Poster: Large Language Model-powered Wi-Fi-based Human Activity Recognition

Yili Ren‡*, Haopeng Zhang†*, Haohan Yuan†, Jingzhe Zhang‡, Yitong Shen‡
University of South Florida‡, ALOHA Lab, University of Hawaii at Manoa†
{yiliren, jingzhe, shen202}@usf.edu, {haopengz, haohany}@hawaii.edu

ABSTRACT

Recent advances in LLMs have shown exceptional reasoning capability. However, their ability to integrate physical model knowledge for real-world signal interpretation remains largely unexplored. We introduce Wi-Chat, an LLM-powered Wi-Fi-based human activity recognition system. By embedding Wi-Fi sensing principles into prompts, we show that LLMs can infer human activities through Wi-Fi signals in a zero-shot manner without complex signal processing.

CCS CONCEPTS

• **Human-centered computing → Ubiquitous and mobile computing systems and tools**; • **Computing methodologies → Artificial intelligence**.

KEYWORDS

Large Language Model, Wi-Fi Sensing, Activity Recognition

ACM Reference Format:
Yili Ren‡*, Haopeng Zhang†*, Haohan Yuan†, Jingzhe Zhang‡, Yitong Shen‡. 2025. Poster: Large Language Model-powered Wi-Fi-based Human Activity Recognition. In *The 23rd ACM International Conference on Mobile Systems, Applications, and Services (MobiSys '25), June 23–27, 2025, Anaheim, CA, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3711875.3734545

1 INTRODUCTION

Large Language Models (LLMs) exhibit remarkable reasoning capabilities and exceptional generalization skills [4]. However, their reliance on Internet-based textual training limits their understanding of the physical world. Meanwhile, the ubiquitous Wi-Fi devices and signals offer an opportunity to extend Wi-Fi beyond communication to sensing physical environments, including people and objects [1]. However, traditional Wi-Fi sensing systems rely on complex signal processing and labor-intensive machine learning. This raises a fundamental question: can we integrate LLMs with Wi-Fi sensing to comprehend the physical world in a zero-shot manner without complex signal processing?

We explore this question by examining LLMs' ability to interpret raw Wi-Fi signals with guidance from physical models of Wi-Fi sensing. Specifically, we introduce Wi-Chat, the first LLM-powered Wi-Fi sensing system for human activity recognition. Wi-Chat processes raw signals using well-known LLMs, such as ChatGPT. Additionally, by integrating Wi-Fi sensing principles into prompts, we enhance LLMs' understanding of human activity in the physical world. We evaluate Wi-Chat by comparing its performance against traditional Wi-Fi-based recognition systems and fundamental machine learning models. Our results demonstrate that LLMs can perform zero-shot human activity recognition directly from raw Wi-Fi signals, achieving an accuracy of 90%.

2 SYSTEM DESIGN

As shown in Fig. 1, a Wi-Fi transmitter emits signals that are received by receivers to probe a person and the system extracts Wi-Fi channel state information (CSI) measurements. We analyze human activities based on the physical models of Wi-Fi sensing and derive the LLM prompts accordingly. *a) Signals of Walking:* Walking is a large-scale activity that occurs over an extended period and causes significant changes in CSI amplitude. These changes can reach both maximum and minimum values, resulting in multiple peaks and troughs in the CSI (Fig. 2(a)). *b) Signals of Falling:* Falling is also a large-scale movement. However, falls typically occur over a short duration, leading to concentrated peaks and troughs in the signal. After smoothing, these rapid fluctuations can often be approximated by a single prominent peak or trough, followed by a potential static period (Fig. 2(b)). *c) Signals of Breathing:* Breathing is a small-scale activity, with chest expansion and contraction typically spanning only a few centimeters. Consequently, the overall CSI amplitude may not reach its maximum and minimum values. Moreover, breathing is a continuous and smooth process that persists over an extended duration (Fig. 2(c)). *d) Signals of No-event:* Since there are no dynamic signal components, the overall CSI

*These authors contributed equally to this work.

MobiSys '25, June 23–27, 2025, Anaheim, CA, USA

ACM ISBN 979-8-4007-1453-5/25/06.
https://doi.org/10.1145/3711875.3734545

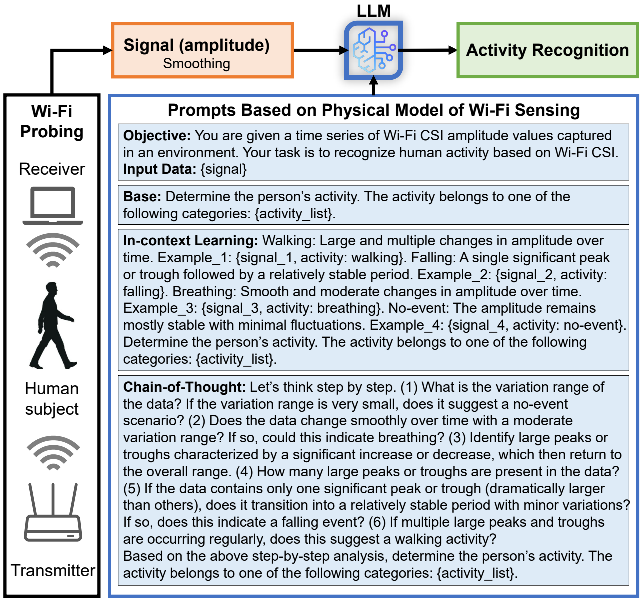

Figure 1: System overview.

amplitude is determined solely by static signal components and remains nearly constant over time (Fig. 2(d)).

Next, we outline different prompting strategies for leveraging LLMs in Wi-Fi-based human activity recognition. *a) Base:* The LLM receives raw CSI amplitude data and is prompted to infer human activity directly. *b) In-context Learning (ICL):* LLMs exhibit strong few-shot learning capabilities across various tasks, known as ICL. By using Wi-Fi sensing knowledge and exemplars in inference, models recognize signal patterns and improve accuracy. *c) Chain-of-Thought (CoT):* CoT enhances interpretability by incorporating explicit knowledge and intermediate steps into prompts, helping the model better capture relationships between signal patterns and human activities. The formulated prompts are shown in Fig. 1.

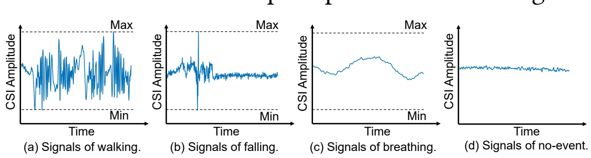

Figure 2: Signals of human activities.

3 PERFORMANCE EVALUATION

We conducted experiments using commodity Wi-Fi devices with Intel 5300 NICs as both transmitters and receivers for data collection. Participants performed one of four activities: walking, falling, breathing, or no event. The dataset includes over 1,965,000 Wi-Fi CSI packets from individuals of varying heights, weights, and ages, collected across three real-world environments (i.e., bedroom, kitchen, and living room), over two months. We compare Wi-Chat with the following baselines: 1) Traditional Wi-Fi-based systems, such as CARM [2] and E-eyes [3], which rely on complex signal processing and model construction. 2) Machine learning models, including CNN, RNN, and SVM, that process raw signals. Baselines are evaluated in both supervised and zero-shot settings. In the supervised setting, we randomly split the dataset into 70% for training and 30% for testing.

Method	Accuracy	F1-score
Zero/Few-shot		
E-eyes (zero-shot)	0.26	0.26
CARM (zero-shot)	0.24	0.24
SVM (zero-shot)	0.27	0.27
CNN (zero-shot)	0.23	0.23
RNN (zero-shot)	0.26	0.26
Wi-Chat: GPT-4o: base (zero-shot)	0.47	0.42
Wi-Chat: GPT-4o: ICL (4-shot)	0.77	0.73
Wi-Chat: GPT-4o: CoT (zero-shot)	**0.90**	**0.90**
Signal Processing + Supervised		
CARM	0.98	0.98
E-eyes	1.00	1.00

Table 1: Performance comparison of different systems.

Table 1 summarizes the results across different systems. In the zero-shot setting, traditional Wi-Fi-based systems and machine learning models show relatively low performance. The LLM model GPT-4o achieves an average accuracy of 0.47, outperforming traditional models. The ICL approach significantly improves the performance by integrating Wi-Fi sensing knowledge, reaching an accuracy of 0.77. Additionally, GPT-4o combined with knowledge-based CoT achieves the highest accuracy of 0.90. Notably, this is comparable to that of traditional Wi-Fi-based systems using complex signal processing and supervised models.

4 CONCLUSION

In this work, we introduced Wi-Chat, the first LLM-powered Wi-Fi-based human activity recognition system that integrates the reasoning capabilities of large language models with the sensing capabilities of wireless signals. Our experimental results demonstrate the promising potential of LLMs in enabling zero-shot Wi-Fi sensing.

REFERENCES

[1] S. Tan, Y. Ren, J. Yang, and Y. Chen. 2022. Commodity WiFi sensing in ten years: Status, challenges, and opportunities. *IEEE Internet of Things Journal* 9, 18 (2022), 17832–17843.

[2] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu. 2015. Understanding and modeling of wifi signal based human activity recognition. In *Proceedings of the 21st annual international conference on mobile computing and networking*. 65–76.

[3] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu. 2014. E-eyes: Device-free location-oriented activity identification using fine-grained WiFi signatures. In *Proceedings of the 20th annual international conference on Mobile computing and networking*. 617–628.

[4] J. Wei, X. Wang, D. Schuurmans, M. Bosma, E. Chi, Q. Le, and D. Zhou. 2022. Chain of thought prompting elicits reasoning in large language models. *arXiv preprint arXiv:2201.11903* (2022).