u-GIS 환경에서 OpenAPI와 매쉬업 가능 서비스에 대한 통합 검색 기법 개발

(Development Integrated Retrieval Methods for OpenAPIs and Mashup Capable Services in u-GIS Environments)

천 동 석* 차 승 준** 김 경 옥*** 이 규 철****
(Dong-Suk Chun) (Seung-Jun Cha) (Kyong-Ok Kim) (Kyu-Chul Lee)

요 약 웹의 양상이 '웹 2.0'으로 변화해감에 따라, '웹 2.0'의 핵심요소인 OpenAPI의 사용이 늘어가고 있다. OpenAPI란 자신들의 사이트에서만 이용할 수 있는 서비스를 외부에서도 이용할 수 있도록 공개한 프로그래밍 인터페이스이다. u-GIS 국토정보도 이러한 OpenAPI를 활용하여 여러 벤더들이 제공하는 서비스를 매쉬업하여 제공할 수 있다. 하지만, OpenAPI는 이미 많이 존재하며 빠르게 증가하기 때문에 사용자가 원하는 서비스를 정확하게 찾는 것은 어렵게 되었다. 또한 여러 개의 OpenAPI를 연결하여 새로운 서비스를 만드는 매쉬업 서비스를 위한 서비스 검색에 어려움이 있다. 본 논문은 이러한 서비스 검색의 문제점을 해결하기 위해서 웹서비스 검색 엔진을 확장하여 통합서비스 정보모델을 정의하고, 정보모델을 바탕으로 통합 검색과 매쉬업 가능 서비스 검색 기법을 개발하였다. 또한 개발한 검색 기법을 관계형 데이터베이스와 JSP를 통해 구현함으로써 유사도 기반의 순위화 된 검색 결과, OpenAPI 통합검색, 카테고리 검색, 매쉬업 가능 서비스 검색을 제공함을 확인하였다.

키워드 : OpenAPI, 매쉬업, 유사도 기반 검색, u-GIS

Abstract As the trend of the Web is changing toward 'Web 2.0', OpenAPIs, Web 2.0's core technology, are used in many web sites. In the past, services in websites are used in its own, but recently it is possible to use services in other websites by using OpenAPI. In u-GIS many vendors also can provide combined service by using OpenAPI. There are already lots of OpenAPIs and the numer of OpenAPI increases very fast. So it is difficult to find a service that we want to use, and also difficult to find services for mashup. In this paper, we developed retrieval methods for OpenAPIs and mashup capable services based on similarity. First we define the integrated service information model to cover various protocols of OpenAPI, then developed a retrieval methods based on it. By implementing system according these methods by using relational database and JSP, we prove that the system can provide an ranked result sets based on similarity, OpenAPI's integration retrieval results and mashup capable service retrieval results.

Keywords : OpenAPI, Mashup, Similarity-based Retrieval, u-GIS

1. 서 론

현재 웹의 양상은 '웹 2.0'으로 변화하고 있다. '웹 2.0'이란 고립된 정보와 기능의 저장소에서 웹서비스의 플랫폼으로 변화를 의미하며, 웹 컨텐츠를 만들고 유통하는 과정에서 사용자들이 적극적으로 참여함을 의미한다.

OpenAPI(Open Application Program Interface)는 이러한 '웹 2.0'의 플랫폼화를 실현하는 핵심요소로, 기업들은 자신들의 사이트에서만 이용할 수 있는 서비스를 외부에서도 이용할 수 있도록 공개한 프로그래밍 인터페이스이다[1]. 현재 구글, 마이크로소프트, 야후, 네이버 등 이미 여러 회사에서 OpenAPI를 제공하고 있다.

더구나 사용자들은 OpenAPI로 제공되는 한가지의 서비스뿐만 아니라, 여러개의 OpenAPI를 기반으로 매쉬업

† 본 연구는 건설교통부 첨단도시기술개발사업 - 지능형국토정보기술혁신 사업과제의 연구비지원(07국토정보C05)에 의해 수행되었습니다.

* 삼성전자 DMC 총괄 연구원, ikarus1004@cnu.ac.kr
** 충남대학교 컴퓨터공학과 박사과정, junii@cnu.ac.kr
*** 한국정보통신연구원, kokim@etri.re.kr
**** 충남대학교 컴퓨터공학과 교수, kclee@cnu.ac.kr(교신저자)

논문접수 : 2009.02.09 수정일 : 1차 2009.03.09 심사완료 : 2009.03.11

서비스를 통하여 새로운 서비스를 창출할 수 있다. 매쉬업(Mashup)이란 '두 개 이상의 OpenAPI를 조합하여 새로운 가치를 만드는 서비스[3]'를 말한다. 매쉬업은 폴 래드머처(Paul Rademacher)가 만든 하우징맵(HousingMaps)[4]을 통해 대중에게 널리 알려지게 되었다.

국토정보 서비스를 제공함에 있어서, 과거에는 하나의 제공업체가 지도뿐만 아니라 관련된 자료들을 모두 저장하고 있어야 했다. 하지만 OpenAPI와 매쉬업이 활성화됨에 따라 관련된 자료를 자신의 사이트에서 정보를 가지고 제공하는 것이 아닌, 필요한 자료는 다른 제공사이트를 통해 얻음으로써 손쉽게 여러 가지의 서비스를 작성할 수 있다. 이것이 u-GIS에서 서비스를 제공하는 중요한 기술이 된다.

이러한 OpenAPI와 매쉬업에 대한 정보를 제공해주는 웹사이트인 Programmableweb을 통해 살펴보면, OpenAPI와 매쉬업에서 이용되는 주된 서비스는 지리정보 서비스이다. 이는 지도 위에 부동산 정보를 표시하고, 쇼핑몰에 채팅을 결합할 수 있으며, 또한 멀티미디어 포탈도 만들 수 있음을 보여준다. 지도정보를 제공하는 대표적인 OpenAPI는 구글 맵 OpenAPI로, 구글 맵은 자바스크립트만을 이용하여 웹에 지도 정보를 표시하는 혁신적인 서비스다. 이 서비스는 AJAX란 용어가 구글 맵을 통해 탄생했을 정도로 대중에게 많은 주목을 받았다. 구글 맵은 웹에 지도를 표시하고 위치 정보를 제어할 수 있는 기능을 OpenAPI로 제공한다. 위치 정보란 실생활에 밀접하게 사용되는 인프라 데이터이기 때문에 지도를 이용한 매쉬업은 가장 활발하게 생산되고 있다.

이미 Programmableweb에 등록된 OpenAPI의 수는 1100개가 넘는다. 이와같이 많고 또 빠르게 증가하고 있는 OpenAPI에서 사용자가 실제 원하는 서비스를 정확하게 찾는 것은 점점 어렵게 되었다. 또한 매쉬업 서비스를 작성할 때도 마찬가지로 검색된 서비스와 매쉬업이 가능한 서비스를 검색하는 것은 어려움이 있다. 따라서 OpenAPI의 효율적인 이용 및 매쉬업을 위하여 서비스 검색이 반드시 필요하다.

Programmableweb[2]에서 다양한 형태의 OpenAPI 검색을 제공하고 있지만, 사이트에서 제공되는 검색 범위는 서비스 이름, 설명정보, 카테고리 정보등 제한적이고, 실제 기능을 수행하는 오퍼레이션 정보는 실제 제공 사이트에서만 확인 가능하다.

마이크로소프트의 대표적인 매쉬업 도구인 Popfly[5]에서도 매쉬업 가능 서비스 검색을 제공하고 있지만, Popfly에서는 제한된 매쉬업 정보(키워드 질의를 통해 검색)만 제공한다는 문제점이 존재하고, 매쉬업 가능 서비스의 위치를 제공하지 않아 사용자가 선택한 결과의 값이 이전 혹은 다음 서비스인지 사용자가 정의를 내려야 한다.

이러한 제약사항을 해결하기 위하여 본 논문에서는 OpenAPI와 매쉬업 가능한 서비스에 대한 검색 기법을 개발하였다. 우선 웹서비스 검색엔진을 바탕으로 통합 서비스 정보 모델을 개발하고, 이 모델을 기반으로 한 통합 OpenAPI 검색과 매쉬업 가능 서비스 검색 기법을 개발하였다. 또한, 순위화 된 결과값을 제공하기 위하여 OpenAPI 등록 정보에 대하여 유사도 기반 검색을 통해 얻어진 유사도를 이용하여 벡터 공간 모델[6]을 통한 유사도 비교 방법을 이용한다. 유사도 측정시 서비스 정보 모델과 같은 계층적인 문서 구조를 검색 결과에 반영할 수 있는 방법을 지원한다. 지원하는 검색 방법은 두 가지로, 통합 OpenAPI 검색과 매쉬업 가능 서비스 검색을 지원한다.

이렇게 개발된 OpenAPI 통합 검색 시스템은 기존의 시스템에서 제공하는 검색 방법보다 정확하며 유사도 기반 검색을 통하여 순위화된 검색을 제공하고, 매쉬업 가능 서비스 검색을 통하여 사용자는 효율적으로 매쉬업 서비스를 구축 할 수 있게 도와준다.

이에 따른 본 논문의 구성은 다음과 같다. 2장에서는 관련연구로 Programmableweb, Popfly, 웹서비스 검색시스템에 대해 살펴본다. 3장에서는 통합서비스 정보 모델에 대해 설명하고, 이를 바탕으로 4장에서는 통합검색 기법을 설명한다. 또한 5장에서는 매쉬업 가능 서비스 검색 기법에 대해 설명한다. 6장에서는 실제 구현된 시스템에 대해 설명하며, 7장에서는 다른 시스템과 비교 / 분석한다. 8장에서는 결론 및 향후 연구에 대해 살펴본다.

2. 관련 연구

2.1 Programmableweb

Programmabelweb은 OpenAPI와 매쉬업과 관련된 웹사이트가 최근 빠른 속도로 늘어남에 따라 해당 정보들을 저장 및 관리해주는 대표적인 웹사이트이다. 이곳에서는 키워드를 기반으로 OpenAPI를 검색하거나, 미리 분류해 놓은 카테고리에서 검색을 제공한다. 또한 이미 열렸거나 최근에 열릴 예정인 매쉬업 콘테스트에 대한 일정도 소개한다[7].

키워드 검색의 결과는 그림 1과 같다. 사용자는 OpenAPI를 검색하기 위하여 키워드를 입력하면 이름, 설명정

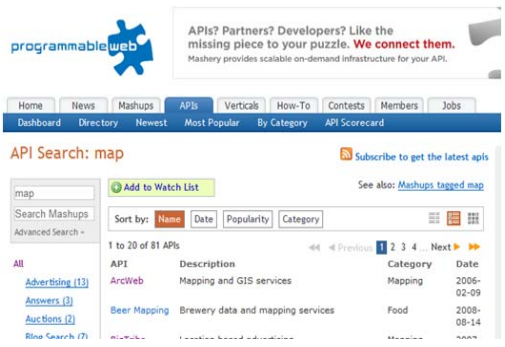


그림 1. 키워드 검색 방법

보, 카테고리에 저장된 내용을 비교하여 매칭이 되는 단어에 대한 검색 결과를 제공한다.

또한, Programmableweb에서는 모두 57개의 카테고리 정보를 제공한다. 각각의 카테고리에는 카테고리명과 관련된 OpenAPI들이 등록되어 있다. 카테고리 검색을 위해 저장되어있는 카테고리를 선택하면 해당 카테고리에 해당되는 OpenAPI 서비스를 제공한다.

2.2 Popfly

마이크로소프트에서는 매쉬업을 손쉽게 할 수 있도록 실버라이트(Silverlight)를 기반으로 하는 Popfly 서비스를 제공한다. Popfly는 OpenAPI를 하나의 블록으로 취급하여 블록간의 연결을 통해 쉽게 OpenAPI들을 매쉬업하는 서비스한다. 이를 위해 OpenAPI 검색과 매쉬업 가능한 서비스 검색을 제공하고 있다.

Popfly에서의 OpenAPI 검색은 모두 11개로 이루어진 카테고리 검색을 기반으로 한다. 각각의 카테고리에는 카테고리명과 관련된 OpenAPI들이 블록 형태로 저장되어 있다. 사용자는 자신이 원하는 OpenAPI를 관련된 카테고리에서 검색할 수 있다.

매쉬업 가능 서비스 검색이란 현재 선택된 서비스와 매쉬업이 가능한 모든 서비스를 검색 하는 것이다. Popfly에서 매쉬업을 하기 위해서는 먼저 사용할 블록을 선택해야 한다.

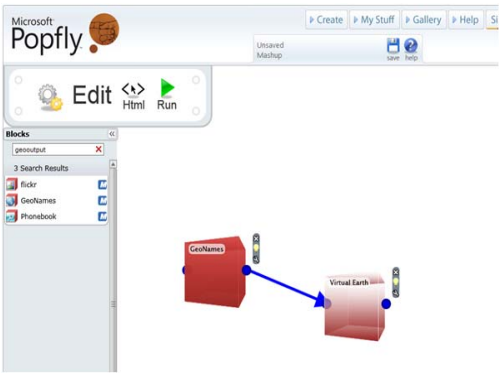


그림 2. Popfly의 매쉬업 가능 서비스 검색

매쉬업 가능 서비스를 검색한 결과는 그림 2와 같다. 선택한 서비스에 대해 매쉬업이 가능한 서비스들이 왼쪽의 검색 결과에 제공되는 것을 확인할 수 있다.

Popfly에서는 OpenAPI를 블록으로 저장할 때 매쉬업을 위한 코드들도 함께 저장한다. 이러한 코드를 바탕으로 매쉬업 가능 서비스를 검색하여 연결할 경우, 새로운 매쉬업의 동작을 바로 확인할 수 있다.

2.3 웹서비스 검색 시스템

키워드 질의를 이용한 순위화 된 웹서비스 검색 시스템은(이하 웹서비스 검색 시스템) 키워드 질의를 이용하여 웹서비스(web services)에 대하여 유사도 연산을 통해 사용자에게 순위화 된 결과를 보여주는 시스템이다[8].

웹서비스 검색 시스템은 관계형 데이터베이스를 이용하여 구현되었으며, 웹서비스 레지스트리인 UDDI[9]와 링크로 연결되는 WSDL[11] 문서를 가져와 XML 형태의 서비스 정보 모델을 구축하고, 이 문서들을 파싱하여 테이블과 이들에 걸쳐지는 인덱스들로 구성된 데이터베이스에 저장하고 SQL 연산을 통해 검색을 수행한다.

웹서비스 검색 시스템의 사용자는 키워드 질의를 통해 순위화 된 서비스 검색 결과를 제공받는다. 서비스 검색에 있어 기존 기법들과 달리 UDDI 레지스트리에 저장된 서비스 설명 정보와 WSDL 서비스 정의 모두를 활용한다. 또한 시스템은 기존 관계형 데이터베이스의 기능을 이용하여 구현됨으로써 의미적인 서비스 발견에 비해 구현이 용이하고, 데이터베이스의 여러 장점을 취할 수 있다.

2.4 기존 연구의 문제점

Programmableweb에서는 키워드와 카테고리 등 키워드 기반을 이용한 OpenAPI 검색을 제공한다. 하지만 서비스 정보와 실질적인 기능을 하는 오퍼레이션 정보에 관하여 검색을 지원하지 않으며. 또한 순위화 된 결과 값을 제공하지 않아 검색된 결과가 사용자가 찾고자 하는 값과 얼마나 유사한지 알 수 없다.

Popfly에서는 키워드 기반을 통하여 매쉬업 가능 서비스를 검색했지만, 검색된 매쉬업 가능 서비스의 위치를 제공하지 않아 이전/다음 서비스인지 구분을 하지 못했다. 또한, 순위화 된 결과 값을 제공하지 않아 검색된 결과가 사용자가 찾고자 하는 값과 얼마나 유사한지 알 수 없다.

웹서비스 검색 시스템에서는 웹서비스를 대상으로 유사도 기반 검색을 이용한 순위화 된 결과값을 지원을 한다.

본 논문에서는 OpenAPI에 대한 통합 검색 및 매쉬업 가능 서비스 검색을 위하여 웹서비스 검색 시스템을 활용 및 확장하였다. 이는 웹서비스도 OpenAPI의 다양한 프로토콜 중 하나(SOAP)에 속해 있으며, 이용된 서비스 정보 모델 역시 다른 OpenAPI를 수용할 수 있기 때문이다. 또한 기본적으로 유사도를 기반한 순위화 된 결과를 지원하기 때문에 사용자들은 찾고자하는 서비스에 대한 자세한 결과값을 제공받을 수 있다.

3. 통합서비스 정보 모델 구축

OpenAPI는 REST, JavaScript, SOAP, RSS, XML-RPC 등의 다양한 프로토콜에 의해 웹 사이트들이 상호작용을 할 수 있도록 하는 기술을 이용하여 작성된다. 각각의 기술의 통신 방식이나 메시지가 다르지만 본 논문에서는 이러한 다양한 기술에 모두 적용할 수 있는 통합 서비스 정보모델을 제안하였다. Programmableweb에 의하면 REST, JavaScript, SOAP 방식을 이용하는

OpenAPI들이 전체의 95%임을 확인 할 수 있다. 따라서 본 논문에서는 이러한 각기 다른 형태의 OpenAPI에 대한 통합 검색을 위하여 REST, JavaScript, SOAP 에 대해 살펴보고 검색을 위한 통합 서비스 정보 모델에 필요한 기본 요소들을 추출하였다.

REST는 HTML 기반 프로토콜이다. 요청 URL과 각각의 요청 변수를 HTTP 위에서 전달하면 결과 값을 응답 변수로 이루어져 있는 XML 형태로 사용자에게 전달한다. REST 기반의 프로토콜을 이용하는 대표적인 OpenAPI는 대표적으로 Naver 검색 OpenAPI, flickr, Google chart, Google Notebook등이 있다.

JavaScript는 벤더들이 데이터를 접근하기 위한 방법인 JavaScript 라이브러리를 제공하여 사용자들은 라이브러리에서 제공하는 함수를 이용하여 클라이언트에서 구현하는 방식이다. JavaScript 라이브러리는 클래스와 메소드로 구성되어있다. 클래스는 각각의 메소드를 포함하고 있고, 사용자는 실제 기능을 하는 메소드를 호출하고 메소드의 리턴 값이 사용자에게 전달된다. JavaScript 기반의 프로토콜을 이용하는 대표적인 OpenAPI는 Google Maps, Naver Maps등이 있다.

SOAP은 단순한 정보 전달용 XML 메시지로 구성되고, 다른 분산 컴퓨팅 프로토콜(CORBA, RMI, DCOM)처럼 원격 프로시저(Remote Procedure Call)을 하는 XML 메시지로 구성된다[12]. SOAP은 요청 파라미터를 포함하는 요청 메시지를 전송하면 결과 값을 XML 형태의 응답 메시지로 사용자에게 전달한다. SOAP 기반의 프로토콜을 이용하는 대표적인 OpenAPI는 Amazon EC2, Amazon eCommerce등이 있다.

웹서비스 검색 시스템에서의 정보모델은 UDDI에 저장된 서비스 정보와, WSDL에 저장된 오퍼레이션의 정보를 통해 XML 문서 형식으로 구성된다. UDDI에서 BusinessEntity, BusinessService 정보를 저장하고 해당서비스에 대한 오퍼레이션 정보를 WSDL을 통해 읽어와 Service, Port, Operation 정보를 저장한다.

OpenAPI의 SOAP 프로토콜은 웹서비스 검색 시스템에 저장되는 내용과 동일하다. 따라서 OpenAPI의 검색을 위해 필요한 내용으로는 회사정보, 서비스정보, 분류정보, 오퍼레이션 정보로 나눌 수 는데 이러한 정보는 그림 3과 같이 웹서비스 정보 모델을 이용해 저장할 수 있다.

회사 정보에서는 OpenAPI를 제공하는 회사 이름, 회사의 설명등의 정보들로, 웹서비스 검색 시스템에서의 BusinessEntity에 이름, 설명정보를 기술한다. 서비스 정보는 BusinessEntity에서 기술한 회사에서 제공하는 OpenAPI 서비스의 이름, OpenAPI 서비스의 설명, 서비스의 위치 정보이다. 이는 웹서비스 검색 시스템에서의 BusinissService, Service 부분에 이름과 설명정보로 저장된다. 분류정보는 유사한 오퍼레이션들에 대한 분류정보를 기술한 것으로 웹서비스 검색 시스템의 Port에 저장된다. OpenAPI에 사용되는 입력, 출력 인자들은 매쉬업 가능 서비스 검색을 위한 필수적인 요소로 웹서비스

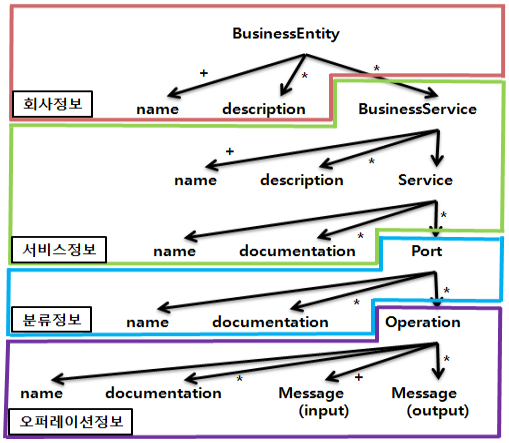


그림 3. 통합 서비스 정보 모델

검색 시스템의 Operation에 저장된다. Operation에서는 오퍼레이션의 이름과 설명정보 및 입력, 출력 인자들에 대해 저장한다.

예를 들어, Naver Maps OpenAPI를 통합 서비스 정보 모델에 적용하면 표 1과 같다.

표 1. OpenAPI를 통합 서비스 정보 모델에 적용 예

엘리먼트	내 용
B.E/name	Naver
B.E/description	NHN(주)에서 제공하는 인터넷 포털사이트로, 1999년 6월에 시작되었다. 일반적인 통합검색 서비스는 물론이고, 사전, 지도, 동영상, 이미지 검색 등 다양한 서비스를 제공하고 있다.
B.S/name(S/name)	Naver maps OpenAPI
B.S/description (S/documentation)	네이버 지도 API를 이용하시면 여러분의 웹 사이트에 지도를 표시할 수 있으며, 지도 상의 원하는 위치에 정보창을 통해서 관련 정보를 표시할 수 있습니다.
P/name	NMap
P/documentation	새로운 지도 객체를 생성한다.
O/name	setCenter
O/documentation	현재 지도의 중앙점(center point)을 설정한다.
O/input	point
O/output	none

웹서비스 검색 시스템의 서비스 정보 모델의 요소와 본 논문의 서비스 정보 모델의 기본 요소는 대부분 일치하지만 웹서비스 검색 시스템의 서비스 정보 모델은

OpenAPI의 카테고리 정보, OpenAPI 서비스의 URL, 서비스에서 사용되는 프로토콜(Protocol)을 저장하는 부분이 존재하지 않았다. 이를 위해 category, accesspoint와 protocol를 추가하였다.

category는 카테고리내 서비스를 검색하기 위해 필요한 정보로 기존 웹서비스 검색시스템에서는 존재하지 않으며, OpenAPI 통합 정보 모델에 함께 저장된다. accesspoint는 각 서비스별로 저장되어야 하는 서비스에 접근할 수 있는 주소로 서비스의 URL 정보가 저장된다. 이는 실제 서비스 정보가 저장되는 BusinessService의 하위 요소로 저장된다. protocol 정보는 앞에서 분류한 OpenAPI의 프로토콜에 대한 정보이다. 그래서 오퍼레이션의 카테고리를 저장하는 Port부분에 추가하였다.

위와 같이 추가한 정보를 포함하여 수정된 통합 서비스 정보 모델은 그림 4와 같다. 모델의 수정을 통하여 본 논문의 서비스 정보 모델은 OpenAPI의 카테고리 정보, URL 정보와 기능을 수행하는 오퍼레이션의 프로토콜 정보를 제공할 수 있게 되었다.

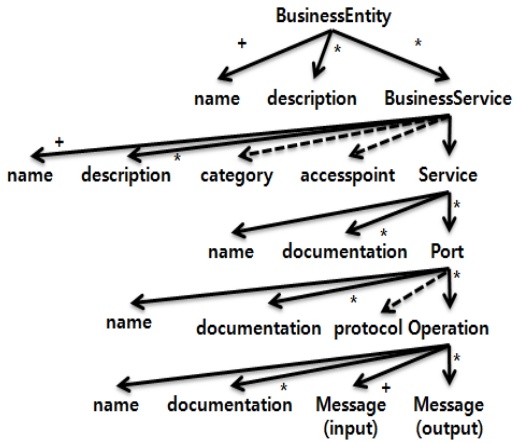


그림 4. 확장된 통합 서비스 정보 모델

4. 통합 검색 기법 개발

통합 검색 기법이란 제시한 통합 서비스 정보 모델을 바탕으로 키워드를 기반으로 OpenAPI를 검색하는 것이다. 통합 서비스 정보 모델은 웹서비스 검색 시스템의 정보모델을 확장하였기 때문에 통합 검색 기법에서는 역시 웹서비스 검색 시스템에서 제공하는 키워드 검색 기법을 확장하였다.

기본적으로 통합 서비스 정보모델은 XML의 구조화된 문서로 저장되기 때문에 질의어에 대한 유사도를 검색하기 위해서 정보검색에서 가장 일반적인 가중치 부여 기법인 TF*IDF를 수정한 TF*IEF를 이용한다[8]. 구해진 가중치를 바탕으로 VSM[6]을 이용해 서비스 유사도를 측정하여 사용자에게 순위화된 검색 결과를 제공한다.

키워드 검색 기법은 크게 특정 범위 검색과 확장 범위 검색으로 나뉜다. 특정 범위 검색은 사용자가 설정한 범위(Business, Service, Operation) 안에서 입력된 키워드를 바탕으로 검색한다. 확장 범위 검색은 사용자가 설정한 범위 뿐만 아니라 상/하위로 전파된 유사도를 바탕으로 검색한다.

또한, OpenAPI의 특성에 맞추어 좀 더 자세한 검색을 위하여 가중치 부여 검색과 카테고리 검색을 추가하였다.

4.1 가중치 부여 검색

웹서비스 검색 시스템에서 제공하는 확장 범위 검색은 입력된 키워드에 대해 사용자가 설정한 범위 뿐만 아니라 상/하위로 전파된 유사도를 바탕으로 검색을 한다. 예를 들어, 사용자가 Service을 선택하였을 경우 검색 대상을 BusinessService뿐만 아니라 BusinessEntity, Operation에 대해서도 전파된 유사도를 바탕으로 검색을 한다.

하지만, 확장 범위 검색은 상/하위 전파도는 기본값으로 같이 고정되어 있다. 검색어와 검색결과에 따라 전파해야 하는 범위가 다를 수 있다. 예를 들면 사용자는 'map'이라는 단어를 포함하는 서비스를 찾고 싶을 때, 실제로 business의 값으로는 네이버, 구글, 야후와 같이 관련이 적기 때문에 더 적은 전파도를 가져야 할 것이고, 오히려 Operation 부분에 더 많은 전파도가 필요하다. 따라서 본 논문에서는 전파도를 사용자가 수정할 수 있도록 확장하였다.

4.2 카테고리내 검색

카테고리란 사물의 개념을 분류할 때 그 이상 일반화할 수 없는 가장 보편적인 단어를 말한다. 즉, OpenAPI를 표현하는데 있어서 가장 기본적이고, 근본적인 정보가 될 수 있다. 예를 들어 사용자가 맵에 관련된 정보를 찾고 싶지만 아무런 정보도 없을 경우 맵에 관련된 서비스를 찾는 것은 어렵다. 하지만 카테고리 검색을 제공함으로써 맵과 관련된 서비스의 정보를 얻을 수 있다.

본 논문에서 OpenAPI에 대한 카테고리는 기준은 Programmableweb의 기준을 준용한다. 현재 Programmableweb에서는 이미 57개의 카테고리를 제공을 한다.

OpenAPI의 카테고리 정보는 OpenAPI 메타정보를 입력할 때 사용자의 입력을 통하여 통합서비스 정보 모델의 /BusinessEntity/BusinessService/category에 저장이 된다. 사용자는 저장된 정보를 통하여 사용자는 카테고리내에 OpenAPI 검색이 가능하다.

5. 매쉬업 가능 서비스 검색 기법 개발

5.1 이전 매쉬업 가능 서비스 검색

이전 매쉬업 가능한 서비스 검색은 현재 선택된 서비스의 이전에 매쉬업 될 수 있는 서비스를 찾아주는 것이다. 이전 매쉬업 가능한 서비스(Input값을 제공하는 서비스)를 검색하기 위해서는 선택한 서비스의 입력값을 출력값으로 가지고 있으며 입력받은 키워드를 포함하는 유사한 기반 서비스 검색 방법을 이용한다.

5.2 다음 매쉬업 가능 서비스 검색

다음 매쉬업 가능한 서비스 검색은 현재 선택된 서비스의 다음에 매쉬업 될 수 있는 서비스이다. 다음 서비스(Output값을 제공받는 서비스)를 검색하기 위해서는 선택한 서비스의 출력값을 입력값으로 가지고 있으며 입력받은 키워드를 포함하는 유사도 기반 서비스 검색 방법을 이용한다.

5.3 교차 매쉬업 가능 서비스 검색

교차 매쉬업 가능한 서비스 검색은 매쉬업이 가능하지 않은 두 개의 서비스를 선택했을 경우 두 개의 서비스 사이에서 두 서비스와 매쉬업이 가능한 서비스를 찾는 것이다. 처음에 오는 서비스를 전행 서비스라 하고 마지막으로 오는 서비스를 후행 서비스라 한다.

교차 매쉬업 가능한 서비스 검색은 그림 5와 같이 전행 서비스의 출력값을 입력값으로 가지고 후행 서비스의 입력값을 출력값으로 가지는 서비스를 검색한다.

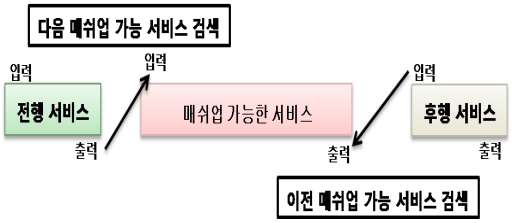


그림 5. 교차 매쉬업 가능 서비스 검색

6. 시스템 구현

본 논문에서 개발한 시스템은 웹서비스 검색 시스템을 수정 / 보완하여, 다양한 OpenAPI 프로토콜에 적용 가능하도록 개발하였다.

6.1 시스템 구조

본 논문에서 구현한 시스템 구조는 그림 6과 같다. 통합 서비스 정보 모델의 저장소는 데이터베이스 테이블들과 이들에 걸쳐지는 인덱스들로 구성된다. OpenAPI에 대한 데이터베이스 구축을 위해 저장되어있는 OpenAPI 정보 모델을 바탕으로 데이터베이스에 저장한다. 또한 UDDI에 저장된 내용을 바탕으로 웹서비스 검색 시스템에 필요한 내용을 생성한다.

저장된 통합 서비스 정보 모델은 유사도를 계산하기 용이하게 서비스 저장소에 저장된다. 서비스 저장소에 저장된 내용을 바탕으로 사용자는 질의를 하게 되고 순위화된 결과를 제공한다.

6.2 저장 구조

그림 7은 본 연구에서 서비스 정보 모델 저장을 위한 데이터베이스 테이블 구조이다. 본 연구에서는 단어-기반 역인덱스를 데이터베이스 테이블로 구축한 후, 각 단어에 대한 n-그램[13]들을 별도의 테이블로 구성하고, 이들 n-그램들과 용어 테이블간의 동등-조인을 통해 부분 문자열 검색을 지원하였다. 이는 n-그램 자체를 용어로 인식하고 n-그램을 기준으로 역인덱스를 구축하고 가중치를 부여하는 [14]의 연구와는 다르게 n-그램에서 n의 크기는 용어 가중치의 계산에는 영향을 미치지 않는다. 다만 검색 문자열의 최소 길이는 |n|으로 제한된다.

Lexicon과 Posting, Ngrams, NtoL 테이블은 각각 단어의 정보를 저장하고, Maxfreq와 NumOfType에는 TF*IEF 계산에 필요한 정보를 저장하고, Link 테이블에는 OpenAPI에 대한 정보를 저장한다.

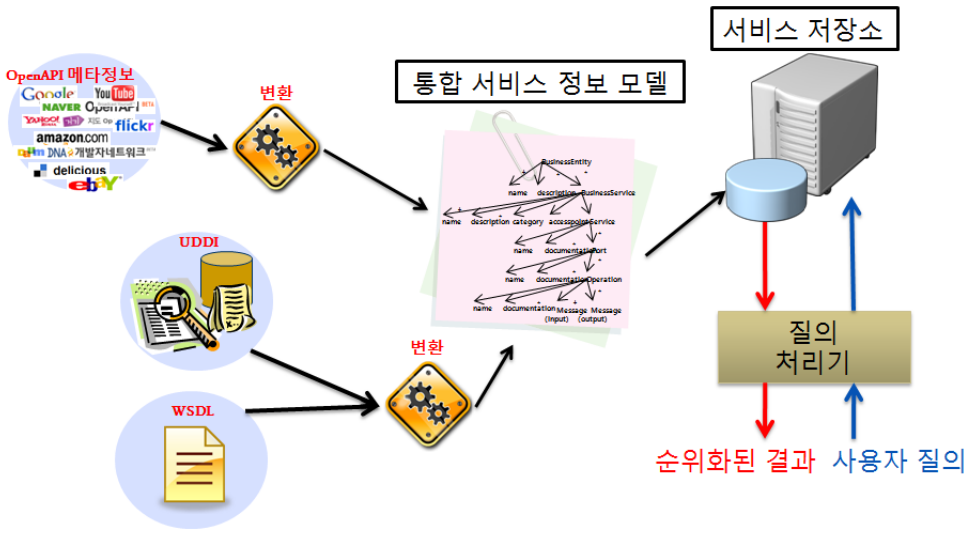


그림 6. 시스템 구조

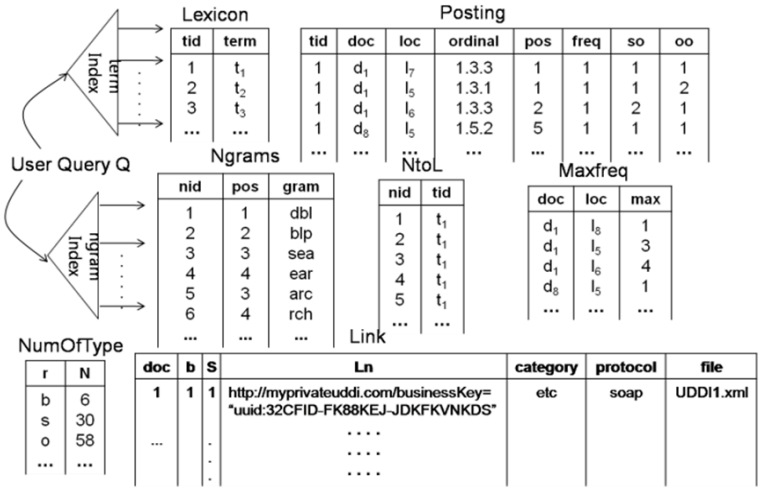


그림 7. 저장 구조

6.3 구현 환경

본 논문에서 개발한 검색 시스템의 구현 환경은 다음과 같다.

· 운영체제 : Fedora release 8
· 웹 서버 : Tomcat 5.0, Apache, JDK 1.6
· 데이터베이스 : MySQL 5.0.51
· 개발 언어 : Java, JSP

운영체제로는 리눅스인 Fedora 8을 사용하였고, 데이터베이스는 MySQL 5를 이용하였다. 본 시스템은 자바로 구현하였기 때문에, 웹에서 JSP로 동작하였고, 서블릿으로는 Tomcat 5를 이용하였다.

6.4 사용자 인터페이스

6.4.1 OpenAPI 입력

그림 8은 OpenAPI 통합 정보 모델에 기반하여 데이터를 입력하는 화면이다. 데이터 입력 폼은 업체정보-서비스정보-클래스정보-메소드정보의 계층형 구조로 이루어져 있다. 각각은 추가버튼을 통해 하위의 정보를 입력할 수 있고 또한 삭제 버튼을 통해 삭제할 수 있다. 이렇게 정보를 입력한 후 입력완료 버튼을 누르게 되면 입력한 정보들이 XML 형식의 설명정보에 맞춰 서버에 저장되게 된다.

그림 8. OpenAPI 입력폼

6.4.2 특정/확장 범위 검색

그림 9와 같이 특정 범위 검색은 사용자가 직접 검색 범위(Business, Service, Operation)를 지정하여 사용자가 검색어를 입력하면 지정한 검색 범위 내에서 검색이 이루어진다.

그림 9. 특정 범위 검색 결과

확장 범위 검색의 검색 범위는 특정 범위 검색처럼 Business, Service, Operation을 선택할 수 있지만 선택

된 범위만을 검색하는 것이 아닌 상/하위 개념 간 서비스 유사도의 전파를 고려하여 검색한다. 즉, Service를 선택할 경우, 상/하위 유사도 전파를 통해 해당 검색어를 가진 Business와 Operation을 모두 검색한다. 또한 상/하위 유사도의 가중치를 사용자가 임의로 지정할 수 있어 사용자가 더 중요하다고 생각하는 곳에 가중치를 더 부여 할 수 있다.

특정/확장 범위 검색으로 검색된 결과는 유사도를 기반으로 순위화 되어 출력되며, "문서 구조 보기"를 통해 매쉬업 검색에 사용된다.

6.4.3 이전/다음 조합 가능 서비스 검색

특정/확장 범위 검색의 검색된 결과에서 선택한 문서의 구조를 트리형식으로 확인하여 사용자는 Operation의 정보를 얻을 수 있다. 이 중 검색의 대상이 되는 Operation을 선택하여 매쉬업 가능 서비스를 검색할 수 있다. 이전 매쉬업 검색 서비스일 경우에는 선택한 서비스의 입력인자가 찾고자 하는 서비스의 출력인자와 일치하고 추가 검색어를 가지고 있는 서비스를 검색하며, 다음 매쉬업 가능 서비스일 경우에는 선택한 서비스의 출력인자가 찾고자 하는 서비스의 입력인자와 일치하고 추가 검색어를 가지는 서비스를 검색한다. 이렇게 검색을 수행한 결과는 그림 10과 같이 출력된다. 검색된 결과와 다시 매쉬업 될 수 있는 서비스 검색도 가능하다.

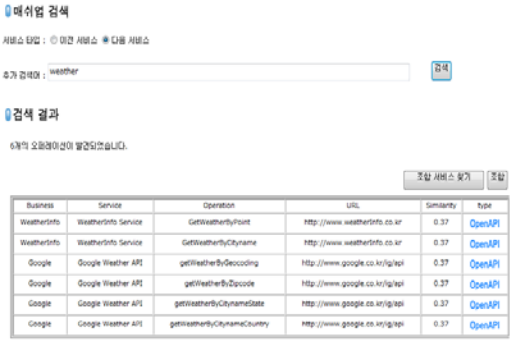


그림 10. 다음 매쉬업 가능 서비스 검색 결과

6.4.4 교차 조합 가능 서비스 검색

교차 매쉬업 가능한 서비스는 서비스와 서비스간에 매쉬업을 직접 할 수 없는 경우에 서비스 간에 다리가 되어 줄 수 있는 서비스를 찾는 것이다. 이를 위해 사용자는 전행 서비스와 후행 서비스에 대하여 특정 범위 검색을 한다. 2개의 서비스에 대하여 검색한 후 각각의 서비스에 대하여 문서 구조 보기를 통하여 실제 기능을 하는 오퍼레이션을 선택한 후 등록을 한다.

두 개의 서비스를 등록 한 뒤 교차 가능한 서비스 검색 버튼을 누르면 그림 11과 같이 전행 서비스와 후행 서비스 사이에 매쉬업 될 수 있는 서비스들을 검색한다.

그림 11. 교차 매쉬업 가능 서비스 검색 결과

7. 타 시스템과 비교

표 2는 기존 시스템과 비교 분석 결과이다. 기능에 있어서의 분석 기준은 검색 범위, 유사도 검색, 매쉬업 검색, 특정/확장 범위 검색의 유무의 기준을 가지고 비교하였다.

표 2. 타 시스템과의 비교

	Programmableweb	Popfly	개발한 시스템
유사도 검색	X	X	O
매쉬업 가능 서비스	X	O	O
카테고리 검색	O	O	O
특정범위 검색	X	X	O

결과를 분석해보면, 검색 범위의 경우 Programmableweb에서는 OpenAPI 검색에 대하여 키워드 질의와 카테고리 검색을 지원한다.

Popfly에서는 OpenAPI 검색 경우 카테고리 검색을 통하여 제공하였고, 미리 저장해둔 키워드를 통하여 매쉬업 가능 서비스 검색을 제공하였다.

본 시스템에서는 OpenAPI, 매쉬업 검색 모두 키워드 질의를 통하여 검색이 가능하고 유사도 검색을 지원하여 사용자에게 순위화 된 결과값을 제공하고 특정 범위 검색을 지원하여 사용자가 원하는 범위의 검색을 가능하게 해준다.

8. 결론 및 향후 연구

본 논문에서는 u-GIS에서 사용될 중요한 기술의 하나인 OpenAPI와 매쉬업 서비스를 활성화하기 위해, 이것들에 대한 검색기법을 개발하였다.

OpenAPI와 매쉬업은 웹 2.0에서의 중요한 기술로 OpenAPI를 이용 및 조합하여 창의적인 사이트를 구축할 수 있고, 수익을 창출할 수 있어 새로운 비즈니스 모델로 각광받고 있다. 특히 지리정보 서비스를 제공해주는 대표적인 업체인 구글, 야후, 네이버 등은 자신들의 서비스를 이미 OpenAPI를 통해 제공하여 사용자들은 자신들의 홈

페이지에 여러 지리 정보를 제공하는 서비스를 제공하고 있다.

다양한 종류의 OpenAPI 및 매쉬업 서비스의 활용을 위해 OpenAPI 검색 및 매쉬업 가능 서비스 검색이 필요하다. 현재는 다양한 웹 기술 사용하는 수많은 OpenAPI들 중에서 자신이 원하는 OpenAPI 검색을 위해서는 찾고자 하는 OpenAPI의 사전 지식이 필요하다. 또한 OpenAPI들 간에 매쉬업을 하기 위해서는 OpenAPI들간에 입/출력 정보를 알아야 가능하다. 본 논문에서는 이러한 문제점을 해결하기 위해 OpenAPI 및 매쉬업 검색 시스템을 개발하였다.

이를 위해 우선 유사도 기반 검색을 지원할 수 있는 통합 OpenAPI 및 매쉬업 가능 서비스 검색을 위한 검색 엔진을 제안하였다. 제안한 검색 엔진은 OpenAPI 등록 정보에 대하여 유사도 기반 검색을 수행할 수 있도록 TF*IEF 및 VSM을 이용하여 유사도를 계산하였다. 또한 검색시 통합 OpenAPI가 등록된 서비스 정보 모델과 같은 계층적인 문서 구조를 검색 결과에 반영할 수 있는 방법을 지원하였다. 지원하는 검색 방법은 두 가지로 통합 OpenAPI 검색과 함께 매쉬업 가능 서비스 검색을 지원한다. 매쉬업 가능 서비스 검색은 특정 범위 검색을 이용한 오퍼레이션의 입/출력값 비교와 오퍼레이션의 이름의 유사도 기반 질의를 통하여 매쉬업 가능 서비스를 검색을 지원한다.

또한, 개발한 검색 엔진은 기존의 검색 시스템에서 검색할 수 없었던 서비스와 오퍼레이션의 내용에 접근하여 검색할 수 있다. 또한 서비스 정보 모델을 구축하여 이에 대한 유사도를 계산하여 검색의 정확도를 향상시켰다.

이러한 시스템은 OpenAPI와 매쉬업 서비스에 대한 검색을 제공하는 Programmableweb, Popfly 등에서 제공하지 못하는 기능을 제공한다는 것에 큰 의미가 있다. 이는 매쉬업을 위해서 OpenAPI를 검색할 경우나 또는 검색된 서비스와 매쉬업이 가능한 서비스를 쉽고 정확하게 찾아줄 수 있기 때문에 관련된 서비스 제작을 더욱 간편히 할 수 있다는 장점이 있다.

이러한 시스템이 직접적으로 매쉬업 도구들과 연결되어 사용한다면 좀더 효율적으로 u-GIS를 위한 서비스를 제작할 수 있을 것이다.

참 고 문 헌

[1] 박지강, “당신은 웹 2.0 개발자입니까?”, 한빛미디어, 서울, 2007

[2] Programmableweb : http://www.programmableweb.com

[3] 이재광, “Web 2.0 and OpenAPI”, NHN, KRNet 2006

[4] HousingMaps : http://www.housingmaps.com

[5] Popfly : http://www.popfly.com

[6] Lee, D., Chuang, H., Seamons, K., “Document ranking and the vector-space model”, IEEE Software, 14(2), 1997, pp. 67-75

[7] 임백준, “프로그래밍은 상상이다”, 한빛미디어, 서울, 2008

[8] 이경하, 이규철, 김경옥, “키워드 질의를 이용한 순위화 된 웹서비스 검색 기법”, 한국전자거래학회, 제13권 제2호, 2008, pp. 213-233

[9] Curbera, F., et al., “Unraveling the Web services web: An introduction to SOAP, WSDL, and UDDI”, IEEE Internet Computing, 6(2), 2002, pp. 86-94

[10] 조대수, “대용량 센서 데이터 아카이빙을 위한 색인 분할 기법”, 한국공간정보시스템학회 논문지, 제9권 제1호, 2007, pp. 31-43

[11] “Web Services Description Language(WSDL) 1.1”, W3C Note, World Wide Web Consortium, http://www.w3c.org/TR/WSDL, 2001

[12] 신민철, “기초에서 실무까지 XML 웹서비스”, 프리렉, 서울, 2004

[13] Navarro, G., Baeza-Yates, R.A., Sutinen, E., Tarhio, J., “Indexing Methods for Approximate String Matching”, IEEE Data Engineering Bulletin, 24(4), 2001, pp. 19-27

[14] Dong, X., Halevy, A., Madhavan, J., Nemes, E., Zhang J., “Similarity Search for Web Services”, Proceedings of VLDB, 2004, pp. 372-383

[15] 임덕성, 홍봉희, “위치 기반 질의 처리를 위한 궤적 보존 색인의 설계 및 구현”, 한국공간정보시스템학회 논문지, 제10권 제3호, 2008, pp. 67-78

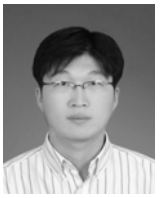
천 동 석

2007년 한남대학교 컴퓨터공학과 학사

2009년 충남대학교 공과대학 컴퓨터공학과 석사

2009년~현재 삼성전자 DMC 총괄 연구원

관심분야는 데이터베이스, 웹서비스, GIS



차 승 준

2006년 충남대학교 공과대학 컴퓨터공학과 학사

2006년~현재 충남대학교 공과대학 컴퓨터공학과 석박사통합과정 재학중

관심분야는 데이터베이스, 웹서비스, GIS, 웹 아카이빙

김 경 옥

1976년 서울대학교 학사
1987년 미국 오하이오 주립대학 전산학석사
1998년 충남대학교 컴퓨터공학과 박사
1981년~1984년 미국 오하이오 주립대학 기술원
1984년~1988년 미국 오하이오 주립대학 연구조교
1988년~1997년 시스템공학연구소 선임연구원
1998년~현재 한국전자통신연구원 텔레매틱스연구그룹 공간정보연구팀장
관심분야는 멀티센서(위성, 항공영상, 레이다, SAR, 현장센서 등) 데이터를 이용한 분석 및 3차원 건물, 지형 등을 추출, 3차원 모니터링 및 맞춤형 국토정보 제공기술 웹 서비스 시스템 구축

이 규 철

1984년 서울대학교 공과대학 컴퓨터공학과 학사
1986년 서울대학교 공과대학 컴퓨터공학과 석사
1990년 서울대학교 공과대학 컴퓨터공학과 박사
1994년 미국 IBM Almaden Research Center 초빙 연구원
1995년~1996년 미국 Syracuse University 초빙 교수
2001년~현재 전자상거래 표준화 통합 포럼 전자거래 기반 기술위원회 위원장
2003년~현재 한국전자거래학회 편집이사
2003년~현재 웹 코리아 포럼 부위원장
현재 충남대학교 공과대학 컴퓨터공학과 교수
관심분야는 데이터베이스, XML 웹 서비스, 시맨틱 웹 서비스, 유비쿼터스 컴퓨팅, GIS