From Personal Ontologies to Socialized Semantic Space

Jason Jung, Jérôme Euzenat

From Personal Ontologies to Socialized Semantic Space

Jason J. Jung
INRIA Rhône-Alpes
ZIRST 655 avenue de l'Europe, Montbonnot
38334 Saint Ismier cedex, France
j2jung@intelligent.pe.kr

Jérôme Euzenat
INRIA Rhône-Alpes
ZIRST 655 avenue de l'Europe, Montbonnot
38334 Saint Ismier cedex, France
Jerome.Euzenat@inrialpes.fr

ABSTRACT

We have designed a three-layered model which involves the networks between people, the ontologies they use, and the concepts occurring in these ontologies. We propose how relationships in one network can be extracted from relationships in another one based on analysis techniques relying on this network specificity. For instance, similarity in the ontology layer can be extracted from a similarity measure on the concept layer.

Keywords

Semantic web, social network

1. INTRODUCTION

Social network is built from the explicit assertion by users that they have some relation with others or by the implicit evidence of such relations (e.g., co-authoring). In order to support efficient collaboration between users, we propose a three-layered architecture that is capable of capturing the semantics emerged from communities. These semantics are discovered from analyzing the user's social activities on the semantic space. In order words, while building the personal ontologies, the social activities such as linking to a certain user and referring to a domain ontology can represent the corresponding user's semantic preferences. In contrast, as a related work, a tripartite model (actors, concepts, and instances) in [2] has focused on the personal activities based on tagging instances. Also, the similarity measurement for socializing the users is done by co-occurrence analysis with instances and concepts applied by them.

2. THREE-LAYERED ARCHITECTURE

We have designed the three-layered architecture composed of social, ontology, and concept layer. In **social layer** ($\mathcal{S}$), nodes are representing people, and relations are the connections between peoples. It is a directed graph $\langle N_S, E_S^{knows} \rangle$, where N_S is a set of person and $E_S^{knows} \subseteq N_S \times N_S$ the set of relations between these persons.

Demos and Posters of the *3rd European Semantic Web Conference (ESWC 2006)*, Budva, Montenegro, 11th - 14th June, 2006

Ontology layer $\mathcal{O}$ is a network $\langle N_O, E_O^i \rangle$, in which N_O is a set of ontologies and $E_O^i \subseteq N_O \times N_O$ the relationships between these ontologies. Two main kind of relations are $i)$ *import* when some ontology explicitly import another ontology, and $ii)$ *refer* when some ontology uses some concept defined in another ontology. The objective relationship from the $\mathcal{S}$ to the $\mathcal{O}$ is through the explicit usage of ontology by a user which can be expressed by a relation: $Use \subseteq N_S \times N_O$.

In **concept layer** ($\mathcal{C}$), nodes are concepts, and links are the numerous kinds of links that can be found in ontologies. It is a network $\langle N_C, E_C^i \rangle$, in which N_C is a set of entity of an ontology (classes, properties, individuals) and $E_C^i \subseteq N_C \times N_C$ the relationships between these entities. This time the relationships are far more numerous and depends on the kind of entity considered. Such relationships are $i)$ *subClass* linking a class to its subclasses; $ii)$ *superClass*(=subClass^{-1}) linking a class to its super classes; $iii)$ *sibling* linking a class to its siblings; $iv)$ *disjoint* linking a class to the classes it is explicitly disjoint with; $v)$ *property*(=domain^{-1}) linking a class to its properties; $vi)$ *range*$^{-1}$ linking a class to the properties that refer to it. The objective relationship from the $\mathcal{O}$ to the $\mathcal{C}$ is through the definition of concept in an ontology which can be expressed by a relation: $Defines \subseteq N_O \times N_C$. However, this notion of definition is not easy to catch: it could be based on either the assertion of a constraint on some ontology entity or the namespace in which entity belongs. Wewill consider the namespace in the following.

3. INFERRING RELATIONSHIPS

This three-level semantic social network does not bring in itself new improvement for our peer-to-peer sharing application. In order to provide new insight in the possible collaborations it is necessary to analyze these networks and to propagate information from one layer to another. It is assumed that user behaviors is semantically socialized. We explain how, starting from the lower concept layer, it is possible to enrich the upper ontology layer and social layers with new relations from which social network analysis helps finding relevant peers. Besides the numerous relationships that can be found by construction of the concept layer, new relationships can be inferred between the entities. One particular relationship that will be interesting here is *similarity*. In order, to find relationship between concepts from different ontologies, identifying the entities denoting the same concept is a very important feature. As a matter of fact, most of the matching algorithms use some similarity measure or distance in order to match entities. In the spirit of network analysis, they can be defined from the structure of the network. For instance, [1], defines all possible similarities (e.g., Sim_C, Sim_R, Sim_A) between classes, relationships, attributes, and instances. Given a pair of classes c and c' from two different ontologies, $Sim_C \in [0,1]$ is defined as

$$Sim_C(c,c') = \sum_{E \in \mathcal{N}(C)} \pi_E^C MSim_Y(E(c), E(c')) \qquad (1)$$

where $\mathcal{N}(C) \subseteq \{E^1 \dots E^n\}$ is the set of all possible relationships in which classes participate, e.g., subclass, instances, or attributes. The weights π_E^C are normalized (i.e., $\sum_{E \in \mathcal{N}(C)} \pi_E^C = 1$). Thus, if we consider class labels (L) and three relationships in $\mathcal{N}(C)$, which are the superclass (E^{sup}), the subclass (E^{sub}) and the sibling class (E^{sib}), Equ. 1 is rewritten as:

$$\begin{aligned} Sim_C(c,c') &= \pi_L^C sim_L(L(A_i), LF(B_j)) \\ &+ \pi_{sup}^C MSim_C(E^{sup}(c), E^{sup}(c')) \\ &+ \pi_{sub}^C MSim_C(E^{sub}(c), E^{sub}(c')) \\ &+ \pi_{sib}^C MSim_C(E^{sib}(c), E^{sib}(c')). \qquad (2) \end{aligned}$$

where the set functions $MSim_C$ compute the similarity of two entity collections. A distance between two set of classes can be established by finding a maximal matching maximising the summed similarity between the classes:

$$MSim_C(S,S') = \frac{\max(\sum_{\langle c,c' \rangle \in P(S,S')} (Sim_C(c,c'))}{\max(|S|,|S'|)}, \qquad (3)$$

in which P provides a matching of the two set of classes. Methods like the Hungarian method allow to find directly the pairing which maximises similarity. The OLA algorithm is an iterative algorithm that compute this similarity [1]. This measure is normalised because if Sim_C is normalised, the divisor is always greater or equal to the dividend.

A normalized similarity measure can be turned into a distance measure by taking its complement to 1 ($E_C^{dist}(x,y) = 1 - Sim_C(x,y)$). Such a distance introduces a new relation E_C^{dist} in the concept network $\mathcal{C}$. This relation in fact defines a distance network as introduced above.

Then, it can be used for computing a new distance at the ontology level. Again, a distance between two ontologies can be established by finding a maximal matching maximising similarity between the elements of this ontology and computing a global measure which can be further normalised. Thus, **ontology distance** can be computed. Given a set of ontologies N_O, a set of entities N_C provided with a distance function $E_C^{dist} : N_C \times N_C \longrightarrow [0\ 1]$ and a relation $D : N_O \times N_C$, the distance function $E_O^{dist} : N_O \times N_O \longrightarrow [0\ 1]$ is defined as

$$E_O^{dist}(o,o') = \frac{max(\sum_{\langle c,c' \rangle \in P(D(o),D(o')} E_C^{dist}(c,c'))}{max(|D(o)|,|D(o')|} \qquad (4)$$

which is the measure that is used in the OLA algorithm for deciding which alignment is available between two ontologies [1]. However, other distances can be used such as the well known single, average and multiple linkage distances.

This ontology distance introduces a new relation on the ontology layer. This measure provides a good idea of the distances between ontologies. These distances, in turn, are a clue of the difficulty to find an alignment between ontologies. It can be used for choosing to match the closest ontologies with regard to this distance. This can help a newcomer in a community to choose the best contact point: the one with who ease of understanding will be maximised.

Once these measure on ontologies are obtained, this distance can be further used on the social layer. As we proposed it is possible to think that people using the same ontologies should be close to each other. It is possible to measure the **affinity** between people from the similarity between the ontology they use. Given a set of people N_S, a set of ontologies N_O provided with a distance $E_O^{dist} : N_O \times N_O \longrightarrow [0\ 1]$ and a relation $Uses : N_S \times N_O$, the affinity is the similarity measure defined as

$$E^{aff}(p,p') = \frac{1 - max\left(\sum_{\langle o,o' \rangle \in P(Use(p),Use(p'))} 1 - E_O^{dist}(o,o')\right)}{max(|Use(p)|,|Use(p')|)}$$

Since this measure is normalised, it can be again converted to a distance measure through complementation to 1. Introducing the distance corresponding to affinity in the social network allows to compute the affinity relationships between people with regard to their knowledge structure.

4. CONCLUSION AND FUTURE WORK

In order to improving the collaborative sharing and exploitation of this knowledge, we have proposed a three-layered architecture for constructing socialized semantic space from personal ontologies. This space not only supports the relations within a layer, but also the propagation of relations between layers. We have provided the principles for extracting similarity between concepts and propagating this similarity to a distance and an alignment relation between ontologies. This distance relation can be used for discovering affinity in the social network.

There remains important issues to be investigated: all these networks are not equal and their exploitation with classical social network analysis tools can be meaningless (in the same sense that considering the “loves” and “hates” relations as the same would lead to problems). It is thus important to characterise the various relations that were provided with regard to the measures that can be used on them.

5. REFERENCES

[1] J. Euzenat and P. Valtchev. Similarity-based ontology alignment in OWL-Lite. In R. L. de Mántaras and L. Saitta, editors, *Proc. of 16th ECAI*, pages 333–337. IOS Press, 2004.

[2] P. Mika. Ontologies are us: A unified model of social networks and semantics. In *Proc. of 4th ISWC*, pages 522–536, 2005.