning system (IPS) named GraphIPS that dynamically obtains accurate radio maps. First, GraphIPS combines crowdsourced data to formulate a graph. The nodes in the developed graph depict the positions of devices or user steps in a time series. The edges are also approximated using the Wi-Fi RSS, IMU accelerometer, and AoA data. Second, the multidimensional scaling (MDS) algorithm is utilized to efficiently resolve the formulated graph and calculate the positions of the user’s steps. Radio maps may be created without the need for calibration or map information by utilizing APs as anchors. The testing findings demonstrate that GraphIPS’s accuracy is similar to that produced by calibration-based technique but better than that of the PDR or trilateration-based methods.
6.1.2. Monitoring and Guidance

Wireless network technologies advancement and the demand for mobile internet devices have enabled the creation of participatory sensing networks that can exploit mobile user’s mobility, smartphone built-in sensor and communication heterogeneity of existing infrastructure to obtain ubiquitous sensing results in a low-cost and improved manner [71]. Smartphones with numerous built-in sensors can be utilized to collectively retrieve environmental, infrastructural and transportation conditions to form a body of information. Specifically, using smartphones together with IoT localization device can effectively collect current location and distribution of indoor individuals for smart application.

Chen [72] proposed a mobile system called iBaby for monitoring and finding children who drift away from their parents using IoT devices (wearables) to actively detect if they are being led away by strangers; and utilizing neighboring smartphones to collaboratively find a missing child via IoT localization technologies. The system has two modes—monitoring and finding modes. In the monitoring mode, the children are provided with wearable devices that contain a mobile iBeacon and 3-axis accelerometer for calculating the distance between the parent and the child, detecting stranger holding, informing parents of stranger holding behavior and broadcast the signal of the mobile iBeacon to neighboring smartphones to initiate crowdsourced finding. The parents, on their side, can monitor the current state of the child, receive warning alert from wearable devices, as well as view current trajectories of missing children using smartphones/mobile tablets. In the finding mode, the BLE signals of the iBeacon is continuously broadcasted by the wearable device to the neighboring BLE scanning enabled smartphones to make internet connection via Wi-Fi access point or 4G-based stations. The signal of the mobile iBeacon received by the neighboring smartphones is utilized to seamlessly find the present location position via outdoor GPS/GNSS as well as indoor IoT/Wi-Fi localization. A posture recognition method based on multi-feature, artificial neural network and convolutional neural network (CNN) is designed to precisely detect the holding behaviors of the stranger. The system was implemented using an iOS-baby prototype together with Arduino wearable device and a fixed iBeacon which proved the feasibility and precision of a system. Figure 28 shows the system architecture of iBaby.

![Figure 28. System architecture of iBaby [72].](image)

Chen [73] designed and implemented EasyFind, a mobile crowdsourced guiding system (utilizing smartphones) to guide indoor people to make timely emergency evacuation as well as find lost items via IoT localization technologies. Figure 29 shows the system
architecture of EasyFind. In an ideal situation, the proposed platform can utilize the density of indoor users in all areas to provide the quickest grouping parts that has the shortest/lowest moving time to a destination position. In an emergency situation, the proposed system can utilize spatial and temporal mobility of indoor users to evacuate all users in the lowest/shortest over escaping time. In addition, EasyFind system can collaboratively detect a lost item filled with a dynamic iBeacon node via crowd sourced sensing network comprised of mobile smartphone users in position with fixed iBeacon nodes. To achieve precision, in localizing the lost items, the system addressed six items localization cases to minimize the positioning error associated with both various numbers of smartphones finding the lost item and static iBeacon nodes surrounding them. EasyFind was implemented in an android based prototype and experiment proved its effectiveness.

Yu [74] proposed a situation-enabled method for detecting and blacklisting malicious phone calls/numbers based on crowdsourcing and edge computing. First, the authors developed a self-loading teaching system which fetched a phone user’s everyday activity data and utilized it to train a Long Short-Term Memory (LSTM) model, which is then employed to predict the phone user’s future activity. The tagging system is enabled to have reference to the history of the caller ID upon the prediction, and if the call time is inappropriate, the caller ID will be automatically marked as malicious, but user’s ability to adjust the tag will be sustained. Second, a local blacklist is created, and each device classifies the blacklist into two: 1. A group of caller IDs (that shares the same) whose area code match that of the prefix of call recipient. 2. A group of caller IDs whose area code differ from that of the caller recipient. The two classes of blacklist are shared by each device to the edge server and utilized to develop a regular list of caller IDs in two categories (edge local and foreign blacklist) which is intermittently disseminated to user devices at

![System architecture of EasyFind](image_url)

**Figure 29.** System architecture of EasyFind [73].
the back-end to effectively improve their tagging system. Evaluation conducted through simulation and implementation verified that the proposed system is feasible.

Huang [75] developed an open-source tool called IoT inspector for harvesting labelled network traffic from smart home appliance from within real-life home network based on crowdsourcing. Unlike Netalyzer, [76] and HomeNet Profiler [77], IoT inspector is designed to build datasets of smart home network traffic on a large scale. It has the ability to collect labelled network traffic from different categories of device and vendors. The aim is to make available to the public a large dataset that will drive new research into smart homes. Furthermore, the authors conducted a study on security and privacy of smart home device using labelled network traffic generated by IoT inspector from 44,956 smart home devices, 4322 users covering 13 groups and 53 vendors. The result of their study revealed that the majority of device vendors use out-of-date Transport Layer Security (TLS) versions as well as advertise weak ciphers. It also revealed different third party advertisers as well as tracking domains on smart TVs. IoT Inspector’s data can also drive research in areas such as health care and network management.

6.1.3. Video & Live Streaming

The widespread availability of IoT cameras, including dash mounted cameras, smartphone cameras, drone cameras, and Google Glass, has supported a plethora of beneficial applications [78]. Video and photo crowdsourcing, which takes advantage of the power of IoT devices and everyday mobile users for large-scale sensing, has been utilized by researchers in many applications areas such as 3D reconstruction, live streaming, etc.

Yu [79] proposed a framework for photo crowdsourcing based on Edge computing. Figure 30 shows the proposed edge computing-based crowdsourcing framework. Their aim is to utilize ubiquitous mobile and IoT devices within the network edge to harvest a collection of representative photos for real-time 3D model reconstruction in cognizance to monetary cost and network resource issue. The framework consists of three modules: (1) photo/participant module which utilizes photo crowdsourcing for collection of photos at reduced monetary cost. (2) a 3D reconstruction module, which can utilize the selected photos contained in the MEC server side to reconstruct the needed 3D modules. (3) a caching module that makes dynamic decisions of whether to cache the reconstructed 3D reconstruction in the MEC server or not. First, the authors recommended a monetary-based motivation strategy which promotes pricing of photos based on data size, freshness, resolution and wireless channel state of their linked participants. This is to ensure the selection of enough and superior fresh photos from closet participants which guarantees great results of 3D reconstruction. They also proposed a photo selection strategy to choose a representative set of photos with low cost to meet the expected coverage requirement. However, finding an optimal set of photos with low cost is proven to be NP-hard. The author therefore presented a grid-based approximation algorithm in order to get a near-optimal solution. Furthermore, the author designed an optimal resource allocation strategy to reduce the highest uploading delay. An evaluation of the proposed framework through experimenting with real-world datasets revealed that the framework is superior to existing schemes.
Figure 30. Proposed edge computing-based crowd sourcing framework [79].

Zhu [80] demonstrated a system that employed commercial fleets as Vehicular Fog Computing (VFC) nodes for real-time video crowd sourcing and processing from vehicles within the communication domain. The feasibility of the system is assessed by first, investigating the availability of fog nodes with respect to the spatial temporal dispersal of the buses. A real-life dataset generated from city of Luxembourg traffic network was used for the simulation experiment. The networking performance of the proposed VFC-based video crowd sourcing system was evaluated using Simulation of Urban Mobility (SUMO) [81] and VeinsLTE [82] at differing time intervals over 24 h. The simulation result revealed that over 90% client vehicles are positioned lower than 200 m from the nearest vehicular fog node, which implies that within a short transmission distance, they can stream videos in real-time to the VFC node. However, when there is crowd in the traffic (i.e., over 200 client vehicles plying a designated area), more than 80% of client vehicles would be able to connect to a VFC node within 100 m.

Wang [83] provided a background review of current crowdcast solution in uniform crowdcast as well as personalized crowdcast, and highlighted their challenges and opportunities. They further designed an intelligent framework based on edge-learning (EL cast) which combines CNN and deep reinforcement learning (DRL) in an edge computing architecture to enable personalized crowdcast. A trace-driven experiment showed that the proposed framework was more effective than current approaches.

Zahoor [84] presented an overlay network architecture to supplement the enhanced multimedia broadcast multimedia service (emBMS) standard architecture in order to allow crowd sourcing live video providers to multicast. Figure 31 illustrates the Multicast at edge architecture. The authors designed a Virtual Network Function (VNF) in order to identify possible multicast scenarios, using user request for a live video from a particular locality. Information collection, validation of a multiple multicast framework as well as initialization of emergency multicast service is executed by the VNF application server on the fly. The authors executed a simulation evaluation of the proposed architecture using Facebook live video datasets. The result showed a considerable advantage of the proposed architecture in respect to efficiency, cost and quality of experiment (QoE), and also in saving bandwidth at the backhand, transit and RAN links.
Wang [85] proposed an edge-assisted crowdcast framework called Deepcast which carries out intelligent decisions using massive amount of viewers and network real-time data to aid personalized QoE with reduced system cost. To deal with the high computation complexity involved in this situation, the authors designed a data-driven DRL based solution capable of learning the most appropriate approaches to schedule and transcode users request automatically. The proposed framework was evaluated with a trace-driven experiment using various real-world datasets. The result revealed that the proposed framework and its DRL-based solution was superior.

6.1.4. Semantics & Data Mining

Semantic information has been a critical link in connecting human intelligence to media data as well as big data in IoT contexts. As the IoT is becoming more social, computing based on crowdsourcing has emerged as a significant research subject in a variety of fields [86]. Many crowdsourcing frameworks have been created to investigate and address important issues in semantic information processing.

Guo [87] proposed a crowdsourcing scheme for semantic fusion and studied the critical challenges of processing semantic information of crowdsourced socialized IoT data. Social media data semantic fusion is confronted with new problems such as Source heterogeneity, Storage heterogeneity, multimodal expression and dissemination socialization. In order to adjust to the latest prerequisites of social media semantic fusion in crowdsourcing environment, the authors designed an efficient fusion strategy for multimodal semantics together with a distributed scheme to wholly ensure sharing of semantic information as the social media files are distributed. The proposed framework is divided into three modules—semantic extraction, multimodal semantic fusion and semantic storage and distribution module. In the first stage, crowdsourcing users extract semantic information from heterogeneous media from various sources. They defined an extraction workflow that includes an interaction interface for the users. Tasks are suggested to appropriate people based on their semantic information. The system monitors extraction quality and modifies jobs as needed. Figure 32 illustrates the system framework.
In the next stage, semantic items are disseminated. To avoid duplicate information, multimodal semantic items are synthesized to a single format while also minimizing the number of dimensions. They next extract some intelligence from the semantic data and improve it to remove any incorrect semantic data. The semantic data merged in the second stage is saved and disseminated with regards to the media files in the third stage. Because semantic information is dynamic, the authors proposed a reading and writing mechanism for semantic updating. Finally, the authors created an HBase-based database [88] to save semantic data and ensure synchronization. An experimental evaluation of the proposed framework on a categorized and an uncategorized dataset, revealed that the algorithm performed better on categorized dataset than in uncategorized dataset.

Liu [131] proposed a topic modelling framework based on crowdsourcing to enable various persons to contribute their knowledge and then utilized the outcome of their interaction to deduce meaningful topics. Figure 33 shows the service marching architecture for IoT based on crowd-LDA. First, the proposed model extracts raw topics from raw description texts and then requests users to cluster topic words using their own cognition so as to reduce the sparseness problem. The authors designed a crowd clustering algorithm to resolve the weakness among users and create accurate word clusters. Consequently, the author presented a new modelling method that integrates word clusters to extract feature vectors from raw data description text. Furthermore, they built a service matching algorithm which utilized features to discover service. Experimental study of the proposed method using real-world dataset demonstrates its effectiveness in terms of signature extraction and service matching in IoT.
Rahman [87] designed a fog-cloud hybrid framework that enables energy efficiency among large ad-hoc crowd, comprising of large distributed IoT devices and social network within a smart city context. A framework based on fog computing is presented to incorporate IoT devices which interface with ad-hoc crowd to ensure energy efficiency. The aim of the proposed framework is to collect massive crowdsourced, social media and IoT spatio-temporal multimedia data and process them at the edge and cloud ends in real-time to infer phenomena and disseminate the outcomes among the large crowd using personal queries, as well as storing the queries and data load in a big data cloud repository. The results for each requester are generated by the proposed framework in a visualized and personalized fashion. The suggested architecture improves energy efficiency by easing the cloud of the need to be “always active” as majority of the processing is done at the client edge. Further, the suggested framework assists sustainability through the incorporation of an incentive model, effective distribution of IoT devices inside a smart city environment, and efficient support of crowd sensing and crowd sourcing. The framework was tested extensively during Hajj 2018 when a significant amount of pilgrims took advantage of the proposed system for smart city context-aware services. The authors in [89] proposed a different approach for improving energy efficiency for multimedia sensor data using low-memory video compression architectures.

6.2. Specific Applications

Crowdsourcing IoT has been adopted to provide solutions to problems in a number of fields critical to our society. This subsection reviews the utilization of Crowdsourcing IoT in research in these areas, which includes disaster, emergency and risk management, transportation and smart cities, ocean/marine and environmental, recruitment and workers, finance, budget and pricing, and health/medical.

6.2.1. Disaster, Emergency and Risk Management

Disaster is a prevalent phenomenon in our world today. Every country and human is vulnerable to natural or man-made disasters [90]. During such an emergency situation, gathering and analyzing accurate information in a timely manner is of great importance especially when lives and critical infrastructures are in danger. Crowdsourcing and IoT
paradigms have been employed to develop applications to respond to disaster and emergency situations [91].

Lui [92] designed a disaster rescue framework called Echo for searching missing people in a disaster situation. The framework employed edge servers (ECV) computing, communication and storage abilities to preprocess and analyze in real-time, crowdsourced disaster photos to generate more precise visuals that will immensely assist in conducting emergency response and rescue operation. In order to conserve bandwidth, Echo allowed for only the extraction and uploading of significantly associated photos to the cloud center for successive facial recognition. Furthermore, the author developed an adaptive photo detector to make optimal use of the valuable and unsteady communication bandwidth while still ensuring photo sensing precision and recall rate. Echo prototype was implemented, and its performance was evaluated using a real-world disaster photo set. Using the search for missing persons as an example, tests showed that Echo can assist with the real-time and effective search for missing persons.

Rauniyar [93] applied crowdsourcing based and fog computing model to detect disasters in actual-time as well as disseminate timely warning to ensure the safety of the public. A data offloading mechanism is employed for the model to ensure the transfer of IoT data related to disaster to the fog platform even when a direct connectivity is unavailable. The suggested model as well as its data offloading mechanism showed great effectiveness in detecting disaster in actual time, and in broadcasting timely reportage to the public when compared to traditional cloud computing-based models for detecting disaster situations.

Zhao [94] proposed a content-aware crowdsourcing method called COCO which takes advantage of edge computing to enable image sensing in real-time disaster environment. The framework takes advantage of mobile edge computing technologies to deliver dynamic computing resources close to the data source. To overcome the resource limitation on mobile devices, the proposed method filters crowdsourced pictures at the data source and transmits only images that include relevant items to the application. To maximize the upload speed under varied bandwidth conditions, the authors discovered the best processing speed of the computer vision detectors under a given bandwidth and presented an adaptive technique to make the speed of processing optimal with varying bandwidth. A coco prototype was implemented and its performance was evaluated with a real-world image set. The result of the experiment showed that COCO can regulate the content of photos sent to the cloud as well as manage information loss within a certain limit. Furthermore, as compared to methods with adaptive technique, the adaptive technique fast-tracks image uploading by 21% to 35%.

Garcia [95] designed RAMBLE, an architecture for geo-referenced content sharing in territories with limited communications infrastructure such as those seen during rescue operation after a natural disaster. RAMBLE employed mobile edge-clouds, network of neighboring mobile devices, as well as light-weight cloudlets that service a small geographical region. The authors designed an Android app which rambling users utilized to create geo-referenced materials (such as sensor reading, video, text messages, or photos) and distributed them opportunistically to neighboring devices, cloudlets, or even cloud servers, as enabled by irregular wireless connections. Every device equipped with RAMBLE generated information, used up information that it showed interest in to its neighbors, and acted as other devices’ cache in an opportunistic manner. A real-world experiment of RAMBLE was carried out at Porto’s Botanical Garden and included a crowdsourced information collecting scenario with participants having Android smartphones, as well as deployed cloudlets. The result of the experiment revealed that mobile devices equipped with RAMBLE and modest cloudlets were effective in intelligence gathering.

Dubey [96] designed a theoretical integrated framework that can assist disaster relief staff in improving their coordinaton by utilizing relevant information generated from a complete crowdsourcing framework. The author adopted the case research decision tree
of [97] in their study and recognized that training and development, as well as leadership styles, make a substantial impact on the volume, variety, velocity, validity and value of data. They further proposed that IoT can be most effectively used if trust can be built between IoT and humans. Their analysis also showed that reducing behavioral ambiguity among partners will aid the development of successful coordination. However, the integrated framework (CS-10) was not empirically tested.

Raja and Thomas [98] proposed SAFER, a monitoring system for disaster situations based on crowdsourcing and software defined fog computing. Their aim was to reduce service delay and increase quality of service (QOS), which could be achieved by lowering transmission and computational delays in disaster situations. SAFER therefore employed SDN to provide dynamic network path to reduce data transmission delay and also utilized optimized fog computation to reduce computation delay. Simulation evaluation of SAFER architecture using crowdsourced heterogeneous data demonstrated that data were processed in less time, allowing for the early discovery of disaster situations. This result could allow first responders to pull spontaneous actions, which would be impossible with traditional networks owing to delay.

Han [99] designed a crowdsourcing and IoT integration model using crowdsourced data collected from social media and RFID devices, to improve disaster response. In addition, the authors adopted the SIR model to further build their model.

6.2.2. Transportation & Smart Cities

One significant IIoT application scenario is smart cities, which attempt to improve public services in urban areas and address urbanization issues including traffic congestion, environmental pollution, and energy consumption. The major problem in smart cities is divided into two parts. The first is how to successfully gather and capture huge volumes of dynamic data in ubiquitous settings. Another issue is the approach to evaluate this multisource diverse spatio-temporal data and then accurately create data profile for certain activities. As an effective and critical IoT technique, crowdsourcing focuses on connecting a wide range of mobile devices equipped with various sensing, actuation, and computational abilities to the wireless network, thereby supporting decentralized pervasive applications and services in a smart city environment [100].

Kong [101] proposed a smart city shared bus profiling framework using diverse mobile crowdsourced data. First, the authors used a shared bus application (APP) for mobile phones called Futurefleet to collect traveler’s order data as well as driver’s GPS data. Second, once the shared bus data has been obtained, the critical challenge is determining an approach to study and comprehend the requirements for travel of residents from different perspectives and subsequently creating precise travelers profiles. To address the challenge, the authors presented a travel profiling (TP) to set out resident travel, which they further refined into loss tolerance, delay tolerance, seat utilization rate (SUR), waiting time and journey time. Lastly, a multi-constraint evolution algorithm (MCEA) was designed to improve the routes based on the TP. The experimental results showed that TRProfiling performed excellently in terms of meeting traveler’s requirements.

Mathur [102] designed a mobile system called ParkNet that used designated vehicles to obtain information about parking space occupancy. Each vehicle used for the ParkNet system was provisioned with a GPS receiver and an ultrasonic range finder attached to the passenger’s side to ascertain parking space occupancy. A central server was used to aggregate data and created a real-time map of available parking spaces, applying fingerprinting method to realize better location accuracy. The system could provide parking availability information to parking search queries made by clients. A system test of ParkNet demonstrated a parking spot count accuracy of 95% and occupancy maps accuracy of 90% based on 300 km of road-side parking data gathered over two months by San Francisco taxicabs. The result also revealed that ParkNet provided enough coverage as well as cost effectiveness by an estimated ratio of 10–15 if implemented in city taxi cabs.
compared to a network of sensors with dedicated sensors at different parking spaces. Figure 34 shows the parking space framework using mobile sensors.

![Sensed scene](image1.png) ![ParkNet output](image2.png) ![Applications](image3.png)

**Figure 34.** Parking space framework using mobile sensors [102].

Bousios [103] designed two complementary mobile applications called CityCare and CityCareW that aimed to enhance citizen’s quality of life in smart city via crowdsourcing and participatory sensing principles. CityCare enables citizens to report concerns relating to the environment in which they reside, and also receive notifications about the progress in processing their complaints. On the other hand, CityCareW was designed for use by municipal personnel and allows them to organize, prioritize and process citizens’ request. Aside from easing information sharing between citizens and services, the two apps focused on the effective administration and processing of both citizens and services, the two apps focused on the effective administration and processing of both citizens and services to the benefit of the smart city. A trial assessment revealed that the two apps are highly user-friendly and pleased their end-user to a large extent.

Olariu [104] proposed a vehicular crowdsourcing framework named VACCS for congestion control support in smart cities and developed its theoretical basis. The author suggested that the in-built computational resources (together with smartphones and other IoT devices) contained inside vehicles trapped in city traffic can be harnessed to aid transportation authorities with dispersing congestion via large-scale signal re-timing. VACCS will support transportation authorities to improve vehicular movement at the level of smart city by dynamically re-timing signals, using vehicular network probe data, instead of only providing signals re-timing at the corridor level. VACCS will make traffic lights more sensitive to real situations instead of being reliant on historical volume counts. By designing timing plans that adapt to actual traffic circumstances, general vehicular movement will be boosted, and carbon emissions level would be decreased. Financial implications associated with congestion, with regard to high amount of fuel usage as well as the number of missed productivity hours will be minimized.

Montori [105] proposed a collaborative IoT architecture named SenSquare that integrates disparate data sources and makes them accessible to end user through an open interface. Figure 35 illustrates the architecture of Sensquare Machine learning methods used to classify and homogenize heterogeneous data streams, which are often not well-formed. It then provides the users the option of creating their own services using a web interface or a mobile application. A variety of service templates are available to the end user, who decides to utilize them or begin afresh in developing his or her own template. As a result, such a service is tied to a specific domain for each user, who will eventually be alerted of developments about it. The data required to satisfy the service requirement
are acquired from reputable sources if accessible and other untrusted sources. Any template service developed by one user for a particular area could be extended for use in another area using appropriate data. Finally, user-created services are made available to the whole community, allowing any member to make use of them or expand on them.

![Diagram of Sensquare architecture](image)

**Figure 35.** Architecture of Sensquare [105].

Yavari [106] developed IoT-based contextualization approaches that take into account the whole spectrum of data generated in smart cities and utilize such data to give hyper-personalized information to each user. Their proposed contextualization solution was exemplified in a smart parking spot recommender application/service. The experimental evaluation of their system demonstrates that contextualization of IoT data decreases query durations for IoT services (for example smart parking spot recommender) by more than three times when compared to a context-agnostic query.

6.2.3. Ocean/Marine & Environmental

In the smart ocean, the Internet of Things (IoT) has grown quickly. It can monitor and manage marine complex data by integrating data and information generated from various mobile maritime terminals (including buoys, boats, cargo, and offshore platforms). Clearly, IoT for the smart ocean presents a viable way to solve those pressing issues by controlling and exploiting ocean data [107]. Mobile crowdsourcing is an efficient methodology for collecting IoT data and has been employed by researchers and industries in developing solutions to smart ocean problems.

Yu [52] constructed a crowdsourcing privacy protection scheme technique to solve the issue participants’ privacy threats in a crowdsourced-IoT network for the smart ocean, using multi-authority ciphertext-policy attribute-based encryption (CSMA-ABE). The proposed system model comprises of requesters, semi-trusted third party, workers, a crowdsourcing platform, and a set of attribute authorities. The authors applied an independent key component distribution method via multiple authorities, taking into account the cooperation and independence of the various attributes, to enable the crowdsourcing platform to successfully disseminate security responsibility. They also introduced into the platform the concept of partial decryption in order to decrease mobile users’ computation cost. Furthermore, they completed the attribute revocation method which ensured the authenticity of subsets of workers’ attributes as well as the task search function, which allowed employees to swiftly perform searching jobs on the fly for the crowdsourcing platform. Simulation results demonstrated that their proposed method outperforms other comparable schemes in terms of lowering the time of consumption.
Luccio [108] presented a framework called DYNAMO (Distributed leisure Yachts sensor Network for Atmosphere and Marine Observation) for collecting coastal data generated by sensors and devices installed in maritime machinery in a crowdsourcing manner, in accordance with the Internet of Floating Things (IoFT) paradigm for environmental model simulation. The aim of the authors is to illustrate how the IoFT, as well as data on coastal marine life collected from instruments, in-built vessels, and sensors may be used to enhance the numerical prediction of the diffusion and transportation of pollutants for the quality evaluation of marine farm’s products. As a result, utilizing the crowdsourced seabed depth data from the Bay of Pozzuoli in Naples, Italy with their proposed system as a software environment test-bed, they demonstrated the enhancements made in a practical implementation. The bathymetry data were utilized to generate an enhanced form of the numerical representation of the seafloor which the models employed to improve the general outcome for estimating pollutant distribution details.

Huang [109] researched on the real-time and fine-grained air quality at the city scale using crowdsourced vehicles and their on-board sensors which substantially enhanced the sensing system’s practicality. The idea behind their study was that air component concentration inside a vehicle is extremely comparable to that of its surrounding environment when the vehicles are open due to the reality that air will interchange between the inside and outside of the car through the open window. As a result, they introduced an intelligent algorithm to identify vehicle air exchange status, and the extracted pollutant concentrations when the concentration trend converges after the windows are opened. The sensed convergent value is referred to as the comparable degree of air quality in the surrounding environment. Streams of real-time air quality data gathered from around the city were analyzed in a data center using their IoT cloud platform and used to produce a fine-grained city air quality level map. The suggested method was feasible for gathering air quality data in city areas, according to the algorithm’s evaluation and performance when compared to ground truth.

Montella [110] presented a framework for processing sensor data (GPS location and depth sounder) generated by leisure boat to create and update a comprehensive 3D seabed (bathymetry) map. The data collection was executed through DYNAMO, an IoT ecosystem, while the cloud hosted FACE-IT Galaxy workflow engine was used to organize and combine the data gathered through DYNAMO. The system employed CUDA-accelerated algorithms on virtual machine (VM) instances that support NVIDIA CUDA to process the highly voluminous data. GVirtus was used to virtualize the CUDA calls, allowing the interpolation method to run on ordinary VMs, and the CUDA-enhanced algorithms to run on GPU-enabled VMs when needed in order to minimize cost. The result of the experiment indicated that the system successfully collected and processed data and can readily scale to handle bigger inputs at moderate cost.

Kodali and Sahu [111] constructed a low-cost weather monitoring system that obtains weather conditions for any place from a cloud database management system and displays the results on an OLED display. The suggested system employed an ESP8266-EX microcontroller-based Wemos D1 board that was built on the Arduino platform to fetch the data from the cloud.

6.2.4. Recruitment & Workers

As an emerging intelligent data sensing model used in a wide-range of sensor applications for instance in the industrial IoT, mobile crowd sensing (MCS) allocates industrial sensing jobs to workers in order to collect and share data, creating a promising future for developing a robust industrial system and enhancing industrial efficiency (IIoT). The concept of MCS has been utilized to solve several challenges of industrial relations.

Hamrouni [112] proposed a low-complexity collaborative Mobile Crowdsourcing (CMCS) team recruitment strategy for forming and hiring a group of socially linked employees with the necessary abilities to complete a CMCS job. Figure 36 shows the typical CMCS platform. The proposed strategy inspired by Swarm intelligence (genetic algorithm)
allowed for project matching and virtual team building based on four key fuzzy-logic-based criteria—degree of knowledge, recruitment cost, strength of social relationship, and platform confidence level. The strategy aimed at forming a team who are not just talented but also socially connected. The results of the proposed system evaluation using Monte Carlo Simulations demonstrated that the proposed CMCS recruitment algorithm realized loss outcomes to the baseline optimal Integer Linear Program (ILP) strategy while saving notable computational time and outperformed the Particle Swarm Optimization (PSO).

![Figure 36. Typical CMCS platform [112].](image)

Hamadi [113] illustrated an automated crowdsourcing method to coordinate collaboration between financial advisers and investors. The system provides a recruiting platform to connect financial advisers to investors based on their profile and attributes, as well as the prior activities of their peers, without interfering with their private portfolio optimization strategies. An unsupervised approach was applied to the crowdsourcing database to group advisers with a high degree of similarity into clusters, reducing the search space. Following that, a machine learning regression model was trained and estimated the matching score that an investor could get if he or she selected a specific investor. In addition, a double-phased many-to-many maximum weight matching method was utilized to convert the scores to bipartite graph weights to find an appropriate investor-advisor pair. The performance evaluation of the proposed approach through simulation revealed a higher profit for both the investors and financial advisors compared to other conventional approaches.

Lu [114] researched the challenges of least workers selection to help big MCS systems execute sensing tasks more effectively and achieve specified coverage while satisfying certain restrictions. They introduced a Many-objective worker selection technique (MaOWS) to optimize the profit of the workers and reduce payment from the requester, while increasing platform profit. They also presented a many-objective optimization technique utilizing enhanced differential evolution algorithm (EDEA) for a quick extraction of possible worker selection viable result scope while creating the required tradeoff between the objectives. The outcomes of the experiments demonstrated the proposed technique’s effectiveness on real-life data.

Khanfor [115] suggested a Spatial Mobile Crowdsourcing (SMCS) recruitment method utilizing automated Social Internet of Things (SIoT) service discovery for the se-
lection of reliable employees that satisfy the requirements of the requester. Figure 37 illustrates the SMCS high-level architecture. The system provided an employee filtering stage that applied the Louvain community detection algorithm (CD) on the SIoT relation graphs to reduce the employee’s scope of the search for a selection of potentially reliable candidates. It also employed an ILP in the selection process to identify the final selection of devices/workers. The ILP improved employee efficiency metric taking into account the degree of skill/specs, cost of recruitment, and level of trustworthiness of IoT devices recruited. Experimental analysis of the suggested CD-ILP algorithm performance based on real-life data proved to be superior compared to existing stochastic algorithm.

![SMCS high-level architecture](image)

**Figure 37.** SMCS high-level architecture [115].

Hamrouni [116] formulated a SMCS recruitment framework that jointly optimize the matching and scheduling of tasks and workers. First, the matching quality is optimized as a function of weighted scores of various recruitment criteria using a Mixed Integer Linear Program (MILP) while defining the trajectory of all picked workers doing a job. The system further presented a heuristic SMCS technique to allow for solving a weighted bipartite graph problem to achieve a sub-optimal matching and recruitment solution. The outcomes of the simulation demonstrated that the suggested SMCS algorithm performed better than existing greedy recruitment techniques.

Yang [117] introduced fog platform as a service to detect resourceful employees via learning their performance information. Employees past performance data are recorded at the local fog server and valuable employees are detected and selected to execute the tasks based on this data. Considering balancing exploration and exploitation, the author presented an online algorithm which promotes employees that have not been fully explored while optimizing the long-term platform utility with a limited budget. Simulations conducted on real-world datasets proved that the algorithm was effective.

6.2.5. Finance, Budget & Pricing

Crowdsourcing is viewed as a viable technique in the IoT paradigm for offering effective sensing, computing, and processing services to a specific job produced by consumers, utilizing the power of the crowd in an efficient manner. Within the crowdsourcing-based IoT paradigm, to appropriately price tasks of customers is critical for platforms to optimize their utility. Crowdsourcing-based IoT has been leveraged to develop solutions that improve efficiency, performance, quality of service (QoS), as well as customer satisfaction.

Zhang [118] constructed a blockchain-enabled crowdsourcing platform that combined resource lending function and task assignment through a pair of customized smart
contracts to ensure the security, as well as dependability of interacting users while overcoming the limitations of centralization. Figure 38 shows the blockchain-based hierarchical crowdsourcing platform. The platform is comprised of two smart contracts (task matching contract and resource marketing contract), workers, customers, and resource providers. Using economic modelling methodologies and heterogeneous agent theory, the authors developed a dynamic hierarchical pricing system to maintain stability of financial transaction within the blockchain-enabled market. It also ensures that the job complexity received is matched with the appropriate worker’s competence. The system is therefore applied to dynamically divide the market into several levels making use of customer compensation offer and the market worth of the employees’ resources. Additional restrictions are recommended in the market to avoid potentially harmful training behavior by resource lending workers. Extensive experimental outcomes revealed the platform’s feasibility and correctness, and proved that the pricing mechanism could sustain market stability.

![Figure 38. Blockchain-based hierarchical crowdsourcing platform [118].](image)

Using VCG and quality of all-pay, Hao [119] presented a collaborative crowdsourcing reverse pricing system (VCG-QCP). The authors examined a real crowdsourcing scenario using the VCG method and introduced the idea of quality all-pay to assess the quality of job of employees that execute the tasks. The price algorithm was then built based on a generic reverse pricing model through mathematical modeling. Simulations indicated that the suggested method could realize greater algorithm effectiveness, improved job completion quality, suitable equal advantages among employers and employees, as well as ensure the veracity of employees’ bids.

Haghhighatdoost [120] presented a method based on crowdsourcing for determining fair pricing of wireless service in the IoT. The authors focused on an oligopoly where the regulation dynamically establishes a maximum permitted cost of service to avoid anti-trust conduct and unfair policy of pricing services. As a result, they suggested a three-tier pricing model in which the regulator, wireless network providers, as well as clients, are all players and considers client’s preferences to find fair service pricing slightly above each network provider’s marginal costs. Their results demonstrated that the proposed model was not susceptible to complicity and could, at most, converge when one network provider declared the fair price.

Li [121] investigated the utility maximization of crowdsourcing platforms by determining the best pricing policy for requested jobs. They formulated the pricing strategy as a finite sum optimization problem where the nodes attempt to realize a general agreement pertaining to the pricing policy of every job. The authors further employed the stochastic gradient descent (SGD) type algorithm non-convex primal-dual splitting with exact min-
imization (NESTT-E) to resolve the optimization problem in a distributed and asynchronous manner. The numerical outcomes revealed that NESTT-E achieved better effectiveness compared to synchronous ADMM and traditional SGD in terms of greater number of active nodes.

Chi [122] applied game theory to illustrate a long-term incentive strategy for motivating participating workers to supply continuous and long-term rich data for MCS systems. The proposed method viewed cooperation among workers and requester as a recurring game and then generates a long-term incentive using past data and a discount factor. Furthermore, the evolutionary game theory is utilized together with the Wright–Fisher model to assess the development of participants’ strategies. The result of simulation experiment proved that the presented mechanism provided participants in MCS systems with a long-term incentive effect.

Singh [123] designed a truthful budget feasibility strategy for the task allocation problem (TUBE-TAP) that provides accurate information regarding the quality of IoT devices. Simulation results demonstrated that TUBE-TAP met various economic characteristics, including truthfulness and budget feasibility.

6.2.6. Health/Medical

The IoT technology is also being leveraged in the healthcare monitoring system to provide excellent emergency assistance to patients. It is also utilized as an E-health application in a variety of areas, including early identification of medical problems, emergency alerts as well as computer-assisted rehabilitation. Smartphones have become an important part of people’s everyday lives, and they are linked with sensors to monitor the subject’s health [124]. Crowdsourcing-based IoT paradigm enables the collection of large amounts of health-related data from patients, hospital equipment, environment, and so on, and is used in several applications for quality healthcare delivery.

Using city air quality data generated via meteorological sites, IoT sensing, MCS, and user’s body signal, Chen [125] introduced a city healthcare system called UH-BigDataSys to offer urban residents exhaustive and correct air quality services and real-time physiological index monitoring. An air quality indicator (AQI) sensing is implemented at three levels of the network, and a multi-dimensional AQI (M-AQI) big data integration method is proposed to integrate data at the three levels. Crowdsourced AQI data are gathered at the first level, while AQI data fusion is considered at the edge-cloud level. Lastly, the AQI data were uploaded to a remote cloud or meteorological super computing platform. Furthermore, the physiological big data and M-AQI big data are processed and analyzed by the system to offer health advisory support to users on respiratory diseases, sleep quality control, outdoor travel, among other things. Figure 39 shows the system architecture of UH-BigDataSys.
Figure 39. System architecture of UH-BigDataSys [125].

Fernandez-Carames [126] designed and implemented an IoT-enabled system to improve commercial Continuous Glucose Monitor (CGM) for remote monitoring of patients and notifying them of potentially dangerous situations. The proposed system employed smartphones to gather blood glucose data from CGMs and transmit them to either a distributed fog or remote cloud computing nodes. The system incorporated a decentralized storage system for receiving, processing and storing of allocated data to ensure a reliable, trustworthy and cyber secure exchange of data with medical doctors, scientists and caretakers. The system also incorporated a digital crypto currency called GlucoCoin that utilizes blockchain to execute smart contracts so as to automate CGM sensor purchases as well as reward users who contribute data to the system. Experimental results showed the effectiveness of the system.

Servadei [127] designed and prototyped Medical Monkeys, a collaborative crowdsourcing web application for annotating medical images. The platform comprised of medical students, researchers, online gamers, and patients under a virtual and collective advantageous collaboration for enhancing machine learning algorithms on medical images. The result of the system testing promised improved automatic segmentation and detection algorithms, as well as improved clinical decision support systems and a reduction in human-based inaccuracy in diagnostic and therapeutic evaluation.

Zaman [128] leveraged the concept of crowdsourcing and IoT-based network to design a medical emergency system to improve existing e-Health monitoring system. They created a mathematical model to represent the uplink and downlink scenarios and evaluated the proposed system. Moreover, a mobile application was developed to practically simulate the system in the Bangladesh context.

Different application areas require the use of different kinds of sensing data, and also encounter many domain constraints in gathering data. Researchers should acquaint themselves of application domains’ constraints to sensing data as this would guide in choosing the appropriate techniques, sensing devices, communication network, and so on, to deploy in any implementation.

7. Conclusions

This paper has given a comprehensive survey of the emerging research area of crowdsourcing Internet of Things. First, we reviewed different architectures proposed for crowdsourcing systems. Having identified that no architecture was specifically designed for crowdsourcing IoT, we proposed an architecture termed Crowd-IoT to effectively capture the structure and components, and the interactions that exist within crowdsourcing
IoT systems. The security and privacy of individuals and devices that participate in crowdsourcing tasks are considered highly vital, in order to motivate confidence in the crowd to participate in crowdsourcing tasks. As a result, we presented the current security and privacy solutions developed for crowdsourcing systems. This paper further discussed the role of edge computing in crowdsourcing IoT and the techniques mitigating its limitations to achieve fast computation, effective resource sharing and data caching, and energy consumption reduction. Finally, we presented studies that employed crowdsourcing and IoT in several application areas.

Overall, this survey paper has given a theoretical discussion from several different perspectives (architectures for Crowd-IoT, trustworthy, privacy and security for Crowd-IoT, resources, sharing, storage and energy considerations for Crowd-IoT, and applications for Crowd-IoT) to motivate and stimulate continuing developments and innovations from industry and the research community towards Crowdsourcing IoT paradigm. The authors have also discussed preferences in the survey Sections 3–6 for researchers who would imagine appropriate implementation.

The limitation of this work is the lack of a practical implementation to test the effectiveness of our proposed architecture. Our future work will focus on implementing our proposed architecture in some application areas. Specifically, we intend to implement a more efficient smart city waste transport system and an emergency information management system. The results of our implementations would enable us to evaluate our architecture and guide us on whether to affect any improvements.

Author Contributions: Conceptualization, K.L.M.A. and J.S.K.P.; writing—original draft preparation, K.L.M.A., J.S.K.P. and E.N.; writing—review and editing, J.S.K.P., K.L.M.A. and E.N.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References


