A Probabilistic Model for Intrusive Recommendation Assessment

Imen Akermi
IRIT, University of Toulouse III
Toulouse, France
imen.akermi@irit.fr

Mohand Boughanem
IRIT, University of Toulouse III
Toulouse, France
mohand.boughanem@irit.fr

Rim Faiz
LARODEC, University of Carthage
Carthage, Tunisia
rim.faiz@ihec.rnu.tn

ABSTRACT

The overwhelming advances in mobile technologies allow recommender systems to be highly contextualized and able to deliver recommendation without an explicit request. However, it is no longer enough for a recommender system to determine what to recommend according to the users' needs, but it also has to deal with the risk of disturbing the user during recommendation. We believe that mobile technologies along with contextual information may help alleviate this issue. In this paper, we address intrusiveness as a probabilistic approach that makes use of the several embedded applications within the user's device and the user's contextual information in order to figure out intrusive recommendations that are subject to rejection. The experiments that we conducted have shown that the proposed approach yields promising results.

CCS CONCEPTS

• **Information systems → Mobile information processing systems**; **Recommender systems**;

KEYWORDS

Intrusiveness assessment; proactive recommender systems; context-awareness; risk-awareness; mobile applications

ACM Reference Format:
Imen Akermi, Mohand Boughanem, and Rim Faiz. 2018. A Probabilistic Model for Intrusive Recommendation Assessment. In *Twelfth ACM Conference on Recommender Systems (RecSys '18), October 2–7, 2018, Vancouver, BC, Canada.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3240323.3240403

1 INTRODUCTION

In certain situations, the user may choose to reject a recommendation regardless of its content. This abstinence may not concern the recommended information itself but it takes part in the situation where the user may be in and during which the user does not want to be interrupted or disturbed. Therefore, it is crucial to assess the intrusiveness level of the user's situation before sending any recommendation. In this paper, we address intrusiveness as a probabilistic decision making process within a proactive recommendation approach that takes into account not only the user's agenda activities but also the user's contextual information with its several level of representation along with other applications and technologies embedded within the user's mobile device.
Indeed, the several sensors and information provided by mobile devices can help keep track of the situations during which recommendations are subject to rejection.

This paper's main contributions can be summarized as the following:

- A probabilistic intrusiveness assessment approach that integrates contextual information along with the applications and sensors embedded within the user's mobile device.
- An extensive evaluation framework for intrusive recommendation assessment

The paper is organized as follows. We introduce in section 2, a literature review about intrusiveness/risk assessment within proactive recommender systems. We detail in section 3 the proposed approach. Section 4 describes the experiments that we conducted using a user study and we conclude in section 5 with thoughts for future work.

2 RELATED WORK

The work presented in [5] was the first to integrate intrusiveness as a phase in which they assess the risk of disturbing the user before recommending. They define a "*critical*" or "*risky*" situation as a situation in which a user does not want to be disturbed. For each situation, they compute a risk score that depends on the risk-level of the concept describing the user's activity extracted from his/her agenda. They believe that a situation is deemed risky if its risk score exceeds a pre-defined threshold. They assume that the user indicates, for each activity he/she may undertake, the time and the location information, which is not always the case in real life. Bedi et al. [4] proposed a situation assessment approach for restaurants recommendation, that uses fuzzy logic as an inference technique that depends on *distance*, *time*, *budget* and *reachability*, to assess the context level (i.e. intrusiveness level) of a given situation. For example, they predefine the fuzzy sets for the context level and the distance attribute as: `Distance={Near,Moderate,Far}`, `Context-level={Low,Medium,High}`. The context attributes are used in a rule-based approach to infer the context-level, example:

```
IF(Distance IS 'Near') AND (Time IS 'In-Time') AND
(Budget IS 'Affordable') AND (Reachability IS 'High')
THEN Context-level IS 'High'.
```

In [9], the authors tackled non-intrusiveness as a privacy permission issue. They actually ask users explicitly before recommending to choose among three options : "*agree*", "*reject*" or "*agree only this*

RecSys '18, October 2–7, 2018, Vancouver, BC, Canada

ACM ISBN 978-1-4503-5901-6/18/10…$15.00
https://doi.org/10.1145/3240323.3240403

time" to get a recommendation. The work presented in [7] explored the user's receptivity to notifications/interruptions and suggested that the opportune moments for a user to get interrupted, regardless of his/her situation or the notification's content, is only at activity transitions, which means that they only consider a notification as non-intrusive when the user is practically transitioning from the use of a given application to another. The work proposed in [6] considered intrusiveness in a recommendation approach as a classification problem which aims at identifying whether a given context is "good" or "bad" to trigger the recommendation process. They collected mobile data over a three weeks user study in order to learn the classification model.

These works and several others [3, 8, 10, 12–15] tend to deal with the intrusiveness issue from an implicit user profiling angle and they depend on the user's explicit feedback to figure out if a recommendation is appropriate or not, forgetting that the large amount of applications embedded in the user's phone could be the issue itself. Therefore, we propose in this paper, to assess intrusiveness from a situation assessment perspective, not only in terms of context as generally defined by time and location, but also considering the actual user' activity inferred from the applications that a user is using at a given situation and the sensors installed within the user's mobile device.

3 ASSESSING A SITUATION'S INTRUSIVENESS LEVEL

3.1 Situation modelling

The probabilistic intrusiveness assessment method that we propose is integrated within a proactive recommendation approach that covers various domains like POI (Points of Interests), News, Movies, etc. It consists in recommending information that match a user's situation and preferences without waiting for the user to initiate any interaction with his/her device [1]. The recommendation of an item is only sent after assessing the intrusiveness level of the user's current situation.

We model the user's daily routine as a set of situations described by the user's actual activity and the spatio-temporal contextual factors. In a more formal way, a situation is represented as $S = (F_t, F_w, F_l, F_a)$, where: F_t is *time of the day* , F_w is *day of the week* , F_l is *the user's actual location* and F_a is *the user's activity*. Therefore, the recommendation process entails a context model that figures out what and when to recommend the relevant information (news, movies, a place to visit, a restaurant, ...) to the user according to his/her situation. Indeed, we believe that a situation, with its different level of representation, defines the changing user's information need. In this paper, we do not focus on the type of information that may be recommended to the user and that was addressed in [1], but we rather expose a probabilistic approach for balancing the recommendation process against intrusive interruptions. Indeed, as we are working within mobile devices framework, the risk of disturbing the user becomes increasingly high. Indeed, according to the 2017 Mobile Usage Report[1], people spend more than 50% of their digital time on smartphone applications. The desktop represents only 34% of the digital time and it was also observed that entertainment and communication applications (social networks, instant messaging) account for almost two-thirds of the time spent. We believe that the amount of information provided by mobile devices not only provide valuable information about the user's interests but also can help figure out the user's situation in general and the user's activity in particular, which can alleviate the problem of assessing whether a situation is conducive to recommendation's rejection or not. In order to assess intrusiveness in this approach, we only consider F_t, F_w and F_a as the features describing a situation $S = (F_t, F_w, F_a)$. The time feature takes into account two levels: *time of the day* and *the week day*. We choose to divide a daily routine into four periods (*morning*, *midday*, *afternoon* and *evening*) that are framed within 24 hours intervals.

$F_t \in \{morning[07:00, 12:00], midday[12:00, 14:00], afternoon[14:00, 18:00], evening[18:00, 00:00]\}$.

Assuming that the user lives in a Western country, the weekdays can be partitioned as following:

$F_w \in \{work\_days\{monday, \ldots, friday\}, rest\_days\{saturday, sunday, public\_holiday\}\}$.

This partition is automatically changed according to the user's location. Indeed, while Saturdays and Sundays may be rest days in most Western countries, this is not the case for Middle-Eastern countries, where Friday is typically a rest day and Sunday is not.

At a given situation S_i, the system takes a snapshot of the user's current activity F_a such as driving, texting messages, tweeting or browsing, using the sensors and the applications embedded in the user's mobile device. For example, we can figure out if the user is in a meeting according to his agenda or if the user is driving or jogging by checking the accelerometer sensor.

3.2 Intrusiveness assessment

We assess intrusiveness as a probabilistic function measuring the natural logarithm of the conditional probability of recommendation acceptance knowing the user's actual situation upon the conditional probability of the recommendation rejection:

$$\theta = \ln \frac{P(R|S = S_i)}{P(\neg R|S = S_i)} \tag{1}$$

Where R stands for the act of sending a recommendation and S_i refers to the i^{th} situation, for a given user.

According to the θ's score, the system decides if the user's situation is appropriate to recommendation, i.e. non-intrusive, or not. As we explained earlier, a situation is modelled as $S = (F_t, F_w, F_a)$ referring respectively to *time of the day, day of the week* and *the user's current activity*. Thus, the intrusiveness score can be estimated as :

$$\theta = \ln \frac{P(R|F_t^i, F_w^i, F_a^i)}{P(\neg R|F_t^i, F_w^i, F_a^i)} \tag{2}$$

Where $F_k^i, k \in \{t, w, a\}$, is the k^{th} feature of situation S_i.

After applying the Bayes rule, the conditional probability of sending a recommendation R knowing the features of situation S_i (respectively $\neg R$) is written as:

$$P(R|F_t^i, F_w^i, F_a^i) = \frac{P(F_t^i, F_w^i, F_a^i|R) \times P(R)}{P(F_t^i, F_w^i, F_a^i)} \tag{3}$$

[1] https://www.comscore.com/Insights/Presentations-and-Whitepapers/2017/The-2017-US-Mobile-App-Report

Assuming that the situation's features are independent and that $P(F_t^i, F_w^i, F_a^i)$ is uniform, then $P(R|F_t^i, F_w^i, F_a^i)$ (respectively $P(\neg R|F_t^i, F_w^i, F_a^i)$) for a given user can be expressed as:

$$P(R|F_t^i, F_w^i, F_a^i) \propto \prod_{k \in \{t,w,a\}} P(F_k^i|R) \times P(R) \tag{4}$$

Where $P(F_k^i|R)$ refers to the probability of having feature F_k^i (*time of the day*, *day of the week* or *activity*) within situations where recommendations were accepted R (respectively rejected for $P(F_k^i|\neg R)$).

$$P(F_k^i|R) = \frac{|S_{user}^{R,F_k^i}|}{|S_{user}^R|} \tag{5}$$

Where :
$|S_{user}^{R,F_k^i}|$ is the number of situations that occurred to the given user and during which feature $f_k = f_k^i$ of situation S_i and recommendation was accepted, i.e. R(respectively rejected for $\neg R$).
$|S_{user}^R|$ is the number of situations during which the user accepted a recommendation (respectively rejected for $\neg R$).
In order to avoid zero equal probabilities, we consider Dirichlet smoothing [11] that combines the use of a parameter μ along with judgements of other users regarding a given situation. Therefore, $P(F_k^i|R)$ (respectively $P(F_k^i|\neg R)$) can be estimated as:

$$P(F_k^i|R) = \frac{|S_{user}^{R,F_k^i}| + \mu \times \frac{|S_{all}^{R,F_k^i}|}{|S_{all}^R|}}{|S_{user}^R| + \mu} \tag{6}$$

Where :
$\mu \in [1, 1000]$ is a parameter that is set according to experiments.
$|S_{all}^{R,F_k^i}|$ is the number of situations that occurred to other users and during which feature $f_k = f_k^i$ of situation S_i and recommendation was accepted (i.e. R).
$|S_{all}^R|$ is the number of situations that occurred to other users and during which the recommendation was accepted (respectively rejected for $\neg R$).

Deciding whether a recommendation will be perceived by a given user as intrusive or not is determined by the θ score (see Eq 1). If $\theta > 0$, the recommendation is deemed non-intrusive.

4 EXPERIMENTS

Since work on intrusiveness assessment evaluation within recommender systems (RS) is still at its early stages, we could not find a suitable dataset to evaluate the approach that we propose. Therefore, we opted for constructing a user study. Indeed, this kind of evaluation allows a subjective assessment of the system as surveys can be conducted along with the experiments.

4.1 User study

The user study that we conducted integrates automatically generated situations simulating real life ones. These situations are described by four features: *time of the day*, *the week day*, *the user's current activity* and *the type of information* that can be recommended (news, POI, restaurant, TV program, etc.). The possible values that can be taken by the first three features were gathered using a survey that we carried within our lab (colleagues from different backgrounds and age range). We made sure that these values could be extracted through the applications and sensors embedded within the user's mobile device. We addressed the type of information that can be recommended according to a user's given situation in a previous work [1]. The set of the generated situations is then filtered in order to take out those that are not likely to happen, like *having a meeting at late night at home*. We choose to work with 100 situations to avoid participants getting bored and then bias the study. More than 1400 participants took part in this study. Given a situation they might be in, participants were asked if they accept to get a recommendation or not regardless of the content of the information being recommended. They were also asked to express their opinion with reference to the information they were recommended according to their situation. Here is an example of a situation that can be suggested to a participant:

```
It is Saturday,Afternoon and you are doing the following
activity : Walking
Would you accept to get a notification :
    • YES
    • NO
Given this situation, do you think that recommending
this type of information is interesting : Points
of interests(Park, Museum, ...), Concert Theaters
Program,Movie Theaters Program:
    • YES
    • NO
Comments: (please comment your answers)
```

We used the Crowdflower[2] platform to run the user study. In order to avoid any bias, we configured several quality control mechanisms such as speed traps measuring the time spent by a participant to answer the questions of the study. We also made sure that the participants understand perfectly English and the questions they were asked. In this paper, we only considered the first section of this study that addresses the issue of accepting to receive a recommendation or not regardless of its content.
We used the Mean Average Precision (MAP) metric to evaluate the intrusiveness assessment accuracy across all the participants and against two baselines.

4.2 Results

We run a K-cross-validation evaluation (K=10) in order to put forward the accuracy of the proposed approach for intrusiveness assessment:

$$MAP = \frac{\sum_{u=1}^{U} AveP(u)}{U} \tag{7}$$

$$AveP(u) \frac{\frac{\sum_{k=1}^{K} rel\_s}{S}}{K} \tag{8}$$

Where U is the number of users, rel_s is the number of correctly assessed situations for each run, K is the number of runs ($K = 10$)and S is the number of situations.

[2]https://www.crowdflower.com/

Then, a mean over all users for every possible feature combination is calculated. As we are using the Dirichlet smoothing in the proposed approach, we varied μ within the range of $[1, 1000]$ while computing the MAP.
The results that we obtained have shown that the proposed approach for intrusiveness assessment performs better when μ equals 1, which is normal since we do not address a large volume of data like it is typically addressed within Information Retrieval tasks.

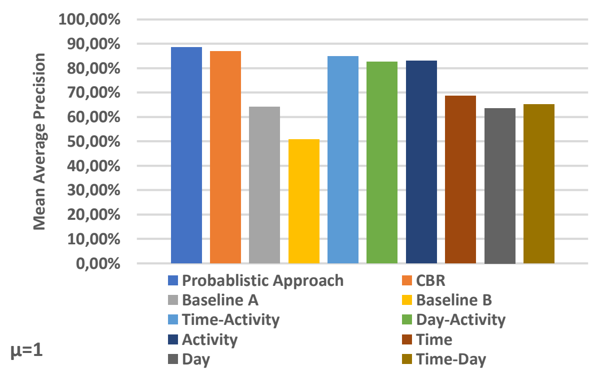

Figure 1: MAP evaluation for the Intrusiveness assessment approach

As illustrated in Figure 1, the proposed approach using the situation's features that we considered scores *88.5%* for the MAP compared to other features' combinations, to two baseline approaches and to a Case Based Reasoning approach that we proposed in a previous work [2]. Baseline A which sends recommendations without considering the user's interruptibility scores *64%* for the MAP.
Baseline B which consists in not sending a recommendation when an application is ON, scores *50.81%*. The CBR approach proposed in [2] consists in using the user's analogous past situations, that are most similar to the actual situation, to figure out if we could interrupt the user's current activity and send a recommendation. The CBR approach scored a MAP of 87% which is slightly less accurate than the proposed probabilistic approach. Indeed, while the CBR approach only uses the user's past situations, the probabilistic approach combines the user's and other users' judgements.
We also note that, for the probabilistic approach, the combinations that entail the activity feature, like *Activity-Time*, *Activity-Day* and *Activity*, scores a high precision that is quite similar to the MAP of the approach that makes use of all the features. Then, we can assume that users mainly tend to reject recommendations based on what they are doing at a given situation and do not always consider the other features like time of the day or the week day. However, based on a study that we conducted on the same dataset, we noted that these features are important for users to decide whether a given recommended type of information is interesting or not, given the situation they might be in [1]. Besides, the performance of Baseline B, which considers that a recommendation should not be sent when an application is ON, proves that approaches that automatically consider the use of any random application at a given situation as a hinder to sending a recommendation, are not effective. It actually depends on the type of the application being used and on the user's behaviour. These findings are also proved by the analysis that we conducted on the users' responses regarding recommendations and according to the undertaken activity. We measured the proportion of users who considered recommendations, while undertaking some activities, as intrusive or not. Figure 2 puts forward some of the activities that we predefined for this study. We notice that more

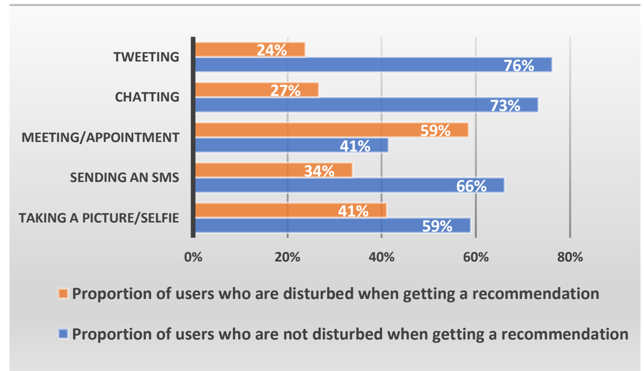

Figure 2: Correlation between intrusiveness and activities

than 70% of the participants did not consider recommendations as intrusive while chatting or tweeting. This can be explained by the fact that people may want to share the recommended information with friends. We also note that 59% of the participants against 41% were not disturbed when getting a recommendation while taking a picture which could be somehow interpreted as not making sense because we normally expect users to get annoyed if they were interrupted while typing a message or using the device's camera. These findings make us believe that the content of the information has to be considered also as a feature. In fact, a user may not want to be disturbed usually when working but perhaps work related news is still acceptable. Indeed, even though the user chose not to be disturbed, the recommended information might be worth being interrupted for, such as breaking news or an accident that happened on the user's way home.
We believe that such trade-off need to be studied. Therefore, we are currently working on integrating into the approach we proposed, a trade-off between the importance of the information to be recommended and the risk of disturbing the user. We also made the conducted user study available[3] for the RS research community as a dataset for intrusiveness assessment. This may help alleviate the datasets shortage and might provide a framework for different approaches to be compared on a same basis.

5 CONCLUSION

We put forward, in this paper, a probabilistic approach for intrusiveness assessment for proactive recommendation that exploits the applications and sensors installed within the user's device along with other contextual information. The user study that we conducted for the experiments has proven that the proposed approach yielded promising results regarding intrusiveness assessment and recommendation accuracy.

[3] contact authors

REFERENCES

[1] Imen Akermi, Mohand Boughanem, and Rim Faiz. 2016. Just-In-Time Recommendation Approach within a Mobile Context. In *2016 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2016, Omaha, NE, USA, October 13-16, 2016*. 636–639.

[2] Imen Akermi and Rim Faiz. 2018. Situation assessment for non-intrusive recommendation. In *12th International Conference on Research Challenges in Information Science, RCIS 2018, Nantes, France, May 29-31, 2018*. 1–11.

[3] Bahram Amini, Roliana Ibrahim, Mohd Shahizan Othman, and Ali Selamat. 2014. Capturing scholar's knowledge from heterogeneous resources for profiling in recommender systems. *Expert Systems with Applications* 41, 17 (2014), 7945–7957.

[4] Punam Bedi and Sumit Kr Agarwal Agarwal. 2012. A Situation-Aware Proactive Recommender System. In *Proceedings of the 12th International Conference on Hybrid Intelligent Systems (HIS)*. 85–89.

[5] Djallel Bouneffouf, Amel Bouzeghoub, and Alda Lopes Gançarski. 2012. A contextual-bandit algorithm for mobile context-aware recommender system. In *Proceedings of the International Conference on Neural Information Processing*. Springer, 324–331.

[6] Nofar Dali Betzalel, Bracha Shapira, and Lior Rokach. 2015. Please, not now!: A model for timing recommendations. In *Proceedings of the 9th ACM Conference on Recommender Systems*. ACM, 297–300.

[7] Joyce Ho and Stephen S Intille. 2005. Using context-aware computing to reduce the perceived burden of interruptions from mobile devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 909–918.

[8] Karen Kay-Lynn Liu. 2004. *A personal, mobile system for understanding stress and interruptions*. Ph.D. Dissertation. Massachusetts Institute of Technology.

[9] Rui Liu, Jiannong Cao, Kehuan Zhang, Wenyu Gao, Junbin Liang, and Lei Yang. 2016. When privacy meets usability: unobtrusive privacy permission recommendation system for mobile apps based on crowdsourcing. *IEEE Transactions on Services Computing* (2016).

[10] Fengji Luo, Gianluca Ranzi, Weicong Kong, Zhao Yang Dong, Shu Wang, and Junhua Zhao. 2017. Non-intrusive energy saving appliance recommender system for smart grid residential users. *IET Generation, Transmission & Distribution* 11, 7 (2017), 1786–1793.

[11] David JC MacKay and Linda C Bauman Peto. 1995. A hierarchical Dirichlet language model. *Natural language engineering* 1, 3 (1995), 289–308.

[12] Adem Sabic. 2016. Proactive Recommendation Delivery. In *Proceedings of the 10th ACM Conference on Recommender Systems*. ACM, 459–462.

[13] James Schaffer, Tobias Höllerer, and John O'Donovan. 2015. Hypothetical Recommendation: A Study of Interactive Profile Manipulation Behavior for Recommender Systems.. In *FLAIRS Conference*. 507–512.

[14] Daniel P Siewiorek, Asim Smailagic, Junichi Furukawa, Andreas Krause, Neema Moraveji, Kathryn Reiger, Jeremy Shaffer, and Fei Lung Wong. 2003. SenSay: A Context-Aware Mobile Phone. In *Proceedings of the Seventh IEEE International Symposium on Wearable Computers*. 248–249.

[15] Reto A Stucki, Prabitha Urwyler, Luca Rampa, René Müri, Urs P Mosimann, and Tobias Nef. 2014. A web-based non-intrusive ambient system to measure and classify activities of daily living. *Journal of medical Internet research* 16, 7 (2014).