
Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2025 SCPE. Volume 26, Issues 2, pp. 821–828, DOI 10.12694/scpe.v26i2.4010

THE EDGE COMPUTING EXTENSIVE DATA PROCESSING FRAMEWORK AND ALGORITHM FOR THE INTERNET OF THINGS

LUYAO GE*

Abstract. The paper studies a monitoring pedestrian recognition method driven by big data for edge-cloud collaboration. It extends the original centralized computing to edge and cloud collaborative processing. Firstly, the image boundary node N0 is preprocessed, and the extracted image is expressed in multiple levels. Then, the RGB-D multimodal image learning modeling method is applied to the edge nodes of the network using cloud computing. The boundary node uses the existing learning mode to perform action identification and uploads the identified action information to the cloud to form the final action classification. The method of bone surface fitting and dense trajectories is combined to achieve robust, dense human posture feature extraction. The directed principal component histograms of 3D stereo structures are obtained using dense point cloud data. The features of the spatial and temporal neighborhood 3D gradient histogram are extracted from the apparent texture. Through experiments, it is verified that the proposed method can significantly reduce the shortcomings of traditional centralized algorithms in data transmission and cloud storage and improve the pattern recognition accuracy by 2.2% based on edge cloud collaboration.

Key words: Edge cloud collaboration; behavior recognition; multimodal characteristics; internet of things; edge computing

1. Introduction. A traffic video monitoring system is a security system based on computer, image recognition, network communication, information processing and other technologies, including central control equipment, station control equipment, image intake and display, video signal transmission and other equipment, which can provide visual information of monitoring area for dispatchers in the control center, station duty personnel, train drivers, etc. It plays a significant role in ensuring the safe operation of traffic and high-quality services [1]. With the continuous expansion of traffic lines and stations, monitoring scenarios and target types are becoming increasingly complex, with massive scale and wide distribution. This challenges the monitoring system's intelligent, networked, real-time and maintenance management. Road video surveillance systems have experienced analog to digital video surveillance systems, which are the current intelligent video surveillance systems.

The combination of motion feature extraction and video content analysis can effectively solve the current security problems and has important practical significance for constructing a "smart city." The current cloud computing model has been unable to thoroughly and efficiently handle large-scale monitoring video data, so we extend the cloud computing model to edge cloud collaboration to improve monitoring efficiency. Literature [2] researched network-oriented resource allocation and task scheduling in an edge-cloud-oriented collaborative computing environment, conducted integrated forecasting of potential tasks from horizontal and vertical spatiotemporal scales, and optimized their configuration. Literature [3] shows that edge-cloud collaboration has advantages such as bandwidth saving, delay reduction and privacy protection, which is significant for developing power grids. Literature [4] breaks through the traditional single-cloud model and studies the flexible, intelligent manufacturing method of end-edge-cloud collaboration to adapt to SMT production lines' flexible and intelligent production requirements. Literature [5] proposes an optimization algorithm that integrates hardware and software to minimize the weighted delay of each node, aiming at migrating multi-way collaborative tasks in cloud edge collaboration.

This project will study a monitoring pedestrian recognition method driven by big data for edge-cloud collaboration. Extend the original centralized computing to edge and cloud collaborative processing. Then, the RGB-D multimodal image learning modeling method is applied to the edge nodes of the network using cloud computing.

*Dongchang College, Liaocheng University, Liaocheng, Shandong, 252000, China (gly10082024@126.com)

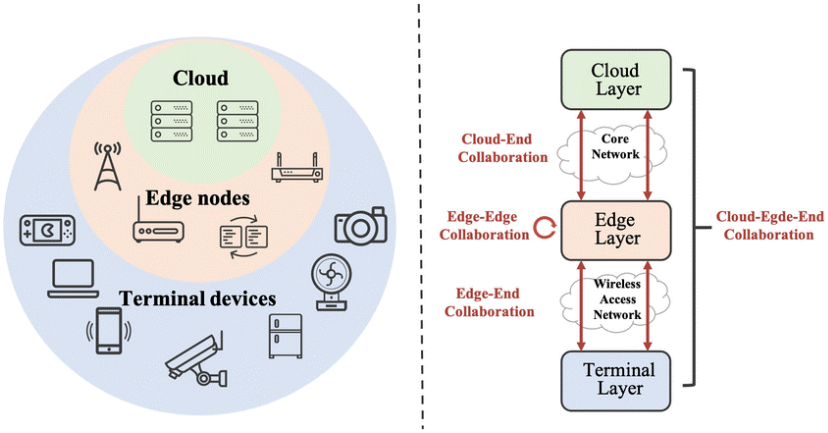

Fig. 2.1: Application architecture of cloud edge collaboration technology for traffic video surveillance system.

2. Surveillance video analysis architecture design based on cloud edge collaboration. This project takes the cloud platform of video surveillance as the research object and researches critical issues such as resource virtualization, video target recognition, surveillance scene classification, and monitoring big data processing of video surveillance, focusing on key technologies such as video image data preprocessing and surveillance scene pre-recognition based on edge computing [6]. By constructing an intelligent collaboration and collaboration framework, the video surveillance cloud and edge network can be fully coordinated to achieve the overall and global performance of video surveillance. The application architecture of cloud-edge collaboration technology for traffic video surveillance systems is shown in Figure 2.1.

The system consists of a data acquisition layer, edge node layer, network transmission layer, cloud computing center layer and service application layer. The line will install cameras in some crucial places according to the needs of the operation and public security departments, which are used to collect and perceive the video images of the original monitoring area, such as passenger flow density, personnel behavior, and on-off order. There are usually several control boxes with AI computing functions in the boundary layer. It includes intelligent hardware, software, and models. It is used for resource scheduling, target identification, scenario analysis, service scheduling, etc. The information transmission is completed in the network transmission layer, and the monitoring devices of the control center, station, yard and other regions are connected through the transmission network. A large amount of video monitoring data is uniformly stored and intelligently analyzed using edge network collaboration, and through collaborative control of the edge network, a full range of cooperation is carried out in resources, data, intelligent analysis, application services and other levels [7]. At the service application level, various road video surveillance system scenarios are intelligently identified, analyzed, and coordinated with other safety subsystems.

3. RGB-D video multimodal feature extraction.

3.1. Feature extraction of binary dense moving human posture. In this paper, the normalized skeletal surfaces are fitted based on the reference of feature-dense trajectories and MP characteristics. Then, it is integrated with moving dense trajectories to obtain dense postural features that can describe human movement [8]. First, the $U_i^{t_0}$ of each frame surface sampling is fitted in the case that the given tracking track length is the frame S. The motion trajectories of several lower layers are extracted to correlate feature P. Then, the average clustering algorithm is used to get the corresponding dictionary C_{MP}. Each path in P is sorted by Euclidean distance to obtain multiple semantic multiple features D consistent with the number of dictionary primitives and other dimensions. The basic tracing features of the authorized MP feature are as follows

$$P = (P_1, P_2, \cdots, P_{M_s}) \tag{3.1}$$

Among

$$\begin{aligned} &M_s = M_{MP}\left\lfloor \frac{M_F}{S} \right\rfloor, P_t = \left(U_i^t, U_i^{l+1}, \cdots, U_i^{i+S-1}\right) \\ &1 \le i \le M_{\sin p\,\rho} \end{aligned} \tag{3.2}$$

$M_{M\rho}$ The number of dictionary primitives extracted from the underlying features; M_F is the number of frames left after the extra time is removed. Where $M_{\sin p}$ is the fitting number of sample points on the surface. The descriptor $U_i^{\sigma_0}$ represents the sampling point of the 3D behavior action at time t_0. The combination of the three-dimensional posture $f_i(t_0)$ and the first and second derivative $\zeta f_i(t_0)$ and $\zeta^2 f_i(t_0)$ of the normalized human skeleton

$$U_i^{\sigma_0} = \left[f_i(t_0), v_j f_i(t_0), \varphi_s^{-2} f_i(t_0)\right] \tag{3.3}$$

v and φ are the associated weights of the velocity and acceleration components of the multipoint motion characteristic descriptor of Dense MP.

3.2. Feature extraction of directionally sparse principal component histogram. This project takes the pedestrian apparent surface as the research object. The apparent surface of the human body is characterized by the main direction of the divergence matrix based on a dense point cloud, and pedestrian apparent features are extracted similarly to plane regular [9]. Firstly, the dispersion of the point cloud in the spacetime neighborhood of the three-dimensional point cloud sequence and its corresponding feature vector are mapped to aspect M, thus obtaining the quantization histogram HOPC as follows:

$$y_t^{\rho *} = \left[x_1^T x_2^T x_3^T\right]^T \in R^{\beta \times M} \tag{3.4}$$

The space adjacent to the skeleton movement is segmented by the space-time pyramid method. The algorithm is integrated with sparse coding technology to obtain the HOPC signal dictionary C, the corresponding coefficient matrix, the HOPC feature of the weighted coefficient and the difference of the visual dictionary [10]. The n_f class adaptive space-time pyramid method is used to decompose the $n_H \times n_W \times 7$ type space-time basis elements and combine them with the weighted average and maximum values of the spatial axis weights of each basis element to produce corresponding feature vectors. Finally, several corresponding flight paths are joined to form SHOPC features (where D is the main direction component of the dispersion matrix). $n_H \times n_w$ is the number of grids in the range of interest and n_f is the number of time cone levels [11]. The weight of the optimal basis function and ELM's structural parameters are obtained using the maximum likelihood and least square method. Combine ELM with the weight of the base unit and add modular and non-negative constraints to the weight vector of the base unit:

$$\begin{aligned} &\min_i \min_{\varphi,\infty} \frac{1}{2}\|\varphi\|_G^2 + \frac{1}{2} z \sum_{i=1}^{M} \|\omega_i\|^2 \\ &\text{st. } \varphi^T \phi(x_i; \delta) = y_i - \omega_i, \forall i \\ &\sum_{f=1}^{M_K} \delta_f = 1\delta_f \ge 0, \forall f \end{aligned} \tag{3.5}$$

Among

$$\begin{aligned} &\varphi^T = [\varphi_1, \varphi_2, \cdots, \varphi_{M_K}] \in R^{\left(\hat{\phi}(-)+++\varphi_k, K(-)\right)_k)} \\ &\varphi_f \in R^{\phi_i(-)}, f = 1, 2, \cdots, M^K \end{aligned} \tag{3.6}$$

$\omega_i = [\omega_{1i}, \omega_{2i}, \cdots, \omega_{Ti}]^T, 1 \le i \le M$ is column i of ω. Where $Y = [y_1, y_2, \cdots, y_M] \in R^{U \times M}$ is the class matrix. In M training samples, according to formula (3.2), the kernel matrix corresponding to the multi-mode visual

Table 4.1: Hardware parameters of edge node and cloud server.

Platform	Hardware	Computing resource	Internal memory	Hard disk
Edge Node	PC	CPU: i5 8400 GPU: RTX2060	16GB	512GB
High in the clouds Server	Rack type Server	CPU: Xeno 4116 GPU: Tesla100	128GB	1TB

features is:

$$\mu(\cdot,;;r) = \mu\left(x_i, x_j; \delta\right) = \phi\left(x_i; \delta\right)^T \phi\left(x_j; \delta\right) = \sum_{f=1}^{M^K} \delta_f K_f\left(x_i, x_j\right) \tag{3.7}$$

The optimal kernel function of the multi-kernel ELM model is combined with the weight vector to obtain the optimal binding kernel function and the optimal binding kernel function in the training sample [12]. Iterate repeatedly to get the best construction parameters v^* and δ^* of the best ELM, and finally get the decision vector for the test sample C:

$$g_c(x) = \sum_{i=1}^{M} v_{ic}^* \sum_{f=1}^{m} \delta_f^* K_f\left(x_i, x\right) \tag{3.8}$$

$v^* \in R^{M \times U}$ is the best subset of Lagrange. Where δ^* is the best core weight vector. The subscript c is the classification mark for the test sample U.

4. Experimental analysis. Simulation experiments were carried out to test the effectiveness of the behavior identification mechanism under edge-cloud collaboration [13]. The test system uses 6 PCS and 1 frame server on the boundary and cloud sides, respectively. The problem of the limited computing power of boundary nodes is simulated using the limitations of VMWare virtual machines on computing and storage. Table 4.1 lists hardware data for border nodes and cloud computing servers.

4.1. Data Set.

4.1.1. TU-RGB+D120 data set. This database has the most extensive collection of bone samples in the world. The sample size was 114,480, and the behavior types were 120. The shooting angles are increased to 155, and 106 targets are tested [14]. The database provides two test criteria for cross-topics and cross-settings. The test criterion based on cross-topic will take 53 individuals as the research objects and take the training set and the test set as the research objects. Number the training set as even and mark the test set as odd in the cross-setting test criteria.

4.1.2. Kinetics Database. This database includes 400 categories of human behavior extracted from YouTube, with at least 300 video clips of each action of approximately 10 seconds. The range of activities is extensive, including the interaction between people and things and between people [15].

4.2. Test the advantages of edge cloud collaborative computing. This article compares using performance in an edge-cloud collaboration environment with the following approach [16]. The system's performance is mainly reflected in three dimensions: network transmission volume, total task time and device energy consumption. In the single boundary operation, the operation is carried out on the boundary without any network transmission. In a single cloud computing environment, when the amount of data increases, the data traffic increases with the increase of the data scale (figure 4.1) [17]. The extracted feature information must be uploaded to the cloud first, then the cloud computing learns and distributes it to the edge nodes, thus reducing the initial network overhead. In the following identification, only the identified data needs to be uploaded to the cloud to reduce the growth of network traffic [18]. With the increased number of tasks, the advantages of edge cloud collaboration are more prominent. Power consumption refers to the proportion of CPU, memory and

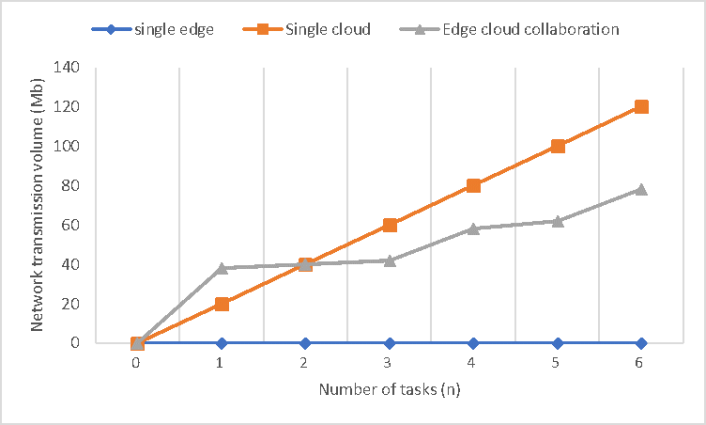

Fig. 4.1: Comparison of data traffic of the three mechanisms under the same number of jobs.

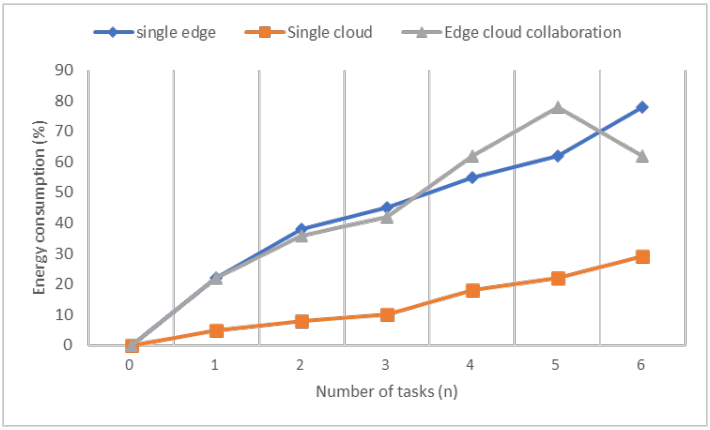

Fig. 4.2: Comparison of energy consumption of the three methods under the same number of operations.

hard disk each machine uses to operate [19]. Under the premise of ensuring the integrity of the identification work, the computing resources of cloud computing and the memory capacity of the edge have been maximized. Figure 4.2 shows the comparison of energy consumption of the three schemes under the same number of tasks.

The results of the comparative test of operation time in the three plans are shown in figure 4.3. Because of the network's limited storage space and CPU resources, single identification takes too long in the single boundary operation mode. And the more tasks there are, the more time it takes. The boundary end indicates that the task is terminated when the capacity limit is reached [20]. More time is needed to learn and deploy the cloud model before the first run under edge cloud collaborative computing. So, the time of first use is more than the time of a single use, but when the number of uses increases, the time of use will be less than the other two methods, which can prove the advantages of edge cloud collaboration.

This article selects a classic case with a resolution of 1280x720, a time length of 2 seconds, and a frame rate of 20 frames [21]. When the number of tasks increases, the more data the video uploads, the more likely it is to cause network congestion [22]. The method in this paper does not need to upload the video to the cloud, but only requires edge cloud collaboration in the range of 10×50x16b=8 kb. This dramatically reduces the network traffic for uploading videos. Edge cloud collaboration shows better performance in realizing human movements considering the impact of all aspects.

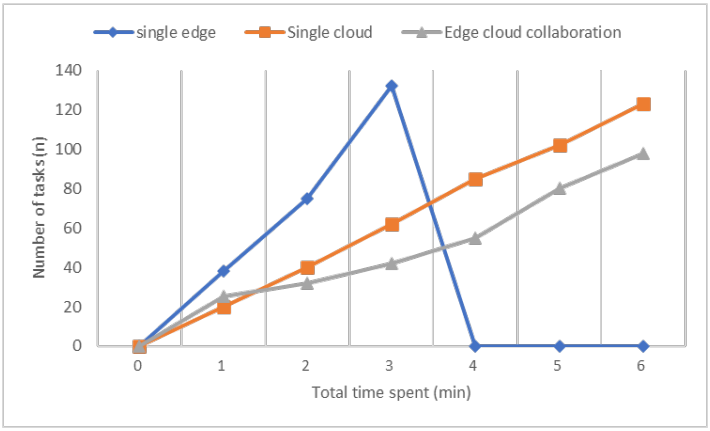

Fig. 4.3: Comparison of the total time of the three plans for the same number of jobs.

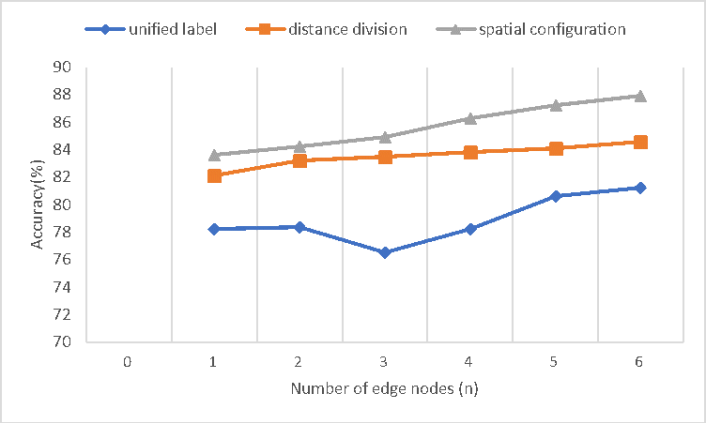

Fig. 4.4: Influence of the number of boundary nodes on identification accuracy.

4.3. Comparison of classification accuracy of RGB-D multi-mode images by three classification methods. The classification accuracy of RGB-D multi-mode images under three classification strategies is studied based on edge-cloud collaboration. The change in identification accuracy can be observed by adjusting the number of boundary nodes in the experiment [23]. The model is classified correctly after using a different number of nodes and division methods. Figure 4.4 shows the influence of the number of edge nodes on the recognition accuracy.

The number of boundary nodes gradually increases under the three partition methods, and the recognition accuracy is gradually improved until all boundary nodes are adopted. However, the segmentation method is based on a single market that lacks local feature differences when processing bone sequences. This leads to a decline in accuracy. Compared with region segmentation, the spatial layout partitioning method can obtain better recognition [24]. Therefore, this project proposes a multi-model modeling method based on RGB-D multimodal images to enhance its practical value in complex environments.

4.4. Cloud fusion effect verification. The identification effect of multiple boundary nodes in cloud computing is tested using two sets of different samples [25]. Firstly, at the edge node N_0, the similar frames in the image are eliminated, and the pose is estimated [26]. Secondly, the action identification is carried out at $N_1 \sim N_m$, and finally, the identification results are uploaded to the cloud computing for data processing. Tables

Table 4.2: Joint identification accuracy of edge nodes and cloud computing based on NTU RGB+D120.

	Cross Subject	Cross View
N1	84.58%	90.73%
N2	86.04%	91.67%
N3	84.06%	90.52%
N4	85.31%	91.04%
N5	85.94%	91.35%
Single cloud	85.52%	91.56%
Merge.	87.40%	93.44%

Table 4.3: Accuracy of boundary node and cloud fusion identification in Kinetics database.

	Top 1	Top 5
N1	86.88%	88.75%
N2	78.54%	89.90%
N3	85.52%	88.23%
N4	84.27%	90.42%
N5	85.10%	88.44%
Single cloud	86.04%	89.06%
Merge.	88.02%	91.88%

4.2 and 4.3 are performance tests conducted on the Kinetics platform in a mix of single-edge, single-cloud and cloud-based modes based on NTU- RGB+D120.

Under the cross-subject test criterion, the accuracy of a single cloud image reached 85.52%, and the accuracy of a single cloud image was 84.06% 86.04%. The mean value is 85.10%. In the cloud computing environment, the experimental effect reaches 87.40%, about 2.29% higher than the average recognition accuracy in a single cloud computing environment. The accuracy of the single cloud model based on cross-view is 91.56%, and the accuracy of the single cloud model is 90.52% to 91.67%. The average was 91.04%, while the cloud computing model reached 93.13%, an increase of 2.40%. It is found that motion recognition accuracy can be improved by 2% based on cloud computing.

The cloud server can perform identity authentication faster and improve accuracy in a single cloud computing environment. However, a large number of video and image uploads will lead to data congestion, which will also cause a significant burden on the performance of the cloud server. There may be some problems in data transmission, such as missing frames, which seriously reduces the system's performance. On the premise that the energy consumption of the edge cloud collaboration algorithm studied in this project is comparable to that of a single cloud algorithm, it not only reduces the data transmission overhead, reduces the cloud server's data storage burden but also effectively plays the computing performance of the edge node and improves the identification accuracy of the system.

5. Conclusion. Firstly, the collaborative computing architecture of edge cloud is studied to realize the effective configuration and efficient use of edge and cloud. Secondly, multi-mode image features based on RGB-D are constructed, and image processing is carried out in the cloud. Finally, the image information of each region is uniformly processed on the cloud platform. The experiment verifies that this project will improve cognitive accuracy and significantly improve the network transmission capacity, equipment energy consumption, total task time and other indicators.

6. Acknowledgements. The Project is supported by Research on Data Governance of Education Enrollment Examination (2023QY038). The project information is 2023 project of the "14th Five-Year Plan" of Education Science in Shandong Province.

REFERENCES

[1] Cui L, Su X, Ming Z, et al. CREAT: Blockchain-assisted compression algorithm of federated learning for content caching in edge computing. IEEE Internet of Things Journal, 2020, 9(16): 14151-14161.
[2] Wang X, Ning Z, Guo S. Multi-agent imitation learning for pervasive edge computing: A decentralized computation offloading algorithm. IEEE Transactions on Parallel and Distributed Systems, 2020, 32(2): 411-425.
[3] Lakhan A, Mohammed M A, Rashid A N, et al. Deadline aware and energy-efficient scheduling algorithm for fine-grained tasks in mobile edge computing. International Journal of Web and Grid Services, 2022, 18(2): 168-193.
[4] Lv Z, Chen D, Wang Q. Diversified technologies in internet of vehicles under intelligent edge computing. IEEE transactions on intelligent transportation systems, 2020, 22(4): 2048-2059.
[5] Tang M, Wong V W S. Deep reinforcement learning for task offloading in mobile edge computing systems. IEEE Transactions on Mobile Computing, 2020, 21(6): 1985-1997.
[6] Xiao H, Zhao J, Pei Q, et al. Vehicle selection and resource optimization for federated learning in vehicular edge computing. IEEE Transactions on Intelligent Transportation Systems, 2021, 23(8): 11073-11087.
[7] Mills J, Hu J, Min G. Multi-task federated learning for personalised deep neural networks in edge computing. IEEE Transactions on Parallel and Distributed Systems, 2021, 33(3): 630-641.
[8] Zhang L, Ansari N. Latency-aware IoT service provisioning in UAV-aided mobile-edge computing networks. IEEE Internet of Things Journal, 2020, 7(10): 10573-10580.
[9] Xiong X, Zheng K, Lei L, et al. Resource allocation based on deep reinforcement learning in IoT edge computing. IEEE Journal on Selected Areas in Communications, 2020, 38(6): 1133-1146.
[10] Dai X, Xiao Z, Jiang H, et al. Task co-offloading for D2D-assisted mobile edge computing in industrial internet of things. IEEE Transactions on Industrial Informatics, 2022, 19(1): 480-490.
[11] Liu Y, Wang S, Zhao Q, et al. Dependency-aware task scheduling in vehicular edge computing. IEEE Internet of Things Journal, 2020, 7(6): 4961-4971.
[12] Ning Z, Dong P, Wang X, et al. Distributed and dynamic service placement in pervasive edge computing networks. IEEE Transactions on Parallel and Distributed Systems, 2020, 32(6): 1277-1292.
[13] Luo Q, Hu S, Li C, et al. Resource scheduling in edge computing: A survey. IEEE Communications Surveys & Tutorials, 2021, 23(4): 2131-2165.
[14] Liu X, Yu J, Feng Z, et al. Multi-agent reinforcement learning for resource allocation in IoT networks with edge computing. China Communications, 2020, 17(9): 220-236.
[15] Yang L, Yao H, Wang J, et al. Multi-UAV-enabled load-balance mobile-edge computing for IoT networks. IEEE Internet of Things Journal, 2020, 7(8): 6898-6908.
[16] Fan G, Chen L, Yu H, et al. Multi-objective optimization of container-based microservice scheduling in edge computing. Computer Science and Information Systems, 2021, 18(1): 23-42.
[17] Zhan C, Hu H, Sui X, et al. Completion time and energy optimization in the UAV-enabled mobile-edge computing system. IEEE Internet of Things Journal, 2020, 7(8): 7808-7822.
[18] Yuan Q, Li J, Zhou H, et al. A joint service migration and mobility optimization approach for vehicular edge computing. IEEE Transactions on Vehicular Technology, 2020, 69(8): 9041-9052.
[19] Wang Y, Lang P, Tian D, et al. A game-based computation offloading method in vehicular multiaccess edge computing networks. IEEE Internet of Things Journal, 2020, 7(6): 4987-4996.
[20] Liu X, Yu J, Wang J, et al. Resource allocation with edge computing in IoT networks via machine learning. IEEE Internet of Things Journal, 2020, 7(4): 3415-3426.
[21] Zeng F, Chen Q, Meng L, et al. Volunteer assisted collaborative offloading and resource allocation in vehicular edge computing. IEEE Transactions on Intelligent Transportation Systems, 2020, 22(6): 3247-3257.
[22] Zhang R, Cheng P, Chen Z, et al. Online learning enabled task offloading for vehicular edge computing. IEEE Wireless Communications Letters, 2020, 9(7): 928-932.
[23] Yuan H, Zhou M C. Profit-maximized collaborative computation offloading and resource allocation in distributed cloud and edge computing systems. IEEE Transactions on Automation Science and Engineering, 2020, 18(3): 1277-1287.
[24] Zhan Y, Guo S, Li P, et al. A deep reinforcement learning based offloading game in edge computing. IEEE Transactions on Computers, 2020, 69(6): 883-893.
[25] Wang L, Wang K, Pan C, et al. Multi-agent deep reinforcement learning-based trajectory planning for multi-UAV assisted mobile edge computing. IEEE Transactions on Cognitive Communications and Networking, 2020, 7(1): 73-84.
[26] Bai T, Pan C, Deng Y, et al. Latency minimization for intelligent reflecting surface aided mobile edge computing. IEEE Journal on Selected Areas in Communications, 2020, 38(11): 2666-2682.

Edited by: Hailong Li
Special issue on: Deep Learning in Healthcare
Received: May 22, 2024
Accepted: Jun 25, 2024