

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Abderrahmane Mira de Bejaia

Faculté des Sciences Exactes

Département d'Informatique

ECOLE DOCTORALE RESEAUX ET SYSTEMES DISTRIBUES



Mémoire de Magistère

En Informatique

Option : Réseaux et Systèmes Distribués

Thème

Reformulation de Requêtes pour la découverte des services web sémantiques

Présenté par

MEKHZOUMI Dalila

Devant le jury composé de :

Président	TARI Abdelkamel	Maitre de conférences A à l'Université de Bejaia.
Rapporteur	NOUALI Omar	Directeur de recherche au CERIST, Alger.
Examineur	BALLA Amar	Professeur à l'ESI, Alger.
Examineur	DERHAB Abdlouahid	Maitre de recherche A au CERIST, Alger.
Invitée	ALIANE Hassina	Chargée de recherche au CERIST, Alger.

Promotion 2008 / 2009

Remerciements

Je remercie Dieu le tout puissant qui m'a donné le courage et la force pour aboutir à l'accomplissement de ce travail.

J'adresse tout d'abord mes remerciements à mes encadreurs, pour avoir accepté de m'encadrer et m'avoir aidé à conduire ce travail jusqu'au bout : Madame ALIANE Hassina, Chargée de recherche au CERIST et Monsieur NOUALI Omar, Directeur de recherche au CERIST.

Je remercie aussi Messieurs TARI, BALLA et DERHAB, les membres du jury pour l'honneur qu'ils m'ont attribué en acceptant d'évaluer et de juger ce modeste travail.

Je remercie le directeur du CERIST, le Professeur BADACHE Nadjib pour nous avoir donné cette opportunité de continuer nos études.

Ma plus profonde reconnaissance et remerciement s'adresse à : Chahnez ZAKARIA et Nabila BOUSBIA, qui ont activement participé à l'élaboration de cette recherche et pour m'avoir poussé à faire le magister, encouragé, motivé et surtout leur soutien moral ininterrompu.

Ce travail n'aurait jamais été possible sans le soutien et l'appui moral des membres de ma famille. Je les remercie tous, particulièrement mes parents, mon mari ALLOUNE Farid et ses parents.

Et je tiens à remercier mes amies Samira BOUCHAMA , Imane ZENDJEBIL, Fadila Machouk et Mina BOUHBILA pour leur disponibilité et encouragements.

Enfin, pour être sûr de n'oublier personne, que tous ceux, qui de près ou loin, ont contribué par leurs conseils, leurs encouragements ou leurs amitiés, à l'aboutissement de ce modeste travail, trouvent ici l'expression de ma profonde reconnaissance.

RESUME

Les services web sémantiques ont été introduits pour mieux décrire les services web et améliorer leurs fonctionnements. Plusieurs approches de découverte ont été proposées. La plus utilisée est OWL-S, où la découverte consiste à mettre en correspondance la requête de l'utilisateur et le service profil d'un service OWL-S. Cependant, cette découverte n'offre pas toujours les meilleurs résultats, surtout quand un utilisateur ne connaît pas le contenu et la description du service web dont il a besoin. L'objectif de ce travail consiste à reformuler la requête utilisateur afin de sélectionner tous les services web répondants à cette requête. Pour cela, nous donnons d'abord plus de souplesse à l'utilisateur dans la formulation de sa requête. Ensuite nous reformulons la requête de l'utilisateur en l'enrichissant via l'ontologie générale WordNet selon la famille morphologique, les synonymes du sens le plus connu et les synonymes de tous les sens. Enfin, nous mettons en correspondance la requête et les ontologies des services web sémantique OWL-S en utilisant des mesures de similarité. Nous avons mis en place une évaluation de notre approche et nous avons présenté les résultats.

Mots clés : découverte des services web sémantiques, reformulation par expansion, enrichissement de requête, OWL-S, Matching, WordNet, similarité sémantique.

ABSTRACT

The semantic web services have been introduced to describe better Web services and improve their functioning. Several discovery approaches have been proposed. The most used is the one proposed in OWL-S, where the discovery is to match the user query and service profile of OWL-S service. However, this finding does not always produce the best results, especially when a user does not know the contents and the description of the web service it needs. This work aims at reformulating the user query to select all web services that respond to this query. For this, we first give more flexibility to the user in formulating his request. Then we reformulate the user query by enriching it via the general ontology WordNet according to the morphological family, synonyms meaning the best known and the synonyms of all senses. Finally, we match the query and the semantic web service ontologies OWL-S using similarity measures. We have implemented an evaluation of our approach and we presented the results.

Keywords: discovery of semantic web services, matching, query expansion, WordNet, semantic similarity.

المخلص

لقد تم إدخال خدمات الويب الدلالية لوصف أفضل لخدمات الويب وتحسين أدائها. وقد اقترحت عدة أساليب اكتشاف لهذه الخدمات. الأكثر استخداما هو مقترح OWL - S، حيث الاكتشاف يتمثل في مطابقة طلب المستخدم و خدمة تعريف خدمة OWL-S. ومع ذلك، هذه الطريقة في اكتشاف خدمات الويب لا تقدم دائما أفضل النتائج، خصوصا عندما يجهل المستخدم محتويات ووصف خدمة الويب التي يحتاجها. هذا العمل يهدف إلى إعادة صياغة طلب المستخدم لتحديد خدمات الويب التي تستجيب لهذا الطلب. لهذا، فإننا نقدم أولا مرونة أفضل للمستخدم في صياغة طلبه. ثم نقوم بإعادة صياغة هذا الطلب عن طريق إثرائه بواسطة الأنطولوجيا العامة WordNet وفق: الأسرة المورفولوجية، مرادفات المعنى الأكثر شهرة والمرادفات لجميع المعاني. وفي الأخير، فإننا نطابق طلب المستخدم و انتولوجيات خدمات الويب الدلالية OWL-S باستخدام قياسات التشابه. قمنا بتقييم منهجنا المقترح وقدمنا النتائج.

الكلمات الرئيسية : اكتشاف خدمات الويب الدلالية، إعادة صياغة طلبات المستخدم، WordNet، التشابه اللفظي.

Table des Matières

Introduction générale	1
1. Contexte général	1
2. Problématique et objectifs de l'étude	1
3. Contributions	2
4. Organisation du mémoire	3
Chapitre 1 : Le web sémantique et les ontologies	4
1. Définition du web sémantique	4
2. Architecture du web sémantique	4
3. Les ontologies	6
4. Exemples d'ontologie : WordNet	9
5. Les mesure des similarités entre concepts	11
5.1 Similarité entre concepts	11
5.2. Les approches des mesures de similarité	12
5.3. Comparaison et évaluation	14
6. Conclusion	15
Chapitre 2 : Les services web sémantiques	16
1. Définition et caractéristiques des services web	16
2. Architecture des services web	17
3. Définition des services web sémantiques	18
4. Approches pour les services web sémantiques	19
4.1. OWL-S	19
4.2. WSMO	20
4.3. METEOR -S	22
4.4. Etude comparative	23
5. Conclusion	24
Chapitre 3 : Stratégies de découverte des services web	25
1. Les modèles syntaxique basés sur UDDI/WSDL	25
2. Les modèles sémantiques basés sur OWL-S	27
3. Quelque travaux existants	29

4. Conclusion	30
Chapitre 4 : Reformulation de requête basée sur une ontologie	31
1. Processus de base de recherche d'information	31
2. Utilisation des ontologies dans un système de recherche d'information	32
3. Reformulation de requête via une ontologie	33
4. Quelques travaux sur la reformulation de requête.....	35
5. Conclusion	36
Chapitre 5 : Contribution	37
1. Schéma et description générale de l'approche proposée	38
2. Le format de la requête de l'utilisateur:.....	39
3. Traitement de la requête :	40
3.1. La lemmatisation	41
3.2. Uniformiser les termes.....	42
3.3. Suppression des mots vides	42
3.4. Traduction des termes :.....	43
4. Enrichissement de la requête :	44
5. Prétraitement sur les ontologies OWL-S	46
5.1. Extraction des termes.....	46
5.2. Lemmatisation.....	47
6. Mise en correspondance	47
6.1. Principe de mise en correspondance.....	48
6.2. Les mesures de similarités utilisées.....	50
6.2. Similarité entre requête utilisateur et service profil d'une ontologie OWL-S	53
6.3. Sélection des services pertinents	54
7. Conclusion	55
Chapitre 6 : Evaluation	56
1. Les corpus et leur usage pour l'évaluation	56
2. Les mesures d'évaluation	57
2.1. Rappel.....	57
2.2. Précision	57
2.3. F-mesure	57
3. Présentation du corpus.....	58
4. Scénario de l'évaluation	60

4.1. Expérimentation du seuil de similarité	60
4.2 Thèmes	63
4.3. Dictionnaire.....	67
5. Synthèse.....	68
Conclusion.....	70
Perspectives	71
Références Bibliographiques.....	73
Annexe : Evaluation.....	81

Liste des Figures

Figure 1.1: L'architecture du Web Sémantique.....	5
Figure 1.2 : Exemple de sous hiérarchie dans WordNet correspondant au concept "car".	11
Figure 2.1 : Architecture des services web. [21]	17
Figure 2.2 : L'ontologie OWL-S [26].	20
Figure 2.3 : Les éléments de WSMO [63].	21
Figure 3.1 : Exemple de Matching de type subsumé.	28
Figure 4.1: Recherche d'information en réponse à une requête [50].	32
Figure 4.2 : L'ontologie insérée dans le processus de recherche d'information [9].	33
Figure 5.1 : Notre approche de reformulation de requête pour la découverte des services web sémantiques.	39
Figure 5.2: Traitement de la requête.	41
Figure 5.3: Exemple d'une phrase lemmatisé avec TreeTagger.	42
Figure 5.4: Les méthodes d'enrichissement de requête utilisateur.	45
Figure 5.5: Exemple d'un service profil d'une ontologie OWL-S.....	47
Figure 5.6 : Schéma de mise en correspondance.	49
Figure 5.7: Schéma de calcul de similarité entre le paramètre Output de la requête et les paramètres Output d'une ontologie OWL-S.	50
Figure 5.8: Exemple de la richesse du paramètre « <i>textDescription</i> », d'un service profil d'une ontologie OWL-S.....	52
Figure 6.1: Taille du corpus par thème.	58
Figure 6.2: Exemple de service web sémantique « book price ».	59
Figure 6.3: Précision, Rappel et F mesure des quatre méthodes pour le thème « Voiture ».	68
Figure 6.4: Précision, Rappel et F mesure des quatre méthodes pour le thème « Livre ».	69

Liste des Tableaux

Tableau 1-1 : Tableau comparatif des différentes mesures [49].	14
Tableau 2-1: Comparaison en les approches Services Web Sémantiques.....	23
Tableau 6-1: Résultats du seuil de similarité pour la requête 1.	62
Tableau 6-2: Résultats du seuil de similarité pour la requête 2.	62
Tableau 6-3: Résultats du thème « Hôpital » pour les requêtes 3 et 4.	63
Tableau 6-4: Résultats de la recherche par thème pour les requêtes 5, 6, 7, 8, 9, 10, 11 et 12.....	66
Tableau 6-5: Résultats de la recherche par langue pour les requêtes 5, 6, 7, 8, 9, 10, 11 et 12.....	67