

통합 개념 모델에 기반한 시맨틱 웹 서비스 탐색

두화준⁰, 이경호

연세대학교 컴퓨터과학과

duhj@icl.yonsei.ac.kr khlee@cs.yonsei.ac.kr

Semantic Web Services Discovery based on the Integrated Concept Model

Hwa-Jun Du⁰, Kyong-Ho Lee

Dept. of Computer Science, Yonsei University

요 약

웹 서비스는 소프트웨어 컴포넌트의 재사용 및 상호 운용을 지원하여 서비스의 통합을 가능케 하는 기술이다. 웹 서비스를 이용하기 위해서는 웹 서비스 탐색 과정이 필수적이다. 최근에는 웹 서비스의 정확한 탐색을 위해 온톨로지를 이용한 시맨틱 웹 서비스 탐색 방법이 제안된다. 그러나 기존 연구는 온톨로지 해석 시 개념 모델에 따라 발생하는 의미적 차이를 고려하지 않는다. 따라서 본 논문에서는 일관된 온톨로지 해석을 통해 탐색 결과의 정확성을 향상시키는 통합 개념 모델에 기반한 시맨틱 웹 서비스 탐색 방법을 제안한다.

1. 서 론

웹 서비스(Web Services)는 프로그래밍 언어, 운영체제에 상관없이 어플리케이션의 상호 운용을 가능케 하는 소프트웨어 기술이다. 웹 서비스는 SOAP, WSDL(Web Services Description Language), UDDI(Universal Description Discovery and Integration)와 같은 표준 인터페이스를 통해 서비스의 재사용 및 서비스 간의 상호 작용을 지원한다. 웹 서비스를 구성하는 요소에는 서비스 제공자, 서비스 요청자, 서비스 레지스트리가 포함된다. 웹 서비스가 제공되는 단계는 서비스 제공자가 웹 서비스를 서비스 레지스트리에 등록하는 출판 단계와 서비스 요청자가 레지스트리를 검색하는 탐색 단계 그리고 서비스 요청자가 탐색된 서비스를 이용하는 요청 단계로 구성된다. 서비스 요청자는 원하는 기능의 적합한 서비스를 검색하는 것이 중요하고, 서비스 제공자는 다양한 서비스 제공을 위한 웹 서비스 조합 시 복합 서비스(Composite Service) 구성을 위한 정확한 탐색이 필요하다. 따라서 웹 서비스를 정확하게 탐색해주는 웹 서비스 탐색 기술은 매우 중요하다.

현재 웹 서비스 탐색 표준에는 UDDI가 있다. 그러나 UDDI는 키워드 기반의 탐색 방식으로 인해 서비스의 구체적인 기능 정보를 통해서 서비스를 탐색할 수 없다. 기존 연구들은 이를 개선하기 하기 위해 온톨로지 기반의 의미적 매칭을 지원하는 시맨틱 웹 서비스 탐색 방법을 제안한다. 시맨틱 웹 서비스 탐색 방법은 온톨로지를 통해 데이터 정보들을 온톨로지의 개념 정보와 매핑하여 서비스 탐색의 정확성을 높인다.

그러나 온톨로지는 해석 방식에 따라 개념 간 관계의 의미가 달라질 수 있다. 즉 동일한 개념이라 해도 서비스의 기능 정보를 다른 의미로 해석 할 수 있다. 기존 연구는 온톨로지 해석 시 발생하는 의미 차이를 고려하지 않으므로 탐색 결과의 정확성을 보장할 수 없다. 또한 기존 연구는 서비스와 개념 정보 매핑 시 온톨로지의 클래스(Class)만을 개념 정보로 사용한다. 그러나 일반적으로 클래스가 가지는 구체적인 데이터 정보는 프로퍼티(Property)가 가지는 데 이를 매핑할 방법을 제공하지 않는다.

본 논문에서는 기존 연구의 개념 간 매칭 시에 발생하는 부정확성을 개선하기 위해 온톨로지의 개념 간 관계를 정립한 통합 개념 모델을 제안하며 이를 통해 매칭의 정확성을 높인다.

※ 이 논문은 2004년도 한국학술진흥재단의 지원에 의하여 연구되었음. (2004-041-D00613)

또한 서비스 기술과 사용자 요구 사항 기술의 유연성을 제공하기 위해 온톨로지 매핑 개념을 프로퍼티로 확장한다.

2. 시맨틱 웹 서비스 탐색

시맨틱 웹: 서비스 탐색은 온톨로지의 개념 정보를 기반으로 서비스를 탐색하는 방법이다. 온톨로지는 특정 도메인의 개념 정보를 클래스로 나타내고 개념의 특징 정보를 프로퍼티로 나타내어 클래스 간의 관계를 구조화 한 것이다. 이를 기술하기 위한 언어로는 DAML+OIL, OWL 등이 있다. 서비스 제공자는 서비스의 데이터 정보들을 온톨로지의 개념 정보와 매핑하여 서비스를 기술하는데 이때 애노테이션된 WSDL, OWL-S 등을 사용할 수 있다. 사용자는 온톨로지를 기반으로 원하는 서비스 기능의 개념을 통해 사용자 요구 사항을 기술하고 서비스를 탐색한다. 탐색 시에는 서비스 제공자가 등록한 온톨로지의 개념과 사용자가 원하는 개념 사이의 매칭을 수행한다. <그림 1>은 시맨틱 웹 서비스 탐색의 예이다.

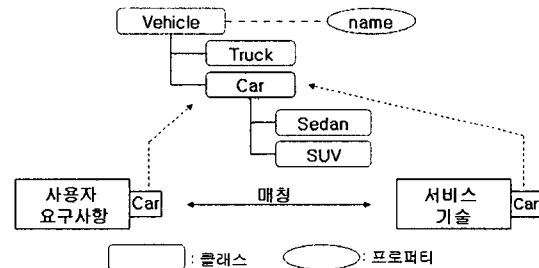


그림 1. 시맨틱 웹 서비스 탐색의 예.

3. 관련연구

시맨틱 웹 서비스 탐색에 관한 기존 연구는 개념 모델에 따라 구분된다. 개념 모델은 개념 간 포함 관계에 대한 온톨로지의 해석 방법에 따라 범위 모델과 상속 모델로 구분할 수 있다. 범위 모델은 현재 개념이 자식 개념을 제공해주어 일반적인 상위 개념이 구체적인 하위 개념의 요청을 만족하는 모델이고, 상속 모델은 하위 개념이 상위 개념으로부터 상속되어 상위 개념의 요청을 하위 개념이 만족하는 모델이다.

Jaeger 등 [1]은 OWL-S로 기술된 서비스와 사용자 요구 사항에 대해 매칭을 수행한다. 서비스 탐색 시에는 입력, 출력,

서비스 카테고리 정보에 대해 상속 모델 기반의 매칭을 수행한다. 기존 매칭 모듈 이외에 사용자 정의 기반의 매칭 모듈을 추가할 수 있으며 이를 기반으로 서비스의 랭킹을 계산한다.

Kawamura 등 [2]은 서비스 탐색 시 필터를 통해 유사 서비스를 찾는다. 필터링은 동일 온톨로지의 사용 여부, 유사 단어 빈도수, 입출력 타입 매칭, 제약 조건 매칭을 통해 수행한다. 입출력 타입 매칭 시에는 범위 모델에 기반하여 유사도를 평가한다. 그러나 매칭 시에 프로퍼티를 고려하지 않는다.

Srinivasan 등 [3]은 OWL-S로 기술된 서비스와 사용자 요구 사항 사이에 범위 모델 기반의 매칭을 수행한다. 또한 OWL-S와 UDDI와의 통합을 위해 OWL-S를 UDDI에 매핑 시키는 방법을 제안한다. 그러나 도메인 정보나 카테고리 정보 등은 고려하지 않으며 개념 간 매칭 시에 프로퍼티 정보도 고려하지 않는다.

Aggarwal 등 [4]은 서비스와 사용자 요구 사항을 도메인, 위치, 오퍼레이션, 입력, 출력, 예외, 선행조건, 사후조건을 통해 기술한다. 서비스 매칭 시에는 두 개념 간의 프로퍼티를 기반으로 공통 부모, 형제, 부모 지식 관계를 이용하여 범위 모델의 매칭을 수행한다. 그러나 개념 간 공통 부모 관계를 고려하기 때문에 의미적으로 연관이 적은 개념이 가까운 형제 위치에 있는 경우에도 매칭이 성립되는 문제점이 있다.

Lin 등 [5]은 탐색 시에 서비스와 사용자 요구 사항 사이의 오퍼레이션, 입력과 출력에 대해서 매칭을 수행한다. 매칭 시에는 상속 모델을 기반으로 클래스 간 포함 관계와 프로퍼티 관계를 통해 유사도를 계산한다.

Sriharee 와 Senivongse [6]는 온톨로지와 서비스에서 추출된 사실 (Fact) 정보를 기반으로 추론을 통해 사용자 요구 사항과 만족되는 서비스 집합을 찾는다. 그러나 사용자 요구 사항과 서비스에 대한 유사도를 계산하지 않고 매칭 여부만을 파악하기 때문에 랭킹 처리를 하지 않는 문제점이 있다.

Cardoso 와 Sheth [7]는 웹 서비스 조합을 위해 워크 플로우에 필요한 서비스를 탐색한다. 탐색 시에는 서비스 명세 간의 유사도, QoS 정보의 만족도, 서비스의 입력과 출력에 대하여 유사도를 계산한다. 매칭 시에는 클래스 간 공통 프로퍼티에 기반한 범위 모델의 매칭을 수행한다.

기존 연구는 온톨로지 개념 간 관계에 대해 한 가지 접근 방법만을 제공한다. 따라서 서비스 제공자가 출판 시에 선택한 개념 모델과 서비스 요청자가 요청 시 선택한 개념 모델이 동일하지 않을 수 있다. 결국 사용자는 동일 개념을 개념 모델에 따라 다른 의미로 기술할 수 있으며 탐색 시 결과의 정확성을 보장할 수 없다. 또한 기존 연구는 서비스와 개념 정보 매핑 시 온톨로지의 클래스만을 개념 정보로 사용한다. 그러나 클래스가 가지는 구체적인 데이터 정보는 프로퍼티가 가지게 된다. 따라서 구체적인 데이터 정보를 가지고 개념 매핑을 수행하려 할 때 프로퍼티가 아닌 프로퍼티가 속한 클래스와 매핑해야 한다. 이 경우 클래스가 다양한 프로퍼티를 가질 수 있으므로 개념이 불분명해지는 문제가 발생한다. 이러한 기존 연구의 문제점을 보완하기 위해서는 온톨로지 해석 시의 문제점을 해결하고 서비스 기술과 사용자 요구 사항 기술의 유연성을 제공하기 위해 온톨로지 매핑 개념을 프로퍼티로 확장할 필요가 있다.

3. 제안된 웹 서비스 탐색 방법

제안된 웹 서비스 탐색 방법은 도메인 온톨로지를 서비스 분류 온톨로지와 서비스 데이터 정보 온톨로지 로 구분하여 OWL로 기술하고 이를 기반으로 OWL-S를 사용해 서비스의 분류 정보와 서비스의 입출력을 기술한다. 탐색 시에는 사용자 요구 사항과 출판된 서비스 사이에 서비스 분류 정보 매칭, 서비스 데이터 정보 매칭을 수행함으로써 만족하는 서비스를 찾는다. 서

비스 분류 정보 매칭, 서비스 데이터 정보 매칭은 개념 간 매칭을 기반으로 하여 제안된 탐색 방법은 <그림 2>와 같다.

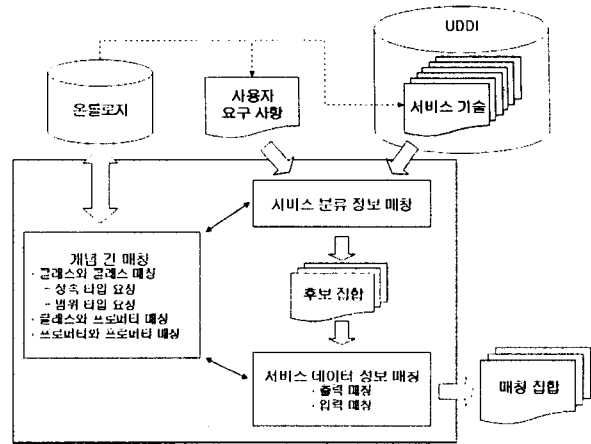


그림 2. 제안된 웹 서비스 탐색 방법.

3.1 개념 간 매칭

개념 간 매칭은 출판된 서비스와 사용자 요구 사항에 매핑된 온톨로지 개념을 매칭하는 것으로 클래스와 클래스, 클래스와 프로퍼티, 프로퍼티와 프로퍼티의 매칭으로 구분한다. 클래스와 클래스 매칭 시에는 통합 개념 모델을 사용한다.

3.1.1 통합 개념 모델

본 논문에서는 일관된 온톨로지 해석을 제공하기 위해 통합 개념 모델을 제안한다. 그러나 기존의 범위 모델과 상속 모델은 각각 개념에 대한 접근 방식과 매칭 방식이 다르기 때문에 개념 모델을 단순히 결합할 수 없다. 따라서 제안된 방법은 개념 모델을 유지하고 매칭의 일관성을 제공하기 위해 하나의 클래스를 상속 타입과 범위 타입으로 정의하여 이를 기반으로 매칭을 수행한다. 상속 타입은 해당 클래스의 제한된 정보만을 제공한다. 단일 개념 집합이며, 범위 타입은 현재 클래스와 자식 클래스를 모두 제공해주는 다중 개념 집합으로서 각각의 원소들에 대해서는 상속 모델 기반의 매칭이 수행된다. 제안된 통합 개념 모델은 아래와 같다.

A를 클래스라고 가정했을 때
 · 상속 타입 : $Inheritance(A) = \{ A \}$
 · 범위 타입 : $Coverage(A) = \{ A \} \cup \{ S_1 \dots S_n \}$
 (단, $S_1 \dots S_n = subClassOf(A)$)

예를 들어 <그림 1>의 Vehicle의 경우 $Inheritance(Vehicle)$ 은 $\{Vehicle\}$ 로서 Vehicle 타입 하나를 갖는 집합이 되고, $Coverage(Vehicle)$ 은 $\{Vehicle, Car, Truck\}$ 의 여러 타입을 갖는 집합이 된다.

제안된 방법에서는 개념 간 매칭 시 요청 개념에 대한 서비스 개념의 매칭 수준을 네 가지로 구분하여 계산한다. 'exact'는 완전히 같은 개념 간의 매칭 정도, 'altertype'은 유사 개념으로서 서비스 개념이 요청의 개념을 대체할 수 있는 매칭 정도이다. 예를 들어 사용자가 <그림 1>의 Vehicle을 요청하고 서비스가 Car로 출판된 경우처럼 상위 개념의 요청을 상속된 하위 개념의 서비스가 제공하는 경우에 해당한다. 'partialtype'은 전체 중 일부의 정보만이 제공되는 낮은 수준의 매칭 정도로서 사용자가 Car를 요청하고 서비스가 Vehicle로 출판된 경우처럼 구체적인 하위 개념의 요청에 대해 일반적인 상위 개념의 서비스가 일부 정보를 제공하는 경우이다. 'fail'은 포함 관계가 존재하지 않아 매칭이 이루어지지 않은 경우이다. 매칭

수준의 유사도는 각각 1, 0.75, 0.25, 0으로 계산한다.

3.1.2 클래스와 클래스의 매칭

클래스와 클래스의 매칭은 요청한 개념과 출판된 개념에 대해 개념 타입에 따라 매칭을 수행한다. 매칭 수준은 네 가지 매칭 수준에 따라 계산하며 요청 개념을 Req, 출판된 개념을 Adv로 가정하고 Req에 대한 Adv의 만족도를 계산한다. 특히 Adv와 Req 사이에 공통 부모 관계는 고려하지 않으며 포함 관계에 있는 경우에 대해서만 매칭 정도를 계산한다.

상속 타입으로 요청하고 상속 타입으로 출판된 경우의 매칭 정도는 다음과 같다. 특히 Adv가 Req를 포함하는 경우의 매칭 정도는 개념 간 거리가 멀어짐에 따라 유사도가 낮아지므로 이를 반영한다.

개념 간 관계	매칭 수준
Req와 Adv가 동일 개념	exact
Adv가 Req의 자식 클래스	altertype
Req가 Adv를 포함	altertype
Req가 Adv의 자식 클래스	partialtype
Adv가 Req를 포함	(partialtype) ^d d= distance(Req,Adv)

상속 타입으로 요청하고 범위 타입으로 출판된 경우의 매칭 정도는 다음과 같다. 특히 Req가 Adv의 자식 클래스인 경우 Adv에 Req의 타입이 포함되어 있으므로 exact로 본다.

Req와 Adv가 동일 개념	exact
Req가 Adv의 자식 클래스	exact
Adv가 Req의 자식 클래스	altertype
Req가 Adv를 포함	altertype
Adv가 Req를 포함	(partialtype) ^{d-1}

범위 타입으로 요청한 경우에는 개념의 집합을 요구한다. 따라서 원소를 각각에 대해 서비스의 개념들과 매칭을 수행한다. 범위 타입의 매칭 정도는 요청한 각 원소들의 매칭 정도의 평균으로 구한다.

Req와 Adv사이 포함 관계가 존재할 때	$\frac{\sum_{n} \max(\text{match}(R_n, Adv))}{n}$ (단, $\forall R_n \in \text{Coverage}(Req), \exists Adv \in \text{AdvList}, n = \text{number}(\text{Coverage}(Req))$)
-------------------------------	--

3.1.3 클래스와 프로퍼티의 매칭

서비스 요청이 클래스이고 출판된 개념이 프로퍼티인 경우는 요청된 개념이 프로퍼티의 정의역인 경우와 치역인 경우로 나누어서 매칭 정도를 계산한다. 특히 Req가 Adv의 정의역인 경우에는 요청한 개념이 가지는 프로퍼티에 대해 하나만 만족하므로 요청한 개념의 프로퍼티의 수로 나눈다.

Req가 Adv의 치역인 경우	exact
Req가 Adv의 정의역인 경우	$\frac{\text{exact}}{n_p}, n_p = \text{number}(Req)$

반대로 서비스 요청이 프로퍼티이고 출판된 개념이 클래스인 경우는 다음과 같다.

Adv가 Req의 정의역인 경우	exact
Adv가 Req의 치역인 경우	altertype

3.1.4 프로퍼티와 프로퍼티의 매칭

서비스 요청과 출판된 개념이 모두 프로퍼티인 경우는 프로퍼티의 포함관계를 고려하여 매칭을 수행한다.

Req와 Adv가 동일 프로퍼티	exact
Adv가 Req의 자식 프로퍼티	altertype
Req가 Adv의 자식 프로퍼티	partialtype

3.2 서비스 분류 정보 매칭

서비스 분류 정보 매칭은 서비스 분류 정보 온톨로지를 기반으로 기술된 서비스 분류 정보와 사용자 요구 사항의 분류 정보 간에 개념 간 매칭을 수행한다. 매칭 시에는 사용자가 요구하는 분류 개념을 서비스가 얼마나 만족시키는 지를 평가한다.

3.3 서비스 데이터 정보 매칭

서비스 데이터 정보 매칭은 사용자 요구 사항과 서비스 사이에 입출력 간 매칭을 수행한다. 입출력 매칭도 개념 간 매칭을 기반으로 한다. 먼저 출력 매칭은 사용자 요구 사항의 출력 리스트와 서비스의 출력 리스트 간에 매칭 정도를 계산한다. 매칭 시에는 사용자가 요구하는 모든 출력 개념을 서비스가 얼마나 만족시키는 지를 평가한다. 입력 매칭은 사용자 요구 사항의 입력 리스트와 서비스의 입력 리스트 간에 매칭 정도를 계산한다. 매칭 시에는 서비스가 필요로 하는 모든 입력 개념을 사용자가 제공해 줄 수 있는 지를 평가한다. 그러나 입력 매칭은 정보를 제공하는 쪽이 사용자가 되고 정보를 제공받는 쪽이 서비스가 되므로 출력 매칭의 반대 개념으로 매칭을 수행한다.

4. 결론

본 논문에서는 개념 간 매칭 시에 정확성을 향상시키기 위해 온톨로지의 개념 모델을 통합한 통합 개념 모델을 제안한다. 또한 서비스 기술의 유연성을 제공하기 위해 온톨로지 매핑 개념을 프로퍼티로 확장한다. 이를 바탕으로 서비스 분류 온톨로지와 서비스 데이터 정보 온톨로지를 기반의 서비스 분류 정보 매칭과 서비스 입출력 매칭을 수행한다. 따라서 제안된 방법은 서비스 기술 시의 유연함과 통합 개념 모델을 통해 의미적으로 정확한 탐색을 제공한다.

5. 참고문헌

[1] Michael C.jaeger, Gregor Rojec-Goldmann, Christoph Liebruth, Gero Muhl and Kurt Geihls, "Ranked Matching for Service Descriptions using OWL-S," Proc. of KIVS, pp. 91-102, 2005.
 [2] Takahiro Kawamura, Tetsuo Hasegawa, Massimo Paolucci and Katia Sycara, "Web Service Lookup: A Matchmaker Experiment," IT Professional Vol. 7, No 2, pp. 36-41, 2005.
 [3] Naveen Srinivasan, Massimo Paolucci and Katia Sycara, "An Efficient Algorithm for OWL-S Based Semantic Search in UDDI," Proc. of SWSWPC(LNCS 3387), pp. 96-110, 2005.
 [4] Rohit Aggarwal, Kunal Verma, John Miller and William Milnor, "Constraint Driven Web Service Composition in METEOR-S," Proc. of Int'l conf. on Services Computing, pp. 23-30, 2004.
 [5] Chuan Lin, Zhaohui Wu, Shuiguang Deng and Li Kuang, "Automatic Service Matching and Service Discovery Based on Ontology," Proc. of Workshops on Grid Cooperative Computing(LNCS 3252), pp. 99-106, 2004.
 [6] Natenapa Sriharee and Twittie Senivongse, "Discovering Web Services Using Behavioural Constraints and Ontology," Proc. of DAIS(LNCS 2893), pp. 248-259, 2003.
 [7] Jorge Cardoso and Amit Sheth, "Semantic E-Workflow Composition," Journal of Intelligent Information Systems, Vol.21, No.3, pp. 191-225, 2003.