

효율적인 서비스 조합을 위한 시맨틱 웹 서비스 탐색

두화준^o, 이경호

연세대학교 컴퓨터과학과

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

Semantic Web Services Discovery for Efficient Service Composition

Hwa-Jun Du^o, 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 서비스 기술을 이용하면 기존에 존재하는 서비스를 통합하여 새로운 서비스를 제공해주는 것이 가능하다. 서비스 조합을 위해 정확한 서비스 탐색은 필수적이다. 본 논문에서는 서비스 조합을 효율적으로 제공하기 위해 오퍼레이션 중심의 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 상호작용을 지원한다. 따라서 웹 서비스는 기존에 구축된 소프트웨어 컴포넌트를 재사용하여 확장하거나 이종의 시스템 간에 통합을 가능케 하여, 시스템 개발에 소요되는 비용과 시간을 줄일 수 있다. 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. <그림 1>은 웹 서비스 구성 요소 간의 관계를 나타낸다.



그림 1. 웹 서비스 구성 요소.

서비스 요청자의 관점에서는 원하는 기능의 적합한 서비스를 검색하는 것이 가장 중요하고, 서비스 제공자의 관점에서는 제공하는 서비스가 정확하게 요청자에게 검색되는 것이 중요하다. 또한 변화하는 사용자의 요구에 맞추기 위해서는 기존의 여러 웹 서비스를 조합한 복합 서비스를 빠르게 제공할 수 있어야 하며, 이를 위해서는 웹 서비스 조합 시에 원하는 기능의 적합한 서비스를 빠르게 검색해야 한다. 따라서 웹 서비스를 빠르고

정확하게 탐색해주는 웹 서비스 탐색 기술은 매우 중요하다.

현재 웹 서비스 탐색 표준에는 UDDI가 있다. UDDI는 비즈니스 정보를 표현하는 구조적인 템플릿을 제공하며 비즈니스 명, 서비스 명, 서비스 카테고리, tModel을 이용한 키워드 기반의 탐색을 제공한다. 그러나 UDDI는 키워드 기반의 탐색 방식으로 인해 서비스의 구체적인 기능이나 서비스의 프로세스 정보를 통해서 서비스를 탐색할 수 없다. 기존 연구들은 이를 개선하기 하기 위해 온톨로지를 통해 의미적 매칭을 지원하는 시맨틱 웹 서비스 탐색 방법을 제안한다. 또한 UDDI는 중앙 집중형의 인덱싱 방식으로 인해 확장성을 제공하지 못하며 심각한 퍼포먼스 병목현상을 발생시킬 수 있다. 따라서 기존 연구들은 분산 레지스트리를 통해 확장성을 제공하고 저장 공간과 처리 속도에서 효율성을 가지는 P2P 탐색 방법을 제안한다. 그러나 기존 연구들은 효율적으로 웹 서비스를 조합하기 위한 탐색 방법에 초점을 맞추고 있지 않다.

본 논문에서는 서비스 조합을 효율적으로 제공하기 위해 입력, 출력, 선행조건 (pre-condition), 사후조건 (post-condition)을 이용한 오퍼레이션 중심의 시맨틱 웹 서비스 탐색 방법을 제안한다. 제안된 방법은 서비스 조합의 효율을 높여주기 위해 오퍼레이션들의 관계를 통해 전체적인 서비스를 파악할 수 있는 오퍼레이션 그래프를 제공하고 오퍼레이션 클러스터를 통해 유용한 오퍼레이션에 대해서는 좀 더 빠른 탐색을 지원하는 장점을 가진다.

2. 관련 연구

웹 서비스 탐색과 관련된 기존 연구들은 의미 정보의 사용 여부에 따라 키워드 기반 검색과 시맨틱 검색으로 구분되고, 네트워크 구성 방식에 따라 중앙집중형과 P2P 방법으로 구분된다.

Schmidt와 Parashar[1]는 부분 키워드와 와일드 카드를 지원하는 P2P 기반의 탐색 방법을 제안한다. 그러

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

나 공통의 키워드를 가지는 서비스를 인접 서비스로 간주하여, 의미적으로 관련이 적은 서비스가 탐색 되는 문제를 포함한다. 예를 들어 computer 서비스와 company 서비스가 관련 서비스로 탐색되는 오류를 범한다.

Zhuge와 Liu [2]는 서비스 간 및 tModel 간의 관계와 유사도를 그래프 형태로 저장하고 유사도를 통해 해당되는 서비스를 검색하고, 검색 결과가 존재하지 않을 경우 기존 UDDI를 통해 검색한 결과를 반환하는 중앙집중형 탐색 방법을 제안한다. 그러나 키워드 기반의 탐색만을 지원하여 다른 키워드를 가지는 유사 서비스를 탐색하기 어렵다.

Emekci 등 [3]은 기존의 서비스 기능 중심의 탐색 방법이 아니라 서비스 행위 중심의 탐색 방법을 제안한다. 서비스 기능 및 프로세스 행위를 기술하기 위해 웹 서비스를 유한 오토마타 형태로 나타내고 이를 정규식으로 변환한 후, P2P 네트워크에 매핑하여 출판하고 탐색한다. 그러나 서비스의 행위가 전혀 의미 정보를 가지고 있지 않기 때문에 서비스 오퍼레이션의 명칭이 동일하지 않은 경우 서비스가 검색되지 않는 단점이 있다.

Verma 등 [4]은 WSDL과 온톨로지의 매칭을 통해 의미 정보가 애노테이션된 WSDL을 생성하고 탐색 시에 서비스 템플릿을 생성하는 P2P 기반의 탐색 방법을 제안한다. 서비스 템플릿은 찾고자 하는 서비스 정보를 기술하는 검색 포맷으로서 입력 정보, 출력 정보와 서비스 명 등을 온톨로지에 존재하는 개념과 매핑 시킨다. 서비스 매칭 시에는 서비스 템플릿과 애노테이션된 WSDL 사이의 개념을 매칭하여 찾는다. 그러나 매칭 시에 두 개의 개념이 완전히 일치하는 경우 만을 탐색하는 한계를 갖는다.

Hausmann 등[5]은 온톨로지를 기반으로 서비스 제공자와 서비스 요청자의 규칙(Rule)을 UML 형태로 기술하고, 규칙 매칭 기법을 사용하는 중앙집중형 탐색 방법을 제안한다. 그러나 탐색 시에 웹 서비스의 입력과 출력 정보를 고려하지 않는 문제가 있다.

Srinivasan 등[6]은 OWL-S로 기술된 웹 서비스와 서비스 요청 사이에 입력 및 출력 개념을 매칭하여 서비스를 탐색하는 중앙집중형 탐색 방법을 제안한다. 그러나 입력과 출력 개념만을 매칭하여 서비스를 탐색하는 것은 동일한 입력과 출력 개념을 가진 다른 기능의 서비스를 탐색할 수 있다는 문제가 있다.

기존 연구들은 효율적으로 웹 서비스를 조합하기 위한 탐색 방법에 초점을 맞추고 있지 않다. 웹 서비스 조합 시에는 오퍼레이션의 구체적인 입력과 출력 정보가 필요하며 해당 오퍼레이션이 수행되기 위해 선행되어야 하는 오퍼레이션 정보와 이후에 수행되어야 하는 오퍼레이션 정보가 반드시 필요하다. 따라서 오퍼레이션을 중심으로 서비스를 조합하기 위한 웹 서비스 탐색 방법이 제안되어야 한다.

3. 제안된 방법

본 논문에서는 서비스 조합을 효율적으로 제공하기 위해 입력, 출력, 선행조건, 사후조건을 이용한 오퍼레이션 중심의 탐색 방법을 제안한다. 제안된 방법은 출판 단계와 탐색 단계의 두 단계로 구성되며 전체적인 시스템의 구

조는 <그림 2>와 같다.

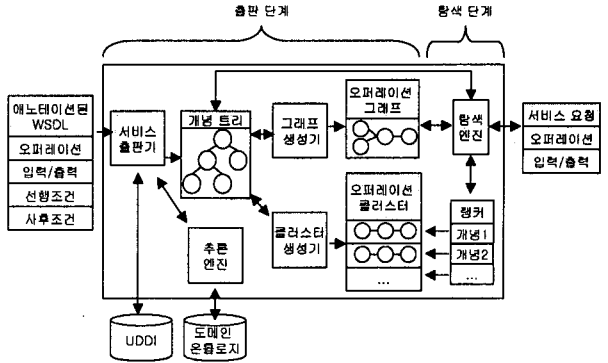


그림 2. 시스템 구조.

3.1 출판 단계

웹 서비스의 시맨틱 탐색을 위해서는 웹 서비스가 의미 정보에 기반하여 출판되어야 한다. 따라서 출판 단계에서는 먼저 의미적으로 기술된 웹 서비스 명세를 생성하는 단계를 거친다. 또한 생성된 웹 서비스 명세를 UDDI에 출판한 이후에는 탐색 시에 필요한 정보를 사전에 추출하는 단계를 수행한다.

3.1.1 의미적으로 애노테이션된 WSDL 파일의 생성

본 논문에서는 의미 정보를 기반으로 웹 서비스를 탐색하기 위해 [7]에서 제안된 의미적으로 애노테이션된 WSDL에 기반한다. 애노테이션된 WSDL에는 도메인 온톨로지의 개념 정보를 토대로 오퍼레이션, 입력, 출력, 선행조건 그리고 사후조건 개념 정보를 저장한다. 입력과 출력은 온톨로지에 존재하는 데이터의 개념과 매핑되고, 선행조건과 사후조건은 사전에 수행되거나 이후에 수행되어야 할 오퍼레이션의 개념과 매핑된다. 서비스 제공자는 애노테이션된 WSDL 형식에 따라 서비스 명세를 생성하고 의미적으로 애노테이션된 WSDL 파일을 서비스 출판기를 통해 UDDI에 등록한다.

3.1.2 개념 트리 생성

서비스 출판 시에 시맨틱 매칭 정보를 추론하여 저장해 놓으면 탐색 시간을 줄이는 것이 가능하다. 이를 위해 온톨로지의 개념과 출판된 서비스와의 매칭 정보를 저장하는 개념 트리를 생성한다. 개념 트리에는 출판된 서비스의 오퍼레이션, 입력, 출력, 선행 조건, 사후조건 개념 정보가 저장된다. 개념 트리 생성 단계는 [6]에서 제안된 방법을 확장한다. 개념 트리에 저장되는 매칭 정보는 두 가지의 개념에 대해서 네 가지 매칭 레벨로 정의한다. 먼저 서비스 요청의 개념을 R, 출판된 서비스의 개념을 A로 가정했을 때, R이 A와 같거나 R이 A의 바로 아래의 하위클래스이면 'exact' 매칭으로 정의하고, R이 A의 손자 클래스이면 'plug-in'으로 정의한다. 또한 R이 A보다 상위클래스인 경우에 'subsume'으로 정의하고, 마지막으로 둘 사이에 관계가 없을 경우에 'fail'로 정의한다. 이러한 매칭 방법에 기반하여 개념 트리의 노드에

는 현재 자신의 개념과 관련이 있는 서비스에 대한 매칭 정도를 <서비스명, 매칭정도> 형태로 저장한다. 예를 들은 톨로지에 'Buy' 개념과 하위 클래스인 'BuyBook' 개념이 존재한다고 가정했을 때, 'BuyBook'의 개념으로 '도서 구매 서비스' 출판하면 개념 트리의 'BuyBook' 노드에는 <도서 구매 서비스, exact>가 저장되고 'Buy' 노드에는 <도서 구매 서비스, subsume>이 저장된다. 또한 개별 노드에는 다중의 서비스가 매칭 될 수 있으므로 매칭 정보를 벡터 방식으로 저장한다. [6]에서는 개념 트리의 노드에 입력과 출력 매칭을 위한 벡터 정보만 유지하지만 본 논문에서는 이를 확장하여 오퍼레이션과 선행조건, 사후조건을 위한 벡터 정보를 추가한다. 저장된 매칭 정보는 탐색 시에 이용된다.

3.1.3 오퍼레이션 그래프 생성

사용자가 서비스 조합을 효율적으로 하기 위해서 오퍼레이션을 탐색할 때에는 현재 사용된 오퍼레이션뿐만 아니라 선행되어 수행되어야 할 오퍼레이션과 이후에 수행되어야 할 오퍼레이션에 대한 정보를 얻는 것이 중요하다. 본 논문에서는 이러한 가정 하에 오퍼레이션 정보를 제공하는 선행조건과 사후조건을 이용한 오퍼레이션 그래프를 생성한다. 오퍼레이션 그래프는 서비스 출판 시에 선행조건과 사후조건 매칭에 의하여 오퍼레이션들의 링크를 형성한 그래프이다. 서비스 탐색 시 사용자가 검색한 오퍼레이션에 해당하는 오퍼레이션 그래프를 제공하여 서비스 조합의 효율을 향상 시킨다. 예를 들어 도서 검색 개념을 중심으로 사용자 로그인 개념의 선행조건과 도서 구매 개념의 사후조건이 존재한다면, <그림 3>과 같이 도서 검색 개념의 선행조건 및 사후조건을 기반으로 각각에 대해 다시 선행조건과 사후조건을 고려하여 오퍼레이션 그래프를 생성할 수 있다.

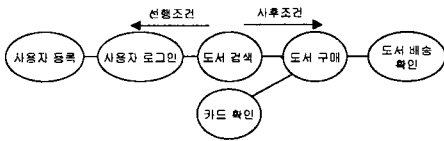


그림 3. 오퍼레이션 그래프.

3.1.4 오퍼레이션 클러스터 생성

자주 검색되는 오퍼레이션 개념은 또한 재검색될 가능성이 높다. 따라서 자주 검색되는 오퍼레이션 개념을 위해 랭킹 정보를 유지하는 오퍼레이션 랭커를 둔다. 또한 랭커에 포함된 개념 중에 유사 오퍼레이션 개념을 가지면서 선행조건과 사후조건, 입력과 출력이 모두 매칭되는 오퍼레이션들에 대해 오퍼레이션 클러스터를 생성한다. 이 작업은 출판 단계에서 자동적으로 수행된다. 만약 출판 시에 새로 출판된 서비스가 랭킹 정보에 포함된 개념이라면 오퍼레이션 집합 중에서 선행조건, 사후조건, 입력 그리고 출력이 같은 집합을 찾아 저장한다.

3.2 탐색 단계

서비스 사용자는 오퍼레이션, 입력, 출력의 개념 정보를 통해 서비스를 탐색한다. 서비스 요청이 들어오면 탐색

엔진은 먼저 오퍼레이션 랭커에서 해당하는 개념이 있는지를 살펴보고 있는 경우에 오퍼레이션 클러스터 집합에서 탐색을 먼저 수행하여 해당 결과를 사용자에게 제공한다. 만약 오퍼레이션 랭커에 해당 개념이 없다면 요청에서 기술된 오퍼레이션 개념과 매칭되는 집합을 개념 트리를 통해 찾는다. 다음으로 해당 집합에서 요청의 출력 개념과 유사한 집합을 찾아내고 이후에 요청의 입력 개념과 유사한 집합을 찾는다. 요청의 입력과 출력 개념 매칭 시에는 개념들을 모두 만족시키는 교집합을 선택한다. 탐색된 오퍼레이션에 대해서는 해당 오퍼레이션 그래프가 제공되어 서비스 조합의 효율을 향상시킨다.

4. 결론

본 논문에서는 서비스 조합을 효율적으로 제공하기 위해 입력, 출력, 선행조건, 사후조건을 이용한 오퍼레이션 중심의 시맨틱 웹 서비스 탐색 방법을 제안하였다. [6]에서는 서비스 탐색 시에 입력과 출력 개념 정보만을 비교하여 동일한 입력과 출력을 갖는 다른 기능의 서비스가 검색결과로 출력되는 문제가 있었으며 서비스 조합에 필요한 정보를 얻기 어려웠다. 이에 반해 본 논문에서는 입력과 출력 정보 뿐만 아니라 선행조건과 사후조건을 추가하여 오퍼레이션 그래프를 통해 서비스 조합의 효율을 높여주고 오퍼레이션 클러스터를 통해 유용한 오퍼레이션에 대해서는 좀 더 빠른 탐색을 지원하는 장점을 갖는다.

5. 참고 문헌

- [1] Cristina Schmidt and Manish Parashar, "A Peer-to-Peer Approach to Web Service Discovery," World Wide Web: Internet and Web Information Systems, 7, pp. 211-229, 2004.
- [2] Hai Zhuge and Jie Liu, "Flexible Retrieval of Web Services," Journal of Systems and Software, Volume 70, Issues 1-2, pp. 107-116, 2004.
- [3] Fatih Emekci, Ozgur D.Sahin, Divyakant Agrawal, Amr El Abbadi, "A Peer-to-Peer framework for Web Service Discovery with Ranking," Web Services, 2004. Proceedings. IEEE International Conference on, 6-9 July 2004 pp. 192-199.
- [4] Kunal Verma, Kaarthik Sivashanmugam, Amit Sheth, Abhijit Patil, Swapna Oundhakar, John Miller, "METEOR-S WSDI : A Scalable P2P Infrastructure of Registries for Semantic Publication and Discovery of Web Services," Journal of Information Technology and Management, Volume 6, Number 1, pp.17-39, 2005.
- [5] Jan Hendrik Hausmann, Reiko Heckel, Marc Lohmann, "Model-based Discovery of Web Services," Proceedings. IEEE International Conference on, 6-9 July 2004 pp. 324-331, 2004.
- [6] Naveen Srinivasan, Massimo Paolucci, Katia Sycara, "An Efficient Algorithm for OWL-S Based Semantic Search in UDDI," SWSWPC2004, LNCS 3387, pp.96-110, 2005.
- [7] Kaarthik Sivashanmugam, Kunal Verma, Amit Sheth, John Miller, "Adding semantics to Web services standards," Proceedings of the International Conference on Web Services(2003), pp.395-401, 2003.