

이동통신 자료 전수화를 통한 존재인구 산정 방안

김경태 · 이인목 · 곽호찬* · 민재홍

한국철도기술연구원 녹색교통물류시스템공학연구소

An Estimation of Occupancy Population Using the Expanded Mobile Phone Data

KIM, Kyoung Tae · LEE, Inmook · KWAK, Ho-Chan* · MIN, Jae Hong

Green Transport and Logistics Institute, Korea Railroad Research Institute, Gyeonggi 16105, Korea

*Corresponding author: kwak01@krii.re.kr

Abstract

Recently, mobile phone data was applied in travel demand modeling as a new source of dynamic population movement. This study is also aimed to estimate "occupancy population" during a given period of time within a given spatial region using mobile phone data. An occupancy population was defined as the number of people residing or moving within a given time and space. In case of Seoul Metropolitan area, we divided the area into a number of administrative districts as zones for analysis and estimated the occupancy population of each zone by mobile phone data collected by SK telecom Co., a wireless telecommunication provider in Korea. For the expansion of mobile phone data, a new concept of "communication probability" was introduced and applied in the estimation of occupancy population of each zone by the hour. We compared the estimated number with the daytime population and the daytime population index referred by the Statistics Korea. The results showed that a positive correlation existed between the estimated number and the statistical number by nationwide survey. It was concluded that mobile phone data could be more cost-effective sources than a conventional survey method to estimate the pattern of population movement by the hour or by the day.

Keywords: communication probability, daytime population, daytime population index, expanded mobile phone data, occupancy population, occupancy population index

초록

본 연구에서는 이동통신 자료를 활용하여 존재인구를 산정하기 위한 방안을 제시하였다. 존재인구란 특정 시간 및 공간 단위에 거주하거나 활동하는 인구로 정의하였다. 이를 위하여 SKT에서 제공하는 이동통신 자료를 활용하였으며, 통신확률 개념을 적용하여 이동통신 자료의 전수화를 통한 존재인구 산정 방안을 제시하였다. 서울시 개별 구를 대상으로 시간대별 존재인구를 산정한 후, 이를 활용하여 일간 존재인구 및 존재인구지수를 도출하였으며, 통계청에서 제공하고 있는 주간인구 및 주간인구지수와 비교분석을 통해 두 값 사이에 강한 양의 상관관계가 있음을 제시하였다. 이처럼 시간대별, 일간 존재인구 산정에 이동통신 자료를 활용할 경우 기존의 조사 방법에 비해 비용 효율적이며, 합리적이고 일관성 있는 결과를 도출할 수 있을 것으로 판단되며, 본 연구에서 제시된 시간대별, 요일별 패턴 분석 등을 통해 보다 합리적인 정책 수립과 행정력 배분이 가능할 것으로 사료된다.

주요어: 통신확률, 주간인구, 주간인구 지수, 이동통신 자료 전수화, 존재인구, 존재인구지수

J. Korean Soc. Transp.
Vol.34, No.3, pp.222-233, June 2016
<http://dx.doi.org/10.7470/jkst.2016.34.3.222>

pISSN : 1229-1366
eISSN : 2234-4217

Received: 13 January 2016

Revised: 18 April 2016

Accepted: 30 June 2016

Copyright ©
Korean Society of Transportation

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

서론

현대사회에서는 직장과 주거의 불일치율이 계속 높아지고 있기 때문에 많은 사람들이 직장이나 학교 등 자기가 거주하는 지역이 아닌 일터에서 하루 대부분의 시간을 보내는 경우가 많다(Eun, 2001). 따라서 주간활동인구 자료는 도시의 동적 활동을 파악하는데 효과적인 지표로서, 도시 및 교통 등 공공분야 뿐만 아니라 홍보, 영업 등 민간 영역에서도 활용 가치가 높다. 하지만 이는 실제 활동하고 있는 인구에 대한 자료이기 때문에 조사 및 측정에 한계가 있다. 현재 통계청에서는 개개인이 통상적으로 거주하는 지역을 기준으로 인구를 파악하는 개념인 상주인구 외에, 낮 시간 동안 통근이나 통학 등 직업 및 학업과 관련하여 이동하는 인구를 의미하는 주간인구 자료를 배포하고 있다. 하지만 주간인구의 이동은 통근 및 통학 외 사교적 레크레이션의 목적, 쇼핑과 비즈니스의 목적에 의해서도 발생하며(Foley, 1954), 통계청에서 제공하는 주간인구의 개념은 이들 목적의 통행은 반영되지 않기 때문에 서비스 경제가 차지하는 비율이 높은 서울과 같은 대도시의 주간활동인구를 파악하는 데에는 한계가 있다. 이러한 통근자와 통학자만을 기준으로 한 유입자와 유출자 간의 차이를 파악하는 주간인구 개념의 한계를 보완하기 위하여 Byun and Seo(2011)의 연구에서는 유동인구의 개념을 제안하였다. 유동인구를 특정 지역 내에서 일정시간 동안 이동한 총 보행량으로 정의하였으며, 개별 거리 위주로 측정되는 일반적인 보행량에 비해 유동인구는 개별 거리들을 포괄하는 단위 지역 내에서 발생하는 보행량의 총합으로 개념화하였다.

이처럼 주간활동인구 지표의 중요성이 증대되고 있는 현 시점에서 본 논문에서는 국민 1인이 평균적으로 1대 이상의 이동전화를 보유하고 있는 점(2014년 기준 1.13대/인), 그리고 이동전화 통신기록 자료에 위치정보를 포함하고 있다는 점에 착안하여 이동전화 통신기록 자료를 주간활동인구의 추정에도 활용하고자 한다. 구체적으로는 이동전화 통신에서 발생하는 데이터를 활용하여, 기존 표본조사 기반 자료의 신뢰성 한계를 극복하고 더 나아가 특정 시간대 및 공간에 존재하고 있는 존재인구의 개념을 제안하고, 이를 추정하기 위한 방법을 제시하고자 한다. 이를 통하여 기존의 주간인구나 유동인구 등의 개념과는 차별화된 지표를 구축할 수 있을 것이다.

Kim et al.(2015a)의 연구에서는 통신사에서 제공하고 있는 셀(50m×50m 단위로 구분된 공간) 단위의 이동통신 자료를 행정구역 단위로 집계하여, 이를 통계청의 주간인구 및 서울시 유동인구(보행통행량) 조사 결과와 비교하였으며, 그 결과 이동통신 자료를 유동인구 추정에 활용할 수 있는 가능성을 제시하였다. 비교 결과 이동통신 자료를 통해 추정되는 유동인구는 도로변 보행통행량보다는 해당 지역에 현존하는 인구의 개념으로 통계청 주간인구의 개념과 더 유사함을 알 수 있었다.

일 단위로 추정되는 통계청 주간인구에 비해 이동통신 자료의 경우 시간대별로 인구분포 특성을 파악할 수 있고, 통근 및 통학 목적 외 쇼핑, 비즈니스, 사교 등의 목적을 모두 포함하는 개념으로서 본 연구에서는 특정 시간대에 해당 지역에 거주하거나 활동하는 인구를 존재 인구라는 용어를 정의하고 이를 산정하기 위한 방안을 제시하였다. 또한 현재 일부 통행 목적에 대한 샘플링을 통해 산출되는 통계청 주간인구 및 주간인구지수의 한계를 극복하기 위해 본 연구에서 제시된 시간대별 존재인구를 활용하여 일간 존재인구 및 존재인구지수를 산정했으며, 이를 기존의 통계청 주간인구 및 주간인구지수와 비교하였다.

이론적 고찰

통계청은 도시의 동적 활동을 파악하기 위해 1990년부터 5년 단위로 인구주택총조사의 표본조사로써 12세 이상의 통근 및 통학을 대상으로 유입·유출 인구를 조사하여 주간인구와 주간인구지수 자료를 시·군·구 단위로 제공하고 있으며, 이에 대한 산정 방식은 Equation 1 및 Equation 2와 같다(Kim et al., 2015).

$$\text{주간인구} = \text{상주인구} + \text{유입인구} - \text{유출인구} \quad (1)$$

$$\text{주간인구지수} = (\text{상주인구} + \text{유입인구} - \text{유출인구}) / \text{상주인구} \times 100 \quad (2)$$

여기서, 상주인구는 해당 시·군·구에 거주하는 인구 수를 뜻하며, 유입인구는 해당 시·군·구에 활동을 위해 유입된 인구 수, 유출인구는 다른 시·군·구로 유출된 인구의 수를 의미한다.

통계청 조사 자료의 경우, 2010년에는 10% 표본율로 주간인구와 주간인구지수를 산출하였으며, 통근 및 통학 외 기타 목적의 통행은 반영되지 않는 한계가 있다.

Byun and Seo(2011)의 연구에서는 도시의 주간활동인구를 파악하기 위하여 유동인구의 개념을 활용하였으며 서울시 유동인구 조사 방법 및 주요 분석결과를 제시하였다. 이는 10,000개 지점에 대한 보행통행량 조사를 통해 유동인구를 산출하여 정책의 활용도를 높였으나, 본 연구에서 제안하는 존재인구와는 개념상 차이가 있다. 또한 이는 표본조사에 의한 방법으로 조사지점의 선정 방식 및 시기에 따라서 조사의 표준화가 어렵고, 조사 자원 및 질의 한계에 따라 시공간적으로 유동적일 수 있다는 한계가 있다.

Kim et al.(2015b)의 연구에서는 국토교통 분야에 대한 모바일 빅데이터의 활용성을 제시하였다. 국토 분야에서는 도시공간 구조 및 사회조사 비용 절감, 공간 위계별 동적 행정수요 파악, 그리고 개인 활동을 고려한 정책수립 지원 등의 시사점을 도출하였으며, 교통 분야에서는 공간위계별 이동패턴의 실시간 파악, 교통량 보완자료, 기종점 통행량 구축의 신뢰도 제고 등에 대한 활용성을 제시하였다. 하지만 이는 일부 샘플을 활용한 사례분석을 기반으로 모바일 빅데이터 활용성에 대한 정책적 시사점을 도출하였을 뿐, 구체적인 분석 결과를 제시하지는 않았다.

Kim et al.(2015a)의 연구에서는 유동인구 산정에 이동통신 자료의 활용 방안을 검토하였으며, 서울시 자료를 활용하여 사례분석을 수행하였다. 하지만 이는 통신이 발생한 데이터만을 대상으로 유동인구를 산정하였으며, 통신이 발생하지 않은 대상에 대한 전수화 과정이 누락되어 실제 유동인구와는 차이가 있다. 따라서 본 연구에서는 통신이 발생하지 않은 대상에 대한 전수화 과정을 포함하여 특정 공간 및 시간에 존재하는 인구로서 존재인구를 정의하고, 이를 활용하여 통근 및 통학에 한정된 통계청에서 제공하는 주간인구 및 주간인구지수의 의미를 보다 확장할 수 있는 방법론을 제시하고자 한다.

자료 수집

본 연구에서는 이동통신사(SKT)에서 제공되는 시간대별 통신 자료를 기반으로 분석을 수행하였다. SKT는 전체 통신 인구의 약 50% 가량을 점유하고 있어 자료의 전수화가 용이하다는 장점이 있다. 개인정보 보호 문제로 서비스 가입자에 대한 직접적인 정보를 구득하기에는 한계가 있어 기지국 단위로 집계된 자료를 분석에 활용하였으며, 본 연구에서 활용된 자료의 구성은 Table 1과 같다.

Table 1. Structure of mobile phone data

Columns	Contents	Types	Descriptions
1	Date	String	4-digit year, 2-digit month, 2-digit day (Example: 20140513)
2	X coordinate of cell	Number	UTM-K coordinate system
3	Y coordinate of cell	Number	UTM-K coordinate system
4-27	Communicating population	Number	Hourly number of mobile phone communicated in the cell
28	Origin code	String	8-digit code of "dong(TAZ)" unit

시간대별 통신 자료는 셀 단위로 집계되며, 셀은 기본적으로 50m×50m 크기로 설정되어 있다. 2014년 5월 기준 전국의 셀 수는 3,712,040개이며, 서울시의 셀 수는 175,448개로 구성되어 있다. 산악이나 강 등 통신이 전혀 발생하지 않는 지역은 셀이 존재하지 않으며, 셀의 X, Y 좌표는 UTM-K 좌표계를 따른다.

시간대별 통신인구는 1일을 24시간으로 구분하여 각 시간대에 해당 셀에서 통신한 사람의 수를 의미한다. 여기서

통신이라 함은 문자 및 통화 수발신, 데이터 통신 기록을 기준으로 하며, 와이파이 통신은 집계에서 제외된다. 또한 동일 시간대에 여러 번 통신한 기록은 한 번으로 계산되고, 시간대가 다를 경우에는 중복 계산한다. 즉, 한 대의 휴대 전화에 대해 하루 최대 24회까지 기록될 수 있다.

본 연구에서는 서울시, 인천시, 경기도에 대해 2014년 5월 10일부터 5월 16일까지 1주일 동안 구득된 자료를 분석에 활용하였으며, 존재인구 및 존재인구지수 산출의 공간적 범위는 서울시로 한정하였다(Table 2). 해당 기간 동안 수도권에서는 일평균 119,748천인의 통신이 이루어졌으며, 지역별로는 서울시, 경기도, 인천시 순으로 통신이 많이 일어난 것으로 나타났다. 토요일의 경우 평일 대비 81-85% 수준의 통신이 수도권에서 이루어지는 것으로 나타났다. 일요일에는 평일 대비 57-60% 수준의 통신이 이루어지는 것으로 분석되었다. 또한 시간대별 분포를 살펴본 결과 12-21시 사이에 통신이 집중적으로 이루어지는 것으로 나타났으며, 다음으로 09-12시, 21-24시의 순으로 통신량이 많은 것으로 분석되었다.

Table 2. Summary of communicating population data

Date	Region	Communicating population (1,000 persons)								Total
		00-03	03-06	06-09	09-12	12-15	15-18	18-21	21-24	
10 May, 2014	Seoul	2,495	969	2,706	8,905	10,520	10,151	9,565	5,921	51,232
	Incheon	497	186	805	2,378	2,676	2,618	2,472	1,425	13,058
(Sat)	Gyeonggi	1,644	614	2,727	8,332	9,289	8,986	8,598	4,814	45,003
	Sum	4,636	1,769	6,239	19,615	22,484	21,755	20,635	12,160	109,292
11 May, 2014	Seoul	1,926	738	1,722	5,253	7,389	7,690	7,091	4,047	35,857
	Incheon	443	178	514	1,428	1,988	2,007	1,794	936	9,288
(Sun)	Gyeonggi	1,416	554	1,751	5,057	6,786	7,034	6,396	3,324	32,317
	Sum	3,785	1,470	3,987	11,738	16,164	16,730	15,281	8,307	77,462
10 May, 2014	Seoul	2,495	969	2,706	8,905	10,520	10,151	9,565	5,921	51,232
	Incheon	497	186	805	2,378	2,676	2,618	2,472	1,425	13,058
(Sat)	Gyeonggi	1,644	614	2,727	8,332	9,289	8,986	8,598	4,814	45,003
	Sum	4,636	1,769	6,239	19,615	22,484	21,755	20,635	12,160	109,292
11 May, 2014	Seoul	1,926	738	1,722	5,253	7,389	7,690	7,091	4,047	35,857
	Incheon	443	178	514	1,428	1,988	2,007	1,794	936	9,288
(Sun)	Gyeonggi	1,416	554	1,751	5,057	6,786	7,034	6,396	3,324	32,317
	Sum	3,785	1,470	3,987	11,738	16,164	16,730	15,281	8,307	77,462
12 May, 2014	Seoul	995	474	3,361	11,313	12,529	14,549	12,298	6,437	61,954
	Incheon	210	110	972	2,694	2,972	3,460	2,965	1,438	14,820
(Mon)	Gyeonggi	712	368	3,399	9,375	10,086	11,930	10,405	5,069	51,344
	Sum	1,917	952	7,731	23,382	25,587	29,939	25,667	12,944	128,119
13 May, 2014	Seoul	1,656	670	3,500	11,255	12,127	14,195	12,226	6,537	62,167
	Incheon	313	136	1,006	2,647	2,832	3,368	2,910	1,466	14,677
(Tue)	Gyeonggi	1,075	451	3,468	9,287	9,757	11,768	10,379	5,177	51,362
	Sum	3,044	1,257	7,974	23,189	24,716	29,331	25,516	13,180	128,206
14 May, 2014	Seoul	1,688	651	3,411	11,150	12,255	14,236	12,388	6,640	62,419
	Incheon	319	130	971	2,589	2,826	3,363	2,921	1,497	14,615
(Wed)	Gyeonggi	1,105	439	3,396	9,218	9,930	11,825	10,438	5,246	51,597
	Sum	3,111	1,219	7,778	22,957	25,011	29,424	25,747	13,383	128,631
15 May, 2014	Seoul	1,743	693	3,748	11,629	13,118	14,493	12,515	6,738	64,678
	Incheon	333	141	1,051	2,655	2,986	3,532	2,998	1,487	15,184
(Thu)	Gyeonggi	1,140	471	3,718	9,337	10,207	12,095	10,538	5,277	52,783
	Sum	3,216	1,306	8,516	23,621	26,311	30,121	26,052	13,503	132,645
16 May, 2014	Seoul	1,756	671	3,538	11,518	12,431	14,354	12,909	7,467	64,643
	Incheon	338	138	1,052	2,762	2,920	3,425	3,098	1,715	15,448
(Fri)	Gyeonggi	1,141	428	3,577	9,574	10,111	12,036	10,954	5,972	53,793
	Sum	3,235	1,236	8,167	23,854	25,461	29,815	26,960	15,155	133,884

존재인구 및 존재인구지수 산정 방법론

1. 통신 자료의 특성 및 전수화 방안

앞서 제시된 시간대별 통신인구 자료는 해당 시간 통신이 이루어진 인구에 대한 자료로서, 보다 정확한 존재인구 산정을 위해서는 통신이 이루어지지 않은 인구에 대한 전수화 과정이 필요하다. 이에 본 연구에서는 통신확률이라는 개념을 도입하여 통신이 이루어지지 않은 인구에 대한 전수화를 수행하였다. 본 연구에서 정의하는 통신확률은 전체 인구 대비 통신이 이루어진 인구의 비율을 뜻하며, 분석 자료로부터 구득되는 통신인구를 통신확률로 나누어 줌으로써 통신이 이루어지지 않은 인구를 포함하는 전수화 과정을 수행하였다.

통신 자료로부터 통신확률을 도출하기 위해서는 특정 지역 내 기지국에 포함되는 전체 이동전화 수 대비 통신이 발생한 이동전화 수의 비율을 산정하기 위한 자료가 구축되어야 한다. 통신이 발생한 이동전화에 대한 집계자료의 경우 통신사 입장에서 요금 부과와 관련 있는 부분이기 때문에 이에 대한 자료가 구축되어 있지만, 기지국 범위 내의 전체 이동전화에 대한 집계 자료는 관리비용 대비 활용도가 미미하기 때문에 현재 집계되지 않고 있다. 따라서 본 연구에서는 이에 대한 대체변수로서 특정 지역 내 주민등록상 인구수를 설정하여 통신확률을 산정하고자 한다.

이러한 방법으로 통신확률 개념을 존재인구 산정에 적용하기 위해서는 세 가지 가정이 필요하다. 먼저, 통신확률 산정에 활용되는 지역의 유입·유출량은 동일하다고 가정하였다. 이는 분석대상 지역의 전체 인구는 일정하게 유지되며, 이에 따라 분석대상 지역 기준으로 유출 및 유입량은 같다는 것을 의미한다.

KTDB에서 제공하는 수도권 OD자료에 따르면 2012년 기준 서울에서 기타 지역으로의 통행량은 1일 3,972,701통행, 기타 지역에서 서울시로의 통행은 3,822,158통행으로, 그 차이는 150,543통행으로서 2014년 서울시 추계인구인 9,890,661인의 1.5% 수준을 차지한다. 수도권으로 그 범위를 확대하면, 수도권에서 기타 지역으로의 통행은 727,024통행, 기타 지역에서 수도권으로의 통행은 786,974통행으로서 그 차이는 59,950통행이며, 2014년 수도권의 추계인구인 25,029,687인의 0.2% 수준에 불과하다. 이처럼 수도권의 경우 대부분 수도권 내부에서 통행이 이루어지고 외부와의 통행량은 무시할 수 있을 정도로 작기 때문에 본 연구에서는 수도권 지역을 대상으로 통신확률을 산정하였다.

또한 본 연구에서는 이동전화 소지자의 위치에 관계없이 수도권 내 모든 인구의 통신확률은 동일하다고 가정하였다. 실제적으로는 수도권 내 각 지역별 인구 특성에 따라 통신확률 역시 차이를 보이겠지만, 본 연구에서는 자료 구득의 한계 상 통신확률을 수도권 내 전체 평균값으로 산정하였다. 마지막으로 1인당 1대의 이동통신 기기를 이용한다고 가정하였다.

이러한 가정에 의거하여 각 시간대별 수도권 지역의 통신확률은 Equation 3과 같이 표현될 수 있다.

$$\begin{aligned} (\text{통신확률})_{\text{수도권}, t} &= \frac{(\text{통신전화수})_{\text{수도권}, t}}{(\text{전체전화수})_{\text{수도권}, t}} \\ &= \frac{(\text{통신인구})_{\text{수도권}, t}}{(\text{상주인구})_{\text{수도권}}} \end{aligned} \quad (3)$$

여기서, $(\text{통신확률})_{\text{수도권}, t}$ 는 t 시간대 수도권의 통신확률, $(\text{통신전화수})_{\text{수도권}, t}$ 는 t 시간대 수도권에서 통신이 이루어진 이동전화 수, $(\text{전체전화수})_{\text{수도권}, t}$ 는 t 시간대 수도권의 전체 이동전화 수, $(\text{통신인구})_{\text{수도권}, t}$ 는 t 시간대 수도권의 통신인구 수, 그리고 $(\text{상주인구})_{\text{수도권}}$ 는 수도권의 상주인구수를 의미한다.

또한, 수도권의 시간대별 통신확률 산정을 위하여 본 연구에서는 수도권 통신인구는 1주일치 자료의 평균값을, 상주인구는 2014년 기준 서울시, 인천시, 경기도 인구의 합을 각각 활용하였다. 이에 따라 산정된 수도권 지역의 통신인구 및 통신확률은 Figure 1과 같다.

시간대별 통신확률을 살펴보면 00-09시 및 22시 이후는 20% 이하의 낮은 통신확률을 가지는 것으로 나타났으

며, 이 경우 통신인구 및 통신확률을 활용하여 시간대별 존재인구를 산출하기에는 추출되는 표본 비율이 너무 낮아 전수화의 신뢰도가 높지 않을 수 있다. 따라서 본 연구에서는 20% 이상의 통신확률을 가지며 사람들의 활동이 가장 활발히 이루어지는 09-21시까지 12시간의 자료를 기반으로 존재인구를 산정하였다.

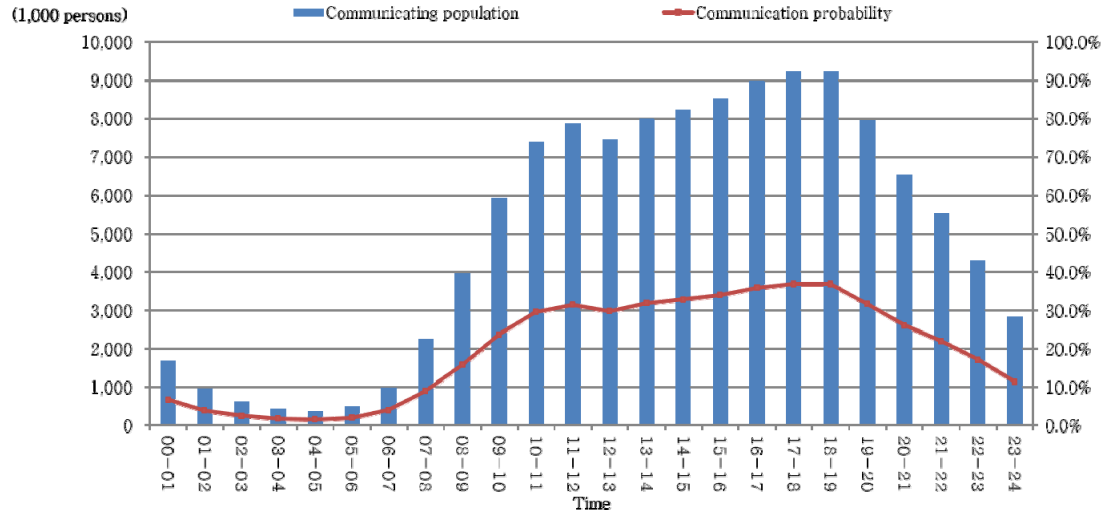


Figure 1. Hourly communicating population and communication probability in Seoul metropolitan area

2. 시간대별 존재인구 및 존재인구지수 산정

앞서 언급되었듯이, 시간대별 존재인구는 통신인구를 통신확률로 나누어 줌으로써 계산되며, 서울시 구별 시간대별 존재인구는 Equation 4에 의해 계산된다.

$$(\text{존재인구})_{i,t} = \frac{(\text{통신인구})_{i,t}}{(\text{통신확률})_{\text{수도권},t}} \quad (4)$$

여기서, $(\text{존재인구})_{i,t}$ 는 서울시 i 구의 t 시간대 존재인구 수, $(\text{통신인구})_{i,t}$ 는 서울시 i 구의 t 시간대 통신인구 수, 그리고 $(\text{통신확률})_{\text{수도권},t}$ 은 수도권의 t 시간대 평균 통신확률을 의미한다.

이에 따른 서울시 시간대별 존재인구 산정 결과는 Table 3에 제시하였다. 그 결과를 지역별로 살펴보면, 강남구, 서초구, 송파구, 영등포구, 그리고 중구의 순으로 존재인구가 많은 것으로 나타났으며, 양천구, 도봉구, 강북구 등의 지역이 가장 적은 것으로 분석되었다. 또한 서울시 전체의 시간대별 존재인구는 11-15시 사이에 비교적 높은 것으로 나타났으며, 17시 이후에는 비교적 일정한 수준을 유지하는 것으로 분석되었다. 이 결과를 토대로 서울시의 경우 사무실이 밀집되어 있는 지역, 그리고 점심시간 전후에 가장 활발한 활동이 이루어지는 것으로 판단할 수 있다.

앞서 산정된 시간대별 존재인구를 기반으로 본 연구에서는 상주인구 대비 시간대별 존재인구의 비율을 의미하는 시간대별 존재인구지수를 산출하였다. 시간대별 존재인구지수는 시간대별 존재인구를 상주인구로 나누어줌으로써 계산된다. 즉, 존재인구지수가 100일 경우 해당 시간대 존재인구와 상주인구의 수가 같음을 뜻하며, 존재인구지수가 100보다 클 경우 상주인구 대비 존재인구수가 더 많고, 100보다 작을 경우 상주인구 대비 존재인구수가 더 적음을 의미한다.

$$\begin{aligned}
 (\text{존재인구지수})_{i,t} &= \frac{(\text{존재인구})_{i,t}}{(\text{상주인구})_i} \\
 &= \frac{(\text{통신인구})_{i,t}}{(\text{통신확률})_{\text{수도권},t}} / (\text{상주인구})_i
 \end{aligned}
 \tag{5}$$

여기서, $(\text{존재인구지수})_{i,t}$ 는 서울시 i 구의 t 시간대 존재인구지수, $(\text{존재인구})_{i,t}$ 는 서울시 i 구의 t 시간대 존재인구 수, $(\text{상주인구})_i$ 는 서울시 i 구의 상주인구 수, $(\text{통신인구})_{i,t}$ 는 서울시 i 구의 t 시간대 통신인구 수, 그리고 $(\text{통신확률})_{\text{수도권},t}$ 은 수도권의 t 시간대 평균 통신확률을 의미한다. 본 연구에서는 상주인구로서 2014년 기준 서울시 구별 인구수 자료를 활용하였다. 이에 따라 산정된 서울시 구별 시간대별 존재인구지수는 Table 4와 같다.

강남구, 종로구, 그리고 중구의 경우 상주인구 대비 2배 이상의 존재인구를 가지는 것으로 나타났으며, 서초구, 영등포구 등도 비교적 높은 값을 가지는 것으로 분석되었다. 강서구, 도봉구, 양천구, 은평구 등은 상주인구 대비 존재인구의 수가 더 적은 것으로 나타났으며, 이들 지역은 비교적 서비스 경제가 발달하지 않은 지역으로 판단된다. 시간대별 흐름을 살펴보면, 주간에 지수가 높은 지역의 경우 야간으로 갈수록 낮아지는 경향을 가지고 있으며, 주간에 지수가 낮은 지역은 야간에 다시 높아지는 경향을 보이는 것으로 나타났다. 이는 주간에 업무지구나 상업지구에서의 용무가 끝난 인구가 야간에 집으로 돌아오는 현상으로 해석될 수 있다.

Table 3. Time-based occupancy population by district in Seoul (unit: 1,000 persons)

District	Time											
	09-10	10-11	11-12	12-13	13-14	14-15	15-16	16-17	17-18	18-19	19-20	20-21
Gangnam	1,284	1,372	1,441	1,376	1,424	1,451	1,435	1,402	1,375	1,363	1,291	1,231
Gangdong	449	448	442	440	436	438	439	445	442	450	476	490
Gangbuk	266	263	261	269	270	266	265	269	269	273	300	313
Gangseo	384	378	380	380	382	383	386	377	374	377	383	384
Gwanak	434	436	441	453	452	447	444	449	450	464	504	530
Gwangjin	410	414	416	422	426	424	422	424	424	433	464	476
Guro	475	481	476	474	475	473	466	466	466	484	490	471
Geumcheon	336	338	336	327	340	335	331	331	330	342	340	325
Nowon	475	479	476	491	483	487	498	513	517	513	543	569
Dobong	257	254	245	250	247	248	250	257	258	260	282	293
Dongdaemun	431	438	442	455	455	443	437	436	432	427	436	439
Dongjak	364	366	363	366	369	364	358	365	369	380	393	402
Mapo	485	500	518	501	515	525	512	511	511	526	522	518
Seodaemun	316	322	330	335	329	333	330	329	327	329	325	338
Seocho	835	876	900	850	873	886	873	859	843	836	794	748
Seongdong	420	417	415	408	412	407	400	400	399	399	397	391
Seongbuk	421	420	424	435	442	427	428	434	433	431	454	472
Songpa	789	774	762	741	752	759	761	761	756	763	773	785
Yangcheon	173	175	177	183	184	187	192	192	189	189	199	203
Yeongdeungpo	687	711	723	683	704	704	694	679	672	693	650	606
Yongsan	343	355	359	345	350	356	356	347	342	336	329	308
Eunpyeong	297	292	287	292	289	287	286	298	294	296	319	333
Jongno	450	495	545	533	540	535	516	492	467	445	388	348
Jung	574	630	697	669	692	683	661	630	602	578	504	452
Jungnang	348	345	339	348	348	339	343	350	352	358	405	414
Seoul	11,703	11,979	12,196	12,026	12,191	12,187	12,082	12,016	11,893	11,944	11,961	p11,839

Table 4. Time-based occupancy population index by district in Seoul

District	Time											
	09-10	10-11	11-12	12-13	13-14	14-15	15-16	16-17	17-18	18-19	19-20	20-21
Gangnam	248	265	278	266	275	280	277	271	265	263	249	238
Gangdong	98	98	97	96	95	96	96	97	97	98	104	107
Gangbuk	83	82	82	84	85	83	83	84	84	85	94	98
Gangseo	71	70	71	71	71	71	72	70	70	70	71	71
Gwanak	86	86	87	89	89	88	87	88	89	91	99	104
Gwangjin	115	116	117	119	120	119	119	119	119	122	130	134
Guro	119	121	119	119	119	119	117	117	117	121	123	118
Geumcheon	147	148	147	143	149	146	145	145	144	150	148	142
Nowon	82	83	82	85	83	84	86	89	89	89	94	98
Dobong	75	74	72	73	72	72	73	75	75	76	82	85
Dongdaemun	128	130	131	135	135	132	130	129	128	127	130	130
Dongjak	94	94	94	94	95	94	92	94	95	98	101	104
Mapo	134	139	143	139	143	145	142	142	141	146	145	143
Seodaemun	103	105	108	110	108	109	108	108	107	108	107	111
Seocho	216	227	233	220	226	230	226	223	219	217	206	194
Seongdong	146	145	144	142	143	141	139	139	138	138	138	136
Seongbuk	94	94	95	97	99	95	95	97	97	96	101	105
Songpa	124	122	120	117	119	120	120	120	119	120	122	124
Yangcheon	37	38	38	40	40	40	42	42	41	41	43	44
Yeongdeungpo	184	191	194	183	189	189	186	182	180	186	174	163
Yongsan	158	164	165	159	162	164	164	160	158	155	152	142
Eunpyeong	67	66	65	66	65	65	64	67	66	67	72	75
Jongno	301	330	364	356	360	357	344	328	311	297	259	232
Jung	492	540	598	574	593	586	567	540	516	495	432	387
Jungnang	88	87	86	88	88	86	87	88	89	90	102	105
Seoul	123	125	128	126	128	128	127	126	125	125	125	124

3. 요일별 존재인구 및 존재인구지수 산정

존재인구 산정을 위해서 통신 자료를 활용할 경우, 해당 지역의 시간대별 평균적인 패턴 뿐 아니라 요일별 패턴에 대한 분석을 통해 보다 합리적인 정책 수립이 가능하기 때문에 본 연구에서는 요일별 분석을 수행하였다. 요일별 분석 결과는 09-21시 사이의 시간대별 존재인구 및 존재인구지수에 대한 평균으로 산정하였다. 이에 따른 요일별 분석 결과는 Table 5에 제시하였다. 서울시 전체의 요일별 변화를 살펴보면, 존재인구는 목요일에 가장 높은 값을 가지는 것으로 나타났으며, 금요일, 월요일의 순으로 높은 것으로 분석되었다. 특히, 주말인 토요일과 일요일에는 평일에 비해 상대적으로 낮은 값을 보이는 것으로 나타났으며, 일요일의 경우 존재인구지수가 100 이하로 일주일 중 유일하게 상주인구에 비해 존재인구의 수가 적은 것으로 분석되었다. 이는 일요일에 서울 외곽으로 빠져나가는 유출인구의 비중이 유입인구에 비해 상대적으로 높기 때문인 것으로 판단된다.

서울시 존재인구의 요일별 변동성을 살펴보기 위하여 본 연구에서는 변동계수를 활용하였다. 변동계수는 해당 샘플의 표준편차를 평균으로 나누어준 값으로 샘플들의 상대적인 변동률을 살펴볼 수 있는 지표이다. 이에 따라 서울시 구별 존재인구의 변동계수에 대한 산정결과를 Table 6에 제시하였다.

서울시 전체의 경우 평균 존재인구는 약 1,200만 명으로 나타났으며, 평균 존재인구 대비 요일별 변동률은 약 20% 정도로 분석되었다. 서울시 구별 존재인구 변동률을 살펴본 결과, 중구의 요일별 변동률이 약 34%로 가장 큰 것으로 나타났으며, 종로구(30%), 강남구(28%)의 순으로 요일별 변동이 큰 것으로 분석되었다. 특히, 존재인구지수가 100이하로 유출인구가 유입인구에 비해 많은 지역의 경우 존재인구의 변동률 역시 상대적으로 낮은 수준인 것으로 나타났다.

Table 5. Daily occupancy population and index by days of the week in Seoul

District	Daily occupancy population (1,000 persons)							Daily occupancy population index						
	Sat	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri
Gangnam	1,011	655	1,590	1,570	1,570	1,603	1,595	195	126	307	303	303	309	308
Gangdong	425	316	481	470	475	492	489	93	69	105	103	104	108	107
Gangbuk	268	206	289	282	283	296	291	84	65	91	88	88	93	91
Gangseo	360	263	418	402	402	402	418	67	49	78	75	75	75	78
Gwanak	449	345	487	472	477	496	486	88	68	96	93	94	98	96
Gwangjin	413	291	457	448	453	476	470	116	82	128	126	127	134	132
Guro	399	295	535	519	506	532	537	100	74	134	130	127	133	135
Geumcheon	267	191	381	369	369	378	383	117	83	167	161	161	165	167
Nowon	496	372	528	518	525	550	536	86	64	91	90	91	95	93
Dobong	255	198	272	265	267	280	272	74	58	79	77	78	82	79
Dongdaemun	392	271	487	473	478	494	480	116	81	145	140	142	147	143
Dongjak	346	250	402	391	394	414	404	89	64	104	101	102	107	104
Mapo	439	291	559	557	563	594	580	122	81	155	154	156	165	161
Seodaemun	286	203	357	355	358	377	363	94	67	117	116	117	123	119
Seocho	665	431	969	958	955	977	981	172	112	251	248	247	253	254
Seongdong	348	225	453	444	447	463	459	121	78	157	154	155	160	159
Seongbuk	408	296	465	459	464	486	468	91	66	104	102	103	108	104
Songpa	693	485	835	813	818	858	851	109	76	132	128	129	135	134
Yangcheon	182	128	197	194	197	207	202	39	28	43	42	43	45	44
Yeongdeungpo	527	364	779	766	769	793	788	141	98	209	205	206	213	211
Yongsan	289	195	378	377	380	393	397	133	90	174	174	175	181	183
Eunpyeong	296	224	320	314	303	318	308	67	50	72	71	68	72	69
Jongno	352	206	546	553	557	575	568	235	137	364	369	372	384	379
Jung	411	228	716	725	729	744	745	353	196	614	622	625	638	639
Jungnang	353	264	380	365	373	384	382	89	67	96	92	94	97	97
Seoul	10,329	7,192	13,281	13,057	13,113	13,581	13,455	108	75	139	137	137	142	141

Table 6. Day-to-day variation of occupancy population

(unit: 1,000 persons)

District	AVG	SD	CV
Gangnam	1,370	382	0.28
Gangdong	450	63	0.14
Gangbuk	274	31	0.11
Gangseo	381	56	0.15
Gwanak	459	52	0.11
Gwangjin	430	65	0.15
Guro	475	93	0.20
Geumcheon	334	75	0.23
Nowon	504	61	0.12
Dobong	258	28	0.11
Dongdaemun	439	82	0.19
Dongjak	371	58	0.16
Mapo	512	110	0.21
Seodaemun	329	62	0.19
Seocho	848	216	0.25
Seongdong	405	89	0.22
Seongbuk	435	66	0.15
Songpa	765	135	0.18
Yangcheon	187	27	0.14
Yeongdeungpo	684	170	0.25
Yongsan	344	75	0.22
Eunpyeong	297	33	0.11
Jongno	479	144	0.30
Jung	614	208	0.34
Jungnang	357	43	0.12
Seoul	12,001	2,399	0.20

이처럼 존재인구 산정에 통신 자료를 활용할 경우, 특정 시간대에 인구의 평균적인 공간적 분포 뿐 아니라, 요일별, 그리고 특정 기간대의 날짜별 인구의 유입 및 유출 패턴에 대한 파악을 통해 보다 효율적인 정책 수립 및 행정력 배분이 가능하다는 장점이 있다.

4. 산정 결과 비교

본 연구의 분석 결과를 통계청 주간인구 및 주간인구지수와 비교한 결과를 Table 7에 제시하였다. 통신 자료를 활용하여 산정한 서울시 존재인구 및 존재인구지수는 각각 12,001천 인, 126으로 나타났으며, 통계청에서 발표한 주간인구 및 주간인구지수는 각각 10,370천 인, 109로 분석되었다. 즉, 본 연구에서 제시한 존재인구의 절대적인 값이 통계청에서 산정한 주간인구에 비해 더 큰 것으로 나타났는데, 이는 통계청 자료의 경우 통근·통학 인구만을 반영했다는 측면에서 본 연구의 결과는 합리적인 것으로 판단된다.

본 연구에서 도출한 서울시 구별 존재인구 및 존재인구지수를 통계청의 주간인구 및 주간인구지수와 비교한 결과, 존재인구 및 주간인구의 상관관계는 0.9186으로 아주 강한 양의 상관성을 가지는 것으로 나타났으며, R2값은 0.8438로 분석되었다(Figure 2). 존재인구지수 및 주간인구지수 역시 두 자료 사이의 상관관계는 0.9857로 강한 양의 상관성을 가지는 것으로 나타났으며, R2값은 0.9717로 분석되었다(Figure 3). 특히, 앞서 언급하였듯이 본 연구에서 제시한 존재인구의 경우 통근·통학 인구 뿐 아니라 모든 활동에 대한 인구를 반영할 수 있다는 측면에서 본 연구에서 구축한 존재인구 지표의 활용성은 매우 높다고 할 수 있다.

Table 7. Comparison between the results of this study and the Statistics Korea

District	This study		Statistics Korea	
	Population (1,000 persons)	Index	Population (1,000 persons)	Index
Gangnam	1,370	265	953	184
Gangdong	450	98	379	83
Gangbuk	274	86	265	83
Gangseo	381	71	468	87
Gwanak	459	90	423	83
Gwangjin	430	121	336	94
Guro	475	119	417	105
Geumcheon	334	146	271	118
Nowon	504	87	502	87
Dobong	258	75	274	80
Dongdaemun	439	131	379	113
Dongjak	371	96	360	93
Mapo	512	142	407	113
Seodaemun	329	108	344	113
Seocho	848	220	548	142
Seongdong	405	141	297	103
Seongbuk	435	97	427	95
Songpa	765	121	602	95
Yangcheon	187	41	384	83
Yeongdeungpo	684	183	514	138
Yongsan	344	159	270	125
Eunpyeong	297	67	359	81
Jongno	479	320	358	239
Jung	614	527	406	348
Jungnang	357	90	317	80
Seoul	12,001	126	10,370	109

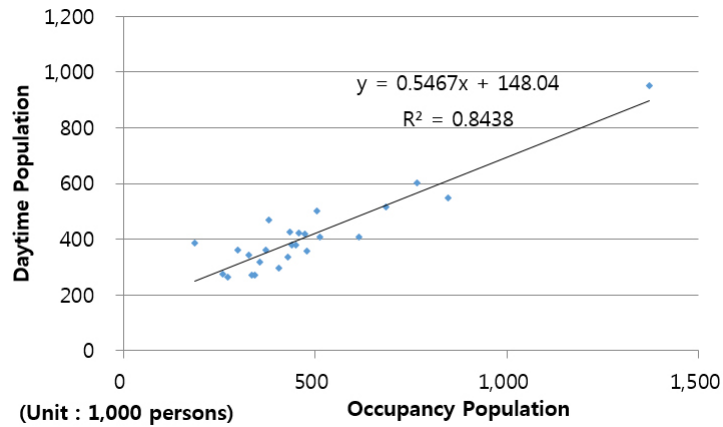


Figure 2. Correlation of the populations

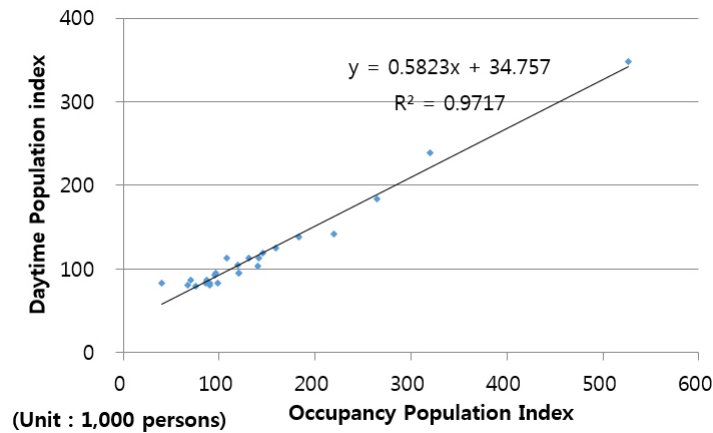


Figure 3. Correlation of the population indices

결론

본 연구에서는 특정 시간대에 특정 지역에 거주하거나 활동하는 인구의 개념으로서 존재인구라는 용어를 정의하고, 통신 자료를 활용하여 존재인구 및 지수를 산정하기 위한 방안을 제시하였다. 이를 위하여 통신확률 개념을 통해 통신 자료의 전수화 방안을 제시하였으며, 본 연구에서 산정된 시간대별 존재인구를 기반으로 통신량이 많은 12개 시간대의 평균을 이용하여 일간 존재인구 및 상주인구 대비 존재인구의 비율을 의미하는 존재인구지수를 산출하였다.

존재인구 산정을 위하여 SKT에서 제공하는 통신 자료를 활용하였으며, 서울시를 대상으로 구별 시간대별 존재인구 및 존재인구지수를 산정한 후, 통계청에서 제공하고 있는 주간인구 및 주간인구지수와 비교분석을 위해 09시에서 21시 시간대의 평균값을 사용하였다. 그 결과 존재인구 및 주간인구의 상관관계는 0.9186, 존재인구지수 및 주간인구지수의 상관관계는 0.9857로 강한 양의 상관관계를 가지는 것으로 나타났다. 특히, 본 연구에서 제시한 존재인구 및 존재인구지수가 통계청에서 제시하고 있는 주간인구 및 주간인구지수에 비해 절대값이 약 15.7% 가량 높은 것으로 나타났으며, 이는 통계청 주간인구의 경우 통근통학 인구만을 반영하고 있기 때문인 것으로 사료된다. 이처럼 존재인구 산정에 통신 자료를 활용할 경우 보다 비용 효율적이고 정확도 높은 결과를 도출할 수 있을 것으로 판단되며, 본 연구에서 제시된 요일별 패턴 분석 등을 통해 보다 합리적인 정책 수립과 행정력 배분이 가능할 것으로

사료된다.

하지만 본 연구에서 제시한 존재인구 및 존재인구지수에 대한 신뢰도를 높이고 실제적으로 적용하기 위해서는 추가적인 연구가 수행되어야 한다. 우선, 통신자료의 전수화를 위해 활용된 통신확률의 보다 합리적인 추정 방안에 대한 연구가 필요하다. 앞서도 언급되었지만, 통신확률은 지역별 인구 특성에 따라 차이가 있을 것으로 판단되지만, 본 연구에서는 자료 구득 상의 어려움으로 인해 이러한 특성을 반영하지 못하고 몇 가지 가정을 통해 통신확률을 산정하였다. 따라서 이동전화 통신에 대한 추가적인 자료 구득이나 보다 합리적인 가정을 통해 보다 신뢰성 있는 통신확률을 산정할 필요가 있다.

다음으로 본 연구에서 제시된 존재인구 및 존재인구지수에 대해 적용성 및 활용성을 높이기 위하여 거주인구와 활동인구를 명확히 구분할 수 있는 자료 및 방법론에 대한 검토가 필요하며, 기지국 위치에 따른 자료 상의 오류 등에 대한 추가적인 검토를 통해 추정의 정확도를 높이기 위한 노력이 필요하다. 이러한 연구를 통해 본 연구 결과는 교통 시설물의 공급 및 운영에 대한 시사점을 도출하고, 수요자 중심의 정책 수립 등에 유용하게 활용될 수 있을 것이다.

ACKNOWLEDGEMENT

This research was supported by a grant from R&D Program of the Korea Railroad Research Institute, Republic of Korea.

REFERENCES

- Byun M., Seo U-. S. (2011), How to Measure Daytime Population in Urban Streets?: Case of Seoul Pedestrian Flow Survey, *Survey Research*, 12(2), 27-50.
- Eun K-. S. (2001), Social Stratification of the Great Seoul Area: A Comparative Study Using Two Types of Population, *Korea Journal of Population Studies*, 24(1), 41-65.
- Foley D. L. (1954), Urban Daytime Population: A Field for Demographic-Ecological Analysis, *Social Forces*, 32(4), 323-330.
- Kim J. H., Ko Y. S., Kim J. K. (2015b), Application and Implication of Mobile Big Data in Land and Transportation Fields, *KRIHS Policy Brief*, 499, 1-6.
- Kim K. T., Lee I. M., Kwak H-. C., Min J. H. (2015a), Application Study of Telecommunication Record Data in Floating Population Estimation, *Seoul Studies*, 16(3), 181-191.