

## 유동인구 추정 시 통신 자료의 활용에 관한 연구

김경태\* · 이인묵\*\* · 곽호찬\*\*\* · 민재홍\*

### Application Study of Telecommunication Record Data in Floating Population Estimation

Kyoung Tae Kim\* · In Mook Lee\*\* · Ho-Chan Kwak\*\*\* · Jae Hong Min\*

**요약** : 유동인구 자료는 도시의 동적 활동을 파악하기에 효과적인 지표로서 다양한 활용이 가능하지만, 이는 실제 움직이는 인구에 대한 자료이기 때문에 조사에 어려움이 있다. 최근 휴대전화의 사용이 보편화 되면서 국민 대부분이 사용하고 있기 때문에 휴대전화를 통해서 생성되는 통신 자료는 비용효율성 및 분석의 연속성 측면에서 기존의 표본 조사 방법과 더불어 유동인구를 산정하는 데 활용성이 높을 것으로 사료된다. 이에 본 연구에서는 유동인구 추정에 있어 통신 자료의 활용성에 대한 분석 결과를 제시하였다. 본 연구에서는 서울시 사례를 대상으로 분석을 수행하였으며, 분석 결과 통신 자료를 활용한 유동인구는 통계청에서 집계하는 주간인구 자료와 상관성이 상당히 높은 것으로 나타나, 통신 자료를 활용한 유동인구 추정 방안이 기존의 표본조사를 일부 대체할 수 있는 가능성을 확인하였다. 또한 통신 자료는 유동인구의 유입지에 대한 분석이 가능하기 때문에 향후 추가적인 연구를 통해 교통시설물의 공급 및 운영 방안에 대한 분석도 가능할 것으로 사료된다.

**주제어** : 통신 자료, 유동인구, 휴대전화

**ABSTRACT** : A floating population is an useful index to analyze dynamic activities of the cities. Especially the telecommunication record data collected from cell phone could be applied in estimation of floating population usefully, because of generalization of cell phone. Therefore, an estimation of the floating population using telecommunication record data was proposed in this study. The floating population in Seoul was estimated as a case. The results presented that the floating population using telecommunication record data has a high correlation with the daytime population collected from the National Statistical Office. Thus, the telecommunication record data can partly replace sample surveys to estimate floating population. Also it will be possible to analyze supply and operation of transportation facilities from the origin analysis of the floating population.

**Key Words** : telecommunication record data, floating population, cell phone

\* 한국철도기술연구원 녹색교통물류시스템공학연구소 책임연구원(Principal Researcher, Green Transport and Logistics Institute, Korea Railroad Research Institute)

\*\* 한국철도기술연구원 녹색교통물류시스템공학연구소 선임연구원(Senior Researcher, Green Transport and Logistics Institute, Korea Railroad Research Institute)

\*\*\* 한국철도기술연구원 녹색교통물류시스템공학연구소 선임연구원(Senior Researcher, Green Transport and Logistics Institute, Korea Railroad Research Institute), 교신저자(E-mail: kwak01@krii.re.kr, Tel: 031-460-5495)

## I. 서론

최근 교통 부문에서는 교통카드 자료 등의 빅데이터 활용에 대한 연구가 활성화되고 있는 상황이다(김경태 외, 2015; 김관호 외, 2013). 빅데이터의 경우 추가적인 조사 없이 개별 통행자 수준의 대규모 자료가 실시간으로 발생하기 때문에 비용효과적인 측면과 정밀도 측면에서 이전 자료에 비해 강점을 가진다.

2015년 통계청 자료에 따르면 우리나라는 이동통신 서비스의 발달로 국민 1인당 휴대전화 보급률이 2014년 기준으로 1.13대/인원으로 국민 1인당 1대 이상의 휴대전화를 가지고 있는 것으로 나타났다. 휴대전화의 경우 대부분의 사람들이 몸에 지니고 다닌다는 측면에서 휴대전화의 통신에 따라 발생하는 데이터는 전수에 가까운 자료로서, 해당 휴대전화 소유자의 통행패턴을 파악하는 데 상당히 유용하게 활용될 수 있다. 이러한 특성을 활용하여 일부 통신사에서는 통신 자료를 활용하여 관광지나 축제가 열리는 곳을 방문하는 고객들이 어느 곳에서 유입되는지에 대한 정보나 상권의 유동인구 등을 포함한 다양한 서비스를 제공하고 있다.

따라서 국민 대부분이 활용하고 있는 휴대전화를 통해서 생성되는 통신 자료는 여러 부문에서 매우 유용한 정보를 제공할 수 있을 것으로 판단되며, 통신 자료의 특성을 분석하고 활용 가능한 분야를 발굴하려는 노력이 필요한 시점이다. 그 일환으로서 본 연구에서는 통신사에서 제공하는 시간대별 유동인구<sup>1)</sup> 추정자료의 유용성을 인지하고 도시·교통 분야에서의 활용 가능성에 대한

분석을 하고자 한다.

## II. 이론적 고찰

최근에는 도시의 기능 분화로 유동인구가 특정 지역에 집중되면서 상주인구와 유동인구의 차이가 커지고 이에 따라 도시의 동적 활동을 파악하기에 효과적인 지표로서 유동인구에 대한 파악이 중요해지고 있다(정윤영·문태현, 2014).

이에 통계청에서는 1990년부터 5년 단위로 인구주택총조사의 표본조사로 12세 이상 통근 및 통학을 대상으로 하여 유입인구와 유출인구를 조사하여 주간인구와 주간인구지수 자료를 시·군·구 단위로 제공하고 있다. 주간인구는 해당 지역의 상주인구에 타 지역으로부터 유입된 통근·통학인구(주간유입인구)를 더하고 타 지역으로 유출되는 통근·통학인구(주간유출인구)를 뺀 것으로 정의하며, 이 외 사교적 레크리에이션, 쇼핑 및 비즈니스 목적으로 유출·유입하는 인구는 포함되지 않는다. 통계청에서 제공하는 주간인구와 주간인구지수의 산정 방식은 다음과 같다.

$$\text{주간인구} = \text{상주인구} + (\text{주간유입인구} - (\text{주간유출인구}))$$

$$\text{주간인구지수} = (\text{상주인구} + \text{유입인구} - \text{유출인구}) / \text{상주인구} \times 100$$

하지만 서울시와 같이 서비스업이 발달한 대도시의 경우, 통근·통학 외 기타 목적의 통행이 차지하는 비중이 상당히 높을 것으로 사료되는 바, 통근자와 통학자만을 기준으로 하는 주간인구 개

1) 유동인구의 개념은 이를 활용하는 주체에 따라서 달라질 수 있는데, 본 연구에서는 특정 지역에서 기지국과 통신한 휴대전화의 수를 이용하여 특정 지역에 존재하는 인구를 파악하려고 하였다.

념의 한계를 보완하기 위하여 변미리·서우석(2011)의 연구에서는 유동인구 개념을 제안하였다. 유동인구는 '특정 지역 내에서 일정 시간 동안 이동한 총 보행량'으로 정의하였으며, 서울시에서 2009년에 실시한 유동인구 조사 방법 및 주요 분석결과를 제시하였다. 또한 유동인구 조사자료와 센서스 기반 주간인구에 대한 비교분석을 통해 유동인구 조사자료의 특성을 규명하였다.

이처럼 유동인구는 실제 움직이는 인구에 대한 자료이기 때문에 조사를 위해 많은 시간과 비용이 소요되며, 조사 시기와 분석 시기의 격차로 인해 보다 효율적인 정책 수립에 어려움이 있을 수 있다. 하지만 휴대전화로부터 발생하는 통신 자료의 경우 추가적인 조사 없이 비교적 미시적 수준의 자료들이 실시간으로 제공된다는 측면에서, 기존의 조사 방법론에 비해 상당히 비용효율적이며 시간의 격차 없이 즉각적인 분석이 가능하다는 장점을 가지고 있다. 이에 본 연구에서는 서울시 유동인구 추정에 있어 휴대전화에서 발생하는 통신 자료의 활용성을 제안하고 검토하고자 한다.

### Ⅲ. 통신 자료의 특성 분석

#### 1. 자료 수집

통계청에 따르면 우리나라 2014년 추계인구는 50,423,955인 수준<sup>2)</sup>이다. 이에 비하여 이동전화서비스 가입자는 2014년 1월 54,823,782인에서 12월에는 57,207,957인 수준으로 증가하여 인구 1인당 이동전화서비스 가입자는 1.13으로 인구를 상회하고 있는 실정이다.

따라서 통신 자료는 국민 대다수가 이용하는 거의 전수에 가까운 자료라고 할 수 있으며, 이로부터 여러 가지 유용한 정보를 도출할 수 있다. 프라이버시 보호 문제로 이동전화서비스 가입자에 대한 직접적인 정보나 위치 정보는 파악할 수 없으나 기지국 단위로 집계화된 통신 자료를 활용하는 것만으로도 의미 있는 정보를 추출할 수 있다.

본 연구에서는 약 50%의 점유율을 가지고 있는 SKT 통신사가 제공하고 있는 시간대별 유동인구 자료를 토대로 하여 그 의미를 살펴보고, 현재 통계청에서 제공하는 주간인구와 서울시 유동인구

〈표 1〉 시간대별 유동인구 통신 자료의 구성

컬럼번호	내용	형식	설명
1	일자	문자열	4자리 연도, 2자리 월, 2자리 일 (예: 20140513)
2	셀 <sup>1)</sup> 의 X좌표	수	UTM-K 좌표체계 (소수점 4자리까지)
3	셀의 Y좌표	수	UTM-K 좌표체계 (소수점 4자리까지)
4~27	시간대별 유동인구	수	소수점 2자리까지 표시되며, 해당 셀에서 해당 시간대에 통신한 휴대전화 수로서 동일 시간대에서는 중복을 허락하지 않으며, SKT 통신사의 점유율을 고려하여 전수화한 자료임
28	유입지코드	문자열	전국의 법정동 기준 코드로서 8자리 체계

주 1) : 셀은 기본적으로 50m×50m 기준으로 구분되어 있으며, 전국 셀의 수는 2014년 5월 기준 3,712,040개이며 서울특별시는 175,448개임(산악, 강 등 통신이 전혀 발생하지 않는 곳은 고려되지 않음).

2) 통계청 국가통계포털([www.kosis.kr](http://www.kosis.kr)) 홈페이지 자료 인용

조사자료 등과 비교하여 그 활용성을 파악하고자 한다. SKT에서 제공하는 시간대별 유동인구 자료의 구성은 <표 1>과 같다.

## 2. 분석 자료의 특성

본 연구에서 분석한 통신 자료의 특성을 컬럼 별로 살펴보고, 기타 필요한 사항에 대한 설명을 추가하였다. 우선 일자는 2014년 5월 13일 화요일 자료를 이용하여 분석하였다. 자료의 수는 131,871,048 개 행 수준으로서 전체 자료를 분석하기에는 연산 시간이 많이 소요되어 이 중에서 일부 자료를 추

출하는 방식으로 분석하였다. 자료추출의 기준은 통신량이 가장 많은 시간대인 17~18시, 18~19시 시간대에 통신량이 0이 아닌 행을 선별하였으며, 추출된 자료는 모두 26,499,214개 행으로서 전체 자료의 약 20.1% 수준이다. 그러나 추출된 자료에 대한 유동인구의 합은 55,880,564로, 2014년 5월 13일 화요일 전체 유동인구인 62,179,623의 약 90% 수준으로 나타났으며, 이는 추출된 자료가 전체 자료를 충분히 대표할 수 있음을 의미한다.

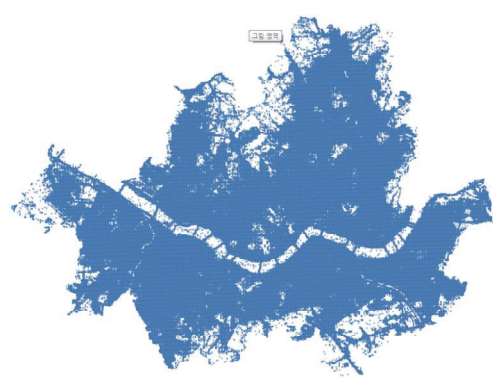
셀의 X, Y 좌표는 UTM-K 좌표체계를 따르며 소수점 4자리까지 표시된다. <표 1>에서 언급한 바와 같이 서울특별시의 셀 수는 175,448개 수준

<표 2> 서울특별시 구별 셀 수 및 면적당 셀 수(2014년 5월 기준)

구분	셀 수	면적(km <sup>2</sup> )당 셀 수	구분	셀 수	면적(km <sup>2</sup> )당 셀 수	구분	셀 수	면적(km <sup>2</sup> )당 셀 수
강남구	11,951	302	관악구	7,267	246	성동구	5,739	341
송파구	10,945	323	강동구	7,141	291	광진구	5,585	327
서초구	10,396	221	구로구	6,416	319	동대문구	5,432	382
강서구	9,887	239	양천구	6,368	366	도봉구	5,426	261
노원구	8,852	250	용산구	6,358	291	강북구	5,204	221
영등포구	7,992	325	종로구	6,092	255	금천구	4,401	338
은평구	7,621	257	동작구	5,883	360	중구	3,774	379
성북구	7,539	307	서대문구	5,876	334			
마포구	7,435	311	중랑구	5,868	317	합계	175,448	290



(a) 서울시 유동인구 조사지점 분포도(변미리·서우석, 2011)



(b) 본 연구의 셀 분포도

<그림 1> 서울시 유동인구 조사지점과 본 연구의 셀 비교

〈표 3〉 본 연구와 선행 조사 자료의 비교

구분	통신 자료 기반 유동인구 (본 연구)	통계청 통근·통학 조사	서울시 유동인구(보행통행량) 조사
조사방식	- 휴대전화 통신기록 이용	- 인구주택총조사 중 표본조사 항목	- 조사지점별 보행통행량 조사
표본 수	- 전체 휴대전화 가입자의 약 50%*	- 인구주택총조사 대상 중 10%	- 서울시 10,000개 지점
특징	<ul style="list-style-type: none"> <li>- 전국 약 370만 개의 셀로 자료 제공이 가능하며, 본 연구에서는 서울시 175,448개 셀</li> <li>- 이미 기록된 정보를 활용하므로, 조사기간의 제한은 없음</li> <li>- 셀별 유동인구를 유입지(주로 거주하는 곳) 기준으로 분류 가능</li> </ul>	<ul style="list-style-type: none"> <li>- 전국 시·군·구별 주간인구, 주간인구지수 제공</li> <li>- 13세 이상 연령대별로 제공</li> <li>- 주거지에 대한 조사는 있으나 통근·통학의 범위 내에서만 조사하여 업무용 출장, 쇼핑, 여가 등의 목적 방문은 제외됨</li> <li>- 5년 단위의 조사로서 시계열적 활용 제약</li> </ul>	<ul style="list-style-type: none"> <li>- 조사비용이 많이 소요되어 상시조사가 힘들고, 예산상황에 따라 조사지점이 유동적일 수 있음</li> <li>- 가로망상에서 조사지점을 선택하여, 행정구역 단위의 대포값 산정이 용이하지 않음</li> </ul>

\* 본 연구에서 제공받은 자료 기준임

이며, 구별로는 〈표 2〉와 같이 구성되어 있다. 강남 3구의 셀 수가 가장 많은 것으로 분석되었으며, 산지 및 강 등 통신이 발생하지 않는 지형을 많이 포함하고 있을수록 셀의 밀도가 떨어지는 것으로 분석되었다. 또한 셀의 경우 50m×50m의 정해진 면적을 커버하기 때문에 이에 따른 자료의 정확도 문제는 큰 영향을 미치지 않을 것으로 판단된다.

시간대별 유동인구는 1일을 24시간으로 구분하여 각 시간대별로 해당 셀에서 통신한 사람의 수<sup>3)</sup>를 나타낸다. 동일 시간대에서는 여러 번 통신한 기록이 있더라도 1번으로 계산하고, 시간대가 달라질 경우에는 중복하여 계산한다. 예를 들면, 6~7시 시간대에 10번의 통신 기록이 있더라도 1번으로 산정하며, 한 대의 휴대전화에 대해서 하루 최대 24회까지 기록될 수 있다.

마지막으로 유입지코드는 해당 셀에서 통신한 사람이 전달 심야시간대에 동일한 곳에 16일 이상

머문 곳을 법정동 기준으로 분류한 것으로서 전국을 5,050개 법정동으로 구분하였다. 위 기준에 부합하지 않는 경우는 불명으로 처리하였으며, 불명 자료의 합은 전체 자료의 약 13.3% 수준이다.

### 3. 기존 조사자료와의 차이점

기존의 분석자료와의 차이점은 공간적인 측면과 유동인구의 정의와 관련된 부분으로 정리하였다. 우선 통계청의 주간인구는 시·군·구 기준으로 작성되고 있고, 서울시에서 조사한 유동인구의 경우는 서울시 10,000개 지점을 대상으로 조사한 것이다. 그러나 본 자료는 서울시를 175,448개의 셀로 구분하였기 때문에 자료 수집 지점의 상세도는 서울시 유동인구 조사에 비해서 17.5배 수준이다.

통계청 주간인구는 상주인구와 통근 및 통학의 유출입을 이용하여 산정하고 있으며, 표본조사

3) 엄밀한 의미에서는 각 셀에서 통신(문자/통화 수발신, 데이터 통신 기록 기준이며, Wifi 통신은 제외됨)한 휴대전화 수이다.

〈표 4〉 통신 자료를 이용한 서울특별시 시간대별 유동인구

(단위: 천 명)

구분	시간대				합계	구분	시간대				합계
	00~06시	06~12시	12~18시	18~24시			00~06시	06~12시	12~18시	18~24시	
강남구	296	1489	2896	1955	6,636	서대문구	55	338	614	478	1,483
강동구	78	477	808	673	2,037	서초구	132	934	1719	1152	3,939
강북구	51	282	489	429	1,251	성동구	73	451	790	584	1,899
강서구	63	409	695	546	1,715	성북구	77	454	802	668	2,000
관악구	102	469	819	737	2,127	송파구	159	851	1413	1115	3,540
광진구	78	441	773	659	1,952	양천구	29	190	354	286	856
구로구	65	525	921	696	2,205	영등포구	102	786	1387	949	3,224
금천구	47	379	669	479	1,575	용산구	52	369	658	450	1,527
노원구	83	516	930	795	2,325	은평구	52	322	542	471	1,386
도봉구	44	272	463	407	1,185	종로구	73	540	1041	564	2,218
동대문구	81	481	852	637	2,050	중구	114	726	1406	760	3,005
동작구	66	393	670	575	1,704	중랑구	69	378	642	569	1,658
마포구	92	540	995	757	2,383	합계	2,134	13,010	23,353	17,387	55,881

(2010년 조사의 경우 표본율 10%)를 통하여 제공하고는 있으나, 표본율이 10% 수준으로서 자료의 신뢰성은 충분히 확보할 수 있는 것으로 판단된다. 그러나 주간인구의 산정에서 통근과 통학을 제외한 기타 목적의 유출입 수요는 고려할 수 없다는 단점이 존재한다. 예를 들면 명동 일대는 보행여건이 매우 열악한 혼잡지역이지만 이곳을 방문하는 사람들의 목적은 대부분 통근이나 통학보다는 쇼핑수요일 가능성이 높기 때문에 통계청의 주간인구에서는 이러한 현상을 반영할 수 없다는 한계가 있다.

서울시 유동인구(보행통행량) 자료는 구별로 10,000개 지점에 대해서 선정하는 기준을 6가지로 제시하고는 있으나, 이러한 공간표집은 전체 모집단을 대표하기 위한 표집을 체계적으로 할당하는 것이 불가능하다는 점을 인식하고 분석하였다. 또한 조사에 너무 많은 비용이 소요되어 연속적으로

조사하기에는 한계가 있고, 향후 조사 시에 가로망 구조가 변경되는 등의 여건 변동으로 조사지점이 달라진다면 이에 대한 해석이 연속적이지 않을 수도 있다.

따라서 통신 자료의 경우 자료 구득에 있어 기존 조사자료에 비해 비용효율적이며 시간적, 공간적 측면의 분석에 있어 연속성을 가진다는 장점이 있다. 또한 유동인구의 정의 측면에서 통신 자료의 특성은 상주인구와 보행자, 그리고 타 교통수단 이용자를 포함하고 있다는 측면에서 기존 연구(변미리·서우석, 2011)에서 제시된 보행자만을 대상으로 하는 유동인구와는 개념적 차이가 존재한다. 즉, 본 연구에서 제시한 유동인구의 개념은 통계청에서 제시하고 있는 주간인구의 개념과 보다 유사한 측면이 있으며, 통계청 자료에서 제시하지 못하는 통근·통학 외 기타 목적의 통행까지를 포함하고 있는 개념이다.



#### IV. 통신 자료를 이용한 유동인구 추정

##### 1. 통신 자료를 이용한 유동인구 추정 결과

본 연구에서는 통신 자료 중 50m×50m의 공간을 커버하는 셀 단위의 시간대별 통신 자료를 활용하여 서울시 행정구역별 유동인구를 추정하였으며, 데이터의 분석을 위하여 Oracle 프로그램을 사용하였다. SKT에서 제공하는 시간대별 통신 자료는 각 셀별 유입지별 해당 시간대에 통신을 발생시킨 인구를 하나의 행으로 제공하고 있다. 이에 본 연구에서는 각 셀의 좌표를 이용하여 이를 행정구역에 매칭시켰으며, 행정동 및 행정구 단위로 각 시간대별 유동인구를 추정하였다.

앞서 언급했듯이 본 연구에서 제시한 유동인구의 개념은 통신 자료의 특성을 감안하여, 상주인구와 보행자, 그리고 타 교통수단 이용자를 모두

포함하고 있다. 이를 근거로 서울특별시 구별 시간대별 유동인구를 추정한 결과는 <표 4>와 같다.

분석 결과 강남구의 유동인구가 다른 구에 비해 거의 1.5배 이상 높은 것으로 나타났으며, 서초구, 송파구, 영등포구, 중구 등의 순으로 유동인구가 많은 것으로 분석되었다. 또한 양천구, 도봉구, 강북구, 은평구 등의 지역은 서울에서 유동인구가 가장 낮은 수준으로 나타났다. 이는 대규모 업무 통행과 쇼핑 통행을 유발하는 지역이 유동인구가 많고, 주로 주거 기능을 담당하고 있는 지역의 경우 유동인구가 적음을 보여주고 있다.

시간대별 유동인구의 지역별 분포를 살펴본 결과 하루 단위의 분석결과와 큰 차이를 보이지는 않는 것으로 분석되었다. 강남구, 서초구, 송파구 등이 시간대별로도 상위 그룹을 형성하고 있었으며, 강북구, 도봉구, 양천구 등은 하위 그룹에 속하는 것으로 분석되었다.

<표 5> 통신 자료를 이용한 서울특별시 성별 유동인구

(단위: 천 명)

구분	성별		합계	구분	성별		합계
	남성	여성			남성	여성	
강남구	2,367	1,738	4,105	서대문구	494	435	929
강동구	627	519	1,146	서초구	1,500	1,056	2,556
강북구	350	338	687	성동구	664	475	1,139
강서구	826	669	1,496	성북구	601	542	1,143
관악구	660	560	1,220	송파구	1,162	923	2,085
광진구	676	536	1,212	양천구	592	532	1,125
구로구	711	548	1,259	영등포구	1,174	794	1,968
금천구	497	351	848	용산구	599	402	1,002
노원구	670	619	1,289	은평구	382	354	735
도봉구	343	310	653	종로구	771	553	1,324
동대문구	666	510	1,176	중구	1,057	730	1,787
동작구	562	470	1,032	중랑구	482	405	887
마포구	831	615	1,446	합계	19,263	14,986	34,248

시간대별 유동인구 분포를 분석해본 결과, 12~18시 사이 오후 시간대에 전체 유동인구의 약 40% 정도가 집중되는 것으로 나타났으며, 18~24시 사이 저녁 시간대에 약 30% 내외, 06~12시 사이 오전 시간대에 약 22~24% 정도 집중되는 것으로 분석되었다. 00~06시 심야 시간대에는 약 3~5% 정도의 유동인구 집중률을 보이는 것으로 나타났다.

종로구와 중구의 경우 다른 지역에 비해 오후 시간대 유동인구의 집중률이 상당히 높은 것으로 나타났으며, 오후 시간대와 저녁 시간대 유동인구의 비율 차이가 20% 이상 발생하는 것으로 분석되어 다른 통행에 비해 업무통행의 비율이 상당히 높은 것으로 추정할 수 있다.

통신 자료의 경우 가입자 정보를 통해 해당 휴대전화를 보유한 사람의 개별 특성을 집계된 형태로 이용 가능하며, 이를 통해 성별, 나이대별 유동

인구 패턴 분석이 가능하다. 이용자의 개별정보는 아직 시간대별로 제공하고 있지 않기 때문에 본 연구에서는 일 단위를 기준으로 성별, 나이대별 유동인구를 추정하였다. 시간대별 유동인구 추정과 마찬가지로 각 셀의 좌표를 활용하여 행정구역 단위로 각 셀을 매칭하였으며, 행정동 및 행정구 단위로 해당 유동인구를 추정하였다. <표 5>는 휴대전화 가입자 정보를 통해 구축한 성별 일간 유동인구를 나타내고 있으며, <표 6>은 나이대별 일간 유동인구를 분석한 결과이다.

성별 유동인구 분석결과를 살펴보면, 여성의 경우 남성에 비해 강서구, 노원구 등 주거인구가 비교적 밀집된 지역의 유동인구가 상대적으로 많은 것으로 나타났으며, 이는 여성의 경우 남성에 비해 근로자의 비중이 높지 않기 때문인 것으로 사료된다. 또한 연령대별 유동인구 분석결과를 살펴보면, 10대의 경우 다른 연령대에 비해 노원구, 강

<표 6> 통신 자료를 이용한 서울특별시 연령대별 유동인구

(단위: 천 명)

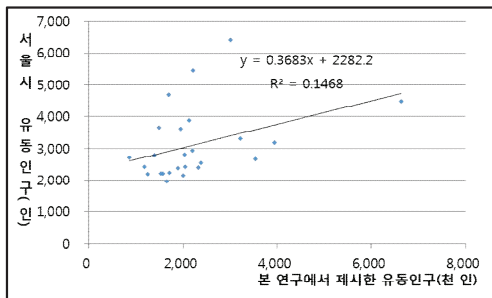
구분	연령대						합계	구분	연령대						합계
	10대	20대	30대	40대	50대	60대			10대	20대	30대	40대	50대	60대	
강남구	153	579	1,088	994	763	519	4,105	서대문구	58	177	184	199	184	127	929
강동구	74	125	239	271	274	162	1,146	서초구	95	320	613	632	533	362	2,556
강북구	49	85	133	157	156	107	687	성동구	57	161	253	266	244	159	1,139
강서구	102	165	328	363	342	196	1,496	성북구	89	184	230	256	231	154	1,143
관악구	77	208	271	250	244	170	1,220	송파구	117	247	469	506	463	282	2,085
광진구	72	190	254	269	260	168	1,212	양천구	100	115	226	288	256	140	1,125
구로구	73	157	294	295	268	171	1,259	영등포구	73	239	478	488	421	270	1,968
금천구	44	104	199	210	184	107	848	용산구	36	121	235	241	214	154	1,002
노원구	116	166	241	309	286	172	1,289	은평구	58	77	148	171	167	115	735
도봉구	48	72	119	152	158	103	653	종로구	46	188	304	314	282	190	1,324
동대문구	66	170	236	259	261	183	1,176	중구	45	239	441	441	372	247	1,787
동작구	66	178	223	219	205	141	1,032	중랑구	57	101	175	206	213	134	887
마포구	68	227	342	342	284	184	1,446	합계	1,841	4,535	7,733	8,098	7,265	4,716	34,248



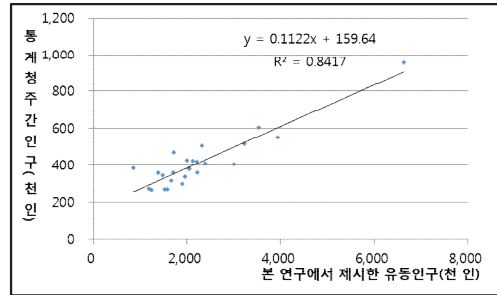
서구, 양천구 등 주거지역에서의 유동인구 순위가 높은 것으로 나타났으며, 20대의 경우 마포구, 관악구 등 대학가 주변의 유동인구가 상대적으로 높게 나타나는 특성을 보였다. 이는 초·중·고교 및 대학교 등에 등·하교하는 인구가 반영된 결과라 할 수 있다. 나머지 연령대의 경우 중심업무지구 및 일반적인 상업지구들이 밀집된 지역의 유동인구가 많은 것으로 분석되었다. 이처럼 통신 자료는 통행자 특성에 따른 유동인구 분석이 가능해 이를 통한 보다 효과적인 정책 수립에 활용될 수 있을 것이다.

## 2. 기존 자료와의 비교 분석

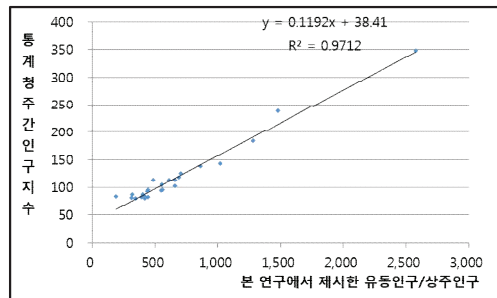
자료 작성의 시점은 다르지만 서울시의 유동인구 조사자료(2009년 시행)와 통계청의 주간인구(2010년 기준) 자료를 본 연구에서 도출한 자료와 비교하였다.<sup>4)</sup>



〈그림 2〉 본 연구와 서울시 유동인구의 상관관계



〈그림 3〉 본 연구와 통계청 주간인구의 상관관계



〈그림 4〉 본 연구와 주간인구지수의 상관관계

먼저 서울시 유동인구 결과 중에서 구별 평일 평균과 본 연구에서 도출한 구별 시간대별 자료 합계와의 상관계수를 분석한 결과 0.3832에 그치는 것으로 분석되었으나 통계청의 주간인구 자료는 상관계수가 0.9174로 매우 강한 양의 상관관계를 보이고 있음을 알 수 있다. 또한 본 연구에서 구축한 시간대별 유동인구의 합과 서울시 유동인구 사이의 결정계수 값은 0.1468 정도로 약한 상관관계를 가지지만, 통계청의 주간인구 자료와의 결정계수 값은 0.8417로 비교적 높은 상관관계를 가지는 것을 확인하였다. 이는 서울시 조사 자료의 경우 보행자를 분석대상으로 하고 있으나, 통계청 주간인구 자료 및 본 연구 결과는 보행자 및 상주

4) 서울시 유동인구(보행통행량) 조사에서는 통계청의 2005년 주간인구 자료와 비교하였으므로 이 자료를 활용할 수도 있으나, 이를 시점이 비슷한 2010년 통계청의 주간인구 및 주간인구지수와 비교하여도 주간인구와의 상관계수는 0.2597, 주간인구지수와의 상관계수는 0.7950으로서 비슷한 결과를 보여주고 있어, 통계청의 주간인구 자료는 최근 조사자료인 2010년 자료와 비교하였다.

인구를 포함하고 있기 때문인 것으로 사료된다.

또한 본 연구에서 도출한 유동인구와 2010년 통계청의 주간인구지수와의 상관계수는 0.4291로서 상관성이 떨어지는 것으로 분석되었지만 이것은 지표의 산정 방식에 따른 결과이다. 본 연구에서 도출한 유동인구를 상주인구로 나누어 준 값과 통계청의 주간인구지수를 비교하면 상관계수는 0.9855 수준이며, 결정계수 값은 0.9712 수준으로

매우 높은 일치성을 보여주고 있다. 즉, 통신 자료를 활용하여 추정된 본 연구의 결과는 통계청에서 제공하는 주간인구 및 주간인구지수를 장기적으로 대체·보완할 수 있을 것으로 판단되며, 특히 통근·통학을 제외한 다른 통행 특성을 반영할 수 있다는 측면에서 통계청 자료에 비해 활용성이 높을 것으로 판단된다.

## V. 결론

본 연구에서는 유동인구 추정에 있어 통신 자료의 활용성에 대한 다양한 분석을 통해 통신 자료가 기존의 설문조사 또는 직접조사 방식에 의존하던 것을 일부 대체할 수 있는 가능성을 확인하였다. 그러나 본 연구에서 제시한 부분은 통신 자료를 통해서 얻을 수 있는 정보의 극히 일부분에 불과할 것으로 판단된다. 향후 통신 자료를 활용할 수 있는 다양한 가능성을 열어놓고 이에 대한 분석이 필요할 것으로 판단된다.

앞서 언급된 것처럼 통신 자료는 유동인구 추정에 유용하게 활용될 수 있을 것으로 판단되나, 유동인구에 대한 보다 자세한 특성을 밝히기 위해서는 기존의 조사 방법 및 추가적인 자료 구축이 병행되어야 할 필요가 있다. 또한 유동인구 추정에 있어 해당 시간대에 동일 지점에 머무르지만 휴대전화 통신을 하지 않음으로써 해서 제외된 이용자 및 교통수단을 이용해 단순히 통과하는 이용자를 확인하여 자료를 보완하는 방안, 그리고 통신 자료의 전수화 과정에 있어 목적통행에 따른 편차 검토 및 지역별, 시간대별 특성 반영에 대한 추가적인 연구가 수행될 필요가 있다.

본 연구에서는 구축된 통신 자료의 일부만을 추출하여 분석을 수행하였기 때문에 빅데이터 활

〈표 7〉 서울특별시 구별 본 자료와 통계청 자료 정리  
(단위: 천 명)

구분	유동인구 (A)	통계청 주간인구	통계청 상주인구 (B)	$A \div B$ $\times 100$	통계청 주간 인구지수
강남구	6,636	953	518	1,281	184
강동구	2,037	379	457	445	83
강북구	1,251	265	320	391	83
강서구	1,715	468	538	319	87
관악구	2,127	423	508	419	83
광진구	1,952	336	355	549	94
구로구	2,205	417	399	552	105
금천구	1,575	271	229	688	118
노원구	2,325	502	578	402	87
도봉구	1,185	274	343	345	80
동대문구	2,050	379	337	609	113
동작구	1,704	360	387	440	93
마포구	2,383	407	361	660	113
서대문구	1,483	344	305	486	113
서초구	3,939	548	386	1,021	142
성동구	1,899	297	288	658	103
성북구	2,000	427	449	446	95
송파구	3,540	602	634	558	95
양천구	856	384	461	186	83
영등포구	3,224	514	373	865	138
용산구	1,527	270	217	704	125
은평구	1,386	359	444	312	81
종로구	2,218	358	150	1,480	239
중구	3,005	406	117	2,577	348
중랑구	1,658	317	396	419	80
합계	55,881	10,262	9,550	585	109

주: 본 연구의 유동인구 자료는 2014년 5월 13일 기준이며,  
통계청 자료는 2010년 자료임

용의 장점을 극대화하는 데는 한계를 가지고 있다. 따라서 통신 자료 전체에 대한 보다 광범위한 분석이 수행될 필요가 있으며, 시간적 범위도 하루치보다는 넓게 1주일 또는 1개월 정도의 자료를 이용하여 요일별 변동 등의 특성에 대한 분석이 필요하다.

마지막으로 유동인구 추정에 있어 통신 자료의 활용성을 극대화하기 위해서는 다음과 같은 두 가지 연구가 수행될 필요가 있다. 먼저, 통계청의 주간인구나 주간인구지수와 비슷한 형태의 지수를 만들어서 기존의 표본조사에 의존하던 주간인구 지수를 보다 현실적으로 반영할 수 있는 지수를 만드는 것이 필요하다. 또한, 유동인구의 유입지 자료를 활용하여 유동인구의 OD 구축에 대한 연구를 통해 유동인구에 대한 보다 구체적인 분석 및 정책 결정이 이루어질 수 있을 것이다. 이러한 연구를 통해 통신 자료의 활용은 교통시설물의 공급 및 운영에 대한 시사점을 도출하고, 수요자 중심의 교통 인프라 시설 구축 및 행정력의 배분 등에 유용하게 활용될 수 있을 것이다.

## 참고문헌

- 김경태 · 이인목 · 박호찬 · 민재홍, 2015, “스마트카드 자료를 활용한 서울시 버스 노선의 굴곡도 평가”, 『서울도시연구』, 16(1): 33~42, 서울연구원.
- 김관호 · 오규협 · 이영규 · 정재윤, 2013, “스마트카드 빅데이터를 이용한 서울시 지하철 이동패턴 분석”, 『한국전자거래학회지』, 18(3): 211~222, 한국전자거래학회.
- 변미리 · 서우석, 2011, “도시 거리의 주간활동인구 측정과 해석: 서울시 유동인구 조사 사례”, 『조사연구』, 12(2): 27~50, 한국조사연구학회.
- 정윤영 · 문태현, 2014, “유동인구 자료를 이용한 서울시 도시공간구조 분석 연구: ‘2030 서울플랜’과 비교연구”, 『한국지역개발학회지』, 26(3): 139~158, 한국지역개발학회.

<http://kosis.kr/>

원 고 접 수 일 : 2015년 3월 20일  
1차심사완료일 : 2015년 5월 13일  
2차심사완료일 : 2015년 7월 21일  
최종원고채택일 : 2015년 8월 3일