International Congress on Environmental Modelling and Software

9th International Congress on Environmental Modelling and Software - Ft. Collins, Colorado, USA - June 2018

Jun 27th, 3:40 PM - 5:00 PM

A Hybrid Recommender System for Industrial Symbiotic Networks

Anna Gatzioura
gatzioura@cs.upc.edu

Miquel Sànchez-Marrè
miquel@cs.upc.edu

Karina Gibert
karina.gibert@upc.edu

iEMSs International Congress on Environmental Modelling and Software

9th International Congress on Environmental Modelling and Software
Fort Collins, Colorado, USA, Mazdak Arabi, Olaf David, Jack Carlson, Daniel P. Ames (Eds.)
https://scholarsarchive.byu.edu/iemssconference/2018/

A Hybrid Recommender System for Industrial Symbiotic Networks

Anna Gatzioura[a,b], Miquel Sànchez-Marrè[a,b], Karina Gibert[a,c]

[a]Knowledge Engineering & Machine Learning Group (KEMLG-UPC)
Intelligent Data Science and Artificial Intelligent Research Centre (IDEAI-UPC)
[b]Dept. of Computer Science
[c]Dept. of Statistics and Operations Research
*Universitat Politècnica de Catalunya · Barcelona**Tech** (UPC)*
Barcelona, Catalonia
(gatzioura@cs.upc.edu, miquel@cs.upc.edu, karina.gibert@upc.edu)

Abstract: Various solutions enabling the realization of synergies in Industrial Symbiotic Networks have been proposed. However, incorporating intelligence into the platforms that these networks use, supporting the involved actors to automatically find possible candidates able to cover their needs, is still of high importance. Usually, the actors participating in these networks act based on previously predefined patterns, without taking into account all the possible alternatives, usually due to the difficulty of finding and properly evaluating them. Therefore, the recommendation of new matches that the users were not aware of is a big challenge, as companies many times are not willing to change their established workflows if they are not sure about the outcome. Thus, the ability of a platform to properly identify symbiotic alternatives that could provide both economic and environmental benefits for the companies, and the sector as a whole, is of high importance and delivering such recommendations is crucial. In this work, we propose a hybrid recommender system aiming to support users in properly filtering and identifying the symbiotic relationships that may provide them an improved performance. Several criteria are taken into account in order to generate, each time, the list of the most suitable solutions for the current user, at a given moment. In addition, the implemented system uses a graph-based similarity model in order to identify resource similarities while performing a hybrid case-based recommendation in order to find the optimal solutions according to more features than just the resources' similarities.

***Keywords**:* Hybrid Recommender Systems; Industrial Symbiotic Networks; Case-Based Reasoning; Waste Optimization.

1 INTRODUCTION

In the majority of current application domains where users come across a lot of item alternatives and information about them, that have to discover, process and use, *Intelligent Decision Support Systems (IDSS)* and especially *Recommender Systems (RS)* have been identified by Ricci et al. (2011) among the most promising techniques able to support users in their decision making processes.
Industrial Symbiotic Networks (ISN) form an example of such domain, as usually the actors participating in these networks act based on predefined patterns due to the difficulty of properly identifying all possible symbiotic alternatives. *Industrial symbiosis (IS)* has been defined by Chertow (2007) as "engaging traditionally separate industries in a collective approach to competitive advantage involving physical exchange of materials, energy, water, and by-products". Furthermore, ISN have proven as successful in waste treatment and re-use as well as in creating new business opportunities. Therefore, the recommendation of effective possible *synergies*, being "matches between companies interested in

providing and using their waste", that the users weren't aware of, would lead to both economic and environmental benefits for the companies and the sector as a whole. Therefore, the ability of a platform to deliver such recommendations is crucial, as companies many times are not willing to change their established workflows if there are not sure about the outcome.

In this work we describe a hybrid RS (Bruke, 2002) designed and developed within the SHAREBOX project in order to support users in identifying possible symbiotic relationships of which they were not aware and may provide them improved performance.
The SHAREBOX project aims to develop and bring to market a secure platform for the flexible management of shared process resources with intelligent decision support tools. To provide plant operations and production managers with the robust and reliable real-time needed to optimise symbiotic connections (plant, energy, water, residues and recycled materials) with other companies in a symbiotic ecosystem. More details can be found at SHAREBOX project website (http://sharebox-project.eu).

The implemented recommender uses a graph-based similarity model in order to identify resource similarities while performing a hybrid case-based recommendation in order to find the optimal solutions according to more than the resource similarities. Various criteria, like the user request, the current situation of the system, as well as proper resource characteristics are taken into account in order to generate the list of the most suitable solutions for a user at a given moment. Although it has been implemented for a specific platform, this system could be easily extended in order to be used as a complementary tool in other Industrial Symbiosis IDSS.

2 BACKGROUND

2.1 Recommender Systems (RS)

RS are software tools and techniques for information retrieval and filtering, used to provide meaningful suggestions for items to be used by a user. The term "*item*" refers to the type of entity being recommended by each recommender (ex: products, resources etc.), while as "*user*" we refer to the entity, (ex. a person or another system), that through a specific interface interacts with the system in order to get recommendations. The central development idea behind a RS is that a user is trying to find the items that are most possible to maximize his/her utility under given circumstances. Therefore, the problem to be solved by a RS, has been defined by Adomavicious and Tuzhilin (2005) as, given a scope and a user to identify those items that seem most capable of maximizing this utility.
Furthermore, as items we treat the resources needed or generated in the form of *waste* being the result of an industrial process, while a user of the ISN platform would be a registered *company*, that accesses the platform to find a specific resource, a waste that would need to use or to distribute. In addition, apart from looking for a waste with specific characteristics, there are more attributes related to the company and the different situations that affect the formalization of the utility function in this problem (Deshpande and Karypis, 2003; Melville and Sindhwani, 2010; Ricci *et al*, 2011).

2.2 Case-Based Reasoning (CBR)

Case-Based Reasoning (CBR) (Richter and Weber, 2013; Kolodner, 1993) is a problem solving paradigm that uses old experiences in order to solve new problems. It is based on the following sentence, also known as the CBR assumption, *"Similar problems have similar solutions",* therefore new problems are solved by properly adapting successful past solutions to a new situation. The CBR problem solving methodology, as described by López de Mántaras (2001) and Leake (1996), is closely related to the human way of thinking, when facing new problems. Therefore, it seems as a more natural method for building intelligent reasoning systems that can be applied to many domains, as it does not require an explicit domain model. CBR reasoners derive their reasoning from complete past cases, rather than decomposing them into rules. Therefore, they have the ability to adapt and improve their problem solving performance over time.

The CBR solving and learning process can be summarized as the *CBR cycle* or "*the four REs*" meaning:

- *Retrieve*: the most relevant cases from the case memory.

- *Reuse*: the information provided by those in order to solve the new problem.
- *Revise*: the solution obtained
- *Retain*: the parts that are likely to be used for future purposes.

2.3 Case-Based Recommender Systems

Case-Based RS follow the CBR reasoning process and rely on its core concepts of similarity and retrieval. A CBR recommender uses the user query as a problem specification while the set of existing items, with their descriptions, form the cases in the case base that based on their similarity to the user's request, defined in the same space, the space of items' characteristics, may be retrieved and recommended. Depending on the type and the characteristics of the items to be recommended, different local similarity measures are used and aggregated into a different global similarity function. This function is used for the final item selection and has been found by Lorenzi and Ricci (2005) to highly affect the outcome of this process. Case-based RS, as presented by Bridge et al. (2006) implement a type of content-based recommendation that relies on a structured representation of items that allows them to make judgements about product similarities and semantic relations and based on those to provide recommendations, more than simply using the ratings assigned to products by various users. They have mainly emerged as an alternative to collaborative RS intending to overcome the shortcomings that those systems come with.

3 PROPOSED SOLUTION

3.1 Resolution Workflow

The scope of the designed system can be described as follows:

"*Given a list of companies, with their waste offers and/or demands, each being a case c described with its id, sector, location, the waste type, quantity, frequency etc.; and a query q that consists of, at least, the company Id its sector, location, waste Id, quantity, availability dates and the definition of whether it is a needed or offered waste, to generate the list of top–N suggested synergies, those matching in a locally optimal way the given query.*"

Therefore, supposing that each case is described in terms of *Quantity of Resource r_i needed at time t_i* and/or *Quantity of Waste w_j generated at time t_j*, by a specific company, along with some geographical, transportation-related and cost data, in order to generate the list of the top–N matching candidates, the resolution workflow can be divided into the following basic steps that refer mainly to the three first REs of the CBR cycle:

1. A user inserts a request about a specific waste, in terms of related key words or by its European Waste Code (EWC codes) of the desired or offered resource and then the quantity, and availability frequency or a specific date.

2. The recommender retrieves the initial candidates' set that could address the user request, in terms of resource properties. More specific, for a given user query apart from the EWCs that are associated with the given keywords also the most similar to the requested resource code, ones, are retrieved. Under these codes items with same properties or closely similar can be found, although being classified under different codes. Then an initial filtering is performed, based on the request *want/have* parameter of the user request. For instance, if the query is about a *wanted* waste only the records of "have" waste(s) will be evaluated and vice versa.

3. Following, the resource availability is consolidated. As the requested date is a strict constraint, among the candidates' list, only those that may ensure that the waste will be delivered/received on time, will be further evaluated.

4. Finally, the ranked list with the closest matchings is generated using a distance function comparing the request and the candidate resources based on their locations, quantities and their similarity degree. More specific, the total semantic distance among the query *q* and a candidate case *c* is of the following general form:

$$d(q,c) = f\left(d(loc_q, loc_c), d(q_q, q_c), d(w_q, w_c)\right) \qquad (1)$$

Where loc_i is the location of resource *i*; q_i is the quantity of resource *i*; w_i is the waste *i*, and $d(loc_q, loc_c)$ is the geographical distance between the two locations, $d(q_q, q_c)$ compares the requested and the offered quantities, as in some cases a combination of more than one provider may be acceptable while $d(w_q, w_c)$ is used to calculate the semantic distance of the requested and the candidate wastes, given that not only exact, but also closely similar, resources may be used.

3.2 Resource Model

In the majority of the systems, based on the EWC(s) submitted by the user, only the exact matches (if any) would be presented as possible solutions. Thus, an important functionality of the designed system, able to deliver an added value is its ability to identify alternative resource candidates that may be further evaluated based on their specific characteristics.

As the EWC categorization is based on the processes and the industrial sectors from which waste comes, and not on its recourse properties, waste with identical properties may be classified under different codes. Therefore, in order to identify additional item candidates based on their similarity, a graph model, mining the keywords of an existing thesaurus related with EWCs, and describing EWCs as distributions over keywords, like in figure 1, was built.

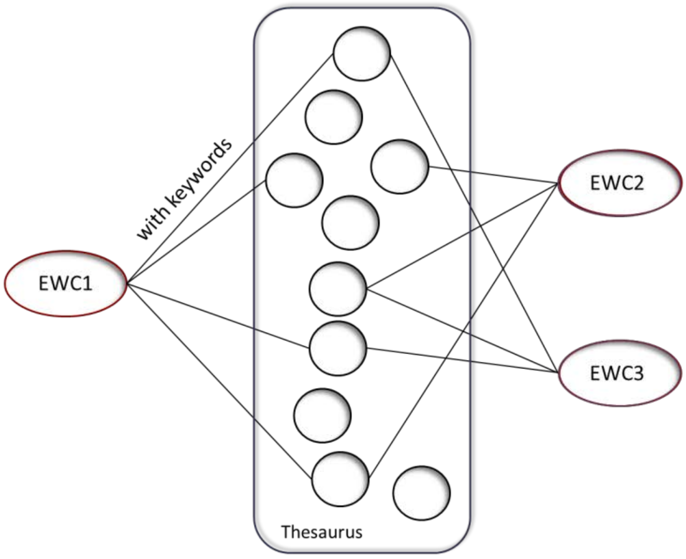

Figure 1: EWC graph-based similarity model.

Then, for two EWCs, e_a and e_b, with $n(a \cap b)$ tags in common, while $n(a \backslash b)$ and $n(b \backslash a)$ being the number of those associated only with e_a and e_b, respectively, the similarity is calculated based on the density of their common descriptions as proposed by (Gatzioura and Sànchez-Marrè, 2017), as:

$$sim(e_a, e_b) = 1 - log_2\left(1 + \frac{n(a \backslash b) + n(b \backslash a)}{n(a \cap b) + n(a \backslash b) + n(b \backslash a)}\right) \qquad (2)$$

where $sim(e_a, e_b)$ is the similarity degree between two EWCs (e_a and e_b)

3.3 Preliminary Evaluation

The initial tests have revealed that the proposed similarity model is able to identify EWCs having the same tag descriptions as well as identifying closely similar but not equal codes, even if these come from different industrial processes. In addition, these similarity calculations may be done during an offline process and being updated only when resources with new codes are added to the system or some descriptions change.

In addition, as the keywords used for tagging the EWCs come from the Thesaurus, a proper dictionary, these permit us model some dimensions of the traded resources while avoiding the fuzziness of user generated tag clouds. Therefore, resources can be treated as distributions over keywords based on which their degree of similarity is calculated and recommendations can be made. This modelling permits us to identify the clusters of item codes with common descriptions that could be automatically recommended to serve the same request, as well as to identify groups of similar items, sharing same part of their descriptions, that could possibly be used, if a user decides so.

4 CONCLUSIONS AND FUTURE WORK

In this document, the hybrid case-based recommender system developed for needs of the SHAREBOX platform has been presented along with its basic functionalities and the resolution process followed. The implemented recommender system is in charge of generating a list of the top-N synergy candidates according to a given request, for a needed or an available resource with specific characteristics. As a key functionality for such a system is the ability to identify closely similar resources, being also able to overcome user and system related constraints, on a given moment, a graph based model has been designed to capture those based on the similarities of the EWCs related with the resources. Among the current research issues is the extension of this model in order to describe also resources not related with an EWC.

ACKNOWLEDGMENTS

The authors acknowledge the funding support of this work by the European Union project SHAREBOX (Secure Management Platform for Shared Process Resources), with code H2020-SPIRE-2015-680843, and the support of the Research Consolidated Groups/Centres Grant (2017 SGR 574) from the Catalan Agency of University and Research Grants Management (AGAUR), from Catalan Government.

REFERENCES

Adomavicius, G., Tuzhilin, A., 2005. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17:734–749.

Bridge, D., GöKer, M.H., McGinty, L., Smyth, B., 2006. Case-based recommender systems. *The Knowledge Engineering Review* 20, 315.

Burke, R., 2002. Hybrid Recommender Systems: Survey and Experiments, in: *User Modeling and User-Adapted Interaction*, vol.12, issue 4, pp.331-370.

Chertow, MR., 2007. Uncovering Industrial Symbiosis. *Journal of Industrial Ecology* 11(1):11–30.

Deshpande, M., Karypis, G., 2003. Item-based top-n recommendation algorithms. DTIC Document.

Kolodner, J., 1993. Case-Based Reasoning. San Mateo, CA: Morgan Kaufmann.

López de Mántaras, R., 2001. Case-Based Reasoning, in: *Machine Learning and its Application*, vol. 2049, pp.127-145.

Gatzioura, A, Sànchez-Marrè, M., 2017. A Case-Based Reasoning Framework for Music Playlist Recommendations. Proceedings of *4th IEEE International Conference on Control, Decision and Information Technologies (CoDIT'17).*

Leake, D.B., 1996. CBR in context: The present and future. Case-based reasoning: experiences, lessons, and future directions 3–30.

Lorenzi, F., Ricci, F., 2005. Case-based recommender systems: A unifying view, in: *Intelligent Techniques for Web Personalization*. Springer, pp. 89–113.

Melville, P., Sindhwani, V., 2010. Recommender Systems, in: *Encyclopaedia of Machine Learning*, ch. 00338

Ricci, F., Rokach, L., Shapira, B., 2011. Introduction to Recommender Systems Handbook, in: Ricci, F., Rokach, L., Shapira, B., Kantor, P.B. (Eds.), *Recommender Systems Handbook*. Springer US, Boston, MA, pp. 1–35.
Richter, M.M., Weber, R.O., 2013. *Case-Based Reasoning: A Text-Book*. Springer-Verlag.