

공간정보 플랫폼 구축을 위한 전자지도와 POI 정보의 매칭 방법 Matching Method of Digital Map and POI for Geospatial Web Platform

김정옥* · 허 용** · 이원희*** · 유기윤****
Kim, Jung Ok · Huh, Yong · Lee, Won Hee · Yu, Ki Yun

要 旨

최근 포털사이트 및 여러 분야에서 제공하는 지도서비스로 쉽게 공간정보를 이용할 수 있어 다양한 기관에서 제공하는 공간정보의 통합에 대한 연구가 주목을 받고 있다. 이에 따라 본 연구에서는 이종의 공간정보를 통합하는 연구를 진행하였다. 공간정보의 통합과정은 다른 제공기관에서의 공간정보 추출과 동일좌표계로의 변환 그리고 대응되는 공간객체를 찾는 단계로 구성되며, 본 연구에서는 마지막 단계인 실세계의 대응되는 공간객체를 매칭하는 연구를 중심으로, 전자지도와 POI와 같은 이종의 공간정보를 효율적으로 통합하는 방법을 개발하였다. 이를 위해 네 개의 매칭 종류와 그에 따른 기준을 제시하였다. 본 연구의 주된 목적은 이종의 공간객체에 대한 매칭을 위해 공간정보 및 부가적인 속성정보를 이용하는 구체적인 방법과 절차를 제시하는 것이다.

핵심용어 : 수치지도, POI 매칭, 지도 API, 상호운용성, Web 2.0

Abstract

Recent growth of the geospatial information on the Web has made it possible to easily access a wide variety of geospatial information. An integration of different geospatial objects consists of the following three steps; extracting geospatial objects from the maps, converting the coordinate system and discovering pairs of objects that represent the same real-world entity in the two maps. This paper deals mainly with the third step to correspond conjugate objects and four matching types and criteria is presented. The techniques designed and developed can be utilized to efficiently integrate distributed heterogeneous spatial databases such as the digital maps and POIs from other data sources. To achieve the goal, we presented four types and criteria for the matching schema. The main contributions of this paper are as follows. A complete process of integrating data from maps on the Web is presented. Then, we show how attributes of the objects can be used in the integration process.

Keywords : Digital map, POI matching, Map API, Interoperability, Web 2.0

1. 서 론

최근 포털사이트 및 여러 분야에서 제공하는 지도서비스에서 지도의 가치는 단순한 하나의 콘텐츠가 아니라 다양한 콘텐츠를 수용할 수 있는 플랫폼(location based contents platform)으로 인식되고 있다. 국외의 경우 일찍이 구글, 마이크로소프트, 야후 등의 대형업체들이 지도를 플랫폼으로 개발하여 전 세계 위성사진과 지도 서비스를 통합하여 선보이고 있으며, 국내의

경우도 보유하고 있는 뛰어난 지도 서비스들을 바탕으로 맞춤형 공간정보의 활용을 극대화할 수 있는 기술을 선보이고 있다.

플랫폼으로써의 공간정보를 효과적으로 활용하기 위해서는 다양한 원천자료를 통합하는 것이 무엇보다도 중요하며, 특히 공공 및 민간의 공간정보를 통합하는 기술 개발이 필요하다(김은형, 2009). 공공부문은 많은 비용과 시간을 들여 구축한 NGIS(National Geographic Information) 데이터를 보유하고 있으며, 이에

2009년 10월 9일 접수, 2009년 10월 30일 채택

* 정희원 · 서울대학교 대학원 건설환경공학부 박사과정(geostar1@snu.ac.kr)

** 서울대학교 대학원 건설환경공학부 박사과정(hy7808@snu.ac.kr)

*** 서울대학교 대학원 건설환경공학부 BK21 계약교수(wlee2432@snu.ac.kr)

**** 교신저자 · 정희원 · 서울대학교 공과대학 건설환경공학부 부교수(kiyun@snu.ac.kr)

대한 데이터 갱신도 지속적으로 이루어지고 있다. 예를 들어 서울시의 경우 도로관리시스템, 새주소관리시스템 등을 통해 수치지형도의 수치 갱신된 자료의 취득이 가능하다(강영욱, 2006). 그러나 수치지형도의 데이터 갱신주기가 민간업체에서 활용하기에는 실제 수요자의 요구를 만족시키지 못하고 있다. 민간분야의 경우 공간 정보를 수치로 갱신하는 것이 체계적으로 관리되고 있다. 특히 지도검색서비스에서 제공하는 POI(Point Of Interest, 관심지점)정보 및 POI정보 갱신 전략은 공공보다 민간부문에 비해 빠르게 유지 관리되고 있으며, 최근 참여와 공유를 추구하는 웹 2.0의 사상을 반영한 Open API(Application Program Interface)로 쉽게 접근할 수 있다. 그러나 GIS를 포함한 LBS 시장에서는 각 업체별로 기본도를 갱신하고 있어 5년 주기로 갱신된 NGIS 데이터와 5년 동안 자체 수치갱신체계에 의해서 갱신된 지도가 불일치하는 경우가 발생하여 두 데이터를 통합하는 문제가 주요한 애로사항으로 언급된다. 또한 위치정보 제공의 바탕이 되는 주차장, 버스정류장, 신호등, 도로표지판, 도로개통, 재난통제 정보와 같은 공공 정보 수집을 위해 유지 관리 비용을 지급하고 있어 동일한 콘텐츠에 대해 공공 및 민간에서 각각 유지 관리 비용을 중복 투입하고 있다.

공공과 민간의 공간정보 통합으로 공공기관에서는 지도정보를 갱신하고 대시민 서비스를 향상하기 위하여 투입되는 많은 예산을 민간의 노하우를 이용해 비용 절감과 서비스의 질을 향상하고, 민간부문에서는 지형 지물과 공공정보 수집을 위한 유지 관리 비용의 중복 투입 방지를 통해 실질적인 가치를 창출하는 위치정보 서비스가 필요하다. 따라서 본 연구에서는 더 많은 정보가 보다 정확하고 다양한 방법으로 맵핑(mapping)되는 기술을 개발하기 위해 공공에서 구축한 수치지형도와 민간의 POI정보의 공간객체 단위 매칭을 위한 방법을 제안한다.

2. 선행연구 분석

강영욱(2006)의 연구에서 민간부문을 상대로 공공부문과의 협업 시 요구사항을 인터뷰한 결과를 보면 민간업체에서 공공으로부터 제공받기를 바라는 데이터는 행정안전부의 새주소 사업과 국토해양부의 유통망 사업으로 구축된 데이터이다(표 1).

이러한 요구사항 및 공공서비스 측면과 민간 콘텐츠 측면에서 추구하는 바가 상당한 차이가 있으므로 이를 고려하여 본 연구에서는 갱신주기가 짧고 공공 및 민간 모두에서 활용도가 높은 건물객체에 대한 통합 방법을

표 1. 공공부문과의 협업에 관한 민간업체 인터뷰 내용

| 업체명 | 요구사항 |
|-------------|---|
| 우리사이 맵 | <ul style="list-style-type: none"> • 새주소 데이터를 포함한 지번데이터의 합법적 이용 • 공유되는 데이터의 질과 갱신주기 확보 |
| 콩나물 | <ul style="list-style-type: none"> • 새주소사업 및 유통망사업으로 구축된 데이터 • 국가 데이터의 POI나 랜드마크의 선별 |
| 팅크웨어(INAVY) | <ul style="list-style-type: none"> • 새주소데이터, 건축물 동별 정보, 상습정체구간, 거주자 우선 주차구역 정보 등 • 지속적인 협업 유지체계 필요 |
| SK텔레콤 | <ul style="list-style-type: none"> • 새주소데이터 모바일 관련 데이터, 주차장 정보, 항공사진, 신호등, 도로표지판, 도로개통정보, 재난통제정보, BIS 데이터의 공유 원함 |

연구하였다. 따라서 행정안전부에서 제공하는 도로명주소(새주소) 전자지도와 공공보다 민간부문에 비해 빠르게 유지 관리되고 있는 POI정보를 연구 자료로 선정하고 이들의 통합 방법을 연구하였다.

2.1 도로명주소 전자지도 개요

행정안전부(2009)의 도로명주소 전자지도는 기존의 지번 표시 주소 외에 도로명 주소를 표기한 전국 각지 건물의 위치정보를 담은 전자지도로 도로명 자료 이용 신청서를 작성 후 제공되고 있다. 특정 주소의 건물과 관련한 위치 정보는 물론 건물용도, 상호, 건물설계도, 비상구의 위치정보 등이 단계적으로 정부부처 및 지방자치단체 등 공공분야는 물론 민간에도 제공돼 재난 대응, 신속한 물류처리 등의 목적으로도 활용 가능할 것으로 기대된다. 도로명주소 전자지도는 도로명 주소안 내를 위해 건물, 건물이력, 건물군, 실폭도로, 도로구간, 도로구간 이력, 도로명판, 기초구간, 출입구, 연결선, 교차로, 지하철선로, 지하철역사, 지하철출입구, 철도역사, 철도선로, 터널, 교량, 하천호수, 공원, 고가도로, 자동차전용도로, 지하철도, 기타시설물, 기타표지판 이상 25개의 레이어를 포함한다. 또한 새주소안내시스템에서는 테스트 목적의 새주소와 지번의 매칭테이블을 제공한다.

2.2 민간부문의 POI 제공 현황

POI는 주요 시설물, 역, 공항, 터미널, 호텔 등을 좌표로 전자지도에 표시하는 데이터이다(네이버, 2009b). 보통 목적지 검색에 사용되는 검색 데이터와 바탕 화면에 표시만 되는 바탕 데이터로 구분할 수 있으며, ASCII 포맷으로 독립적으로 제공되기도 하나 POI 제

표 2. 주요 내비게이션 전자지도 제작업체

| 구분 | 업체명 | 지도브랜드명 |
|----|---------|--------|
| 국내 | 엠펙소프트 | 맵피, 지니 |
| | 팅크웨어 | 아이나비 |
| | SK 에너지 | 엔나비 |
| | 파인디지털 | 파인맵 |
| | 트윈클리틀스타 | 콩나물 |
| 국외 | 나브텍 | 나브텍 |

표 3. 포털사이트의 지도서비스 현황

| 업체 | 서비스명 | 특징 | 지도 API |
|-----|---|---|--------|
| 구글 | 구글어스 (earth.google.com) 구글맵스 (maps.google.com) | 스트리트뷰, 실시간 교통정보, 지형모습 | 제공 |
| 네이버 | 네이버지도 (map.naver.com) | 국내 최대 지역 DB 확보, 2m급 위성사진, 수도권 50cm급 항공사진 | 제공 |
| 다음 | 다음지도 (local.daum.net/map) | 50cm급 항공사진, 스카이뷰, 로드뷰 | 제공 |
| 야후 | 야후 거기 지도 (kr.gugi.yahoo.com/map) | 하이브리드 지도(위성지도+일 반정보지도)실시간 교통정보 | 제공 |
| 파란 | 파란지도 (local.paran.com/map) | 리얼스트리트 | - |

매일경제(2008) 수정

작의 기본이 되는 전자지도와 함께 제공된다(표 2).

구글, 야후, 네이버, 다음 등의 포털사이트의 지도서비스는 사용자의 웹 사이트에 지도를 표시할 수 있는 지도 오픈 API를 제공한다(표 3). 이는 지도상의 원하는 위치에 여러 가지 정보를 표시할 수 있으며, 포털의 지도서비스에서 제공하는 기능을 손쉽게 웹 사이트에 추가할 수 있는 기능을 제공한다. 또한 지역서비스에 등록된 각 지역별 업체 및 관광지 등의 검색결과를 지도서비스와 손쉽게 연동하여 활용할 수 있다. 즉, 지도 API와 함께 POI 검색 API를 사용하면, 국내는 물론 전세계 유명한 지역명을 한글 및 영문으로 검색 가능하다(네이버, 2009a; 다음, 2009, 야후, 2009).

2.3 연구동향

POI로부터 주기 및 검색안내 데이터화와 다른 객체 간의 상호연관성에 대한 연구는 주로 POI의 공간 및

속성정보를 이용한다. 즉, 동일좌표계로 변환 후 포인 트와 폴리곤의 중첩분석을 수행하거나, 두 객체 사이의 유클리드(euclidean) 및 하우스도르프(Hausdorff) 거리가 최소인 객체와 매칭하는 방법을 사용한다(Beeri et al., 2004; Beeri et al., 2005). 또한 좌표를 기준으로 주소를 파악하는 지오코딩(geocoding)이나 주소를 바탕으로 좌표를 추적하는 역지오코딩(reverse geocoding) 방법을 이용하기도 한다(Hild et al., 1998).

그러나 한번 구축 및 가공된 POI정보는 특정 수치지형도 사양에 맞추어 제작되기 때문에 이를 이중지도 혹은 사양이 다른 지도에 적용하기가 매우 힘들었다. 기존의 연구들은 이러한 과정이 거의 수작업에 의해서 진행되었기 때문에 다양한 어플리케이션용 POI정보로 활용하기란 시간과 비용적 측면에서 어려웠다(김대식 외, 2002).

또한 지도상에 표시되는 바탕데이터로의 POI는 많은 오차를 포함한 좌표정보를 저장하고 있다. 이는 POI를 전자지도 상에 심볼로 표시할 경우 지도의 축척에 따라 심볼의 겹침을 방지하기 위해 심볼위치는 실제 객체의 위치가 아닌 곳에 배치되고 해당 위치를 좌표로 저장하기 때문이다. 이 경우 정교한 좌표변환을 수행한다 하여도 정확한 위치 파악이 어려워 기존의 좌표정보만을 이용한 공간분석으로 매칭을 시도하면 오매칭되는 사례가 다수 포함된다.

다른 연구에서는 이를 보완하기 위해 명칭, 동번호, 지번 등과 같은 POI 속성을 활용하기도 하나 이 역시 특정 사양에 맞추어 속성정보가 제작될 경우 자동 매칭이 제한적이고, 매칭 방법에 대한 구체적 절차와 기준이 성립되어 있지 않다(류재현, 2007; Fu et al., 2005; Safra et al., 2006).

3. 도로명주소 전자지도와 POI 정보의 매칭 알고리즘

도로명주소 전자지도와 POI정보의 매칭 순서도는 그림과 같다. 우선 각각의 데이터를 단일 좌표계로 변환하고, 매칭을 위한 속성정보의 공통요소를 정규화 한다. 공간 및 속성정보를 토대로 4개의 매칭유형으로 구분하고, 이에 따른 기준을 설정하여 매칭 한다. 본 연구에서는 매칭 종류에 따른 구체적인 절차와 기준을 제시함으로써 자동 매칭을 실현하고자 하였다.

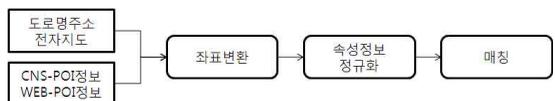


그림 1. 매칭 순서도

3.1 입력변수 정규화

매칭 방안을 제시하기 위하여 도로명주소 전자지도의 건물레이어(테이블명 : TN_SPBB02), CNS(Car Navigation System) 전자지도 상의 POI정보, 포털사이트 야후의 POI검색 API를 이용해 제공되는 POI정보의 차이를 비교하였다(표 4). 현재까지 공공분야와 민간에서 사용하고 있는 전자지도는 초기 제작 당시에는 국가 기본도 및 1/1,000 수치지형도를 바탕으로 출발했지만 갱신이 거듭될수록 제작 및 활용방법 등에서 차이를 보

표 4. 실험 데이터셋의 차이

| | | A | B | C |
|----|--------------|---|---|---|
| 속성 | • 건물관리번호(ID) | ○ | ○ | ○ |
| | • 도로관리번호 | ○ | × | × |
| | • 명칭 | ○ | ○ | ○ |
| | • 건축물용도/분류코드 | ○ | ○ | ○ |
| | • 지상/지하층수 | ○ | ○ | ○ |
| | • 우편번호 | ○ | × | × |
| | • 주소(지번 포함) | ○ | ○ | ○ |
| | • 읍면동코드 | ○ | ○ | × |
| | • 전화번호 | × | ○ | ○ |
| | • 홈페이지 | × | × | ○ |
| 도형 | • 기하정보 | 면 | 점 | 점 |

A : 도로명주소 전자지도, B : CNS POI, C : Web POI

표 5. 유사성 척도 정의

| 구분 | 정보 | 내용 |
|-----------|----|---|
| 속성 정보 유사도 | 명칭 | $\sigma(N_1, N_2) \begin{cases} 1.0 & \text{if } N_1 = N_2 \\ 0 & \text{otherwise} \end{cases}$ N_1 : 도로명주소 전자지도의 건물명 N_2 : POI의 명칭 |
| | 지번 | $\sigma(L_1, L_2) \begin{cases} 1.0 & \text{if } L_1 = L_2 \\ 0 & \text{otherwise} \end{cases}$ L_1 : 도로명주소 전자지도의 지번 L_2 : POI의 지번 |
| 공간 정보 유사도 | 중첩 | $\sigma(O_1, O_2) \begin{cases} 1.0 & \text{if } O_1 \cap O_2 = O_2 \\ 0 & \text{if } O_1 \cap O_2 = \emptyset \end{cases}$ O_1 : 도로명주소 전자지도의 건물객체 O_2 : POI 객체 |
| | 거리 | $\sigma(L_1, L_2) \begin{cases} 1.0 & \text{if } dis(D_1, D_2) \leq tol \\ 0 & \text{otherwise} \end{cases}$ D_1 : 도로명주소 전자지도의 건물객체 D_2 : POI 객체, tol : 허용거리 dis : 두 객체 사이의 거리 |

이고 있다.

명칭과 지번을 포함한 주소와 도형정보가 매칭에 이용 가능한 요소로 분석되었다. 따라서 본 연구에서는 도로명주소 전자지도와 CNS POI 및 Web POI의 매칭을 위해 공간정보 유사성(spatial similarity)과 속성정보 유사성(thematic similarity)을 표 5와 같이 정의한다.

3.2 매칭종류 및 기준

본 연구에서 제안하는 도로명주소 전자지도와 POI의 매칭은 객체의 공간 및 속성정보를 토대로 그림 2와 같

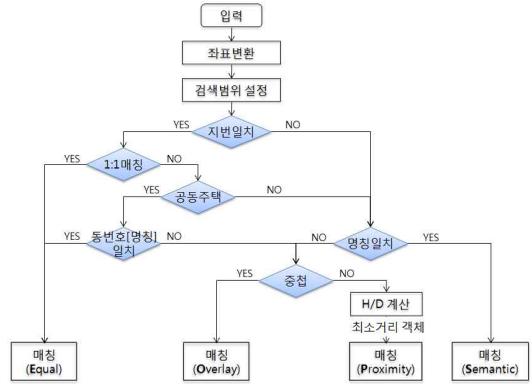


그림 2. 매칭 단계 및 종류

표 6. 매칭 종류 및 기준

| 매칭종류 | 매칭기준 | |
|-------------|---|--|
| 1 Equal | <ul style="list-style-type: none"> • 중심간 거리 : 설정 검색범위(m) 이내 • 지번 : 동일 (유일) • 집합건물의 경우 지번과 동번호 일치 | |
| 2 Semantic | <ul style="list-style-type: none"> • 중심간 거리 : 설정 검색범위(m) 이내 • 명칭/동번호 유사도 : $Max(T_1, T_2, \dots, T_n)$ | |
| 3 Overlay | O-E | <ul style="list-style-type: none"> • 중심간 거리 : 설정 검색범위(m) 이내 • 건물 폴리곤 내부에 매칭 대상 포인트 포함 • 지번 : 동일 |
| | O-S | <ul style="list-style-type: none"> • 중심간 거리 : 설정 검색범위(m) 이내 • 건물 폴리곤 내부에 매칭 대상 포인트 포함 • 명칭/동번호 유사도 : $Max(T_1, T_2, \dots, T_n)$ |
| | O-O | <ul style="list-style-type: none"> • 중심간 거리 : 설정 검색범위(m) 이내 • 건물 폴리곤 내부에 매칭 대상 포인트 포함 |
| 4 Proximity | <ul style="list-style-type: none"> • 중심간 거리 : 설정 검색범위(m) 이내 | |

이 4개(Equal, Overlay, Proximity, Semantic)로 구분되며, 자세한 매칭기준은 표 6과 같다. 매칭 하고자하는 POI의 좌표를 도로명주소 전자지도와 같은 좌표계로 변환 후 지도에 표시하고, 50m 반경의 검색범위를 설정하여 범위에 포함되는 객체를 매칭 후보객체군으로 한다. 이때 후보객체는 2개 이상이어야 하며, 이 조건이 만족될 때까지 반경을 10m 단위로 증가시킨다.

3.2.1 Equal 매칭

POI와 도로명주소 전자지도 검색범위 내의 매칭후보 건물객체의 지번정보를 비교하여 같은 지번정보를 갖는 유일한 객체쌍이 존재하면 해당 관리번호로 매칭 테이블을 작성한다. 만약 다수의 후보객체와 지번정보가 일치하고, 건물군 번호를 부여받은 아파트와 같은 집합 건물이라면 구체적인 동정보를 이용하여 1:1 매칭을 시도한다. 어떠한 조건도 만족하지 못한다면 Semantic 매칭 단계로 넘어간다(그림 3).

3.2.2 Semantic 매칭

Semantic 매칭은 비교객체의 명칭정보의 유사도를

```

Algorithm: EqualMatching(P, MC)
01: P_Lot ← GetPoiLotNumber(P)
02: FOR each object of MC DO
03:   B_Lot ← GetBuildingLotNumber(MC)
04:   IF P_Lot is equal to B_Lot THEN
05:     DecisionField ← 1
06:   END IF
07: END FOR
08: FOR each object of MC DO
09:   Sum ← The sum total of DecisionField
10:   IF Sum is equal to 1 THEN
11:     RESULT MakeMTable(P, MC(D DecisionField(1)))
12:   ELSE
13:     IF MC is Building group THEN
14:       FOR each object of MC(D DecisionField(1)) DO
15:         P_Number ← GetPoiNumber(P)
16:         B_Number ← GetBuildingNumber(MC)
17:         IF P_Number is equal to B_Number THEN
18:           RESULT ← MakeMTable(P, MC, Equal)
19:         END IF
20:       END FOR
21:     END IF
22:   END IF
23: END FOR
24: RETURN RESULT
    
```

그림 3. Equal 매칭 알고리즘

분석하여 유사도가 가장 높은 객체쌍에 대해 매칭 테이블을 작성한다(그림 4). 이때 본 연구에서는 CNS POI의 별칭(alias) 데이터베이스를 활용하여 동의어집합(synset)을 작성하였다(그림 5). 예를 들어 CNS의 경우 주민센터를 일컫는 용어가 매우 다양하여 표 7과 같은 별칭 테이블을 별도 작성 후 검색에 활용하고 있다. 별칭 데이터베이스가 포함하고 있는 별칭은 약 350만개로 검색에 사용될 모든 공간객체에 대한 별칭을 각각 저장하고 있어 비효율적이다. 이에 본 연구에서는 별칭 데이터베이스를 분석하여 동일한 의미를 갖는 용어를 대분류 18개, 소분류 155개의 동의어집합으로 구축하고, Semantic 매칭 시 활용하였다.

```

Algorithm: SemanticMatching(P, MC)
01: P_Name ← GetPoiName(P)
02: FOR each object of MC DO
03:   B_Name ← GetBuildingName(MC)
04:   SemanticSimField ← Similarity score Between P_Name and B_Name
05: END FOR
06: FOR each MC of SemanticSimField DO
07:   IF MAX(SemanticSimField) > threshold THEN
08:     RESULT ← MakeMTable(P, MC, Semantic)
09:   ELSE
10:     OverlayMatching(P, MC)
11:   END IF
12: END FOR
13: RETURN RESULT
    
```

그림 4. Semantic 매칭 알고리즘

표 7. POI 별칭의 예

| POI_ID | SEQ | NAME |
|--------|-----|-----------|
| 3079 | 0 | 대학동주민센터 |
| 3079 | 1 | 대학동동사무소 |
| 3079 | 2 | 동사무소 |
| 3079 | 3 | 대학동사무소 |
| 3079 | 4 | 대학동자치센터 |
| 3079 | 5 | 대학동주민자치센터 |
| 3079 | 6 | 대학동주민센터 |

| | | |
|-----|--|-----------------------|
| 관공서 | {[H: 경찰서, 경찰소], [H: 공무원교육원, 지방공무원, 공무원연수원], [H: 국군휴양소, 계통스파텔]} | {[H: 지구대, 파출소, 치안센터]} |
| | {[H: 유성, 계통]} | |
| | {[H: 지방법원, 지방], [H: 대법원]} | {[M: 통기소, 통기서]} |
| | {[H: 가정법원]} | |
| | {[H: 민사무소, 민주민센터]} | |
| | {[H: 세무서]} | {[M: 지사, 지서]} |
| | {[H: 선거관리위원회, 선관위, 선관위사무소]} | |
| | {[H: 소방서, 소방소, 소방파출서, 소방파출소, 119안전센터, 119안전센터]} | |
| | {[H: 우정국, 우편취급소, 우편취급국]} | |
| | {[H: 주민센터, 동사무소, 자치센터, 주민자치센터, 사무소, 민원센터, 민원사무소]} | |
| | {[H: 하수처리사업소, 사업소, 하수정화처리장, 하수처리장, 환경사업소]} | |

그림 5. 구축한 동의어집합의 일부

```

Algorithm: OverlayMatching(P, MC)
01: FOR each object of MC DO
02:   IF MC overlaps P THEN
03:     RESULT ← MakeMTable(P, MC, Overlay)
   ELSE
04:     ProximityMatching(P, MC)
   END IF
END FOR
05: RETURN RESULT
    
```

그림 6. Overlay 매칭 알고리즘

```

Algorithm: ProximityMatching(P, MC)
01: FOR each object of MC DO
02:   Calculate hausdorff_distance between P and MC
03:   RESULT ← MakeMTable(P, MC(MIN(hausdorff_distance)), Proximity)
END FOR
04: RETURN RESULT
    
```

그림 7. Proximity 매칭 알고리즘

3.2.3 Overlay 매칭

지번 및 명칭정보로 매칭여부를 판별할 수 없는 경우 공간객체의 좌표정보를 이용하였다. 후보객체인 건물 폴리곤과 POI 포인트의 중첩여부를 확인하여 POI 포인트를 포함하는 건물과 매칭 테이블을 작성한다(그림 6).

3.2.4 Proximity 매칭

POI 포인트와 후보객체 폴리곤과의 하우스도르프 거리를 계산하여 가장 가까운 건물객체와 매칭 테이블을 작성한다(그림 7).

3.3 매칭실험

실험에 사용된 데이터셋은 경기도 수원시 팔달구 지역의 도로명주소 전자지도의 건물객체(22,090개)와 동일지역의 CNS POI(2,588개), 팔달구 음식점이라는 검색어로 야후의 POI 검색 API를 통해 취득한 POI(100개)로 아래 그림과 같다. 본 연구에서는 실험의 편의를 위해 우선 CNS POI 중 전자지도 건물 폴리곤과 중첩되는 POI로 제한하여 실험을 진행하였으며, 알고리즘은 Python으로 ArcGIS 지오프로세싱 기능을 이용하였다.

3.4 실험결과

도로명주소 전자지도와 POI의 매칭 종류별 분포 점유율은 그림 9와 같다. 그림은 E(Equal), S(Semantic), O(Overlay), P(Proximity) 매칭의 비율을 보여준다.

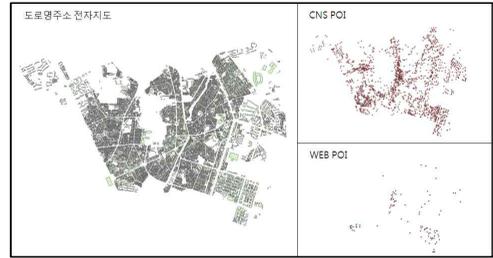


그림 8. 실험 데이터셋

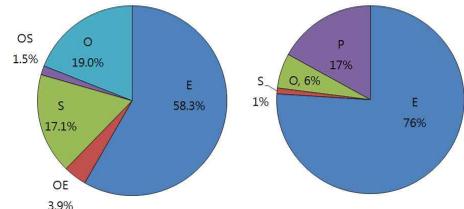


그림 9. CNS POI 매칭(좌)과 Web POI 매칭(우) 실험 결과

Web POI의 경우 모든 실험 데이터가 지번정보를 포함하고 있어 지번정보만으로도 다수의 매칭이 이루어졌다. 따라서 Semantic 매칭의 사례는 극히 드물었으며, OE(Overlay-Equal)와 OS(Overlay-Semantic) 매칭은 나타나지 않는 특징을 보였다. CNS POI의 경우 Web POI와 달리 모든 데이터가 지번정보를 가지고 있지 않다. 따라서 Equal 매칭의 사례가 Web POI 보다 낮은 비율을 보였으며, Semantic 매칭의 비율이 높았다. 특히 Semantic 매칭에서 기존의 단순 DBMS의 LIKE 연산자를 이용한 방법보다 본 연구에서 구축한 동의어집합을 이용한 유사도 분석을 통해 2배의 매칭 효과를 볼 수 있었다. CNS POI의 Proximity 매칭은 앞서 설명한 대로 건물폴리곤에 중첩되는 POI만을 실험 데이터셋으로 선정하였기 때문에 매칭 사례는 없다.

매칭종류 별 신뢰도 조사를 위해 매칭 데이터셋을 수동으로 구축하였다. 이를 바탕으로 분석한 결과 지번정보의 정비작업이 이루어졌다는 가정 하에 Equal 매칭은 100%의 신뢰도를 Semantic 매칭의 경우 약 98.5%, Overlay 매칭의 경우 약 85%의 신뢰도를 도출할 수 있었다. Web POI는 모든 데이터가 지번정보를 포함하고 있으며 도로명주소 전자지도 역시 모든 건물객체에 지번정보를 가지고 있다. 이런 사례에서 지번정보로 매칭이 이뤄지지 않는다는 것은 지번정보 변동에 대한 최신성이 확보되지 않은 경우이다. 실험에서는 신축건물에 대한 정보가 도로명주소 전자지도에 포함되지 않은 경

우 Proximity 매칭이 수행됨을 확인하였다. 따라서 Proximity 매칭의 신뢰도는 매우 낮아 Web POI와의 통합 이용에 활용될 수 없다고 판단된다.

4. 결론

본 연구에서는 이중의 공간객체를 매칭하는 기술을 개발하기 위해 공공에서 구축한 도로명주소 전자지도와 민간 POI정보의 공간객체 단위 매칭을 제안하였다. 객체매칭의 기준으로는 공간정보로 두 객체사이의 거리와 포인트 객체가 폴리곤 객체에 포함되는지의 여부와 지번 등의 주소 및 명칭의 속성정보를 사용하였다.

제안한 매칭방법은 매칭을 위해 사용되는 공간 및 속성정보에 따라 Equal, Semantic, Overlay, Proximity의 네 종류로 구분된다. Equal 매칭은 매칭후보 객체 간의 거리가 설정범위 이내이고 주소정보가 일치하는 경우이며, Semantic 매칭은 명칭의 유사도가 높은 경우, Overlay 매칭은 주소 및 명칭정보가 일치하지 않으면서 건물 폴리곤 내부에 매칭 대상 포인트가 위치하는 경우이고, 마지막으로 Proximity 매칭은 후보객체 중 가장 가까운 위치에 있는 경우이다.

본 연구에서는 실세계의 공간객체를 바탕으로 실험을 진행하였다. 도로명주소의 건물객체와 CNS 및 야후 지도에서 제공하는 POI의 실험 결과 제안한 방법이 기존의 좌표정보만을 이용한 매칭방법보다 정확율이나 재현율에서 높은 결과를 보였다.

향후 다양한 실세계 자료에 대한 실험을 통해 정교한 방법론 제시한다면 공공기관에서는 지도정보를 갱신하고 대시민 서비스를 향상하기 위하여 투입되는 많은 예산을 민간의 노하우를 이용해 비용절감과 서비스 질의 향상이 가능하고, 민간부문에서는 지형지물과 공공정보 수집을 위한 유지 관리 비용의 중복 투입 방식을 통해 실질적인 가치를 창출하는 위치정보 서비스 제공할 수 있을 것이다. 또한 제안한 매칭 종류 및 기준을 이용하여 자동 매칭에 의한 정보의 통합으로 공간정보 플랫폼 구축의 기반을 제공할 수 있을 것으로 기대된다.

감사의 글

본 연구는 국토해양부 첨단도시기술개발사업 - 지능형국토정보기술혁신사업과제의 연구비지원(07국토정보 C04)에 의해 수행되었습니다.

참고문헌

1. 강영옥 외, 2006, 서울시 GIS 현황 분석 평가 및 3단계 GIS 기본계획 수립연구, 서울시정개발연구원, pp.99-116, 291-294.
2. 김대식, 신소영, 서동권, 박현열, 최장원, 2002, "수치지도 주기 데이터의 가공 및 자동배치에 관한 연구", 2002년 추계학술대회논문집, *한국GIS학회*, pp.127-135.
3. 김은형, 2009, "Geospatial Web 플랫폼 기술 분석 및 기술개발전략", 한국GIS학회지, *한국GIS학회*, 제17권, 제2호, pp.171-181.
4. 네이버, 2009a, 오픈API, <http://dev.naver.com/openapi/tutorial>
5. 네이버, 2009b, 용어사전, <http://terms.naver.com/item.nhn?dirId=209&docId=22703>
6. 다음, 2009, 지도 API 2.0, <http://dna.daum.net/apis/maps>
7. 류재현, 2007, 수치지도상의 건물정보 데이터 통합 장치 및 방법, 10-2007-002127, 특허청
8. 매일경제, 2008, 인터넷포털의 지도검색 대전, <http://news.mk.co.kr/newsRead.php?year=2008&no=704137>
9. 야후, 2009a, POI 검색 API, <http://kr.open.gugi.yahoo.com/document/poisearch.php>
10. 행정안전부, 2009, 새주소 안내시스템, <http://www.juso.go.kr/>
11. Beerli, C. Kanza, Y. Safra, E. and Sagiv, Y., 2004, "Object fusion in geographic information systems", *In VLDB*, pp.816-827
12. Beerli, C. Doytscher, Y. Kanza, Y. Safra, E. and Sagiv, Y., 2005, "Finding corresponding objects when integrating several geo-spatial datasets", *In ACM-GIS*, pp.87-96.
13. Fu, G. Jones, C. B. Abdelmoty, A. A., "Building a Geographical Ontology for Intelligent Spatial Search on the Web", *Databases and Applications 2005, IASTED*, pp.167-172.
14. Hild, H. and Fritsch, D., 1998, "Integration of Vector Data and Satellite Imagery for Geocoding", *IAPRS*, Vol. 32, pp.246-251.
15. Safra, E. Kanza, Y. Sagiv, Y. and Doytscher, Y., 2006, "Integrating Data from Maps on the World-Wide Web", *W2GIS 2006, LNCS 4295*, pp.180-191.