A Prototype Design of LLM-Based Autonomous Web Crowdsensing

Zhengqiu Zhu[1], Yatai Ji[1], Sihang Qiu[1,2,✉], Yong Zhao[1], Kai Xu[1], Rusheng Ju[1,2], and Bin Chen[1,2]

[1] College of Systems Engineering, National University of Defense Technology, Changsha 410073, Hunan Province, China
{zhuzhengqiu12,jiyatai_1209,zhaoyong15,xukai09,chenbin06}@nudt.edu.cn
[2] Hunan Institute of Advanced Technology, Changsha 410073, Hunan Province, China sihangq@acm.org,jrscy@sina.com

Abstract. The expanded demands of complex sensing campaigns involving Cyber-Physical-Social spaces have brought forth a multitude of challenges for web crowdsensing applications, such as substantial human efforts, potential user privacy breaches, and sluggish system response. Significant advancements have occurred in the application of Large Language Models (LLMs) for various tasks, such as conversational engagement, social simulation, and decision-making. Despite this, their potential to empower web crowdsensing activities is under-explored. To bridge this gap, we explore the design of LLM-based autonomous web crowdsensing and correspondingly propose a novel sensing approach, *i.e.*, AutoWebCrowds. Moreover, we use a practical case related to flooding data collection to demonstrate the efficacy of AutoWebCrowds. To facilitate future research, we also propose potential avenues and discuss the challenges. We believe that this approach can mitigate the workload and professional demands on individuals, thereby facilitating faster response and broader popularization of web crowdsensing applications.

Keywords: Web Crowdsensing · Cyber-Physical-Social · Large Language Models · Mobile Web Applications.

1 Introduction

The rapid development of Web of Things and mobile web technologies has fostered the emergence of a social, open, and large-scale sensing paradigm, namely web crowdsensing (also known as social sensors) [6]. It harnesses the collective intelligence of individuals and organizations to enable efficient collection and sharing of real-time information within web applications like Twitter. However, the expanded demands of complex sensing tasks have brought forth a multitude of challenges and issues for it, such as substantial human efforts, sluggish system response, and personal privacy protection concerns [15]. Consequently, these factors impede the broader adoption of this sensing paradigm.

In this context, a novel generation of web crowdsensing is imperative to enable autonomous, intelligent, dependable, and interactive sensing. Large language models (LLMs), such as GPT-4 [1], have gained significant attention due to their exceptional abilities in natural language processing. Recent studies have extended the capabilities of these LLMs beyond text generation, positioning LLMs as versatile agents capable of conversational reasoning and task completion. By leveraging LLMs, web crowdsensing could enhance intelligence throughout the entire process and facilitate interactions among humans, humans and AI, as well as AI with AI, thereby resulting in more effective task comprehension and completion. However, the potential of LLM-based AI agents in empowering web crowdsensing activities remains untapped.

Therefore, we posit that the integration of LLM-based AI agent technology into web crowdsensing holds promise for uncovering valuable insights pertaining to autonomous sensing and the development of mobile web applications. To that end, we investigate the design of LLM-based autonomous web crowdsensing and propose a novel sensing approach called AutoWebCrowds, featuring a four-layer architecture. To answer how this approach works, we conduct a case study on flooding data collection and flooding prediction. Our findings indicate that LLM-based AI Agent technology has the potential to revolutionize web crowdsensing by alleviating human workload and enhancing system response efficiency.

2 Background and Motivation

2.1 Web Crowdsensing

As the usage of web grew over time, the demand for interactivity among users also increased. Consequently, it has morphed into a comprehensive application platform that encompasses desktop, mobile devices, and beyond [13]. The advancement of web technology has given rise to the proliferation of web crowdsensing (also known as social sensors). Countless users actively express their views and leave digital traces on social web platforms like Twitter and Instagram. Meanwhile, web crowdsensing collects, processes, and analyzes multi-source data to offer diverse services, including earthquake detection by Twitter users [9], web-based traffic sentiment analysis [2], interesting place finding by microblogs [6], and etc. Although it has been applied in many fields, the broader adoption of this sensing paradigm is impeded by challenges such as potential user privacy breaches, diminished participation interest, and substantial human involvement.

2.2 Autonomous Agents

In recent two years, prior research has explored the utilization of LLMs as autonomous agents capable of executing tasks via self-directed planning and actions. An exemplary instance is AutoGPT [11], which automates task completion by decomposing tasks into manageable subtasks, supplemented by web searching and information gathering. WorkGPT [5] is another significant contribution, as it serves as an agent framework. Upon receiving an instruction, WorkGPT engages in iterative dialogues with LLMs to effectively accomplish the specified task. Additionally, several studies have explored the potential of autonomous agents to tackle complex tasks and realize advanced reasoning through equipping with external tools or operating smartphone applications, such as GITAGENT [4] and AppAgent [16]. Although these efforts have demonstrated impressive capabilities of AI agents in specific domains, effectively enabling decision-making and task completion in web crowdsensing activities still remains unexplored.

3 The Architecture of AutoWebCrowds

To realize AutoWebCrowds, a prototypical four-layer architecture is proposed in this vision paper, as shown in Fig. 1.

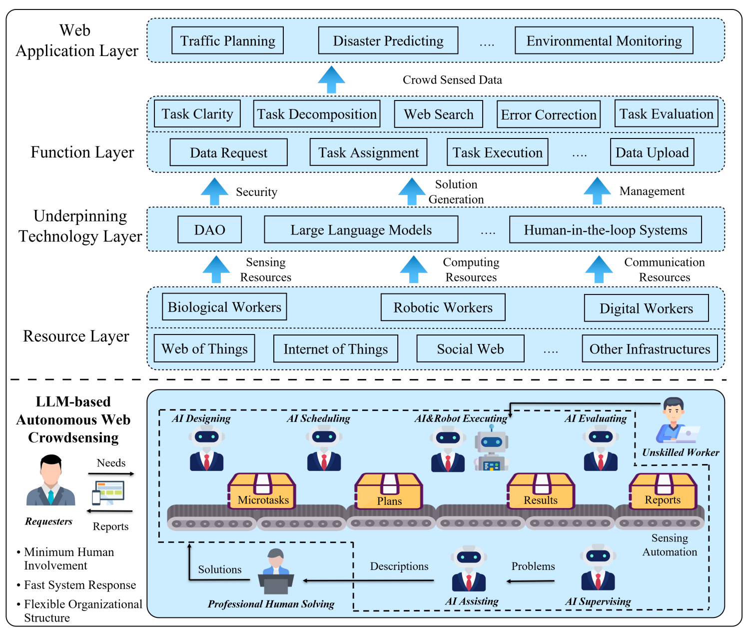

Fig. 1. The architecture of AutoWebCrowds.

3.1 Resource Layer

The resource layer encompasses essential resources required for AutoWebCrowds, including the workforce responsible for executing web-based sensing tasks, web of Things for connecting these resources, and other relevant infrastructure.

Based on previous work [18], the workforce can be divided into three categories: digital, robotic, and biological workers. Specifically, biological workers, who are human individuals actively engaged in the sensing process, contribute their expertise, knowledge, and physical capabilities to perform tasks that necessitate human intervention. While robotic workers are designed to assist biological humans to perform complex real-world tasks. They are often equipped with advanced sensors and shaped with embodied intelligence to autonomously complete sensing tasks that would otherwise be challenging or dangerous for humans. Digital workers serve as virtual representations of both biological and robotic workers. They encompass AI programs, Apps, agents, web assets, and other elements that mainly operate in cyberspace. Digital workers play a vital role in sensing activities by collecting data in cyberspace, creating sensing plans, conducting computational tasks, providing guidance to robotic and biological workers, and enabling online incentive distribution.

3.2 Underpinning Technology Layer

With the emergence and advancement of a series of technologies such as LLMs, autonomous web crowdsensing has emerged as a result. These technologies collectively form the underpinning technology layer.

(1) *Decentralized Autonomous Organizations and Operations (DAO)*. It is an organizational structure that operates based on smart contracts and blockchain technology [17]. In contrast to traditional hierarchical and centralized structures, DAO integrates various sensing resources within distributed, decentralized, autonomous, organized, and orderly communities, thereby facilitating collective decision-making and resource management among a diverse workforce. (2) *Large Language Models*. LLMs obtain extensive knowledge and strong generalization abilities through pre-training on vast corpora [12]. Furthermore, by fine-tuning them with domain-specific data, we can enhance their capacity to tackle a diverse range of complex tasks within web crowdsensing activities, thereby bolstering the autonomy of sensing processes, such as facilitating decision-making and generating solutions for task decomposition. (3) *Human-in-the-loop Systems*. It represents an advanced version of management and computer operating systems designed to facilitate seamless communication and interaction between humans and their digital or robotic counterparts [10]. It effectively alleviates the burdens associated with both physical and cognitive tasks for human individuals [7].

3.3 Function Layer

The current web crowdsensing framework, however, has not gained widespread adoption in real-world applications due to its inherent inflexibility and lack of adaptability. The novel AutoWebCrowds approach aims to establish a highly autonomous and decentralized web-based sensing paradigm, leveraging the aforementioned technologies, in order to overcome the limitations associated with traditional approaches.

As autonomous agents built upon LLMs has already achieved an initial level of human-like intelligence [14], it is believed that a multitude of tasks within web crowdsensing can be effectively accomplished through LLM-based AI agents, allowing for the replacement or assistance of humans and facilitating the automation of crowdsensing campaigns. It autonomously operates in a human-centric manner to address task clarity, task decomposition, web search, task evaluation, and more, but without substantial human efforts. As a result, the response time of web applications can also be significantly reduced. While humans primarily handle unexpected situations or complex problems via HOOS without the need for extensive professional knowledge and skills, which enables widespread participation of non-experts at a reduced cost. Furthermore, DAO facilitates the formation and organization of autonomous web crowdsensing communities by assembling individuals who share common objectives, driven by specific mechanisms for value creation and incentive distribution [3].

3.4 Application Layer

The crowd-sensed data can be applied to various fields including traffic planning, disaster prediction, environmental monitoring, and etc. For instance, in the field of disaster prediction and management, WEBCS-AGENT runs in an autonomous manner and employs tools or algorithms to complete microtasks in different phases. Details can be found in case study (Section 4).

4 Case Study

In this section, a case is dedicatedly designed and conducted to illustrate how AutoWebCrowds works. We first fine-tune the Qwen-7B model[1] and ensure that it is capable of decomposing the complex tasks of web crowdsensing. Additionally, we construct a tool set of web crowdsensing to assist the completion of assigned online microtasks automatically. As shown in Fig. 2, we take the improvement of flooding prediction model of coastal city as an example by leveraging the proposed AutoWebCrowds approach. The hierarchical task decomposition strategy is used to accurately invoke tools to complete sensing-related microtasks. Specifically, the AI agent sequentially performs three phases of work: ***Search***, ***Evaluation***, and ***Correction***.

During the ***Search*** stage, the AI agent searches for relevant data on social media or mobile web applications based on the task objective. In our case, WEBCS-AGENT initially processes the collected dataset, extracting the necessary coordinates and flood depth data for accurate flooding prediction while simultaneously performing essential format conversions. Subsequently, the AI agent utilizes API interface of Baidu map or Google Map to retrieve street view map data corresponding to the given coordinate location. Finally, the street view map data is stored. During the ***Evaluation*** stage, the AI agent extracts and

[1] https://github.com/QwenLM/Qwen-7B

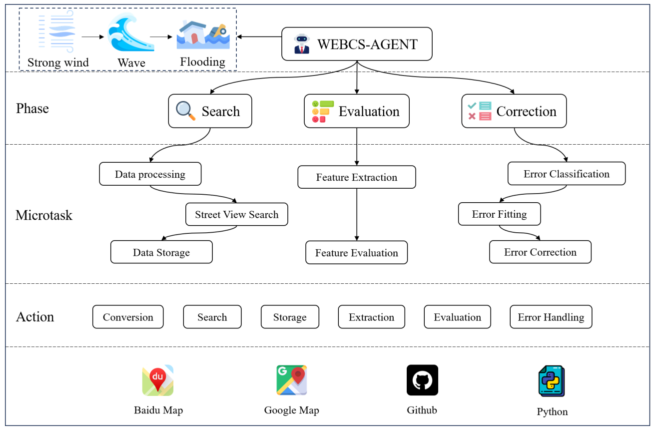

Fig. 2. Illustration of the hierarchical task decomposition strategy of WEBCS-AGENT.

evaluates the stored data. In our case, WEBCS-AGENT would identify the relevant feature elements of street view images (e.g., building density, road width, etc.) and generate specific text descriptions. Based on these descriptions, Based on these descriptions, the AI agent assesses the impact of each feature on flooding errors. During the ***Correction*** stage, the AI agent corrects the flooding error based on the evaluation results. In our case, WEBCS-AGENT employs the decision tree model to classify the actual flooding errors at different coordinates based on the evaluation results of features. The classification outcomes indicate the varying degrees of influence that different features have on flooding errors. WEBCS-AGENT selects the features with high influence to train a support vector regression (SVR) model for predicting flooding errors. Finally, the AI agent utilizes the obtained fitting results from the SVR model to rectify the error outcomes.

Taking the ***search*** stage as an example, we show the actions performed by WEBCS-AGENT after hierarchical task decomposition in Fig. 3. Under the identical action environment, the performance of WEBCS-AGENT is compared with the Qwen-7B model and other open-source LLMs of equivalent size in the web crowdsensing task of flooding prediction of coastal cities. The results demonstrate that our fine-tuned LLM-based AI agent exhibits enhanced accuracy in task decomposition and execution of corresponding web-based microtasks. Since most sensing activities are completed by LLM-based AI agents in this way, thereby greatly reducing human participation as well as facilitating rapid responses and wide application of web crowdsensing systems.

5 Discussions and Conclusions

In this work we investigate the design of LLM-based autonomous web crowdsensing and propose a four-layer sensing approach called AutoWebCrowds. A case

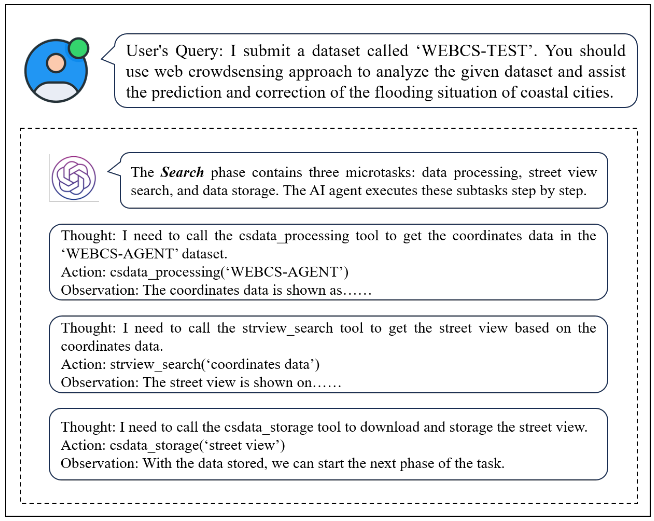

Fig. 3. Example of WEBCS-AGENT execution process in the ***search*** phase.

study is conducted on flooding data collection and flooding prediction, revealing the capability of LLMs empowering the entire life cycle of web-based sensing activities in an autonomous mode with minimal human intervention. The current work, however, still has certain limitations including both inherent characteristics of LLMs themselves and challenges associated with their integration within the autonomous sensing workflow. Inherent limitations of LLMs include issues such as hallucination and high resource requirements during inference and training. The application limitations primarily manifest in the requirement for LLMs to adapt and respond according to different application scenarios. Moreover, this work solely focuses on a solitary AI agent role and has not yet explored the potential of collaborative diverse workforce, which would pose challenges to the organization and operation of diverse workforce.

The future is expected to witness the emergence of a novel human-oriented web crowdsensing paradigm through multi-agent dialogue, which will become the mainstream [8]. Thus, the current study can be extended from many aspects, such as conversational multi-agent system design, complex social network-based web crowdsensing, DAO-based workforce organization. Additionally, we will further explore the integration of LLMs and web crowdsensing, and enrich the crowdsensing tool set by releasing it as an open-source solution. In this way, the organization of various activities in web crowdsensing will be reshaped, giving rise to a more autonomous, intelligent, interactive, and human-centered sensing and computing.

References

1. Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F.L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., et al.: GPT-4 technical report, https://arxiv.org/pdf/2303.08774.pdf
2. Cao, J., Zeng, K., Wang, H., Cheng, J., Qiao, F., Wen, D., Gao, Y.: Web-based traffic sentiment analysis: Methods and applications. IEEE transactions on Intelligent Transportation systems **15**(2), 844–853 (2013)
3. Li, J., Wang, F.Y.: The TAO of blockchain intelligence for intelligent web 3.0. IEEE/CAA Journal of Automatica Sinica **10**(12), 2183–2186 (2023)
4. Lyu, B., Cong, X., Yu, H., Yang, P., Qin, Y., Ye, Y., Lu, Y., Zhang, Z., Yan, Y., Lin, Y., et al.: Gitagent: Facilitating autonomous agent with github by tool extension, https://arxiv.org/pdf/2312.17294.pdf
5. OPENPM, T.: WorkGPT, https://github.com/team-openpm/workgpt
6. Pandharipande, A.: Social sensing in iot applications: A review. IEEE Sensors Journal **21**(11), 12523–12530 (2021)
7. Qiu, S., Bozzon, A., Birk, M.V., Gadiraju, U.: Using worker avatars to improve microtask crowdsourcing. Proceedings of the ACM on Human-Computer Interaction **5**(CSCW2), 1–28 (2021)
8. Qiu, S., Gadiraju, U., Bozzon, A.: Just the right mood for hit! analyzing the role of worker moods in conversational microtask crowdsourcing. In: Web Engineering: 20th International Conference, ICWE 2020, Helsinki, Finland, June 9–12, 2020, Proceedings 20. pp. 381–396. Springer (2020)
9. Sakaki, T., Okazaki, M., Matsuo, Y.: Earthquake shakes twitter users: real-time event detection by social sensors. In: Proceedings of the 19th international conference on World wide web. pp. 851–860 (2010)
10. Shen, Y., Liu, Y., Tian, Y., Na, X.: Parallel sensing in metaverses: Virtual-real interactive smart systems for "6S" sensing. IEEE/CAA Journal of Automatica Sinica **9**(12), 2047–2054 (2022)
11. SIGNIFICANT-GRAVITAS: AutoGPT, https://github.com/Significant-Gravitas/Auto-GPT
12. Tian, Y., Li, X., Zhang, H., Zhao, C., Li, B., Wang, X., Wang, F.Y.: VistaGPT: Generative parallel transformers for vehicles with intelligent systems for transport automation. IEEE Transactions on Intelligent Vehicles (2023)
13. Vepsäläinen, J., Hellas, A., Vuorimaa, P.: The rise of disappearing frameworks in web development. In: International Conference on Web Engineering. pp. 319–326. Springer (2023)
14. Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., Zhao, W.X., Wei, Z., Wen, J.R.: A survey on large language model based autonomous agents, https://arxiv.org/pdf/2308.11432.pdf
15. Wu, W., Yang, W., Li, J., Zhao, Y., Zhu, Z., Chen, B., Qiu, S., Peng, Y., Wang, F.Y.: Autonomous crowdsensing: Operating and organizing crowdsensing for sensing automation. IEEE Transactions on Intelligent Vehicles (2024)
16. Yang, Z., Liu, J., Han, Y., Chen, X., Huang, Z., Fu, B., Yu, G.: Appagent: Multimodal agents as smartphone users, https://arxiv.org/pdf/2312.13771.pdf
17. Yao, S., Yu, J., Yu, Y., Xu, J., Dai, X., Li, H., Wang, F.Y., Lin, Y.: Towards integrated traffic control with operating decentralized autonomous organization. arXiv preprint arXiv:2308.03769 (2023)
18. Zhao, Y., Hu, C., Zhu, Z., Qiu, S., Chen, B., Jiao, P., Wang, F.Y.: Crowd sensing intelligence for its: Participants, methods, and stages. IEEE Transactions on Intelligent Vehicles (2023)