

Article

Semantic Search Enhanced with Rating Scores

Anna Formica * , Elaheh Pourabbas  and Francesco Taglino 

National Research Council (CNR), Istituto di Analisi dei Sistemi ed Informatica “Antonio Ruberti”,
Via dei Taurini 19, I-00185 Rome, Italy; elaheh.pourabbas@iasi.cnr.it (E.P.); francesco.taglino@iasi.cnr.it (F.T.)

* Correspondence: anna.formica@iasi.cnr.it

Received: 11 March 2020; Accepted: 13 April 2020; Published: 15 April 2020



Abstract: This paper presents *SemSim^e*, a method based on semantic similarity for searching over a set of digital resources previously annotated by means of concepts from a weighted reference ontology. *SemSim^e* is an enhancement of *SemSim* and, with respect to the latter, it uses a frequency approach for weighting the ontology, and refines both the user request and the digital resources with the addition of rating scores. Such scores are *High*, *Medium*, and *Low*, and in the user request indicate the preferences assigned by the user to each of the concepts representing the searching criteria, whereas in the annotation of the digital resources they represent the levels of quality associated with each concept in describing the resources. The *SemSim^e* has been evaluated and the results of the experiment show that it performs better than *SemSim* and an evolution of it, referred to as *SemSim_{RV}*.

Keywords: similarity reasoning; semantic search; reference ontology; semantic annotation

1. Introduction

The most significant improvement within the Semantic Web research area pertains to reasoning and searching abilities. In this perspective, semantic similarity reasoning, which relies on the knowledge coded in a reference ontology [1], is a different technique with respect to the well-known deductive reasoning used in expert systems. In [2], we proposed *SemSim*, a semantic search method based on a *Weighted Reference Ontology (WRO)*. In *SemSim*, both the resources in the search space and the requests of users are represented by means of an *Ontology Feature Vector (OFV)*, which is a set of concepts from the *WRO*. We distinguish the user request, also denoted as *Request Vector (RV)*, from the description of a resource, also referred to as *Annotation Vector*, indicated by *AV*. In the search process, *SemSim* contrasts the *RV* against each *AV*, and the result is a ranking of the resources that exhibit the highest similarity degree with respect to the request defined by the user.

In [2], we analyzed two different approaches in order to weigh the reference ontology, namely the frequency-based and the uniform probabilistic approaches. In the experiment described in that paper, we show that *SemSim* by the *frequency-based* approach outperforms the *SemSim* by the *uniform probabilistic* approach, as well as the most representative similarity methods from the literature.

In this work, we present a new method, referred to as *SemSim^e*. It relies on the frequency-based approach and revises *SemSim* along two directions. According to the first direction, in contrasting the *RV* with the *AV*, *SemSim^e* takes into consideration the cardinality of the set of the concepts (features) in the user request rather than the maximal cardinality of the compared *OFV*. This choice allows us to give more relevance to the features which are requested by the user rather than the extra features contained in the annotation vectors available in the search space. Along the second direction, *SemSim* has been enhanced with the rating scores *High (H)*, *Medium (M)*, and *Low (L)* in the *OFV*, with regard to both the request and the search space resources. Within the request, rating scores denote the preferences given by the user to the concepts of the *WRO* used to specify the query whereas, within the annotation vectors, rating scores represent the levels of quality associated with the concepts when they describe

the resources. Consider an example rooted in the tourism domain, where the user is searching for a vacation package by specifying the following features: *InternationalHotel* (*H*), *LocalTransportation* (*M*), *CulturalActivity* (*H*), and *Entertainment* (*L*). On the basis of the given rating scores, he/she gives a high preference to resorts which are international hotels offering cultural activities, and less priority to the remaining features, in particular to the entertainments. Analogously, a holiday package annotated with *HorseRiding* (*H*), *Museum* (*M*), and *ThaiMeal* (*L*) is characterized by a high quality level with regard to the horse riding service, rather than the facilities in visiting museums or having Thai meals at lunch or dinner. Note that, in [3], a proposal concerning rating scores was given, where the concepts of the *WRO* are weighted according to the *uniform probabilistic* approach [4], rather than the *frequency*-based one. Furthermore, in our approach we assumed that, given a facility (for instance, *HorseRiding*) included in a tourist package, the higher the user's priority about that facility, the higher the expectancy about the quality of the same facility and, therefore, the greater the availability of the user for considering more expensive solutions.

In this paper, we have experimented *SemSim^e* in the domain of tourism and we have compared it to the *SemSim* method defined in [2] and a further evolution of *SemSim*, referred to as *SemSim_{RV}*. Essentially, *SemSim_{RV}* is the original *SemSim* method where, in line with the first direction adopted in *SemSim^e* illustrated above, more priority has been given to the features indicated by the user in his/her request. The results of the experiment show that *SemSim^e* outperforms both these methods.

The organization of the rest of the paper is the following: Section 2 is about the related work; Section 3 shows the *SemSim^e* method and recalls *SemSim*; Section 4 illustrates the experimental results; Section 5 presents the conclusion and future activities.

2. Related Work

In the literature, several proposals regarding semantic similarity reasoning have been defined, as for instance [5–9]. In general, the matchmaking between vectors of concepts is computed on the basis of their intersection as performed in Dice and Jaccard methods [10], without considering neither the information content of the concepts in the ontology nor the hierarchical relationship among them. In [11], the Weighted Sum method has been proposed where the similarity of hierarchically related concepts is evaluated, although by using a fixed value (i.e., 0.5). In [12], the similarity between sets of ontology concepts is computed on the basis of the shortest path in a graph by applying an extended version of the Dijkstra algorithm [13]. Finally, other proposals, such as [14], consider the inverse document frequency (IDF) method, and combine it with the term frequency (TF) approach.

With respect to the mentioned papers, in this work the semantic matchmaking method is performed in accordance with the *information content* approach defined by Lin in [15], which is an evolution of [16]. The Lin's method results in a better correlation with human judgment when compared with other approaches, such as the edge-counting [17–19]. Furthermore, concerning the evaluation of the similarity between vectors of concepts, we borrowed the Hungarian algorithm for solving the *maximum weighted matching* problem in bipartite graphs [20].

In [21], the representation of users' preferences relies on the definition of users' profiles, which are built according to an ontology-based model. The ontology has been exploited in order to establish relationships among user profiles and features. Then, the level of interest of the user is modeled by taking into account both features and historical data. Reference [22] copes with the evaluation of the similarity between users' profiles. In particular, it proposes a collaborative filtering method in order to identify similar users rating in a given set of items, and to analyze the ratings associated with these users. In [23], the traditional feature-based measures for computing similarity are used, and the authors present an ontology-based approach for evaluating similarity between classes where attributes replace features. The underpinning of this approach is that the more attributes two classes have in common, the higher similarity they have.

Finally, it turns out that $SemSim^e$ is a novel approach because, as opposed to the above mentioned proposals, it introduces rating scores for enriching the semantic annotations of both user requests and resources.

3. The Semantic Similarity Method

The $SemSim^e$ method is an extension of $SemSim$ proposed in [2,24], which is based on the information content approach applied to ontologies [15]. According to this approach, concepts of the ontology are associated with weights such that, along the hierarchy, as weights of concepts decrease their information content increase. A *Weighted Reference Ontology* (WRO) is a pair:

$$WRO = \langle Ont, w \rangle$$

and, in particular:

- $Ont = \langle C, H \rangle$, where C is a set of *concepts*, also referred to as *features*, and H is the set of pairs of concepts of C that are related according to the *ISA* hierarchy (specialization hierarchy);
- w is a function, referred to as *weight*, such that, given a concept c in the ontology, $w(c)$ is a number in $[0, \dots, 1]$.

In Figure 1, a WRO related to the tourism domain is shown, whose weights are defined by using the frequency-based approach.

Given a WRO and a concept $c \in C$, the information content $ic(c)$ is defined as [25]:

$$ic(c) = -\log w(c)$$

The *Resource Space* represents all the searchable and available digital resources. To define the semantic content of a given resource, a structure gathering a set of concepts from the WRO is associated with it. This structure will be referred to as *Ontology Feature Vector* (OFV) and is represented as follows:

$$ofv = (c_1, \dots, c_n), \text{ where } c_i \in C, \quad i = 1, \dots, n$$

Similarly, an OFV describes a user request. As mentioned in the Introduction, in the following, *AV* (*Annotation Vector*), and *RV* (*Request Vector*) will denote the semantics of a resource and a user request, respectively.

In order to search for the resources in the Resource Space, the $SemSim$ method has been defined. This method evaluates the semantic similarity between an *AV* and a *RV* by using the $semsim$ function which relies on the $consim$ function. The latter allows the evaluation of the similarity between concepts, say c_i, c_j , as defined below:

$$consim(c_i, c_j) = \frac{2 \times ic(lca(c_i, c_j))}{ic(c_i) + ic(c_j)} \tag{1}$$

where lca (*lowest common ancestor*) is the least abstract concept in the hierarchy subsuming c_i and c_j . Given an instance of *AV*, say av , and an instance of *RV*, say rv , $semsim$ computes $consim$ for each pair of concepts belonging to the Cartesian product of av and rv . Indeed, we aim at identifying the set of pairs of concepts from av and rv that maximizes the sum of $consim$ according to the Hungarian algorithm for the *maximum weighted matching* problem in bipartite graphs [20]. In particular, given:

$$rv = \{c_1^r, \dots, c_m^r\}$$

$$av = \{c_1^a, \dots, c_n^a\}$$

where $\{c_1^r, \dots, c_m^r\} \cup \{c_1^a, \dots, c_n^a\} \subseteq C$, let S be defined as the following Cartesian product:

$$S = rv \times av$$

and let $\mathcal{P}(rv, av)$ be defined as:

$$\mathcal{P}(rv, av) = \{P \subset S : \forall (c_i^r, c_j^a), (c_h^r, c_k^a) \in P, c_i^r \neq c_h^r, c_j^a \neq c_k^a, |P| = \min\{n, m\}\}.$$

Thus, $semsim(rv, av)$ is defined as follows:

$$semsim(rv, av) = \frac{\max_{P \in \mathcal{P}(rv, av)} \{\sum_{(c_i^r, c_j^a) \in P} consim(c_i^r, c_j^a)\}}{\max\{n, m\}} \tag{2}$$

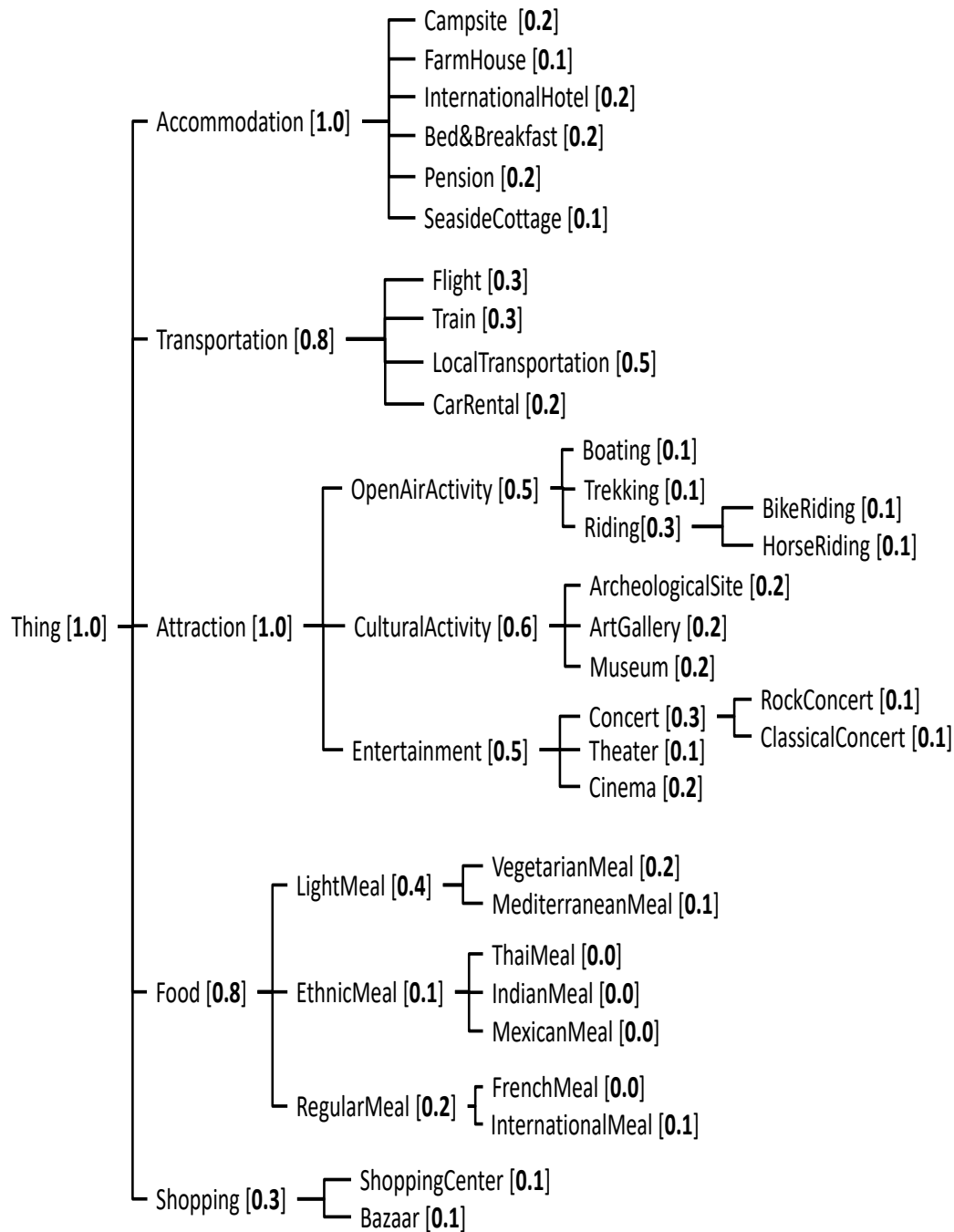


Figure 1. A Weighted Reference Ontology (WRO) in the tourism domain.

SemSim^e

The *SemSim^e* method, which is based on *SemSim*, has been conceived in order to manage a more refined representation of an *OFV*, where concepts are associated with scores. In *SemSim^e*, an *OFV* is indicated as *OFV^e* and, similar to the previous section, *RV^e* and *AV^e* as well.

In our approach, the presence of the scores in the *AV^e* and *RV^e* represents a refined annotation and a refined request vector, respectively. In the case of the *AV^e*, each score, denoted by *s^a*, stands for the quality level of the feature characterizing the resource, while in the case of *RV^e*, denoted by *s^r*, it indicates the priority degree assigned by the user to the requested feature. As mentioned above, these scores are: *H* (*High*), *M* (*Medium*), or *L* (*Low*). Therefore, a given *OFV^e*, say *ofv^e*, is represented as follows:

$$ofv^e = \{(c_1, s_1), \dots, (c_n, s_n)\}$$

where *s_i* is *s_i^a* or *s_i^r*, *c_i* is a *c_i^a* or *c_i^r*, *i* = 1, . . . , *n*.

The *semsim^e* function evaluates the similarity between an *AV^e* and a *RV^e*, represented as *av^e* and a *rv^e*, respectively, by coupling their elements. The value calculated by *semsim^e* is achieved by multiplying the *consim* value by the *score matching* value, namely *sm*, illustrated in Table 1, on the basis of both *s^a* and *s^r* scores. Formally, given:

$$rv^e = \{(c_1^r, s_1^r), \dots, (c_m^r, s_m^r)\}$$

$$av^e = \{(c_1^a, s_1^a), \dots, (c_n^a, s_n^a)\}$$

semsim^e(*rv^e*, *av^e*) is defined as follows:

$$semsim^e(rv^e, av^e) = \frac{\sum_{(c_i^r, c_j^a) \in P_M} consim(c_i^r, c_j^a) * sm(s_i^r, s_j^a)}{m} \tag{3}$$

where $(c_i^r, s_i^r) \in rv^e$, $(c_j^a, s_j^a) \in av^e$, and $P_M \in \mathcal{P}(rv, av)$ is utilized to evaluate the *semsim* value between *rv* and *av*, without considering the rating scores of both *rv^e* and *av^e*.

Table 1. The score matching function *sm*.

		<i>s^r</i>		
		<i>H</i>	<i>M</i>	<i>L</i>
<i>s^a</i>	<i>H</i>	1.0	0.8	0.5
	<i>M</i>	0.8	1.0	0.8
	<i>L</i>	0.5	0.8	1.0

As mentioned above, in the experimentation presented in the next section, we also address an evolution of *SemSim*, referred to as *SemSim_{RV}*, where in contrasting the *RV* with the *AV*, analogously to *SemSim^e*, the number of features defined in the user request are considered rather than the maximal cardinality of the compared *OFV*. Formally, the *semsim_{RV}*(*rv*, *av*) function is defined as the *semsim*(*rv*, *av*) (see formula (2)), where *max{n,m}* is replaced by the cardinality of the request vector *m*, i.e.,

$$semsim_{RV}(rv, av) = \frac{\max_{P \in \mathcal{P}(rv, av)} \{ \sum_{(c_i^r, c_j^a) \in P} consim(c_i^r, c_j^a) \}}{m} \tag{4}$$

Note that *max{n,m}* in Equation (2) has been replaced in both Equations (3) and (4) by *m*, which is the cardinality of the request vector, in order to give relevance to the user request and, in particular, to the features searched by the user.

4. SemSim^e Evaluation

In the experimentation, we asked 15 colleagues at work to identify their privileged tourist packages (request vector, rv) by choosing 4 or (at most) 5 concepts of the ontology illustrated in Figure 1. In particular, assume the user desires to reside in an international hotel, in a location connected by flight and local transportation services, and the opportunity of enjoying cultural activities and entertainments. The related request vector is given below:

$$rv = (InternationalHotel, LocalTransportation, CulturalActivity, Entertainment, Flight)$$

By associating a score among H , M , and L with each concept in the rv , the user can indicate the level of his/her priority for that concept. Accordingly, the user can express his/her priority in the above rv by the following:

$$rv^e = (InternationalHotel (H), LocalTransportation (M), CulturalActivity (H), Entertainment (L), Flight (H))$$

where *InternationalHotel*, *CulturalActivity* and *Flight* received a higher preference with respect to *LocalTransportation* and *Entertainment*, the latter with the lowest score.

Then, we requested our colleagues to examine the 10 packages reported in Table 2, where each concept has been associated with a score among H , M , and L , specifying, similar to *TripAdvisor*, the quality level of the offered services. Our colleagues were also asked to choose 5 packages, out of 10, more similar to their rv , and to assign a score to each chosen one representing the degree of similarity between the package and their rv (Human Judgment—HJ). The request vectors defined by the colleagues are shown in Table 3.

Table 2. Tourist packages.

<i>Packages</i>	<i>AV^e</i>
P1	(Campsite (H), LocalTransportation (H), Concert (M), Flight(M), ArcheologicalSite (H), Food (L))
P2	(FarmHouse (H), Trekking (M), HorseRiding (H), MediterraneanMeal (H))
P3	(InternationalHotel (M), LocalTransportation (M), ArcheologicalSite (H), Flight (H), RegularMeal (M))
P4	(InternationalHotel (H), CarRental (H), Museum (H), Cinema (L), Flight (H), InternationalMeal (L), ArtGallery (H), ShoppingCenter (L))
P5	(Bed&Breakfast (M), LocalTransportation (H), Train (H), RockConcert(H), Cinema (M), BikeRiding (L))
P6	(Bed&Breakfast (H), Bazaar (H), Museum (M), Theater (M), OpenAirActivity (L))
P7	(SeasideCottage (M), Riding (H), Boating (H), Train (M), VegetarianMeal (H))
P8	(Pension (M), CarRental (H), Shopping (L), CulturalActivity (H), LightMeal L)
P9	(Campsite (H), LocalTransportation (M), OpenAirActivity (H), Train (M), VegetarianMeal (H))
P10	(Pension (M), ArtGallery (H), EthnicMeal (H), ClassicalConcert (H), LocalTransportation (L), Entertainment (L))

Table 3. Request vectors.

<i>Users</i>	<i>RV^e</i>
U1	(InternationalHotel (M), Train (M), ArcheologicalSite (H), MediterraneanMeal (H))
U2	(Campsite (H), BikeRiding (M), Trekking (H), MediterraneanMeal (M))
U3	(InternationalHotel (H), Flight (M), LocalTransportation (H), InternationalMeal (H), Attraction (H))
U4	(Bed&Breakfast (M), Flight (H), Museum (M), EthnicMeal (L))
U5	(Bed&Breakfast (M), Riding (H), Boating (M), VegetarianMeal (H), Shopping (L))
U6	(InternationalHotel (M), Flight (M), Trekking (H), MediterraneanMeal (H), ShoppingCenter (L))
U7	(InternationalHotel (H), LocalTransportation (M), CulturalActivity (H), Entertainment (L), Flight (H))
U8	(Farmhouse (H), Train (H), Boating (H), MediterraneanMeal (M))
U9	(Bed&Breakfast (H), Train (H), LocalTransportation (H), OpenAirActivity (M), EthnicMeal (M), RockConcert (M))
U10	(Bed&Breakfast (H), Museum (M), MediterraneanMeal (H), Bazaar (H))
U11	(Bed&Breakfast (M), Car Rental (H), ArcheologicalSite (H), MediterraneanMeal (H), ShoppingCenter (M))
U12	(InternationalHotel (L), Flight (M), ArcheologicalSite (H), EthnicMeal (M))
U13	(Campsite(H), Train (L), Trekking (M), RockConcert (M), Food (L))
U14	(InternationalHotel(M), Flight(H), LocalTransportation(H), EthnicMeal(M), OpenAirActivity(H))
U15	(Bed&Breakfast(H), OpenAirActivity(M), EthnicMeal(M), Bazaar(M), Entertainment(M))

For the purpose of illustration, let us consider the request vector defined by the user *U3*. This user desires to get to his/her destination by flight, reside in an international hotel, commute by public transportation, enjoy international meals and local attractions. As we observe, in the request vector of *U3* all the features have been refined by the score *H*, except *Flight* which has been refined by *M*.

Table 4 shows the experimental results related to the user *U3*. The first column contains the packages chosen by the user, that are *P1*, *P3*, *P4*, *P5*, *P8*, and, respectively, the related similarity values 0.95, 0.8, 0.5, 0.4, 0.2. The other columns illustrate the 10 packages ranked, respectively, according to *SemSim*, *SemSim_{RV}* and *SemSim^e*, with the associated similarity values with *U3*.

Table 4. User *U3*.

<i>HJ</i>	<i>SemSim</i>	<i>SemSim_{RV}</i>	<i>SemSim^e</i>
P1 0.95	P3 0.76	P3 0.76	P3 0.61
P3 0.8	P4 0.40	P4 0.64	P4 0.50
P4 0.5	P1 0.36	P1 0.44	P1 0.42
P5 0.4	P9 0.26	P9 0.26	P5 0.23
P8 0.2	P5 0.20	P5 0.24	P9 0.22
	P10 0.18	P10 0.22	P10 0.12
	P7 0.07	P7 0.07	P7 0.06
	P8 0.07	P8 0.07	P8 0.05
	P2 0.02	P2 0.02	P2 0.02
	P6 0.00	P6 0.00	P6 0.00

In Table 5 *SemSim*, *SemSim_{RV}*, and *SemSim^e* correlations with *HJ* are given. We observe that *SemSim^e* has higher values in 73% of cases. Table 6 shows the precision and recall for all the users, where “-” stands for undefined because the set of retrieved resources is empty. Note that with respect to *SemSim* the precision of *SemSim^e* increases in 5 cases and decreases in 1 case, while the recall increases in most of the cases. Regarding to *SemSim_{RV}* the precision of *SemSim^e* is the same while the recall increases in 4 cases and is the same for the remaining cases.

Table 5. Correlation.

<i>User</i>	<i>Correlation</i>		
	<i>SemSim</i>	<i>SemSim_{RV}</i>	<i>SemSim^e</i>
U1	0.62	0.71	0.71
U2	0.74	0.80	0.91
U3	0.73	0.74	0.82
U4	0.22	0.19	0.53
U5	0.56	0.45	0.81
U6	0.80	0.58	0.69
U7	0.93	0.98	0.98
U8	0.94	0.90	0.89
U9	0.75	0.75	0.83
U10	0.43	0.41	0.54
U11	0.75	0.72	0.71
U12	0.50	0.51	0.57
U13	0.61	0.79	0.58
U14	0.25	0.53	0.57
U15	0.72	0.80	0.78

Table 6. Precision and Recall.

<i>User</i>	<i>Precision</i>			<i>Recall</i>		
	<i>SemSim</i>	<i>SemSim_{RV}</i>	<i>SemSim^e</i>	<i>SemSim</i>	<i>SemSim_{RV}</i>	<i>SemSim^e</i>
U1	-	1	1	0	0.33	0.33
U2	1	1	1	0.33	0.33	0.66
U3	1	1	1	0.33	0.66	1
U4	-	0.5	0.5	0	0.33	0.33
U5	1	1	1	0.33	0.33	0.33
U6	-	1	1	0	0.33	0.33
U7	0.5	0.75	0.75	0.33	1	1
U8	1	1	1	0.33	0.66	1
U9	1	1	1	0.66	0.66	0.66
U10	1	1	1	0.33	0.33	0.33
U11	1	1	1	0.33	0.33	0.66
U12	1	0.66	0.66	0.33	0.66	0.66
U13	1	1	1	0.33	0.66	0.66
U14	-	0.33	0.33	0	0.33	0.33
U15	1	1	1	0.33	0.33	0.33

Note that, due to the presence of the rating scores, the similarity values computed according to *SemSim^e* never increase with respect to *SemSim* and *SemSim_{RV}*. For this reason, the evaluation of the precision and the recall has been performed on the basis of two different thresholds: 0.5 with regard to *HJ*, *SemSim*, and *SemSim_{RV}*, as indicated in Table 4, and 0.4 concerning *SemSim^e*. This distinction allows us to balance the decrease of the similarity values obtained by *SemSim^e* due to the presence of rating scores. In particular, considering *U3* in Table 4, the precision remains invariant (i.e., 1), while the recall increases (0.33 vs. 0.66 vs. 1.00). Both the precision and the recall are equal to 1 because *SemSim^e* retrieves all and only the resources relevant to the user whereas, for instance, the recall of *SemSim_{RV}* is equal to 0.66 because, among *P1*, *P3*, and *P4*, *P1* is not retrieved.

Overall, in selecting resources more similar to users' requests, the experimental results show that *SemSim^e* improves the performances of both *SemSim* and *SemSim_{RV}*. This improvement is achieved because, with respect to *SemSim* and *SemSim_{RV}*, *SemSim^e* relies on the additional knowledge carried by the preference and quality rating scores.

5. Conclusions and Future Work

In this work a semantic similarity method, referred to as *SemSim^e*, has been introduced. It relies on a previous proposal of the authors, namely *SemSim*. In particular, *SemSim^e* has been endowed with the rating scores *High (H)*, *Medium (M)*, *Low (L)* in the *OFV* for representing both the preferences of the user and the quality of the resources in the search space. An experimentation of *SemSim^e* has been performed in the tourism domain by comparing it with the original method *SemSim*, and a variant of it, referred to as *SemSim_{RV}*. The experiment reveals that *SemSim^e* outperforms both *SemSim* and *SemSim_{RV}*, with respect to precision, recall and correlation.

Currently, we are planning to apply the proposed approach on a large dataset with regard to the annotated resources, and also by considering a large participation in the human judgment activity. In particular, we are setting up an experiment by focusing on the Digital Library of the Association for Computing Machinery (ACM). In this context, the reference ontology is represented by the ACM Computing Classification System (ACM-CCS), and the semantic annotations of the resources (more than one thousand papers) are the sets of keywords selected by the authors according to the ACM-CCS.

As a future investigation, *SemSim^e* will be applied to the automotive domain, in the framework of a collaboration with an Italian SME working in this sector, in order to improve the decision making process from both the dealer and the customer sides. In this domain, the idea of giving more relevance to the offer rather than the request will be investigated in order to enable the customer to be aware about the new features from car companies.

Author Contributions: A.F., E.P. and F.T. contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gruber, T.R. A translation approach to portable ontologies. *Knowl. Acquis.* **1993**, *5*, 199–220. [[CrossRef](#)]
2. Formica, A.; Missikoff, M.; Pourabbas, E.; Taglino, F. Semantic search for matching user requests with profiled enterprises. *Comput. Ind.* **2013**, *64*, 191–202. [[CrossRef](#)]
3. Missikoff, M.; Formica, A.; Pourabbas, E.; Taglino, F. Enriching semantic search with preference and quality scores. In *Encyclopedia with Semantic Computing and Robotic Intelligence*; World Scientific: Singapore, 2017.
4. Formica, A.; Missikoff, M.; Pourabbas, E.; Taglino, F. *Weighted Ontology for Semantic Search*; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2008; Volume 5335, pp. 1289–1303.
5. Alani, H.; Brewster, C. Ontology ranking based on the Analysis of Concept Structures. In Proceedings of the K-CAP'05, Banff, AB, Canada, 2–5 October 2005.
6. Euzenat, J.; Shvaiko, P. *Ontology Matching*; Springer: Heidelberg, Germany, 2007.
7. Fang, W.-D.; Zhang, L.; Wang, Y.-X.; Dong, S.-B. Towards a Semantic Search Engine Based on Ontologies. In Proceedings of the 4th International Conference on Machine Learning, Guangzhou, China, 18–21 August 2005.
8. Madhavan, J.; Halevy, A.Y. Composing Mappings among Data Sources. In Proceedings of the VLDB'03—International Conference on Very Large Databases, Berlin, Germany, 9–12 September 2003, pp. 572–583.
9. Maguitman, A.G.; Menczer, F.; Roinestad, H.; Vespignani, A. Algorithmic Detection of Semantic Similarity. In Proceedings of the WWW'05—World Wide Web Conference, Chiba, Japan, 10–14 May 2005.
10. Maarek, Y.S.; Berry, D.M.; Kaiser, G.E. An Information Retrieval Approach for Automatically Constructing Software Libraries. *IEEE Trans. Softw. Eng.* **1991**, *17*, 800–813. [[CrossRef](#)]
11. Castano, S.; De Antonellis, V.; Fugini, M.G.; Pernici, B. Conceptual Schema Analysis: Techniques and Applications. *ACM Trans. Databases Syst.* **1998**, *23*, 286–333. [[CrossRef](#)]
12. Cordi, V.; Lombardi, P.; Martelli, M.; Mascardi, V. An Ontology-Based Similarity between Sets of Concepts. In Proceedings of the WOA'05, Camerino, Italy, 14–16 November 2005; pp. 16–21.

13. Dijkstra, E.W. A note on two problems in connexion with graphs. *Numerische Mathematik* **1959**, *1*, 269–271. [[CrossRef](#)]
14. Manning, C.D.; Raghavan, P.; Schütze, H. *Introduction to Information Retrieval*; Cambridge University Press: New York, NY, USA, 2008.
15. Lin, D. An Information-Theoretic Definition of Similarity. In Proceedings of the 15th International Conference on Machine Learning, Madison, WI, USA, 24–27 July 1998; Shavlik, J.W., Ed.; Morgan Kaufmann: San Francisco, CA, USA, 1998; pp. 296–304.
16. Resnik, P. Using information content to evaluate semantic similarity in a taxonomy. In Proceedings of the IJCAI'95—The 14th International Joint Conference on Artificial Intelligence, Montreal, QC, Canada, 20–25 August 1995; pp. 448–453.
17. Pourabbas, E.; Taglino, F. A semantic platform for enterprise knowledge interoperability. In *Shaping Enterprise Interoperability in the Future Internet 5, Proceedings of the I-ESA'12—Enterprise Interoperability V, Valencia, Spain, 20–23 March 2012*; Poler, R., Doumeingts, G., Katzy, B., Chalmeta, R., Eds.; Springer: London, UK, 2012.
18. Rada, L.; Mili, V.; Bicknell, E.; Bletler, M. Development and application of a metric on semantic nets. *IEEE Trans. Syst. Man Cybern.* **1989**, *19*, 17–30. [[CrossRef](#)]
19. Wu, Z.; Palmer, M. Verb semantics and lexicon selection. In Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics, Las Cruces, NM, USA, 27–30 June 1994; pp. 133–138.
20. Dulmage, A.L.; Mendelsohn, N.S. Coverings of bipartite graphs. *Can. J. Math.* **1958**, *10*, 517–534. [[CrossRef](#)]
21. Ali, G.; ElKorany, A. Semantic-based Collaborative Filtering for Enhancing Recommendation. In Proceedings of the KEOD'14—International Conference on Knowledge Engineering and Ontology Development, Rome, Italy, 21–24 October 2014; pp. 176–185.
22. Eckhardt, A. Similarity of users' (content-based) preference models for Collaborative filtering in few ratings scenario. *Expert Syst. Appl.* **2012**, *39*, 11511–11516. [[CrossRef](#)]
23. Akmala, S.; Shih, L.H.; Batres, R. Ontology-based similarity for product information retrieval. *Comput. Ind.* **2014**, *65*, 91–107. [[CrossRef](#)]
24. Formica, A.; Missikoff, M.; Pourabbas, E.; Taglino, F. Semantic Search for Enterprises Competencies Management. In Proceedings of the KEOD'10—International Conference on Knowledge Engineering and Ontology Development, Valencia, Spain, 25–28 October 2010; pp. 183–192.
25. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).