

# 基于安全与低能耗的传感云边缘协同优化策略



赵庶旭, 张占平\*, 王小龙, 韩淑梅, 元琳, 张家祯

(兰州交通大学电子与信息工程学院 兰州 730071)

**【摘要】**多传感器传感网数据采集效率低下,且大量数据在传感云处理存在数据泄露风险。基于此,首先设计了一种安全、节能及高效的分布式边缘协同传感网资源选择架构,提出了一种边缘协同分析节点选择(ECANS)方案。通过对用户请求的分析,获取传感网节点的选择策略,以降低传感节点数据采集的时延和能耗。其次,构建了一种最大化隐私熵的边缘协同传感网隐私保护数据卸载模型,并通过智能启发式算法得到隐私熵最大的边缘资源选择策略。实验结果表明,与ENS数据采集方案相比,ECANS方案使节点时延与能耗分别降低了56.71%和57.66%;在边缘资源选择阶段,与GA资源选择方案和PSO资源选择方案相比,最大化隐私熵模型使系统隐私熵分别提高32.07%及15.36%;与不引入no-EC相比,传感网节点时延和能耗平均降低了46.92%与11.26%。

**关键词** ECANS; 边缘协同分析; 多传感器传感网; 隐私保护; 资源选择  
中图分类号 TP39 文献标志码 A doi:10.12178/1001-0548.2022009

## Collaborative Optimization Strategy of Edge Sensor Cloud Based on Security and Low Energy Consumption

ZHAO Shuxu, ZHANG Zhanping\*, WANG Xiaolong, HAN Shumei, YUAN Lin, and ZHANG Jiazhen

(School of Electronic and Information Engineering, Lanzhou Jiaotong University Lanzhou 730071)

**Abstract** There are two problems to be solved in multi-sensors wireless sensor networks: low data collection efficiency and the risk of data leakage when a large amount of data is processed in sensor cloud. Owing to these reasons, we devise a safe, energy-saving, and efficient distributed edge collaborative sensor network resource selection architecture firstly. Secondly, to address first problem, an edge analysis node selection (edge collaborative analysis node selection, ECANS) algorithm is proposed. Through the analysis of user requests, the best strategy of sensor network nodes is obtained to reduce the node's delay and energy consumption of data collection. Aiming at the second problem, an edge collaborative sensor network privacy protection data offloading model is constructed to maximize privacy entropy, and the edge resource selection strategy with the largest privacy entropy is gained through intelligent heuristic algorithm. At last, experimental results show that ECANS algorithm can reduce node delay and energy consumption by 56.71% and 57.66% compared with effective node sensing (ENS) data collection methods. In the edge resource selection stage, the maximum privacy entropy model makes the system privacy entropy increased by 32.07% and 15.36%, compared with genetic algorithm (GA) resource selection scheme and particle swarm optimization (PSO) resource selection scheme. The latency and energy consumption of the sensor network were reduced by 46.92% and 11.26% compared with no-EC.

**Key words** ECANS; edge collaboration analysis; multisensor network; privacy protection; resource selection

无线传感器网络(wireless sensor network, WSN)具有部署方便、抗干扰能力强、安装成本低等特点,它能扩展人们收集外界信息的能力<sup>[1]</sup>。将传感器节点部署在不同的环境中,实时、便捷地获取所

需数据,实现了信息世界与物理世界的无缝融合。因此,传感网已被广泛应用于环境监测、智能家居、智能制造、智慧医疗等领域<sup>[2]</sup>。近年来,日益增长的应用需求和用户规模对传感网数据的多样

收稿日期: 2022-01-04; 修回日期: 2022-03-03

基金项目: 甘肃省重点研发计划(20YF8GA123); 甘肃省优秀研究生“创新之星”(2021CXZX-557)

作者简介: 赵庶旭(1977-),男,博士,教授,主要从事边缘计算、智能交通、深度学习等方面的研究。

\*通信作者: 张占平, E-mail: 0219665@stu.lzjtu.edu.cn

性、实时性和安全性提出了新的要求。传感云的诞生为之提供了有效的解决方案。传感云作为云计算与传感网结合的一种新技术，给传感网注入了新的活力，催生了新的应用与服务，同时拓展了传感网在数据处理和存储方面的能力<sup>[3]</sup>。但考虑到云服务商 (cloud service provider, CSP) 提供的服务机制存在数据隐私和数据完整性等潜在问题<sup>[4]</sup>，无论 CSP 采取何种可靠措施，基础设施都可能出现漏洞。同时，信誉度低的 CSP 可能会泄露或隐藏用户存储的私人数据<sup>[5]</sup>。边缘传感云能为用户提供低时延、低能耗和高安全性的服务，其中边缘协同传感云资源分配已成为当前研究的重点问题。

文献 [6] 从数据采集、存储、处理及安全性等方面对传感云和边缘协同传感网进行对比分析，研究表明基于中心化的传感云易发生数据泄露，且传感网与云端频繁的数据交换易导致核心网络阻塞及能耗损失加剧；同时该文献指出边缘协同传感网是一种扩展传感网计算能力和存储能力的有效技术。文献 [7] 针对精准农业中数据采集冗余会增加 WSN 能耗和时延问题，提出一种基于边缘计算的低时延和低能耗多传感器数据采集方法，但没有考虑到传感节点密集型数据的处理。文献 [8] 在智慧家居场景中针对传感云架构易出现数据丢失及篡改的问题，提出了一种结合边缘计算与区块链技术的数据处理架构，通过在边缘层部署异常数据检测算法，以保证传感网数据质量，但没有考虑到时延及能耗的优化。文献 [9] 针对时空变化场景中 WSN 感知质量与节点能耗不平衡的问题，提出了一种分层的边缘自适应框架，通过主成分分析和稀疏贝叶斯学习算法得到了感知质量和传感节点能耗的权衡策略，但未考虑感知时延的优化。文献 [10] 在智慧物流场景中，针对分布式仓库管理系统响应时延较长的问题，构建了一种基于边缘计算的 WSN 快速响应系统，并通过一种快速响应算法优化系统响应时延，但该方法单方面考虑时延会导致 WSN 节点电量消耗过快。

上述研究主要集中于边缘协同传感网数据质量、时延及能耗方面的优化，忽略了密集型数据卸载到边缘服务器的隐私泄露问题。为解决该问题，本文设计了一种最大化隐私熵的分布式边缘协同资源选择 (distributed edge collaborative resource selection, DECERS) 联合优化方案，针对优化问题建立数学模型，并采用智能启发式算法得到时延、能耗及隐私熵联合优化的边缘资源选择策略。同时，考虑到多

传感器传感网数据采集中冗余数据导致节点时延和能耗增加问题，提出一种边缘协同分析节点选择 (edge collaborative analysis node selection, ECANS) 方案，在最大化数据质量条件下得到数据采集策略。

## 1 系统描述

本文系统模型如图 1 所示。该模型由传感器网络、边缘服务器、边缘优化器和应用程序组成。其中， $N=\{N_1, N_2, \dots, N_\zeta\}$  表示传感网集合， $N_\zeta=\{n_1, n_2, \dots, n_\psi\}$  表示传感网  $\zeta$  的传感节点集合， $X=\{x_1, x_2, \dots, x_\psi\}$  表示传感网  $\zeta$  中每个传感节点的位置信息， $\Delta=\{l_1, l_2, \dots, l_k\}$  为区域信息，且满足式 (1) 和式 (2)。

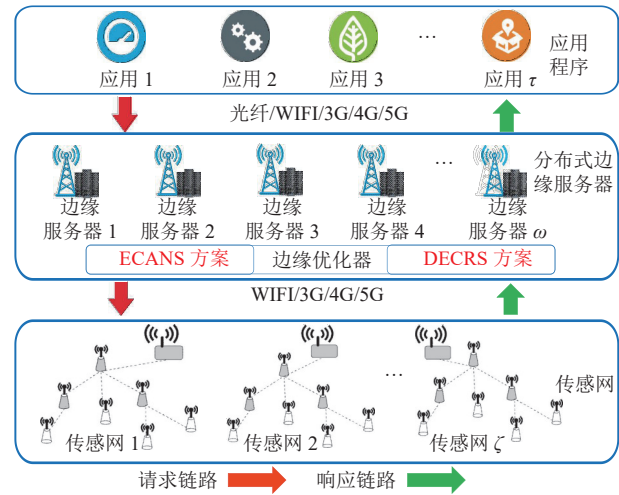


图 1 系统模型

$$l_i \cap l_j = \emptyset \quad 1 \leq i \leq \kappa, 1 \leq j \leq \kappa \quad (1)$$

$$\sum_{i=1}^{\kappa} l_i = \Delta \quad 1 \leq i \leq \kappa \quad (2)$$

传感器集合由  $P=\{p_1, p_2, \dots, p_\xi\}$  表示，规定相同区域内不同传感节点的传感器不完全相同，且能采集声音、视频、图像、湿度、温度等数据。边缘服务器由集合  $Q=\{q_1, q_2, \dots, q_\omega\}$  表示。考虑到传感网节点资源受限，当应用程序集合  $R_{app}=\{R_1, R_2, \dots, R_\tau\}$  中的一个或多个应用请求计算密集型数据时，需要把数据分块，然后卸载到边缘服务器。与其他物联网设备不同，传感网中的数据可分为直接量和间接量，直接量不需要额外计算，从传感器读取后能直接使用，如温度、湿度等；而间接量需要把采集的数据处理后将结果提供给应用程序，如图像、视频目标检测、语音识别等<sup>[11]</sup>。若强行在传感网节点计算间接量，会消耗大量节点能耗，导致传感网

节点过早死亡。

本文架构的处理流程主要由数据采集和边缘资源选择两阶段组成, 如图 2 所示。在数据采集阶段, 为降低数据采集的时延与能耗, 提出了 ECANS

方案获得数据采集策略; 在边缘资源选择阶段, 提出了一种最大化隐私熵的 DECERS 方案, 先根据优化目标建立模型, 然后通过智能启发式算法得到最大化隐私熵下的最优边缘资源选择策略。

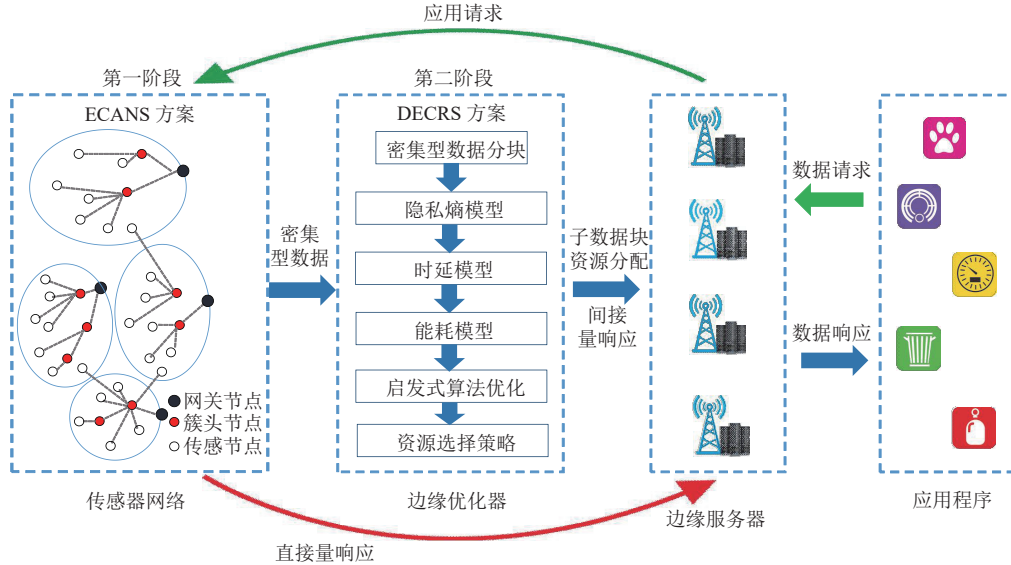


图 2 系统处理流程

### 1.1 ECANS 方案

在多传感器传感网数据采集中, 不合理的数据采集策略不但消耗大量的节点能耗, 还会增加数据采集时延。本文考虑到现有方法存在多传感器传感网数据采集中采集数据冗余度大的问题, 设计了基于字母编码 (letter coding, LC) 的 ECANS 方案, 通过 ECANS 算法得到数据质量最大化下的节点选择策略。定义单个传感器与节点的关系由式 (3) 表示, 则所有节点与传感器的连接关系由矩阵  $\Pi$  表示:

$$\pi_{\psi,\xi} = \begin{cases} 1 & \text{传感节点}\psi\text{与传感器}\xi\text{连接} \\ 0 & \text{其他} \end{cases} \quad (3)$$

$$\Pi = \begin{bmatrix} \pi_{1,1} & \pi_{1,2} & \cdots & \pi_{1,\xi} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{\psi,1} & \pi_{\psi,2} & \cdots & \pi_{\psi,\xi} \end{bmatrix} \quad (4)$$

用户通过应用程序请求的数据集合由  $R_\tau = \{Rd_1, Rd_2, \dots, Rd_\sigma\}$  表示, 单个传感节点提供的数据由  $n_\psi = \{s_1, s_2, \dots, s_\rho\}$  表示, 其中  $s_\rho$  为一个三元组  $\langle id, type, data\ volum \rangle$ , 且  $type (type \in \{0, 1\})$  为数据类型, 0 表示直接量 (无需卸载), 1 表示间接量 (需要卸载),  $data\ volum$  为数据量。考虑到应用程序请求数据与传感器提供数据存在长度差异问题, 本文提出一种基于 LC 的数据编码方法。已知请求数据长

度为  $\sigma$ , 响应数据长度为传感节点上传感器的个数。单个传感节点获取有效数据的过程为:

- 1) 获取应用数据请求  $Rd_\sigma$ , 使用 LC 编码, 并将 LC 编码转换为二进制编码得到  $Val_\sigma^{R\_Code}$ ;
- 2) 获取节点传感器数据信息, 使 LC 对传感器编码, 经二进制编码转换得  $Val_\psi^{S\_Code}$ ;
- 3) 将  $Val_\sigma^{R\_Code}$  与  $Val_\psi^{S\_Code}$  进行逻辑“与”运算, 得到单个节点的有效数据编码, 编码流程如图 3 所示。



图 3 LC 数据编码流程

所有节点提供的数据与用户应用程序请求数据须满足式 (5):

$$\frac{\sum_{i=1}^{\psi} \sum_{j=1}^{\xi} \pi_{i,j} \text{Val}_j^{\text{S-code}}}{\sum_{k=1}^{\sigma} \text{Val}_k^{\text{R-code}}} = 1 \quad (5)$$

$$\text{DQ}_{\xi,\tau} = \frac{\sum_{j=1}^{\rho} \pi_{\psi,j} \text{Val}_j^{\text{S-code}}}{\sum_{k=1}^{\sigma} \text{Val}_k^{\text{R-code}}} \quad (6)$$

**定义** 有效数据与总数据的比值为数据质量 (data quality, DQ), 数据采集中 DQ 值越大, 表明节点选择策略越好<sup>[7]</sup>。节点 $\xi$ 的数据质量由式 (6) 表示。为得到节点选择策略, 本文设计了 ECANS 算法, 描述如算法 1 所示, 其中 ID 函数用于返回该数据编码对应的节点序号。

#### 算法 1 ECANS

输入: 数据请求 $\text{Rd}_\sigma$ , 传感网节点, 传感网区域 $\Delta$ , 节点位置矩阵 $\mathbf{X}$ , 传感器与节点关系矩阵 $\mathbf{II}$ , 用户请求区域 $\Delta_{\text{user}}$ ;

输出: 满足数据 $\text{Rd}_\sigma$ 的最佳传感节点集合 $\text{Node}_{\text{out}}$ ;

初始化 $\text{Node}_{\text{temp}} \leftarrow N$ ,  $\text{Node}_{\text{out}} \leftarrow \emptyset$ ,  $\text{Val}_{\text{temp}} \leftarrow \emptyset$ ,  $\text{Val}_{\text{temp}}^{\text{R-code}} \leftarrow \emptyset$ ,  $\text{Val}_{\text{temp}}^{\text{S-code}} \leftarrow \emptyset$ ;

for 节点 $i$ 从 1 到 $\psi$ , 执行

if  $\Delta_{\text{user}} \notin x_i$ , 执行

从 $\text{Node}_{\text{temp}}$ 中剔除节点 $n_i$ ;

end if

end for

对请求 $\text{Rd}_\sigma$ 使用 LC 编码, 得到 $\text{Val}_\sigma^{\text{R-code}}$ ;

对 $\text{Node}_{\text{temp}}$ 中的每个节点使用 LC 方法编码, 得到 $\text{Val}_\psi^{\text{S-code}}$ ;

for 节点 $j$ 从 1 到 $\text{Node}_{\text{temp}}$ 长度, 执行

由式 (6) 计算 $\text{Node}_{\text{temp}}$ 中单个节点的数据质量, 把结果保存到 $\text{Val}_{\text{temp}}$ ;

end for

对 $\text{Val}_{\text{temp}}$ 中的节点按 DQ 降序排序;

for 节点 $k$ 从 1 到集合 $\text{Val}_{\text{temp}}$ 长度, 执行

$\text{Val}_{\text{temp}}^{\text{code}} \leftarrow \text{Val}_{\text{temp}}^{\text{code}} \cup \text{Val}_k^{\text{S-code}}$

$\text{Node}_{\text{out}} \leftarrow \text{ID}(\text{Val}_k^{\text{S-code}})$

if  $\text{Val}_\sigma^{\text{R-code}} \subseteq \text{Val}_{\text{temp}}^{\text{code}}$ , 执行

break;

end if

end for

返回 $\text{Node}_{\text{out}}$

## 1.2 分布式资源选择隐私保护模型

在传感网节点密集型数据处理中, 基于中心化的计算模式增加了数据泄露风险。为降低该风险, 提出一种基于隐私保护的 DECERS 方案。首先将单个密集型数据分块; 然后基于智能启发式算法得到最大化隐私熵条件下的边缘资源选择策略; 最后通过该策略把数据卸载到分布式边缘服务器 (distributed edge server, DES) 上。该方法有效提高了数据块在 DES 上计算的分布, 从而达到降低数据泄露的目的。

密集型数据的资源选择策略由矩阵 $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_\eta]$ 表示, 其中 $\mathbf{y}_\eta$ 由式 (7) 表示, 当子数据块 $\theta$ 在边缘服务器 $q_\omega$ 计算时 $b_{\theta,\omega}$ 为 1, 否则 $b_{\theta,\omega}$ 为 0。对于应用请求 $\text{Rd}_\sigma$ , 经 ECANS 方案节点选择及数据采集后所得的密集型数据由集合 $\text{Data} = \{\text{data}_1, \text{data}_2, \dots, \text{data}_\eta\}$ 表示。该集合中密集型数据 $\text{Data}$ 与资源选择策略 $\mathbf{Y}$ 的关系由式 (8) 表示。且 $\text{data}_\eta$ 可被随机分成 $\theta$ 个子块, 如集合 $\text{data}_\eta = \{d_{\eta,1}, d_{\eta,2}, \dots, d_{\eta,\theta}\}$ , 为保证数据的完整性,  $\theta$ 需满足式 (9) 约束, 其中 $\omega$ 为边缘服务器数量。 $\text{data}_\eta$ 的子数据块 $d_{\eta,\theta}$ 卸载到边缘服务器的概率 $\text{pr}_{\eta,k}$ 服从参数为 $\chi$ 的泊松分布<sup>[12]</sup>, 如式 (10) 所示。子数据块 $d_{\eta,\theta}$ 与其资源分配概率的关系如式 (11) 所示。

$$\mathbf{y}_\eta = \begin{bmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,\omega} \\ \vdots & \vdots & \ddots & \vdots \\ b_{\theta,1} & b_{\theta,2} & \dots & b_{\theta,\omega} \end{bmatrix} \quad (7)$$

$$\begin{pmatrix} \text{Data} \\ \mathbf{Y} \end{pmatrix} = \begin{pmatrix} \text{data}_1 & \text{data}_2 & \dots & \text{data}_\eta \\ \mathbf{y}_1 & \mathbf{y}_2 & \dots & \mathbf{y}_\eta \end{pmatrix} \quad (8)$$

$$\theta = \text{random}(1, \omega) \quad (9)$$

$$\text{pr}_{\eta,\theta} = \frac{\chi^\theta}{\theta!} e^{-\chi} \quad \theta = 1, 2, \dots, \theta \quad (10)$$

$$\begin{pmatrix} \text{data}_\eta \\ \text{Pr}_\eta \end{pmatrix} = \begin{pmatrix} d_{\eta,1} & d_{\eta,2} & \dots & d_{\eta,\theta} \\ \text{pr}_{\eta,1} & \text{pr}_{\eta,2} & \dots & \text{pr}_{\eta,\theta} \end{pmatrix} \quad (11)$$

数据 $\text{data}_\eta$ 的隐私熵由式 (12) 计算, 则所有数据 $\text{Data}$ 的资源选择策略 $\mathbf{Y}$ 与隐私熵 $K$ 的关系由式 (13) 表示,  $\text{Data}$ 中所有数据卸载的平均隐私熵 $K$ 由式 (14) 表示。

$$k(\text{data}_\eta) = - \sum_{\theta=1}^{\theta} \text{pr}_{\eta,\theta} \log_2(\text{pr}_{\eta,\theta}) \quad (12)$$

$$\begin{pmatrix} \mathbf{Y} \\ \mathbf{K} \end{pmatrix} = \begin{pmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_\eta \\ k(\text{data}_1) & k(\text{data}_2) & \cdots & k(\text{data}_\eta) \end{pmatrix} \quad (13)$$

$$K = \frac{1}{\eta} \sum_{j=1}^{\eta} k(\text{data}_j) \quad (14)$$

### 1.3 分布式资源选择时延、能耗模型

为降低边缘资源分配中的时延和能耗, 建立时延和能耗的优化模型。在场景中, 边缘服务器的计算速率由集合  $F = \{f_1^{\text{mec}}, f_2^{\text{mec}}, \dots, f_\omega^{\text{mec}}\}$  表示, 不同计算速率的边缘服务器计算时产生的时延和能耗不同。

#### 1) 边缘资源选择时延模型

分布式边缘资源选择的时延包括传感网数据采集时延、传感网、核心网的数据传输时延和边缘服务器上的计算时延。设集合  $\text{DS}_{\text{Data}} = \{\text{ds}_1, \text{ds}_2, \dots, \text{ds}_\eta\}$  表示 Data 的数据量。数据在 WSN 的传输速率满足香农定理:

$$C_{\text{sensor}} = B_{\text{sensor}} \log_2 \left( 1 + \frac{P_s \delta_s}{\varphi_s^2} \right) \quad (15)$$

式中,  $C_{\text{sensor}}$  表示传感网传输速率;  $B_{\text{sensor}}$  表示传感网带宽;  $P_s$  为传输功率;  $\delta_s$  为天线增益;  $\varphi_s$  表示信道噪声。数据在传感网的传输时延由式 (16) 计算, 其中  $\text{ds}_i$  为第  $i$  个传感网节点密集型数据的数据量。

节点数据从 WSN 到达边缘服务器需经过核心网, 其接入方式分为无线网络接入和有线网络接入<sup>[13-14]</sup>。本文选择天线阵列 (antenna array, AA) 接入, 因此数据在核心网的传输速率仍然满足香农定理, 如式 (17), 所有数据 Data 在核心网的传输时延由式 (18) 计算。

$$t_{\text{sensor}} = \frac{\sum_{i=1}^{\eta} \text{ds}_i}{C_{\text{sensor}}} \quad (16)$$

$$C_{\text{wireless}} = B_{\text{wireless}} \log_2 \left( 1 + \frac{P_{\text{wireless}} \delta_{\text{wireless}}}{\varphi_{\text{wireless}}^2} \right) \quad (17)$$

式中,  $B_{\text{wireless}}$  为无线核心网的带宽;  $P_{\text{wireless}}$  为无线核心网传输功率;  $\delta_{\text{wireless}}$  为无线核心网天线增益;  $\varphi_{\text{wireless}}$  为无线核心网噪声功率。

$$t_{\text{wireless}} = \sum_{i=1}^{\eta} \text{ds}_i / C_{\text{wireless}} \quad (18)$$

多个子数据块  $d_{\eta,\theta}$  在边缘服务器上并行计算, 计算时延  $t_{\eta,j}$  如式 (19) 所示, 且  $\text{data}_\eta$  的计算时延由多个子块计算时延的最大值  $t_\eta$  决定, 如式 (20) 所

示。传感节点的数据采集时延及计算时延如式 (21) 和式 (22) 所示。

$$t_{\eta,j} = b_{j,i} O^j \frac{\text{ds}_{\eta,j}}{f_i^{\text{mec}}} \quad 1 \leq j \leq \theta, 1 \leq i \leq \omega \quad (19)$$

$$t_\eta = \max \{t_{\eta,1}, t_{\eta,2}, \dots, t_{\eta,\theta}\} \quad (20)$$

$$t_{\text{collect}} = \sum_{i=1}^{\eta} t_i^{\text{sensing}} \quad (21)$$

$$t_{\text{MEC}} = \sum_{j=1}^{\eta} \sum_{k=1}^{\theta} t_{j,k} \quad (22)$$

式中,  $t_i^{\text{sensing}}$  表示  $\text{data}_i$  数据采集中的感知时延。由于不同类型时延均满足线性关系, 则总时延  $T$  为:

$$T = t_{\text{sensor}} + t_{\text{wireless}} + t_{\text{MEC}} + t_{\text{collect}} \quad (23)$$

#### 2) 边缘资源选择能耗模型

在分布式边缘资源选择中, 系统能耗主要来源于数据采集、传输和计算过程。系统能耗由数据量、传输功率和计算功率决定, 且不同的边缘服务器由于计算能力的不同, 其计算功率也存在差异。多个边缘服务器单位数据的计算功率由集合  $P_{\text{MEC}} = \{p_1, p_2, \dots, p_\omega\}$  表示, 子数据块  $\text{ds}_{\eta,j}$  的计算能耗  $w_{\eta,j}$  由式 (24) 计算, 则  $\text{data}_\eta$  和 Data 中全部数据计算所消耗的能耗  $w_\eta$ 、 $W_{\text{MEC}}$  分别由式 (25)、式 (26) 表示, 其中  $p_i$  表示边缘服务器  $i$  计算的功率。

$$w_{\eta,j} = p_i b_{j,i} O^j \frac{\text{ds}_{\eta,j}}{f_i^{\text{mec}}} \quad 1 \leq j \leq \theta, 1 \leq i \leq \omega \quad (24)$$

$$w_\eta = \sum_{j=1}^{\theta} w_{\eta,j} \quad (25)$$

$$W_{\text{MEC}} = \sum_{k=1}^{\eta} w_k \quad (26)$$

数据在传感网和核心网的传输能耗由式 (27) 和式 (28) 计算, 传感节点数据采集的能耗由式 (29) 计算, 则边缘资源选择的总能耗  $W$  由式 (30) 计算。

$$W_{\text{sensor}} = P_s \frac{\sum_{r=1}^{\eta} \text{ds}_r}{C_{\text{sensor}}} \quad (27)$$

$$W_{\text{wireless}} = P_{\text{wireless}} \frac{\sum_{r=1}^{\eta} \text{ds}_r}{C_{\text{wireless}}} \quad (28)$$

$$W_{\text{collect}} = \sum_{i=1}^{\eta} w_i^{\text{sensing}} \quad (29)$$

$$W = W_{\text{sensor}} + W_{\text{wireless}} + W_{\text{MEC}} + W_{\text{collect}} \quad (30)$$

式中,  $W_{\text{sensor}}$  和  $W_{\text{wireless}}$  分别表示密集型数据 Data 在传感网和核心网的传输能耗;  $W_{\text{MEC}}$  和  $W_{\text{collect}}$  表示所有数据的采集能耗和数据在边缘服务器的计算能耗。

## 2 边缘协同优化及求解

### 2.1 问题描述

在分布式边缘协同传感网场景下, 时延和能耗是资源选择中最常用的两个指标, 隐私熵用于衡量资源分配的安全性。本文综合考虑这 3 个因素, 构建联合优化函数如下:

$$\text{Cost} = \alpha T + \beta W + (1 - \alpha - \beta) \frac{1}{K} \quad (31)$$

式中,  $\alpha, \beta \in [0, 1]$  用于调节系统资源选择中对时延、能耗及隐私熵的权衡, 若  $\alpha$  值越大, 系统资源选择对时延越敏感, 若  $\beta$  值越大, 则系统资源选择对能耗越敏感, 若  $1 - \alpha - \beta$  值越大, 则表明系统对数据隐私越敏感。最终, 把边缘协同优化问题转化为系统隐私熵  $K_{\text{req}}$ 、时延  $T_{\text{req}}$ 、能耗  $E_{\text{req}}$  约束下寻找最佳资源选择策略  $\mathbf{Y}^*$  问题, 使得式 (32) 取最小值, 其中  $\mathbf{Y}^*$  为  $\mathbf{Y}$  的最优策略。

$$\mathbf{Y}^* = \arg \min_{\text{s.t.}} (\text{Cost}(\mathbf{Y})) \quad (32)$$

$$K \geq K_{\text{req}} \quad (33)$$

$$E \leq E_{\text{req}} \quad (34)$$

$$T \leq T_{\text{req}} \quad (35)$$

### 2.2 求解方法

针对上述问题, 首先通过 ECANS 方案得到节点选择策略, 并根据该策略采集数据。之后在资源选择阶段使用 DECRS 方案得到条件约束下的最优卸载策略。由于求解策略矩阵  $\mathbf{Y}^*$  是一个 NP 问题, 其凹凸性不确定, 用数学方法无法在有限时间内得到满意解, 本文使用智能启发式蚁群算法求解边缘资源选择策略。但由于传统的蚁群算法每次把数据块分配给大信息素的边缘服务器, 将导致算法停滞, 即算法很难找到最优或近似最优解。

#### 算法 2 DECRS

输入: 密集型数据集 Data, 蚂蚁数量 AntNum, 迭代次数 IterNum, 边缘服务器计算速率集合  $F$ ;

输出: 最大化隐私熵的资源选择策略矩阵  $\mathbf{Y}^*$ ;  
由式 (9) 得到随机分块数  $\theta$ , 并对数据 Data $_{\eta}$  分块;  
初始化资源分配策略矩阵  $\mathbf{Y}$  和信息素矩阵  $\mu$ ;  
for  $i$  从 1 到 IterNum, 执行

for  $j$  从 1 到 AntNum, 执行

由式 (36) 得边缘资源分配方式的临界值  $m$ ;

for  $k$  从 1 到 Len(Data $_{\eta}$ ), 执行

if  $j < m$ , 执行随机分配边缘资源

else

执行最大化分配边缘资源

由式 (14) 计算当前数据隐私熵;

由式 (23) 计算当前数据时延;

由式 (30) 计算当前数据能耗;

由式 (31) 计算当前策略总开销;

end for

end for

生成可能的资源分配策略  $\mathbf{Y}$ ;

更新信息素矩阵  $\mu$ ;

查找迭代资源选择中开销最小的策略, 设置矩阵元素  $b_{j,k}$  为 1;

end for

输出资源分配策略最优矩阵  $\mathbf{Y}^*$ ;

为提高边缘资源选择策略的质量, 本文提出一种改进的蚁群算法, 在  $[1, \text{AntNum}]$  间生成随机整数  $m$ , 如式 (36) 所示, 并命名  $[1, m]$  蚂蚁为“发现者”, 按随机信息素分配边缘资源, 而  $[m, \text{AntNum}]$  蚂蚁为“跟随者”, 按最大化信息素分配边缘资源, 该方法可有效提高解的质量。

$$m = \text{random}(1, \text{AntNum}) \quad (36)$$

## 3 实验与结果分析

### 3.1 实验环境及参数

在仿真实验环境下进行实验, 操作系统为 MAC OS15.6, CPU 为 Intel i9 2.30 GHz, 内存为 16 GB, 硬盘为 512 GB, 编程平台为 Pycharm 2019, 编程语言为 Python。实验包括数据采集和边缘资源选择两部分。根据本文所提方案, 首先在不同区域传感网中分别采用多种方法采集数据, 并分析不同方法的时延与能耗; 然后, 针对密集型数据边缘资源选择问题, 通过 DECRS 方案、基于遗传算法 (genetic algorithm, GA) 方案<sup>[14]</sup>、基于粒子群算法 (particle swarm optimization, PSO) 方案<sup>[15]</sup> 及随机选择 (random resource selection, RRS) 方案, 得到不

同方案资源选择策略的时延、能耗及隐私熵, 并对其进行分析。实验的其他参数见表1。

表1 实验参数

参数	参数范围
边缘服务器数量( $\omega$ )/个	15
请求数据长度( $\tau$ )/个	[5,10]
传感网数量( $\zeta$ )/个	3
每个WSN的节点数量( $\rho$ )/个	[5,10]
边缘服务器计算速率( $F$ )/flops·MB <sup>-1</sup>	[1,3.5]
边缘服务器计算能力( $o$ )/Gflops·s <sup>-1</sup>	1×10 <sup>9</sup>
传感器采集数据量( $DS_{Data}$ )/MB	[1,10]
边缘服务器计算功率( $P_{MEC}$ )/W	[5,10]
无线传感网传输功率( $P_s$ )/W	[0.1,0.5]
无线核心网传输功率( $P_{wireless}$ )/W	[0.5,1]
传感节点感知时延( $t_{sensing}$ )/s	[0.01,0.1]
传感节点感知功率( $W_{sensing}$ )/W	[0.002,0.02]
无线传感器网络传输速率( $C_{sensor}$ )/Mbps	[1,5]
无线核心网传输速率( $C_{wireless}$ )/Mbps	[5,10]
信道高斯白噪声( $\varphi$ )/W	3×10 <sup>-10</sup>
蚁群数量/只	35
迭代次数/次	50

### 3.2 实验结果分析

#### 1) ECANS 方案时延及能耗分析

为验证 ECANS 方案的数据采集性能, 将本文方案与文献 [7] 的有效节点选择方案 (effective node sensing, ENS)、区域节点周期性采集方案 (periodically sensing with all nodes, PSAN) 从时延性能和能耗性能方面进行对比分析。图4为数据采集的时延测试, 实验表明本文 ECANS 方案数据采集时延最短, 且平均采集时延为 1.45 s; ENS 方案更趋向于满足应用请求的节点组合, 而没有考虑冗余数据对数据采集时延的影响, 平均时延为 3.35 s; PSAN 方案采集整个区域节点的数据, 因此数据冗余度更大, 从而导致数据采集时延最大且每次都相同<sup>[16]</sup>, 平均时延为 5.00 s。ECANS 方案较 ENS 和 PSAN, 平均时延分别降低了 56.71% 和 71.00%。

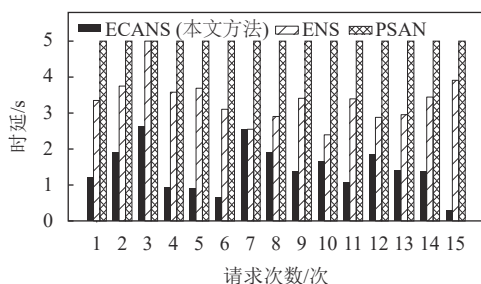


图4 传感节点数据采集时延

图5为能耗对比分析。在多次数据采集中本文所提方法的能耗最小, 为 10.39 J; ENS 方案通过剔除无关节点减少了节点采集数据的数量, 从而降低了数据采集的能耗, 能耗为 24.54 J; 而 PSAN 方案在数据采集时, 需要采集区域内所有节点上传传感器的数据, 这将导致数据采集能耗最大, 为 44.70 J。ECANS 方案较 ENS 和 PSAN, 平均能耗分别降低了 57.66% 和 76.76%。

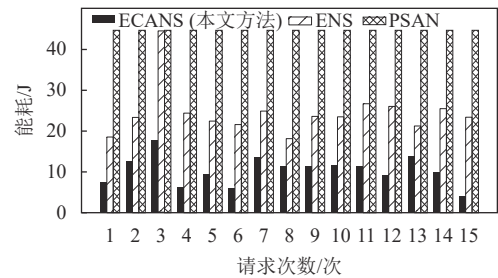


图5 传感节点数据采集能耗

图6为不同数据采集方案的数据质量, 由图可得本文的 ECANS 方案所得的平均数据质量最高, 为 57.66%, 经分析得本方案有效剔除了无关及相关度小的传感节点, 得到最佳的数据采集策略, 从而使数据质量最大化。而 ENS 方案和 PSAN 方案所得平均数据质量分别为 30.40% 和 22.93%, 这是由于采集无关传感节点的数据导致数据质量降低。

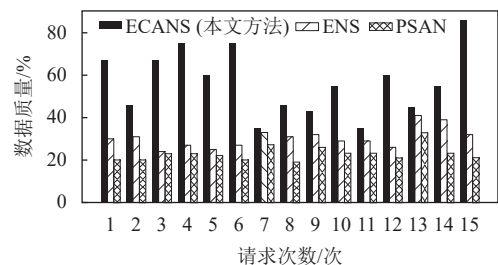


图6 传感节点数据采集数据质量

#### 2) DECRS 方案隐私熵分析

在相同的数据分块方式下, 不同边缘资源选择方案的隐私熵对比如图7所示。图中 RRS 方案的隐私熵在迭代次数为 25 时达到最大 3.90, 由于随机资源选择的不确定性大, 但不能使隐私熵持续最大化。对于其他方案, 随着迭代次数的增加, GA 方案的隐私熵收敛于 2.90, PSO 方案的隐私熵收敛于 3.32, 而 DECRS 方案最终达到 3.83。分析得出 DECRS 方案使用的改进蚁群算法收敛速度快, 不易陷入局部最优, 而 PSO 方案和 GA 方案的核心算法在本场景收敛速度慢、效率低, 更易陷入局部最优。

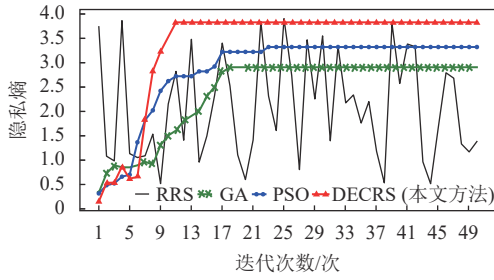


图 7 不同方法的隐私熵

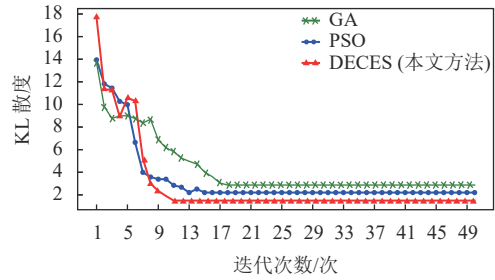


图 9 KL 散度对比

针对不同分块数量对 DECRS 方案隐私熵的影响进行仿真, 结果如图 8 所示。实验中分别设置分块数为 5、10、15 块, 并对其进行边缘资源分配。实验结果表明, 在边缘服务器数量为 15 台的条件下, 随着分块数增加, 隐私熵增大, 分别为 2.29、3.27、3.83。经分析得知, 随着分块数量越接近边缘服务器数, 子数据块被分配到边缘服务器越均匀, 被攻击者得到完整数据的概率越小, 从而数据泄露风险越小<sup>[17]</sup>。

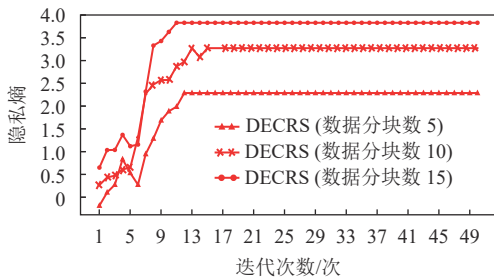


图 8 不同分块方式的隐私熵

为评价隐私保护强度, 本文采用 Kiullback-Leibler(KL)<sup>[18-21]</sup> 衡量不同方案的隐私泄露风险, 其值越低, 表明隐私保护强度越高, 数据越不易泄露。

两个特征集  $\{(x_i, y_i) | x_i, y_i \in R^n\}$  和  $\{(x'_i, y_i) | x'_i, y_i \in R^n\}$  最多有一条记录不同, 采用 KL 散度, 可将不同传输模式下的隐私泄露 CPL 定义为:

$$CPL = \sum_{i=1}^N P\{M(x_i, y_i) = O\} \ln \left( \frac{P\{M(x_i, y_i) = O\}}{P\{M(x'_i, y_i) = O\}} \right) \quad (37)$$

以随机资源选择所得的最大隐私熵作为标准, 分别计算不同方案在每次资源选择的 KL 散度, 如图 9 所示, 随着迭代次数的增加, GA 方案、PSO 方案及 DECRS 方案的 KL 散度都呈下降趋势, 这是因为 3 种方案在资源选择时以最大化隐私熵为目标资源选择。但 3 者达到收敛分别收敛于 2.87、2.20、1.49, 有  $CPL(GA) > CPL(PSO) > CPL(DECRS)$ , 表明 DECRS 方案所得的策略隐私泄露风险较低。

### 3) DECRS 方案时延、能耗分析

不同资源选择方案的时延如图 10 所示, 随着迭代次数的增加, RRS 的时延在 [20.25, 83.34] 内变化, 而随着迭代次数增加, GA、PSO、DECRS 方案的总时延呈现下降趋势, 但 GA 方案和 PSO 方案的收敛效率低, 且分别收敛于 14.57 s 和 12.23 s; DECRS 方案收敛于 9.89 s, 表现最好。经分析, DECRS 方案得到解的效率及精度高, 且可使数据块在最小时延约束下在多个边缘服务器上并行计算, 通过减少计算时延降低资源分配的总时延。

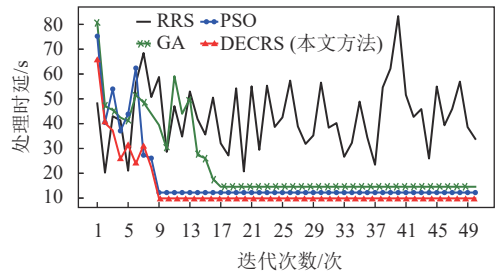


图 10 不同方案的资源选择时延

不同资源选择方案的能耗如图 11 所示, 由于 RRS 资源选择的随机性, 其能耗在 [376.49, 3 765.64] 内变化。而 DECRS 方案收敛于 807.36 J, PSO 方案收敛于 872.99 J, GA 方案收敛于 979.96 J。DECRS 方案较 GA 及 POS 方案, 能耗分别降低了 17.56% 和 7.51%。经分析可得, DECRS 方案收敛速度快, 能找到近似最优解, 而 GA 方案和 PSO 方案得到解的质量相对较差。

考虑不同天线阵列模式对资源选择时延的影响, 将 DECRS 方案在随机天线选择模式 (random antenna selection, RAS)、联合天线模式 (combining selection, CS)、最大传输模式 (maximum antenna transmission, MAT) 下进行对比<sup>[22-23]</sup>。从图 12 可知, 在 3 种天线模式下随着迭代次数增加时延降低, 且 CS 模式收敛于 10.07 s, MAT 模式收敛于 15.20 s, RAS 模式收敛于 17.33 s; CS 模式与 RAS



模式和 MAT 模式相比, 传输时延分别降低了 41.89%、33.75%。经分析得, CS 模式下所有天线被开启并传输数据, 使传输带宽增加, 从而使传输时延降低, 而 RAS 模式和 MAT 模式的速率均小于 CS 模式, 因此, 传输时延相对较大。

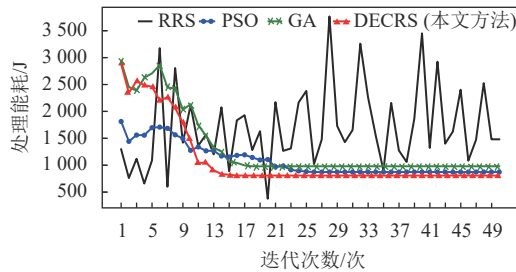


图 11 不同方案的资源选择能耗

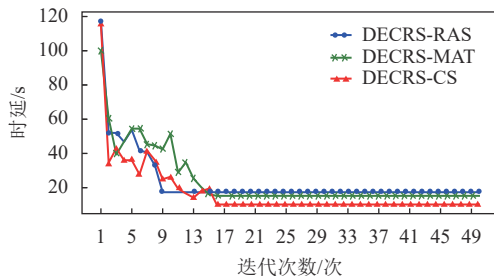


图 12 不同传输模式的资源选择时延

不同天线模式的能耗如图 13 所示。RAS 模式能耗收敛于 923.59 J, MAT 模式能耗收敛于 991.37 J, CS 模式收敛于 1 145.09 J。RAS 模式能耗较 MAT 及 CS, 分别降低了 19.34% 和 6.84%。经分析得, CS 模式下所有天线被开启后, 随着传输速率的增大, 系统能耗也会增大, 而 RAS 模式随机选择单个天线传输, 从而导致总能耗减小。

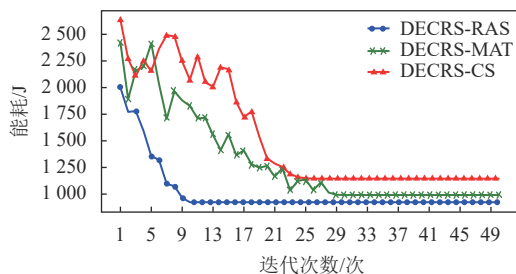


图 13 不同传输模式的资源选择能耗

随机生成权重因子 (本文  $\alpha = 0.4, \beta = 0.3$ ) 的总开销如图 14 所示, 在相同传输模式下, RRS 方案的总开销在 [83.81, 767.22] 范围内变化, 其他 3 种方案的总成本均低于 RRS 方案。经分析得, 资源选择中, DECRS 方案、PSO 方案及 GA 方案考虑

到延迟、能耗及隐私熵倒数的联合最小, 寻找最优的资源选择策略, 且分别收敛于 166.96、180.82 和 202.98。但 DECRS 方案所得的总开销较 PSO 方案及 GA 方案有显著减少。同时, 该参数下联合优化的总开销变化趋势与能耗相似, 这是由于能耗在总开销占的比重较大, 而时延和隐私熵占比较小。

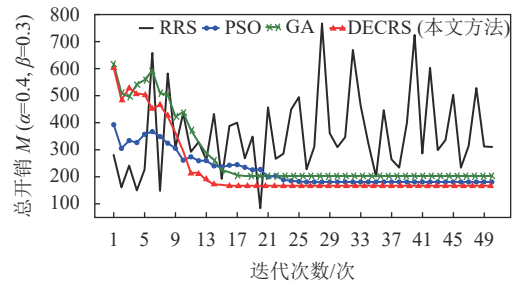


图 14 资源选择联合优化

#### 4) 传感网性能对比

针对本文所提的分布式边缘协同方法对传感网性能提升问题, 从时延及能耗方面进行验证与分析。图 15 针对传感网数据响应时延, 系统中引入 ECANS 和 DECRS 方案后, 传感网响应时延明显降低, 这是由于本文方法在数据采集阶段通过 ECANS 方案得到最优数据采集策略, 从而降低了数据采集时延; 而在计算阶段, 本文 DECRS 方案考虑到用户时延、能耗和隐私熵的要求, 得到三者联合最优策略, 通过降低数据计算的时延和传输时延, 降低传感网响应时延。图 16 为传感网能耗的优化结果, 使用本文方法后, 传感网能耗明显降低。这是由于 ECANS 方案在数据采集阶段减少了传感节点数量, 从而减少数据采集及传输能耗, 导致传感网能耗降低。在节点密集型数据的卸载阶段, 通过 DECRS 方案得到最优的资源选择策略, 并将密集型数据卸载到边缘, 从而使节点端数据量减少, 导致传感网平均时延和能耗分别降低了 46.92% 和 11.26%。

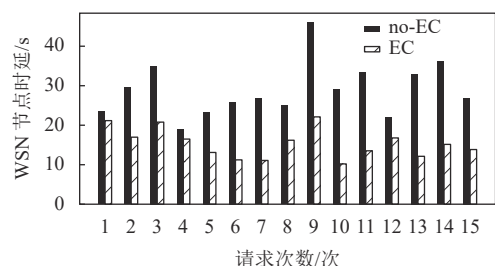


图 15 边缘协同传感网节点时延

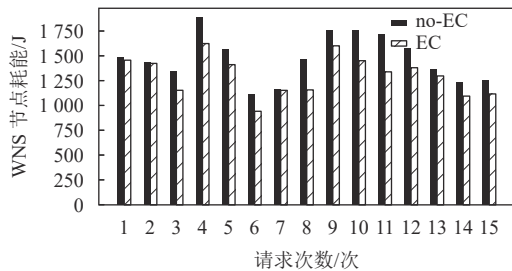


图 16 边缘协同传感网节点能耗

## 4 结束语

本文建立了一种分布式边缘协同传感网安全、节能和高效的资源选择模型。在传感网数据采集阶段,提出一种 ECANS 方案,得到最佳数据采集策略,通过降低数据冗余度,降低节点数据采集的时延与能耗。在密集型数据计算阶段,本文基于启发式智能算法和信息熵理论提出一种分布式边缘协同 WSN 安全、节能及高效的 DECRS 方案,在降低数据泄露风险和能耗的同时提高计算效率。下一步工作,将基于 KubuEdge 及嵌入式硬件实现本文所提系统及方法,并在真实场景部署,验证本文方法的性能。

## 参 考 文 献

- [1] ALHAYANI B, ABBAS S T, MOHAMMED H J, et al. Intelligent secured two-way image transmission using corvus corone module over WSN[J]. *Wireless Personal Communications*, 2021, 5: 1-36.
- [2] GULATI K, BODDU R S, KAPILA D, et al. A review paper on wireless sensor network techniques in internet of things (IoT)[J]. *Materials Today: Proceedings*, 2022, 51: 161-165.
- [3] DWIVEDI R K, KUMARI N, KUMAR R. Integration of wireless sensor networks with cloud towards efficient management in IoT: A review[J]. *Advances in Data and Information Sciences*, 2020, 1: 97-107.
- [4] LIU Q, WANG G, LI F, et al. Preserving privacy with probabilistic indistinguishability in weighted social networks[J]. *IEEE Transactions on Parallel and Distributed Systems*, 2016, 8(5): 1417-1429.
- [5] WU Y, JIANG Z L, WANG X, et al. Dynamic data operations with deduplication in privacy-preserving public auditing for secure cloud storage[C]//2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC). [S.l.]: IEEE, 2017: 562-567.
- [6] WANG T, LU Y, CAO Z, et al. When sensor-cloud meets mobile edge computing[J]. *Sensors*, 2019, 19(23): 5324.
- [7] LI X, ZHU L, CHU X, et al. Edge computing-enabled wireless sensor networks for multiple data collection tasks in smart agriculture[J]. *Journal of Sensors*, 2020, 25: 2020.
- [8] ROBERTO C, JAVIER P, JUAN M, et al. Blockchain framework for IoT data quality via edge computing [C]//Proceedings of the 1st Workshop on Blockchain-

- enable-d Networked Sensor Systems. [S.l.]: ACM, 2018: 19-24.
- [9] GUPTA V, DE S. Energy-Efficient edge computing framework for decentralized sensing in WSN-assisted IoT[J]. *IEEE Transactions on Wireless Communications*, 2021, 20(8): 4811-4827.
- [10] WANG Y, MAN K L, HUGHES D, et al. A rapid response approach applying edge computing for distributed warehouses in WSN[J]. In *Advanced Multimedia and Ubiquitous Engineering*, 2019, 590: 183-189.
- [11] DE ARCANGELIS L, IACONO M, LIPPIELLO E, et al. Towards a multiparadigm approach to model energy management in wsn for Iot based edge computing applications[C]//The 34th International Ecms Conference on Mod-Elling And Simulation (ECMS). [S.l.]: Springer, 2020, DOI: [10.7148/2020-0361](https://doi.org/10.7148/2020-0361).
- [12] XU X, HUANG Q, YIN X, et al. Intelligent offloading for collaborative smart city services in edge computing[J]. *IEEE Internet of Things Journal*, 2020, 7(9): 7919-7927.
- [13] NUMAN M, SUBHAN F, KHAN W Z, et al. A systematic review on clone node detection in static wireless sensor networks[J]. *IEEE Access*, 2020, DOI: [10.1109/ACCESS.2020.2983091](https://doi.org/10.1109/ACCESS.2020.2983091).
- [14] ALHABOB A A, DOBRE O A, ARMADA A G, et al. Task scheduling for mobile edge computing using genetic algorithm and conflict graphs[J]. *IEEE Trans-Actions on Vehicular Technology*, 2020, 69(8): 8805-8819.
- [15] ZHOU W, CHEN L, TANG S, et al. Offloading strategy with PSO for mobile edge computing based on cache mechanism[J]. *Cluster Computing*, 2021, 2: 1-3.
- [16] SHAHRAKIA, TAHERKORDIA, ELIASSENF. Clustering objectives in wireless sensor networks: A survey and research direction analysis[J]. *Computer Networks*, 2020, 180: 107376.
- [17] GUO Y, ZHAO Z, ZHAO R, et al. Intelligent offloading strategy design for relaying mobile edge computing networks[J]. *IEEE Access*, 2020, 8: 35127-35135.
- [18] ZHAO H, XIAO M J, et al. Differentially private unknown worker recruitment for mobile crowdsensing using multi-armed bandits[J]. *IEEE Transactions on Mobile Computing*, 2021, 20(9): 2779-2794.
- [19] FAROKHI F, KAAFAR M A. Modelling and quantifying membership information leakage in machine learning[EB/OL]. [2021-10-13]. <https://arxiv.org/pdf/2001.10648v2.pdf>.
- [20] XU W, LIU J, JIN S, et al. Spectral and energy efficiency of multipair massive MIMO relay network with hybrid processing[J]. *IEEE Transactions on Communications*, 2017, 65(9): 3794-3809.
- [21] 彭长根, 丁红发, 朱义杰, 等. 隐私保护的信息熵模型及其度量方法[J]. *软件学报*, 2016, 27(8): 1891-1903.
- [22] PENG C G, DING H F, ZHU Y J, et al. Information entropy model of privacy protection and its measurement method[J]. *Journal of Software*, 2016, 27(8): 1891-1903.
- [22] ZHAO J H, YANG T, GOPNG Y, et al. Power control algorithm of cognitive radio based on non-cooperative game theory[J]. *China Communications*, 2013, 10(11): 143-154.
- [23] NI S, ZHAO J, YANG H H, et al. Enhancing downlink transmission in MIMO HetNet with wireless backhaul[J]. *IEEE Transactions on Vehicular Technology*, 2019, 68(7): 6817-6832.