A label knowledge graph powered multi-task framework for crowdsourcing and mobile crowd sensing tasks

Published Version:
Liu, Y., Yu, Z., Li, N., Guo, B., Helal, S. (2025). A label knowledge graph powered multi-task framework for crowdsourcing and mobile crowd sensing tasks. EXPERT SYSTEMS WITH APPLICATIONS, 270, 1-15 [10.1016/j.eswa.2025.126562].

Availability:
This version is available at: https://hdl.handle.net/11585/1051850 since: 2026-02-27

Published:
DOI: http://doi.org/10.1016/j.eswa.2025.126562

A Label Knowledge Graph Powered Multi-task Framework for Crowdsourcing and Mobile Crowd Sensing Tasks

Yimeng Liu[a] (ymliu@cs.ecnu.edu.cn), Zhiwen Yu[b,c] (zhiwenyu@nwpu.edu.cn),

Nuo Li[d] (linuo@mail.nwpu.edu.cn), Bin Guo[c] (guob@nwpu.edu.cn),

Sumi Helal[e] (sumi.helal@unibo.it)

[a] School of Computer Science and Technology, East China Normal University, Shanghai, 200062, China

[b] School of Computer Science, Harbin Engineering University, Harbin, 150006, Heilongjiang, China

[c] School of Computer Science, Northwestern Polytechnical University, Xi'an, 710072, ShaanXi, China

[d] School of Computer Science, Fudan University, Shanghai, 200438, China

[e] Department of Computer Science and Engineering, University of Bologna, Bologna, 40126, Italy

Corresponding author: Yimeng Liu

Email: ymliu@cs.ecnu.edu.cn

Address: School of Computer Science and Technology, East China Normal University, Shanghai, 200062, China.

A Label Knowledge Graph Powered Multi-task Framework for Crowdsourcing and Mobile Crowd Sensing Tasks

Yimeng Liu[a,*], Zhiwen Yu[b,c,*], Nuo Li[d], Bin Guo[c] and Sumi Helal[e]

[a]*School of Computer Science and Technology, East China Normal University, Shanghai, 200062, China*
[b]*School of Computer Science, Harbin Engineering University, Harbin, 150006, Heilongjiang, China*
[c]*School of Computer Science, Northwestern Polytechnical University, Xi'an, 710072, ShaanXi, China*
[d]*School of Computer Science, Fudan University, Shanghai, 200438, China*
[e]*Department of Computer Science and Engineering, University of Bologna, Bologna, 40126, Italy*

ARTICLE INFO

Keywords:
Heterogeneous Label Knowledge Graph
Multi-task Hybrid Reasoning
Crowdsourcing
Mobile Crowd Sensing Tasks

ABSTRACT

Crowdsourcing and Mobile Crowd Sensing (MCS) platforms have revolutionized data collection, harnessing the collective intelligence of crowdsourced sensing. Accurately classifying and extracting core information such as heterogeneous sensors in MCS tasks plays a key role in the platform execution efficiency. However, existing methods struggle with extracting pivotal information from task descriptions that are open-domain, implicitly expressed, and linguistically diverse, ultimately hindering the efficiency of task assignment and execution. To overcome these challenges, we propose LKG-MF, a Label Knowledge Graph-powered Multi-task Framework, to achieve better core information mining performance in crowdsourcing and mobile crowd sensing tasks. Specifically, we first construct an MCS task dataset comprising over 10,000 real tasks from 7 platforms. Then we devise a label knowledge graph to capture heterogeneous semantics and relationships among labels and enhance label representation. Further, we present a multi-granularity feature extraction network to capture precise task-specific features. To optimize performance across disparate tasks, we incorporate a task-adaptive loss function that adeptly balances their optimization rates. Experimental results show that LKG-MF outperforms baselines average by 2.3%, significantly improving multi-task classification accuracy. Notably, when we integrate the LKG-MF model into MCS platforms, the task assignment efficiency is improved by 38.6% and the task completion time is reduced by 45.1%, which demonstrates the practical impact and effectiveness of our model in improving the performance of MCS platforms.

1. Introduction

In recent years, with the popularity of portable terminals and sensing devices, crowdsourcing ideas (Bazaluk et al., 2024) and mobile crowd sensing technology (Suhag and Jha, 2023) have developed rapidly. By publishing tasks on the Crowdsourcing and Mobile Crowd Sensing (MCS) platforms and recruiting the public to execute them, it helps to obtain urban data at low cost, high efficiency, and high coverage, and effectively solve the issues of municipal and livelihood (Wang et al., 2024), environmental monitoring (Chen et al., 2023; El Hafyani et al., 2024), public safety (Liu et al., 2023) and other urban issues. However, most existing MCS platforms struggle with inadequate analysis and comprehension of task nuances, obstructing effective task classification (Tian et al., 2023) and heterogeneous core information mining, further leading to low efficiency in platform task assignment and data collection quality. Therefore, this article aims to apply Natural Language Processing (NLP) and Knowledge Graph (KG) technologies (Zhou et al., 2024; Fu et al., 2024; Liang et al., 2021; Sun et al., 2022; Li et al., 2025) to achieve better reasoning performance in crowdsourcing and mobile crowd sensing tasks.

For crowdsourcing and mobile crowd sensing tasks (which are uniformly named MCS tasks or MCST), the Task Type (TP) and Crowd sensing Agents (CAs) are the core information needed to identify. The task type affects the user's interest in the task and willingness to participate, while the crowd sensing agents dictate whether individuals possess the necessary prerequisites to contribute data. Nevertheless, mining TP and CAs from open-domain task descriptions (as illustrated in Fig.1) in MCS poses a pivotal challenge. This challenge manifests primarily in three aspects: 1) **The diversity of TP and CAs representation**. MCST comes from diverse task publishers, where individuals exhibit unique styles and preferences in describing tasks. Consequently, identical TP and CAs may be phrased differently, leading to a variety of expressions. Moreover, the expression of TP and CAs is often indirect and implicit, necessitating

*Corresponding author

ymliu@cs.ecnu.edu.cn (Y. Liu); zhiwenyu@nwpu.edu.cn (Z. Yu); linuo@mail.nwpu.edu.cn (N. Li); guob@nwpu.edu.cn (B. Guo); sumi.helal@unibo.it (S. Helal)

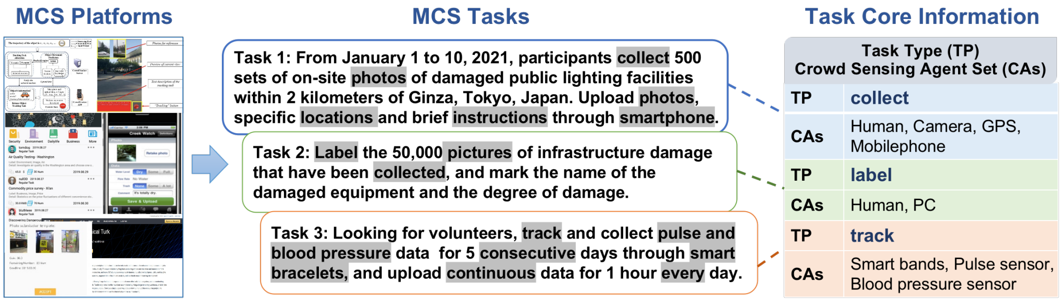


Figure 1: MCST Description and Core Information.

contextual inference and analysis of their inherent attributes. For instance, for Task 1 in Fig.1, the 'camera'(CA) is not stated explicitly but inferred from the term 'photo'. 2) **The ambiguity and interference of TP and CAs representation**. For example, for Task 2 in Fig.1, if there are both 'collect' and 'label' words, which one is the true TP category, we need to combine multiple factors such as TP trigger word context and local structure to determine. 3) **The diverse relationships among TP and CAs**. For example, there is an inter-layer correlation between TP and CAs, and an intra-layer correlation between CAs. For instance, in Fig.1, the similarity between 'PC' and 'Mobilephone' is greater than the similarity between 'PC' and 'Human'. 4) **Lack of ubiquitous datasets in the MCS field**. The non-uniformity and specificity of platforms pose significant challenges in aggregating MCS task data, exacerbating the scarcity of comprehensive datasets that encompass tasks from multiple MCS platforms. This scarcity hinders advancements in research aimed at universal MCS task analysis.

The prevailing approaches for extracting core information in MCS tasks can be broadly categorized into two groups: label-enhanced single-task methods and multi-task methods. The label-enhanced single-task methods enhance information mining accuracy by uncovering associations among implicit data and labels (Zhang et al., 2023; Li et al., 2024; Sun et al., 2024). Nevertheless, these techniques predominantly focus on pre-trained label semantics, and label co-occurrence frequencies, overlooking semantic topologies and intricate relationships between labels, thus struggling with the ambiguity problem outlined in challenges 1 and 3. On the other hand, the label-enhanced multi-task methods (Ma et al., 2023) target similar tasks in nature, difficulty, and scale, making them suitable for mining heterogeneous tasks like TP and CAs that exhibit substantial differences in complexity. However, they disregard the unique attributes of MCS tasks, such as varied expressions and semantic ambiguities, limiting their ability to address the diversity and ambiguity issues presented in challenges 1 and 2. In summary, the existing methods fall short in addressing the multifaceted challenges faced by current MCS platforms in task comprehension.

To address the above challenges, this article proposes a Label Knowledge Graph-powered Multi-task Framework, LKG-MF, to achieve better reasoning performance in MCS tasks. Specifically, we first construct a Label Knowledge Graph incorporating Heterogeneous Semantics and Relationships (HSR-LKG) to deeply understand the implicit expression and diverse description of tasks and accurately classify them. Then we introduce a label-guide multi-granularity feature extraction approach to extract fine-grained features of tasks at multiple granularities and incorporate semantic correlations between labels to obtain more precise task-specific features. Subsequently, we design a task-adaptive multi-task loss method to balance the optimization speed of different tasks and improve the efficiency of the multi-task model. Finally, to verify the proposed LKG-MF and promote the development of MCS platforms, we also construct a MCS task dataset (CrowdTask dataset (Liu et al., 2024)). It contains a total of 10,000+ tasks from 7 MCS platforms, including 4 task types and 14 types of crowd sensing agents. For more details about the dataset, please refer to Section 3 and Section 6. The main contributions of this article can be summarized as follows:

- We introduce a label knowledge graph-powered multi-task framework, LKG-MF, to achieve better classification performance in MCS tasks.
- We devise a label heterogeneous knowledge graph, HSR-LKG, to capture heterogeneous semantics and relationships among labels and enhance label representation.
- We present a label-guide multi-granularity feature extraction method to capture precise task-specific features, and propose a task-adaptive loss function to enhance the performance across multiple tasks.

- Experimental results demonstrate that LKG-MF outperforms baselines. Notably, when integrated into MCS platforms, LKG-MF leads to a 38.6% improvement in task assignment efficiency, validating its effectiveness and practical applicability.

2. Related Work

In this section, we analyze existing crowdsourcing and mobile crowd sensing systems and platforms, compare how current MCS platforms handle tasks, and summarize the current semantic reasoning methods.

2.1. Task Understanding in MCS Platforms

The existing MCS platform allocates tasks reasonably and efficiently based on the current task type and the user's capabilities (such as microphones, cameras, and other sensors). For example, for tasks such as sensory data collection, participants need to meet the sensor requirements proposed in the task description. However, most platforms, such as AMTurk (Buhrmester et al., 2011), FlierMeet (Guo et al., 2015), WeSense (Liu et al., 2020), MPS (Mo et al., 2024), Food Rescue (Shi et al., 2021), just play the role of a bulletin board to display tasks for participants to choose, which leads to low quality of collected data. Only a small number of research platforms take into account participant factors in task assignment, such as space-time trajectory (Wang et al., 2019b), location (Xie et al., 2023) and energy consumption (Xiang et al., 2023), etc., to increase the probability of participants performing MCST. None of these platforms do an in-depth analysis and understanding of the task requirements (Zhao et al., 2022), and do not consider whether the interests and skills of the participants match the tasks. Only by increasing the bilateral matching between tasks and participants can the data collection rate and result quality be improved. Therefore, it is urgent to find an efficient task core information mining method.

2.2. Semantic Reasoning Methods

This article aims to gain a deeper understanding of the task characteristics in MCS platforms, to achieve a reasonable and efficient semantic reasoning. The essence of this process lies in accurately identifying the key task information such as task types and sensors in MCS tasks, to provide a solid foundation for subsequent task management and resource allocation. Therefore, this section comprehensively outlines the current semantic reasoning methods, aiming to discover effective technical methods suitable for MCS task classification, and provide theoretical support and practical guidance for building a precise and efficient MCS semantic reasoning system.

The label-enhanced semantic reasoning method can achieve better classification performance (Feng et al., 2024; Liu et al., 2025), and the current label-enhanced semantic reasoning method can be split into two categories: label-enhanced single-task methods and multi-task methods. 1) The label-enhanced single-task methods enhance information mining accuracy by uncovering associations among implicit data and labels (Zhang et al., 2023; Li et al., 2024; Sun et al., 2024; Liang et al., 2024), and then mined features are input into the ST_a and ST_b classifiers, respectively. For example, Zhang et al. (Zhang et al., 2021) design auxiliary-label co-occurrence prediction tasks to enhance label correlation learning. CBi-Attention (Lu et al., 2023) utilizes CNN and BiLSTM to extract global semantics, and model the Attentive mechanism to select the most pertinent features for each label, achieving superior classification performance. LSAE (Wang and Gao, 2024) leverages the graph convolution network to capture the relationships among labels, and then model the multi-dimension semantic interactions between these labels and text features, to enhance the representation of the text, ultimately boosting the text classification performance. 2) The label-enhanced multi-task methods, on the basis of mining associations among implicit data and labels, additionally consider the relevant information(features, labels, etc.) of other tasks to improve the performance of the target task. For example, Shen et al. (Shen et al., 2022) focus on multi-task classification issues, where related classification tasks share the same label space and can be learned simultaneously. MHCAN (Lu et al., 2022) utilizes the cross-attention method to model the text and hierarchical labels, and subsequently leverages the iterative hierarchical-attention module to capture the relationships among levels, enhancing classification performance. And LED (Ma et al., 2023) employs a hierarchy-aware attention mechanism to represent label semantics, and leverages multi-task learning to effectively capture label correlation, enhancing the classification accuracy.

To sum up, the current MCS platforms do not have sufficient in-depth processing of crowd tasks and lack analysis and understanding of the tasks. Because the platform does not conduct in-depth analysis and reasoning on the core information of the tasks that affect the efficiency of task assignment and the quality of the result data, it is difficult to efficiently and accurately match the task to the appropriate participant. Based on the development of current technology,

Table 1
Frequently Used Notations

Notation	**Description**
MCST	Crowdsourcing and Mobile Crowd Sensing Tasks
TP	MCS Task tyPe
CAs	Crowd sensing Agent set
HSR-LKG ($\mathcal{G}$)	Label Knowledge Graph incorporating Heterogeneous Semantics and Relationships
LKG-MF	Label Knowledge Graph-powered Multi-task Framework
E, R	The set of entities; The set of relationships between entities
e_{ls}, e_{sa}, e_{fa}	Label semantic entity; Entity synonym attribute; Function and purpose attribute
$en_1, en_2, \dots, en_L$	The heterogeneous LKG nodes
$\mathbf{en}_u^{(l)}$	Final representation of node u, where l denotes the number of layers updated
$\mathbf{en}_u^{r\prime(l)}$	The aggregation characterization of node u under the condition of relationship r
$\mathbf{en}_u^{r_{1-3}\prime(l)}$	Representation of the label nodes under three different relationship conditions
ST_a, ST_b	The multi-class classification task; The multi-label classification task
$\mathcal{N}_u^r$	Neighbors obtained by sampling node u using the relationship r
γ_i, $\mathbf{h}_i$, $\mathbf{W}_c$	Weight of the word i in MCS task; Latent representation of i; Trainable parameters
V_B, V_C	Sentence-level feature representation; Local phrase structure feature representation
$\mathbf{Z}^{a,b}$	The final private feature representation for the task ST_a and ST_b

some methods can assist in the mining of core task information. However, due to the lack of annotated MCS task datasets, model design and training are difficult to achieve. Secondly, for MCS tasks, existing semantic analysis and reasoning methods cannot establish the correlation between heterogeneous core information, and it is difficult to achieve joint reasoning on the core information of the tasks. Therefore, it is necessary to construct a universal MCS dataset and find a way to conduct unified analysis and reasoning on the MCS tasks in each MCS platform or system, to mine the core information that improves task assignment efficiency.

3. Problem Definition

In this section, we mainly introduce the definition of mobile crowdsensing tasks, the construction of a label knowledge graph incorporating heterogeneous semantics and relationships, and our research objectives. The frequently used notations are illustrated in Table 1 .

Definition 1: Mobile Crowd Sensing or Crowdsourcing Tasks (MCST). The tasks published by task publishers in the platform are usually presented in the form of natural language descriptions, that is $RT = (w_1, w_2, ..., w_n)$, where w_i represents a word, n is the number of words in task descriptions. As shown in Fig.1, MCST contains core elements such as task type and crowd sensing agent set directly or implicitly.

We collect more than 10,000 published MCST from 7 public MCS platforms (Section 5.1). The task assignment (Wang et al., 2018) plays a key role in MCS platforms, and core elements that affect task assignment are the type of task (TP) and the crowd sensing agent set (CAs). According to the MCST analysis in multiple MCS platforms, this article locates the TP in 4 categories, that is, $Label_1$={*Collect, Label, Survey, Track*}, and locates CAs in 14 categories, i.e., $Label_2$ ={*Human, PC, Mobile Phone, Smart bands, UAV, Mobile Robot, GPS, Camera, Microphone, Acceleration sensor, Pulse sensor, Blood pressure sensor, Air quality sensor, Water quality sensor*}. CAs can be defined as 2 levels, where the first level is sensors with comprehensive computing capabilities, such as the 1st to 6-th CAs in $Label_2$, and the second level is specific sensors, such as the 7-th to 14-th CAs in $Label_2$.

Definition 2: Label Knowledge Graph incorporating Heterogeneous Semantics and Relationships (HSR-LKG). HSR-LKG ($\mathcal{G}$) includes the label entity set and the relationships between entity elements obtained by abstracting and modeling MCST labels, such as semantic entities, attribute entities, and heterogeneous relationships between entity elements. We set E to be the set of entities, and R to be the set of relationships between entities. The information in E includes: label semantic entity (e_{ls}), entity synonym attribute (e_{sa}), and entity function and purpose attribute (e_{fa}). The information in R includes three types: r_1 - the relationship between the semantic entity and its attributes, where attributes include synonym attributes and functional attributes. r_2 - the relationship between $Label_1$ entity elements and $Label_2$ entity elements; r_3 - the relationship between elements inside the $Label_2$ entity set. Let $\{(e^1_{ls}, r, e^2_{ls}) | e^1_{ls}, e^2_{ls} \in E, r \in R\}$ be a collection of triplets in $\mathcal{G}$, where each (e^1_{ls}, r, e^2_{ls}) triplet indicates that there is some associations between entity e^1_{ls} and entity e^2_{ls}. Therefore, HSR-LKG can capture and express more deep implicit and explicit information.

Research Objectives: 1) Construct the label knowledge graph HSR-LKG to express the heterogeneous semantics and hierarchical relationships between label entities. 2) Based on the constructed HSR-LKG, to design a hybrid reasoning neural network to accurately and quickly mine two types of implicit information contained in MCST, namely the type of task (TP) and the type of crowd sensing agent set (CAs). 3) Integrate and verify the model on MCS platforms.

4. Methodology

We design a Multi-task Hybrid Reasoning Model integrating HSR-LKG (LKG-MF), which consists of four parts (see Fig. 2). 1) The Task and Label Preprocessing module, which aims to obtain a low-dimension vector representation of the MCST tasks and label element sets. 2) HSR-LKG Construction module. This knowledge graph not only adds attribute information such as corresponding synonyms and usage methods to label semantic entities, but also covers heterogeneous and diversified relationships between label entity elements. Compared with the traditional method of constructing knowledge graphs only by co-occurrence frequency, HSR-LKG can mine the finer-grained and deeper semantic information and relationships of labels, which is more suitable for open-domain scenarios. 3) Label-guide Multi-granularity Feature Extraction module. It aims to utilize the label semantic to guide the feature extraction process, and more accurately mine the private features of two tasks ST_a and ST_b. 4) Task-adaptive Multi-task Loss module. Considering the difference in the difficulty of the two tasks, this module aims to better balance the optimization speed of different tasks and improve the efficiency of multi-task models.

4.1. Task and Label Preprocessing

Task description preprocessing. The BERT (Bidirectional Encoder Representations from Transformers) pre-training model incorporates a large amount of contextual semantic information and plays a huge role in the NLP field (Devlin et al., 2019). Therefore, after performing basic data cleaning, we choose the BERT pre-training model to preprocess the text description RT of MCST. That is, $RT = \{x_1, x_2, \ldots, x_R\}$ is represented as $ET = \{ex_1, ex_2, \ldots, ex_R\}$, where R is the number of words in each MCST after compensation, $ex_i \in \mathbb{R}^{1\times d}$, and d=768.

Label information preprocessing. In the process of processing MCST tasks, there are relationships between the labels of different tasks, and relationships between the labels within the task. The traditional data label processing methods simply map data and labels, ignoring the semantic topology and deep associations between labels. Some studies use the label co-occurrence structure to simply model relationships between labels. These methods are too simple to comprehensively learn and represent the semantic information of MCST labels. To mine the potential implicit semantic association of labels, we construct HSR-LKG to represent and model label entities. In the preprocessing stage, we use the BERT pre-training model to process the words, phrases, and sentences involved in the label element set, and obtain the initial label element set vector representation V_{label_set}. Then use the method introduced in Section 4.2 to construct the characterization vector $\mathcal{G} = \{en_1, en_2, \ldots, en_L\}$ of the heterogeneous LKG nodes, where L represents the number of heterogeneous LKG nodes (that is, the total number of labels), $en_i \in \mathbb{R}^{1\times d}$, and d=768.

4.2. HSR-LKG Construction

This knowledge graph not only adds attribute information such as corresponding synonyms and usage methods to label semantic entities, but also covers heterogeneous and diversified relationships between label entity elements. Compared with the traditional method of constructing knowledge graphs only by co-occurrence frequency, HSR-LKG can mine the finer-grained and deeper semantic information and relationships of labels, which is more suitable for open-domain scenarios.

In the MCST text descriptions of the open-domain scenarios, there are relationships between the labels of different task types (ST_a and ST_b), and there are also relationships between the different labels in ST_b. Therefore, we need to analyze the heterogeneous hierarchical associations between multi-task labels in MCST, and construct a knowledge

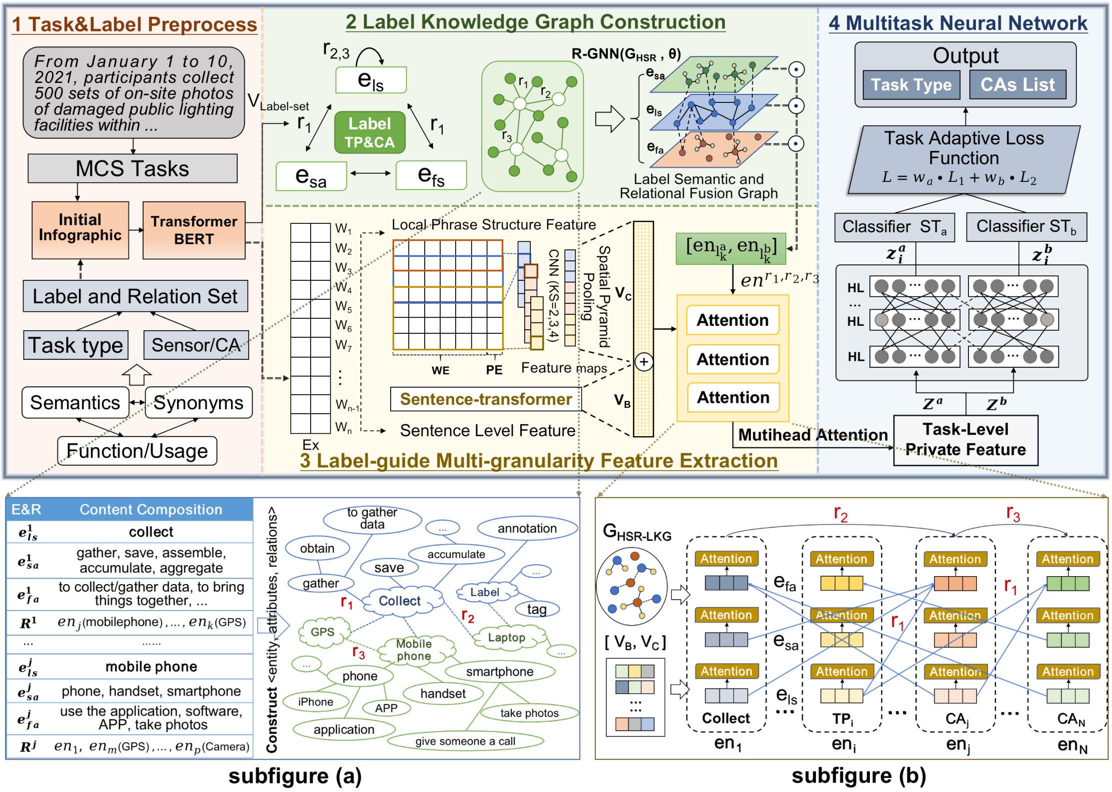


Figure 2: Multi-task Hybrid Reasoning Model Integrating HSR-LKG. (subfigure (a) An example of a label knowledge graph with heterogeneous semantics and relationships. subfigure (b) Label-guide subtask private feature extraction with multi-head attention mechanism)

graph that incorporates label hierarchical relationships and various attributes, to accurately mine the explicit and implicit relationships between labels.

The overall topology structure and content of HSR-LKG are shown in Fig.2 subfigure (a). R-GNN in Fig. 2 means the abstract representation of HSR-LKG. To better aggregate and update node u, the aggregation characterization of node u under the condition of relationship r is as follows:

$$\mathbf{en}_u^{r\prime(l)} = \frac{1}{K}\sum_{k=1}^{K}\sum_{i=1}^{|\mathcal{N}_u^r|} \alpha_i^{rk(l)} \cdot \mathbf{en}_i^{(l-1)}, \qquad \alpha_i^{rk(l)} = \frac{\exp\left(\mathbf{W}_a^{rk}\left(\mathbf{en}_i^{(l-1)} \oplus \mathbf{en}_u^{(l-1)}\right)\right)}{\sum_{j=1}^{|\mathcal{N}_u^r|}\exp\left(\mathbf{W}_a^{rk}\left(\mathbf{en}_j^{(l-1)} \oplus \mathbf{en}_u^{(l-1)}\right)\right)} \tag{1}$$

Where l denotes the number of layers updated, $\alpha_i^{rk(l)}$ indicates the weight of the node i obtained by the k-th attention mechanism during the l-th layer update under the condition of relationship r, $\mathcal{N}_u^r$ represents the neighbors obtained by sampling node u using the relationship r, $\mathcal{N}_u^r = \{(v) \mid (u, r, v) \in \mathcal{G}\}$, $\mathbf{W}_a^{rk}$ denotes the trainable parameters, and $\oplus$ represents the concatenation operation.

Finally, we will get the vector representation of the label nodes under three different relationship conditions, $\mathbf{en}_u^{r_1\prime(l)}$, $\mathbf{en}_u^{r_2\prime(l)}$ and $\mathbf{en}_u^{r_3\prime(l)}$. Different types of relationships have different weights for the vector representation of each label node. Therefore, when constructing label nodes, we need to weight and sum the characterization results under different relationships based on the target-guided adaptive weighting method to obtain the final representation, namely:

$$\mathbf{en}_u^{(l)} = \sum_{i=1}^{3}\beta_i^{(l)} \cdot \mathbf{en}_u^{r_i\prime(l)}, \qquad \beta_i^{(l)} = \frac{\exp\left(\mathbf{W}_b\left(\mathbf{en}_u^{r_i\prime(l)} \oplus \mathbf{en}_u^{(l-1)}\right)\right)}{\sum_{j=1}^{3}\exp\left(\mathbf{W}_b\left(\mathbf{en}_u^{r_j\prime(l)} \oplus \mathbf{en}_u^{(l-1)}\right)\right)} \tag{2}$$

Where $\beta_i^{(l)}$ denotes the weight of the relationship i when updating the layer l, and $W_b \in \mathbb{R}^{1\times 2d}$ represents the trainable parameters.

Here are more details of the graph construction process. We abstractly define HSR-LKG, specifying node entities, attributes, and three heterogeneous relations R among nodes. The method based on our proposed R-GNN can better capture the complex, implicit set of triples. We represent label graph nodes and R through operations such as network embedding, aggregation, and update. We utilize pre-trained models to obtain initial graph vector representations and relational topologies. As a neighbor aggregation strategy, R-GNN can simultaneously encode node structure, attributes, and relationships. The node vector representation is calculated through cyclic aggregation and transmission, and the node aggregation method can refer to Eq. (1) and Eq. (2). Multi-head attention calculation is performed on the label node vector obtained by R-GNN and the multi-granularity fusion feature matrix, and then the input label-guided task private features of the multi-task network are obtained.

4.3. Label-guide Multi-granularity Feature Extraction

MCS tasks are described and uploaded to the platform through personalized user descriptions, resulting in personalized language descriptions. To accurately express the task, we extract the features of different granularities of MCST: 1) Phrase-level features. 2) Sentence-level features. 3) Label-guide task private features.

Local Phrase Structure Feature. The local phase structure feature of sentences plays a huge role in the task description. For example, when there are *collect picture* and *label the collected picture* in two tasks, although the descriptions both contain the words *collect*, they belong to different TP. The former represents the TP->$collect$, while the latter is the TP->$label$. Therefore, it is necessary to combine local information for task representation. Convolutional neural network (CNN) (Albawi et al., 2017) can automatically combine and filter N-gram features to obtain semantic information of different granularity, so we utilize CNN to learn N-gram information of task texts. Specifically, we set the convolutional kernel sizes 2,3 and 4 to extract bi-gram, tri-gram, and four-gram information, respectively. Then, we set the number of convolutional kernels to all 3 and apply the max pooling method to obtain the final vector representation of the task text: $V_C = ||_{i=1}^{3} CNN\,[f(i=2,3,4)]$, $V_C \in \mathbb{R}^{9\times d}$.

Sentence-level Feature. It is mainly used to represent the syntactic information of the task and the spatial structure characteristics of the sentence. Since BERT (Devlin et al., 2019) only relies on position embedding to tell the model the position information of the input token, it weakens the position information between words. It can be seen from Fig.1 that although the word $collect$ appears in both $Task$ 1 and $Task$ 3, $Task$ 3 also contains prominent persistence information such as $track$, $every$, and $continuous$. Therefore, the vector representation of $collect$ in the two tasks is different. Since the contextual content and word order information are crucial to the task representation, we use the BiLSTM network (Huang et al., 2015) based on the BERT pre-trained word embeddings to learn the dependencies on the task sequence, i.e. Eq. (3),

$$
\begin{aligned}
i_t &= \sigma\left(W_{i_1} ex_i + W_{i_2} h_{t-1} + b_i\right) \\
f_t &= \sigma\left(W_{f_1} ex_i + W_{f_2} h_{t-1} + b_f\right) \\
c_t &= i_t \tanh\left(W_{c_1} ex_i + W_{c_2} h_{t-1} + b_c\right) + f_t c_{t-1} \\
o_t &= \sigma\left(W_{o_1} ex_i + W_{o_2} h_{t-1} + b_o\right) \\
h_t &= o_t \tanh\left(c_t\right)
\end{aligned}
\tag{3}
$$

Where i_t, o_t and f_t denote the input, forget and output gate of the t-th word respectively. W_{i_1}, W_{i_2}, W_{f_1}, W_{f_2}, W_{c_1}, W_{c_2}, W_{o_1} and W_{o_2} are weight matrices. b_i, b_f, b_c and b_o are bias vectors. c_t and h_t are cell state and hidden state respectively. Then the hidden state of BiLSTM on i-th word is as: $h_i' = \left[h_i \oplus \overleftarrow{h_i}\right]$, where $\oplus$ denotes the addition of corresponding elements, and $\overleftarrow{h_i}$ denotes the hidden representation of i-th of the reverse LSTM. Consequently, the final representation can be shown as $V_B = ||_{i=1}^{R} h_i'$, $V_B \in \mathbb{R}^{2R\times d}$, h_i' denotes representations of words.

Label-guide Task Private Feature. This module aims to calculate the task private feature $\mathbf{Z}^a$ and $\mathbf{Z}^b$ for ST_a and ST_b based on the constructed HSR-LKG and achieve V_B and V_C. Different tasks and task labels have different concerns, that is, they have different weights for different features. Therefore, we leverage a label-guide attention mechanism to handle two different granularity tasks, guiding and strengthening the extraction and fusion representation of private features through the semantics and relationships of heterogeneous labels. Specifically, we first can obtain the label representation $\mathbf{en}_{l_k^a}$ based on the constructed HSR-LKG, then apply it to achieve the sentence-level feature $\mathbf{V}_B^{l_k^a}$, i.e. Eq.

(4), and further apply it to obtain the final private feature representation $\mathbf{Z}^a$ for the task ST_a, i.e. Eq. (5).

$$\mathbf{V}_B^{l_k^a} = \sum_{i=1}^{2L} \gamma_i^a \cdot \mathbf{h}_i, \quad \gamma_i^a = \frac{\exp\left(\mathbf{W}_c\left(\mathbf{h}_i \oplus \mathbf{en}_{l_k^a}\right)\right)}{\sum_{j=1}^{R} \exp\left(\mathbf{W}_c\left(\mathbf{h}_j \oplus \mathbf{en}_{l_k^a}\right)\right)} \tag{4}$$

Where γ_i^a denotes the weight of the word i in MCST descriptions, $\mathbf{h}_i$ denotes the latent representation of the word i, and $\mathbf{W}_c$ represents the trainable parameters.

Therefore, the private feature $\mathbf{Z}^a$ of the task ST_a can be expressed as:

$$\mathbf{q} = \left[\mathbf{V}_B^{l_k^a}, \mathbf{V}_C\right], \quad \mathbf{Z}^a = \sum_{i=1}^{10} \delta_i^a \cdot \mathbf{q}_i, \quad \delta_i^a = \frac{\exp\left(\mathbf{W}_d\left(\mathbf{q}_i \oplus \mathbf{en}_{l_k^a)}\right)\right)}{\sum_{j=1}^{10} \exp\left(\mathbf{W}_d\left(\mathbf{q}_j \oplus \mathbf{en}_{l_k^a}\right)\right)} \tag{5}$$

Where q denotes the fused representation incorporating the multi-scale CNN features and label-guide Bi-LSTM features, δ_i^a represents the weight of the representation i.

For the task ST_b, since each MCST task corresponds to a variable number of CA entities, we first use a neural network to obtain a label semantic representation, i.e:

$$\mathbf{en}_{l_k^b} = \sigma\left(\mathbf{W}_f \frac{1}{\left|l_k^b\right|} \sum_{i=1}^{\left|l_k^b\right|} \mathbf{en}_i\right) \tag{6}$$

The task ST_a has a unique label and can be directly represented as l_k^a, so its latent representation is $\mathbf{en}_{l_k^a}$. However, The number of labels in task ST_b is variable, so in this article, we apply $\left|l_k^b\right|$ to represent the number of labels and apply $\mathbf{en}_{l_k^b}$ to denote the latent vectors of labels in ST_b. The sentence-level feature calculation of task ST_b is shown Eq. (7):

$$\mathbf{V}_B^{l_k^b} = \sum_{i=1}^{2L} \gamma_i^b \cdot \mathbf{h}_i, \quad \gamma_i^b = \frac{\exp\left(\mathbf{W}_c\left(\mathbf{h}_i \oplus \mathbf{en}_{l_k^b}\right)\right)}{\sum_{j=1}^{R} \exp\left(\mathbf{W}_c\left(\mathbf{h}_j \oplus \mathbf{en}_{l_k^b}\right)\right)} \tag{7}$$

Where γ_i^b denotes the weight of the word i in MCST descriptions, $\mathbf{h}_i$ denotes the latent representation of the word i, and $\mathbf{W}_c$ represents the trainable parameters. Hence, the private feature $\mathbf{Z}^b$ of the task ST_b can be denoted as:

$$\mathbf{p} = \left[\mathbf{V}_B^{l_k^b}, \mathbf{V}_C\right], \quad \mathbf{Z}^b = \sum_{i=1}^{10} \delta_i^b \cdot \mathbf{p}_i, \quad \delta_i^b = \frac{\exp\left(\mathbf{W}_d\left(\mathbf{p}_i \oplus \mathbf{en}_{l_k^b}\right)\right)}{\sum_{j=1}^{10} \exp\left(\mathbf{W}_d\left(\mathbf{p}_j \oplus \mathbf{en}_{l_k^b}\right)\right)} \tag{8}$$

Where p denotes the fused representation incorporating the multi-scale CNN features and label-guide Bi-LSTM features, γ_i^b represents the weight of the representation i.

4.4. Task-adaptive Multi-task Loss Function

The main goal of ST_a is to explore the task type of MCST, which is a coarse-grained and relatively simple classification subtask. ST_b needs to extract all the heterogeneous CAs that may be used in the MCST, which is a fine-grained subtask with a relatively complicated process. Therefore, we propose a task-adaptive multi-task loss function to achieve the optimal solution of this multi-task model, to balance different tasks and improve the efficiency of the multi-task reasoning network.

For the multi-class classification task ST_a, the loss function is shown in Eq. (9):

$$\mathcal{L}_1\left(y^a, \hat{y}^a\right) = -\frac{1}{N} \sum_{i=1}^{N} \sum_{j=0}^{K} \mathrm{I}\left\{y_i^a = j\right\} \log \hat{y}_i^a \tag{9}$$

Where y^a denotes the real type of ST_a, $\hat{y}^a$ means the probability of the predicted value of ST_a, $\hat{y}_i^a = \mathrm{softmax}\left(z_i^a\right)$, $z_i^a = \sigma\left(\mathbf{W}_g \mathbf{Z}_i^a + b^a\right)$, and N represents the total number of MCST. For the multi-label classification task ST_b, the loss function is shown in Eq. (10):

$$\mathcal{L}_2\left(y^b, \hat{y}^b\right) = -\frac{1}{N}\frac{1}{M} \sum_{i=1}^{N} \sum_{j=1}^{M} \left[y_{i,j}^b \log \hat{y}_{i,j}^b + \left(1 - y_{i,j}^b\right) \log\left(1 - \hat{y}_{i,j}^b\right)\right] \tag{10}$$

Algorithm 1: Multi-task Hybrid Reasoning Model Integrating HSR-LKG (TRAINING)

Input: $\{RT_t\}_{t=1}^{T}$: Training sets of all MCST; T: Number of tasks; $\mathcal{E}$: Shared multi-granularity feature extractor; R is the number of words in each MCST after compensation; l denotes the number of layers updated in HSR-LKG; C: Number of label nodes; $\mathbf{C_t}$:Task-specific classifiers; λ: Task learning rate.
Output: LKG-MF network and all parameters.

Randomly initialize various $\mathbf{W}$ parameters in the model, such as $\mathbf{W_c}$, $\mathbf{W_e}$, $\mathbf{W_f}$.
Use BERT pre-training RT, initialize label entity cluster, including label semantics (e_{ls}), synonyms and functional description attributes (e_{sa}, e_{fa}).
while not converged **do**
  Sample a batch of each MCST description RT, and embed each instance into the feature space $\mathcal{E}(x)$, get their embeddings ET.
  for *i=1 to l* **do**
    Get the aggregation characterization of node u ($\mathbf{en}_u^{r\prime(l)}$) under the condition of relationship r by Eq. (1).
    Represent the neighbors ($\mathcal{N}_u^r$) obtained by sampling node u using the relationship r, $\mathcal{N}_u^r = \{(v) \mid (u, r, v) \in \mathcal{G}\}$.
    Construct label nodes $\mathbf{en}_u^{(l)}$ by Eq. (2).
    Get the HSR-LKG representation.
  end
  Get the spatial distance dependent feature V_B and local phrase structure feature V_C.
  for *j=1 to T* **do**
    Use label-guide attention mechanism to handle ST_a and ST_b, guiding the extraction and fusion representation of private features through HSR-LKG by Eq. (3).
    Obtain the final private feature representation $\{\mathbf{Z}^a, \mathbf{Z}^b\}$ for $\{ST_a, ST_b\}$ by Eq. (5) and Eq. (8).
  end
  Compute ST_a and ST_b loss function $\mathcal{L}_1(y^a, \hat{y}^a)$ and $\mathcal{L}_2\left(y^b, \hat{y}^b\right)$ by Eq. (9) and Eq. (10).
  Calculate the weight of the task i: ω_i.
  Use the task-adaptive multi-task loss function $\mathcal{L}$ to achieve the optimal solution of the multi-task model by Eq. (11) and Eq. (12).
  Update network parameters.
end
return LKG-MF network

Where y^b denotes the real labels of ST_b, $\hat{y}^b$ means the predicted multi-label probability value of ST_b, $z_i^b = \sigma\left(\mathbf{W}_h \mathbf{Z}_i^b + b^b\right)$, $\hat{y}_i^b = \text{sigmoid}\left(z_i^b\right)$, M represents the total number of labels.

However, the learning of ST_b is more difficult than that of ST_a. On the one hand, ST_a is a coarse-grained task, while ST_b is a hierarchical fine-grained task. On the other hand, ST_a belongs to a multi-class classification task, while ST_b is a multi-label classification task, containing an uncertain number of labels and interdependent labels. Therefore, inspired by (Guo et al., 2018), we propose a task-adaptive loss function:

$$\mathcal{L} = \omega_a \times \mathcal{L}_1(y^a, \hat{y}^a) + \omega_b \times \mathcal{L}_2\left(y^b, \hat{y}^b\right) \tag{11}$$

Where ω_i denotes the weight of the task i, that is, the difficulty of the task i, can be expressed as:

$$\omega_i = -\left(1 - p_i\right)^{\gamma_0} \log\left(p_i\right) \tag{12}$$

Where p_i indicates the related metric of task i, such as accuracy, which is generally in [0,1]. γ_0 allows the adjustment of the weight of specific tasks. The higher p_i, the better the learning of the task, and the smaller the weight γ_0. For ST_a, a multi-class classification task, the higher the confidence when the predicted $\hat{y}^a$ is close to 0 or 1. Therefore, we employ the entropy method inspired by the article (Wang et al., 2019c) to define p_a, that is, $p_a = 1 - H\left(\hat{y}_i^a\right) = 1 + \sum \hat{y}_i^a \log \hat{y}_i^a$. For ST_b, a multi-label classification task, the predicted value $\hat{y}_i^a$ indicates the distribution of labels, and then the confidence degree p_b can be calculated by the distance between the real label distribution and the predicted label distribution, that is $p_b = \left\|\hat{y}^b - y^b\right\|_2$.

Algorithm 1 details the training process of the LKG-MF method.

Table 2
CrowdTask Dataset Introduction.

Attributes	Details
Number of samples	10403
TA Type	4 types, i.e., Collect, Label, Survey, Track
CAs Type	14 types, i.e., Human, PC, Mobile Phone, Smart bands, UAV, GPS, Mobile Robot, Camera, Microphone, Acceleration sensor, Pulse sensor, Blood pressure sensor, Air quality sensor, Water quality sensor
Data sources	From 7 platforms, i.e., AMTurk (Buhrmester et al., 2011), Common Sense (Dutta et al., 2009), CrowdTracker (Jing et al., 2018), FlierMeet (Guo et al., 2015), CrackSense (Wang et al., 2019a), CrowdWatch (Wang et al., 2016), WeSense (Liu et al., 2020)
Support Language	English
MCS Task Format	Open-domain; Unstructured text

5. Experimental Evaluation

5.1. Experimental Setting

Dataset description. We collect more than 10,000 published MCST from 7 public MCS platforms and uniformly mark these tasks by multiple experts. Then we built an open-domain dataset CrowdTask. These are all very well-known and continuously operating open platforms in the MCS field. The number of tasks in this dataset is scalable and the amount of data continues to increase. We briefly introduce the attributes and corresponding detail descriptions of MCST in Table 2. The dataset and instruction manual can be found in (Liu et al., 2024).

Baseline Methods. To verify the efficiency of our approach, we select 9 different state-of-the-art baselines (shown in Table 3) for experimental analysis. The SN1-SN5 are single-task methods, completing one task at a time. The SN6-SN9 are multi-task methods, which share parameters and perform ST_a and ST_b simultaneously. The details are listed as follows:

- LSAE (Wang and Gao, 2024) leverages the graph convolution network to capture the relationships among labels, and then model the multi-dimension semantic interactions between these labels and text features.

- Bi-Attentive (Liu et al., 2021) introduces an innovative approach that integrates label embeddings and a bi-directional attentive mechanism into BERT's classification framework.

- CBi-Attention (Lu et al., 2023) utilizes CNN and BiLSTM to extract and global semantics, and model the Attentive mechanism to select the most pertinent features for each label.

- CoocNet (Li et al., 2024) proposes the label denoising attention approach to model the local label co-occurrence relationships, and then utilizes the contrastive learning strategy to capture global label co-occurrence relationships.

- ML-Reasoner (Wang et al., 2021) utilizes a binary classifier to simultaneously predict all labels, and then applies an iterative reasoning mechanism to effectively model utilize the high-order correlations among these labels.

- LocalAtt (Xu et al., 2021) utilizes the label embedding and a modified TF-IDE matrix to extract the task-specific features, and then utilizes the label embedding-based attention weight learning to enhance the text representation.

- LED (Ma et al., 2023) employs a hierarchy-aware attention mechanism to represent label semantics, and leverages the multi-task learning to effectively capture label correlation.

- MHCAN (Lu et al., 2022) utilizes the cross-attention method to model the text and hierarchical labels, and subsequently leverages the iterative hierarchical-attention module to capture the relationships among levels.

Table 3
Comparative Evaluation with Baselines.

SN	Methods	P(+)		R(+)		F1(+)		HL(-)
		ST_a	ST_b	ST_a	ST_b	ST_a	ST_b	ST_b
1	LSAE (Wang and Gao, 2024)	0.950	<u>0.770</u>	0.936	<u>0.748</u>	0.943	<u>0.759</u>	<u>0.0325</u>
2	Bi-Attentive (Liu et al., 2021)	0.941	0.472	0.925	0.459	0.933	0.465	0.0677
3	CBi-Attention (Lu et al., 2023)	0.938	0.675	0.919	0.652	0.928	0.663	0.0383
4	CoocNet (Li et al., 2024)	0.946	0.697	0.931	0.682	0.938	0.689	0.0425
5	ML-Reasoner (Wang et al., 2021)	0.935	0.627	0.919	0.611	0.927	0.619	0.0498
6	LocalAtt (Xu et al., 2021)	0.943	0.526	0.928	0.518	0.935	0.522	0.0586
7	LED (Ma et al., 2023)	<u>0.951</u>	0.785*	0.942*	0.751*	0.946*	0.768*	0.0309*
8	MHCAN (Lu et al., 2022)	0.948	0.763	<u>0.938</u>	0.746	0.943	0.754	0.0357
9	LACO (Zhang et al., 2021)	0.953*	0.726	0.936	0.718	<u>0.944</u>	0.722	0.0371
	LKG-MF	**0.962**	**0.802**	**0.958**	**0.782**	**0.960**	**0.792**	**0.0229**

- LACO (Zhang et al., 2021) introduces the document-label joint embedding to model the more discriminative document and label representation, and designs the correlation aware multi-task learning to enhance label correlation.

Evaluation Metrics. Our research objective is to infer two types of implicit core information from the MCST, namely TP and CAs. We define the problem as a multi-task joint reasoning problem combining multi-class and multi-label classification. Therefore we utilize the precision rate (P), recall rate (R), F1-Score, and Hamming Loss (HL) as evaluation indicators. From the perspective of system application, we compare the differences in task assignment efficiency and task completion time between MCS platforms before and after integrating our model.

Implementation Settings. All approaches are implemented in Pytorch 1.10.2 and Python 3.9. For fairness, we uniformly set the parameters of all models. We set the optimizer to Adam, the batch size to 256, the learning rate to 0.001, and the dropout ratio to 0.8. While the embedding dimension is 768, the convolution kernel size is 2, 3, 4. To obtain more reliable experimental results, we run all approaches five times and take the average results.

5.2. Comparative Analysis with Baselines

We select 9 current state-of-the-art models to deal with the MCST multi-task reasoning problem on the CrowdTask dataset. The experimental results are shown in Table 3, where the single-task algorithm trains ST_a and ST_b through two networks respectively, and the algorithm trains heterogeneous tasks simultaneously through a multi-task method in one network. To make the benchmarks applicable to ST_a and ST_b, we make adjustments to the classifiers for some models.

Analyzing Table 3, we have the following findings: 1) we analyze the methods that incorporate label representation and observe that those modeling label correlation (SN1, SN4, SN5, SN7, SN8, and SN9) outperform those only modeling label representation (SN2, SN3, and SN6). In particular, our method models explicit and implicit correlations between labels and establishes correlations between labels across multiple tasks, demonstrating optimal performance and highlighting the importance of label correlations in multi-label classification. It also demonstrates the important correlation between labels in MCS tasks. 2) On the whole, the multi-task approach slightly outperforms the single-task approach. This can be attributed to two factors: firstly, feature extraction can be shared across tasks; secondly, the strong correlations between tasks facilitate the extraction of both shared and private features, resulting in better performance. 3) For both tasks, task ST_a achieves good performance in different methods, because task ST_a is a multi-classification task that is relatively simple, while task ST_b involves multi-label classification requiring the acquisition of fine-grained CAs, making it more complex. 4) Our method achieves greater performance improvement in task ST_b. This mainly stems from our utilization of a knowledge graph to model the fine-grained semantics of labels and their correlations. In summary, our method, LKG-MF, achieves remarkable performance, significantly outperforming other approaches in P, R, F1, and HL metrics, thereby demonstrating the effectiveness of our approach.

5.3. HSR-LKG Effectiveness Analysis

To verify the effectiveness of HSR-LKG, we carry out the ablation comparison verification of HSR-LKG, that is, only adjust the HSR-LKG part in the LKG-MF model, and ensure that other modules and parameters remain unchanged.

Table 4
HSR-LKG Effectiveness Analysis for ST_a and ST_b

VN	Variants of HSR-LKG	P(+)		R(+)		F1(+)		HL(-)
		ST_a	ST_b	ST_a	ST_b	ST_a	ST_b	ST_b
1	w/o HSR-LKG	0.937	0.673	0.921	0.658	0.929	0.665	0.0372
2	HSR-LKG (e_{ls})	0.951	0.782	0.935	0.759	0.943	0.770	0.0258
3	HSR-LKG (e_{ls}+e_{sa})	<u>0.957</u>	<u>0.785</u>	<u>0.947</u>	<u>0.763</u>	<u>0.952</u>	<u>0.774</u>	<u>0.0249</u>
4	HSR-LKG (e_{sa}+e_{fa})	0.949	0.780	0.931	0.754	0.940	0.767	0.0327
5	HSR-LKG (1-head)	0.953	0.781	0.944	0.762	0.948	0.771	0.0284
6	HSR-LKG (5-head)	0.960*	0.792*	0.951*	0.780*	0.955*	0.786*	0.0241*
7	**LKG-MF**	**0.962**	**0.802**	**0.958**	**0.782**	**0.960**	**0.792**	**0.0229**

Table 5
Functional Module Ablation Test for ST_a and ST_b.

FN	Functional Module	P(+)		R(+)		F1(+)		HL(-)
		ST_a	ST_b	ST_a	ST_b	ST_a	ST_b	ST_b
1	w/o LSTM	0.927	0.726	0.910	0.704	0.918	0.715	0.0397
2	w/o CNN	0.921	0.731	0.903	0.709	0.912	0.720	0.0442
3	w/o LSTM+CNN	0.901	0.708	0.879	0.673	0.890	0.690	0.0491
4	w/o LSTM+CNN+LGF	0.882	0.683	0.861	0.652	0.871	0.667	0.0675
5	HSR-LKG (w/o e_{sa}+e_{fa})	<u>0.951</u>	<u>0.782</u>	<u>0.935</u>	<u>0.759</u>	<u>0.943</u>	<u>0.770</u>	<u>0.0258</u>
6	w/o HSR-LKG	0.937	0.673	0.921	0.658	0.929	0.665	0.0372
7	w/o Task-adaptive Loss	0.959*	0.791*	0.949*	0.768*	0.954*	0.779*	0.0235*
8	**LKG-MF**	**0.962**	**0.802**	**0.958**	**0.782**	**0.960**	**0.792**	**0.0229**

Table 4 shows these experimental results on ST_a and ST_b. The VN1 row represents the model evaluation results when HSR-LKG is not included in the framework. VN2 indicates that the label semantics in HSR-LKG only use e_{ls} information. VN3 and VN4 represent the multi-layer semantic information of $e_{ls} + e_{sa}$ and $e_{sa} + e_{fa}$ respectively integrated into the label entity. VN5-VN7 indicates the utilization of the K-head attention mechanism (K=1, 5, 8) when performing label entity (en_u) representation. VN7 (our approach) is the (K=8) model used in this article. From the comparison results of VN7 and VN1, it can be seen that the existence of HSR-LKG can increase the values of indicators P, R, and F1 by 19.2%, 18.8%, and 19.1% respectively.

To verify the effectiveness of label attribute fusion and multi-head attention mechanism in the process of building HSR-LKG, we compare the performance of the final model after diversifying the HSR-LKG module (VN2-VN6). From the comparison of VN2, VN3, and VN4, it can be seen that adding hierarchical attribute auxiliary information can improve the performance of the model, and the contribution of e_{ls} is greater than e_{sa}+e_{fa}. VN5 shows the effect of using single-head attention, and the effect of 8-head attention (VN7) used in this article is better than VN5. After a comprehensive analysis, the experimental indicators of our approach (VN7) are higher than those of VN1-VN6, for example, the F1 value is 0.8%-19.1% higher than other model variants. All these verify that our approach can express the MCST task more accurately.

5.4. Function Module Ablation Study

To evaluate the effectiveness of different modules in the LKG-MF framework and their functional effects, we design ablation experiments for different functional modules. The first column FN in Table 5 indicates the function number. The second column indicates the model operation results obtained after removing different modules or module combinations in the LKG-MF framework, such as removing the LSTM module, CNN module, CNN+LSTM module, HSR-LKG module, and task adaptation module. The symbol w/o stands for the word without. LGF means Label-guide Features.

Analyzing Table 5, we have four findings: 1) Comparing FN1, FN2, FN3, and FN8, we can see that the multi-dimensional task features of MCST tasks are mined through modules such as CNN and LSTM, and are further used as the input of the final multi-task classification network. It can effectively improve the accuracy of task execution. 2) By comparing FN5, FN6, and FN8, it can be seen that the design of HSR-LKG and its internal structure helps to express

Table 6
Performance of LKG-MF model with different loss weights for ST_a and ST_b.

Variants of Loss	P(+)		R(+)		F1(+)		HL(-)
	ST_a	ST_b	ST_a	ST_b	ST_a	ST_b	ST_b
ω_a=0.2, ω_b=0.8	0.959	0.798	0.951	0.776	0.955	0.787	0.0234
ω_a=0.4, ω_b=0.6	0.961*	0.800*	0.955*	0.779*	0.958*	0.789*	0.0230*
ω_a=0.5, ω_b=0.5	0.959	0.791	0.949	0.768	0.954	0.779	0.0235
ω_a=0.7, ω_b=0.3	0.955	0.784	0.943	0.760	0.949	0.772	0.0242
LKG-MF(Adaptive Loss)	**0.962**	**0.802**	**0.958**	**0.782**	**0.960**	**0.792**	**0.0229**

the inner correlation of labels at a deeper level, thereby improving the P, R, and F1 values of the task. 3) By comparing FN4 and FN8, it shows the multi-dimensional hybrid feature extraction module can greatly improve the accuracy of feature representation for different tasks. 4) Comparing FN7 and FN8, it means that the existence of the task adaptive module can help the model obtain the most suitable optimization weights for ST_a and ST_b, thereby accelerating the model convergence speed and improving performance indicators.

5.5. Task-adaptive Multi-task Loss Analysis

Section 5.4 has verified the effectiveness of the task-adaptive multi-task loss function as a whole, this section aims to further illustrate the performance of the different loss functions in LKG-MF. Specifically, we set four comparative experiments to analyze the performance and training loss of the model under different weights ω_a and ω_b in Eq.(11), to assess the effectiveness and learning speed of the model. And the experimental results are shown in Table 6 and Fig. 3. Table 6 demonstrates the superiority of our adaptive loss in performance, while Fig. 3 shows the superiority of our adaptive loss in performance training speed.

From the Table 6, it's obvious that our model with adaptive loss can achieve the most performance, and the model with ω_a=0.4 and ω_b=0.6 can get the suboptimal performance. This is because when ω_a is small, the weight of task ST_a is relatively small, and the overall model focuses on the fitting of task ST_b, which can improve the overall performance and learning speed. Therefore, the performance of the model with ω_a=0.2 and ω_a=0.4 in Table 6 is good, and the training losses of these models in Fig. 3 decrease rapidly, as indicated by the green and orange lines. In addition, for the model with ω_a=0.2 and ω_b=0.8, perhaps it is because the model focuses too much on task ST_b, resulting in good performance on task ST_b, but overall it still falls short of the model with ω_a=0.4 and ω_b=0.6. On the other hand, theoretically speaking, the larger the ω_a, the better the performance of task ST_a. However, from Table 6, it can be found that the overall performance of the model with ω_a=0.7 and ω_b=0.3 in tasks ST_a and ST_b is relatively low, which may be due to overfitting of the model to simple task ST_a, resulting in a decrease in the performance of task ST_a; At the same time, the model has insufficient fitting for difficult task ST_b, resulting in slightly poorer performance for task ST_b. This model requires more rounds of training to reduce the training loss of difficult task ST_b, as indicated by the purple line in Fig.3.

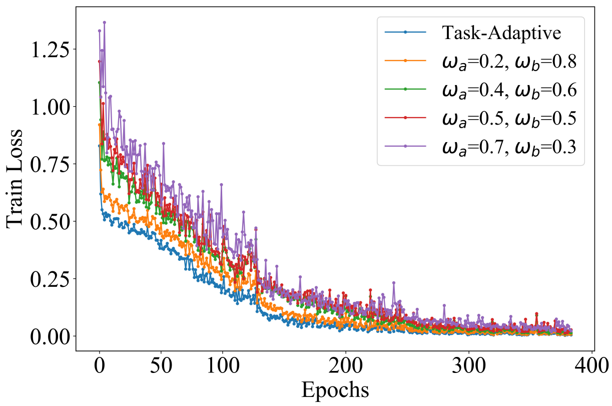
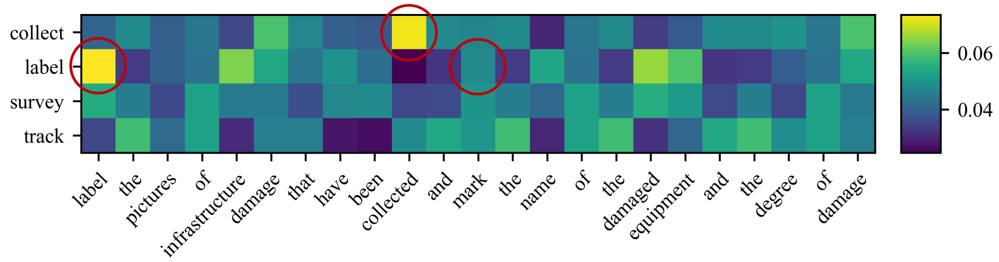
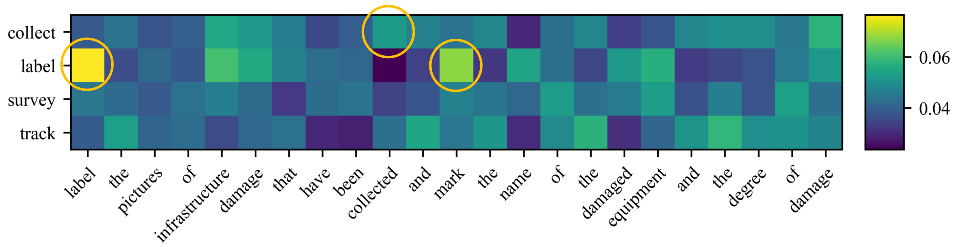


Figure 3: The training loss within 400 epochs.



(a) Similarity matrix between $Task2$ and TP nodes (for ST_a, without HSR-LKG)



(b) Similarity matrix between $Task2$ and TP nodes (for ST_a, with HSR-LKG)

Figure 4: Comparison of word attention weights for TP in $RT - Task\ 2$ with or without HSR-LKG fusion.

5.6. Visual Analysis of Task Instances

Here we show the similarity matrix between RT tasks and label graph nodes.

Task 2: Label the 50,000 pictures of infrastructure damage that have been collected, and mark the name of the damaged equipment and the degree of damage.

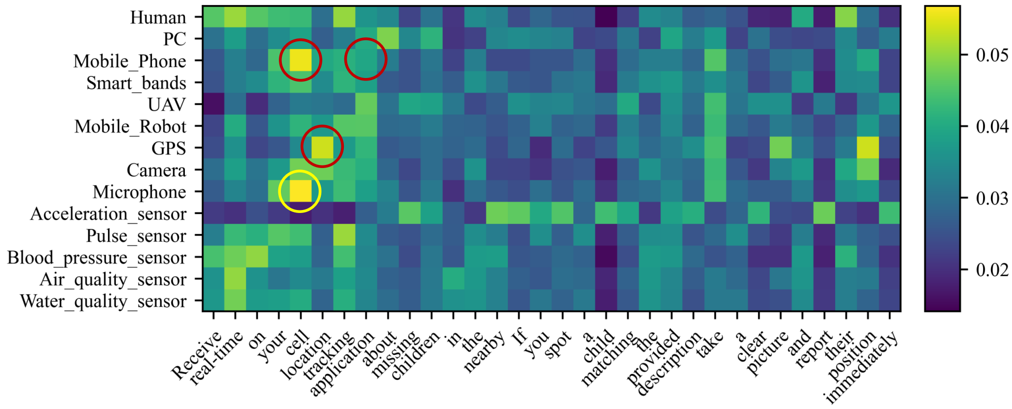
Task 9: Receive real-time notifications on your cell's location tracking application about missing children in nearby. If you spot a child matching the provided description, take a clear picture and report their position immediately.

We take two specific MCSTs as examples, that is, $Task\ 2$ and $Task\ 9$ above, observe and record the actual effect of using different HSR-LKG variants in discriminating TP and CAs trigger words. Since there are two TP trigger words ($collect$ and $label$) appear in $Task\ 2$, but a task only belongs to one TP, the model needs to combine the context to determine which trigger is important to $Task\ 2$. In addition, there are 4 CAs trigger words ($human$, $mobile\ phone$, GPS, $camera$) appeared in $Task\ 9$.

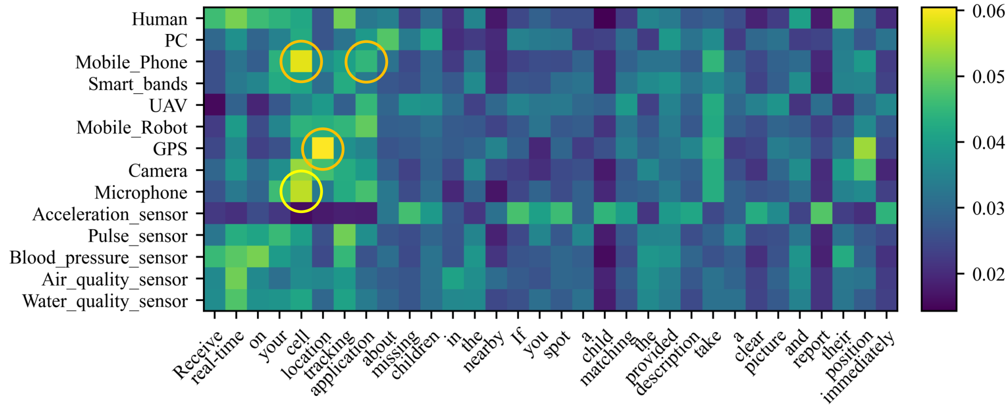
Fig.4 shows the degree of correlation between the vector representation of RT and different types of labels through a heat map. Among them, Fig.4 (a) and Fig.4 (b) compared the attention weight between RT and TP labels for $Task\ 2$ before and after adding HSR-LKG. For a detailed comparison of the attention weights between RT and CAs labels for $Task\ 9$, please see Fig.5. From the comparative analysis, it can be seen that the HSR-LKG based on the multi-head attention mechanism effectively integrates the use and scene information, especially when the attention mechanism with head K=8 (full model) is used (Fig.4 (b), Fig.5 (b)), the model can clearly identify real task-type trigger words. When not using HSR-LKG (Fig.4 (a)) or using the castrated version of HSR-LKG, the task feature is unable to distinguish true and false TP trigger words ($label$ and $collect$). However, when using our model to process $Task\ 2$ (Fig.4 (b)), the word $label$ has a significantly higher weight than $collect$. Through the comparison of Fig.5 (a) and Fig.5 (b), we can also see that the fusion of HSR-LKG and RT multi-granularity features effectively strengthens the expression strength of TP and CAs core words or phrases. The accuracy and effectiveness of RT private features for ST_a and ST_b are greatly improved. The above analysis verifies the accuracy and effectiveness of HSR-LKG and the multi-head attention mechanism for the expression of task-critical information.

5.7. Evaluation Results on the AAPD Dataset

To further verify the generalization and universality of our model, we conduct experiments on the public dataset AAPD. AAPD is a general dataset for multi-label learning, which aims to assign relevant topics to papers. It contains 55,840 abstracts of papers in the computer field and 54 related topics. The use of this dataset is to verify the effectiveness of HSR-LKG, and we do not need to consider the task adaptive loss module. Table 7 shows the results of the comparative



(a) Similarity matrix between $Task9$ and CAs nodes (for ST_b, without HSR-LKG)



(b) Similarity matrix between $Task9$ and CAs nodes (for ST_b, with HSR-LKG)

Figure 5: Comparison of word attention weights for CAs in $RT - Task\ 9$ with or without HSR-LKG fusion.

Table 7
Comparative Evaluation with Baselines on ST_b on AAPD Dataset.

SN	Methods	P(+)	R(+)	F1(+)	HL(-)
1	LSAE (Wang and Gao, 2024)	0.818*	0.710*	0.760*	0.0214
2	Bi-Attentive (Liu et al., 2021)	0.752	0.611	0.674	0.0259
3	CBi-Attention (Lu et al., 2023)	0.785	0.675	0.726	0.0231
4	CoocNet (Li et al., 2024)	0.806	0.692	0.745	0.0220
5	ML-Reasoner (Wang et al., 2021)	0.726	0.708	0.717	0.0238
6	LocalAtt (Xu et al., 2021)	0.762	0.658	0.706	0.0254
7	LED (Ma et al., 2023)	0.815	0.683	0.743	0.0215
8	MHCAN (Lu et al., 2022)	0.792	0.702	0.744	0.0221
9	LACO (Zhang et al., 2021)	0.802	0.696	0.745	0.0213*
	LKG-MF	**0.826**	**0.712**	**0.765**	**0.0210**

evaluation with baselines on ST_b on the AAPD dataset. For the parameter settings, we set the optimizer to Adam, the batch size to 32, the learning rate to 5e-5, the embedding dimension to 768, the maximum total input sequence length to 320, etc. It can be seen that the method of modeling label correlation is better than the method without modeling correlation in various indicators. Moreover, our method is better than the traditional model which depends on the frequency of label co-occurrence pattern, which also verifies the effectiveness and generalization of our model.

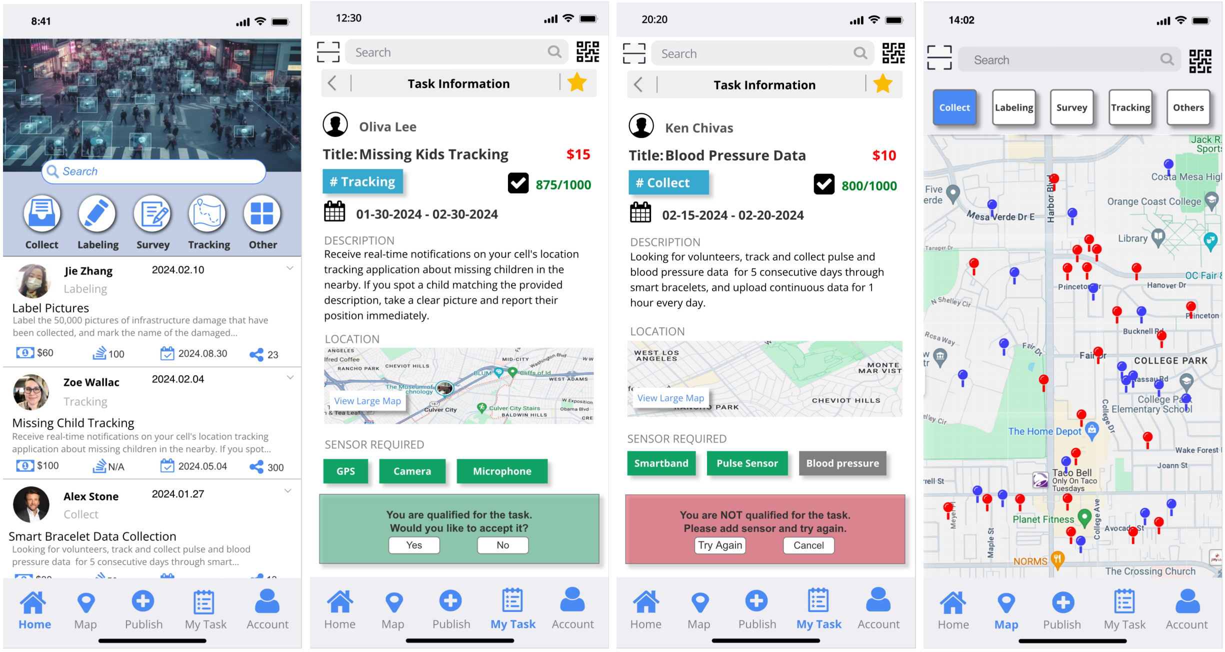
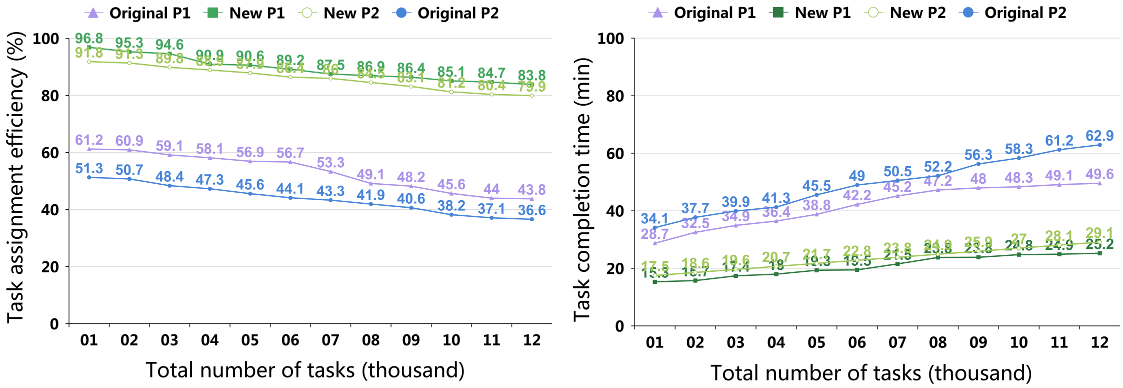


Figure 6: Function pages display (MCS platform integrated with LKG-MF). (a) Home - current task lists in the system. (b) Task Details - display the detailed information of a Tracking type task (Qualified). (c) Task Details - display the detailed information of a Collect type task (Not Qualified). (d) Map - task distribution map.

5.8. Evaluation of LKG-MF integrated into MCS platforms

We further verify how the model proposed in this article can help MCS systems or platforms improve task assignment efficiency. We also clarify the detailed aspects from which the LKG-MF serves the existing MCS platforms. The experiment is carried out by embedding the LKG-MF model in the Open Source Crowd-Sensing Platform SafeCity (Liu et al., 2023), and then testing it by publishing tasks and collecting results of the tasks. The new system can extract the core information from the original task description, and complete reasonable and efficient task assignment based on the information, thereby quickly obtaining efficient task result data. Fig.6 shows the system interaction interfaces that integrate the LKG-MF model. Fig.6 (a) shows the tasks published on the platform, and the platform can automatically identify the type of task. Fig.6 (b) is the task information seen by the participants. It can be seen that the figure includes the mining and display of task types and sensor information used. This enables the system to review the qualifications of participants, and participants who do not meet the qualifications cannot perform the task. Fig.6 (c) shows that when the participant does not meet the basic requirements of the current task, the participant can not perform the task. Fig.6 (d) shows the task map near the current location of the participant. The red marks are MCS tasks to which participants can be matched after being screened by LKG-MF, and the blue marks indicate tasks to which current participants are not qualified and cannot participate. From the above analysis, it can be seen that the role of LKG-MF in improving task processing efficiency is mainly from the following aspects: 1) LKG-MF deeply understands and analyzes the tasks entering the platform, digs out the core information (CAs and TP) of the task and feeds it back to the system; 2) The system optimizes the task allocation and participant selection algorithm based on these core information, and gives priority to participants who meet the task execution conditions to complete the task; 3) Through the above screening, the quality of the results obtained and the completion speed have been greatly improved. LKG-MF has greatly optimized the original platform processing flow.

Through comprehensive comparison and verification, integrating the LKG-MF model into the original MCS platforms can help existing MCS platforms quickly screen and match qualified participants, thereby obtaining high-quality task results. As shown in Fig. 7, through 2,000 sets of actual task tests and 10,000 sets of simulation tests on the open source platforms SafeCity (Liu et al., 2023) (Original P1) and WeSense (Liu et al., 2020) (Original P2),



(a) Task assignment efficiency evaluation.

(b) Task completion time evaluation.

Figure 7: Comparative evaluation of MCS platforms before and after integrating the LKG-MF model.

we find that MCS platforms integrating the LKG-MF model (New P1, New P2) can improve the task assignment efficiency (the matching degree between tasks and participants and task completion degree) by an average of 38.6%, reduce task completion time (the time from task release to task result submission) by an average of 45.1%, and greatly improves the quality of collected task results (including the total amount of result data, data completeness, data type correctness, etc.). The quality of the task results collected by WeSense (Liu et al., 2020) and SafeCity (Liu et al., 2023) platforms before and after the integration of LKG-MF (Original Platforms and New Platforms) is currently evaluated by the following indicators. 1) The total amount of result data. In the new platforms, since the MCS tasks are more accurately assigned to the appropriate participants, the result data collected is richer and takes less time than before; 2) Data completeness. Since the task participants in the new platform have the sensors required for the task through the auxiliary screening of LKG-MF during the task assignment stage, the sensing data provided is more complete than the original platforms. 3) Data type correctness. Due to the understanding and analysis of MCST, the type of sensing data collected in the new platform is more in line with the original intention of the task. Since the current MCS platforms usually does not provide an online evaluation function for the quality of task results, the specific process of this evaluation is completed by the joint calculation of manual evaluation and automatic detection after the platforms collect the result data. In order to optimize and uniformly evaluate MCS tasks, an intelligent evaluation and feedback plug-in for multiple MCS tasks can be developed in the future to assist in the detection and evaluation of result quality. We describe this part in Discussions and Future Work Section. This experiment not only proves the correctness and effectiveness of the method proposed in this article, but also further verifies the significance of this research for MCS platforms, which can effectively improve task assignment efficiency and promote the quality of task results.

6. Discussion

6.1. Results and Implications

In this article, we introduce a label knowledge graph-powered multi-task framework, LKG-MF, to achieve better classification performance in crowdsourcing and mobile crowd sensing tasks. In contrast to existing multi-label methods that rely solely on pre-trained label semantics, co-occurrence modeling, or hierarchical modeling of labels, which are inherently constrained by label limitations and struggle in open-domain contexts characterized by diverse and implicit expressions, LKG-MF introduces a label knowledge graph that additionally includes incorporated label's attributes and relationships (HSR-LKG). This innovative graph captures heterogeneous semantics and relationships among labels, enabling our method to excel in MCS scenarios marked by implicit, indeterminate semantics and complex, diverse relationships. In addition, we present a label-guide multi-granularity feature extraction method to model the personalized text descriptions and capture precise task-specific features. Furthermore, we propose a task-adaptive loss function to enhance the performance across multiple tasks and improve the efficiency of the multi-task reasoning network. Extensive experimental results on our constructed dataset demonstrate that LKG-MF outperforms baselines average by 2.3%, significantly improving multi-task classification accuracy. Results on open-source dataset demonstrate that LKG-MF also outperforms SOTA, expressing the effectiveness and generalization of our LKG-MF.

In addition, we integrate our LKG-MF model into the existing MCS platforms through the API interface provided by the model, empowering them to swiftly filter and match optimal task participants, thereby yielding superior task outcomes. This integration has yielded remarkable results, with a notable 38.6% improvement in task assignment efficiency, a 45.1% reduction in task completion time, and a substantial enhancement in the quality of collected data. This result indicates the practical significance and effectiveness of our LKG-MF method in enhancing the performance of MCS systems, underscoring its potential to revolutionize the field.

6.2. CrowdTask Dataset and Deployment Details

The CrowdTask dataset is being expanded and maintained continuously. The CrowdTask MCS dataset (Liu et al., 2024) constructed in this article contains rich task data in the MCS field that can currently be collected. The core information of MCS tasks in the existing data set can be attributed to 4 types of TP and 14 types of CAs. When constructing the dataset, we not only referred to the seven platforms (Table 2) mentioned before, but also considered multiple research platforms that have been released in the MCS field. Since these platforms have not been released to the public, we have not collected enough public tasks, so we cannot use these platforms as the task source of CrowdTask. However, the core factors that affect task assignment and execution efficiency in the platform, namely the core information in MCST, have basically been included in this dataset, that is, task types and sensors that may have been covered. Therefore, the CrowdTask dataset constructed in this article has good generalization ability.

In addition, since most MCS platforms currently use English as the basis for communication, we have not yet collected enough MCS tasks in other languages, which is the main reason why our dataset currently only contains English tasks. If more tasks based on other languages can be collected in the future, we will collect these tasks to enrich the overall dataset, while correspondingly modifying the construction of the knowledge graph and the input mode of the LKG-MF method to adapt to data in different languages.

Since the core of the model is to understand the MCS task and identify the task type and sensor information, it is usually deployed in the cloud together with the MCS platforms. The model can be deployed and run normally in the cloud computing platform without additional computing and memory resources. The minimum operating requirements include sufficient idle CPU and memory capacity, such as system memory of not less than 12G. Basic software such as the python compiler, as well as the nltk package and gensim package have been installed to ensure that the model can run normally. Data transmission is stable and the delay is within an acceptable range.

6.3. Model Generalization Analysis

The LKG-MF model has good generalization capability in non-MCS fields. Firstly, the LKG-MF model constructed a label knowledge graph incorporating Heterogeneous Semantics and Relationships, which includes diverse representations and relationships of labels, enabling the model to more accurately identify label content with different expressions and enhance its generalization ability. For example, the entity set E in the HSR-LKG includes label semantic entities, entity synonym attributes, and entity function and purpose attributes. Therefore, even if new types coming, our model can still accurately recognize and has strong generalization ability. Secondly, the LKG-MF model included a Task-adaptive Multi-task Loss Function, which enables the model to adaptively learn multiple tasks, improving the model's generalization ability. Lastly, we verified the ability of the LKG-MF model to process non-MCS tasks. Specifically, we adjusted and modified the model to adapt to the AAPD dataset, and conducted a detailed comparative analysis of the model and multiple baselines on this dataset. The experiment fully verified that the LKG-MF model can also be migrated to other multi-label tasks. Table 7 shows the results of the comparative evaluation with baselines on ST_b on the AAPD dataset. It can be seen that our method achieves the best performance, superior to other methods with modeling label correlation, which also verifies the effectiveness and generalization of our model.

As for the generalization of the model to multilingual tasks, the LKG-MF framework has the potential to generalize to multilingual settings. The main reasons are as follows: The core structure of our model, the HSR-LKG, is designed to capture the heterogeneous semantics and relationships among labels. This design principle is not bound by a specific language. For example, the HSR-LKG Construction module, Label-guide Multi-granularity Feature Extraction module, and Task-adaptive Multi-task Loss Function module, are all designed at the semantic level and independent of the language level, meaning that in theory, the model has good generalization ability for diverse language tasks.

6.4. Discussions and Future Work

During the deployment process, we found that the current MCS platforms lack a universal result verification mechanism. Therefore, we can explore and develop an advanced and universal intelligent evaluation algorithm, build it as a plug-in to assist MCS platforms, and enhance the reliability and effectiveness of the result quality in various MCS platforms. For example, deep learning models can be introduced to achieve a more accurate adaptive result verification mechanism, and a closed-loop feedback function can be developed to provide real-time guidance and feedback to participants during task execution, further helping participants collect high-quality result data.

Since most MCS platforms currently use English as the basis for communication, we have not yet collected enough MCS tasks in other languages, which is the main reason why our dataset currently only contains English tasks. If more tasks based on other languages can be collected in the future, in order to make the model better adapt to multilingual tasks, the following modifications can be made: Firstly, in the data preprocessing stage, we need to incorporate multilingual text processing techniques. This includes using multilingual word embeddings or language-specific tokenizers to handle the unique characteristics of different languages. For example, for languages with rich morphological variations, appropriate morphological analyzers can be integrated to better understand the words. Secondly, when constructing the HSR-LKG, we should expand the knowledge base to cover semantic and relational information from multiple languages. This can be achieved by collecting multilingual data sources related to the tasks and extracting common semantic patterns and relationships across languages. Finally, the model's training process may need to be adjusted. We can use multilingual datasets for training to expose the model to different language expressions and improve its generalization ability. Additionally, techniques such as transfer learning can be explored to leverage the knowledge learned from one language to facilitate learning in other languages.

Compared with large language models (LLMs), our model has certain advantages in theoretical design, training cost, and time and space resource consumption. In the preprocessing stage, LKG-MF is based on a pre-trained large model, and then further models the explicit and implicit multi-level relationships between labels, mining multi-granularity fusion features, and obtaining a more accurate and rich task representation. This model better copes with the challenges faced by heterogeneous crowd sensing tasks such as open-domains, diverse language expressions, an uncertain number of labels, and multi-task reasoning. In the context of LLMs, we observe a limitation in their ability to thoroughly model multi-label classification tasks compared to their proficiency in handling multi-class classification tasks like the traditional text prediction tasks. For multi-label tasks, LLMs tend to focus primarily on capturing the core entities or concepts mentioned in the text, neglecting the broader context and related entities that surround them. For instance, in the text 'Collect data during high-intensity exercises with the smart bracelet to assess the maximum heart rate achieved and optimize exercise intensity', LLMs can accurately identify 'Smart bands', but fail to encompass the crucial role of 'Human' and 'Pulse sensor'. However, retraining the parameters of LLMs requires a lot of time and space resources, and the time and space consumption of LKG-MF models is much lower than that of LLMs. In the future, we hope to leverage the reasoning capabilities of LLMs and integrate the constructed knowledge graph into it without the need for additional training, thus achieving fast and accurate reasoning.

7. Conclusion

In this article, we construct a real-world CrowdTask dataset containing more than 10,000 MCS tasks. Then, we propose LKG-MF, a label knowledge graph-powered multi-task framework, to achieve better core information mining performance in crowdsourcing and mobile crowd sensing tasks, futher cope with problems such as implicit expressions and diverse descriptions in MCS tasks. LKG-MF integrates the diverse relationships between attributes and label entities into the semantics of labels to enhance the representation capabilities of labels. LKG-MF also utilizes a label-guided multi-granularity feature extraction network to achieve accurate representation of multi-task private features, further improving multi-task performance. The experimental results show that LKG-MF can accurately mine the core information in MCS tasks and is better than the state-of-the-art methods in multiple indicators. In addition, LKG-MF is integrated into the open-source MCS platforms for evaluation. The results show that compared with the platform that does not further analyze and understand the MCS tasks, the task assignment efficiency of the MCS platform integrated with LKG-MF is improved by 38.6%, and the quality of task results is also greatly optimized.

CRediT authorship contribution statement

Yimeng Liu: Writing – original draft, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Zhiwen Yu**: Writing – review & editing, Supervision, Resources, Conceptualization. **Nuo Li**: Writing – review & editing, Methodology, Resources, Formal analysis. **Bin Guo**: Writing – review & editing, Resources. **Sumi Helal**: Writing – review & editing, Supervision.

Data availability

I have shared the link about the data or code on the manuscript.

Acknowledgment

This work was supported by the Natural Science Foundation of Shanghai (No. 24ZR1418500), the Education and Scientific Research Project of Shanghai (No. GSC2024017), and the National Natural Science Foundation of China (No. 62032020).

References

Albawi, S., Mohammed, T.A., Al-Zawi, S., 2017. Understanding of a convolutional neural network, in: 2017 International Conference on Engineering and Technology (ICET), pp. 1–6. doi:10.1109/ICEngTechnol.2017.8308186.

Bazaluk, O., Rahman, M.A., Zayed, N.M., Faisal-E-Alam, M., Nitsenko, V., Kucher, L., 2024. Crowdsourcing review: the crowd workers' perspective. Journal of Industrial and Business Economics , 1–20.

Buhrmester, M., Kwang, T., Gosling, S.D., 2011. Amazon's mechanical turk: A new source of inexpensive, yet high-quality, data? Perspectives on psychological science 6, 3–5.

Chen, Z., Li, R., Mao, D., 2023. Truth discovery technology for mobile crowd sensing in water quality monitoring. Wireless Communications and Mobile Computing 2023, 4900238.

Devlin, J., Chang, M.W., Lee, K., Toutanova, K., 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv:1810.04805.

Dutta, P., Aoki, P.M., Kumar, N., Mainwaring, A., Myers, C., Willett, W., Woodruff, A., 2009. Common sense: participatory urban sensing using a network of handheld air quality monitors, in: Proceedings of the 7th ACM conference on embedded networked sensor systems, pp. 349–350.

El Hafyani, H., Abboud, M., Zuo, J., Zeitouni, K., Taher, Y., Chaix, B., Wang, L., 2024. Learning the micro-environment from rich trajectories in the context of mobile crowd sensing: Application to air quality monitoring. Geoinformatica 28, 177–220.

Feng, Z., Mao, K., Zhou, H., 2024. Adaptive micro-and macro-knowledge incorporation for hierarchical text classification. Expert Systems with Applications 248, 123374.

Fu, Y., Chen, X., Miao, D., Qin, X., Lu, P., Li, X., 2024. Label-semantics enhanced multi-layer heterogeneous graph convolutional network for aspect sentiment quadruplet extraction. Expert Systems with Applications , 124523.

Guo, B., Chen, H., Yu, Z., Xie, X., Huangfu, S., Zhang, D., 2015. Fliermeet: A mobile crowdsensing system for cross-space public information reposting, tagging, and sharing. IEEE Transactions on Mobile Computing 14, 2020–2033. doi:10.1109/TMC.2014.2385097.

Guo, M., Haque, A., Huang, D.A., Yeung, S., Fei-Fei, L., 2018. Dynamic task prioritization for multitask learning, in: Proceedings of the European conference on computer vision (ECCV), pp. 270–287.

Huang, Z., Xu, W., Yu, K., 2015. Bidirectional lstm-crf models for sequence tagging. arXiv:1508.01991.

Jing, Y., Guo, B., Wang, Z., Li, V.O.K., Lam, J.C.K., Yu, Z., 2018. Crowdtracker: Optimized urban moving object tracking using mobile crowd sensing. IEEE Internet of Things Journal 5, 3452–3463. doi:10.1109/JIOT.2017.2762003.

Li, C., Liu, X., Wang, C., Liu, Y., Yu, W., Shao, J., Yuan, Y., 2025. Gtp-4o: Modality-prompted heterogeneous graph learning for omni-modal biomedical representation, in: European Conference on Computer Vision, Springer. pp. 168–187.

Li, Y., Shen, J., Mao, Z., 2024. Coocnet: a novel approach to multi-label text classification with improved label co-occurrence modeling. Applied Intelligence , 1–17.

Liang, Y., Tohti, T., Hamdulla, A., 2024. Contrastive classification: A label-independent generalization model for text classification. Expert Systems with Applications 245, 123130.

Liang, Z., Rong, Y., Li, C., Zhang, Y., Huang, Y., Xu, T., Ding, X., Huang, J., 2021. Unsupervised large-scale social network alignment via cross network embedding, in: Proceedings of the 30th ACM International Conference on Information & Knowledge Management, pp. 1008–1017.

Liu, B., Zhou, A., Wei, B., Wang, J., Tsoumakas, G., 2025. Oversampling multi-label data based on natural neighbor and label correlation. Expert Systems with Applications 259, 125257.

Liu, N., Wang, Q., Ren, J., 2021. Label-embedding bi-directional attentive model for multi-label text classification. Neural Processing Letters 53, 375–389. doi:10.1007/s11063-020-10411-8.

Liu, Y., Liu, J., Li, N., 2024. Crowdtask data set. https://github.com/nobody3008/CrowdTask.git. Lasted accessed March 10, 2024.

Liu, Y., Yu, Z., Cui, H., Helal, S., Guo, B., 2023. Safecity: A heterogeneous mobile crowd sensing system for urban public safety. IEEE Internet of Things Journal .

Liu, Y., Yu, Z., Guo, B., Han, Q., Su, J., Liao, J., 2020. Crowdos: A ubiquitous operating system for crowdsourcing and mobile crowd sensing. IEEE Transactions on Mobile Computing .

Lu, G., Liu, Y., Wang, J., Wu, H., 2023. Cnn-bilstm-attention: A multi-label neural classifier for short texts with a small set of labels. Information Processing & Management 60, 103320.

Lu, J., Zhang, H., Shen, Z., Shi, K., Yang, L., Xu, B., Zhang, S., Lin, H., 2022. Multi-task hierarchical cross-attention network for multi-label text classification, in: CCF International Conference on Natural Language Processing and Chinese Computing, Springer. pp. 156–167.

Ma, K., Huang, Z., Deng, X., Guo, J., Qiu, W., 2023. Led: Label correlation enhanced decoder for multi-label text classification, in: ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1–5. doi:10.1109/ICASSP49357.2023.10096210.

Mo, W., Liu, A., Xiong, N.N., Song, H., 2024. Mps: A truth discovery service scheme by using history data to maximize profit for mobile crowd sensing. IEEE Transactions on Services Computing .

Shen, J., Xiao, Z., Zhen, X., Snoek, C., Worring, M., 2022. Association graph learning for multi-task classification with category shifts, in: Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., Oh, A. (Eds.), Advances in Neural Information Processing Systems, Curran Associates, Inc.. pp. 4503–4516.

Shi, Z.R., Lizarondo, L., Fang, F., 2021. A recommender system for crowdsourcing food rescue platforms, in: Proceedings of the Web Conference 2021, Association for Computing Machinery, New York, NY, USA. p. 857–865. URL: https://doi.org/10.1145/3442381.3449787, doi:10.1145/3442381.3449787.

Suhag, D., Jha, V., 2023. A comprehensive survey on mobile crowdsensing systems. Journal of Systems Architecture , 102952.

Sun, G., Cheng, Y., Dong, F., Wang, L., Zhao, D., Zhang, Z., Tong, X., 2024. Multi-label text classification model integrating label attention and historical attention. Knowledge-Based Systems 296, 111878.

Sun, L., Li, C., Ding, X., Huang, Y., Chen, Z., Wang, G., Yu, Y., Paisley, J., 2022. Few-shot medical image segmentation using a global correlation network with discriminative embedding. Computers in biology and medicine 140, 105067.

Tian, Y., Zhou, K., Pelleg, D., 2023. Characterization and prediction of mobile tasks. ACM Transactions on Information Systems 41, 1–39.

Wang, J., Wang, L., Wang, Y., Zhang, D., Kong, L., 2018. Task allocation in mobile crowd sensing: State-of-the-art and future opportunities. IEEE Internet of Things Journal 5, 3747–3757. doi:10.1109/JIOT.2018.2864341.

Wang, L., Yang, C., Yu, Z., Liu, Y., Wang, Z., Guo, B., 2019a. Cracksense: A crowdsourcing based urban road crack detection system, in: 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), IEEE. pp. 944–951.

Wang, L., Yu, Z., Zhang, D., Guo, B., Liu, C.H., 2019b. Heterogeneous multi-task assignment in mobile crowdsensing using spatiotemporal correlation. IEEE Transactions on Mobile Computing 18, 84–97. doi:10.1109/TMC.2018.2827375.

Wang, M., Gao, Y., 2024. A multi-label text classification model with enhanced label information, in: 2024 27th International Conference on Computer Supported Cooperative Work in Design (CSCWD), IEEE. pp. 329–334.

Wang, Q., Guo, B., Peng, G., Zhou, G., Yu, Z., 2016. Crowdwatch: Pedestrian safety assistance with mobile crowd sensing, in: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct, pp. 217–220.

Wang, R., Ridley, R., Su, X., Qu, W., Dai, X., 2021. A novel reasoning mechanism for multi-label text classification. Information Processing & Management 58, 102441. doi:https://doi.org/10.1016/j.ipm.2020.102441.

Wang, X., Li, L., Ye, W., Long, M., Wang, J., 2019c. Transferable attention for domain adaptation, in: Proceedings of the AAAI Conference on Artificial Intelligence, pp. 5345–5352.

Wang, Z., Cao, Y., Jiang, K., Zhou, H., Kang, J., Zhuang, Y., Tian, D., Leung, V.C., 2024. When crowdsensing meets smart cities: A comprehensive survey and new perspectives. IEEE Communications Surveys & Tutorials .

Xiang, C., Zhou, Y., Dai, H., Qu, Y., He, S., Chen, C., Yang, P., 2023. Reusing delivery drones for urban crowdsensing. IEEE Transactions on Mobile Computing 22, 2972–2988. doi:10.1109/TMC.2021.3127212.

Xie, Z., Luo, H., Zhang, X., Xiong, H., Zhao, F., Li, Z., Ye, Q., Rong, B., Gao, J., 2023. Transfloor: Transparent floor localization for crowdsourcing instant delivery. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 6, 1–30.

Xu, Y., Fan, Z., Cao, H., 2021. A multi-task text classification model based on label embedding learning, in: China Cyber Security Annual Conference, Springer. pp. 211–225.

Zhang, X., Zhang, Q.W., Yan, Z., Liu, R., Cao, Y., 2021. Enhancing label correlation feedback in multi-label text classification via multi-task learning, in: Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, pp. 1190–1200.

Zhang, Z., Yu, B., Liu, T., Liu, T., Wang, Y., Guo, L., 2023. Learning structural co-occurrences for structured web data extraction in low-resource settings, in: Proceedings of the ACM Web Conference 2023, Association for Computing Machinery, New York, NY, USA. p. 1683–1692. URL: https://doi.org/10.1145/3543507.3583387, doi:10.1145/3543507.3583387.

Zhao, Y., Chen, X., Deng, L., Kieu, T., Guo, C., Yang, B., Zheng, K., Jensen, C.S., 2022. Outlier detection for streaming task assignment in crowdsourcing, in: Proceedings of the ACM Web Conference 2022, Association for Computing Machinery, New York, NY, USA. p. 1933–1943. URL: https://doi.org/10.1145/3485447.3512067, doi:10.1145/3485447.3512067.

Zhou, B., Zhou, H., Wang, W., Chen, L., Ma, J., Zheng, Z., 2024. Hdm-gnn: A heterogeneous dynamic multi-view graph neural network for crime prediction. ACM Transactions on Sensor Networks .