An Intention Mining Approach using Ontology for Contextual Recommendations

Ramona Elali

Centre de Recherche en Informatique,
University of Paris 1 Panthéon–Sorbonne,
Paris, France
ramona.elali@etu.univ-paris1.fr

Abstract. Most modern Information Systems record user activity logs to extract knowledge and support new and higher quality services. Recommender systems (RS) in particular those based on collaborative filtering, are an example of such new services that have gained tremendous popularity. However, we know that standard activity logs are too poor to properly reflect user activity as they are influenced by many environmental factors (time, location, weather, profile, etc.), out of which a lot is not grasped in standard user activity logs. This makes it very difficult to mine users' intentions, which are key to analyze users' behavior reliably, and therefore key to quality recommendations. Several types of sources, such as sensors, external systems, outside actors, domain knowledge bases, or forecast systems could be used to obtain this information. The theory in this PhD is that the combination of multiple types of sources can extensively contribute to providing better insight into user activity through context-rich intentional process mining and therefore deliver newer and higher quality recommendations. While existing research mainly focuses on log datasets; only a few contributions consider ontologies to gather multiple sources. We propose a novel approach that (a) combines different types of sources into ontologies (b) uses such ontologies for intentional process mining (c) exploits this intentional model for contextual recommendations.

Keywords: Intention Mining, Information System, Logs, Ontology, Context, Recommendations.

1 Introduction

With the growth of the internet and information systems, data has become available everywhere. However, most of this data is neither structured nor well organized [16]. Consequently, a major problem is to identify what users are truly looking for [17]. This problem has been partially solved by innovations in the domain of Recommender

Systems (RS) [18]. RS have become quite inevitable; they are expected to give suggestions to users on what to buy, watch, or read, etc., while pertinently taking into consideration users' preferences and actual interests. The recommended items should somehow match what the users will really choose [17] by filtering out irrelevant alternatives, i.e., those that the user will not select [16].

RS are very important to many businesses; it is shown that they can increase their income significantly, create a competitive advantage, and even create new business opportunities. Known examples are Netflix or Amazon. For instance, Netflix held a competition to offer 1 million dollars for the one who creates a recommender algorithm that outperformed Netflix's algorithm [20].

Many recommendation algorithms have been developed and used; two of them are regularly used in the industry: Collaborative Filtering and Content-based filtering. However, these algorithms are known to have a cold start problem: the suggestions they deliver lack accuracy when input data are absent or sparse [25]. Another concrete issue is due to the fact that many applications or websites allow users to interact anonymously [19]. For example, an unregistered user can browse through an online shop website, add items to his cart and check them out later, or an unlogged user may watch videos on YouTube. The data are then difficult to exploit for these algorithms. We believe that we should always be able to recommend pertinent items to a user, whether or not we have explicit data about him.

Existing recommendation systems tend to study the correlation between the items and users' aspects without focusing on what the users really need, but this issue can be solved by Process mining (PM). PM focuses on the activities that are generated from a business process. Its purpose is to discover, check the conformance and enhance the process models that are based on event logs [1]. In addition, PM can be used as a recommendation technique to show the user what activity should follow according to his current activity [9].

However, users tend to be goal-centric [2] and PM does not address the intentional part of the process as it uses only activity-oriented models. Intentional models are more accurate in representing the user's thinking and detect the human reasoning behind the produced activities. Intention mining (IM) is an emerging field of research that is derived from PM, has the same goals, but precisely addresses intentional process models. Hence, IM is considered a promising technique for recommendation systems since it tackles the intentional part of the user activity [9].

IM has become an important research topic in many areas of computer science domains. The main objective of IM is to identify the user goals and purpose "on the fly" by discovering the intentional process models reflecting the user behavior and strategies. PM states that the actual processes can be extracted from event logs and that they can be different from the prescribed business processes. The main reason for this difference is that the user doesn't rely only on the prescribed processes to complete his activities, but he also uses his own intentions and strategies to complete his tasks. Therefore, the behavior of the user depends on his intention and his goals, which can be defined previously or instantly. PM models are defined as a sequence of steps that don't support variability. In contrast, MAP [3] is an intentional process model that can

be used to model processes according to the user's intentions while providing flexibility to enact different strategies in order to achieve a specific intention.

The user's intention or strategies are affected by multiple factors such as location, time, weekday, weather, profile, season, politics, different kinds of events. Those factors constitute the contextual environment of the intention and its strategies. It is thus crucial to study multiple sources of data to gain all required and additional information that compose the context. Ontology has become an evolving research field in the domain of information systems [30]. It is used by many researchers to classify domain knowledge [4] such as concepts, all types of entities, and the existing relationships between them [30]. Hence, we can structure contextual information in domain ontology by describing the relation between the entities of the domain ontology and the contextual information. Later, IM can rely on these domain ontologies to better build an intentional process model (MAP). Hereafter, to help the user to achieve his goal and his task in an efficient way, the discovered intentional models will be used to provide good recommendations (by giving the user the necessary knowledge and by determining the optimal sequence of actions to achieve his intention). Consequently, useful users' recommendations can be provided using a combination of mined intentional process models, domain ontologies, and contextual information.

Thus, throughout this PhD program, our main objective will be to propose a new approach to make recommendations using IM based on the combination of intentional process models, domain ontologies, and contextual information. The study will include traces of the user activities since those types of data can be fetched from different types of input such as log traces, sensors, etc. In addition, it will include a domain ontology useful to build the intentional process model. Secondly, our objective is to provide the users with recommendations generated from the intentional process model and to propose an adapted recommendation technique guiding the user while enacting a process.

The remainder of the paper will be as follows: the next section will focus on the research methodology; Section 3 will describe our proposal and in Section 4 we will present the related works.

2 Research Methodology

The research work will be conducted in 4 phases as Fig. 1 shows. This research process was elaborated consistently with the key principles of design science research [26]. The main aspect is that this process will be used in a recursive way, for instance, while elaborating the solution, hard sub-problems might need to be addressed raising new research questions and requiring specific literature study and side proposal elaboration/implementation and evaluation. Thus, the apparent sequence is not used in strict order, each activity can be run for a long time while subsequent ones are undertaken, and if our initial contribution attempts fail, it is not excluded to re-enter into a full cycle.

Fig. 1. Research Methodology Phases.

Problem and Research Objectives Definition: as mentioned before, the main goal is to create a novel RS based on domain ontology, contextual information, and users' activities logs using an intentional-oriented model. The main research question can be defined as:

- **RQ**: Can useful users' recommendations be provided using a combination of mined intentional process models, domain ontologies, and context information?

To solve this problem, we have defined the following research questions that need to be addressed in order to build the intentional process model and the RS:

- **RQ1**: How to combine multiple sources of information (domain ontology, context information, etc.) with events logs in a form that can be used as input for IM and recommendation algorithms?
- **RQ2**: How to define an IM algorithm that uses these different sources of logs together with the domain ontology to build an intentional process model?
- **RQ3**: How to create or find a recommendation algorithm that uses these different sources of logs and the intentional process model as input?
- **RQ4**: To what degree the use of domain ontologies improves recommendations?
- **RQ5**: To what degree the use of contextual information improves recommendations?
- **RQ6**: To what degree the use of an intentional process model (like MAP) improves recommendations?

For RQ4 to RQ6, the improvement of the provided recommendation will be compared to a recommendation obtained without the suggested technique, regarding pertinence, completeness, and accuracy.

Literature Study: After identifying the research problem and the objectives of this work, the next step was to build a solid theoretical foundation for the topic by focusing on the existing research work and identify their open issues. In the beginning, the focus was put on the importance of PM, next, on the existing background of IM and the importance of this research field that gives flexibility to the process models. Now, the focus is on the literature related to the ontology-based recommendation system in order to improve the design of this PhD proposal.

Proposal Elaboration and Implementation: The proposal is an approach to provide useful recommendations to the users while enacting the IS, following the defined research goals. We describe the proposal in the next section.

Evaluation and Validation: To validate the proposed method, it is first necessary to check how likely the use of domain ontologies, contextual information, and intentional process models improve the quality of recommendations. The proposed work will be applied to different case studies and real scenarios and results compared to results obtained by other recommendation systems. The proposed method will be validated through different types of evaluations to check the accuracy, consistency, reliability, and confidence of the obtained results. A large set of logs will be used to check the performance of the proposed algorithms; the event logs of CNAM museum will be used in the first application. The domain ontology used in Phase 1 will be validated by experts in the domain field. Data mining techniques (such as Association Rules [5]) will be used to validate the established link between the intention and the domain knowledge. As for the recommendation part, a questionnaire offered to the user will check the consistency with the results. Finally, multiple analyses will be conducted to identify the deviation between the intentional process model and the actual selected user activity.

3 Proposal

The main objective of this PhD research project is to guide the users by providing them with useful recommendations that are based on the discovered intentional process model according to its current traces. We believe intentional models can provide recommendations to users adequately, as they are close to human ways of working and thinking; precisely if these models were built while taking into consideration the contextual environment. In fact, the same user in different contexts will select different strategies or intentions. For instance, the intention of an employee in the morning is to reach his work office on time. His goal can be achieved through different strategies depending on many factors: If it's a sunny day and his work is not too far from his home, he might go walking. But if it is a rainy day, he will use his car. Otherwise, if he is not feeling well and is not able to drive, he will use public transportation. Thus, we propose the following approach that comprises two main phases as Fig.2 shows.

- In Phase 1, logs are collected with their contextual information from different sources (log traces, sensor logs, etc.). Then, an IM algorithm (IM Algorithm as depicted in Fig.2) constructs an intentional model. This algorithm will discover the intentions and strategies of the intentional model from the log traces, with the help of the domain ontology. The model will use MAP formalism.
- In Phase 2, the recommendation engine receives as an input the intentional process model built in Phase 1 besides the current user activities 'logs with its contextual information. Then, while processing the input parameters, the recommendation engine will provide a suitable recommendation for the user according to its current activity.

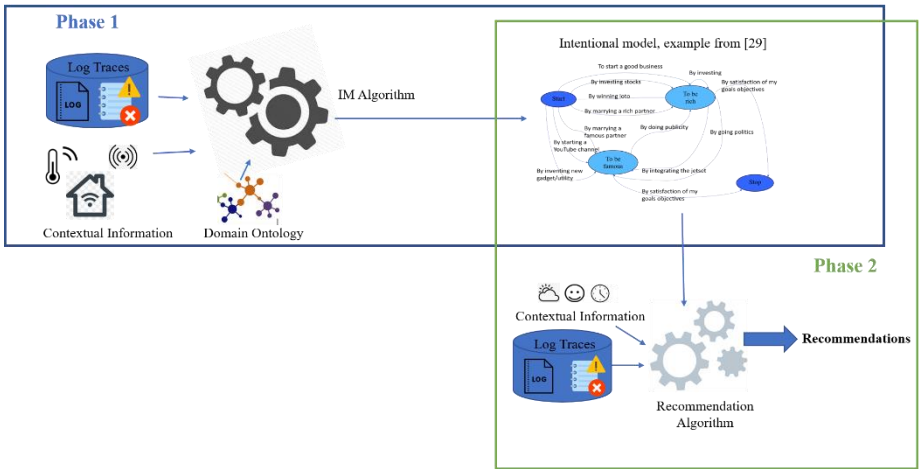

Fig. 2. Framework Proposal.

The proposed framework will be implemented, validated, and tested using real case scenario data. The first application will be implemented using CNAM museum event logs.

To illustrate our proposed framework, let's imagine a log recording all the daily activities of a set of users. This log can be quite enriched with contextual information (how the user is feeling through the day, weather information, time, etc.). This dataset can be combined with a specific ontology, for instance, SPHERE ADL [27] or RADL [28], both specified for daily activities, enriched with other contextual information specific to our dataset. Phase 1 of the proposal will construct the intentional model based on the dataset and the domain ontology. Once the intentional process model is built, when the user is doing an activity (let's say "leave the house"), he will be offered a new recommendation. The recommendation algorithm of Phase 2 based on the model (the user, after leaving his house, can "go to the office" or "go shopping"), the logs analysis (70 percent of the time this specific activity will be followed by the activity "go to the office") and the context (the sun is shining, we are a working day, it is early) will offer him a recommendation (for instance: "walk to the office"). This small example illustrates how the proposed approach can be used to propose pertinent recommendations to users at run time.

4 Related Works

Subsequently, we will present the related background for our research works. There is a difference between the prescribed business process model, which is provided by the organization, and what the users really do to complete their tasks (actual process). For that reason, event logs are considered the most important source of information and the major input for the mining techniques [6]. IM's main objective is to discover the user's intentions while using an information system. In [6], they have identified the distinct elements of intention mining methodology as follows: Event, Event Log, Process

Instance, Process Traces, User Strategies, User Intentions.[6] presented five types of process models from the information systems engineering domain to use them to represent humanities research processes: Activity-oriented process models, Product-oriented process models, Decision-oriented process models, Context-oriented process models, Strategy-oriented process models. [7] think that such types of process models can be a strong basis to achieve thorough research. [8] propose a trace management system adapted to maps that were designed to support recommendation-based guidance. They show how data mining algorithms can be used to find profile clusters in a collection of map enactment traces and to provide recommendations to the users, that can be gathered from the collections of profiles assembled from map enactment traces using methods from the data mining field.

[9] highlights the differences between PM and intention mining. PM purposes to enhance, improve, discover, or check the conformance of activity-oriented process models from event logs. Whereas, while IM has the same objectives as PM, it precisely tackles intentional process models. They worked on a case study where they used the same dataset the techniques of PM and IM. The case study showed that, on the same dataset, both PM and IM techniques are able to discover a process model. However, they mentioned that IM will not be hindered by the same problems branded in PM since intentional process models are flexible.

Unfortunately, the PM approaches do not take into account the hidden aspect of the intentions behind the recorded user activities. By using IM techniques, the intentional process models underlying user activities can be discovered and they can offer better guidance through the processes [10,11].

In [10,11], the authors use HMM (Hidden Markov Models), a probabilistic model that evaluates the most probable intentions behind traces of activities, and compares them to the prescribed intentional process model. [10,11] showed that HMM is an effective model to retrieve intentions from traces of activities. They were able to find the intentions behind the activities with sustaining accuracy, efficiency, and performance. Furthermore, they were able to obtain using HMM, the possibilities of transition from one intention to another and the probabilities for the apparition of activities in each intention, which is the first step to discover the intentional process model. In [13] also, Hidden Markov Model (HMM) was presented as a model that can discover the intentional aspect of activities' sequences, while considering their variability and their probabilistic nature and anticipate the likely set of intentions and/or activities. [12] showcased that HMMs are flexible since they can model the complex structure of temporal dependencies between states and this flexibility can allow adapting the process models to the context in a dynamic way.

[14] converts the conversation of Reddit into logs of verbal behavior to be used in PM technique. They define a taxonomy of 18 classes of speech intentions. The intentions classes are identified by some explicit criteria. They validated the taxonomy with experts and non-experts. And then they have applied PM to discover behavioral models. They used different types of classifiers in order to recognize the speech intention of each class: Logistic Regression, Linear SVM, and random forests. [15] is an exploratory study on conversation analysis. They worked on a digital conversation about a disease on Reddit and they have manually classified the conversation according

to the speech act by using multi-label, a predefined contextual classes annotation. Then they have applied PM using Disco on these data to be able to analyze and detect the sequences of utterances.

In recent years, many recommendations approaches based on ontology have been proposed. Thus, this field appears to be promising if it was combined with the IM research field. In [21] they proposed an ontology-based personalized recommender framework to provide a more precise and personalized recommendation. [22] combined the ontology-based recommender system with machine learning techniques to offer a personalized recommendation for students. [23] developed a collaborative filtering ontology-based recommendation algorithm to achieve a better recommendation in e-Learning systems. While in [24], they have classified research papers using ontology then they have used collaborative algorithms to recommend interesting papers to the users.

Although, there are many approaches that have used IM techniques, ontologies, and contextual information to make recommendations, to the best of our knowledge none of them have combined all these together in order to build intentional process models and provide recommendations.

5 Conclusion

This research paper presents an approach to provide IS users with suitable and relevant recommendations while enacting the system, in an intentional-oriented way. The proposed intention mining approach using ontology for contextual recommendations consists of 2 phases: In the first phase, we build the intentional model by combining the contextual information of the activity's traces using an ontological approach. In the second phase, we guide the user by recommending him the next activity "on the fly" based on the constructed model.

After defining the problem and the PhD research objectives, we are now studying the literature background on the ontology-based recommender systems where we have previously focused on PM, IM, and users activities logs. We plan in near future to start working on the algorithm to create an intentional process model using the different input parameters (phase one of the approach).

Acknowledgement

This PhD thesis is supervised by Prof. Camille Salinesi, Dr. Rebecca Deneckère, and Dr. Elena Kornyshova.

References

1. Van der Aalst, W. M. P.: Process Mining: Data Science in Action. Springer, Heidelberg (2016). ISBN: 978-3-662-49850-7.

2. Cheng, J., Lo, C., Leskovec, J.: Predicting Intent Using Activity Logs: How Goal Specificity and Temporal Range Affect User Behavior. Proceedings of the 26th International Conference on World Wide Web Companion (2017).
3. Rolland, C., Prakash, N., Benjamen, A.: A Multi-Model View of Process Modelling. Requirements Engineering, Springer, Verlag (1999), pp.169 - 187.
4. Mansingh, G., Osei-Bryson, K. M., Reichgelt, H.: Using ontologies to facilitate post-processing of association rules by domain experts. Information Sciences, Volume 181, Issue 3, (2011), Pages 419-434.
5. Zaki, M., Wagner, M.: Data Mining and Analysis: Fundamental Concepts and Algorithms. Cambridge: University Press (2014).
6. Diaz-Rodriguez, O. E., Perez, M., Lascano, J.: Literature Review about Intention Mining in Information Systems. Journal of Computer Information Systems (2019), 1-10.
7. Hug, C., Salinesi, C., Deneckère, R., Lamassé, S.: Process modeling for Humanities: tracing and analyzing scientific processes. Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA 2011), 2011, Beijing, China. pp.245-255.
8. Hug, C., Deneckère, R., Salinesi. C.: Map-TBS: Map process enactment traces and analysis. International Conference on Research Challenges in Information Science (RCIS), May 2012, Valencia, Spain. pp.204-209.
9. Khodabandelou, G., Hug, C., Deneckère, R., Salinesi, C.: Process Mining Versus Intention Mining. Lecture Notes in Business Information Processing. (2013). 147. 466-480.
10. Khodabandelou, G., Hug, C., Deneckere, R., Salinesi, C.: Supervised Intentional Process Models Discovery using Hidden Markov Models. Seventh International Conference on Research Challenges in Information Science, May 2013, Paris, France. pp.1-11.
11. Khodabandelou, G., Hug, C., Deneckere, R., Salinesi, C.: Unsupervised discovery of intentional process models from event logs. 11th Working Conference on Mining Software Repositories, May 2014, Hyderabad, India. pp.282-291.
12. Khodabandelou, G., Hug, C., Deneckere, R., Salinesi, C.: Supervised vs. Unsupervised Learning for Intentional Process Model Discovery. Business Process Modeling, Development, and Support (BPMDS), Jun 2014, Thessalonique, Greece. pp.1-15.
13. Khodabandelou, G. Contextual Recommendations using Intention Mining on Process Traces. Doctoral Consortium, Seventh International Conference on Research Challenges in Information Science, May 2013, Paris, France.
14. Epure, E.V., Compagno, D., Salinesi, C., Deneckere, R., Bajec, M., Žitnik, S.: Process models of interrelated speech intentions from online health-related conversations, Artificial Intelligence in Medicine, Volume 91,2018, Pages 23-38.
15. Compagno, D., Epure, E., Deneckere, R., Salinesi, C.: Exploring Digital Conversation Corpora with Process Mining Corpus Pragmatics, Springer, 2018, 2, pp.1-23.
16. Sielis, G., Tzanavari, A., Papadopoulos, G.: A Review of Recommender Systems: Types, Techniques and Applications, (2014).
17. K. Sorde, R., Deshmukh, S.: Comparative Study on Approaches of Recommendation System. International Journal of Computer Applications. (2015), 118. 10-14.
18. Karypis, G.: "Evaluation of item-based top-n recommendation algorithms." 10th international conference on Information and knowledge management. ACM, 2001.

19. Osadchiy, T., Poliakov, I., Olivier, P., Rowland, M., Foster, E.: Recommender system based on pairwise association rules - Expert Systems with Applications, Volume 115, January 2019, Pages 535-542.
20. “Netflix Prize” Netflixprize.Com, 2009, www.netflixprize.com/.
21. Ge, J., Chen, Z., Peng, J., Li, T.: An ontology-based method for personalized recommendation. 2012 IEEE 11th International Conference on Cognitive Informatics and Cognitive Computing, Kyoto, Japan, 2012, pp. 522-526.
22. Obeid, C., Lahoud, I., Khoury, H., Champin, P.A.: Ontology-based recommender system in higher education. 2018 Web Conference Companion (WWW 2018), Apr 2018, France.
23. Zhang, Z., Gong, L., Xie, J.: Ontology-Based Collaborative Filtering Recommendation Algorithm. (2013). 7888. 172-181.
24. Middleton, S., De Roure, D., Shadbolt, N.: Ontology-Based Recommender Systems. (2009).
25. Schafer, J. B., Frankowski, D., Herlocker, J., Sen, S.: Collaborative Filtering Recommender Systems. (2007).
26. Hevner, A. R.: A Three Cycle View of Design Science Research. Scandinavian Journal of Information Systems: Vol. 19: Iss. 2, Article 4. (2007)
27. Woznowski, P. R., Tonkin, E. L., Flach, P. A.: Activities of Daily Living Ontology for Ubiquitous Systems: Development and Evaluation. Sensors. Basel, Switzerland (2018). https://doi.org/10.3390/s18072361.
28. Bae, I. H.: An ontology-based approach to ADL recognition in smart homes, Future Generation Computer Systems (2013).
29. Deneckere, R., Ralyté, J.: Expressing strategic variability of processes: The MAP process modelling approach, RCIS Tutorial, RCIS 2020.
30. Viinikkala, M.: Ontology in Information Systems. (2004).