Context-aware and Ontology-based Recommender System for E-tourism

Gustavo Castellanos, Yudith Cardinale, Philippe Roose

Context-aware and Ontology-based Recommender System for E-tourism

Gustavo Castellanos[1], Yudith Cardinale[2,1] and Philippe Roose[3]
[1]*Universidad Simón Bolívar, Caracas, Venezuela*
[2]*Universidad Católica San Pablo, Arequipa, Perú*
[3]*Université de Pau et des Pays de l'Adour, LIUPPA – T2I64600, Anglet, France*
14-10192@usb.ve, ycardinale@usb.ve, philippe.roose@iutbayonne.univ-pau.fr

Keywords: Tourism, Recommender Systems, Context Awareness, Ontology, Serendipity, User-centric

Abstract: Frequently, travelers try to collect information for planing a trip or when being at the destination. Usually, tourists depend on places' reviews to make the choice, but this implies prior knowledge of the touristic places and explicit search for suggestions through interaction with applications (i.e., PULL paradigm). In contrast, a PUSH approach, in which the application proactively triggers a recommendation process according to users' preferences and when necessary, – seems to be a more reasonable solution. Recommender systems have become appropriate applications to help tourists in their trip planning. However, they still have limitations, such as poor consideration of users' profiles and their contexts, their predictable suggestions, and the lack of a standard representation of the knowledge managed. We propose a user-centric recommender system architecture, that supports both PULL and PUSH approaches, assisted by an ontology-based spreading activation algorithm for context-aware recommendations, with a focus on decreasing predictable outputs and increasing serendipity, based on an *aging-like* approach. To demonstrate its suitability and performance, we develop a first prototype of the architecture and simulate different scenarios, varying users' profiles, preferences, and context parameters. Results show that the ontology-based spreading activation and the proposed aging system provide relevant and varied recommendations according to users' preferences, while considering their context and improving the serendipity of the system when comparing with a state-of-the-art work.

1 Introduction

Tourism is one of the most promising areas for supporting the economy of countries and is becoming an extremely important market (Alghamdi et al., 2016; Artemenko et al., 2017; Kazandzhieva and Santana, 2019; Kayumovich, 2020). When travellers plan to go to a destination, they try to collect information about their new destination "as best as they can" (e.g., by going to the tourist office or by obtaining information about their destination and its surroundings on Internet). This task may be overwhelming due to the big amount of available information sources. Usually, tourists depend on places' reviews to decide where to spend their vacations, their free time, or even to take lunch, but this requires prior knowledge of the existence of the places.

Currently, there is a plethora of applications offering information of historical centres, points of interest (POI), city tours, green areas to rest, etc. The use of such applications implies that the user is firstly aware of them; then installs them in the hope that they are suitable (an average of 3.3% of mobile apps are still being used after 30 days of being installed[1], and from an average of 40 installed apps, the 89% of the time of use is dedicated to 18 of such apps[2]).

Most of these applications follow the PULL paradigm, i.e., the user has to explicitly search for suggestions/recommendations through interaction with the application's graphical user interface. We believe that this is one of the reasons why these types of applications are not successful: they require too many steps, too many interactions, and too many unknowns. On the other hand, a PUSH approach, in which the application proactively triggers a recommendation process according to users' preferences and when necessary (e.g., when the user is close to a POI, when lunch time is approaching, or when weather conditions change) seems to be a more reasonable solution. A first step in this direction has been taken by Google

[1]https://www.apptentive.com/blog/2017/06/22/how-many-mobile-apps-are-actually-used/

[2]https://www.simform.com/the-state-of-mobile-app-usage/

via Google Now[3], which provides very basic information/suggestions to users by detecting in their geographical environment what they need according to their location and time. However, such applications are simplistic and use only a limited amount of information, and moreover, do not take into account the user's activity. A key element is the consideration of the users' context, such as their locations, weather (*is it raining?*), or time (*is it night?*).

Recommender systems are now at the heart of much research and offer real benefits to users, organizations, and the business community (Lim et al., 2019; Leskovec et al., 2020; Hamid et al., 2021). Most tourism recommender systems are able to predict user's preferences based on previous activities, but research on Context-Aware Recommender Systems are still to be further developed (Laß et al., 2017; Raza and Ding, 2019). Imagine a tourist in Paris, fan of museums and parks, who wishes to spend the afternoon in a good place, but it is a rainy day. The experience should not be affected by the bad weather. She/He would like to use an app that recommends a nearby good place for spending this rainy afternoon, preferably a museum instead of a park, hence spending no time searching for possible places to visit. Moreover, if in the next day it is still raining (it rains a lot in Paris!), the tourist would like to visit a different museum or other indoor place.

Additionally, most recommender systems do not offer a standard representation of the knowledge related to tourism (POI, users' information, context information, etc.) and do not consider to vary recommendations– i.e., they become predictable for same users, under the same situations. Thus, there is still a real need for research and design of innovative solutions in this field. Nowadays, the use of ontologies to represent such as knowledge is becoming a powerful tool to offer such as innovative solutions (Borràs et al., 2014; Yochum et al., 2020).

To overcome these limitations, we propose RECESO, a **REC**ommender system architecture for **E**-tourism with **S**erendipity and **O**ntology-based. RECESO is a user-centric recommender system architecture, that supports both PULL and PUSH approaches. This framework is assisted by a semantic representation of the information managed, such as users' preferences, their contexts, and POI, from which an ontology-based spreading activation algorithm for context-aware recommendations is designed. According to the users' wishes, the serendipity of recommendations may increase (e.g., surprise events, no repeated recommendations) (Kotkov et al., 2016), based on an *aging-like* algorithm, which gives less priority to recently recommended places to the user and increasing priority of popular POI, even if they are not in the user's preferences.

[3]https://www.google.com/intl/fr/landing/now/

We describe the architecture and algorithms of RECESO and demonstrate its suitability and performance, through a first developed prototype of the architecture, in different simulated scenarios, with different users' profiles and preferences and different context parameters. Results show how the spreading activation and the proposed *aging* system give relevant and varied recommendations, according to users' preferences, while taking into account their context.

2 Related Work

Depending on the technique used to make suggestions, recommender systems are classified into two categories: memory-based and model-based (Bobadilla et al., 2013). **Memory-based** recommender systems support their suggestions on similarities among users and their shared items, hence recommending similar items to the ones the user likes (i.e., *content-based recommendation*) or recommending items liked by users that are similar to the user (i.e., *collaborative filtering*). **Model-based** recommender systems try to guess how much a user will like an item that they did not consume before, usually through statistical techniques and machine-learning models. Recommender systems that combine both approaches are called **Hybrid** recommender systems.

We survey some relevant and recent studies in the tourism domain, classifying them according to these three categories and considering aspects related to how they manage users' preferences and interests, the context awareness, the use of ontologies, and the variability of recommendations. We compare them in terms of these criteria and highlight the difference with our proposal.

2.1 User information

All recommender systems take into account some information related to the user, such as identification aspects (sex, age, profession, etc.), interest and preferences, or social relations. This information can be obtained explicitly (e.g., questionnaires) (Laß et al., 2017; Jannach and Zanker, 2020) or implicitly (e.g., by analyzing feedbacks in user's social networks) (Lin et al., 2018).

Some proposals are supported on users' interaction to obtain their data (i.e., following a PULL paradigm). In (Rajaonarivo et al., 2019), authors model the users' information considering their gender, age category (e.g., kid, adult, or elderly), and preferences, classified as thematic preferences (e.g., museum, theater) and historical preferences (e.g., 12th century), that have to be provided by them through a user interface. In (Santos et al., 2019), only health, physical, and psychological conditions are required with forms filled by users. Users' preferences and feedback are directly assigned by the users in the proposals presented in (Laß et al., 2017; Bahramian et al., 2017). As in (Arigi et al., 2018), users' interest (i.e., not interested, less interested, and interested enough) on tourism categories are asked to them. SMARTMUSEUM (Ruotsalo et al., 2013), asks users about the desired duration of a visit to a particular location, the motivation for a visit, and ability to consume the content offered by the system. The system proposed in (Shen et al., 2016) requires photos uploaded by users, from which it extracts their travel history; it also asks users to rank POI (their favorite and non-favorite attractions).

Other recommender systems extract users' information, mainly their preferences and interests, from available sources, without asking for explicit interaction. The work presented in (Kesorn et al., 2017), takes from Facebook basic information (e.g., name, age) and check-in data (e.g., visited places) to identify users' interests and preferences. To automatically deduce users' preferences from their social networks, many techniques are used, such as opinion mining (Logesh et al., 2018; Logesh and Subramaniyaswamy, 2019), analysis of implicit/explicit feedback (Hidasi and Tikk, 2016), and analyzing geotagged pictures (Sun et al., 2019). Curumim (Menk et al., 2017) takes from users' social networks their travel history and level of education, and predicts their degree of curiosity. Most of these works follow the PULL approach.

Many other proposals combine both explicit and implicit users' information gathering, also under the PULL paradigm. SPETA (García-Crespo et al., 2009), supports its recommendation on the interest and rating of tourism places that users explicitly provide, and on preferences deduced by analyzing their behavior on social networks. The system presented in (Alonso et al., 2012), takes into account special needs and context-dependent preferences on tourism sites, directly provided by users, as well as explicit/implicit feedback on their social networks.

2.2 Context-awareness

In order to improve suggestions, the trend is to consider context aspects that describe a specif situation in a determined moment for a user, including transportation media, weather, time, or even health conditions. Many works use context-modeling approaches that mainly consider means of transportation, travel time, location, or weather (Bahramian et al., 2017; Kesorn et al., 2017; Arigi et al., 2018; Logesh et al., 2018; Rajaonarivo et al., 2019; Logesh and Subramaniyaswamy, 2019).

SMARTMUSEUM (Ruotsalo et al., 2013) only considers users' location as contextual information captured by the built-in sensors of mobile devices, such as GPS, accelerometers, and RFID readers. The system proposed in (Shen et al., 2016) collects automatically the users' current location (city, latitude, and longitude) and current time, thus recommendations are influenced by the geo-distance to POI. SPETA (García-Crespo et al., 2009) also considers users' location extracted from users' mobile devices, but it takes into account current and the history of past locations. It also gathers from mobile devices other contextual information, such as weather forecast and time. A similar work is proposed in (Laß et al., 2017), that considers previously visited POI, time of the day, day of the week, weather, temperature, and opening hours to recommend a torist trip.

A General Factorization Framework for context-aware recommender systems is proposed in (Hidasi and Tikk, 2016), with the aim of constructing factorization matrices (not matter which and how many context factors are considered) for machine learning techniques. In (Sun et al., 2019), it is used a combination of contextual information like weather, transportation, or textual information, with geotagged pictures taken by the user for building the user model. The system proposed in (Alonso et al., 2012), is the only one among the referenced works that considers context-dependent preferences, as our proposal, related to weather, day, transport, and special needs. It means that users' preferences for POI, are expressed with regards the context (e.g., *indoor places when is raining*, *art exhibitions at night*).

2.3 Serendipity

A useful goal for a recommender system is to be serendipitous, which means that the recommended items must be relevant, novel, and unexpected. A relevant item is an item that the user likes, consumes, or is interested in; a novel item is one that users have never seen or heard about in their life; an unexpected item is significantly different from the profile of the user. Some quite good recent surveys emphasize the importance of such as feature for recommender systems in general (Kotkov et al., 2016; Chen et al., 2019) and in the tourism domain (Tintarev et al., 2010; Menk et al., 2019).

It starts to be a trend to recommend POI. The model considered in (Rajaonarivo et al., 2019), considers previous activity of other users to recommend tours of POI, thus reducing overspecialization and increasing serendipity, however experiments with this feature were not made. Authors of (Shen et al., 2016) assert that the proposed system is able to recommend fresh and surprise POI, based on collective intelligence. From the level of curiosity predicted from users' social networks, Curumim (Menk et al., 2017) adapts the degree of surprise and unexpectedness of a recommended POI, tailored to users' curiosity values.

2.4 Use of ontologies

The huge amount of data that can be managed in recommender systems, related to users' information, users' preferences, context factors, POI, etc., demands the use of more complex knowledge. Even though, recommender system is an area that has been the focus of many studies, thus reaching a very good level of maturity, there is still a lack of standardization to represent such information. In this sense, it is evident the necessity of a well-defined and standard model for representing the knowledge managed by recommender systems. Semantic Web, in particularly the use of ontologies, seems to be a clear solution, from which we can take its organizational and relational capacity. In the context of tourism, ontology-based recommender system is an emerging trend (Borràs et al., 2014; Yochum et al., 2020).

Some systems only count on ontologies to represent tourism POI (García-Crespo et al., 2009; Bahramian et al., 2017; Arigi et al., 2018; Rajaonarivo et al., 2019). Other works consider user profile ontologies, besides tourism ontologies, such in (Ruotsalo et al., 2013). In (Alonso et al., 2012), an ontology to represent user preferences and context factors is proposed.

2.5 Discussion

Table 1 shows a comparative evaluation of the referenced works, in terms of **category** of the system (model-based, hybrid, etc.), the access approach (**PULL or PUSH**), type of **user's information** gathered, **context** factors considered, **ontology** used, how they handle **serendipity**, and the **information extraction** model to obtain user and context data (explicit or implicit). All cited recommender systems take into account some information related to the user, such as identification aspects (sex, age, profession, etc.), interest and preferences, or social relations, following a PULL approach; thus, users ask for recommendations and have to control the system. Only few of them apply techniques to analyze implicit information, mainly from social media; hence, reducing the intervention of users and being able of dynamically updating the respective information, as we propose in the RECESO architecture. Different types of context data is considered for these recommender systems, demonstrating the importance of such aspect. The majority of them do not offer formal representation of the knowledge managed with ontologies neither a level of serendipity, as we do with RECESO.

The study presented by (Rajaonarivo et al., 2019) is the only work that proposes a context-aware system using ontologies and approaching serendipity, as our proposed system RECESO, nonetheless it follows a PULL approach, the extraction of information is based on explicit interaction of users, and authors do not experiment the serendipitous feature. In contrast, RECESO supports the PUSH paradigm by implicitly gathering users' preferences and context and automatically making recommendations (without expecting users asking), supports the PULL paradigm by explicit user interactions, and we experimentally evaluate the serendipity.

3 RECESO Architecture

In order to overcome the limitations of existing recommender systems, we have identified the following requirements for a new proposal:

- Support the PUSH paradigm, thus reducing the user intervention to get recommendations.
- User-centric, thus returning more relevant items to the users.
- Context-aware approach, for recommending more suitable POIs.
- Ontology-based system, in order to offer an interoperable and flexible system.
- Variability in recommendations to increase serendipity and offer more diverse experiences.
- Combine the extraction of explicit and implicit information of users (e.g., from social media, their smartphones), to manage up to date information.

RECESO can be accessed from a front-end that implements either the PULL or PUSH paradigm or both; it uses user's preferences for recommending relevant places; it uses contextual data gathered by the front-end alongside user's preferences for making context-aware recommendations; a tourism ontology is used for modeling the knowledge managed by the recommender system (i.e., users' preferences, context factors, and POI); and it implements an aging system for recommending the less frequent places more. The general architecture of RECESO is mainly composed by the follwing four modules, as shown in Figure 1:

- **Data Gathering Module**: User's preferences are received by the system, explicitly (i.e., direct in-

Table 1: Related work on recommender systems for e-tourism

Reference	Category/ PULL or PUSH	Users' information	Context	Ontology	Serendipity	Information extraction
García-Crespo et al. (2009)	Hybrid PULL	Preferences, Social	Location	Tourism		Explicit, Implicit
Alonso et al. (2012)	Hybrid PULL	Social, Preferences	Weather, Day, Transport, Special Needs	Tourism, Context		Explicit, Implicit
Ruotsalo et al. (2013)	Model based PULL	Duration, Motivation, Ability	Location	Tourism, User		Explicit
Shen et al. (2016)	Model based PULL	Preferences, History	Location, Time			Explicit
Hidasi and Tikk (2016)	Model based PULL	Social, Feedback General	General			Explicit, Implicit
Kesorn et al. (2017)	Hybrid PULL	Social, Check-in data	Transportation, Location, Weather			Implicit
Bahramian et al. (2017)	Hybrid PULL	Preferences	Weather, Time, Day	Tourism, Context		Explicit
Laß et al. (2017)	Memory based PULL	Preferences, Feedback	Weather, Previously visited, Time, Day, Opening hours			Explicit
Menk et al. (2017)	Hybrid PULL	Social, History, Education			Curiosity	Implicit
Arigi et al. (2018)	Model based PULL	Preferences	Transportation, Location, Weather	Tourism		Explicit
Logesh et al. (2018)	Hybrid PULL	Social Opinion mining	Transportation, Location, Weather			Implicit
Rajaonarivo et al. (2019)	Hybrid PULL	Basic, Preferences	Transportation, Location, Weather	Tourism	Collaborative	Explicit
Santos et al. (2019)	Model based PULL	Health				Explicit
Logesh and Subramaniyaswamy (2019)	Hybrid PULL	Social, Opinion mining	Transportation, Location, Weather			Implicit
Sun et al. (2019)	Model based PULL	Social, Pictures Textual	Transportation, Weather,			Explicit
RECESO	Hybrid PULL/PUSH	Preferences	Weather, Time, Day, Location	Tourism, User, Context	Aging	Explicit, Implicit

teraction) or implicitly (e.g., data mining, social networks analysis, user's pictures analysis). These preferences are related to categories of POI, according to the POI ontology.

- **User Interest Module**: User's preferences are propagated from higher classes to lower classes of the POI ontology.
- **Context Module**: The system receives information about the context of the user, explicitly or implicitly (retrieved from an API, mobile information, etc.).
- **Recommendation Module**: The system recommends a set of places to the user.

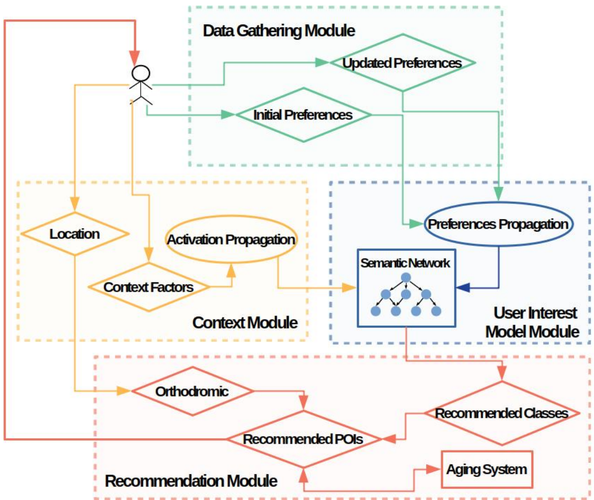

Figure 1: System architecture

3.1 Data Gathering Module (User Interface Client)

Initially, the system receives or calculates initial **preferences**, as values between 0 and 1, for the higher classes of the POI ontology. To obtain these values, the user could explicitly set them by interacting with the system or they could be implicitly determined using data mining, social network analysis, geo-data, similar profile users, friends preferences, etc. During the lifetime of the system, users' preferences are updated, according to their behaviors.

This module represents the client of the system and can be implemented as a mobile apps, web apps, or any other suitable *front-end*, which should be in charge of supporting PULL or PUSH paradigms. PULL approach demands users to explicitly interact with the system to specify her/his data (e.g., personal information, preferences, context) or to ask for a recommendation; while with the PUSH paradigm, the *front-end* gathers information to send to the system, which then returns the recommendation.

3.2 User Interest Model Module

Inspired on the work of Bahramian et al. (2017), we introduce the concept of **semantic network**: a tourism ontology extended with the user's preferences (see Figure 2) and context factor links (see Figure 4). We take advantage of the hierarchy of the semantic network for propagating the preference of superclasses to subclasses. Alongside preferences, each node of the semantic network has a **confidence** related to the user, a value between 0 and 1, that defines how sure is the system about the preference. When a preference is explicitly given by the user, its confidence is 1, but when it is inferred from its ancestors or other kind of analysis, its confidence should be less than 1.

For each user, we compute the preference ($pref_c$) and the confidence ($conf_c$) of each class c, according to Eq. (1) and Eq. (2), respectively; where $ancestors(c)$ is the set of ancestors of the ontology class c, $pref_c$ is the user's preference of the class c, $conf_c$ is the confidence about the user's preference for class c, and α is the *decrease rate* parameter, indicating how much the *confidence* decreases at each level.

$$pref_c = \frac{\sum\limits_{p \in ancestors(c)} conf_p pref_p}{\sum\limits_{p \in ancestors(c)} conf_p} \quad (1)$$

$$conf_c = \frac{\sum\limits_{p \in ancestors(c)} conf_p}{|ancestors(c)|} - \alpha \quad (2)$$

These calculations are applied to each node traversing from the higher classes, whose preferences are obtained from the Data Gathering Module, to the *sinks*, while decreasing confidence values and therefore preference values. This process is called **preference propagation**. Figure 2 shows an example of initial preferences gathered with values 0.85, 0.5, and 0.7 for the ontology classes *Cultural*, *Store*, and *Sport*, respectively. The initial confidence for each preference is 1, since they are gathered explicitly from the user. Then, Figure 3 shows the preference propagation with these values, applying Eq. (1) and Eq. (2), with $\alpha = 0.1$.

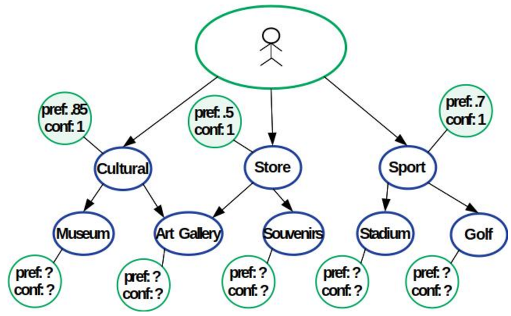

Figure 2: Initial preferences for `Cultural`, `Store`, and `Sport` classes

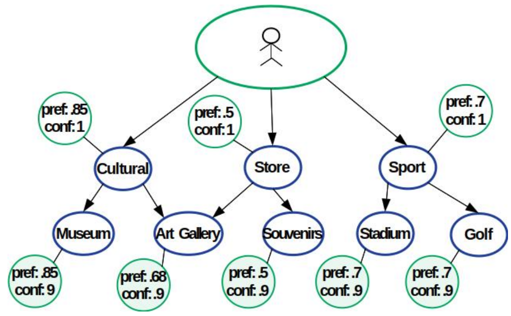

Figure 3: Preference propagation for `Museum`, `Art Gallery`, `Souvenirs`, `Stadium`, and `Golf` classes, with 0.1 as decrease rate (α=0.1)

3.3 Context Module

We define the **activation** of a node of the semantic network as the value that determines how feasible is the user to visit a place that belongs to the category of an ontology class, according to the current user's context. The user's context is determined by the **context factors** (e.g., time, day, weather, transportation, mood), that could affect her/his decision to go to a specific place. To make a recommendation, the system has the context factors per user ($contextFactors$ set), obtained either explicit (provided by users) or implicit (deduced by historical behaviors, data mining, etc.). For example, in Figure 4 the context of the user is *cloudless, night*, and *weekend*.

Let's define f_x the *fulfillment* of a context factor x, where $f_x = 1$ if x fulfills or $f_x = 0$ otherwise, and $r_{c,x}$ as the *relevance* of x for the class c, which is an integer value in $[0,2]$ that specifies how much the context factor can affect the user's decision to go to a POI in c (POI ontology class): 1 means indifference, 2 means the fulfillment increases the wish to go to the POI, 0 means the fulfillment decreases the wish to go the

POI. Hence, we define act_c as the activation for a class c which is directly linked to the user's context factors ($contextFactors$), as shown in Eq. (3), and the activation for subclasses of c is defined in Eq. (4).

$$act_c = \sum_{x \in contextFactors} r_{c,x} f_x \quad (3)$$

$$act_c = \frac{\sum_{c' \in ancestors(c)} act_{c'}}{|ancestors(c)|} \quad (4)$$

.

In the following, we show an example (Figure 4, Figure 5, and Figure 6), where explicit or implicit obtained *relevance* are:

- $r_{Cultural,cloudless} = 1$
- $r_{Cultural,night} = 0.5$
- $r_{Cultural,weekend} = 1.8$
- $r_{Store,cloudless} = 1.5$
- $r_{Store,night} = 1.3$
- $r_{Store,weekend} = 1$
- $r_{Sport,cloudless} = 2$
- $r_{Sport,night} = 0.2$
- $r_{Sport,weekend} = 1.5$
- $f_{cloudless} = f_{night} = f_{weekend} = 1$
- $f_x = 0, x \notin \{cloudless, night, weekend\}$

First, the system deduces (from the user smartphone) the context of the user, which in this example is *cloudless night on the weekend* (Figure 4). Then, using Eq. (3), the activation values for *Cultural*, *Store*, and *Sport* are computed (Figure 5). Finally, using Eq. (4), the activation values for *Museum*, *Art-Gallery*, *Souvenirs*, *Stadium*, and *Golf* are computed (Figure 6). Also, this module detects the user's **location**, which is used for querying near POI.

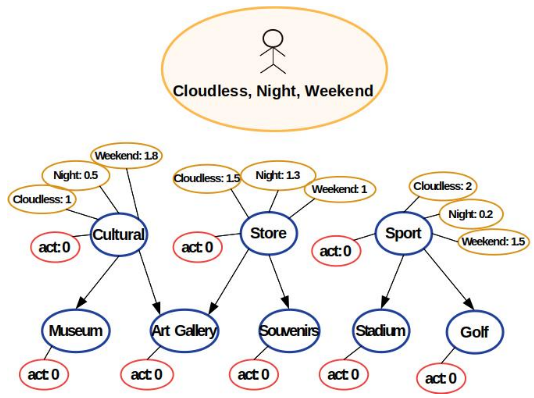

Figure 4: System receives `cloudless, night, weekend` as user context

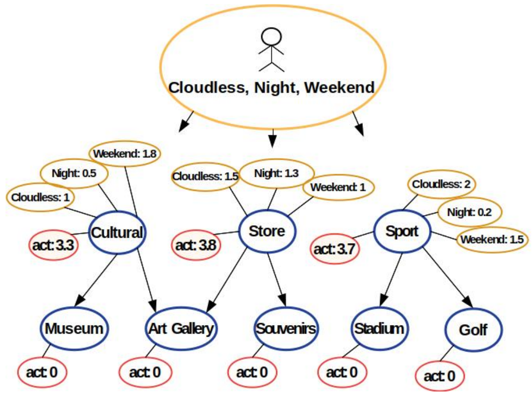

Figure 5: Compute activation for `Cultural, Store`, and `Sport` classes

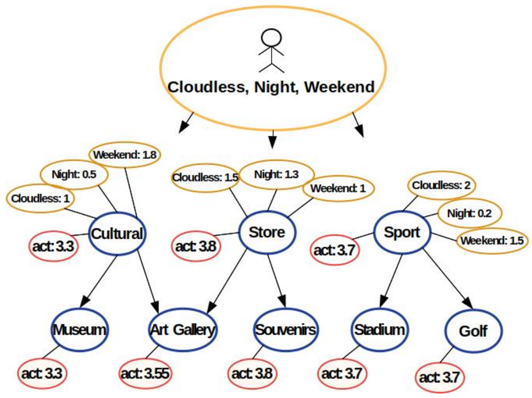

Figure 6: Spread activation to `Museum, Art Gallery, Souvenirs, Stadium`, and `Golf` classes

3.4 Recommendation Module

Besides user's preferences related to POI and context factors, the system bases the recommendations on an *aging-like* algorithm to ensure serendipity, on the distance of the POI to the user's location and its transportation medium, and on calculated scores of POI.

3.4.1 Aging System

Let's define η_p as the POI p's aging, initialized as $\eta_p = 1$. Let's define H as the *aging rate*. Each time a POI p is recommended to the user, η_p decreases by H. When $\eta_p < 0.1$, η_p is reset to 1.

3.4.2 Great-Circle distance

Since the euclidean distance between two points on Earth would cross through the surface, we should use a more convenient measurement of distance: the *great-circle distance* or *orthodromic distance*. It is the shortest distance, along the surface of a sphere, between two points on the surface of the sphere. It is measured with circles on the sphere whose centers coincide with the center of the sphere. Those circles are called *great-circles*. If we assume Earth is a perfect sphere and hence use Great-Circle distance, we get distances with errors no more than 0.5%, according to (Ministry of Defense, London, 1997). The distance between two points i and j on a sphere of radius r is computed as shown in Eq. (5).

$$dist_{i,j} = r \cdot arccos(cos(lat_i) \cdot cos(lat_j) \cdot cos(lon_i - lon_j) + sin(lat_i) \cdot sin(lat_j)) \tag{5}$$

3.4.3 Score

The system calculates the *score* of each instance p of a class c in the ontology, in terms of user's preferences (Eq. (1)), activation level (Eq. (3) or Eq. (4)), the aging value (η_p), and the great-circle distance. Eq. (6) shows how the *score* is calculated; where w_i are configuration parameters of the system that represent the weigh of each term ($\sum w_i = 1$), $pref_c$ is the user's preference on the class c, act_c is the activation level of class c, η_p is the aging factor of the POI p, $dist_{u,p}$ is the great-circle distance between user's location and POI p, and *maxdist* is the maximum great-circle distance that a p should be from user u (this distance depends on the user transportation that can be detected from the user's smartphone sensors, for example).

$$score_p = w_1 \cdot pref_c + w_2 \cdot act_c + w_3 \cdot \eta_p - w_4 \cdot \frac{dist_{u,p}}{maxdist} \tag{6}$$

The system uses this score for returning the list of "top places", whose length could be configured.

4 Study case: experiments and evaluation

To evaluate our system, we implemented a prototype[4], with Java, a Fuseki server as the triplestore to manage the semantic repository with the ontology, Apache Jena for querying the ontology and connecting to the triplestore, and a MySQL database for storing additional information. To have initial data, we requested real users through questionnaires, whose information was used to create simulated users and scenarios for the tests. With this version of RECESO, we evaluate the User Interest Model Module, the Recommendation Module, and the Activation Propagation algorithm from the Context Module, while the Data Gathering Module and the context gathering (Location and Context Factors) are simulated.

[4]https://github.com/gustavoaca1997/datatourisme-recommender

4.1 Ontology Classes

We consider a light version of DATATourisme ontology[5], an open data ontology to manage tourism in France, consisting of the *Point of interest* class, its *Place* subclass, from which we take: *Cultural Site, Food Establishment, Leisure Place, Natural Heritage, Sports,* and *Store*. We slightly modified the DATATourisme ontology to specialize the class *Sports and Leisure Place* into separated *Sports* class and *Leisure Place* class, hence making less ambiguous what kind of places should belong to each category. Figure 7 shows the subset of classes of DATATourisme ontology, considered in the experiments.

The high level classes that are directly linked to the context factors values are:

- *Museum*
- *Interpretation Center*
- *Library*
- *Park and Garden*
- *Archaeological Site*
- *Religious Site*
- *Remarkable Building*
- *City Heritage*
- *Defense Site*
- *Remembrance Site*
- *Technical Heritage*
- *Food Establishment*
- *Natural Heritage*
- *Sports*
- *Leisure Place*
- *Store*

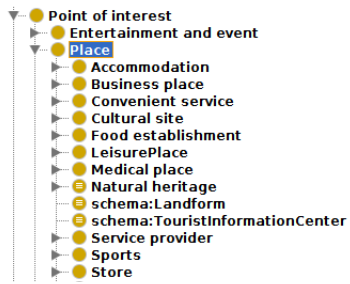

Figure 7: Subset of the modified version of the ontology DATATourisme

4.2 From real users to simulated scenarios

In order to test several scenarios with different user's preferences, profiles, and behaviors, we created synthetic scenarios, however based on data gathered from real people. We asked people to fill a google form with two parts:

[5]http://info.datatourisme.gouv.fr/ontology/core/2.0/

(1) To determine the relevance for each context factor value. The list of asked context factors included weather (with possible values `rainy, cloudless, snowy`), time (with possible values `early morning, morning, afternoon, night`), and day (with possible values `weekday, weekend`). The list of asked types of POI was the high level classes listed in the previous section. The question was "how much the context factor x influences your decision to go to a 'place/event' of the type c (these types are not exclusive)". People should answer an integer value between 0 and 2, where:

- 0 means that if the context factor is met, the user would not go to the place/event.
- 1 means the user does not care if the context factor is met or not.
- 2 means that if the context factor is met, the user would go to the place/event.

Then, for each pair of ontology class, c, and context factor, x, (for example *Museum* with *rainy*), the average relevance is computed and stored as the relevance for the experiments.

(2) To determine users' preferences for tourism categories. In this part of the questionnaires, people provided their preferences of the high level ontology classes (representing POI), and also the genre, country, profession, age, and (optional) social networks to have more information for further work. We got 102 answers from people of different ages, professions, and countries. With this information we generate synthetic tourists, with different behaviors for different context scenarios.

Simulated scenarios. To evaluate the system, we test it with simulated scenarios (simulated contexts and simulated users). The data gathered with questionnaires were grouped into four cluster, with the *Elbow Method* (increasing the number of clusters does not make improvements worth the cost using our small dataset). The centroids of the four clusters represent our four simulated tourists, T_1, T_2, T_3, T_4. Table 2 shows the preferences for each tourist to each tourism category, according to our light version of DATATourise ontology. Figure 8 shows for each tourist how the preferences are distributed, where we can see that T_3 is the one with a more varied set of preferences, including the lowest values between all preferences.

For our experiments, all tourists have the same relevance values, as we show in Table 3. For each tourist T_i, we simulate visits to Paris, Niza, and Lyon. For each combination of context factors (24 combinations), each T_i goes twice to each site; thus 48 visits

Table 2: Simulated tourists' preferences

Class	T_1	T_2	T_3	T_4
Museum	0.7235	0.6476	0.78	0.9421
Interpretation Center	0.6824	0.6476	0.52	0.8737
Library	0.5588	0.46667	0.72	0.7684
Park and Garden	0.865	0.8714	0.7	0.9316
Archaeological Site	0.6882	0.8619	0.82	0.826
Religious Site	0.45882	0.7	0.36	0.7316
Remarkable Building	0.6529	0.8524	0.4	0.7895
City Heritage	0.7412	0.919	0.58	0.8632
Defence Site	0.5412	0.8529	0.68	0.7421
Remembrance Site	0.4941	0.781	0.8	0.7211
Technical Heritage	0.4059	0.7905	0.56	0.5842
Food Establishment	0.9059	0.8905	0.7	0.6632
Natural Heritage	0.9059	0.9143	0.76	0.9211
Sports	0.9353	0.8429	0.22	0.6316
Leisure Place	0.9353	0.8429	0.22	0.6316
Store	0.7647	0.8286	0.28	0.5684

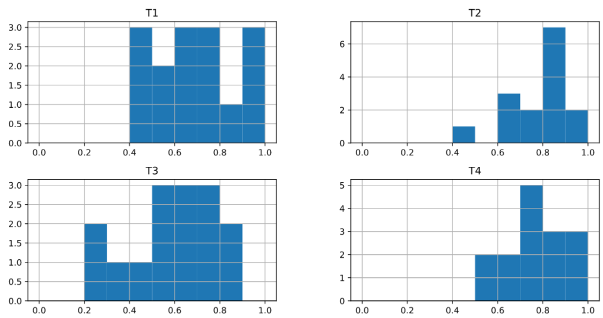

Figure 8: Distributions of initial preferences of T_1, T_2, T_3, T_4

(see Figure 9).

4.3 Metrics

Let R be a recommendation set for a specific user on a specific location and a specific context (combination of context factors), we define some metrics to evaluate how "good" is a *recommendation* R, as follows:

- Average preference of R (Eq. (7)), which denotes how near to the user's preferences are the recommended POIs.

$$pref_R = \frac{\sum_{p \in R} pref_p}{|R|} \tag{7}$$

- Average activation of R (Eq. (8)), which denotes how relevant to the user are the current context factors to the recommended POIs.

$$act_R = \frac{\sum_{p \in R} act_p}{|R|} \tag{8}$$

- Average aging of R (Eq. (9)), which denotes how "aged" to the user are the recommended POIs.

$$\eta_R = \frac{\sum_{p \in R} \eta_p}{|R|} \tag{9}$$

Table 3: Simulated relevance of context factors for T_1, T_2, T_3, T_4

Class	rainy	cloudless	snowy	workday	weekend	morning	afternoon	night	early morning
Museum	1.0	1.394	1.029	1.082	1.488	1.065	1.518	0.788	0.194
Interpretation Centre	0.871	1.312	0.829	1.029	1.424	1.029	1.382	0.771	0.135
Libreary	1.271	1.388	1.188	1.312	1.159	1.324	1.453	0.729	0.271
Park and garden	0.335	1.847	0.894	1.365	1.665	1.388	1.629	0.924	0.494
Archeological Site	0.476	1.488	0.671	0.782	1.471	1.235	1.424	0.682	0.335
Religious Site	0.8	1.018	0.841	0.641	1.106	1.018	0.971	0.529	0.1
Remarkable building	0.912	1.576	0.994	1.165	1.506	1.241	1.453	1.0056	0.324
City heritage	0.665	1.629	0.947	1.088	1.618	1.329	1.565	1.271	0.5
Defense Site	0.712	1.412	0.806	0.835	1.376	1.253	1.388	0.824	0.265
Remembrance Site	0.588	1.318	0.812	0.935	1.306	1.129	1.294	0.6	0.324
Technical Heritage	0.582	1.371	0.682	0.841	1.3	1.265	1.3	0.741	0.3
Food Stablishment	1.488	1.624	1.494	1.635	1.759	1.353	1.671	1.676	0.688
Natural Heritage	0.494	1.776	0.835	1.088	1.659	1.535	1.524	0.953	0.465
Leisure place	1.088	1.541	1.006	1.118	1.588	1.024	1.506	1.418	0.588
Sports	0.1	2.0	0.75	1.118	1.588	1.024	1.506	0.8	0.588
Store	1.312	1.535	1.241	1.388	1.435	1.0588	1.547	1.1647	0.3

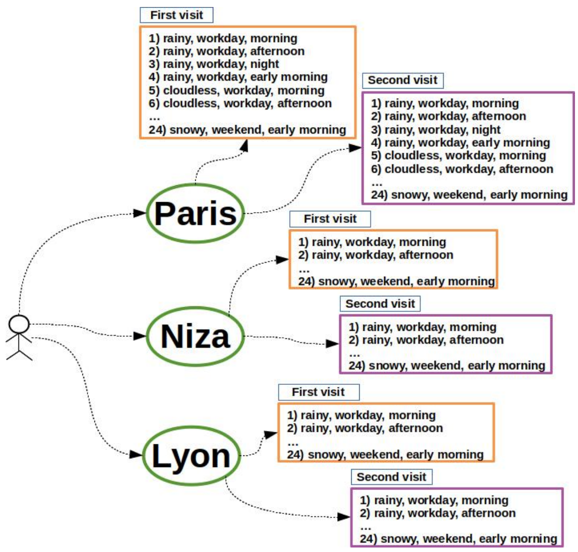

Figure 9: Simulated visits

- Average distance of R (Eq. (10)), which denotes how far are the recommended POIs from the user.

$$dist_R = \frac{\sum_{p \in R} dist_p}{|R|} \quad (10)$$

- Average novelty of R (Eq. (12)), which denotes how novel are the recommended POIs to the user. In Eq. (12), nov_p is calculated by Eq. (11), which is based on the work of Kotkov et al. (Kotkov et al., 2016), where p is a recommended place, rec is the set of already recommended places to the user, c_p is the ontology class to which p belongs, and $dist$ computes the distance of two classes in the ontology (using *Breadth First Search*).

$$nov_p = \min_{q \in rec} (dist(c_p, c_q)) \quad (11)$$

$$nov_R = \frac{\sum_{p \in R} nov_p}{|R|} \quad (12)$$

4.4 Experiments

To calculate the score for each POI p, we need some user-defined parameters, that represent maximum distance (*maxdist*) from the user to the POI and the weighs (w_i) for each term in Eq. (6) (see Section 3.4.3). These user-defined parameters are specified just once, at the installation/configuration of the system. We designed two scenarios with different values of these configuration parameters, as we explain in the following.

4.4.1 First configuration

We consider $maxdist = 8Km$, $w_1 = 0.3611$ for preference, $w_2 = 0.3611$ for activation, $w_3 = 0.25$ for aging, and $w_4 = 0.0278$ for distance from user, resulting on Eq. (13). In this configuration, we give more priority to preference and activation and very little priority to distance from user. With these parameters we start simulating each user and their visits.

$$score_p = 0.3611 \cdot pref_c + 0.3611 \cdot act_c + 0.25 \cdot \eta_p - 0.0278 \cdot \frac{dist_{u,p}}{8} \quad (13)$$

For space limitations, we are showing only some sets of recommended items like Table 4, where each p_i is a recommended place and $score_{p_i} \geq score_{p_j}$ for every $i < j$. Table 4 corresponds to T_1's first visit to Niza, on a *rainy workday in the morning*, where p_1 is Train des Merveilles (from class *TouristTrain*), p_2 is Cinéma de Plein Air (from class *Cinema*), p_3 is Casino de Beaulieu (from class *Casino*), p_4 is Cinéma de Beaulieu (from class *Cinema*), and p_5 is Lyon-style Petanque Fields (from class *BoulesPitch*). Note that since all places are instances of *LeisurePlace* class, which T_1 loves, they have same preference and activation; also, because this is the first recommendation, all aging values are 1.0. Therefore, the distance is the tiebreaker. Despite of a *rainy morning* context, the system recommends a *Tourist Train*, just because it is a Leisure Place.

Table 4: First visit of T_1 to Niza with first configuration and context = (rainy, morning, workday)

Field	p_1	p_2	p_3	p_4	p_5
$pref_c$	0.9353	0.9353	0.9353	0.9353	0.9353
act_c	4.4	4.4	4.4	4.4	4.4
η_p	1.0	1.0	1.0	1.0	1.0
$dist_{u,p}$	1.53	3.99	5.52	5.53	5.58
$score_p$	2.1713	2.1628	2.15745	2.15742	2.1573

At the second visit to Lyon (see Table 5), on a *rainy workday in the morning*, only p_1 is not a Leisure Place but an Interpretation Center, a kind of place T_1 does not like as much as leisure places. However, the activation of *InterpretationCenter* is higher enough to make $score_{p_1}$ greater than $score_{p_2}$. Despite of p_1 being very far away from T_1's location, the little magnitude of w_4 makes it little important to the final score. Something odd on the recommended set is that p_3 and p_4 are the same place, Le Sucre, but reported as belonging to two different ontology classes: *NightClub* and *Theatre*.

Table 5: Second visit of T_1 to Lyon with first configuration and context = (rainy, morning, workday)

Field	p_1	p_2	p_3	p_4	p_5
$pref_c$	0.676	0.9353	0.9353	0.9353	0.9353
act_c	4.818	4.51	4.51	4.51	4.51
η_p	1.0	1.0	1.0	1.0	1.0
$dist_{u,p}$	7.1	4.45	4.54	4.54	4.61
$score_p$	2.2092	2.2015	2.2012	2.2012	2.2010

At the fourth of visit T_1 to Niza, on a *rainy workday in the early morning*, the system recommends what is shown in Table 6, where p_1 is Cinéma de Plein Air (from class *Cinema*), p_2 is Ferronnerie d'art JC Rodriguez (from class *CraftsmanShop*), p_3 is Atelier Hesperida (from class *CraftsmanShop*), p_4 is Horlogerie Foltête (from class *CraftsmanShop*), and p_5 is Galerie Bizet (from class *CraftsmanShop*). Since T_1 loves leisure places more than stores, the predicted preference for Cinéma de Plein Air is greater than the other ones on this visit. However, we can see that η_{p_1} lowers down $score_{p_1}$ to a value not so different from $score_{p_2}$. The recommendations to T_1 returned after this one do not have Cinéma de Plein Air in the recommended POIs until the context *snowy workday afternoon* is reached in the loop, because of the aging of Cinéma de Plein Air and its low activation.

Table 6: Fourth visit of T_1 to Niza with first configuration and context = (rainy, early morning, workday)

Field	p_1	p_2	p_3	p_4	p_5
$pref_c$	0.9353	0.7647	0.7647	0.7647	0.7647
act_c	4.2	4.1	4.1	4.1	4.1
η_p	0.6	1.0	1.0	1.0	1.0
$dist_{u,p}$	3.99	5.21	5.25	5.72	5.737
$score_p$	1.9906	1.9886	1.9884	1.98684	1.98678

While the system recommends same places to T_1 and T_2 for their first visits to Niza, Lyon, and Paris, at *rainy workdays at afternoon*, the sets start to diverge after the second visits. The system recommends to T_1 a set of Leisure Places and one Interpretation Center for the second visit to Lyon, but recommends a set with two Archeological Sites, one Interpretation Center, and two Remarkable Buildings to T_2, in that order (see Table 7). The more diverse initial preferences of T_2 and the aging system are responsible for this.

Table 7: Second visit of T_2 to Lyon with first configuration and context = (rainy, afternoon, workday)

Field	p_1	p_2	p_3	p_4	p_5
$pref_c$	0.852	0.852	0.659	0.843	0.843
act_c	4.659	4.659	4.818	4.6	4.6
η_p	1.0	1.0	1.0	1.0	1.0
$dist_{u,p}$	3.967	5.267	7.0969	3.967	5.267
$score_p$	2.226	2.222	2.203	2.202	2.197

After many visits, again at a *rainy workday at afternoon*, the system recommends a completely different set of places to T_2 when visiting Lyon. It recommends a Theater, an Art Gallery, a Craftsman Shop, and two other Art Galleries, in that order (see Table 8). The system aging is responsible for this variety of recommendations.

Table 8: A visit of T_2 to Lyon with first configuration after many visits and context = (rainy, afternoon, workday)

Field	p_1	p_2	p_3	p_4	p_5
$pref_c$	0.843	0.829	0.829	0.829	0.829
act_c	4.512	4.371	4.371	4.371	4.371
η_p	0.8	1.0	1.0	1.0	1.0
$dist_{u,p}$	7.157	5.425	5.438	5.442	5.462
$score_p$	2.10876	2.10864	2.10859	2.10858	2.10851

With a low value of w_3, which represents aging's weight, there are cases where the recommendations are not very diverse, as is the case of the two visits of T_2 to Paris on *cloudless workdays at early morning*, shown in Table 9 and Table 10, where the second one represents the recommendations made after many visits. Both visits involve Garnier Opera (from class *Palace*), Conciergerie (from class *Palace*), and Le Manoir de Paris (from class *InterpretationCentre*), as p_2, p_3, and p_4 for the first visit and p_1, p_5, and p_2 for the second visit, respectively.

Table 9: First visit of T_2 to Paris with first configuration and context = (cloudless, early morning, workday)

Field	p_1	p_2	p_3	p_4	p_5
$pref_c$	0.829	0.843	0.843	0.659	0.659
act_c	4.171	4.4	4.4	4.571	4.571
η_p	0.8	0.4	0.4	0.4	0.4
$dist_{u,p}$	7.51	4.25	5.06	5.9	6.3
$score_p$	1.97916	1.97869	1.97590	1.96803	1.96665

Table 10: Second visit of T_2 to Paris with first configuration and context = (cloudless, early morning, workday)

Field	p_1	p_2	p_3	p_4	p_5
$pref_c$	0.843	0.843	0.659	0.659	0.659
act_c	4.40	4.40	4.57	4.57	4.57
η_p	0.4	0.4	0.4	0.4	0.4
$dist_{u,p}$	4.25	5.06	3.79	4.55	5.9
$score_p$	1.97869	1.97590	1.97535	1.97273	1.96803

Table 11 shows the average of the recommendation metrics for each tourist using the first configuration. Averages of $pref_R$ of T_1's and T_2's visits are the highest ones, being greater than 0.8. Next is T_4's, being greater than 0.7, which is still a good average preference. T_3's average $pref_R$ is lower, but we can blame the amount of low initial preferences of T_3, as we can see on Figure 8.

Table 11: Averages of recommendations metrics with first configuration

avg	T_1	T_2	T_3	T_4
$avg(pref_R)$	0.8555	0.8213	0.4183	0.7269
$avg(act_R)$	4.286	4.3092	4.314	4.3319
$avg(\eta_R)$	0.7058	0.7111	0.6906	0.6914
$avg(nov_{u,R})$	0.0923	0.1231	0.1329	0.1147
$avg(dist_{u,R})$	4.876	4.8993	4.9101	4.9861

The average act_R of each tourist is greater than 4.0. Theoretically, the maximum possible value of act_R is 6.0, but since it depends of the activation value act_c of the recommended ontology classes (see Eq. (8)), which depends of the relevances $r_{c,x}$ of the user (see Eq. (4)), that do not reach their maximum possible value 2.0 in the survey (see Section 4.2), it is impossible for these experiments to reach the maximum act_R. Therefore, we can say that an average act_R of 4.0 is good enough.

Average η_R is very near 0.7, hence the amount of "young" recommended places is high. Theoretically, the maximum possible value is 1.0, but that would be possible if each recommendation involves POI never recommended before.

Despite of recommending places with distance from user near 8.0 km, the average $dist_{u,R}$ for each tourist is acceptable: greater than 4.0 km and less than 5.0 km. The novelties measured as we proposed on Section 4.3 have bad performance on four tourists.

4.4.2 Second configuration

Now we consider $w_1 = 0.25$ for preference, $w_2 = 0.25$ for activation, $w_3 = 0.472$ for aging, and $w_4 = 0.0278$ for distance from user, resulting on equation 14 and giving now more priority to aging.

$$score_p = 0.25 \cdot pref_c + 0.25 \cdot act_c + 0.472 \cdot \eta_p - 0.0278 \cdot \frac{dist_{u,p}}{8} \quad (14)$$

For the first visit of T_1 to Niza in a *cloudless workday at afternoon*, the η_R is 0.84 (see Table 13), while it is 0.80 using the first configuration (see Table 12). The $pref_R$ improves, from 0.765 to 0.799, and the act_R goes from 4.44 to 4.33.

Table 12: First visit of T_1 to Niza with first configuration and context = (cloudless, afternoon, workday

Field	p_1	p_2	p_3	p_4	p_5
$pref_c$	0.76	0.76	0.76	0.76	0.76
act_c	4.44	4.44	4.44	4.44	4.44
η_p	0.8	0.8	0.8	0.8	0.8
$dist_{u,p}$	5.25	5.72	5.74	5.78	5.79
$score_p$	2.06166	2.06004	2.05998	2.05983	2.05981

Table 13: First visit of T_1 to Niza with second configuration and context = (cloudless, afternoon, workday

Field	p_1	p_2	p_3	p_4	p_5
$pref_c$	0.76	0.76	0.935	0.76	0.76
act_c	4.44	4.44	3.89	4.44	4.44
η_p	0.8	0.8	1.0	0.8	0.8
$dist_{u,p}$	5.20	5.25	5.30	5.72	5.73
$score_p$	1.6612	1.6610	1.6597	1.6594	1.6593

For the second visit of T_4 to Paris on a *snowy weekend at early morning*, the η_R is 0.6 using second configuration (see Table 15), while it is 0.32 using first configuration (see Table 14). This makes sense since we give more weight to aging in the second configuration. The other metrics of R are very similar for both visits.

Table 14: Second visit of T_4 to Paris with first configuration and context = (snowy, early morning, weekend)

Field	p_1	p_2	p_3	p_4	p_5
$pref_c$	0.866	0.866	0.790	0.866	0.866
act_c	4.288	4.288	4.265	4.288	4.288
η_p	0.4	0.4	0.4	0.2	0.2
$dist_{u,p}$	5.899	6.297	4.253	3.791	4.546
$score_p$	1.941	1.939	1.911	1.898	1.895

Table 16 shows the average of the recommendation metrics for each tourist using the second configuration. Again, T_1's and T_2's $avg(pref_R)$ are the highest ones, followed by T_4's, while T_3's is the lowest. We can notice that only T_1's and T_2's $avg(pref_R)$ increase a little from the ones with first configuration, but the other two decrease. Again, all $avg(act_R)$ are

Table 15: Second visit of T_4 to Paris with second configuration and context = (snowy, early morning, weekend)

Field	p_1	p_2	p_3	p_4	p_5
$pref_c$	0.866	0.866	0.738	0.866	0.866
act_c	4.288	4.288	3.81	4.288	4.288
η_p	1.0	1.0	0.6	0.2	0.2
$dist_{u,p}$	3.791	4.546	5.333	5.899	6.297
$score_p$	1.748	1.745	1.402	1.363	1.361

above 4.0, but each one decreases from the first configuration version. As expected, all $avg(\eta_R)$ increase because the second configuration gives more weight to aging.

Table 16: Averages of recommendations metrics with second configuration

avg	T_1	T_2	T_3	T_4
$avg(pref_R)$	0.8658	0.8272	0.3912	0.6987
$avg(act_R)$	4.0738	4.0818	4.0031	4.0415
$avg(\eta_R)$	0.7592	0.7619	0.7850	0.7739
$avg(nov_{u,R})$	0.172	0.1860	0.1986	0.1580
$avg(dist_{u,R})$	4.842	5.0215	4.8999	4.9615

4.4.3 Comparison with related work

We implemented the system proposed in (Bahramian et al., 2017) and executed it using the same score function and same first configuration, we obtain the average metrics shown in the Table 17. The average $pref_R$ for each T_i are very similar; likewise for the average activation values. The average nov_R of RECESO are higher, showing how our proposal improve serendipity.

When we use the second configuration, we obtain the average metrics shown in the Table 18. We can see that for T_1 and T_2, the average preferences returned by RECESO are better, but for T_3 and T_4 the average preferences returned by the system proposed in (Bahramian et al., 2017) are better. The average activation values are worse for RECESO, but as expected, the average nov_R is significantly higher for RECESO, alongside the average η_R.

4.4.4 Discussion

For measuring novelty we use Eq. (11) based on Kotkov et al. (2016), since when novelty increases, so does serendipity. That equation was proposed with the hypotheses that if a user already knows places from class c, then it is probable the user already knows other places from class c. With that hypotheses, our system gives low novelty at first glance: when a place belonging to the ontology class c is recommended, next recommendations with places from c will have lower novelty. But when we consider the comparison with the work of Bahramian et al. (2017), the average novelty of RECESO is significantly higher, implying a better serendipity.

Experiments with the second configuration are focused on increasing average aging by giving a higher value to w_3 on Eq. (6), but the increment of aging is paid with a decrease on activation. Nonetheless, two simulated tourists have an increment for their preferences, because on first configuration the system is trapped between very activated POI but not very relevant, but the heavier aging of the second configuration increases variability, hence lets the system recommend other POI despite of lowering the activation.

Nevertheless, our system lets users choose each w_i of Eq. (6), that way choosing if they want recommendations mainly relevant than contextually activated or varied, or even recommendations with the nearest places ignoring any other attribute, or any other possible configuration.

Comparing the proposed architecture of RECESO with state-of-the-art studies (see Section 2), we overcome some of their limitations: (i) combining PULL and PUSH paradigms reduces users intervention and control over the system and allows making recommendations even if users do not ask the system (e.g., when users are near POI, when the weather change); (ii) analyzing implicit information from users' social media or smartphones makes possible to dynamically update users and context information; thus, manage most current data; (iii) Managing semantic information with ontologies to represent touristic POI, user preferences, and context information, ensures a flexible and formal representation of the knowledge, which in turn enables the implementation of more intelligent and interoperable recommender systems as an alternative of machine learning based approaches; and (iv) as shown by results, the *aging-like* algorithm improves serendipity, compared with a state-of-the-art work.

This first prototype of RECESO demonstrates the feasibility of a recommender system with all these advantages over existing studies. This experience also gives the opportunity of extracting its current limitations and some lessons learnt, as analyzing the score and aging systems to improve serendipity and novelty.

5 Conclusions and future work

In this paper we present RECESO, a context-aware and ontology-based recommender system that is able to recommend varied Points of Interest (POI). We use a spreading activation algorithm to model users and their context supported on an ontology that represents POI and users' preferences. We use a proposed ag-

Table 17: Averages of recommendations metrics with first configuration using RECESO and Bahramian et al. system

avg	T_1		T_2		T_3		T_4	
	RECESO	Bahramian	RECESO	Bahramian	RECESO	Bahramian	RECESO	Bahramian
$avg(pref_R)$	0.8555	0.832	0.8213	0.814	0.4183	0.4290	0.7269	0.74167
$avg(act_R)$	4.286	4.347	4.3092	4.368	4.314	4.3558	4.3319	4.366
$avg(\eta_R)$	0.7058	0.623	0.7111	0.620	0.6906	0.6247	0.6914	0.6206
$avg(nov_R)$	0.0923	0.0699	0.1231	0.0853	0.1329	0.0573	0.1147	0.0741
$avg(dist_R)$	4.876	4.359	4.8993	4.520	4.9101	4.7487	4.9861	4.761

Table 18: Averages of recommendations metrics with second configuration using Bahramian et al. system

avg	T_1		T_2		T_3		T_4	
	RECESO	Bahramian	RECESO	Bahramian	RECESO	Bahramian	RECESO	Bahramian
$avg(pref_R)$	0.8658	0.7416	0.8272	0.8144	0.3912	0.4292	0.6987	0.7416
$avg(act_R)$	4.0738	4.366	4.0818	4.3678	4.0031	4.3556	4.0415	4.3660
$avg(\eta_R)$	0.7592	0.6194	0.7619	0.6211	0.7850	0.6247	0.7739	0.61944
$avg(nov_R)$	0.172	0.07413	0.1860	0.0853	0.1986	0.06154	0.1580	0.07413
$avg(dist_R)$	4.842	4.7292	5.0215	4.5136	4.8999	4.7441	4.9615	4.7292

ing system which gives more priority to POIs that are not frequently recommended. To evaluate the performance of RECESO, we perform some experiments, by simulating visits of four simulated tourists, based on real preferences obtained through a survey. Results show that it is able to recommend different sets of places to the same user at the same context at different moments, still taking into account the user's preferences and improving novelty level. We are developing a more complete version of RECESO, including all its modules. We also plan to integrate it into mobile or web apps able to gather user information and contextual information, including location. We expect to perform more validations with real scenarios.

Acknowledgement

This research was partially funded by FONDO NACIONAL DE DESARROLLO CIENTÍFICO, TEC-NOLÓGICO Y DE INNOVACIÓN TECNOLÓGICA - FONDECYT as executing entity of CONCYTECunder grant agreement no. 01-2019-FONDECYT-BM-INC.INV in the project RUTAS: Robots for Urban Tourism Centers, Autonomous and Semantic-based

REFERENCES

Alghamdi, H., Zhu, S., and El Saddik, A. (2016). E-tourism: mobile dynamic trip planner. In *Intl. Symposium on Multimedia*, pages 185–188. IEEE.

Alonso, K., Zorrilla, M., Iñan, H., Palau, M., Confalonieri, R., Vázquez-Salceda, J., Calle, J., and Castro, E. (2012). Ontology-based tourism for all recommender and information retrieval system for interactive community displays. In *Intl. Conf. on Information Science and Digital Content Technology*, volume 3, pages 650–655. IEEE.

Arigi, L. R. H., Baizal, Z. A., and Herdiani, A. (2018). Context-aware recommender system based on ontology for recommending tourist destinations at bandung. In *Journal of Physics: Conference Series*, volume 971. IOP Publishing.

Artemenko, O., Kunanets, O., and Pasichnyk, V. (2017). E-tourism recommender systems: a survey and development perspectives. *Econtechmod*.

Bahramian, Z., Abbaspour, R. A., and Claramunt, C. (2017). A context-aware tourism recommender system based on a spreading activation method. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-4/W4:333–339.

Bobadilla, J., Ortega, F., Hernando, A., and Gutiérrez, A. (2013). Recommender systems survey. *Knowledge-based systems*, 46:109–132.

Borràs, J., Moreno, A., and Valls, A. (2014). Intelligent tourism recommender systems: A survey. *Expert Systems with Applications*, 41(16):7370–7389.

Chen, L., Yang, Y., Wang, N., Yang, K., and Yuan, Q. (2019). How serendipity improves user satisfaction with recommendations? a large-scale user evaluation. In *The World Wide Web Conference*, pages 240–250.

García-Crespo, A., Chamizo, J., Rivera, I., Mencke, M., Colomo-Palacios, R., and Gómez-Berbís, J. M. (2009). Speta: Social pervasive e-tourism advisor. *Telematics and informatics*, 26(3):306–315.

Hamid, R. A., Albahri, A., Alwan, J. K., Al-qaysi, Z., Albahri, O., Zaidan, A., Alnoor, A., Alamoodi, A., and Zaidan, B. (2021). How smart is e-tourism? a systematic review of smart tourism recommendation system applying data management. *Computer Science Review*, 39:100337.

Hidasi, B. and Tikk, D. (2016). General factorization framework for context-aware recommendations. *Data Mining and Knowledge Discovery*, 30(2):342–371.

Jannach, D. and Zanker, M. (2020). Interactive and context-aware systems in tourism. *Handbook of e-Tourism*, pages 1–22.

Kayumovich, K. O. (2020). Prospects of digital tourism development. *Economics*, (1 (44)).

Kazandzhieva, V. and Santana, H. (2019). E-tourism: Definition, development and conceptual framework. *Turizam: medunarodni znanstveno-strucni časopis*, 67(4):332–350.

Kesorn, K., Juraphanthong, W., and Salaiwarakul, A. (2017). Personalized attraction recommendation system for tourists through check-in data. *IEEE Access*, 5:26703–26721.

Kotkov, D., Wang, S., and Veijalainen, J. (2016). A survey of serendipity in recommender systems. *Knowledge-Based Systems*, 111:180–192.

Laß, C., Herzog, D., and Wörndl, W. (2017). Context-aware tourist trip recommendations. In *Workshop on Recommenders in Tourism co-located with 11th ACM Conference on Recommender Systems*, pages 18–25, Como, Italy.

Leskovec, J., Rajaraman, A., and Ullman, J. D. (2020). *Mining of massive data sets. Chapter 9: Recommendation Systems*. Cambridge university press.

Lim, K. H., Chan, J., Karunasekera, S., and Leckie, C. (2019). Tour recommendation and trip planning using location-based social media: a survey. *Knowledge and Information Systems*, pages 1–29.

Lin, C.-Y., Wang, L.-C., and Tsai, K.-H. (2018). Hybrid real-time matrix factorization for implicit feedback recommendation systems. *IEEE Access*, 6:21369–21380.

Logesh, R. and Subramaniyaswamy, V. (2019). Exploring hybrid recommender systems for personalized travel applications. In *Cognitive informatics and soft computing*, pages 535–544. Springer.

Logesh, R., Subramaniyaswamy, V., and Vijayakumar, V. (2018). A personalised travel recommender system utilising social network profile and accurate gps data. *Electronic Government, an International Journal*, 14(1):90–113.

Menk, A., Sebastia, L., and Ferreira, R. (2017). Curumim: A serendipitous recommender system for tourism based on human curiosity. In *Intl. Conf. on Tools with Artificial Intelligence*, pages 788–795. IEEE.

Menk, A., Sebastia, L., and Ferreira, R. (2019). Recommendation systems for tourism based on social networks: A survey. *arXiv preprint arXiv:1903.12099*.

Ministry of Defense, London (1997). *Admiralty Manual of Navigation: BR 45(1)*. Number v. 1 in BR Series. Stationery Office.

Rajaonarivo, L., Fonteles, A., Sallaberry, C., Bessagnet, M.-N., Roose, P., Etcheverry, P., Marquesuzaà, C., Lacayrelle, A. L. P., Cayèré, C., and Coudert, Q. (2019). Recommendation of heterogeneous cultural heritage objects for the promotion of tourism. *ISPRS International Journal of Geo-Information*, 8(5):230.

Raza, S. and Ding, C. (2019). Progress in context-aware recommender systems—an overview. *Computer Science Review*, 31:84–97.

Ruotsalo, T., Haav, K., Stoyanov, A., Roche, S., Fani, E., Deliai, R., Mäkelä, E., Kauppinen, T., and Hyvönen, E. (2013). Smartmuseum: A mobile recommender system for the web of data. *Journal of Web Semantics*, 20:50–67.

Santos, F., Almeida, A., Martins, C., Gonçalves, R., and Martins, J. (2019). Using poi functionality and accessibility levels for delivering personalized tourism recommendations. *Computers, Environment and Urban Systems*, 77:101173.

Shen, J., Deng, C., and Gao, X. (2016). Attraction recommendation: Towards personalized tourism via collective intelligence. *Neurocomputing*, 173:789–798.

Sun, X., Huang, Z., Peng, X., Chen, Y., and Liu, Y. (2019). Building a model-based personalised recommendation approach for tourist attractions from geotagged social media data. *International Journal of Digital Earth*, 12(6):661–678.

Tintarev, N., Flores, A., and Amatriain, X. (2010). Off the beaten track: a mobile field study exploring the long tail of tourist recommendations. In *Intl. Conf. on Human computer interaction with mobile devices and services*, pages 209–218.

Yochum, P., Chang, L., Gu, T., and Zhu, M. (2020). Linked open data in location-based recommendation system on tourism domain: A survey. *IEEE Access*, 8:16409–16439.