시맨틱 웹 서비스 기반 커뮤니티 정보 검색 시스템

김태환O 전호철 최중민
한양대학교 컴퓨터공학과
{kimthO, hcjeon, jmchoi}@cse.hanyang.ac.kr

Based on Semantic Web Service A Community Information Retrieval System

Taehwan KimO, Hochul Jeon, Joongmin Choi
Department of Computer Science & Engineering
Hanyang University

요 약

웹 기반 자료들이 폭발적으로 증가함에 따라 적합한 자료들에 보다 효과적으로 접근할 수 있는 방법이 요구되고 있다. 이러한 새로운 방법들 중의 하나로 제한 검색이 점보 검색 분야에서 제시되었다. 제한 검색은 현재 입력한 검색어의 검색결과를 줄이고자 할 때 이용하는 검색방식으로 전체 문장을 포함하는 자료나 출판 년도, 특정 저널로 제한하여 검색할 수 있으며 일반적인 검색어로 검색할 경우 제한을 주어 결과물을 최대한 줄일 수 있도록 지원하고 있다. 하지만 이러한 검색 방법은 검색의 범위를 URL에 의해 명시되는 사이트 또는 도메인들로만 제한할 수 있을 뿐이며 의미적으로 관련된 사이트들로 제한할 수 없다.

본 논문에서는 정보의 공유를 목적으로 하는 커뮤니티를 시맨틱 웹 서비스(Semantic Web Services) 기술을 이용하여 플랫폼에 상관없이 사용자 검색 질의와 가장 유사한 커뮤니티를 의미적으로 식별해 내고 커뮤니티 내의 정보 중 질의와 관련된 정보를 검색결과로 도출할 수 있는 구조를 제안한다.

1. 서 론

네트워크 기술의 발달은 인터넷과 월드 와이드 웹(World Wide Web)등의 폭발적인 증가를 가져왔다. 이것은 일반 컴퓨터 사용자들이 다양하고 유용한 정보를 보다 효율적으로 접근할 수 있는 편리함도 제공하지만 아울러 자신이나 기업의 정보를 적은 비용으로 쉽게 제공할 수 있는 환경이 구축되어있다는 더 큰 의미를 부여할 수 있다. 정보 제공 방법이 용이함에 따라 결과적으로 인터넷에서 접근 가능한 정보의 양과 종류의 증가는 사용자로 하여금 자신이 찾고자 하는 정보의 탐색 작업에 많은 시간과 노력을 요구한다. 이러한 문제를 정보과부하[1](information overload)라고 부른다.

정보검색에서 정보 과부하로 인해 나타나는 가장 큰 문제점을 요약하면 다음과 같다. 첫째, 사용자가 원하는 정보와 제공자가 제공한 정보들이 어떤 기준에 의해 단일화된 형태가 아니라 정보를 제공하는 사람의 특색에 따라 정보의 형태가 다양하게 제공된다. 이것은 정보의 형태에 따라 프로그램의 판독을 요구하게 되어 정보를 습득하고 판단하는 시간을 증가시킨다. 둘째, 현재 인터넷에서 사용자가 요구하는 것은 특정 분야에 대한 정보이고 이를 습득하기 위해 주로 검색엔진을 사용한다. 그러나 검색엔진은 실제 사용자가 관심이 있어하는 특정 분야에 대한 정보가 아닌 인터넷 상에 존재하는 모든 정보를 대상으로 사용자가 원하는 정보를 검색한다. 이로 인해 사용자에게 제공되는 정보에 원하지 않는 정보가 포함되게 되어 정보를 선별하는 작업을 한다.

이러한 문제점을 해결하기 위해 질의에 대한 사용자의 의도를 파악하여 특정 분야의 정보만을 검색해주는 웹 에이전트가 필요하다. 특히 정보 형태의 다양화로 인해 문제가 발생하므로 이를 단일화된 형태로 보여줄 수 있어야 한다.

본 논문에서는 시맨틱 웹 서비스를 이용하여 질의와 커뮤니티간의 유사한 특성을 찾아 커뮤니티가 가지는 의미를 사용자의 의도로 부여하고 이를 이용하여 특정 분야에 대해서 검색할 수 있는 구조를 제안한다.

논문의 구성은 2장에서 논문에서 언급된 연구들의 이론적 배경에 대하여 설명한다. 3장에서 OWL을 이용한 커뮤니티 정보 구성 및 처리기법에 대해 기술하고 4장에서 본 논문에서 제시하고 있는 시스템의

구조를 기술한다. 5장에서 검색 사례를 통해 시스템이 어떻게 동작하는지 기술하고 6장에서 향후 연구 방향을 제시하고 결론을 내린다.

2. 관련 연구

본 장에서는 커뮤니티에 대한 정의와 이전에 이용되었던 웹 서비스가 가지는 단점을 기술하고 대안이 되는 시맨틱 웹 서비스에 대해 기술한다.

2.1 커뮤니티

커뮤니티란 원래 지리적으로 근접된 마을, 도시 등 지역사회나 공동체를 뜻하며, 인터넷이 등장하기 이전에도 사회과학에서 중요한 개념으로 자리를 잡아 왔다. 그러나 인터넷의 발달로 인하여 지리적 근접성을 전제로 하였던 전통적인 커뮤니티와는 다른 커뮤니티의 개념이 등장하고 있다. 국내의 경우, 온라인 커뮤니티의 등장은 일반적으로 1990년대 후반이라고 말할 수 있다. 아이러브스쿨과 다음, 그리고 프리챌의 카페 서비스 시작으로 국내 온라인 커뮤니티가 본격적으로 그 모습을 드러냈다. 이와 함께 온라인 커뮤니티에 대한 논의가 본격적으로 시작되었다고 볼 수 있다.

온라인 커뮤니티에 대한 정의는 각 연구자들마다 조금씩 다른 견해를 보이고 있다. 그 중 한 견해는 온라인 커뮤니티는 인터넷 기반의 가상 커뮤니티로서, 인터넷의 기술기반 위에서 공통적 경험과 관심사, 가치를 공유하는 공동체적 모임이라고 말한다[2]. 다른 연구자 또한 온라인 커뮤니티의 설립 목적이 특정한 소재에 관심을 갖는 사람들이 서로의 정보를 교환하고 이와 더불어 서로 만나 친목을 도모하는 것이라고 정의하였다[3,4,5].

이들 연구자의 정의를 정리하면, 온라인 커뮤니티는 '특정한 소재에 관심 있는 사람들이 관심 있는 정보를 공유하기 위한 온라인 상의 공간' 이라고 정의할 수 있을 것이다.

2.2 웹 서비스

웹 서비스는 다양한 하드웨어와 소프트웨어의 문제를 해결하기 위해 SOAP, UDDI, WSDL과 같은 개방형 표준 기술을 기반으로 서비스 제공자(Service Provider), 서비스 저장소(Service Registry) 및 요청자(Service Requester)를 유기적으로 결합한 그림1과 같은 구조를 가진다. UDDI는 웹 서비스의 디렉터리 서비스를 담당하게 되는데 업체가 자사의 웹 서비스를 온라인 디렉터리에 등록 또는 광고하거나 외부에서 웹 서비스를 검색하는데 사용된다. WSDL은 웹 서비스의 서비스를 정의하는 언어로서 프로그램이나 인터페이스 정의를 기술할 때 사용되며 SOAP은 분산된 정보를 교환하는 통신 프로토콜로서 인터넷을 통해 다양한 웹 서비스 사용자가 정보를 교환.전송할 수 있는 기능을 담당한다.

이러한 구조를 가지는 웹 서비스가 진정한 서비스 지향적인 아키텍쳐를 지원하기에는 많은 단점을 지닌다[6]. 현재의 웹 서비스의 브로커인 UDDI는 키워드 기반의 검색을 제공함으로써, QoS나 시맨틱적인 부분까지 검색할 수 없는 단점을 가지고 있다. 이러한 단점들을 해결하기 위해 SWSA[7]는 시맨틱 웹[8]과 웹 서비스 기술을 결합하여 시맨틱 웹 서비스라는 구조를 제안하였다. 시맨틱 웹 서비스의 목적[9]은 웹 서비스를 자연어 수준으로 검색하고 실행을 하며 모니터링, 통합[10]하는 기능까지 지원하는 것이다.

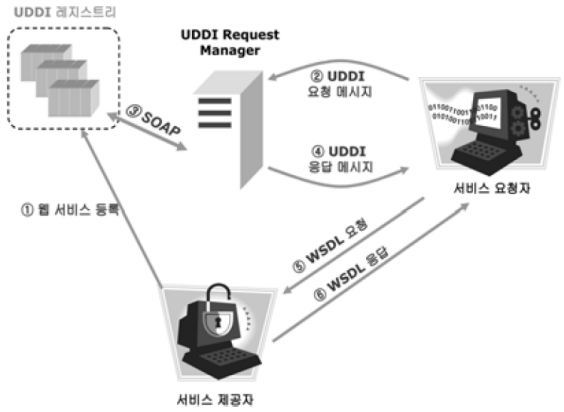

그림 1 웹 서비스 구조

2.3 시맨틱 웹 서비스

시맨틱 웹 서비스는 SOAP, WSDL, UDDI 표준을 기반으로 한 웹 서비스와 RDF, DAML+OIL, OWL 등의 시맨틱 웹 기술을 결합하여 자동화된 웹 서비스의 발견, 실행 및 조합이 가능하다. 즉 시맨틱 웹 서비스는 OWL을 지원하는 에이전트가 WSDL로 기술된 서비스 기술 내용을 읽어서 OWL에 정보를 보내고 OWL은 적당한 온톨로지에 연결하고 이 정보를 검색 엔진에서 제공하게 된다. 장점으로는 RDF와 RDF 스키마 위에서 이루어지므로, 서비스 소비자들이 이러한 프레임워크를 지원하는 어떠한 검색 엔진에서든 자신이 찾는 웹 서비스의 제공자를 찾아 링크할 수 있으며 개인화와 관련된 정보와 협상 능력을 부여할 수도 있다. 결국 에이전트가 웹 서비스 소비자와 제공자의 정보를 갖고 조건을 파악하고, 협상을 통해 계약을 자동으로 이끌어 낼 수도 있게 된다. 현재 대표적 시맨틱 웹 서비스는 W3C(World Wide Web Consortium)의 웹 온톨로지 워킹 그룹에 의해 제안된 OWL을 기반으로 한 OWL-S이다.

이는 2003년 11월에 발표되었으며 2006년 3월에 OWL-S 1.2가 발표되었다. 또한 지능형 웹 서비스를 위한 규칙의 표현과 추론에 관한 표준으로 RuleML[11], SWRL[12]이 현재 진행중이다. OWL-S는 시맨틱 웹 서비스를 구현하기 위한 핵심적인 컴포넌트로 DARPA에 의해 개발된 OWL기반의 서비스 기술을 위한 온톨로지 언어이다.

OWL-S의 온톨로지는 서비스를 기술하기 위해 presents, describeBy, support의 세가지 특성을 구성하고 있으며 이는 "서비스가 하는 일은 무엇인가?"가 의미하는 Service Profile, "서비스는 어떻게 동작하는가"를 나타내는 Service Model, "서비스가 어떻게 접근할 수 있는가?"를 나타내는 Service Grounding의 세 가지 클래스로 그림 2와 같이 구성된다. 즉 OWL-S는 기계가 처리할 수 있는 형태로 서비스를 표현하기 위해 세 가지 클래스를 이용하여 서비스 온톨로지를 기술한다[13]

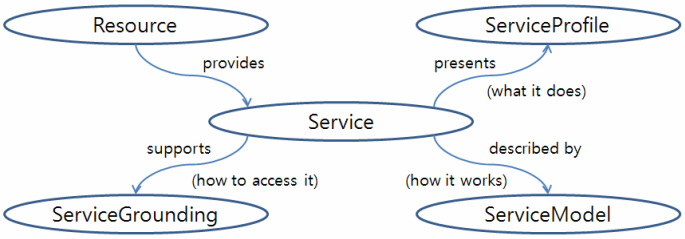

그림 2 시맨틱 웹 서비스 구조

3. OWL을 이용한 커뮤니티 정보 구성 및 처리 기법

정보의 공유를 목적으로 하는 커뮤니티를 시맨틱 웹 서비스(Semantic Web Services) 기술을 이용하여 플랫폼에 상관없이 사용자 검색 질의와 가장 유사한 커뮤니티를 의미적으로 식별해 내고 사용자에게 정보를 제공하기 위하여 정보의 단일화와 서비스 정보 표현기법에 대해 기술하고, 이런 표현 기법을 통해 커뮤니티를 어떻게 분류하고 검색하는지에 대하여 설명한다.

3.1. 정보의 단일화

인터넷의 사용이 보편화되면서 다양한 정보들이 인터넷 상에 분산되어 존재하고 있다. 이러한 정보 자원들은 서로 이질적인 형태로 분산되어 존재하기 때문에, 정보를 얻으려 하는 최종 사용자가 이러한 정보 자원들에 대해 단일화된 형태로 접근할 수 있는 방법이 요구된다. 본 논문에서는 모든 정보의 형태를 HTML 문서로 변환하여 사용자에게 제공한다. 그림 3은 crawler가 문서를 수집할 때 문서안에 링트 되어 있는 파일도 같이 수집하고, 수집된 파일은 문서 파싱을 통해 제목과 내용을 추출한 후 고정된 HTML 플랫폼에 제목과 내용을 적용한 후 사용자에게 결과를 되돌려 준다. 이렇게 파싱된 정보는 웹 서비스의 UDDI를 구축할 때 사용된다.

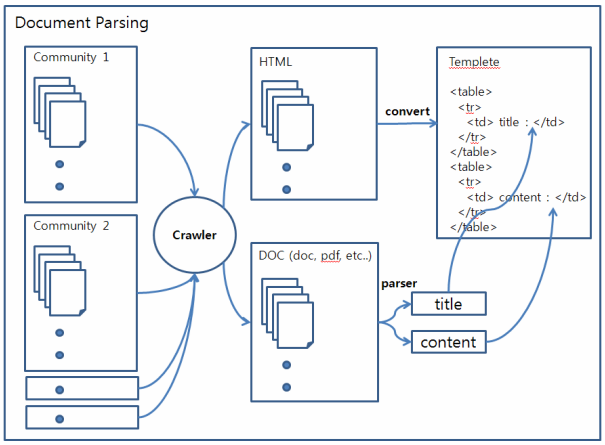

그림 3 정보의 단일화

3.2. 서비스 정보 표현

UDDI는 웹 서비스 정보를 관리하기 위해 White pages, Yellow pages, Green pages를 사용한다. White pages는 비즈니스의 이름, 성명, 연락처, 식별자 정보 등을 가지며, Yellow pages는 표준 분류법에 따른 산업적인 분류 정보를 가진다. Green pages는 웹 서비스 이용에 관련된 기술적 정보를 나타낸다. Green pages는 웹 서비스 이용에 관련된 기술적 정보를 나타낸다. 온톨로지 개념으로 Yellow pages는 정보의 분류를 나타내고있기 때문에 상위 온톨로지라 말할 수 있다. 이는 OWL에서 OWL 스키마에 해당된다. OWL 스키마는 특정 도메인의 상위 개념들 간의 관계및 분류를 정의할 수 있다. White pages와 Green pages는 상위 개념에 따른 하부의 특정 정보를 담고 있기 때문에 OWL에서는 OWL/XML로 표현할 수 있다. 본 논문에서는 일반적인 기업정보에서 표현된 웹 서비스를 커뮤니티에 적용하였다. White pages는 커뮤니티의 관심 주제, URL, 식별자 등의 정보를 가지며, Green pages는 커뮤니티에서 관심있게 다루고 있는 정보에 대한 어휘 정보들을 나타내며, Yellow pages는 커뮤니티의 분류 및 관계를 OWL 스키마에 정의한다.

그림 4는 가상의 정보 저장소에 커뮤니티 서비스 정보를 RDF의 Triple 개념으로 도식화한 것이다. Yellow pages에 해당되는 커뮤니티 분류 및 관계를 OWL schema에 정의하였다. 영화 커뮤니티들을 표현한 것으로 최상위 개념으로 Movie를 두고 그 하위에 영화 제목인 The War, GP506, Candy을 서브클래스로 선언, 커뮤니티 Candy의 Property로 커뮤니티 안에 있는

문서의 정보들을 두었다. OWL/XML Namespace에서는 커뮤니티 내에서 사용되고 있는 문서에서 사용할 속성들에 대해 정의하였고 이를 이용하여 영화 커뮤니티 정보를 구성하였다.

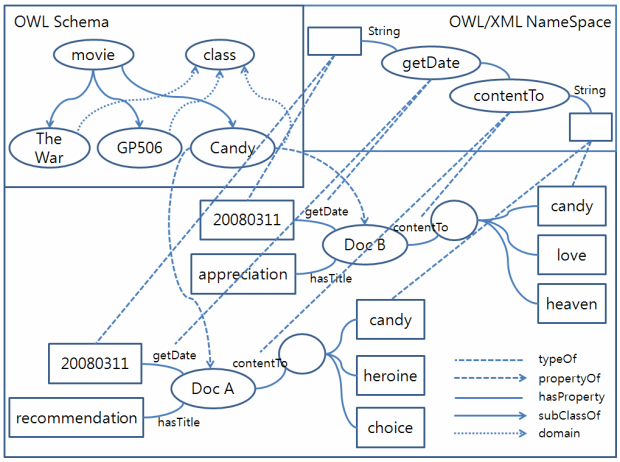

그림 4 영화 커뮤니티의 정보 구성

그림 5는 그림4의 개념을 OWL로 코딩한 커뮤니티 정보의 일부를 나타내고 있다.

```
<?xml version="1.0"?>
<rdf:RDF
    xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
    xmlns:owl="http://www.w3.org/2002/07/owl#"
    xmlns:daml="http://www.daml.org/2001/03/daml+oil#"
    xmlns:dc="http://purl.org/dc/elements/1.1/">
  <owl:Ontology rdf:about="">
    <dc:subject rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >Candy</dc:subject>
    <dc:date rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >February 9, 2007</dc:date>
    <owl:imports rdf:resource="http://purl.org/dc/elements/1.1/"/>
    <owl:imports rdf:resource="http://protege.stanford.edu/plugins/owl/protege"/>
    <dc:creator rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >Neil Armfield</dc:creator>
    <dc:title rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >Candy</dc:title>
    <dc:description rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >candy, love, heaven, heroine, choice</dc:description>
    <owl:versionInfo rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >1.3.1.1</owl:versionInfo>
  </owl:Ontology>
</rdf:RDF>

<!-- Created with Protege (with OWL Plugin 3.2.1, Build 365)  http://protege.stanford.edu -->
```

그림 5 OWL/XML로 표현된 커뮤니티 정보

3.3. 커뮤니티의 분류 및 검색

커뮤니티 기반의 검색시스템은 커뮤니티 내의 정보를 표현하고 있는 온톨로지를 가진다. 커뮤니티 온톨로지는 어떤 정보를 포함하고 있느냐에 따라 웹 서비스들의 서비스 영역에 대해 분류를 담고 있다. 또한 다른 개념 온톨로지에 대한 참조를 통해 각 커뮤니티와 관련이 있는 IOPE(Input Output Precondition Effect) 에 대한 정보를 제공한다. 검색 시스템은 이런 정보를 참조하여 분산 레지스트리를 동적으로 구성하고 질의를 위한 커뮤니티의 선별에 이용한다. 커뮤니티 기반의 시맨틱 웹 서비스 검색 시스템은 먼저 IOPE 와 커뮤니티를 이용한 검색을 통해 정확도의 향상을 목표로 한다. 시스템은 웹 서비스의 커뮤니티가 표현된 온톨로지를 이용해 커뮤니티가 나타내고 있는 정보의 유사도에 따라 웹 서비스를 분류한다. 검색은 커뮤니티를 이용하여 분류된 웹 서비스들에 대해서 이루어진다. 이러한 분류를 통하여 사용자의 요구와 맞지 않은 웹 서비스를 검색의 범위에서 제한함으로 결과의 정확성을 높인다.

커뮤니티 기반의 검색을 위해서는 요청된 서비스에 대한 커뮤니티의 정보를 필요로 한다. 이러한 정보는 사용자의 요청으로부터 또는 요청과 관련성을 가지는 커뮤니티를 추론하여 얻어진다. 커뮤니티 정보가 포함된 질의를 매칭 엔진(Matching Engine)에 요청함으로 관련성을 가지는 커뮤니티의 웹 서비스만을 대상으로 하는 검색이 가능하다.

4. 시스템 구조

제안한 검색 시스템은 그림 6에서 보는 바와 같이, 크게 서비스 인터페이스, 크롤러, 매칭 엔진, OWL-S 레지스트리와 추론부로 구성된다. 서비스 인터페이스는 사용자와의 상호작용과 질의의 생성을 담당한다. 크롤러는 커뮤니티 내에 변화하는 정보를 동적으로 대응하기 위해 커뮤니티에 업데이트 되는 내용을 RSS를 통하여 분석하고 분석된 내용에 맞게 OWL-S 레지스트리에 저장한다. 매칭 엔진은 서비스 인터페이스로부터 받은 질의를 분석하여 레지스트리에 검색을 요청하는 역할을 한다. OWL-S 레지스트리는 분산된 여러 개의 레지스트리로 구성되어 질의에 대한 서비스를 검색하여 결과를 도출한다. 마지막으로 추론부는 입력된 질의와 관련된 커뮤니티 온톨로지를 추론하여 커뮤니티 간의 관계를 생성한다.

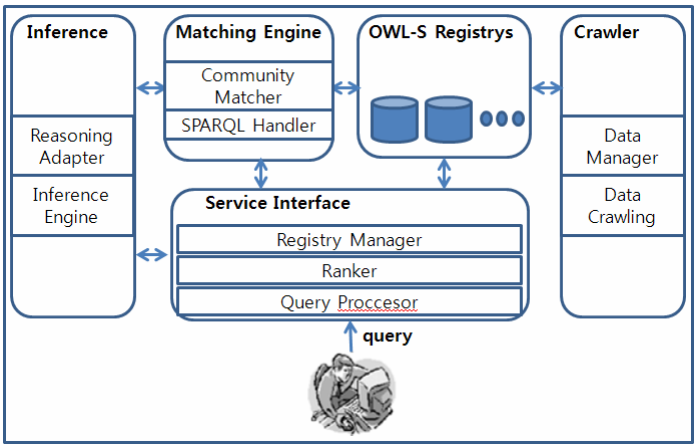

그림 6 전체 시스템 구조

4.1. 서비스 인터페이스

서비스 인터페이스는 사용자로부터 요청을 받고,

요청을 매칭 엔진에 적합한 형태의 질의로 넘겨주는 역할을 한다. 매칭 엔진에는 SPARQL Handler, Community Matcher 두 부분으로 나누며 매칭 엔진과 사용자 상호 작용을 위해 SOAP을 사용한다. 또한 사용자의 질의를 표현하기 위한 질의 프로토콜로는 여러 기술들이 복잡하게 섞여 있는 현재의 시맨틱 웹의 상황에서 표준 온톨로지 쿼리 언어로 선택 될 가능성이 가장 높은 SPARQL을 사용한다. SPARQL의 쿼리는 기본적으로 결과 내역 부분, 데이터 셋 정의 부분, 제약사항 정의 부분으로 구성되어 있다. 결과 내역 부분은 질의의 결과에서 원하는 필드의 내용을 지정 해 주는 부분이다. 이러한 SPARQL의 SELECT 절은 결과 제공할 내용을 변수를 사용하여 나타낸다. 데이터 셋 정의 부분은 정의된 그래프 들을 지정하여 질의의 경로를 알려주는 부분이다. 이러한 SPARQL의 FROM 절과 FROM NAMED 절은 사용할 그래프들을 지정한다. 마지막으로 제약사항 정의 부분은 패턴 매칭 방식의 질의 내용을 정의한 부분이다. 이러한 SPARQL의 WHERE절은 트리플 패턴의 각각에 매칭할 정보들을 자원, 리터럴, 속성등을 이용하여 표현한다. 사용자는 SPARQL 인터페이스를 이용하여 질의를 하고 Community Matcher는 해당 질의에 대한 커뮤니티를 파악하는 역할을 한다. 이는 SPARQL Handler가 사용자의 질의로부터 커뮤니티 정보를 추출할 수 없을 때 호출된다. Community Matcher는 입력 받은 IOPE의 개념과 커뮤니티 온톨로지를 이용해 추론 엔진에 추론을 요청한다. 요청의 결과로 얻어진 커뮤니티를 SPARQL에 추가한다.

4.2. 크롤러

커뮤니티 내에 존재하는 정보를 모두 수집하고 수집된 정보를 단일화 형태로 변환해 준다. 이렇게 변환된 정보는 어휘 분석기와 DATA Parser를 이용해 커뮤니티 별 정보에 대해 OWL-S 레지스트리에 저장한다. 또한 커뮤니티 내에 변화하는 정보를 동적으로 대응하기 위해 커뮤니티에 업데이트 되는 내용을 RSS를 통하여 분석하고 분석된 내용에 맞게 OWL-S 레지스트리에 저장한다. 이렇게 저장된 정보는 매칭 엔진에서 Community Matcher가 사용자 질의 SPARQL에 적합한 Community를 검색할 때 사용된다.

4.3. 매칭 엔진

매칭 엔진은 사용자 질의를 분석/확장하여 레지스트리로부터 웹 서비스 시맨틱 기술(OWL-S)을 검색한다. 이는 크게 Query Processor, Registry Manager, Ranker로 구성된다. Query Processor는 추론 엔진과 레지스트리의 이용을 위한 질의 생성을 담당한다. Query Processor는 각각의 IOPE에 대한 커뮤니티 관계를 추론하기 위한 질의를 생성한다. 질의는 추론 엔진으로 넘겨져 IOPE에 대한 커뮤니티의 관계를 반환한다. 또한 IOPE에 대한 커뮤니티 관계를 이용하여 데이터베이스로부터 웹 서비스 시맨틱 기술(OWL-S)을 검색하기 이한 SQL문을 생성한다. 이러한 SQL문은 Registry Manager에게 전해져 레지스트리로부터 웹 서비스 시맨틱 기술의 검색에 이용된다. Registry Manager는 레지스트리의 관리와 질의를 담당한다. 관리는 커뮤니티 온톨로지사이의 관계를 통해 이루어진다. 또한 Query Process로부터 받은 SQL 문을 실행하여 레지스트리부터 웹 서비스 시맨틱 기술(OWL-S) 을 검색한다. 검색은 커뮤니티를 기준으로 이루어진다. Ranker는 처음 질의와 결과로 얻어진 서비스들이 가지는 의미관계를 통해 유사도를 측정한다. 측정은 사용자 질의의 확장 결과로 도출된 커뮤니티 분류 결과와 레지스트리 검색 결과로 가져온 결과들을 조합하여 이루어진다. 이를 통해 사용자의 질의에 얼마나 유사한 결과를 도출하였는지 판단하는 유사도를 측정하여 결과를 정렬한다.

4.4. OWL-S Registry

OWL-S Registry는 웹 서비스 시맨틱 기술(OWL-S)과 각 기술들의 시맨틱 정보를 저장한다. 각각의 레지스트리에는 OWL-S의 저장경로, 웹 서비스의 경로, OWL-S 파일을 검색하기 위한 IOPE 커뮤니티 정보가 저장된다. OWL-S 레지스트리들은 커뮤니티를 기준으로 분리된다. 각각의 레지스트리는 Registry Manager에 의해 동적으로 분류된다.

4.5. Inference

추론은 Reasoning Adapter와 Inference Engine으로 이루어진다. 추론은 매칭 엔진으로부터 확장쿼리를 입력 받아 관련 온톨로지를 추론하여 관계된 커뮤니티 온톨로지를 정의한다. Reasoning Adapter는 사용자가 입력한 IOPE 데이터타입에 관련도가 높은 커뮤니티 클래스를 추론해 낸다. Inference Engine은 이러한 추론을 도와줄 알고리즘을 제공한다.

5. 검색 사례

사용자가 '테니스'라는 질의를 입력했을 때 테니스에 관한 기본 지식이나 현재 대회가 일어나고 있는 테니스 정보를 검색하기를 원하는데 기존의 Google같은 검색엔진은 이런 자료보다 inlink가 많이 걸려있는 테니스 관련 홈페이지가 더 높은 상위에 링크되게 된다. 이렇게 사용자가 원하는 정보보다 원하지 않는 정보를

검색 결과로 제공해 줌으로써 사용자가 원하는 자료를 찾는 데 더 많은 시간이 걸린다. 이러한 문제를 해결하기 위해 본 논문에서 제안하고 있는 방식은 '특정한 소재에 관심 있는 사람들이 관심 있는 정보를 공유하기 위한 온라인 상의 공간' 이라는 커뮤니티의 특성을 살려 커뮤니티를 수집하고 수집된 커뮤니티가 어떤 소재에 관심이 있는지 판단하고, 판단된 자료를 통해 커뮤니티간의 관계를 OWL-Schema로 정의하는 것이다. 사용자의 질의 '테니스'라는 검색어가 들어오면 커뮤니티 내에 있는 모든 자료를 수집하고 수집된 자료를 가장 많이 포함하고 있는 커뮤니티가 이전 사용자들이 특정한 소재에 관심 있는 정보를 제공했다고 판단하고 그 커뮤니티 뿐 아니라 커뮤니티와 관계를 가지는 커뮤니티 내의 자료도 포함하여 사용자에게 제공해 주는 검색 시스템이다.

6. 향후 과제

커뮤니티에 대한 온톨로지를 구축하여 질의에 대한 의미적 접근을 시도하였다. 하지만 사용자 마다 같은 어휘에 대해서 생각하고 있는 의미가 다를 수 있기 때문에 개인의 성향을 파악할 필요가 있다. 커뮤니티 내에 가입되어있는 회원을 이용하여 개인의 성향을 파악하고 이를 이용하여 커뮤니티 검색에 활용하는 방안에 대한 연구가 필요하다.

참 고 문 헌

[1] P.Maes, "Agents that Reduce Work and Information Overload," Communication of ACM, vol.37, no.7, pp31-40, July 1994.

[2] 서이종, "지식정보사회의 이론과 실제" 서울대학교 출판부, 1998.

[3] Fernback, J., Thompson, B. "Virtual Communities : Abort, Retry, Failure ?, " http ://www.well.com.user/hlr/texts/Vccivil.html, December 1998.

[4] Hagel, Armstrong, "Expanding Markets through Virtual Communities, " Net Gain, 1997

[5] Rheingold, H. "Les communautés virtuelles," Paris : Addison Wesley, 1995.

[6] Mark Turner et al, "Turning Software into a Service," IEEE Computer Society, Oct, 2003.

[7] Semantic Web Service Initiative Architecture Committee(SWSA), http://www.daml.org/services/swsa/, July, 2004.

[8] T.Berenrs-Lee,J. Hendler, and O. Lassila, "The Semantice Web," Scientific American. Vol.284, no.5, pp34~43, 2001.

[9] David Martin, Massimo Paolucci, Sheila McIlraith, Mark Burstein, Drew McDermott, Deborah McGuinness, Bijan Parsia, Terry Payne, Marta Sabou, Monika Solanki, Naveen Srinivasan, Katia Sycara, "Bringing Semantics to Web Services: The OWL-S Approach", Proceedings of the First International Workshop on Semantic Web Services and Web Process Composition (SWSWPC 2004), July 6-9, 2004.

[10] Chris Peltz, Web Services Orchestration and Choreography, IEEE Computer Society, Oct, 2003

[11] OWL-S 1.2 Pre-Release(temporary location at SRI) March, 2006.

[12] "SWRL : A Semantic Web Rule Language Combining OWL and RuleML," http://www.daml.org/2003/11/sw기/

[13] 전종흥, 이원석, 이강찬, 김홍기, "시맨틱 웹서비스 기술 동향 : DAML-S, 한국전자통신연구원 주간기술동향, 제 1123호, pp. 15-28, Nov, 2003.