IJSRIS journal

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH AND INNOVATIVE STUDIES

ISSN: 2820-7157

www.ijsrisjournal.com

June 2024 **Volume 3** **Number 3** **Pages 16-19**

Received Date: April 10, 2024 **Accepted Date:** May 08, 2024 **Published Date:** June 01, 2024

Available Online at https://www.ijsrisjournal.com/index.php/ojsfiles/article/view/151

https://doi.org/10.5281/zenodo.11237705

Understanding User Intention in Pervasive Environments: A Literature Review and Perspectives

Abdelhadi Bouain[1], Mohamed Nezar ABOURRAJA[2], Mohamed Yassine Samiri[3]

[1]A LabSI, Polydisciplinary Faculty of Ouarzazate,
Ibnou Zohr University, Morocco,
a.bouain@uiz.ac.ma

[2]Laboratory of Artificial Intelligence Research and Applications (AIRA)
Smart Factory and Robotics, Hassan 1 University, Morocco,
mnabour@kth.se

[3]Cadi ayyad university, Morocco,
m.samiri@uca.ac.ma

ABSTRACT

Identifying user intention (what the user wishes to achieve within a system) with minimal or ideally no direct user interaction is a major goal in pervasive computing. Achieving this goal requires a clear and consistent definition of intention, a concept widely used but understood differently across various studies. In this work, we first aim to clarify the different interpretations of intention, distinguishing between implicit and explicit intention. Subsequently, we compare various existing approaches from the literature, seeking to reconcile these diverse viewpoints and establish a common foundation for future research efforts.

Key words: Pervasive computing, User intention, User intent prediction, Multimodal Large Language Models**.**

1. INTRODUCTION

In pervasive information systems (PIS), human-machine interaction (HMI) fundamentally differs from traditional information systems (IS) that rely on explicit communication interfaces [1], [2], [3]. Indeed, the principle of invisibility assumes that multiple small-sized devices and computers are distributed and integrated into the environment, interacting with each other to provide a set of services to users. The pervasive information system should enable the discovery of these services with minimal Human-Machine Interaction and maximum transparency, concealing the heterogeneity of services, protocols, and all other technical details from users who may not necessarily be computer experts[1], [2], [3], [4].

Two solutions are possible: The first solution involves using a formal language to express the user's need (user intention). However, this approach can quickly lose its effectiveness in dynamic and open environments with a large number of casual users.

An alternative is to allow users to express their needs (intentions) using non-technical terms, ideally in natural language. The SIP must then reformulate these intentions to facilitate the identification of the services that best meet the needs (intentions) expressed by the users [4], [5]. The second solution involves the system analyzing and identifying the user's need (intention) without the user having to express it explicitly, and then returning the expected result to the user.

The first interpretation views the user's intention as a goal actively expressed by the user, while the second focuses on predicting and anticipating the user's needs. In this article, we explore the evolution of the concept of user intention in pervasive computing. We compare explicit and implicit intentions, examine user intention prediction systems, and ultimately identify the models best suited for such environments.

2. WHAT IS THE "USER'S INTENTION"?

Identifying or anticipating what the user wishes to obtain or achieve during their interaction with a Pervasive Information System (PIS) allows these systems to seamlessly integrate into the daily life and be more proactive, in perfect accordance with the principles of pervasive computing [1], [2], [3]. This is why a clear and concise definition of the user's intention is of crucial importance.

In the literature, various definitions are provided for the concept of intention in pervasive environments. Some define it as an explicit statement by the user regarding a desired goal or state, while others recognize it as an implicit goal inferred from contextual information.

2.1 The intention as an explicit goal

From this perspective, in pervasive environments, users might not be computer experts. Therefore, they should be able to express their needs using natural language, describing their goals without getting bogged down by technical details of existing services. For example, "check unpaid invoices," "look for a hotel reservation in Marrakech," "know the urgency level of a patient", etc.[4], [6]. These goals are referred to as "Intentions". The user's intention in a pervasive environment can also be considered as a high-level description of their needs, specifying what they expect from a service without indicating how to achieve it[5], [7].

2.2 The intention as an implicit goal

Implicit intention aims to identify or anticipate the user's goal or desired outcome, even without explicit user expression. The system deduces this goal or action by observing and analyzing the user's activities and context. According to this definition, the user's intention is an unexpressed goal inferred from contextual information such as the user's profile, history, sensor data, etc. Implicit intention aims to reduce user interaction with the pervasive information system (PIS) and adhere to the principle of invisibility (a key tenet of ubiquitous computing) by anticipating user needs[5], [8].

In general, both explicit and implicit intention definitions can coexist in pervasive environments, complementing each other. Several studies show context, when combined with user-expressed intentions, enhances precision and meaning [5], [7], [9], [10], [11].

3. MODELING USER INTENTIONS

The distinction between implicit and explicit intentions necessitates different approaches for modeling them.

3.1 Modeling explicit intentions

We will first present approaches using non-natural languages to express user intention, and then we will explore those based on natural language.

A. The Bihler Approach

Among the early approaches that focused on user intentions in pervasive environments, the Bihler et al. approach [12] stands out, aiming to translate user intentions into a set of executable actions. To model intentions, this approach uses a formal language called PsaQL (Pervasive Service Action Query Language), a language similar to SQL. Starting from the intention expressed formally in PsaQL (called partial action), Bihler then seeks to find the best combination of services to satisfy this intention using service descriptions, action history, and contextual information.

This approach does not explain how to transform an intention expressed in natural language into a formal language like PsaQL (Pervasive Service Action Query Language). Also, the collection, processing, modeling, and reasoning on contextual information are not detailed in this approach. Additionally, forcing users to express their intentions in PsaQL is not a practical solution, especially in dynamic and open pervasive environments.

B. Natural Language-Based Approaches

Expressing an intention with natural language to interact with a pervasive information system involves the use of NLP (Natural Language Processing) and AI (Supervised Learning, Unsupervised Learning, Collaborative Filtering, Recurrent Neural Networks, Deep Belief Networks, etc.) to recognize, classify user intention, and create intention recognition systems [13], [14], [15]. Several pre-trained models exist in the literature, such as the BERT model (Bidirectional Encoder Representations from Transformers) and its variations (BERT4Rec, BioBERT, RoBERTa, SBERT, etc.), GPT (Generative Pre-trained Transformer), ERNIE (Enhanced Representation through Knowledge Integration), GRU4Rec, BiLSTM (Bidirectional Long Short-Term Memory), etc. In addition, various datasets of varying sizes are used for training and testing different models [13], [14], [15].

The intention modeling from this perspective aims to identify the structure and meaning of the intention, sentiments, and any other useful information that can help detect user intention.

3.2 Modeling implicit intentions

Modeling implicit intention involves developing a predictive system that analyzes the user's history, ongoing activities, and other relevant contextual information to anticipate their needs. Proposing such a system is a complex task, leading several studies to focus solely on specific application domains, such as predicting user intent in search conversations [14], [15], [16]. These studies employ various machine learning models like CNNs, BiLSTMs, and BiLSTM-Context models, enriched with contextual information to enhance their accuracy. Additionally, research on implicit intentions delves into the field of online shopping, where user behavior is tracked and analyzed across different platforms to predict purchase intentions [17].

Unfortunately, these approaches fall short in pervasive environments, characterized by dynamism and openness [2], [3]. In these environments, users should have minimal interaction with the system, which needs to be proactive (collecting and processing various contextual data to predict user needs). Several frameworks address other data types and aspects of pervasive environments. These include eye-tracking to predict the intentions of individuals with limited motor abilities [18] and analysis of pedestrian trajectories, movements, and images to predict their intent to cross the road[8]. Additionally, similar frameworks are used to predict vehicle intentions to change direction or perform maneuvers on the road [19], [20], [21]. However; all these proposals focus on specific domains and typically use data types with well-defined and restricted contextual information.

4. COMPARISON BETWEEN USER INTENT PREDICTION SYSTEMS

User intention prediction systems aim to anticipate user requests and respond based on the analysis of collected environmental information (contextual information, user history and actions, sensors, etc.). In the literature, several user intention prediction systems exist, employing various techniques ranging from analyzing clicks and user history on a website to autonomous driving in cars.

Table 1 presents some recent works on user intention prediction classified according to various criteria: types of collected data, machine learning methods, deep learning, and NLP techniques used, as well as application domains and results obtained.

Table 1: Comparison between some recent works on UIP systems

Work	Type of Data Collected	Methods (ML, DL, NLP) Used	Application Domain	Result
[16]	User profile (basic user information) Historical behavior (user's recent interests)	Attention-based Deep Multiple Instance Learning (MIL, LSTM)	Customer service bot	6.29% improvement in performance compared to the best reference model.
[21]	Roads, lanes, intersections, crossings, traffic signs, traffic lights, etc.	IntentNet (Fully Convolutional Network – FCN)	Self-driving vehicles	Predict lane changes and turns better than reference models.
[8]	Sequential images	PedGNN (GNN-GRU)	Pedestrian intention prediction	F1-score of approximately 92%, lighter and faster than PedGraph+.
[18]	Eye movement	Hidden Markov models (HMMs), Transfer learning, CNN-LSTM, DBSCAN	People with reduced mobility or limited communication	The model achieved an average classification accuracy of 97.42%.
[20]	High-resolution (HR) images	High-Resolution information + Multimodal large language models (HiLM-D)	Autonomous driving system	Improvements of 4.8% in BLEU-4 for caption generation and 17.2% in mIoU for detection.
[15]	Utterances	Traditional machine learning (ML) methods	Predict user intent in information-seeking conversations	Random forest and AdaBoost achieve the best overall performance among all baseline classifiers.

In general, predicting user intention in the pervasive environment requires complex models capable of combining multiple data sources (images, videos, text, sound, gestures, etc.), in addition to collecting information about the user's history and activities. This involves the use of machine learning models, deep learning, and natural language processing. That's why Extended Multimodal Language Models (MLLMs) [19], [20] provide a very compelling solution for modeling intentions in pervasive environments, as they integrate not only text but also non-textual information such as images, videos, audio, and more .

5. CONCLUSION

The concept of user intention in pervasive computing has evolved with advancements in the field of AI (machine learning, deep learning, and NLP), transitioning from a simple expressed goal to anticipating the actions and future needs of a user without direct interaction with them. After clarifying this evolution and presenting some recent works in the field of intention prediction, we are convinced that Multimodal Large Language Models (MLLMs) provide a strong foundation for modeling intentions in pervasive environments, given their significant ability to integrate different types of data (text, image, sound, etc.) and make decisions with performance comparable to that of humans.

REFERENCES

[1] C. Becker, C. Julien, P. Lalanda, et F. Zambonelli, **Pervasive computing middleware: current trends and emerging challenges**, *CCF Transactions on Pervasive Computing and Interaction*, vol. 1, nº 1, p. 10-23, mai 2019, doi: 10.1007/s42486-019-00005-2.

[2] M. Kirsch Pinheiro, P. Roose, L. A. Steffenel, et C. Souveyet, **What Is a Pervasive Information System (PIS)? ,** *in The Evolution of Pervasive Information Systems*, 2023, p. 1-17. doi: 10.1007/978-3-031-18176-4_1.

[3] S. G. Gollagi, M. M. Math, et A. A. Daptardar, **A survey on pervasive computing over context-aware system** , *CCF Transactions on Pervasive Computing and Interaction*, vol. 2, nº 2, p. 79-85, juin 2020, doi: 10.1007/s42486-020-00030-6.

[4] A. Bouain, A. El Fazziki, et M. Sadgal, **Pervasive services vs. Web services: Survey and comparison** , in *2014 International Conference on Multimedia Computing and Systems (ICMCS)*, 2014, p. 552-557. doi: 10.1109/ICMCS.2014.6911148.

[5] A. A. Jaafar, D. N. A. Jawawi, M. Adham Isa, et N. A. Saadon, **Service selection model based on user intention and context** , *Journal of King Saud University - Computer and Information Sciences*, vol. 35, nº 4, p. 209-223, avr. 2023, doi: 10.1016/j.jksuci.2023.03.018.

[6] A. Bouain, A. E. Fazziki, M. Sadgal, et M. N. Abourraja, **Towards a pervasive information system for patient triage and referral in emergency departments**, *IJAHUC*, vol. 25, nº 4, p. 273, 2017, doi: 10.1504/IJAHUC.2017.085132.

[7] I. Choukri, H. Guermah, et M. Nassar, **Towards a Generic Architecture of Context-Aware and Intentional System**, in *Proceedings of the 5th International Conference on Big Data and Internet of Things*, 2022, p. 274-285.

[8] M. N. Riaz, M. Wielgosz, A. G. Romera, et A. M. Lopez, **Synthetic Data Generation Framework, Dataset, and Efficient Deep Model for Pedestrian Intention Prediction** , 2024, doi: 10.48550/ARXIV.2401.06757.

[9] K. Rook, B. Witt, R. Bailey, J. Geigel, P. Hu, et A. Kothari, **A Study of User Intent in Immersive Smart Spaces** , in *2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, Kyoto, Japan: IEEE, mars 2019, p. 227-232. doi: 10.1109/PERCOMW.2019.8730692.

[10] V. Ponce et B. Abdulrazak, **Intention as a Context: An Activity Intention Model for Adaptable Development of Applications in the Internet of Things** , *IEEE Access*, vol. 9, p. 151167-151185, 2021, doi: 10.1109/ACCESS.2021.3126000.

[11] T. Tong, R. Setchi, et Y. Hicks, **Context change and triggers for human intention recognition** , *Procedia Computer Science*, vol. 207, p. 3826-3835, 2022, doi: 10.1016/j.procs.2022.09.444.

[12] P. Bihler, V.-M. Scuturici, et L. Brunie, **Expressing and Interpreting User Intention in Pervasive Service Environments** , *Journal of Digital Information Management*, vol. 4, nº 2, p. 102-106, 2006.

[13] T. Wu, M. Wang, Y. Xi, et Z. Zhao, **Intent recognition model based on sequential information and sentence features** , *Neurocomputing*, vol. 566, p. 127054, janv. 2024, doi: 10.1016/j.neucom.2023.127054.

[14] S. Farshidi *et al.*, **Understanding User Intent Modeling for Conversational Recommender Systems: A Systematic Literature Review**, 2023, doi: 10.48550/ARXIV.2308.08496.

[15] C. Qu, L. Yang, W. B. Croft, Y. Zhang, J. R. Trippas, et M. Qiu, **User Intent Prediction in Information-seeking Conversations**, in *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval*, Glasgow Scotland UK: ACM, mars 2019, p. 25-33. doi: 10.1145/3295750.3298924.

[16] K. Xu *et al.*, **aDMSCN: A Novel Perspective for User Intent Prediction in Customer Service Bots** , in *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, 2020, p. 2853-2860. doi: 10.1145/3340531.3412683.

[17] M. Hendriksen, E. Kuiper, P. Nauts, S. Schelter, et M. de Rijke, **Analyzing and Predicting Purchase Intent in E-commerce: Anonymous vs. Identified Customers**, 2020, doi: 10.48550/ARXIV.2012.08777.

[18] F. Koochaki et L. Najafizadeh, **A Data-Driven Framework for Intention Prediction via Eye Movement With Applications to Assistive Systems**, *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, p. 974-984, 2021, doi: 10.1109/TNSRE.2021.3083815.

[19] C. Cui *et al.*, **A Survey on Multimodal Large Language Models for Autonomous Driving** , 2023, doi: 10.48550/ARXIV.2311.12320.

[20] X. Ding, J. Han, H. Xu, W. Zhang, et X. Li, **HiLM-D: Towards High-Resolution Understanding in Multimodal Large Language Models for Autonomous Driving**. arXiv, 2023. doi:10.48550/arXiv.2309.05186.

[21] S. Casas, W. Luo, et R. Urtasun, **IntentNet: Learning to Predict Intention from Raw Sensor Data**, 2021, doi: 10.48550/ARXIV.2101.07907.