

지적도를 이용한 주소정제 알고리즘 연구

A Study on the Algorithm of Address Symmetric using Cadastral Map

윤훈주*·류중희**·김소희***·김현중****·정종철*****

Yoon, Hun Joo · Ryu, Jung Hee · Kim, So Hee ·
Kim, Hyun Jong · Jeong, Jong Chul

- * 선도소프트 개발팀(hjyoon@sundosoft.com)
- ** 선도소프트 개발팀(jhryu@sundosoft.com)
- *** 선도소프트 개발팀(kimsh@sundosoft.com)
- **** 선도소프트 개발팀(hjkim@sundosoft.com)
- ***** 남서울대학교 교수(jcjeong@nsu.ac.kr)

요旨

본 연구의 목적은 지적기반의 주소 알고리즘의 개발을 통하여 주소체계의 표준화를 확립하고 GIS와 연계하여 전국을 하나의 정보권역으로 설정하기 위한 시도이다. 이를 위해 전자정부의 지적 정보 현황을 분석하고 행정정보와 지적정보의 통합 방안을 제시하여 주소 알고리즘의 활용방안을 모색하였다. 알고리즘 개발은 기존 행정망의 주소와 지적도 상의 주소를 일치시킨 것으로 GUI기반 객체언어로 구현하였으며, 매칭 성공률 분석하여 정확도를 살펴볼 수 있다. 본 연구는 민간과 공공 부문의 효율적인 업무연계를 위한 매개체로 활용되며, 자치단체의 정보화를 더욱 활성화할 것으로 사료된다.

핵심용어 : 지적도, 주소매칭, 전자정부, 지리정보, 공간·행정정보

Abstract

The purpose of this study is to create address standardization through an Algorithm based on cadastral maps, as well as to establish a method of transforming one information area to all of the domestic areas in relation to the geographic information system. To do this, it analyzed the current level of e-government cadastral data and their integration strategy with addresses of both the administrator and the cadastral data. By doing this it is be able to suggest an application method for the Algorithm. The algorithm proposed from this study have a built-in GUI and correlate addresses between the administration and the cadastral map, it can analyze matching success rates. This study is to provide a more efficient method for co-operative work between private sector and public sector. In addition, we can expect that the information strategy of local government will become more active.

Key words : Cadastral map, Address Matching, E-Government, GIS

1. 서론

국가 주도 정보화 산업의 큰 맥락을 이루고 있는 E-Government는 민간주도와 더불어 추진되고 있는 중점사항이다. 이 중 한

분야인 GIS산업은 지적기반에 많은 기초를 두고 있기 때문에 공공단체 및 지자체에서 사용되고 있으며 지적기반 GIS Data는 SOC사업의 초
석으로 자리매김하고 있다. 이러한 추세

에서 E-Government의 도래와 더불어 GIS 산업은 동적인 공간정보를 바탕으로 타부분 산업의 정보를 보다 빠르게 받아들일 수 있

이들의 통합을 보다 효율적으로 이루기 위하여 지적기반 GIS의 공간정보와 행정정보가 긴밀히 이루어져야 한다. 이를 위해 행정정보와 지적정보를 결합하는데 있어서 두 정보의 공통적인 요인을 찾아 서로 대응 관계를 성립시켜주는 주소 매칭이 필요하다. 그러나, 현재 국내에 존재하고 있는 주소 매칭 프로그램은 그 주소 매칭율이 높다고 할지라도 주소매칭의 대상이 공동주택 등으로 제한적인 것이어서 그 신뢰성에 문제가 있다. 또한, 주소일치의 기준을 정하는 것이 모호한 상태이다.

주소정제 기술은 복잡한 행정주소와 법정주소 사이에서의 기준을 제시하고 정보화 기준을 설정하기 위해 반드시 수행되어야 한다.

따라서 이러한 문제를 해결하기 위하여

지적도 기반의 주소 정제 알고리즘을 개발하고 통합하고 활용방안을 살펴보았다.

2. 전자정부의 구축을 통한 FAR 통합방안

현재 구축중인 전자정부는 민원 행정 중심의 대국민서비스에 초점을 맞춘 사업으로 행정 및 경제적 정보의 통합은 비교적 쉽게 이루어지고 있으나, 공간정보에 대한 활용성이 미약하다. 이를 위해 공간정보와의 연계를 위하여 지형 및 지적정보로 구성된 GIS와 연계되어 공간정보 중심의 행정 업무에 초점을 맞출 필요가 있다. 특히, 지적정보는 주민, 환경, 건축, 등기 등 8개의 업무와 공동 활용되어 사용되며, 온라인 자

는 공간정보, 행정정보, 경제정보가 통합된 FAR(Financial Administrative Real World)시스템 체계가 필요하다.

료 전송 주기가 10분 간격으로 이루어져 활용도가 높은 실정이다. 대표적으로 지적도면을 활용하여 현재 구축중인 사업으로 토지정보종합망과통합 중인 KLMIS(Korea Land Information System)와 농촌진흥정보체계가 있다. 이들은 지적기반의 공통 데이터를 가지고 토지행정 및 농지 행정 등 여러 분야에 걸쳐 활용 중이다. 현재 지적정보는 전자정부 전반에서 활용도가 높으며, 이런 상황에서 지적기반의 주소정제는 활용범위가 높을 뿐 아니라 정보화 전략의 핵심요인이 된다. 또한 공간정보 정보 중 주소 부문의 결합은 전자정부의 전반적인 업무를 연결 가능하게 한다. 이는 민간과 관청간의 1:1 맞춤형 정보를 제공하는 의사결정을 할수 있어 FAR통합을 지원하며 업무의 중복을 줄이고 자동화를 가능하게 하여 시간 및 경제적 비용을 절감할 수 있다.

아울러 전자정부의 FAR통합이 지리, 지역적 위치와 관련이 있는 비즈니스정보와 결합할 때 G2B의 연계를 통한 E-Government와 E-Business의 다양한 연계도 고려해 볼 수 있다.

3. 지적기반의 주소매칭 알고리즘

주소 매칭은 법정도 코드와 행정동 코드의 비교를 통해서 이루어지므로 지번 지적 검색을 통해서 상세 정보 획득하고, 적절한 Annotation의 활용을 통해서 수정 및 보완 작업 수행한다. 행정동과 법정동의 표준화를 위한 주요한 비교대상으로 전화번호, 우

편 번호, 주소 등을 활용한다. 활용 기준은 다음과 같다.

- 전화 번호 : 2, 3자리의 국번으로 갱신하며 2004년 6월 DDD갱신을 기준
- 우편 번호 : 2004년 6월 지번 변화를 기준
- 주소 : 개인 소유 주택과 공동주택을 구분하며 경우의 수를 최대한 삼입

정확한 알고리즘을 구현하기 위해서는 내·외부적 민관데이터의 통합, 표준화, 가구화 (Householding)를 통한 데이터의 정제가 선행적 수행되어야 한다. 이를 위해 레퍼런스 테이블의 비교와 매치키를 활용한다.

Reference Table에서는 행정데이터와 법 정데이터의 매치키를 찾아내 서로 비교분석 한다.

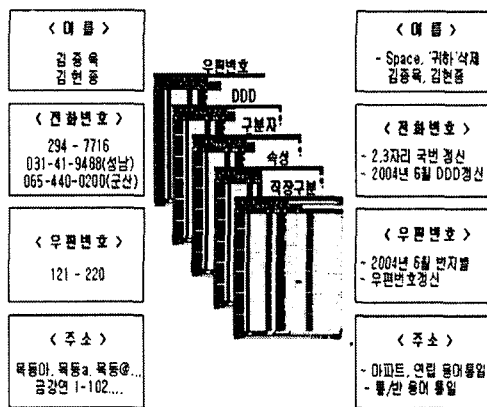


그림 2. 주소매칭을 위한 구성요소

또한 Match Key를 활용하여 데이터를 통합하고, Match된 자료를 가구단위로 Classification하여 하나의 범주로 설정한 후 Match Program을 활용하여 하나의 Data Base로 만든다.

그러나 이들이 완벽하지 않기 때문에 이름/주소/전화번호의 표준화와 오류 수정과정을 거쳐 우편번호를 갱신하고 추가해야

한다. 추가 및 갱신과정에서 Coding, Sorting 과정을 거친 다음 Merge/Purge에 의해 중복자 검색을 수행하며, Householding 기법을 통해서 두개의 영역을 통합한 매칭 프로그램을 완성한다.

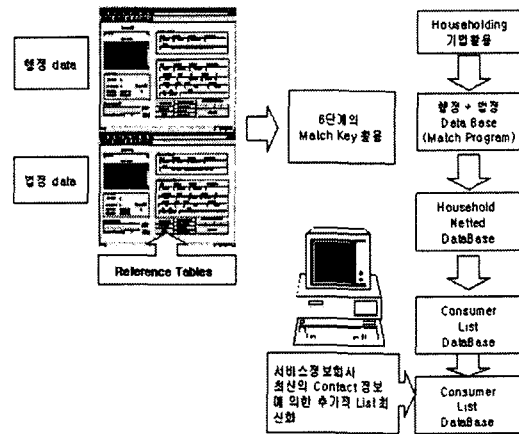


그림 3. 주소정제 과정 프로세스

3.1 주소 정제 알고리즘의 기능

본 연구의 알고리즘 기능은 크게 지번도 매칭부분, 공간데이터 추출부분, 주소매칭 서비스, 점수부분으로 나눌 수 있다. 각 부분의 세부기능을 다음의 표들과 같다.

표 1. 주소정재부분

분류	모듈명	주요기능	관련기술	비고
주소 정재 전처리	원본데이터 연결	1. 다양한 형태의 주소를 가진 원본 데이터를 범용적으로 연결 2. DB 컨넥터에 따른 특성 반영	1. Oracle 2. MSSQL 3. Informix 4. ADODB	
	원본데이터 필드 설정	1. 원본데이터에서 주소관련 필드를 선택 2. 1단계, 2단계, 3단계행정구역, 4단계, 5단계, 지번, 아파트, 건물명, 동, 호, 통, 반, 층수, 우편번호, 새주소, 참고 가능한 주소 필드	테이블 스키마 처리 기술	GUI 화면포함
	주소정재 관련 테이블 검사	1. 해당 원본데이터에 적합한 주소정재 레퍼런스 테이블에 대한 검사 2. 예외처리에 대한 부분 추가 - 예전 행정구역 주소, 완벽하지 않는 주소 유형		
	주소정재 고유키 생성	각각의 개별 레코드별로 고유 키 값 부여		
	인덱스 생성	원본데이터 및 관련 레퍼런스 테이블에 주소 정재관련 필드에 인덱스 반영	DB고유 인덱스 기술 반영	
	주소 처리 옵션 기능	1. 주소 분리 기능 - 행정구역단계별로 구분되지 않은 주소를 구별해 주는 기능 2. 인접지번 기능 - 해당 지번이 지번도에 없을 경우 인접한 지번에 찍어주는 기능	주소 예외 처리기술	GUI 화면포함
행정 구역 분리 기능	분리 단계 아래 단계코드별로 행정구역 표준화 작업 실행 0_주소없음 1_주소정재안됨 2_시도명검색완료 3_시군구명검색완료 4_읍면동명검색완료 5_리명검색완료 6_지번검색완료 7_지번없는아파트 8_지번있는아파트			
예외 처리	주소 정재 예외 처리	1. 행정구역명 변경 반영 2. 잘못 입력한 한글명 수정 기능		

표 2. 지번도 매칭 부분

분류	모듈명	주요기능	관련기술	비고
공간데이터 추출	공간데이터 생성	아래 단계코드별로 지번도 공간 추출 작업실행 - 101_법정동_완전조건_좌표완료 - 102_법정동_인접부번_좌표완료 - 103_법정동_인접본번_좌표완료 - 104_행정동_완전조건_좌표완료 - 105_행정동_인접부번_좌표완료 - 106_행정동_인접본번_좌표완료 - 107_건물_법정동_완전_좌표완료 - 108_건물_법정동_동제외_좌표완료 - 109_건물_행정동_완전_좌표완료 - 110_건물_행정동_동제외_좌표완료	지번도와 속성데이터 실시간 조인 기능	
	인접지번 선택기능	해당 지번도에 없는 지번을 찾음 - 산 번지를 고려한 인접지번 찾기 - 지번의 본번과 부번을 구별하거나 조합하여 검색		
	속성데이터 추가 기능	공간데이터에 부속적으로 필요한 속성데이터를 원본데이터로부터 가져와서 추가시킨다.		

표 3. 공간데이터 추출부분

분류	모듈명	주요기능	관련기술	비고
공간 데이터 추출	공간데이터생성	지번도 매칭결과 X,Y 좌표로 표현된 위치데이터를 공간데이터로 변환	ArcSDE	
	공간데이터 Export	SDE 상에 올라온 공간데이터를 정기적으로 백업 받는 기능		
	락 기능	ArcIMS 등으로 서비스 하고 있는 공간데이터 일 경우 락 모드를 설정하여 현재 서비스에 지장 받지 않도록 설정	공간데이터 처리	

표 4. 주소매칭 서비스

분류	모듈명	주요기능	관련기술	비고
서비스	타이머 기능	시간 간격을 설정하여 해당 시간에 누적된 주소매칭을 처리	윈도우 프로세스 관련 처리 기술	
	로그 기능	1. 주소매칭 시 발생하는 예외처리 및 서비스 시작 및 종료 등에 해당하는 어플리케이션 로그를 저장 2. 결과 엑셀파일이나 텍스트 파일로 Export		GUI 화면 포함
	즉시 실행	타이머에 의해 실행하지 않고 바로 주소정재를 처리하고자 할 때 실행		
DB변동 이력관리 부분	트리거 처리부분	주소관련 테이블에서 발생하는 주소변동에 대한 트리거 발생하고 이를 이력 관련 테이블에 반영	트리거 처리기술	
	동기화 부분	해당 주소 레코드의 추가/수정/삭제에 의해 공간 레코드의 추가/수정/삭제하여 동기화	MapObject	

표 5. 검수부분

분류	모듈명	주요기능	관련기술	비고
검수	주소정재 검수	주소정재 결과 후 각 단계별 결과를 엑셀파일이나 텍스트 파일로 Export할 수 있는 기능	Microsoft Office Contorl	
	지번도 매칭 검수	실제 주소매칭 결과가 지도상에서 제대로 찍혔는지 검수하는 부분	MapObject	
	결과 보정 작업	지번도와 기타 주소 정재 도움 테이블이 시간차에 의해 최신 데이터를 반영하지 못했거나, 오차범위로 인한 오류를 직접 수정할 수 있는 기능	ArcGIS	GUI 화면포함

3.2 FAR 통합을 위한 알고리즘

본 연구의 알고리즘개발에 있어서 가장 중요한 것은 현재 구축되어 있는 주소의 현

황을 파악하는 것이다. 기 구축되어 있는 자료의 형태는 제작부서, 구축기간, 전체 도엽수, 자료의 Format등으로 명시되어 있으

며 자료의 포맷은 CAD파일, Layer단위의 shape파일이 주를 이루고 있다. 여기서 가장 중요한 요소는 매칭과 정제로 이들은 기초적인 data를 제공한다.

알고리즘 개발은 주소 매칭과 주소정제 두 분야에 걸쳐서 이루어지며 민관주소의 수용여부와 데이터가 없는 민관에 대한 임시 위치도를 생성한다. FAR통합을 위한 알고리즘의 개발을 위해 중요한 구성요소는 다음과 같다.

- 법정 및 행정동 코드(DATA - 성명, 전화번호, 주소)
- Household 기법을 활용한 Match Program(자체제작)
- 데이터 통합작업을 위한 알고리즘 개발
- Spatial Data Cleaning 작업 수행

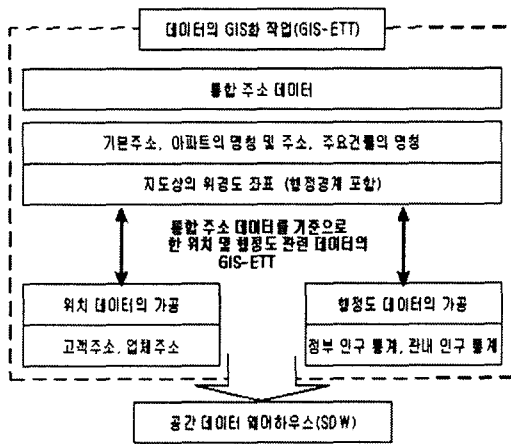


그림 3. FAR통합을 위한 알고리즘

3.3 GUI기반 객체언어로 구현 매칭 성공률 분석

작업의 간결성과 신속성을 위해서 GUI 기반 객체언어로의 구현하였다. 입력데이터의 형식은 DBF, XLS, TXT, SHP등 현재 상용되고 있는 대부분의 파일 형식을 지원하고 있다. 공간에 매칭시킬 주소필드와 우편번호, 전화번호 등을 선택하며 이외에도 서로 다른 속성의 좌표 데이터, 표준주소체

계, 우편번호 등의 형식을 불러올 수 있는 기능을 탑재한다.

프로그램 통해 매칭 과정이 수행되었음을 알리는 창이 생성되면서 통계적 수치를 통해서 주소매칭의 성공률을 화면에 출력한다. 매칭수행 전·후의 결과는 공간속에서 직접 비교해봄으로써 확인 가능하다.

GUI 구현시 고려사항은 다음과 같다.

- 구현원칙 : 간결성, 용이성
- 입력 데이터 형식 : 범용데이터 형식
- 매칭 필드 : 주소필드, 우편번호, 전화번호 등
- 검수방식 : 공간적 속성위치 이동여부
- 최종 결과 : 매칭율(Matching rate(%))

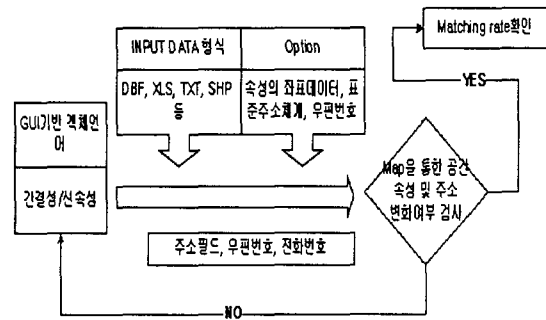


그림 4. GUI기반 객체언어 구현 알고리즘

3.4 매칭 성공률 분석

구현된 알고리즘을 서울시 강남구와 강원도 춘천시를 대상으로 적용하여 매칭 성공률을 분석하여 보았다. 도출된 결과는 다음과 같다.

강남구의 경우 정제 전 일치율은 50% 미만이었으나 주소정제 후 일치율이 80%로 향상되었고, 춘천시는 44%였던 일치율이 주소정제후 약 84% 개선된 것으로 나타났다. 불일치하는 부분은 현 알고리즘이 동기 준까지 정제되므로 법정동 리의 주소는 정제가 불가능하였으며, 행정동·법정동이 섞여있는 경우, 지하상가가 많아, 도로에 찍히는 포인트 많거나, APT, 빌라, 건물의 층 데이터의 부족으로 인함이다.

표 6. 주소매칭 산출 정제률

대상지역		서울특별시 강남구		강원도 춘천시		
단계별 대상	0_주소없음	0	748개 (정제안됨)	0	1,892개 (정제안됨)	
	1_주소정제안됨	0		0		
	2_시도명검색완료	0		0		
	3_시군구명검색완료	0		2		
	4_읍면동명검색완료	0		0		
	5_리명검색완료	158		1834		
	6_지번검색완료	0		0		
	7_지번없는아파트	590		56		
	8_지번있는아파트	0	0			
	사업체 정제수	101_법정동_완전조건_좌표완료	1593	2,971개 (정제됨)	3942	10,156개 (정제됨)
		102_법정동_인접부번_좌표완료	33		378	
103_법정동_인접본번_좌표완료		66	4600			
104_행정동_완전조건_좌표완료		774	220			
105_행정동_인접부번_좌표완료		15	118			
106_행정동_인접본번_좌표완료		109	898			
107_건물_법정동_완전_좌표완료		98	0			
108_건물_법정동_동제외_좌표완료		97	0			
109_건물_행정동_완전_좌표완료		60	0			
110_건물_행정동_동제외_좌표완료		126	0			
대상사업체 총 갯수		3,719개		12,048개		
대상사업체 정제갯수		2,971개		101,586개		
대상사업체 정제율		79.89%		84.30%		
데이터 특성		-행정동, 법정동이 섞여있음 -APT, 빌라, 건물의 층까지 상세한 주소		-법정동 주소에 리가 포함 됨(현주소정제는 동까지 정제) -행정구역명과 본번, 부번 만 있는 간략한 주소.(포 인트 위치정확도가 떨어 짐) -APT, 빌라, 건물의 층데 이터 없음(여러 사업체가 한점에 몰려 생기는 사항 발생) -지하상가가 많아, 도로에 찍히는 포인트 많음		

4. 기대효과

지적정보는 일제시대부터 유래된 부정확성을 내포하고 있기 때문에 사용자 측면에서 많은 어려움을 가지고 왔다.

본 연구의 알고리즘은 지적을 기반으로 하여 행정 및 법정, 전화번호와 통합되어 개발되었으므로 기존 주소체계의 부정확성을 감소시키고, 주소체계의 표준화를 정립할 수 있다.

또한 지적기반의 주소알고리즘은 주민정보, 토지의 이용형태, 용도지역의 지정 등 산재해 있던 주소를 포함하는 여러 정보를 통합하는데 용이하다. 이는 광범위하게 전자정부 전반에서 주민정보와 결합된 주소정부와 지적정보의 연계는 전국을 하나로 통합할 수 있는 기반을 마련한다. 또한 부분적으로 지자체에서 지적정보와 행정동 및 법정동의 잦은 매칭 업무에서 빠른 시간내에 원하는 정보를 찾을 수 있다. 대표적으로 지적도를 이용하는 보호관찰업무, 비지정 문화재 관리, 오염배출업소 관리, 산림병해충 관리 등의 업무가 있다. 이런 업무들은 모두 위치정보를 토대로 이루어진다는 점에서 GIS와 연계하여 지적기반 GIS 정보화방안을 구상 가능하게 한다.

민간 부분에서는 고객정보의 기본이 되는 주소정보를 통하여 지역분석에서 최적 입지의 용이한 선정과 공간의 영역별 특성 분석 및 개발전략 수립에 용이하다. 특히 지리, 지역적 위치와 관련이 있는 비즈니스 정보를 영업적으로 관리, 분석함으로써 지역적인 시장 혹은 민관의 분포현황을 파악하여 시장영역의 분할과 시장변화 예측 및 대책을 수립하는 상권분석을 할 수 있다.

이렇게 지적을 기반으로 이루어지는 주소매칭 프로그램의 활용은 공공 및 민간 부분에서 여러 형태로 활용되며, CRM개념을 도입하여 행정정보와 공간정보와의 통합이 이루어졌을 경우 전 국토가 하나의 정보권역으로 설정될 것이다.

5. 결론

본 연구를 통하여 다음과 같은 결론을 얻었다.

첫째, 전자정부에서 행정·경제·공간정보의 통합이 유기적으로 이루어지기 위하여 이들 간의 매개체로 주소알고리즘이 필요하다.

둘째, 주소 매칭은 행정동 코드와 법정동 코드의 비교와 지적도를 통해 이루어지므로 현주소체계의 표준화를 가능하게 한다.

셋째, 주소 정제알고리즘은 지번도 매칭 부분, 공간데이터 추출부분, 주소매칭 서비스, 검수부분으로 나누어 정제가 이루어진다. 서울시 강남구와 강원도 춘천시를 대상으로 알고리즘의 정확도 분석한 결과로 정제한 산출값은 각각 50%, 44%에서 80%, 84%로 상당히 개선되는 것으로 나타났다.

넷째, 위치 정보를 가지고 있는 주소매칭 알고리즘을 GIS산업과 연계하면 전국을 하나의 정보구역으로 설정할 수 있을 것으로 사료된다.

감사의 글

본 연구는 대한지적공사에서 수행하고 있는 “2004 정보통신 기술 응용분야 연구과제”의 관련 연구임을 밝히며 이에 깊은 감사를 드립니다.

참고문헌

1. 김은형(2003), 전자정부와 GIS연계방안 연구, 경원대학교.
2. 국가지리정보체계추진위원회(2003), 2004년도 국가지리정보체계 시행계획 (안). 건설교통부.
3. 재정경제부·행정자치부·정보통신부·보건복지부·노동부(2003). 전자정부 기반 완성 1주년 전자정부 추진 상황 브리핑 자료
4. 최병남(2002), 필지중심토지정보시스템과 토지관리 정보체계 통합시스템 구축 방안. 국토연구원
5. 더글러스 홈스(2002), 전자정부를 위한 e-비즈니스 전략, 물푸레.