

지능형 위치기반 서비스

전주성, 최준석, 양태준, 이기정
KTF 연구개발원, 차세대서비스개발팀

Intelligent Local Based Service

Jeon Joo-Seong, Choi Jun-Seok, Yang Tae-Joon, Lee Ki-Jeong
KTF R&D Group, Next Generation Service Development Team
E-mail : jsjeon016@ktf.com

요 약

본 논문에서는 이동통신 서비스에 시멘틱 웹 기술을 적용하여 지능형 위치기반 서비스를 제공할 수 있는 시스템 개발에 관한 연구를 기술하였다. 본 연구에서 기술되는 온톨로지 데이터베이스는 지역검색에서 사용자 질의가 많은 14분야의 온톨로지를 구축하여 지능적인 모바일 지역 검색 서비스가 가능하도록 구성하였으며 실제 서비스에 적용하여 그 효과를 입증하였다.

1. 서론

Tim Berners-Lee에 의해 1989년에 처음 제안된 WWW(World Wide Web)은 HTML 언어를 이용하여 편리성을 추구한 덕분에 일반 사용자 누구나 쉽게 정보를 접근하거나 게시할 수 있어 폭발적인 정보의 증가를 가져왔다. 이러한 편리함 때문에 높은 성장을 가져온 반면에 웹(web)의 정보가 감당할 수 없을 정도로 방대해진 현재의 상황에서는 문제점으로 작용하기도 한다. 검색측면에서 보면 현재의 웹 검색엔진은 주로 단어의 빈도수나 어휘 정보를 이용하여 문서의 유사도를 측정하고 랭킹을 매기기 때문에 사용자의 질의와는 관계없는 많은 문서를 결과로 가져올 수 있고 이로 인해 사용자는 불필요한 정보를 걸러내느라 시간을 낭비하게 된다.

현재 웹 환경과는 달리 시멘틱 웹(Semantic Web)

은 웹 상에 존재하는 정보를 사람뿐만 아니라 기계가 의미를 파악하고, 사용자의 요구에 적합한 결과만을 찾아주는 의미 기반 검색을 수행하거나 사람과 기계, 또는 기계와 기계 상호간에 협업을 원활히 수행함으로써 사람을 대신한 자동적인 서비스가 가능하도록 하는 차세대 웹 기술이라고 할 수 있다[1]. 단순한 링크정보로만 연결되어 있는 웹 페이지에 온톨로지(ontology) 기반의 의미 체계를 덧붙으로써 기계가 의미를 파악할 수 있게 되는 시멘틱 웹의 기본 구조를 만들 수 있다. 온톨로지는 컴퓨터를 지능화시킬 수 있기 때문에 시멘틱 웹 뿐만 아니라 지능형 로봇 시스템을 위한 지능형 서비스, e-비즈니스, 의료정보화, 바이오-인포메틱스 등의 다양한 분야에서 활용될 수 있다[2].

Gartner 그룹에서 발표한 통계에 의하면 전세계 이동통신 가입자는 2005년말 기준으로 21억7,581

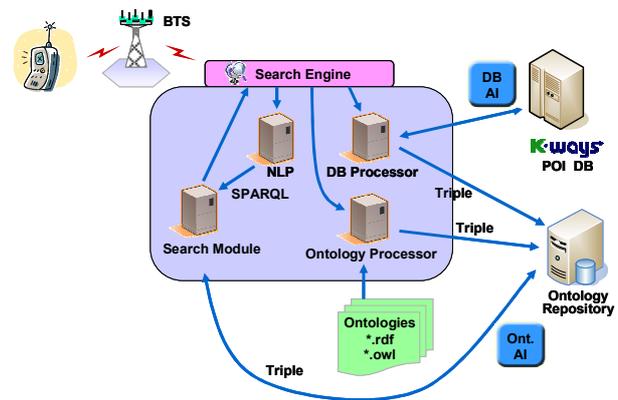
만명으로 보도되었다. 이와 같이 폭발적으로 증가하고 있는 이동통신 시장에서 모바일 검색 서비스는 매력적으로 성장하고 있는 블루 오션이다. 최근 검색엔진의 빅 3라고 할 수 있는 Google과 Yahoo, Micro Soft가 모바일 검색시장의 선점을 위해서 치열한 격전에 돌입한 것으로 보인다. Google은 각국의 이동통신 사업자와의 제휴를 통해서 모바일 검색시장으로 접근하고 있다. 참고적으로 전세계 휴대폰 가입자의 절반 정도의 규모를 가진 퍼스널 컴퓨터 사용자의 2006년 웹 검색 시장 규모가 총 51억 달러라고 보도되었으니 모바일 검색 시장이 본격적인 시장규모가 형성된다면 결코 무시할 수 없는 시장일 것이다.

본 논문에서는 시멘틱 웹 기반의 모바일 지역검색 시스템의 개발에 관한 연구를 기술하였다. 지금까지 시멘틱 웹 기술을 이동통신에 적용한 사례는 보고되지 않았다는 점에서 본 연구가 갖는 의미는 크다고 하겠다. 본 논문에서 기술한 시멘틱 웹 기반의 모바일 지역 검색 시스템은 사용자의 질의가 많은 14분야의 온톨로지를 구축하였다. 14개 분야의 온톨로지에는 2,000개의 class와 350만개의 individual로 구성하였다. 또한 지역 검색 서비스에 네비게이션(navigation) 서비스와 연계하여 모바일로 지도정보를 다운로드 받아 원하는 목적지(Point Of Interest)에 최단시간 내에 이동할 수 있도록 한다. 다운로드 받는 지도정보는 실시간 교통정보가 반영되기 때문에 사용자에게 가장 빠른 최적의 이동경로를 제공할 수 있는 매력적인 서비스이다.

2. 시스템 아키텍처

지금까지의 검색 시스템의 대부분이 관계형 데이터베이스 시스템으로 제공되고 있다. 예를 들면 Oracle, Microsoft SQL(Structured Query Language) 서버에서도 테이블 형태로 데이터를 저장 및 관리하며 SQL라는 인공적인 질의 언어를 통하여 쿼리(query) 하고 답을 얻어야 한다. 그러

나 이 경우에도 SQL 작성에 가입되는 keyword들이 테이블의 field명과 패턴 매칭이 될 수 있도록 일치하여야 한다. 현재의 컴퓨터가 검색을 포함해 지식과 정보 그리고 데이터를 처리하는 방식은 형태의 동일성에 기반한 단순하고 오류 가능성이 많은 메커니즘에 기반하고 있다[3]. 이런 기술의 한계를 극복하고 컴퓨터로 하여금 의미를 이해하고 지식, 정보, 데이터를 처리할 수 있도록 제안된 기술이 시멘틱 웹이라고 알려져 있다.



NLP: Natural Language Processor
 SPARQL: Protocol And RDF Query Language
 DB AI: DB Administrator Interface
 Ont. AI: Ontology Administrator Interface

Fig. 1. 시멘틱 웹 기반 지역검색 시스템 구성도

시멘틱 웹 기반 모바일 지역 검색 서비스를 위한 시스템 구성은 Fig. 1에 나타냈다. 이동통신 네트워크에서 시멘틱 웹기반으로 지역검색을 수행하는 시스템이다. 이동통신 네트워크를 통하여 사용자 휴대폰으로 목적지 검색을 위한 질의문이 전송되면 쿼리를 분석하고 목적지 정보가 온톨로지로 구축된 데이터베이스에 질의한다. 질의 결과로 검색된 데이터들을 공간 검색 모듈을 통해서 거리 정보와 위치정보로 계산하여 상기의 쿼리 요구를 반영하는 목적지 검색결과 데이터를 제공하는 프로세스를 갖는다.

이상과 같이 언급된 지역 검색의 수행 과정을 Fig. 2에서 플로우 차트로 나타냈다. 쿼리가 입력되면 분석모듈에서 쿼리를 형태소 분석하는데 온

톨로지, 데이터베이스, 개체명 사전 순서로 참조하여 분석한다.

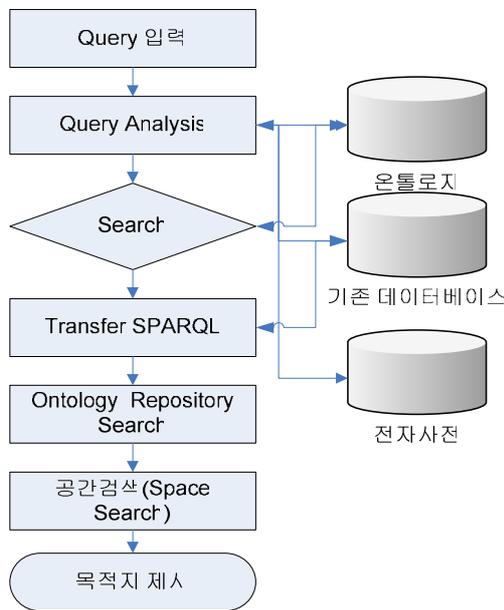


Fig. 2. 지역검색을 위한 프로세스

온톨로지와 개체명 사전, 형태소 분석과 동시에 진행되는 의미분석 결과로 추출되는 의미자질(어간에는 모두 74개의 의미자질, 어미에는 47개의 의미자질이 있다) 그리고 질의문 내의 밀접한 관련이 있는 부분을 묶어서 나누는 구문분석, 마지막으로 주어, 동사, 목적어 등의 문장의 구성성분 분석을 수행한다. 질의문의 내용을 한국어 분석기로 형태소, 의미, 구문, 문장성분으로 분석하여 온톨로지 리파지토리(RDF triple과 OWL 온톨로지 구성)의 종류에 맞는 쿼리 언어로 변환하여 검색을 시도한다.

개체명 사전은 어간사전과 어미사전으로 구성되어 있으며 그 구조는 인덱스 방식의 파일 데이터베이스로 되어 있어 속도가 빠르고 업데이트가 용이하게 설계되었다. 어간사전에는 30만여 어휘와 74개의 의미자질로 구성되며 어미사전에는 5,000여 형태와 47개의 의미자질을 갖도록 구성하였다.

Fig. 3에서 개념 검색 프로세스를 통해서 목적

지가 검색되는 예를 도시하였다. 사용자로부터 휴대폰의 검색 창에 “역삼동 스파게티”를 입력하면 이동통신 네트워크를 통해서 검색엔진에서 쿼리를 분석하여 “스파게티”를 질의어로 OWL ontology에 질의를 하게 된다(1). “스파게티”는 음식점 OWL ontology의 이탈리아 음식점 class의 하위 개념의 메뉴에 포함되어 있다. 따라서 “스파게티”는 개념 검색을 통해 “이탈리아 음식점”의 class가 검색된다(2). 이렇게 추출된 “이탈리아 음식점”을 다시 음식점 OWL ontology에 질의한다. 그러면 이탈리아 음식점들의 individual이 검색된다(3). 이와 함께 공간검색 모듈을 통해서 상기 질의문의 “역삼동 스파게티”에서 역삼동을 공간검색하여 상기의 단계 (3)에서 search된 이탈리아 음식점들 중 역삼동에서 가장 가까운 individual을 계산하고 검색한다. 이때 공간검색에는 RDF triple의 위치정보를 이용한다(4). 이렇게 공간검색과 연동되어 검색된 목적지인 individual ID를 RDF triple에 조회하여 상세한 목적지 정보를 추출한다(5).

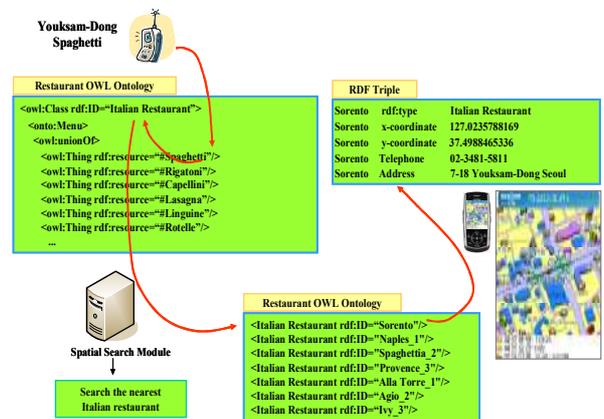


Fig. 3. 개념 검색 처리의 예

이와 같이 본 연구에서 기술한 시멘틱 웹 기반의 지역 검색 시스템은 이동통신 위치기반 서비스에 도입하여 유연한 명칭 검색, 지능형 문장 검색, 개념 검색, 유사 검색 등의 차세대 검색 서비스의 제공이 가능하다.

3. 온톨로지 리파지토리

본 연구에서 기술되는 지역 검색 서비스를 위한 온톨로지 리파지토리는 legacy 데이터베이스를 W3C의 N3(Notation 3) 기법을 적용해 RDF triple로 변환하여 구축하였다. N3 기법은 기존의 관계형 데이터베이스 시스템의 테이블을 “Record-Field-Data” 형식으로 “Subject-Predicate-Object’의 triple 형태로 자동 변환할 수 있도록 하였다. 기존 관계형 데이터베이스에 저장되어 있는 데이터는 온톨로지 인스턴스의 세부정보에 해당하는 바, 이 기법으로 데이터베이스 인스턴스의 자동생성은 물론 무결성을 보장할 수 있다. 기존 데이터베이스의 field(column)는 triple 술어(Predicate)로 매핑되고, record는 주어(Subject), data는 목적어(Object)로 각각 매핑된다.

본 연구에서 기술되는 온톨로지 리파지토리는 지역검색 데이터베이스에서 사용자 질의가 많은 14분야의 온톨로지를 구축하였다. 14 분야의 온톨로지는 Table 1에 나타내었다.

Table 1. 14개 분야의 온톨로지 세부 정보

Ontology	Class	Individual	File size [MBytes]
Administration	204	149	9.6
Government	85	69	0.25
Financial Agency	320	268	18.2
Hospital	87	318	29.8
Transportation	245	225	47.1
Shopping/Gas station	165	462	50.3
Hotel	95	215	4.45
Concert Hall	115	89	6.1
Leisure	165	238	4.95
Education	170	278	37.3
Enterprise	86	419	17.5
Restaurant	207	687	89.5
Event Hall	36	39	3.65
Tourist Resort	20	37	3.63
Total	2,000	3,500,000	322.33

4. 시스템의 성능평가 결과

시멘틱 웹 기반의 모바일 지역 검색 시스템의 성능을 평가한 결과, 시스템의 성능에 가장 큰 영향을 미치는 요소는 문자수이다. 이는 사용자로부터 입력되는 질의문을 SPARQL로 변환하는 과정에서 문자수가 많아지면 변환 정확도가 낮아지기 때문이다. Fig. 4에서 질의문의 문자수와 SPARQL 변환 정확도와의 상관관계를 나타냈다.

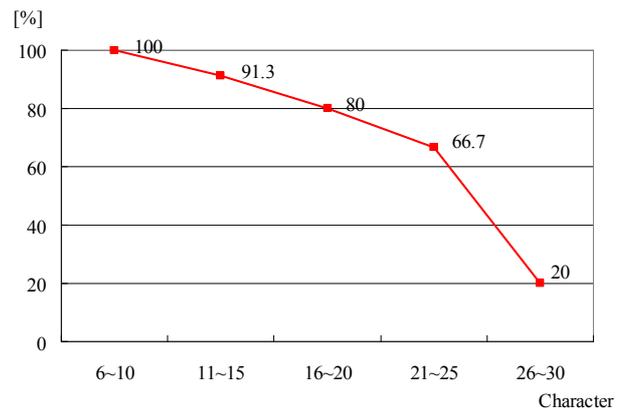


Fig. 4. 질의문의 문자수와 SPARQL 변환 정확도

기존 지역 검색 서비스를 위한 시스템에 저장된 로그 데이터를 사용하여 검색 성공률을 시물레이션하였다. 명칭검색의 경우, 기존 시스템의 검색 성공률은 70.49%이며 본 논문에서 기술한 시멘틱 웹 기반의 모바일 지역 검색 시스템의 검색 성공률은 90.92%의 우수한 결과를 나타냈다. 또한 개념 검색의 시물레이션 결과에서도 검색 성공률이 각각 25.72%와 89%로 비교적 좋은 결과를 나타냈다.

5. 서비스 고도화 기능 구성

지금까지의 네비게이션 서비스에서 가장 큰 단점은 실시간 도로교통 상황과는 무관한 단순 경로만을 제공하는데 있었다. 이러한 단점을 보완하기 위해서 도로교통정보센터로부터 5분 단위로 업데이트되는 교통정보와 네비게이션 서비스에서 초기 설정된 경로를 비교해서 경로변경이 필요한 경우

소통상황을 반영하여 최적경로를 제시한다. Fig. 5에서 이러한 기능에 대한 프로세스를 나타냈다.

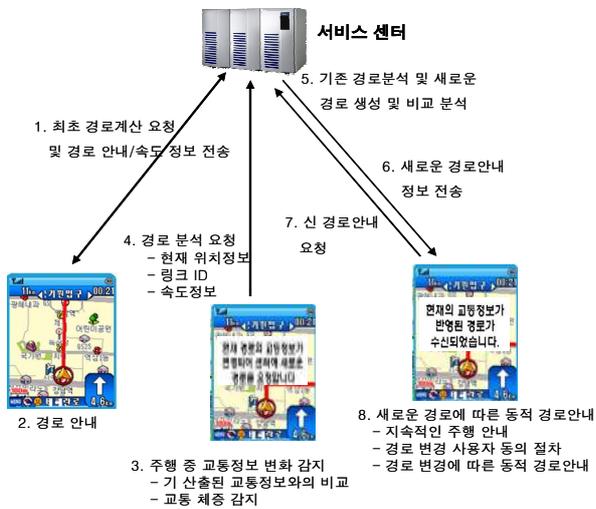


Fig. 5. 다이내믹 네비게이션의 프로세스

기존의 네비게이션 서비스에서는 단순 길안내 서비스에 그쳤지만 다이내믹 네비게이션 서비스는 실시간 교통상황을 반영하여 최종 목적지까지 최적 경로를 안내가 가능하다. 이와 같은 기능은 Stand-alone 타입의 네비게이션 기능과 차별화할 수 있는 기능이라고 할 수 있다.

Table 2에서 현재 이동통신에서 사용되는 측위 방식별 분류를 나타내었다.

Table 2. 이동통신에서 사용되는 측위 방식 분류

구분	개념
MS Assisted GPS (a-GPS)	<ul style="list-style-type: none"> • 휴대폰이 위치계산을 하지 않고, 서버가 위치 계산 <ul style="list-style-type: none"> - 단말이 측위 요청을 할 때, 단말은 자기가 잡고 있는 기지국 정보를 PDE에 송신 - PDE는 기지국 위치를 기준으로 단말 위에 있는 GPS 위성 정보를 단말에 송신 - 단말은 GPS 위성을 확인하여 서버에 송신 - 서버가 단말의 위치를 계산하여 단말로 송신
MS Based GPS	<ul style="list-style-type: none"> • 휴대폰이 위치계산 <ul style="list-style-type: none"> - 단말이 측위 요청을 할 때, 단말은 PDE에 GPS 위성 정보를 요청하고, PDE는 해당정보를 단말에 송신 - 단말은 자기 위에 있는 GPS 위성을 확인하고 자기의 위치 계산
S-GPS (Simultaneous GPS)	<ul style="list-style-type: none"> • 휴대폰이 위치계산 <ul style="list-style-type: none"> - 측위 프로세스는 MS Based GPS와 동일 - 휴대폰 Rx Path가 2개로서, 하나는 호처리리에 활용되고 다른 하나는 GPS Signaling에 활용됨으로써 측위 시간 간격이 1초로 정상 처리됨

본 논문에서 소개되는 이동통신 기반의 네비게이션 서비스가 가능한 것은 이와 같은 측위기술이 뒷받침이 되기 때문이다.

6. 결론

최근 세계 최고의 software 기업인 Microsoft와 Google, 세계 최고의 휴대폰 점유율을 차지하고 있는 Nokia, 미국 navigation 시장의 선두기업인 Garmin에서의 주된 관심사는 지리정보를 활용하여 새로운 콘텐츠 서비스로 고객들에게 접근하여 다양한 business model 창출이라고 한다. 유선 웹 상에서 활성화된 웹 2.0의 소프트웨어 환경, 향상된 전송속도를 제공하는 무선통신환경, 그리고 개인의 위치정보를 활용할 수 있는 모바일 환경 등은 이제 모바일 웹 2.0이라는 새로운 트렌드를 만들어 가고 있다.

본 연구에서 기술한 시멘틱 웹 기반의 모바일 지역 검색 시스템은 이동통신 서비스에 시멘틱 웹 기술을 적용하여 사용자의 검색 시간을 단축시키고 원하는 정보를 정확하게 서비스를 제공할 수 있다. 기존 위치기반 서비스에 비교할 때, 명칭 검색만을 비교할 경우 약 20% 이상의 검색 성공률을 개선하는 좋은 결과를 얻었다. 또한 시스템에서 자연어 처리와 추론기능이 가능하여 유연한 명칭 검색, 지능형 문장 검색, 개념 검색, 유사 검색 등의 차세대 검색 서비스의 제공이 가능하다. 앞으로의 연구방향은 본 연구결과를 연계하여 e-커머스 및 멀티미디어 검색 서비스 등으로 확장하는 것이다.

[참고문헌]

- [1] Daconta, Michael C, Leo Joseph, Smith Kev, "The Semantic Web," John Wiley & Sons Inc, 2003.
- [2] 노상규, 박진수, "인터넷 진화의 열쇠 온톨로지 웹2.0에서 3.0으로," 가즈토이 출판사, 2007.
- [3] 최호섭, 옥철영, "정보검색 시스템과 온톨로지," 정보과학회지, 제22권 4호, 2004.