Measuring semantic centrality based on building consensual ontology on social network

Jason J. Jung and Jérôme Euzenat

INRIA Rhône-Alpes
ZIRST 655 avenue de l'Europe, Montbonnot,
38334 Saint Ismier cedex, France
j2jung@intelligent.pe.kr, Jerome.Euzenat@inrialpes.fr

Abstract. We have been focusing on three-layered socialized semantic space, consisting of social, ontology, and concept layers. In this paper, we propose a new measurement of semantic centrality of people, meaning the power of semantic bridging, on this architecture. Thereby, the consensual ontologies are discovered by semantic alignment-based mining process in the ontology and concept layer. It is represented as the maximal semantic substructures among personal ontologies of semantically interlinked community. Finally, we have shown an example of semantic centrality applied to resource annotation on social network, and discussed our assumptions used in formulation of this measurement.

1 Introduction

We have been focusing on constructing socialized semantic space to efficiently provide semantic collaboration and interoperability between people. With the emergence of semantic web, users (or actors) on social network have been applying their own personal ontologies to annotate the resources for improving interoperability between each other. However, as the number of users and ontologies are dramatically increasing, the structure of these networks are getting complex. Then, people are suffering from sharing and searching for the relevant information from the networks. In order to solve this problem, we have proposed a three-layered architecture for constructing socialized semantic space, (shown in Fig. 2) [1]. This space is designed to propagate the relational information not only within a layer but also 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.

In this paper, we define the notion of semantic centrality, which expresses the power of controlling *semantic* information flow on social network, and propose a novel network analysis method for measuring semantic centrality. Thereby, we need to discover the consensual ontology $\mathcal{CO}$ from personal ontologies applied to annotate the resources in personal information repositories. In fact, social network analysis (SNA) has regarded a consensus implying the central principles underlying the network as an important challenge [2]. With respect to semantic interoperability between heterogeneous information sources, consensual ontology is playing a role of a "semantic pivot" between heterogeneous information sources [3]. Here we assume that the consensual ontology should be simply organized as a set of concepts which are "most commonly" used in personal ontologies, as well as the relations among these concepts.

Basically, data mining methods are to uncover the hidden (more exactly, frequent) patterns from a given dataset like transactional databases. They also have shown the power of analyzing the structured datasets from various domains. Such datasets are not only XML documents [4] but also web link structure (or topology) [5], and protein structures [6].

In the similar way, we are motivated to extend simple frequent pattern mining method (e.g., *Apriori* algorithm) to *semantic substructure mining* (SSM) algorithm building consensus ontology, because ontologies are basically composed of a set of classes (or concepts) $\mathcal{C}$ and relations $\mathcal{R}$ between the classes [7]. In terms of social network, we can exploit the consensual ontology to measure semantic centrality of the participants on the corresponding social network, with respect to the quantity of major semantic information.

The remainder of the paper is organized as follows. Section 2 simply reviews several definitions in previous studies. Section 3 explains the semantic substructure mining algorithm for building consensual ontology, and then section 4 addresses the semantic centrality measurement from a given subgroup in social network. In section 5, we show an example and argue main contributions of this study by comparing with related work. Finally, section 6 will draw a conclusion and mention some issues as our future work.

2 Centrality measures on social network

A social network is denoted as a graph $G = (N, E)$, where N (a set of nodes) and E (a set of edges) represent users and links between users, respectively. In this paper, we consider only directed and labeled graphs. A path $p(i, j)$ between two arbitrary users u_i and u_j in graph G is a sequence of nodes and edges $\langle n_0, n_1\rangle, \langle n_1, n_2\rangle, \ldots, \langle n_{k-1}, n_k\rangle$, beginning with u_i ($= n_0$) and ending with u_j ($= n_k$), such that each edge connects its preceding with its succeeding node. The length of path is the number of edges (here, k), and we denote the set of shortest paths between u_i and u_j as $SP(i, j)$. Thus, the shortest path distance $spd(i, j)$ between two users u_i and u_j is the minimum element from $SP(i, j)$. Additionally, by Bellman criterion [8], let $\sigma_{i,j}(n)$ indicate the number of shortest paths $p(i, j) \in SP(i, j)$ that node $n \in N$ lies on. Basically, the centrality measures of a user are computed by using several features on the social network, and applied to determine the structural power. So far, in order to extract the structural information from a given social network, various measurements such as centrality [9], pair closeness [10], and authoritative [11] have been studied to realize the social relationships among a set of users. Especially, the centrality can be a way of representing the geometrical power of controlling *information flow* among participants on social network.

Table 1 shows four kinds of centrality measurements. Centrality C_C and C_G are based on the distances with the rest of nodes, while C_S and C_B emphasize the medium mediating between a pair of nodes. These are dependent upon the notion of the characteristics of social network. Also, we may apply hybrid approach of topological features, as combining different centrality measurements.

Table 1. Centrality measurements on social network

Closeness centrality C_C [12]	$C_C(n) = \frac{1}{\sum_{t \in N} spd(n,t)}$
Graph centrality C_G [9]	$C_G(n) = \frac{1}{\max_{t \in N} spd(n,t)}$
Stress centrality C_S [13]	$C_S(n) = \sum_{s \neq n \neq t \in N} \sigma_{s,t}(n)$
Betweenness centrality C_B [14]	$C_B(n) = \sum_{s \neq n \neq t \in N} \frac{\sigma_{s,t}(n)}{\lvert SPs,t \rvert}$

3 Discovering consensual ontology

In this section, we explain how to build the consensual ontology. Thereby, we focus on extracting the most frequent and common classes from a set of ontologies on social network. Substructure mining method will be briefly described, and then, we will show how it is applied to discover consensual ontology.

3.1 Background of substructure mining

Basically, data mining process (e.g., *Apriori* algorithm) can find out the correlation between items by statistical analysis of their occurrences in a given database. It consists of two steps; i) generating the candidate combinations, and ii) pruning by evaluating them with user-specified constraints like minimum support and confidence.

As extending to structured datasets, graph (or tree) mining is to discover the maximal frequent substructures from a given graph-structured dataset. For generating the candidates, the topological analysis is needed to justify whether each subgraph G' is a candidate of a given graph G or not. $G' = (N', E')$ is *induced* from $G = (N, E)$, represented as $G' \preceq G$, if and only if there exists a mapping function $\theta : N' \rightarrow N$ such that i) for each node $n' \in N'$, $n = \theta(n')$, and ii) for each edge $\langle n'_i, n'_j \rangle \in E'$, $\langle n_i, n_j \rangle = \langle \theta(n'_i), \theta(n'_j) \rangle$. Hence, only $G' \preceq G$ can be included in a set of candidate subgraphs [15].

Next, each candidate's support is given by $SUP(G') = \frac{Freq_D(G')}{|DB|}$ where $|DB|$ is the number of graphs in a given database DB, and $Freq_D(G') = \sum_{T \in DB} Occur(G')$ counting the frequency of subgraph G'. The G' of which support value is less than minimum support has to be discarded. The candidates over minimum support are joined each other to find out the larger subgraph. After repeating these steps, eventually, the maximal frequent subgraph can be uncovered.

Particularly, in subtree mining, *PatternMiner* and *TreeMiner* propose a level-wise algorithm based on *Apriori* scheme [16] for mining association rules and depth-first searching for using the novel scope-list, respectively [15].

3.2 Semantic substructure mining algorithm

The personal ontologies $\mathcal{PO}$ on social network is the target of this paper. We regard these personal ontologies as graph-structured knowledge, because they are generated by merging the ontology fragments derived from the reference ontologies by the corresponding user's manual (or semi-automatic) coding [1].

Given a set of personal ontologies, we focus on discovering the consensual ontology $\mathcal{CO}$ under *Apriori* assumption. As extending basic idea of data mining (described in the previous section), the semantic substructure mining algorithm SSM follows the three steps;

1. initialization of a set of candidate classes $CDT^1 = \{\ldots, \{c_i\}, \ldots\}$,
2. expansion of CDT^{t-1} to $\widetilde{CDT}^t = \{\ldots, \{c_i, \ldots, c_{i+t-1}\}, \ldots\}$ by join operation, and
3. refinement of CDT^t by evaluation with user-specific minimum support τ_{SUP}

where $\widetilde{CDT}$ and CDT indicate the power sets including the frequent class sets and the candidate class sets, respectively. The second and third steps are repeated until the constraints such as minimum supports are met ($t = T$). It means that we can finally get the consensual ontology which is composed of T classes.

Semantically induced substructure A candidate class is supposed to be a substructure *semantically induced* from the set of ontologies, and it is represented by

$$cdt_i^t \preceq^{\diamond} PO_k \iff SemInd(cdt_i^t, PO_k) \geq \zeta \tag{1}$$

where $cdt_i^t \in \widetilde{CDT}^t$ and $PO_k \in \mathcal{PO}$. For testing this induction, matching two ontologies has to be conducted by using the semantic similarity measurement, proposed in [17], rather than simple string-matching, in order to reduce some lexical heterogeneity problems such as synonyms. Hence, $SemInd$ is given by

$$SemInd(o_i, O) = \max \frac{\sum_{\langle c,c'\rangle \in Pairing(o_i,O)} Sim_C(c, c')}{|o_i|} \tag{2}$$

in which $Pairing$ provides a matching of the two set of classes. It is established by finding the best matching which is maximizing the summation of the similarities between the classes. The basic notion can be described that two entities are more similar, if they have the more similar features. Then, the class similarity measure Sim_C is formulated as

$$Sim_C(c, c') = \sum_{E \in \mathcal{N}(C)} \pi_E^C MSim_Y(E(c), E(c')) \tag{3}$$

$$\begin{aligned} &= \pi_L^C Sim_L(c, c') + \pi_{sup}^C MSim_C(E^{sup}(c), E^{sup}(c')) \\ &+ \pi_{sub}^C MSim_C(E^{sub}(c), E^{sub}(c')) \\ &+ \pi_{sib}^C MSim_C(E^{sib}(c), E^{sib}(c')) \end{aligned} \tag{4}$$

where $\mathcal{N}(C) \in \{E^1, \ldots, E^n\}$ is the set of all relationships in which classes are involved (in this paper, we are considering three relationships; superclass, subclass, and sibling class), and π^C is the normalized weighting factor to the corresponding relationships. Also, similar to Eq. 2, the set function $MSim_C$ is given by

$$MSim_C(S, S') = \frac{\max \sum_{\langle c,c'\rangle \in Pairing(S,S')} Sim_C(c, c')}{\max(|S|, |S'|)}. \tag{5}$$

Finally, $SemInd(cdt_i^t, PO_k)$ is assigned into $[0, 1]$. Thus, cdt_i^t of which similarity with a given ontology PO_k is over ζ is regarded as one of semantically induced substructures from PO_k. When $\zeta = 1$, only candidates exactly matched will be chosen without concerning about the semantic heterogeneity.

Expansion and refinement by evaluation In order to discover the maximal frequent substructure, we have to repeat these two processes; expansion for generating candidates and refinement. Refinement process of candidates induced from personal ontologies, exactly same as in general data mining, is to compare the frequency of the corresponding substructure candidate with user-specific threshold (e.g., minimum support τ_{SUP}). The candidate cdt_i^t extracted through comparing the similarities measured by $SemInd$ with ζ can be counted as the occurrence in the set of personal ontologies $\mathcal{PO}$. Function $Occur^{\diamond}$ returns 1, if $cdt_i^t \preceq^{\diamond} PO_k$. Otherwise, it returns 0. Thus, frequency of a candidate is $Freq_{\mathcal{PO}}(cdt_i^t) = \sum_{PO_k \in \mathcal{PO}} Occur^{\diamond}(cdt_i^t)$, and the support is given by

$$SUP(cdt_i^t) = \frac{Freq_{\mathcal{PO}}(cdt_i^t)}{|\mathcal{PO}|} = \frac{\sum_{PO_k \in \mathcal{PO}} Occur^{\diamond}(cdt_i^t)}{|\mathcal{PO}|}. \tag{6}$$

Only the candidate set of classes cdt_i^t of which support $SUP(cdt_i^t) \geq \tau_{SUP}$ can be chosen to generate the expanded candidates $\widetilde{CDT}^{t+1}$.

After a set of candidate features CDT^1 is initially selected by

$$CDT^1 = \{cdt_i^1 | SUP(cdt_i^1) \geq \tau_{SUP}\}, \tag{7}$$

we have to expand the set of candidate class sets and refine them where $t \geq 2$. Thus, CDT^t is obtained by

$$CDT^t = refine(\widetilde{CDT}^t) \tag{8}$$

$$= refine(expand(CDT^{t-1})) \tag{9}$$

where function $refine$ is to evaluate $\binom{|CDT^{t-1}|}{t}$ set elements generated by function $expand$ where $|CDT^{t-1}|$ is the total number of the single classes in CDT^{t-1}.

3.3 Consensual ontology and semantic subgroup discovery

By using semantic substructure mining algorithm, the maximal semantic substructures were able to be obtained from a given set of personal ontologies. Then, the consensual ontology $\mathcal{CO}$ is represented as $\{cdt_i^T | cdt_i^T \in CDT^T, SUP(cdt_i^T) \geq \tau_{SUP}\}$ when $\widetilde{CDT}^{T+1}$ is an empty set.

However, we have to realize the problem when the target social network is intermingled with semantically heterogeneous communities. Substructure mining algorithm based on counting simple occurrence (or frequency) analysis is difficult to build more than two consensual ontologies at the same time. Thereby, the social network should be fragmented into the communities (or groups [18]) whose semantic preferences are more cohesive with each other than others. In other words, this is similar to *user clustering* based on the semantic cohesion among users on the social network. Thus, let K the number of communities (user groups) on social network. The best combination of user groups is obtained by maximizing the objective function $F_{SubGroup}(UG_1, \ldots, UG_K)$

$$\max F_{SubGroup} = \max \frac{\sum_{k=1}^{K} Distance(\mathcal{CO}_i, \mathcal{CO}_j)}{K} \tag{10}$$

$$= \max \frac{\sum_{k=1}^{K}(1 - Sim_C(c \in \mathcal{CO}_i, c' \in \mathcal{CO}_j))}{K} \tag{11}$$

$$\approx \min \frac{\sum_{k=1}^{K} Sim_C(c \in \mathcal{CO}_i, c' \in \mathcal{CO}_j)}{K} \tag{12}$$

where $\mathcal{CO}_i = SSM(UG_i)$. Function $Distance$ is derived from similarity measure by taking its complement to 1. Through this equation, the underlying communities can be found out. Each time the function $refine$ of SSM algorithm is finished, this process should be conducted.

4 Semantic centrality

As mentioned in Table 1, there have been several centrality indices to measure the power of structural position on social network. However, they are not appropriate to reflect the centrality among the underlying semantic relationships between personal ontologies on the socialized semantic network introduced in our previous work [1].

We define a semantic centrality as the power of semantic bridging on the semantic social network. Suppose that two users s and t are not able to communicate with each other, due to the semantic heterogeneity between their personal ontologies PO_s and PO_t. Thereby, we need to search for the personal ontology PO_i of which semantic centrality is high enough to reconcile these ontologies. It means PO_i is containing some classes matched with the consensual ontology $\mathcal{CO}$. We intuitively assume that a user is assigned higher semantic centrality, as his personal ontology includes more consensual classes in common. Thus, we formulate a semantic centrality of i-th user $C^{\diamond}(i)$ as

$$C^{\diamond}(i) = \frac{|PO_i \cap \mathcal{CO}|}{|PO_i|} \sum_{s \neq t \neq i \in N} \frac{\sigma^{\diamond}_{PO_s, PO_t}(PO_i)}{|SP^{\diamond}(s,t)|} \tag{13}$$

which means the semantic closeness (or coverage) of the personal ontology PO_i to the discovered consensual ontology $\mathcal{CO}$. The denominator $|PO_i|$ is for the normalization by the total number of classes organizing the personal ontology. $SP^{\diamond}$ is a pair of users whose personal ontologies are not semantically interoperable directly. So, C_B can be replaced by C_C or others. More importantly, function $\sigma^{\diamond}$ is to determine the efficiency of reconciliation, and it is given by

$$\sigma^{\diamond}_{PO_s, PO_t}(PO_i) = \frac{|PO_s \cap PO_i| \cdot |PO_t \cap PO_i|}{|PO_s \cap \mathcal{CO}| \cdot |PO_t \cap \mathcal{CO}|} \tag{14}$$

which expresses that the number of matched classes between two ontologies is in linear proportion, in contrast of that of matched classes with consensus ontologies. Additionally, in Eq. 13 and 14, the counting computation of union sets is done by

$$|A \cap B| = count(\langle c, c' \rangle)_{\langle c,c' \rangle \in Pairing(A,B), Sim_C(c,c')=1}. \quad (15)$$

As next issue, we note that there are two kinds of semantic centrality measurements, with respect to the scope and the topologies of communities.

– *Local* semantic centrality $C_L^\diamond$ means the power of semantic bridging between the members within the same community.
– *Global* semantic centrality $C_G^\diamond$ implies the measurement of the bridging power between two communities.

Then, for computing these centrality measurements, the communities on the whole social network should be firstly organized by using SSM algorithm. Local semantic centrality $C_L^\diamond$ is computed by

$$C_L^\diamond(i) = \alpha_T C_B(i) + (1 - \alpha_T) C^\diamond(i) \quad (16)$$

where the first term is for reflecting the effect of physical (or explicit) social linkage of a given community (mentioned in Table 1), and the second term is semantic centrality $C^\diamond$. The coefficient $\alpha_T \in [0, 1)$ is to control the portion of topological effects. This is formulated as linearly combined with topological centrality measurements and semantic centrality in Eq. 13.

On the other hand, global semantic centrality $C_G^\diamond(i, X)$ of i-th user to a certain community X is based on three factors; i) topologically, the betweenness centrality between the people of two communities I, including i-th user, and X, ii) the similarity between the consensual ontology of target communities X and the corresponding personal ontology PO_i, and iii) the corresponding local semantic centrality. Thus, as linearly combined, it is given by

$$\begin{aligned} C_G^\diamond(i, X) = {} & \beta_T C_B(i) \frac{\sum_{x \in X} link(i, x)}{\sum_{u \in I, x \in X} link(u, x)} \\ & + \beta_S MSim_C(PO_i, CO_X) \\ & + \beta_{LS} C_L^\diamond(i) \sum_{j \in X, link(i,j)=1} C_L^\diamond(j) \end{aligned} \quad (17)$$

where $\beta_T, \beta_S, \beta_{LS} \in [0, 1)$ are the coefficients controlling the portion of topological effects, similar to α_T, the similarity effects, and local semantic centrality effects, respectively. For normalization, $\beta_T + \beta_S + \beta_{LS} = 1$. First term simply indicates the ratio of the linkages by the corresponding user to the total linkages with the target community. User j in the target community X is a member linking to the i-th user. In second term, we put the ontology similarity between consensual ontology of community X and personal ontology, because the more similar classes make the mediation powerful. Finally, the third term applies the local semantic centrality. When a user are linked with more "semantically" central users in a community, his global centrality becomes increased.

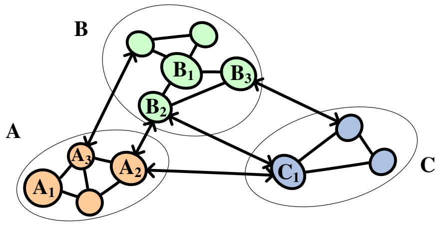

Fig. 1. Two kinds of semantic centrality measurements

We want to show two cases, with respect to $C_L^\diamond$ and $C_G^\diamond$. Let three communities (A, B, and C) organized from a given social network (the number of communities $K = 3$ in Eq. 12), as shown in Fig. 1. We assume that user A_1, B_1, and C_1 are is the highest $C_L^\diamond$ within each community. First case, most possibly, is that the user whose local semantic centrality is highest is also assigned the highest global semantic centrality, e.g., $C_G^\diamond(C_1, B)$ (or $C_G^\diamond(C_1, A)$) calculated as

$$C_G^\diamond(C_1, B) = \beta_T(\frac{1}{3})(\frac{2}{2}) + \beta_S MSIM_C(CO_B, PO_{C_1}) \\ + \beta_{LS} C_L^\diamond(C_1)\{C_L^\diamond(B_2) + C_L^\diamond(B_3)\}$$

where $C_L^\diamond(C_1)$ is assumed to be larger than any other members in community C. Moreover, topologically, C_1 is the only channel to communicate with other communities. As second case, in community B to C, even though B_1's local centrality is the highest, $C_G^\diamond(B_2, C)$ might be higher due to the linkage patterns with C. This is also larger than $C_G^\diamond(B_3, C)$, because C_1 is assigned the highest local semantic centrality.

Here, the scenario is given for explaining how semantic centrality is applied to. Above all, Fig. 2 shows the three-layered architecture of semantic social network, which is composed of a social layer, an ontology layer, and a concept layer. While social layer can simply store the physical connections between users, ontology layer represents the personal ontologies applied to annotate the corresponding user's resources. The classes organizing these ontologies are located in concept layer, so that they are aligned to measure the similarity. Assume that five users are organizing social network and their links are shown in social layer. With respect to the traditional centrality measurements, we

Table 2. Measuring centrality on social network

Centrality	Antoine	Jerome	Jason	Arun	Sebastien
Closeness centrality	1/7	1/6	**1/5**	1/7	1/9
Betweenness centrality	4/10	7/10	**9/10**	4/10	4/10
Semantic centrality		$\star$			

found out that *Jason* is the most powerful user, as shown in Table 2. However, during

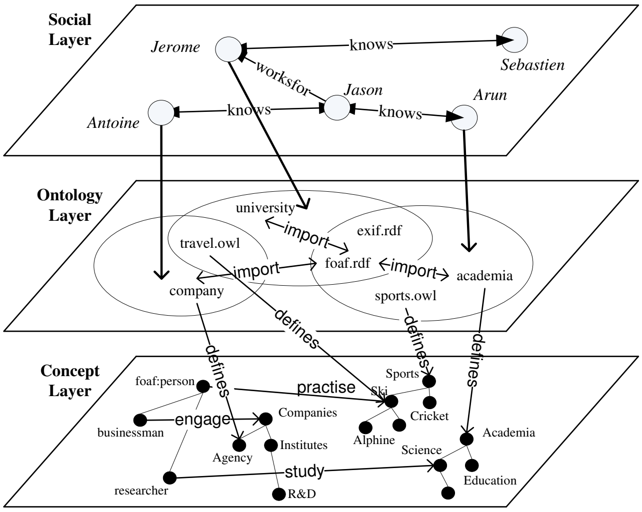

Fig. 2. A three-layered social semantic network

bottom-up inference (which is from concept layer to social layer), the underlying hidden patterns are uncovered. The semantic bridging between *Antoine* and *Arun* is provided only *Jerome*'s personal ontology. People have been trying to annotate the resources (e.g., photos) in their own repository by using personal ontologies. In order to increase the efficiency of annotation, they have been needed to share the semantic information between each other or photos themselves. The problem is caused by semantic heterogeneity of annotations, e.g., between *Antoine* and *Arun*. Thereby, they should search for the users who can most likely play a role of a semantic bridge between them, and ask him (e.g., *Jerome*) to translate the annotations for mutual understanding.

Next, we can consider semantic centrality to propagate semantic information. The semantic information can be newly asserted or changed over time, and it is supposed to be announced to all other users on social network. For the purpose of efficient network management, up-to-date semantic information can be propagated in order of semantic centrality. Somehow, *Arun* acquires some new information and update his personal ontology. Thus, this event has to be notified to *Jerome*, rather to *Jason*, even though he is directly linked with *Jason*.

5 Discussion and related work

Many systems have been interested in information sharing on the distributed systems. With emergent of semantic web environment, rather than content, semantic information has been the target data to be exchanged. Since EDUTELLA introduced an infrastructure for exchanging metadata on peer-to-peer (P2P) environment [19], several querying-based systems have been implemented such as Bibster [20], Oyster [21], and SQAPS [22]. Our goal is also to search for relevant semantic information and share it with other users on distributed information space, but we are more focusing on how peers are interlinked, the so-called social network analysis. Especially, as the contributions of this paper, we proposed two main ideas;

1. consensual ontology discovery by semantic substructure mining algorithm and
2. local versus global semantic centrality measurement.

Firstly, in order to construct consensual ontologies, this paper introduced a pruning-based approach to discover the maximal frequent substructure. In contrast, as the most general approach, Stephens et al. have organized the consensual (or global) ontologies by exhaustive merging of a given set of ontologies [3] in order to retain or maximize the chance to be semantically bridged. Additionally, [23] mentions ontology-based consensus making process from the user communities. It also seems to try to find out the consensus ontology, after organizing the communities from people.

Meanwhile, in order to support the communication between communities, in [24], Breslin et al. proposed the SIOC (Semantically Interlinked Online Communites) ontology[1], rather than discovering the consensus.

Especially, in terms of the efficiency of query propagation on peer-to-peer environment, several studies have introduced the systems based on semantic overlay network (SON). They are based on the broadcasting scheme. After the queries are semantically analyzed, the relevant topics are distilled. Through the multiple overlay network, the queries are propagated to either the set of selected nodes [25], or the super-peers [26]. In contrast, in our method, the queries should be sent to the node whose semantic centrality is largest.

Finally, we want to discuss the personal ontology built by people. In this study, the personal ontology is assumed to play a role of the important evidence reflecting the preferences of corresponding user. However, because users can refer to the upper-level ontologies and even import the other user's personal ontologies, it is too ambiguous to measure the similarity between personal ontologies for the consensus.

6 Concluding remarks and future work

As a conclusion, we put forward a new measurement for semantic centrality, expressing the potential power of semantic bridging among users on social network. Consensual ontologies thereby were built by semantic substructure mining algorithm, and they had the capability to discover the subgroups whose semantic preferences are relatively closer than others. More importantly, the notion of them was designed to be adaptable to the three-layered architecture (social, ontology, and concept layer) for socialized semantic space [1]. The three-layered architecture provides two ways of inference for the hidden relationships between entities; top-down (from social layer to concept layer) and bottom-up (reversely). This paper is related to the bottom-up inference. Especially, we want to mention that the communities on social network are organized with respect to *semantic preference* implicitly reflected during designing and using their own personal ontologies.

[1] SIOC. http://rdfs.org/sioc/

For dealing with this problem mentioned in Sect. 5, we have to track the user actions and interactions. For instance, similar to [27], we may consider only the concepts applied to the specific resources. On the other hand, the "concepts with dust," which is not used for a long time, should be degraded by using machine learning methodologies.

As future work of semantic centrality, we have three main plans to investigate the followings issues

- *semantic subgroup discovery*, to organize the sophisticated user groups with enhancing Eq. 12,
- *query propagation*, to determine the ordering (or route) of potential peers to which the queries will be sent, and
- *semantic synchronization*, to maximize the efficiency interoperability by information diffusion.

Furthermore, we have to consider to enhance the semantic centrality measurement $C^\diamond$ by combining with i) authoritative and hub centrality measurement, proposed in [11], and ii) the modified shortest paths $spd(n, t) = \frac{1}{C^\diamond(n)+C^\diamond(t)}$. Finally, like [28], we have plan to visualize the semantic dynamics on the social network.

References

1. Jung, J.J., Euzenat, J.: From personal ontologies to semantic social space. In: Poster of the 4th European Semantic Web Conference (ESWC 2006). (2006)
2. Wasserman, S., Faust, K.: Social Network Analysis. Cambridge University Press (1994)
3. Stephens, L.M., Gangam, A.K., Huhns, M.N.: Constructing consensus ontologies for the semantic web: A conceptual approach. World Wide Web **7**(4) (2004) 421–442
4. Braga, D., Campi, A., Klemettinen, M., Lanzi, P.L.: Mining association rules from XML data. In Kambayashi, Y., Winiwarter, W., Arikawa, M., eds.: Proceedings of the 4th International Conference on Data Warehousing and Knowledge Discovery (DaWaK 2002), Aix-en-Provence, France. Volume 2454 of Lecture Notes in Computer Science., Springer (2002) 21–30
5. Chakrabarti, S., Dom, B., Kumar, R., Raghavan, P., Rajagopalan, S., Tomkins, A., Gibson, D., Kleinberg, J.M.: Mining the web's link structure. IEEE Computer **32**(8) (1999) 60–67
6. Huan, J., Wang, W., Bandyopadhyay, D., Snoeyink, J., Prins, J., Tropsha, A.: Mining protein family specific residue packing patterns from protein structure graphs. In: Proceedings of the eighth annual international conference on Resaerch in computational molecular biology (RECOMB '04), New York, NY, USA, ACM Press (2004) 308–315
7. Gruber, T.R.: A translation approach to portable ontology specifications. Knowledge Acquisition **5**(2) (1993) 199–220
8. Brandes, U.: A faster algorithm for betweenness centrality. Journal of Mathematical Sociology **25**(2) (2001) 163–177
9. Haga, P., Harary, F.: Eccentricity and centrality in networks. Social Networks **17**(1) (1995) 57–63
10. Girvan, M., Newman, M.E.J.: Community structure in social and biological networks. Proceedings of the National Academy of Sciences **99** (2002) 7821–7826
11. Kleinberg, J.M.: Authoritative sources in a hyperlinked environment. Journal of ACM **46**(5) (1999) 604–632
12. Sabidussi, G.: The centrality index of a graph. Psychometirka **31** (1966) 581–603

13. Shimbel, A.: Structural parameters of communication networks. Bulletin of Mathematical Biophysics **15** (1953) 501–507
14. Freeman, L.C.: A set of measures of centrality based on betweenness. Sociometry **40** (1977) 35–41
15. Zaki, M.J.: Efficiently mining frequent trees in a forest: Algorithms and applications. IEEE Transactions on Knowledge and Data Engineering **17**(8) (2005) 1021–1035
16. Agrawal, R., Srikant, R.: Fast algorithms for mining association rules in large databases. In: Proceedings of the 20th International Conference on Very Large Data Bases (VLDB '94), San Francisco, CA, USA, Morgan Kaufmann Publishers Inc. (1994) 487–499
17. Euzenat, J., Valtchev, P.: Similarity-based ontology alignment in OWL-Lite. In de Mántaras, R.L., Saitta, L., eds.: Proceedings of the 16th European Conference on Artificial Intelligence (ECAI'2004), Valencia, Spain, August 22-27, 2004, IOS Press (2004) 333–337
18. Kleinberg, J.M.: Small-world phenomena and the dynamics of information. In Dietterich, T.G., Becker, S., Ghahramani, Z., eds.: Advances in Neural Information Processing Systems 14 [Neural Information Processing Systems: Natural and Synthetic, NIPS 2001, December 3-8, 2001, Vancouver, British Columbia, Canada], MIT Press (2001) 431–438
19. Nejdl, W., Wolf, B., Qu, C., Decker, S., Sintek, M., Naeve, A., Nilsson, M., Palmér, M., Risch, T.: EDUTELLA: a P2P networking infrastructure based on rdf. In: Proceedings of the 11th international conference on World Wide Web (WWW '02), New York, NY, USA, ACM Press (2002) 604–615
20. Haase, P., Broekstra, J., Ehrig, M., Menken, M., Mika, P., Olko, M., Plechawski, M., Pyszlak, P., Schnizler, B., Siebes, R., Staab, S., Tempich, C.: Bibster - a semantics-based bibliographic peer-to-peer system. In McIlraith, S.A., Plexousakis, D., van Harmelen, F., eds.: Proceedings of the Third International Semantic Web Conference (ISWC 2004),Hiroshima, Japan, November 7-11, 2004. Volume 3298 of Lecture Notes in Computer Science. (2004) 122–136
21. Palma, R., Haase, P.: Oyster - sharing and re-using ontologies in a peer-to-peer community. In Gil, Y., Motta, E., Benjamins, V.R., Musen, M.A., eds.: Proceedings of the 4th International Semantic Web Conference (ISWC 2005), Galway, Ireland, November 6-10, 2005. Volume 3729 of Lecture Notes in Computer Science., Springer (2005) 1059–1062
22. Fernandez-Garcia, N., Sanchez-Fernandez, L., Blazquez, J.M., Larrabeiti, D.: An ontology-based P2P system for query-based semantic annotation sharing. In: Proceedings of Workshop on Ontologies in Peer-to-Peer Communities at the ESWC 2005, May 30, Heraklion, Greece. (2005)
23. Zhdanova, A.V., Martín-Recuerda, F.: Consensus making on the semantic web: Personalization and community support. In Ngu, A.H.H., Kitsuregawa, M., Neuhold, E.J., Chung, J.Y., Sheng, Q.Z., eds.: Proceedings of the 6th International Conference on Web Information Systems Engineering (WISE 2005), November 20-22, 2005. Volume 3806 of Lecture Notes in Computer Science., Springer (2005) 599–600
24. Breslin, J.G., Harth, A., Bojars, U., Decker, S.: Towards semantically-interlinked online communities. In Gómez-Pérez, A., Euzenat, J., eds.: Proceedings of the Second European Semantic Web Conference (ESWC 2005), May 29 - June 1, 2005. Volume 3532 of Lecture Notes in Computer Science., Springer (2005) 500–514
25. Crespo, A., Garcia-Molina, H.: Semantic overlay networks for p2p systems. In Moro, G., Bergamaschi, S., Aberer, K., eds.: Proceedings of the 3rd International Workshop Agents and Peer-to-Peer Computing (AP2PC 2004), July 19, 2004. Volume 3601 of Lecture Notes in Computer Science., Springer (2005) 1–13
26. Löser, A., Naumann, F., Siberski, W., Nejdl, W., Thaden, U.: Semantic overlay clusters within super-peer networks. In Aberer, K., Kalogeraki, V., Koubarakis, M., eds.: Proceedings of the First International Workshop on Databases, Information Systems, and Peer-to-Peer Computing (DBISP2P), September 7-8, 2003. Volume 2944 of Lecture Notes in Computer Science., Springer (2004) 33–47

27. Mika, P.: Ontologies are us: A unified model of social networks and semantics. In Gil, Y., Motta, E., Benjamins, V.R., Musen, M.A., eds.: Proceedings of the 4th International Semantic Web Conference (ISWC 2005), November 6-10, 2005. Volume 3729 of Lecture Notes in Computer Science., Springer (2005) 522–536
28. Jung, J.J.: Visualizing recommendation flow on social network. Journal of Universal Computer Science **11**(11) (2005) 1780–1791