A Graph-Based Semantic Recommender System for a Reflective and Personalised Museum Visit

Extended Abstract

Louis Deladiennée and Yannick Naudet
Luxembourg Institute of Science and Technology (LIST)
Luxembourg
E-mail: {louis.deladiennee, yannick.naudet}@list.lu

***Abstract*—Offering personalised recommendations to visitors of a museum is a complex problem inherent to physical spaces. When at the same time specific applicative or museum objectives have to be taken into account, this becomes even more complicated. We introduce here a graph-based semantic recommender approach relying on ontological formalisation of knowledge about manipulated entities to solve the multi-dimensional recommendation problem encountered in museums.**

I. Introduction

The museum visitor journey can be far from relaxing. Before his visit, he has his own objectives and expectations in mind. Then, when exploring the museum, he is confronted to a certain amount of artefacts (e.g., paintings, sculptures) and when multimedia mediation is used, to a greater amount of information (e.g., signage, audio, video) related to the artefacts. This information can describe an item itself, contextualize its creation, its history, give details about its creator... This can be overwhelming and have a negative impact on the visit experience as there may be a gap between the initial visitor expectations and the reality. On the other side, a museum curator needs to decide what message he wants to deliver to the visitors. He might decide to focus only on a specific aspect of the exhibit (e.g., historical), or to provide an exhaustive overview of the whole knowledge he has about the museum pieces of art. This may be a challenging and frustrating task as it requires to exclude possibly interesting knowledge that may not interest all the visitors.

We suggest that the complexity of computing personalised recommendations for visitors while respecting the museum objectives and constraints can be facilitated by leveraging RDF-based semantic datasets and exploiting their graph structure. The work reported here is inspired by recent research exploiting Linked Open Data (LOD) for recommendations [1] [2]. Our approach is different as it does not seek to complete the recommendation search space with LOD, but considers it is formalised with an RDF-based language (e.g., OWL).

II. Semantic Graph-based Recommender

A. Mapping Semantic Knowledge to a Multi-dimensional Graph

There are numerous ways to describe *The Rockeby Venus* from *Diego Velasquez*: as, e.g., a piece of art produced by Velasquez, a representative of nude art in the 17th century, a baroque painting, a representation of Venus, or a former part of the collection of the Marquis del Carpio. From a more prosaic perspective, this painting is a 122.5 x 177 cm oil canvas currently located in Room 30 of the National Gallery of London. Now let's assume that the knowledge supporting these assertions is part of a semantic web dataset. This dataset would contain the knowledge about the different items that form the collection of a museum, but also knowledge about the latter (e.g., physical organisation, opening hours). The OWL implementation of CIDOC CRM [3], a standard for the cultural heritage domain, would be used to model such dataset.

From a semantic recommender perspective, this classical choice introduces several issues preventing the use of usual semantic similarity measures for ontologies, especially because of the flatness of CIDOC CRM (weak hierarchical structure). As proposed in [4], such a dataset can however be represented as a semantic network $SG = (V, P, E)$, where $V = \{v_1, v_2, ..., v_n\}$ is the set of individuals of the dataset, $P = \{p_1, p_2, ..., p_n\}$ is a set of typed links (a direct mapping of the set of properties defined in the ontology), and $E = \{e_1, e_2, ..., e_n\}$ is the set of instances of these links, i.e. the links between the individuals in the ontology such as $e_i =< p_j, v_a, v_b >$.

Each assertion about Diego Velasquez's painting in our example, illustrates a different facet of the body of knowledge available on this painting. Such facet examples are *Authorship* (i.e. who created the painting), *Art History*, *Style*, *History*, *Material* or *Location*. Let F be a set of predefined facets that are relevant in the domain of paintings. For each facet $fac_i \in F$ there exists a sub-graph SG_{fac_i} of SG containing the knowledge supporting it. The union of these sub-graphs forms a graph $SG_r = \bigcup_{fac_i \in F} SG_{fac_i}$, which constitute the search space of the recommendation problem.

Each sub-graph can be easily defined by a path of recursion P_{fac_i}, as defined in [5]: *A path of recursion* p *with length* i *is a sequence whose first element is a class and the others are relations recursively reachable from the class.* In our case, a path of recursion is a sequence of relations that, starting from one kind of node in the graph SG, allows to explore the graph following one semantically consistent facet and construct the associated sub-graph. $\forall fac_i \in F,\ \exists P_{fac_i} : P_{fac_i} \mapsto SG_{fac_i}$. For example, the facet $Authorship \in F$ which models the authorship of a painting can have the associated path of recursion $P_{Authorship} = \{Painting \rightarrow \{isAuthor\} \rightarrow Person\}$. Using $P_{Authorship}$ it is easy to extract from SG the facet-graph

978-1-5386-0756-5/17/$31.00 ©2017 IEEE

$SG_{Authorship}$ that contains only the knowledge supporting the facet $Authorship$.

B. Personalized Recommendation Search Space

There might be different reasons why a visitor is interested by a POI, which probably also differ from one POI to another. The facets of the semantic graph SG formed by all the POI semantic descriptions can represent these reasons, i.e. specific features of a POI that are of interest for the visitor. Those facets of interests vary from one POI to another and are unique for each visitor.

The key element in a content-based approach to the recommendation problem is to find a good similarity measure to compare items to recommend to a user's interests. As highlighted in [5], similarity measurement in ontological datasets needs to be tailored to the application context. Two applications can use data from the same ontology but may explore it in different ways, following or ignoring some properties or concepts based on their own objectives. Extending this idea, we think the similarity measurement needs to be tailored to each visitor. Thus, for each visitor, we define a specific subset $F_{v_i} \subset F$, representing the facets for which she has an interest.

A personalised recommendation space can be defined for each visitor v_i, as the sub-graph of SG comprising only elements of interest: $SG_{v_i} = \bigcup_{fac_i \in F_{v_i}} SG_{fac_i}$.

C. Initialising the Search Space

To deal with the cold start issue, personal facets can be determined from a combination of different approaches. On one hand, experts can create pre-curated facets based on their own expertise, using results of studies about visitor's behaviours and stereotypes, or that support a certain narrative they want to share. In this case, a recursion path is known for each predefined facet. Then, an explicit profiling asking the user to select, rank or give importance scores on several facets can be applied. Lets call $F_{facet.expl,v_i}$ the obtained facet set.

On the other hand, methods like using a carousel displaying items to rate can provide more precise and complementary information. The set of items rated by the user can be used to generate the visitor initial profile as a graph $G_{v_i,0}$, which contains nodes and links built from the semantic description of the items, where importance weighting is done according to the given rates and propagated sequentially item after item with some weight computation algorithm. After each step of weight propagation, the most important recursion path can be identified, thus generating a corresponding set of facets, $F_{itemExpl,v_i}$.

Using these two techniques, at the start of a visit, each visitor can be represented by a set of facets $F_{v_i,0} = \{F_{facetExpl,v_i}, F_{itemExpl,v_i}, ...\}$, from which graphs can be built (i.e. one set per method used; the two we proposed could be completed by others). The initial search space for the recommender can then be defined for a user v_i as: $SG_{v_i,0} = \bigcup_{fac_i \in F_{v_i,0}} SG_{fac_i}$.

During the visit, the weighting process is constantly running and weights can be refined according to the visitor's actions. This way, users' personal facets constantly evolve: preferred ones are reinforced, sporadic ones can disappear or appear. User profiles evolve accordingly. At time t, we have $G_{v_i,t}$, the updated user profile which generates an updated set of personal facets, $F_{v_i,t}$ and the corresponding search space is $SG_{v_i,t} = \bigcup_{fac_i \in F_{v_i,t}} SG_{fac_i}$.

D. Recommendation approach

The recommendation process using this approach consists in: *(1)* identifying the facets a visitor likes, i.e. personal facets; *(2)* building the graphs corresponding to these facets and combining them into a new graph; *(3)* analysing this graph and building recommendations out of it.

We define the recommendation function for a visitor v_i at time t as follows: $Rec(SG_{vi,t}) = \Phi(Rec_1(SG_{fac_1}), ..., Rec_n(SG_{fac_n})), fac_i \in F_{vi,t}$. It is the result of a fusion Φ of different facet-centric functions Rec_i implementing the algorithm that fits the best their respective facet. Indeed, each facet may not be analysed in the same way. A direct neighbour lookup might be enough in one case while more complicated distance algorithm might be required for others. Facets that support an objective of the system, e.g., respecting user interests, must also be differentiated from those that support a constraint, e.g., the suggested next step must be near the user's current location.

III. Next Steps

We are currently implementing a first system based on this approach, which will be tested on field soon. The next steps to refine the approach presented here will be mainly to choose the suitable weight propagation functions and to find the suitable algorithms to automate the identification of facets from users profiles. A lot of other issues need also to be investigated. For example the visitor will most certainly drift from the suggested path and interact with other POIs different from the recommended ones. How to detect and integrate these deviations is a crucial task, especially since we want our recommender to offer a consistent narrative with the succession of proposed POIs.

Acknowledgment

This project has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement No 693150. www.crosscult.eu.

References

[1] P. Ristoski, M. Schuhmacher, and H. Paulheim, "Using graph metrics for linked open data enabled recommender systems," *Lecture notes in business information processing*, vol. 239, pp. 30–41, 2015.

[2] C. Musto, P. Lops, P. Basile, M. de Gemmis, and G. Semeraro, "Semantics-aware graph-based recommender systems exploiting linked open data," in *Proc. of the 2016 Conf. on User Modeling Adaptation and Personalization*, ser. UMAP '16. NY, USA: ACM, 2016, pp. 229–237.

[3] The ICOM/CIDOC Documentation Standards Group. (2015) Definition of the cidoc conceptual reference model v6.2. [Online]. Available: http://www.cidoc-crm.org/sites/default/files/cidoc_crm_version_6.2.pdf

[4] A. Passant, "Dbrec: Music recommendations using dbpedia," in *Proc. of the 9th Int. Semantic Web Conf. on The Semantic Web - Vol. Part II*, ser. ISWC'10. Berlin, Heidelberg: Springer-Verlag, 2010, pp. 209–224.

[5] R. Albertoni and M. De Martino, *Semantic Similarity of Ontology Instances Tailored on the Application Context*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 1020–1038.