International Journal of Applied Science; Vol. 8, No. 2; 2025
ISSN 2576-7240 E-ISSN 2576-7259
https://doi.org/10.30560/ijas.v8n2p103

Survey on Privacy Preserving in Crowd Sensing

Tian Maoze[1]

[1] School of Information, Guizhou University of Finance and Economics, China

Correspondence: Tian Maoze, School of Information, Guizhou University of Finance and Economics, China.

Received: May 12, 2025; Accepted: May 29, 2025; Published: May 30, 2025

Construction Project of Key Laboratory of Blockchain and Fintech in Colleges and Universities of Guizhou Province (Qianjiao Technology [2023]014).

Abstract

As an emerging sensing technology, crowd sensing has gained wide attention in many application fields and is developing rapidly. However, with the popularization of crowd participation in sensing tasks, the risk of user privacy leakage is also increasing, which becomes an important problem to be solved urgently. When users participate in sensing tasks, they need to upload personal information or sensor data, which often contains sensitive information. Without effective privacy protection measures, user privacy may be leaked or abused. The core goal of privacy protection is to ensure that users' private information will not be leaked when they participate in the task. This paper analyzes the related research progress of privacy protection in the field of crowd sensing, and summarizes the main challenges currently faced.

Keywords: crowd sensing, data quality, privacy protection, data anonymization, encryption techniques

1. Introduction

As an emerging data collection method, crowd sensing has become an important implementation form of the Internet of Things [1]. Relying on widely participating smart mobile devices, this model constructs a distributed, interactive and highly participatory sensing network. These devices are able to collect and transmit information about the surrounding environment in real time, such as temperature, humidity, air quality, and location, providing rich real-time data. In this way, crowd sensing can not only efficiently collect large-scale data, but also give full play to the advantages of crowd intelligence in the process of data quality control, information processing and analysis.

Crowd sensing systems rely on data provided by a large number of participating users, which usually come from different sensing devices and users, reflecting the characteristics of distributed data collection. Although crowd sensing can achieve wide information coverage and instant data collection, users' privacy (such as location, behavior, identity information, etc.) may be exposed in the process of data collection, transmission and analysis. Therefore, privacy protection has become a core challenge in crowd sensing systems. By combining technologies from multiple fields, such as cryptography, distributed computing, and policy design, crowd sensing systems can maximize data availability and social value under the premise of protecting user privacy. Therefore, the discussion on privacy protection and related research issues in crowd sensing has gradually increased, and many research results with important value have been produced. This paper aims to review the research status at home and abroad in this field and provide useful reference for related researchers.

2. Crowd Sensing

The architecture of the crowd sensing system is shown in Figure 1, which is composed of three parts: task publisher, service provider and participating users. Each role assumes different functions and responsibilities in the system. Specifically, the task publisher is responsible for defining and submitting the sensing tasks that need to be collected. These tasks can involve environmental monitoring, location tracking, behavior analysis and other fields, and the purpose is to meet the data requirements in specific application scenarios. Task publishers need to accurately describe task objectives and requirements to ensure that tasks can be effectively delivered to service providers.

Service provider plays a bridge role in the crowd sensing system, which is mainly responsible for distributing the sensing task submitted by the task publisher to the appropriate participating users according to certain rules and

algorithms. Rules may select appropriate users to participate in data collection based on a variety of factors such as geographical location, time requirements, and device capabilities. At the same time, the service provider is also responsible for processing and integrating the data collected from the participating users to ensure the accuracy, integrity and consistency of the data. In addition, service providers need to perform management and coordination functions, including task scheduling, data quality control, user feedback processing, etc., to ensure the smooth operation of the system and the validity of the data.

The participating users are the data providers in the crowd sensing system. After receiving the sensing tasks distributed by the service provider, the users collect data according to the task requirements. Users collect environmental data or personal behavior data through smart devices such as smartphones, sensors, wearable devices, etc., and upload these data to the system server. Participating users not only need to ensure the quality and accuracy of data collection, but also need to comply with the relevant regulations of privacy protection to ensure that the uploaded data does not disclose sensitive information.

The server plays a crucial role in the whole system, it is responsible for the subsequent processing and analysis of the data uploaded by the participating users. The process includes data cleaning, data fusion, data analysis and result feedback. By analyzing the data uploaded by a large number of users, the server can extract valuable information and provide accurate sensing results for the task publisher. At the same time, the server also needs to deal with data storage and security issues to ensure the stability of the system and data security.

In general, through the collaborative work of task publishers, service providers and participating users, the architecture of crowd sensing system can efficiently achieve a wide range of sensing data collection and processing, and provide reliable data support for practical applications. The coordination and cooperation between each component is the key factor to ensure the successful implementation of the system.

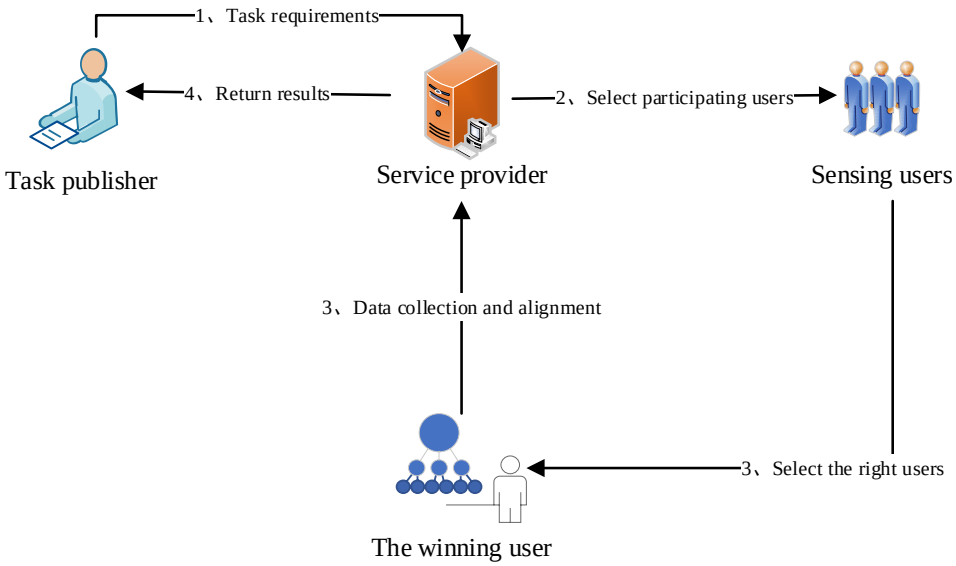

Figure 1. Crowd Sensing System Structure

3. Research Content of Privacy Protection in Crowd Sensing

The research content of privacy protection for crowd sensing covers all aspects from data collection, transmission, storage to processing, aiming to ensure that users' personal information and sensitive data are effectively protected in crowd sensing systems. With the rapid development of information technology and big data applications, the challenge of privacy protection has become increasingly serious. Especially in the complex environment where multi-device, multi-user and multi-source data are intertwined, how to effectively avoid personal privacy leakage, ensure data security and comply with the corresponding laws and regulations has become an important research topic in the field of crowd sensing. The following is a detailed expansion of the main research contents of privacy protection in crowd sensing:

(1) Data encryption and anonymization:

The data sources involved in crowd sensing systems are diverse and highly sensitive. Therefore, how to ensure that these data are not accessed or tampered by unauthorized third parties in the process of collection, transmission and storage has become one of the core issues of privacy protection. Data encryption and anonymization techniques provide effective solutions.

(2) Privacy preserving data sharing and aggregation:

In crowd sensing systems, data usually comes from multiple users and devices. How to realize data sharing and aggregation under the premise of protecting privacy is an important research direction of privacy protection. With the introduction of multi-source data, how to effectively fuse and analyze data without exposing personal privacy has become a complex and key technical challenge.

(3) Privacy-preserving task allocation and data processing

Crowd sensing systems are not only about collecting and processing data, but also about how to properly allocate tasks and data processing processes to avoid leaking privacy. Privacy-preserving strategy of task allocation and data processing is the key to ensure system efficiency and security.

In summary, the research on privacy protection in crowd sensing involves many aspects from data encryption and anonymization to task allocation, data sharing and compliance. With the continuous progress of technology, new privacy protection methods and strategies are constantly emerging, especially with the support of emerging technologies such as multi-source data, decentralized storage and blockchain technology. Privacy protection in crowd sensing systems faces new opportunities and challenges.

4. Research Status of Privacy Preserving in Crowd Sensing

In crowd sensing systems, privacy protection has increasingly become a key research topic. Existing researches mainly focus on protecting users' location, identity and data privacy. Wang et al [1] proposed a privacy protection method based on k-anonymity, which effectively protects the privacy of user location and data by dividing the data into multiple equivalence classes and iteratively uploading within the equivalence classes. Zou et al [2] proposed a decentralized crowd sensing system, CrowdHB, which combines blockchain technology and uses smart contracts to realize the protection of location privacy and enhance the security of data processing. Wu et al [3] ensure the integrity of sensing data and the anonymity of participating users by evaluating the credibility of nodes, selecting reliable transmission paths, using data slicing technology, and combining secret sharing and anonymity strategies. In addition, there are some studies that try to combine privacy protection with practical application scenarios such as incentive mechanisms and task allocation. For example, Ding et al [4] introduced privacy protection in the process of task allocation, adopted homomorphic encryption technology to protect the user's location privacy, and completed task allocation in the ciphertext domain through the cooperation of edge servers. Yan et al [5] proposed an incentive mechanism based on privacy protection, which allocated rewards according to data reliability, combined editable signature and hash function to verify the reliability of data source, so as to effectively protect users' data privacy.

In order to deal with the problem that sensory data may expose user identity and data privacy during transmission, more and more researchers have proposed solutions that use encryption algorithms to encrypt sensory data. The core goal of these encryption algorithms is to prevent data from being stolen or tampered with by malicious attackers while ensuring the security and privacy protection of data transmission. Even if the encrypted data is exposed during transmission, the privacy information of sensing participants can still be effectively protected, so as to ensure the security of the whole system. For example, literature [6] proposed a privacy protection scheme for crowd sensing based on blockchain, which combined zk-SNARK proof and smart contract technology to verify the authenticity and integrity of data through zero-knowledge proof, so as to ensure that the privacy information of any participant would not be leaked during the interaction process. This method can complete the transaction verification without exposing any sensitive data, so as to effectively enhance the ability of privacy protection.

In addition, literature [7] designed a blockchain-based crowd sensing system BSLF, which requires all users to use private keys for authentication and combines Paillier encryption system to protect users' sensitive data. In this scheme, the user's identity and sensing data are encrypted and stored, and only the legitimate user can decrypt and access the relevant data. As a homomorphic encryption technique, Paillier encryption algorithm is able to perform computations on encrypted data without decrypting the data itself, thus protecting user privacy and maintaining the overall security of the system. In addition, literature [8] proposed a privacy-preserving crowdsensing system based on blockchain, which successfully realized the anonymization of transaction records by introducing private chain technology, and effectively solved the problem of user location privacy leakage. In this scheme, all

transaction records are stored and verified through the private chain, which not only ensures privacy, but also reduces the risk of privacy leakage that may be encountered in the traditional public chain.

In summary, combining blockchain technology with crowd sensing system and encrypting sensing data by cryptographic means can effectively realize decentralized management and further protect data privacy and identity security. This combination not only overcomes the security problems in the traditional centralized system, but also provides a more flexible and scalable privacy protection scheme. However, although the existing schemes have achieved certain results in privacy protection, they generally ignore the high storage cost introduced by blockchain technology and the efficiency bottleneck caused by a large number of users verifying signatures at the same time. As the scale of the sensing system increases, these problems may be further aggravated, seriously affecting the performance and response speed of the system. Therefore, how to ensure efficient processing of sensory data and protect user privacy while reducing storage overhead and improving system performance has become a key challenge in current research.

In the research field of stealth protection in task allocation, the existing task allocation methods can be widely divided into two categories: static task allocation and dynamic task allocation. Static task allocation is usually applicable to the situation where task requirements are relatively fixed and workers' abilities are uniform. Among them, WANG et al [9] deeply studied the problem of perceived quality of a single task, and proposed to define the optimization objective of task allocation by introducing a minimum perceived quality threshold, aiming to maximize the system utility by optimizing the matching relationship between tasks and resources. This approach highlights the impact of perceived quality on task completion outcomes and provides a new perspective for the construction of task allocation models. At the same time, WU et al [10] proposed a weighted multi-objective particle swarm optimization algorithm in the edge computing environment to face the multi-objective task allocation problem. The algorithm aims to maximize the comprehensive utility of the platform and workers in the complex edge computing platform, and balances different needs in task allocation by optimizing the weights of multiple objective functions. LI et al [11] proposed two evolutionary algorithms for the multi-task allocation problem with strict time constraints. The core goal of these algorithms is to maximize the utility of the platform, and strive to improve the allocation efficiency while considering the task time limit. However, these static assignment methods often ignore the factors of dynamic changes of tasks and workers in the real environment, resulting in certain limitations in dealing with complex and dynamically changing task requirements.

Compared with the static task allocation, the dynamic task allocation method is more flexible and can respond to the changes of task requirements and worker status in real time. Research on dynamic task allocation can usually be divided into two forms: immediate allocation and batch allocation. TAO et al [12] conducted an in-depth analysis of the temporal and spatial attributes of workers and tasks, and proposed an off-line task allocation method based on ant colony optimization algorithm and a predictive algorithm for online scenarios. The common goal of both is to maximize the overall utility of the platform through a reasonable task allocation strategy in different scenarios. These methods highlight the importance of spatio-temporal attributes in task allocation and further improve the ability of the platform to handle dynamic task requirements. SONG et al [13] focused on the dynamic task allocation problem of multi-skilled workers, and proposed an online greedy algorithm to calculate the optimal matching between new workers and tasks in real time, which could flexibly adjust the task allocation scheme according to the skill changes of workers and the urgency of tasks. XIAO et al [14] focused on the cooperative task allocation problem with average completion time sensitivity, and proposed an online greedy algorithm based on task priority and the principle of earliest user accepting the task, aiming to reduce the completion time of tasks and improve the overall efficiency of the system. MIAO et al [15] proposed a probabilistic model to measure the quality of tasks, combined with the free-rider model to characterize the behavior of workers, and solved the quality maximization problem in online task allocation using an algorithm with polynomial time complexity. These dynamic task allocation strategies greatly enrich the research framework of task allocation, and provide a more flexible and efficient solution for task allocation in practical applications.

However, the existing real-time task allocation methods generally face the problems of low utility and high cost, which makes the scenarios of these methods in practical applications limited. Most batch allocation models still adopt a fixed allocation method, and determine the appropriate batch size through experiments, ignoring the uncertainty and flexibility in dynamic task environments. TO et al [16] proposed the equal batch allocation model, which transformed the task allocation problem into a binary matching problem in each batch, and solved it by the maximum flow algorithm, so as to optimize the efficiency of batch allocation. WANG et al [17] proposed an innovative model combining dynamic and static task allocation framework, by introducing the concept of delay time and using Q-learning to determine whether the current task allocation needs to be delayed, so as to improve the efficiency and adaptability of task allocation. By dynamically adjusting the time of task allocation, this method

avoids the waste of resources in the case of high load and improves the efficiency of the whole system. In general, although the current research has made some progress in task allocation optimization, how to balance the real-time performance and efficiency of task allocation in different scenarios is still a challenge to be solved.

From the existing research and practice results, although the crowd sensing technology has made some important progress, the existing research still faces many challenges and shortcomings. The following is an in-depth discussion of these shortcomings:

(1) Privacy protection and data security issues:

Crowd sensing systems rely on extensive user participation and involve the collection and transmission of large amounts of personal data. These data not only include the user's location information, but also may involve the user's health status, behavior habits and other sensitive information. How to achieve effective data sharing and utilization under the premise of ensuring data privacy has become a major problem in current research. Although a variety of privacy protection techniques have been proposed, such as data encryption and anonymization, there are still significant challenges in balancing privacy protection and data availability in large-scale and low-cost system environments. Especially in the scenario of multi-participant and large-scale data, how to establish an effective data security management system to prevent information leakage and abuse is still a bottleneck problem in the wide application of crowd sensing technology.

(2) Data Quality and Trust Issues:

Crowd sensing systems rely on large-scale data uploaded by users, however, these data often have the problem of unstable quality. Due to the heterogeneity of participants' devices, the complexity of the collection environment, and the improper operation of users, the uploaded data may be erroneous, missing, or forged, which greatly affects the reliability of the data. Current research mainly focuses on how to post-process or filter the data quality through algorithms, such as anomaly detection and data fusion. However, how to ensure the accuracy and consistency of data in a highly heterogeneous system, especially when facing heterogeneous data from different devices and sensor types, there is still a lack of effective standardization methods. In addition, how to improve participants' trust in data quality is also an urgent problem to be solved, and the lack of trust mechanism may lead users to distrust the system, thus affecting their engagement.

(3) Task allocation and resource optimization:

Task allocation in crowd sensing system is a complex optimization problem. Factors such as the geographical location of participants, device performance, and network bandwidth have a profound impact on the efficiency and effectiveness of task execution. Most of the current researches assume that the participants are uniformly distributed, ignoring the device heterogeneity and network volatility in practical applications, which makes many existing task allocation algorithms ineffective in the face of the actual complex environment. Therefore, how to design a task allocation strategy that can adapt to different environments, devices and network conditions is an important research topic in the field of crowd sensing. In addition, the resource optimization in crowd sensing system includes not only the allocation of tasks, but also the efficient utilization of resources such as energy and computing power. How to minimize the consumption of energy and computing resources while ensuring the quality of task completion is still a big problem in system design.

In summary, although crowd sensing, as an emerging technology, has shown broad application prospects in various fields, it still faces challenges in many aspects, such as privacy protection, data quality, resource optimization, real-time performance, and incentive mechanism. With the continuous development of technology and the deepening of research, it is expected that these problems will be gradually solved in the future, thus promoting the application of crowd sensing technology in a wider range of practical scenarios.

5. Prospect and Summary

With the rapid development and wide application of crowd sensing technology, privacy protection has become one of the core issues in this field. Crowd sensing systems essentially rely on a large number of sensors and smart devices to collect and process data in different environments, so the requirements for privacy protection are increasingly stringent. Future privacy protection research will continue to advance in the following directions to cope with the increasingly complex data privacy requirements and technical challenges:

(1) More efficient privacy-preserving algorithms:

In future crowd sensing systems, the efficiency of privacy protection algorithms will become a crucial research direction, especially in the face of large-scale data processing. Crowd sensing systems need to process data from a large number of devices and users in real time. Therefore, how to improve the efficiency of algorithm execution

and reduce the overhead of computing and storage while ensuring privacy will be the key to achieve efficient privacy protection. In order to meet the requirements of real-time, low cost and high efficiency of crowd sensing systems, privacy protection algorithms not only need to improve the execution speed, but also need to have higher adaptability, which can make flexible adjustments for different operation scenarios. In addition, future privacy protection algorithms will pay more attention to the balance between accuracy and privacy protection, so as to ensure that the utilization value of data can be improved as much as possible without affecting the security of user data. For example, by introducing new encryption algorithms, data anonymization techniques, differential privacy and other means, it can effectively improve the analysis value of data and system performance while protecting user privacy.

(2) Adaptive privacy protection mechanism:

As the application scenarios of crowd sensing technology are gradually diversified, the adaptability of privacy protection mechanism is also becoming more and more important. Crowd sensing systems need to flexibly adjust the privacy protection strategy and strength according to different application environments, data characteristics and user requirements. Therefore, future research will pay more attention to building intelligent and adaptive privacy protection mechanisms. These mechanisms can automatically select the most appropriate privacy protection technology according to the changes of specific scenarios, so as to effectively improve the flexibility and user experience of the system. For example, in different environments (such as urban surveillance, smart healthcare, intelligent transportation, etc.), the system will automatically adjust the privacy protection strategy according to the sensitivity of data, user privacy preferences, and legal and regulatory requirements to ensure the optimization of privacy protection effect.

(3) Cross-domain privacy protection:

With the rapid development of crowd sensing technology, the problem of cross-domain transmission and sharing of data has gradually appeared. Cross-domain privacy protection not only involves data transmission and sharing between platforms and devices, but also involves data security management in different network environments. How to effectively protect user privacy in cross-domain context has become an important topic in current research. For example, when the crowd sensing system interacts with data between multiple platforms or devices, how to ensure the security and privacy of data during transmission is an urgent technical challenge to be solved. With the rise of decentralized technologies such as blockchain, this problem may be solved. By providing a decentralized trust mechanism, blockchain technology can effectively ensure the secure transmission and sharing of data between multiple domains, and avoid the problems of data leakage and tampering that may be caused by a single centralized storage. Therefore, future privacy protection research may focus more on how to combine decentralized technologies such as blockchain with existing privacy protection mechanisms to build a more secure and transparent data exchange and protection system.

(4) Quantum computing and privacy protection:

The rapid development of quantum computing technology has brought great challenges and opportunities to the field of privacy protection. Traditional encryption algorithms (such as RSA, AES, etc.) may lose security in front of quantum computers, because quantum computing is able to effectively break current widely used encryption protocols. However, the development of quantum encryption technology provides new solutions for privacy protection. Quantum encryption uses the principles of quantum mechanics to provide highly secure data encryption protection without relying on traditional computing power. In the future, crowd sensing systems will likely incorporate quantum encryption technology to achieve more secure data transmission and processing. Especially in the context of processing massive data and high concurrent access, quantum encryption technology is expected to break through the bottleneck of current encryption technology and provide more robust security. In addition, the parallel processing capability of quantum computing may also have a profound impact on data processing in crowd sensing systems, pushing privacy protection technologies towards more efficient and intelligent directions.

In short, privacy protection in crowd sensing is still a hot field in current research. Although many privacy protection mechanisms have been proposed and applied, there are still many challenges in algorithm efficiency, system adaptability, and cross-domain security. With the continuous progress of technology, privacy protection mechanisms will become more intelligent and flexible in the future, and can better balance the contradiction between data protection and data utilization. In particular, with the continuous development of emerging technologies such as quantum computing and decentralized technology, privacy protection technology is expected to usher in a major breakthrough and further promote the application development of crowd sensing systems. Through continuous technological innovation, crowd sensing systems will better serve various practical application scenarios and bring wider value to society while protecting user privacy.

References

[1] Wang, T., Liu, Y., Jin, X., et al. (2018). Location and data privacy protection method based on k-anonymity in crowd sensing research. *Journal of Communications, 39*(A01), 170–178.

[2] Zou, S., Xi, J., Xu, G., et al. (2021). Crowdhb: A decentralized location privacy-preserving crowdsensing system based on a hybrid blockchain network. *IEEE Internet of Things Journal, 9*(16), 14803–14817. https://doi.org/10.1109/JIOT.2021.3088193

[3] Wu, D., Fan, L., Zhang, C., et al. (2018). Dynamical credibility assessment of privacy-preserving strategy for opportunistic mobile crowd sensing. *IEEE Access, 6*, 37430–37443. https://doi.org/10.1109/ACCESS.2018.2850534

[4] Ding, X., Lv, R., Pang, X., et al. (2022). Privacy-preserving task allocation for edge computing-based mobile crowdsensing. *Computers & Electrical Engineering, 97*, 107528. https://doi.org/10.1016/j.compeleceng.2021.107528

[5] Yan, X., Wang, Y., Zeng, B., et al. (2022). P2SIM: Privacy-preserving and source-reliable incentive mechanism for mobile crowdsensing. *IEEE Internet of Things Journal, 9*(24), 25424–25437. https://doi.org/10.1109/JIOT.2022.3198311

[6] Chatzopoulos, D., Gujar, S., & Faltings, B. (2018). Privacy preserving and cost optimal mobile crowdsensing using smart contracts on blockchain. *IEEE Internet of Things Journal, 10*(9), 442–450. https://doi.org/10.1109/JIOT.2018.2825383

[7] Wang, W., Yang, Y., & Yin, Z. (2022). BSIF: Blockchain-based secure, interactive, and fair mobile crowdsensing. *IEEE Journal on Selected Areas in Communications, 40*(12), 3452–3469. https://doi.org/10.1109/JSAC.2022.3192064

[8] Yang, M., Zhu, T., & Liang, K. (2018). A blockchain-based location privacy-preserving crowdsensing system. *Future Generation Computer Systems, 94*, 408–418. https://doi.org/10.1016/j.future.2018.11.001

[9] Wang, J., Wang, Y., Zhang, D., et al. (2018). Multi-task allocation in mobile crowd sensing with individual task quality assurance. *IEEE Transactions on Mobile Computing, 17*(9), 2101–2113. https://doi.org/10.1109/TMC.2018.2797061

[10] Wu, S., Wang, Y., & Tong, X. (2021). Multi-objective task assignment for maximizing social welfare in spatio-temporal crowdsourcing. *China Communications, 18*(11), 11–25. https://doi.org/10.23919/JCC.2021.11.002

[11] Li, X., & Zhang, X. (2019). Multi-task allocation under time constraints in mobile crowdsensing. *IEEE Transactions on Mobile Computing, 20*(4), 1494–1510. https://doi.org/10.1109/TMC.2019.2901436

[12] Tao, X., & Song, W. (2020). Profit-oriented task allocation for mobile crowdsensing with worker dynamics: Cooperative offline solution and predictive online solution. *IEEE Transactions on Mobile Computing, 20*(8), 2637–2653. https://doi.org/10.1109/TMC.2020.2980951

[13] Song, T., Xu, K., Li, J., et al. (2020). Multi-skill aware task assignment in real-time spatial crowdsourcing. *GeoInformatica, 24*, 153–173. https://doi.org/10.1007/s10707-019-00359-w

[14] Xiao, M., Wu, J., Huang, L., et al. (2016). Online task assignment for crowdsensing in predictable mobile social networks. *IEEE Transactions on Mobile Computing, 16*(8), 2306–2320. https://doi.org/10.1109/TMC.2016.2619666

[15] Miao, X., Kang, Y., Ma, Q., et al. (2020). Quality-aware online task assignment in mobile crowdsourcing. *ACM Transactions on Sensor Networks, 16*(3), 1–21. https://doi.org/10.1145/3398044

[16] To, H., Shahabi, C., & Kazemi, L. (2015). A server-assigned spatial crowdsourcing framework. *ACM Transactions on Spatial Algorithms and Systems, 1*(1), 1–28. https://doi.org/10.1145/2729711

[17] Wang, M., Wang, Y., Sai, A., et al. (2022, November 4). Task assignment for hybrid scenarios in spatial crowdsourcing: A Q-learning-based approach. *[EB/OL]*. Retrieved December 12, 2023.

Copyrights