Research Journal of Applied Sciences, Engineering and Technology 11(7): 674-684, 2015
DOI: 10.19026/rjaset.11.2030
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: March 04, 2013 Accepted: July 14, 2015 Published: November 05, 2015

Research Article

Extending Recommender System by Incorporating Semantic-social Information

[1]Khaled Sellami, [1]Rabah Kassa and [2]Pierre F. Tiako
[1]LMA Laboratory, A/Mira University of Bejaia, Algeria
[2]Center for IT Research & Development CITRD, Langston University, Tiako Corporation, Oklahoma, USA

Abstract: Recommender systems in e-commerce applications have become business relevant in filtering the vast range of information available in web shop (and the internet) to present useful recommendation to user. In this study we combine social network analysis and semantic user profile to provide a new semantic-social recommendation, featuring a two-stage process that relies on a simple formalization of semantic user preferences that contains the user's main interests and heuristically explores the social graph. Given a recommendation request concerning a product, the semantic-social recommendation algorithm compares the user preferences, which are found in the exploration path, with the product preferences by referencing them to domain ontology. Experiments on real-world data from Amazon, examine the quality of our recommendation method as well as the efficiency of our recommendation algorithms.

Keywords: Recommender systems, semantic web, social network, taxonomy, user preferences

INTRODUCTION

The exponential growth of the vast range of information poses challenges and presents new opportunities for recommender system research. Nowadays, recommender systems are widely used in several important domains and in some cases a failure recommendation could cause great losses of time, effort and money.

Recommender systems has three main categories (Melville and Sindhwani, 2010): Content-based (Pazzani and Billsus, 2007) where the users are recommended with items that are similar to those that they liked in the past, collaborative-filtering or social recommendation (Das *et al*., 2007) where the recommendation depends on the user's neighbors' opinions and not on the item itself and hybrid recommendation that combines the content-based and social based recommendation methods (Burke, 2007).

In this study we present a solution to surpass the defects of failure recommendation by presenting semantic-social recommendation algorithm, in which we suppose a set of user and a set of products such as a users are connected through a social network and users and items are described via taxonomy. In this setting, given a product we use a heuristic based search algorithm to search the social network in order to compute a relevant set of users to whom the product can be recommended. This algorithm is concerned with two important aspects, the social aspect by using social network analysis measures and the semantic aspect by using the semantic similarity measures.

LITERATURE REVIEW

The approach described in this study relies on a combination of social network analysis and semantic web for semantic social recommendation. In this section, we explore related works in recommendation systems using these techniques. We also highlight the originality of the approach we propose with respect to the state of the art.

Recommendation systems: The main idea of collaborative filtering recommender systems is to capture the user's tastes, compute the similarity between users and predict the recommendations. Generally all the collaborative filtering algorithms have the main principals, but they differ in the way of computing the similarity between users.

Resnick *et al*. (1994) Grouplens proposed Newsnet, the article recommender system. This algorithm is one of the earliest CF algorithms. It is a user-based and uses Pearson r correlation coefficient to compute the similarity between users. Later, Grouplens implemented this algorithm on Usenet news (Konstan *et al*., 1997).

Shardanand and Maes (1995) authors introduced a personalized recommender system called Ringo, which

Corresponding Author: Khaled Sellami, LMA Laboratory, A/Mira University of Bejaia, Algeria

recommends music and artists to users. For this system the authors implemented and compared four CF algorithms. These algorithms are: the mean squared differences algorithm; which measures dissimilarity between users, the Pearson r algorithm, the constrained Pearson r algorithm and the item-based CF algorithm. Their results showed that the constrained Pearson algorithm gives the best results.

Herlocker *et al*. (1999) Spearman ranking correlation coefficient as another recommendation measure is proposed. Spearman correlation is the same to Pearson correlation, but instead of handling the ratings the algorithm handles the ranking of the ratings. These results proved that Spearman ranking correlation performs as well as Pearson correlation.

Aggarwal *et al*. (1999) authors proposed an intelligent recommendation algorithm called IRA. This algorithm is a graph based collaborative filtering recommendation algorithm, where users are connected via directed graph. The nodes of this graph represent users while the directed edges of this graph represent the horting and predictability relation between these users; horting and predictability relation is mathematically defined in (Aggarwal *et al*., 1999). The algorithm recommends the item j to the user I by computing the shortest path in its entirely between the user i and group of users. Each user in this group should have common rated items with the user i and should have already rated the item j. In this algorithm the author proposed the breadth first search algorithm to compute the shortest paths between users.

Mirza *et al*. (2003) the authors proposed Movie recommender system. In this system three graphs have been defined, the first graph is the bipartite graph. Its nodes are divided into two sets the people set P and the movie set M and the edges E are created between P and M and represents the ratings and viewing preferences between P and M. The second graph is the collaboration network graph which is a one-mode projection graph between the users; two users will have collaboration connection between them, if they have at least one movie in common. The third graph is the recommender graph which is a sum of the social collaboration graph and the bipartite graph. In order to give the recommendation, shortest path algorithm is applied on the recommender graph.

The limitation of the aforementioned works is the tight coupling with the collaborative filtering recommendation. Even if there are several graph based recommender systems, these recommender systems never employ the social network analysis measures in the recommendation algorithm. For that, we propose to involve the social network analysis measures in the recommendation algorithm. Furthermore, we also propose to involve the user's semantic preferences in this recommendation algorithm, in order to have a semantic-social recommendation algorithm.

Social network: Social Networks are networks in which vertices represent people and edges represent social interactions (such as friendship and co authorship) among these people (Newman, 2010). Social network analysis is the study of social networks by understanding their social entities, the people and their relationships. Actually, social network analysis measures are used to study the structural properties of the social network (Abbasi and Altmann, 2011). The most common social network analysis measures are eigenvector centrality (Aggarwal, 2011), pagerank (Aggarwal, 2011), closeness centrality (Abbasi and Altmann, 2011) and betweenness centrality (Newman, 2010). Also, degree centrality is widely used as it is one of the simplest centrality measures. The vertex's degree equals to the number of connected edges to this vertex (vertex degree) (Abbasi and Altmann, 2011).

Furthermore, due to the recent evolution of social networks, social recommender systems are becoming more common such as:

- Finding the user's best co-workers in a social network (Palau *et al*., 2004).
- Recommending friends, using graph based algorithms such as random walk (Konstas *et al*., 2009).
- Proposing music in a social network of connected artists (Cano *et al*., 2006).
- Tagging based recommender system for recommending photos (Rae *et al*., 2010). Bookmarking uses a personalized tag recommendation system for users of bookmarking sites using text mining similarity measures (Byde *et al*., 2007). Baatarjav *et al*. (2008) Presenting a Facebook group recommender system, by using hierarchical clustering and decision tree techniques. Also, Facebook application has been proposed in Bedrick and Sittig (2008) to find colleagues who can work in similar projects.

Semantic social network: As we have seen, the use of software instead of users in the information filtering has certain weaknesses:

- How to represent information complicates communication between agents and between agents and users.
- Reuse of information represented heterogeneously becomes too complicated. With the arrival of the Semantic Web (Tim *et al*., 2001), these deficiencies are mitigated by the improvement and enrichment of the representation of information through the application of these technologies.

Semantic Social Network is the composition of two types of technologies: semantic web technology (Tim *et al*., 2001) and the Social Networks technology

(Downes, 2005). The first research question about the possibility of having a semantic social network was presented in Kim (2002). Later in 2005 Downes (2005), has proposed new type of internet as a network within a network to reshape the internet that we know, this type is based on merging the semantic web technology and the social network (Downes, 2005). Actually, semantics in social networks can lay behind the semantic presentation of user profile (O'Murchu *et al*., 2004). Or can lay behind the RDF (resource Description Framwork) presentation of the social network structure and SPARQL[1] query to present the social network analysis measures (Erétéo *et al*., 2009).

Sellami *et al*. (2012) authors proposed to give recommendations by combining social network analysis and semantic web for semantic social recommendation. The goal of these algorithms is to recommend a group of ranked authors that are similar to a certain request criteria, in a social network of connected authors. In this system, recommendations are based on the social network analysis and the vectorial presentation of user profile.

Combining social network analysis and semantic web for semantic social recommendation, increases filtering precision. First, we integrate semantic information (semantic preferences aspects to represent customers and products) into the user profile. Second, establish collaboration social network, where nodes represent customers and edges (with weights) represent the similarity between these customers. The goal of our proposal is:

- To guarantee the interoperability of system resources and the homogeneity of the representation of information
- To facilitate performance in social networks and collaborative filtering
- To improve the representation and description of different system elements. In the following, we show the main advantages of our approach comparing to related work and illustrate it with a typical scenario from Amazon library eBooks.

MATERIALS AND METHODS

In this section we present our proposed Semantic-Social recommender system. Our approach combines the semantic information on users and products (using semantic user preferences) with social information using the social network analysis. We first give some important concepts about the semantic information part of the model, then we define important concepts on the social information part of the model before explaining the semantic social recommendation algorithm proposed.

Semantic information: The semantic part of our proposed model relies upon the following three fundamental aspects:

User preferences: Users' preferences are grouped in their user profiles which contain all the possible information about users such as activities and main interests. Generally, user profile can be represented in several forms; vector of weighted keywords, semantic form, conceptual form and else more. However, the literature is rich in studies that are related to this field as described in Gauch *et al*. (2007).

Semantic taxonomy: Representation of knowledge about certain domains (domain taxonomy); the taxonomy is defined as a collection of entities that are organized into a hierarchical structure, 'is-a' hierarchy, to describe certain objects in certain domains. In the literature several authors proposed using taxonomy in recommendation systems (Resnik, 1995; Ziegler *et al*., 2008; Cantador and Castells, 2011). Actually, taxonomy representation of information is a very helpful tool to estimate users' preferences in the case of lack of information about users.

Semantic similarity measure: It is used to compute the closeness between any pair of concepts in the ontology. The existing semantic similarity measures have three main approaches:

- Edge based approach (Resnik, 1995); where the similarity computation depends on the maximum length of the taxonomy D and the shortest path between the concepts $len(a, b)$ see the Eq. (1) :

$$sim(a,b) = (2 \times D) - len\ (a,b) \tag{1}$$

- Node based approach (Resnik, 1995; Lin, 1998): where the similarity is based on the information content of the lowest common ancestors of a pair of concepts as described in the Eq. (2).

$$sim(a,b)\ =\ max(-log\ P(d)) \tag{2}$$

- where, $d\ \in$ the set of the lowest common ancestors of the pair a, b and $-log\ P(d)$ is the negative log likelihood of the probability $P(d)$ of being an instance of d.
- And hybrid approach (Jiang and Conrath, 1997): Which combines the both previous approaches.

Moreover, in our approach we intend to use domain taxonomy which represents all the knowledge about products in the entire system. We also intend to attach semantic taxonomy preferences to each user and each product in the system, as well as use a hybrid semantic measure to compute the similarity between the users and the products. The following definitions are necessary to understand the model:

Definition 1: The Semantic Taxonomy Tree STT is a taxonomy of connected terms (these terms represent certain domain). This taxonomy has tree structure. Its nodes t represent the domain terms and its edges h

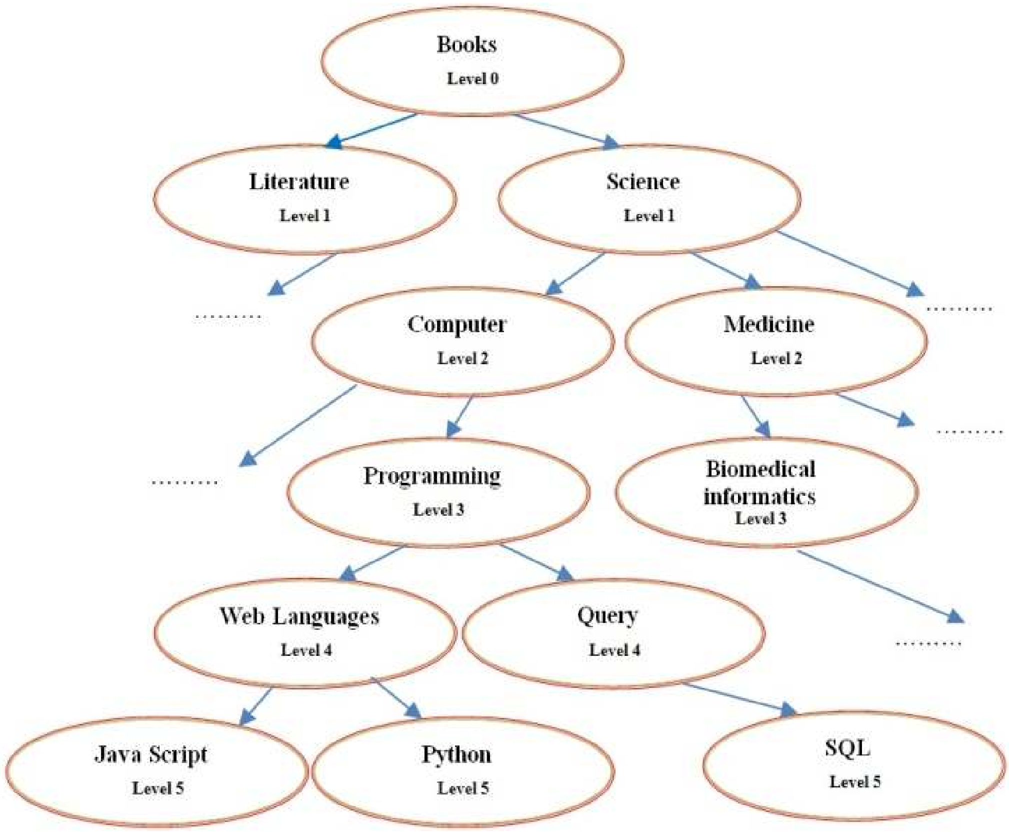

Fig. 1: Books taxonomy example

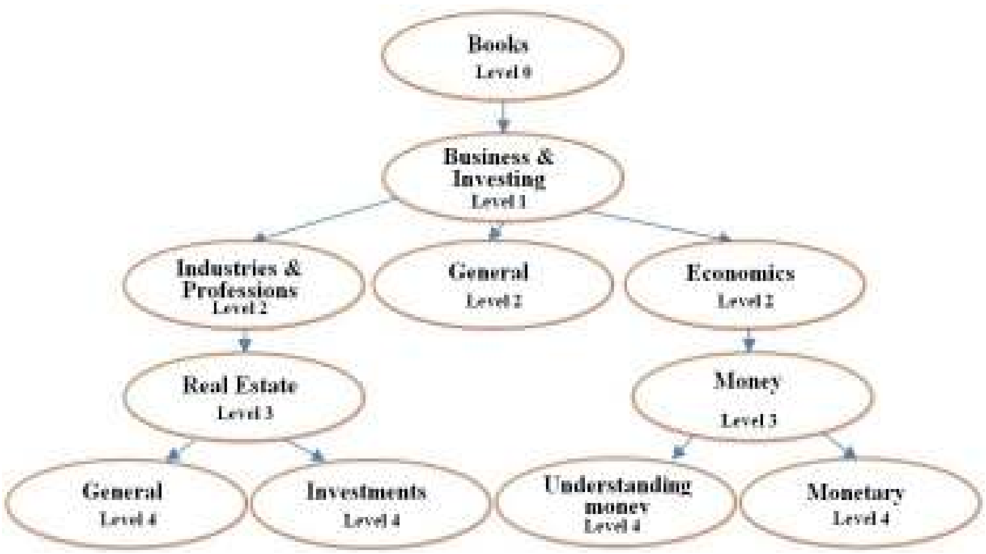

Fig. 2: Product profile

represent the hierarchy between these terms. The hierarchy is described as `is-a` hierarchy. The Semantic Taxonomy tree has n levels; terms in level 0 are the most general terms in the domain, while terms in level $n - 1$ are the most specific terms in the domain. STT is represented via two sets: the terms set and the hierarchy set. The terms set is a set of pairs (term, level) and the hierarchy set is a set of pairs ($term_x$, $term_y$) where the $term_x$ and the $term_y$ are connected via 'is-a' hierarchy and they have the Parent/Child relationships.

Example: Figure 1 is an example of semantic taxonomy. This taxonomy describes books categories and it has $n = 6$ levels. Books category has the nodes of

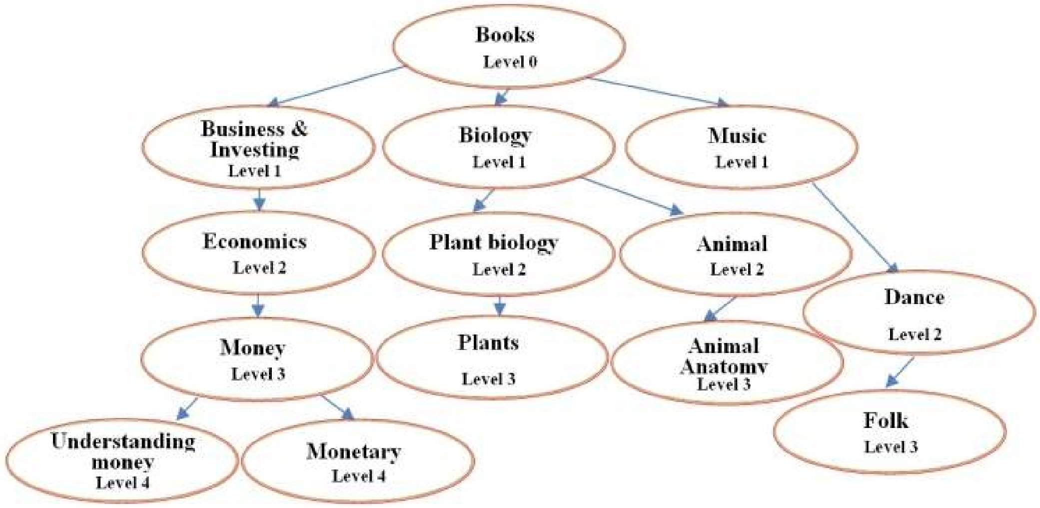

Fig. 3: User profile

this tree and hierarchy between these categories has the edges of this tree. Apparently the terms set for this taxonomy is T = {(*books*, 0), (*Literature*, 1), (*Science*, 1), (*Computers*, 2), (*Medicine*, 2) ...} and the hierarchy H = {(*books*, *Literature*), (*books*, *Science*), (*Science*, *Computers*), (*Science*, *Medicine*) ...}.

Definition 2: Product preferences tree $PPT(x)$ is a tree of connected terms which describe certain product x, This tree is a subtree of the STT, $PPT(x) \subset STT$ and it is represented via two sets: the terms set $P_t(x)$ which is a set of pairs (term, level) and the hierarchy set $P_h(x)$ which is a set of pairs ($term_i$, $term_j$) where the $term_i$ and the $term_j$ are connected via 'is-a' hierarchy and they have the Parent/Child relationships.

Example: Figure 2 we have an example of product preferences PPT that describe the product from the most general term, e.g., Books to the most specific terms, e.g., investments. The product terms set is: $P_t(x)$ = {(*books*, 0), (*business and investing*, 1), (*economics*, 2), (*money*, 3)…} and the product terms hierarchy set is: $P_h(x)$ = {(*books*, *business and investing*), (*business and investing*, *economics*), (*economics*, *money*)…}.

Definition 3: User preferences tree $\mathrm{UPT}(x)$ is defined as set of terms $U_t(x)$ that describe user preferences and the set of hierarchical relations $U_h(x)$ between these terms. In our model, UPT is built based on the historical information about products the user liked in the past. That means the UPT(x) is the union of the taxonomies of all products the user preferred in the past. $\mathrm{UPT}(x) = \cup_{i=0}^{k} \mathrm{PPT}(y_i)$ where, $i \in [0, k]$, k is the number of the user's preferred products and $\mathrm{PPT}(yx_i)$ is the product's preferences tree of the product x_i the user liked in the past. Such as the PPT(y), the UPT(x) is a subset of the semantic taxonomy tree $\mathrm{UPT}(x) \subset \mathrm{STT}$.

Example: Figure 3 shows example of $UPT(x)$ of x user, apparently the user x liked three books: Business book, mathematics book and music book. Each one of these book has its own PPT. $PPT(y_1)$ = {*Books*, *business and investing*, *economics*, *money*}, $PPT(y_2)$ = {*books*, *Biology*, *Plantbiology*, *plants*}, $PPT(y_3)$ = {*Books*, *Music*, *Dance*, *Folk*}, According to our definition the $UPT(x) = PPT(y_1) \cup PPT(y_2) \cup PPT(y_3)$.

This information is very important for the future recommendation.

After giving the three previous definitions, we present user-product similarity measure. This measure is used to compute the semantic similarity between user profile and product profile. Actually, this measure is one of the important bases in the Semantic-Social recommendation algorithm.

User-product semantic similarity: In our model user-product similarity is used to compute the semantic similarity between a certain product and the whole users in the system during the recommendation process. As mentioned in the literature, there are three main categories of semantic similarity measures: node based measure, edge based measure and hybrid based measure.

In this model we propose a hybrid semantic similarity measure adapted and adopted from the literature (Resnik, 1995; Jiang and Conrath, 1997; Lin, 1998; Zuber and Faltings, 2007), but with some modifications in order to take into account the dataset we have and the definitions we suggest. Our proposed measure takes in consideration the content of the ancestor node and the actual level (depth) of the ancestor node, in this case we attach the entire ancestor (not only the lowest common ancestor) with a weight value that represents the actual level of the ancestor, e.g., ancestor in the level 0 has weight equals to 0 and ancestor in the level 10 has weight equals to 10. For that, we suppose, for a given product Y which has a taxonomy preferences $PPT(Y)$ and terms $PPT(Y_t)$= $\{(Y_{t1}, l_{Y_{t1}}), (Y_{t2}, l_{Y_{t2}}), ...\}$ and for a given user X which has a taxonomy preferences $UPT(X)$ and a set of terms UPT(X_t)= $\{(X_{t1}, l_{X_{t1}}), (X_{t2}, l_{X_{t2}}), ...\}$ we present the following function:

$$f\big((X_t, l_{X_t}) \times (Y_t, l_{Y_t})\big) = \begin{cases} l \ if \ X_t = Y_t \ and \ l_{X_t} = l_{Y_t} \\ 0 \ otherwise \end{cases}$$

where, $(Y_t, l_{Y_t}) \in$PPT(Y) and $(X_t, l_{X_t}) \in$UPT(X) and l is the current level of terms X_t and Y_t. In this case, depending on the previous function the similarity measure is described by the Eq. (3):

$$sim\,(\mathrm{UPT(X), PPT(Y)}) = \sum_{n1,i=0} \Big((X_t, l_{X_t})_i \times n2,j=0(Yt, lYt)j \qquad (3)$$

where, n_1 is the number of the elements in the set UPT(X_t) and n_2 is the number of the elements in the set $PPT(Y_t)$.

Social information: The second part of our model is the "Social Information" part which is relied upon by the collaboration social networks proposed in (Ramasco, 2007). Generally, social networks are defined as networks in which vertices represent people and edges represent social interactions among these people (Newman, 2010). Social network analysis in one hand is used to study the social networks based on the following measures:

- Degree centrality which represents the number of connected edges to a certain node
- Degree distribution
- Average shortest path and clustering coefficient (Cano *et al*., 2006).

On the other hand, bipartite networks are defined as particular type of social networks with the vertices divided into two different sets and the edges connected only vertices from different sets (Shang *et al*., 2008), e.g., the user-product bipartite graph has two different sets of vertices, user set and product set. The graph user-product connection could be created if a particular user likes certain a product. Furthermore collaboration networks are extracted from bipartite networks by using one-mode projection (Ramasco, 2007), e.g., in user-product bipartite network if the users u_1 and u_2 like the same product P then these two users are connected in the user's one-mode projection network. Actually the bipartite networks and the one-mode projection have been used in several recommender systems such as the music recommender system (Cano *et al*., 2006) and movies recommender system (Grujific, 2008).

In this study our approach of collaboration social network is based on user one mode projection. The following definitions support our idea:

Definition 4: User-user similarity: Computes the number of common preferred products between any pair of users, with the consideration to the users' ratings. This similarity is described by equation *4* where, $r_{x,i}$ is the rating of the user u_x on the product p_i, $r_{y,i}$ is the rating of the user uy on the same product pi and n is the number of common products p_i between u_x and u_y:

$$sim'\,(u_x, u_y) = \textstyle\sum_{i=0}^{n} (r_{x,i} + r_{y,i}) p_i \qquad (4)$$

Definition 5: The collaboration social network G(U, S): Is a user projection network, that is extracted from the bipartite user-product network, each node represents user U and each edge S represents the similarity between the users as described in the Eq. (4). Actually, the edge s connects the pair of users u_x and u_y if $sim(u_x, u_y) \geq similarity\ threshold.$

Definition 6: Semantic social network: We suppose that the semantic social network is a social network where a semantic user profile is attached to each one of its nodes. In our case the semantic social network is the same to the collaboration social network $G(U, S)$ as mentioned in the definition 2 with simple modification concerning to the user profile. Actually, a semantic user profile is attached to each one of its nodes (users) of $G(U, S)$. The definition of semantic user profile is given by the definition 3.

Recommendation algorithm: The semantic-social recommendation algorithm recommends product P to a group of users $\in SSN$. Actually, this algorithm depends on three main factors: The semantic information as described in the subsection "semantic information", the social information as described in the subsection "social information" and the graph searching algorithm defined. In this subsection, firstly we define our proposed recommendation request then we detail our proposed algorithm.

Recommendation request "The query": The recommendation query is composed of a product p with a semantic profile $U \in SSN$ described by definition 3. According to the semantic-social recommendation algorithm the request is submitted to the semantic social network $SSN(U, S)\ and$ as a result the algorithm gives the set of recommended users $U = \{u_0, u_1, \dots\dots, u_k\}$ where $u_j \in SSN(U, S)$ and $j \in [0, 1, \dots\dots, k]$.

Procedure Heuristic-Search

```
begin
    Procedure: HeuristicSearch(SSN, GraphNode)
    Label GraphNode as visited
    if sim(GraphNode, PPT) > threshold then
        RecommendedUsersList ← GraphNode
        for each edge ∈ SSN.incidentEdges(GraphNode) do
            if edge e is unexplored then
                NodeNeighbor ← G.opposite(GraphNode,e)
                label e as explored edge
                if vertex NodeNeighbor is unexplored then
                    recursively call
                    HeuristicSearch(SSN,NeighborNode)
                end
            end
        end
    end
    else
        Ignore GraphNode and its connections
    end
    ;
end
```

Fig. 4: Heuristic-search procedure

Algorithm 1: Semantic-Social Recommendation Algorithm

```
Input: Product P with the preferences PPT.
Output: Set of recommended users U = {u1, u2, ..., ui}
begin
    Compute the Degree Centrality of all the nodes in the SSN
    CentralNodesVector ← Top N nodes with highest centrality
    degree;
    for i ← 0 to N do
        if sim(CentralNodeVector[i], PPT) > threshold then
            HeuristicSearch(CentralNodeVector[i])
        end
        else
            Ignore CentralNodeVector[i] and ignore its
            connections
        end
        ;
    end
end
```

Fig. 5: Semantic social recommendation algorithm

Semantic-social recommendation algorithm: The Semantic-Social recommendation algorithm mainly depends on the Depth First Search (DFS) algorithm (Tarjan, 1972) with some modifications regarding to DFS explores the entire graph. However, our proposed algorithm never explores the entire graph, it starts the search from the node (user) with the highest value of degree centrality and it uses a heuristic function which is based on product-user similarity measure:

The node's degree: From the definition of the node's degree (the number of its connected nodes) and form the fact that the nodes are part of the collaboration social network, which has a user-user similarity connections as mentioned in the definition 1. We propose to start the search from the node with the highest degree value. In this case, if the node satisfies the heuristic function, then there is a high possibility that this node's connections satisfy the heuristic function so the algorithm continue the DFS, in the other case if the node does not satisfy the heuristic function we ignore this node and its related connections.

The heuristic function: Is the same to the user-product similarity measure as explained in above. If the similarity is more than a fixed threshold then the algorithm applies the procedure Heuristic-Search, in the other case the algorithm ignores the node and its nodes.

See the algorithm 1 and see the procedure Heuristic-Search (Fig. 4 and 5).

RESULTS AND DISCUSSION

In this section we detail our experiments and results. First, we present the "Amazon dataset" dataset we use and define the metrics used to validate our algorithm. Secondly we compare the semantic-social algorithm with two existing recommendation algorithms; the content based and the collaborative filtering. Lastly we discuss the obtained results.

Dataset: We choose Amazon.com[2] data as a real dataset to test our algorithm. Amazon data comprises 7 million customers and more than 500 thousands products. In our experiments we only use 500 products to build the semantic-social network. This social network has 7,997 nodes and about 4.600.000 edges.

Evaluation metrics: In the literature, several metrics have been proposed to evaluate recommendation algorithms; e.g., mean absolute error, mean squared error, precision, recall, f-measure, roc curves and other metrics. In our case, we propose to use precision/recall measure. Precision and recall are categorized as classification accuracy metrics, they measure the frequency of correct and incorrect decisions that could be given by a recommender system. These two metrics are very common and have been used to evaluate several recommendation algorithms (Sarwar *et al*., 2000).

Precision is defined as the probability that a selected item is relevant, while *recall* is defined as the probability that a relevant item is selected.

Evaluation framework: In order to evaluate the Semantic-Social Recommendation algorithm, we compared it to the content based algorithm and the collaborative filtering algorithm (Das *et al*., 2007; Pazzani and Billsus, 2007; Melville and Sindhwani, 2010) using the same Amazon dataset.

Content based recommendation algorithm: The algorithm recommends items to users when the content of these items is similar to the content of the items the user liked in the past. For that, we used the user-product semantic similarity measure to compute the similarity between the product (item) and the user's preferred products (items).

Collaborative filtering: We implemented item based collaborative filtering algorithm, in this algorithm if a user *u* likes item p_1 and the item p_1 is similar to the item p_2 according to the opinion to other users, then we could recommend the item p_2 to the user *u* (Sarwar *et al*., 2001).

Precision is defined as the probability that a selected item is relevant, while *recall* is defined as the probability that a relevant item is selected. We recall below the definitions of these two measures where *TP* (True Positive) is the number of relevant users who have been recommended, *TN* (True Negative) is the number of relevant users who have not been recommend and *FP* (False Positive) is the number of irrelevant users, who have been recommend:

$$Precision = \frac{TP}{TP+FP}\ ,recall = \frac{TP}{TP+TN}$$

Algorithm is also assessed according to the following criteria:

- The percentage of the graph vertices (users) who have been visited by the recommendation algorithms, in order to find the recommended users.
- The time each algorithm takes, to answer a recommendation query.

Results: We developed our algorithms using java 6, we also used JUNG (Java Universal Network/Graph) (Madahain *et al*., 2005) as a framework for social network analysis. Moreover, we performed our experiments on an Intel(R) Xeon(R) CPU E5520 2.27 GHz with 12 Giga of RAM, using Debian Linux as operating system.

We implemented the content based algorithm and the collaborative filtering algorithm and then we compared it to the Semantic-Social recommendation algorithm. Actually, this comparison is based on three aspects: precision and recall values, recommendation time and the number of recommended users. For the content based algorithm our results show that:

Precision and recall: In our case, for 8000 users, each of the two algorithms has very high precision/recall values see Table 1, Fig. 6 and 7. But sometimes the

Table 1: Precision and recall recommendation comparison

	Semantic social algorithm		Content based algorithm		Collaborative filtering algorithm	
Items	Precision	Recall	Precision	Recall	Precision	Recall
Q1	0.6	0.74	0.7	0.98	0.53	0.58
Q2	0.64	0.63	0.76	0.99	0.61	0.6
Q3	0.68	0.71	0.82	0.99	0.56	0.61
Q4	0.68	0.7	0.86	0.84	0.56	0.6
Q5	0.7	0.98	0.86	0.85	0.56	0.55
Q6	0.76	0.99	0.86	0.99	0.45	0.45
Q7	0.76	0.76	0.88	0.99	0.55	0.5
Q8	0.86	0.85	0.89	0.98	0.65	0.59
Q9	0.86	0.99	0.9	0.99	0.59	0.54
Q10	0.86	0.84	0.92	0.91	0.45	0.45
Q11	0.88	0.99	0.95	0.99	0.33	0.42
Q12	0.89	0.98	0.97	0.94	0.47	0.5
Q13	0.9	0.99	0.97	0.99	0.52	0.49
Q14	0.97	0.94	0.98	0.98	0.6	0.52
Q15	0.98	0.99	0.98	0.99	0.5	0.53

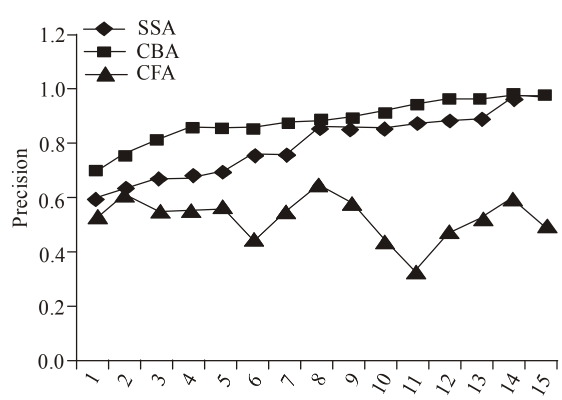

Fig. 6: The precision curves of the three recommendation algorithms Semantic-Social Algorithm (SSA), Content Based Algorithm (CBA), and the Collaborative Filtering Algorithm (CFA)

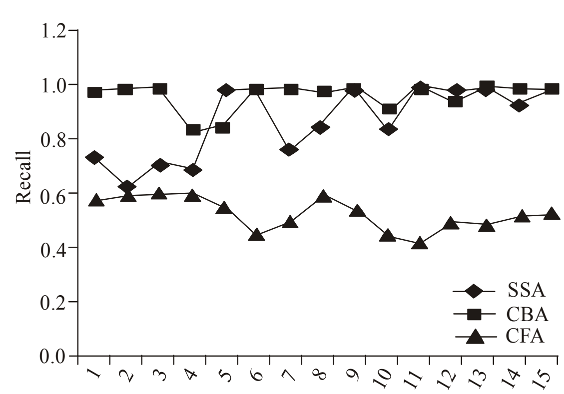

Fig. 7: The recall curves of the three recommendation algorithms Semantic-Social algorithm (SSA), Content Based Algorithm (CBA), and the Collaborative Filtering Algorithm (CFA)

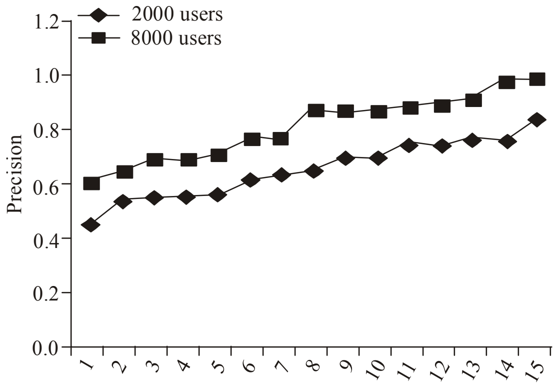

Fig. 8: Evolution of the precision curves for the semantic-social algorithm

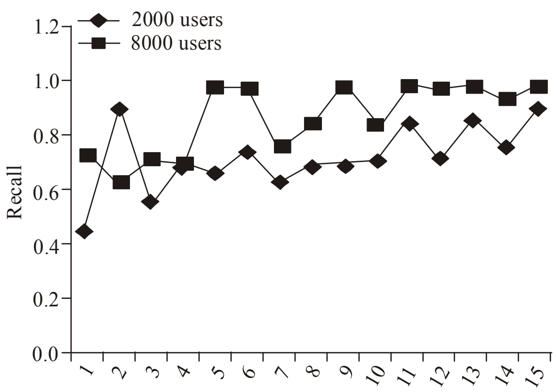

Fig. 9: Evolution of the recall curves for the semantic-social algorithm

Table 2: Evolution of precision and recall recommendation

	2000 users		8000 users	
Items	Precision	Recall	Precision	Recall
Q1	0.45	0.45	0.6	0.74
Q2	0.53	0.91	0.64	0.63
Q3	0.55	0.56	0.68	0.71
Q4	0.55	0.69	0.68	0.7
Q5	0.56	0.67	0.7	0.98
Q6	0.61	0.74	0.76	0.99
Q7	0.63	0.63	0.76	0.76
Q8	0.65	0.69	0.86	0.85
Q9	0.69	0.69	0.86	0.99
Q10	0.69	0.71	0.86	0.84
Q11	0.74	0.85	0.88	0.99
Q12	0.74	0.72	0.89	0.98
Q13	0.76	0.87	0.9	0.99
Q14	0.76	0.76	0.97	0.94
Q15	0.83	0. 91	0.98	0.99

content based algorithm has better values than the semantic-social algorithm. This difference is related to the heuristic nature of the semantic-social algorithm (it explores a part of the data while the content based explores all the data). On the other hand, it is very important to mention that the precision values of the semantic social algorithm becomes better when we increase the number of the connected users in the social network see Table 2, Fig. 8 and 9.

Recommendation time: If we compared the content based algorithm and the semantic-social recommendation algorithm according to the recommendation time, we can conclude that the semantic-social recommendation algorithm achieves the same recommendation in a better time than the content based recommendation algorithm. Without considering the time of building the collaborative social networks, the experiments show that to achieve one recommendation it takes 1 m 27.77 s for the content based algorithm and 5 m 44.955 s for the Semantic-Social algorithm (it takes more than 3 m 44.955 to build the semantic social network).

Number of explored users: As the semantic-social algorithm has a heuristic nature it compares the product semantics with a part of the users according to their position in the social network and their relations with other users in our experiments the semantic social recommendation algorithm explores (between 75 and 85%) explored users.

For the collaborative filtering algorithm our results show that: We compared this algorithm with the SSNA according to the three aspects of precision/recall values, recommendation time and the number of recommended users.

We found that the precision/recall values are very low compared, the recommendation time is better than the semantic-social recommendation 3 m per query, but the algorithm explores all the users in the dataset.

CONCLUSION

In this study, we provided an algorithm that relies on the combination of the social network analysis measures and the semantic user preferences in order to improve the recommendation process. The main feature of the proposed approach is its construction as a two-stage process that relies on a simple formalization of semantic user preferences that contains the user's main interests and heuristically exploration of the social graph. Also, the proposed algorithm compares the user preferences, which are found in the exploration path, with the product preferences by referencing them to domain ontology.

With the aim of achieving representative results, the experiments have been carried out on real dataset "Amazon dataset". The results confirm the quality of our recommendation method as well as the efficiency of our recommendation algorithms.

The semantic-social recommendation algorithm provides a better precision/recall than the collaborative filtering algorithm and the same aforementioned measures with the content based recommendation algorithm. The computation time is better in the semantic-social recommendation algorithm (if the time of building or uploading the social network is not considered). As a heuristic nature the proposed approach algorithm explores between 75% and 85% explored users.

REFERENCES

Abbasi, A. and J. Altmann, 2011. On the correlation between research performance and social network analysis measures applied to research collaboration networks. Proceeding of Hawai International Conference on System Science (HICSS-43). Hawai, January 4-7.

Aggarwal, C.C., J.L. Wolf, K. Wu and P.S. Yu, 1999. Horting hatches an egg: A new graph-theoretic approach to collaborative filtering. Proceeding of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD ’99), pp: 201-212.

Aggarwal, C.C., 2011. Social Network Data Analytics. 1st Edn., Springer, New York, pp: 502.

Baatarjav, E.A., S. Phithakkitnukoon and R. Dantu, 2008. Group recommendation system for facebook. In: Meersman, R., Z. Tari and P. Herrero (Eds.), OTM 2008 Workshop. LNCS 5333, Springer-Verlag, Berlin, Heidelberg, pp: 221-219.

Bedrick, S.D. and D.F. Sittig, 2008. A scientific collaboration tool built on the facebook platform. Proceeding of Annual Symposium proceedings/AMIA Symposium, pp: 41-45.

Burke, R., 2007. Hybrid web recommender systems. In: Brusilovsky, P., A. Kobsa and W. Nejdl (Eds.), the Adaptive Web. LNCS 4321, Springer-Verlag, Berlin, Heidelberg, pp: 377-408.

Byde, A., H. Wan and S. Cayzer, 2007. Personalized tag recommendations via tagging and content-based similarity metrics. Proceeding of the International Conference on Weblogs and Social Media, (March).

Cano, P., O. Celma, M. Koppenberger and M.J. Buldu, 2006. Topology of music recommendation networks. Chaos Interdis. J. Nonlinear Sci., 16(1): 013107.

Cantador, I. and P. Castells, 2011. Extracting multilayered communities of interest from semantic user profiles: Application to group modeling and hybrid recommendations. Comput. Hum. Behav., 27: 1321-1336.

Das, A., M. Datar, A. Garg and S. Rajaram, 2007. GoogleNews personalization: Scalable online collaborative filtering. Proceeding of the 16th International Conference on World Wide Web (WWW’07), pp: 271-280.

Downes, S., 2005. The Semantic Social Network. Published in The Learning Organization: An International Journal, Emerald Group Publishing Ltd., 12: 411-417.

Erétéo, G., F.L. Gandon, O. Corby and M. Buffa, 2009. Semantic social network analysis. CoRR, abs/0904.3701.

Gauch, S., M. Speretta, A. Chandramouli and A. Micarelli, 2007. User Profiles for Personalized Information Access. In: Brusilovsky, P., A. Kobsa and W. Nejdl, (Eds.), the Adaptive Web. LNCS 4321, Springer, Berlin/Heidelberg, pp: 54-89.

Grujific, J., 2008. Movies recommendation networks as bipartite graphs. Proceeding of the 8th International Conference on Computational Science (ICCS ’08), pp: 576-583.

Herlocker, J.L., J.A. Konstan, A. Borchers and J. Riedl, 1999. An algorithmic framework for performing collaborative filtering. Proceeding of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR ’99), pp: 230-237.

Jiang, J.J. and D.W. Conrath, 1997. Semantic similarity based on corpus statistics and lexical taxonomy. Proceeding of the International Conference Research on Computational Linguistics (ROCLING X), pp: 9008.

Kim, H.M., 2002. Ontologies for the semantic web: Can social network analysis be used to develop them? Proceeding of the Conference on Computational Analysis of Social and Organizational Systems (CASOS), 21(3).

Konstan, J.A., B.N. Miller, D. Maltz, J.L. Herlocker, L.R. Gordon and J. Riedl, 1997. Grouplens: Applying collaborative filtering to usenet news. Commun. ACM, 40(3): 77-87.

Konstas, I., V. Stathopoulos and J.M. Jose, 2009. On social networks and collaborative recommendation. Proceeding of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR ’09), pp: 195-202.

Lin, D., 1998. An information-theoretic definition of similarity. In: Shavlik, J.W. and J.W. Shavlik (Eds.), ICML 1998: 15th International Conference on Machine Learning. Madison, Wisconsin, Morgan Kaufmann, San Mateo, pp: 296-304.

Madahain, J., D. Fisher, P. Smyth, S. White and Y. Boey, 2005. Analysis and visualization of network data using JUNG. J. Stat. Softw., 10: 1-35.

Melville, P. and V. Sindhwani, 2010. Recommender Systems. In: Sammut, C. and G.I. Webb (Eds.), Encyclopedia of Machine Learning. Springer, US, pp: 829-838.

Mirza, B.J., B.J. Keller and N. Ramakrishnan, 2003. Studying recommendation algorithms by graph analysis. J. Intell. Inf. Syst., 20(2):131-160.

Newman, M.E.J., 2010. Networks: An Introduction. Oxford University Press, Oxford, New York.

O’Murchu, I., J.G. Breslin, Decker, S. Neumann and M. Sneachta, 2004. The PUII: The semantic social network portal. Proceeding of the 3rd International Semantic Web Conference (ISWC, 2004). Hiroshima, Japan, November 9.

Palau, J., M. Montaner, B. López and J.L. De La Rosa, 2004. Collaboration analysis in recommender systems using social networks. In: Klusch, M. *et al*. (Eds.): CIA, 2004. LNCS 3191, Springer-Verlag, Berlin, Heidelberg, pp: 137-151.

Pazzani, M.J. and D. Billsus, 2007. Content-based recommendation systems. In: Brusilovsky, P., A. Kobsa and W. Nejdl (Eds.), the Adaptive Web. LNCS 4321, Springer-Verlag, Berlin, Heidelberg, pp: 325-341.

Rae, A., B. Sigurbjörnsson and R. van Zwol, 2010. Improving tag recommendation using social networks. Proceeding of Adaptivity, Personalization and Fusion of Heterogeneous Information (RIAO '10), pp: 92-99.

Ramasco, J.J., 2007. Social inertia and diversity in collaboration networks. Eur. Phys. J-Spec. Top., 143: 47-50.

Resnick, P., N. Iacovou, M. Suchak, P. Bergstrom and J. Riedl, 1994. Grouplens: An open architecture for collaborative filtering of netnews. Proceedings of the 1994 ACM Conference on Computer, Supported Cooperative Work (CSCW '94), pp: 175-186.

Resnik, P., 1995. Using information content to evaluate semantic similarity in a taxonomy. Proceeding of the14th International Joint Conference on Artificial intelligence. Morgan Kaufmann Publishers Inc., 1: 448-453.

Sarwar, B., G. Karypis, J. Konstan and J. Riedl, 2000. Analysis of recommendation algorithms for e-commerce. Proceeding of the 2nd ACM Conference on Electronic Commerce (EC '00), pp: 158-167.

Sarwar, B., G. Karypis, J. Konstan and J. Reidl, 2001. Item-based collaborative filtering recommendation algorithms. Proceeding of the 10th International Conference on World Wide Web (WWW '01), pp: 285-295.

Sellami, K., M. Ahmed-Nacer and P. Tiako, 2012. From social network to semantic social network in recommender system. IJCSI, 9(4): 1694-0814.

Shang, M.S., Y. Fu and D.B. Chen, 2008. Personal recommendation using weighted bipartite graph projection. Proceeding of the International Conference on Apperceiving Computing and Intelligence Analysis (ICACIA, 2008). Chengdu, China, Dec. 13-15.

Shardanand, U. and P. Maes, 1995. Social information filtering: Algorithms for automating, word of mout. Proceeding of ACM CHI'95 Conference on Human Factors in Computing Systems, 1: 210-217.

Tarjan, R., 1972. Depth-first search and linear graph algorithms. SIAM J. Comput., 1(2): 146-160.

Tim, B.L., H. James and L. Ora, 2001. The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. Scientific American, May 17.

Ziegler, C., N.G. Lausen and J.A. Konstan, 2008. On exploiting classification taxonomies in recommender systems. AI Commun., 21(2-3): 97-125.

Zuber, V.S. and B. Faltings, 2007. OSS: A semantic similarity function based on hierarchical ontologies. Proceeding of the 20th International Joint Conference on Artifical Intelligence (IJCAI, 07), pp: 551-556.

End note:

[1]: Semantic Web, W3C, http://www. w3.org/2001/sw/.

[2]: The complete dataset can be found in the link http://snap.stanford.edu/data/amazon-meta.html.