



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

지리학석사 학위논문

이동통신 빅데이터를 통해 살펴본
도시의 시공간 점유 패턴에 관한 연구

2018년 8월

서울대학교 대학원

지리학과

김 완 희

이동통신 빅데이터를 통해 살펴본
도시의 시공간 점유 패턴에 관한
연구

지도교수 이 건 학

이 논문을 지리학석사 학위논문으로 제출함
2018년 6월

서울대학교 대학원
지리학과
김 완 희

김완희의 석사 학위논문을 인준함
2018년 8월

위 원 장 손 정 렬 (인)

부위원장 구 양 미 (인)

위 원 이 건 학 (인)

국문초록

본 연구는 시공간적으로 달라지는 도시 유동인구에 따라 유의미한 ‘시공간 점유패턴’이 형성되고 있음을 통계적인 기법을 통해 파악하였다. 기존의 연구는 자료의 문제로 인해 유동인구의 시공간 분포를 체계적으로 탐지하지 못하였으며, 그 결과 역동적으로 변화하는 양상을 정확하게 포착하기 어려웠다. 이에 따라 본 연구에서는 정밀성, 상시성의 문제로부터 자유로운 이동통신 빅데이터를 연구자료로 활용하였고, 불규칙적인 클러스터를 파악하는 데에 효과적인 시공간군집방법인 Spatio-temporal AMOEBA기법을 방법론으로 적용하여 유동인구의 시공간군집 분석을 진행하였다.

분석 결과 및 의의는 다음과 같다. 첫째, 시공간군집기법의 통계적 방법론을 통해 기존의 연구에서는 볼 수 없었던 유동인구의 시공간적 분포를 체계적으로 포착할 수 있었다. 이를 통해 유동인구의 군집을 탐지하였으며, 군집이 시계열적으로 변화하는 모습을 통해 유의미한 점유 패턴이 나타나는 것을 파악하였다. 이러한 시공간적 점유 패턴은 해당 지역의 공간적 특성을 반영하는 것으로 나타났다. 더불어 본 연구에서 활용한 이동통신 빅데이터의 경우 상시적인 특성을 갖고 있으므로, 타 지역에서도 활용 가능할 것이다.

둘째, 성별 및 연령에 따라 유동인구의 분포 및 도시 점유 패턴이 서로 상이하게 분포하고 있음을 확인하였다. 즉 특정 계층에 의한 도시공간의 점유 및 소비에 대한 담론의 실증적인 분석을 제시하였다. 이는 규범적이고 추상적인 부분에서 논의만 이루어져왔던 기존의 연구 분야에 양적인 측면에서 기여를 할 수 있을 것이다.

셋째, 일부 군집은 특정 시간대에서만 분포하고 있음을 확인하였

다. 특정 계층에 의해 일시적으로 공간이 장악된다는 ‘한시적 경관’ 혹은 ‘일시적 도시공간의 점유’ 등의 개념을 양적인 연구를 통해 실증하였다.

하지만 연산의 간소화를 위해 시간단위를 임의적으로 합역하였기 때문에 MTUP를 고려하지 못하였다. 이러한 한계점을 개선하면 보다 시간적 정밀성을 높인 추가적인 연구를 진행할 수 있을 것으로 판단된다.

주요어 : 유동인구, 시공간 점유, 시공간 군집, 이동통신 빅데이터, Spatio-temporal AMOEBA

학 번 : 2016-20205

목 차

제 1 장 서론	1
제 1 절 연구 배경 및 목적	1
제 2 절 연구 자료 및 분석 도구	3
제 3 절 연구 구성	6
제 2 장 선행 연구	8
제 1 절 유동인구를 활용한 도시 내 시공간적 역동성 파악	8
제 2 절 공간빅데이터 분석 방법: 시공간군집분석	13
1. 공간군집분석	13
2. 시공간군집 분석	17
제 3 장 연구 방법	28
제 1 절 연구 설정	28
1. 공간적 사전 설정	28
2. 시간적 사전 설정	29
제 2 절 시공간군집분석: Spatio-temporal AMOEBA	31
1. AMOEBA 기법	32
2. Spatio-temporal AMOEBA 기법	35
제 4 장 유동인구에 따른 도시의 시공간 점유 패턴 분석 결과	39
제 1 절 서울시 유동인구의 도시 점유 패턴	39
1. 서울시 유동인구의 시공간군집 탐색 결과	39
2. 군집별 탐색 결과	42
제 2 절 성별에 따른 유동인구의 도시 점유 패턴	50
1. 서울시 남성 및 여성 유동인구의 시공간군집 탐색 결과	50
2. 군집별 탐색 결과	51
제 3 절 연령에 따른 유동인구의 도시 점유 패턴	63

1. 서울시 연령별 유동인구의 시공간군집 탐색 결과	63
2. 군집별 탐색 결과	66
제 5 장 결론	79
참 고 문 헌	82
영 문 초 록	88

표 목 차

<표 1-1> 서울시 인구 변화 추이	2
<표 1-2> KT 이동통신 빅데이터의 구성	5
<표 2-1> 공간군집분석 분류	15
<표 2-2> 시공간 군집기법의 분류	19
<표 3-1> 시간대별 유동인구	30
<표 4-1> 유동인구의 서울시 점유 공간	41
<표 4-2> 서울시 유동인구 군집의 시간대별 변화 양상 ..	48
<표 4-3> 성별 점유 공간	51
<표 4-4> 남성 및 여성의 점유 공간 1(남 여 공통)	53
<표 4-5> 남성 및 여성의 점유 공간 2(남성 혹은 여성) ·	54
<표 4-6> 서울시 시공간 점유 패턴의 시간대별 변화 양상 (남성)	58
<표 4-7> 서울시 시공간 점유 패턴의 시간대별 변화 양상 (여성)	61
<표 4-8> 연령별 점유 공간	65
<표 4-9> 서울시 시공간 점유 패턴의 시간대별 변화 양상 (20,30대)	71
<표 4-10> 서울시 시공간 점유 패턴의 시간대별 변화 양상 (40,50대)	75
<표 4-11> 서울시 시공간 점유 패턴의 시간대별 변화 양상 (60,70대)	77

그 립 목 차

<그림 1-1> 유동인구의 조사지점	4
<그림 1-2> 연구 구성	7
<그림 2-1> 시공간군집 개념	18
<그림 2-2> 시공간 군집분석 분류(소급적 분석방법, 전망적 분석방법)	20
<그림 3-1> 연구지역 및 토지이용도 지도	30
<그림 3-2> 시간축 설정	32
<그림 3-3> AMOEBA 알고리즘(Aldstadt and Geits, 2006)	34
<그림 3-4> Spatio-temporal AMOEBA 모식도	39
<그림 4-1> 도시의 시공간 점유 지역 탐색 결과(시공간 군집)	41
<그림 4-2> 점유 지역의 유동인구 변화 (전체군집 대비 유동인구 수)	43
<그림 4-3> 유동인구에 기반한 서울시의 시공간 점유 패턴 1	49
<그림 4-4> 유동인구에 기반한 서울시의 시공간 점유 패턴 2	50
<그림 4-5> 점유 지역의 남성 및 여성의 인구변화 (전체 군집 대비 유동인구 수)	52
<그림 4-6> 남성 유동인구에 기반한 서울시의 시공간 점유 패턴 1	59
<그림 4-7> 남성 유동인구에 기반한 서울시의 시공간 점유 패턴 2	60
<그림 4-8> 여성 유동인구에 기반한 서울시의 시공간 점유 패턴 1	62
<그림 4-9> 여성 유동인구에 기반한 서울시의 시공간 점유 패턴 2	63
<그림 4-10> 점유 지역의 연령별 인구변화 (전체군집 대비 유동인구 수)	66

<그림 4-11> 20,30대 유동인구에 기반한 서울시의 시공간 점유 패턴 1	72
<그림 4-12> 20,30대 유동인구에 기반한 서울시의 시공간 점유 패턴 2	73
<그림 4-13> 40,50대 유동인구에 기반한 서울시의 시공간 점유 패턴1	75
<그림 4-14> 40,50대 유동인구에 기반한 서울시의 시공간 점유 패턴 2	76
<그림 4-15> 60,70대 유동인구에 기반한 서울시의 시공간 점유 패턴 1	78
<그림 4-16> 60,70대 유동인구에 기반한 서울시의 시공간 점유 패턴 2	79

제 1 장 서론

제 1 절 연구 배경 및 목적

공간상에서 발생하는 현상들은 시간의 흐름에 따라 그 경향이 변화한다. 시공간은 인간이 현상을 이해하고 예측함에 있어 필수적이고 본질적인 요소이므로 (Fotheringham et al., 2015a) 시간과 공간은 동시에 고려되어야할 차원에 있다. 이러한 맥락에서 ‘도시의 시공간적 역동성’은 다양한 학문적 관점에서 다루어졌은 흥미로운 분야이다. 도시의 시공간적 역동성에 대한 연구는 경제지리학 분야, 정치지리학 분야, 문화지리학 분야 사례, 도시지리학 등의 지리학 분야 전반에서 진행되어왔다(정수열·임석희, 2012; 박선영·김영호, 2012; 이경주, 2013; 김오석, 2014; 신정엽, 2015; 이견학·김감영, 2015).

전술한 사례들에서 도시의 시공간적 역동성을 관찰하는 관점은 시공간적 규모에 따라 거시적·미시적 관점으로 구분할 수 있다. 거시적인 관점에서는 시계열 변화에 따라 도시가 번영 또는 쇠퇴를 겪으면서 도시 전체가 공간적으로 확장 또는 축소되는 형태와 도시의 다양한 특성들과의 연관성을 주로 다루며(Clark, 1965; 이경주, 2013; 김오석, 2014), 미시적으로는 도시 내 요소들의 시공간적 이합집산 등을 다룬다(정수열·임석희, 2012; 박선영·김영호, 2012; 신정엽, 2015; 이견학·김감영, 2015).

한편 도시 내 요소들의 시공간적인 이합집산에 대한 논의는 주로 산업 분야에서 다루어졌으며, 해당 요소들의 시공간적인 분포 및 그 변화를 다루는 연구 또한 산업분야를 중심으로 발전되어왔다. 하지만 비단 산업 분야만 아니라, 다양한 인문·사회현상에서 군집의 유무와 시공간적인 변화는 중요한 의미를 지니며, 이를 포착하는 것에 대한 논의가 여러 분야에서 확산되고 있다(Simmons and Yeates, 1998; Tsai, 2005; 이광국·양위주, 2007; 김완희 외, 2018).

이러한 논의는 ‘인구’의 측면에서도 간과될 수 없다. 인구 역시 공간상에서 균등하게 나타나지 않고, 특정시간, 특정 공간에 집중하는 경향이 있다. 대도시의 성장에 따른 공간의 확장과 변화는 도시 공간의 다양화,

기능적 분절화를 가져왔다. 특히 도시의 역동성이 증가하면서 도시가 가지는 기능적 특성이 다양해졌고, 이것은 도시 내 공간구조가 더욱 다양화하게 한다(이건학 외, 2013). 이는 곧 인구관련 지표에 반영이 되고 있다. <표1-1>에서처럼 공간적인 구조가 다핵화되고 다양화되면서 상주인구와 주간인구간의 차이는 지속적으로 벌어지고 있다.

<표 1-1> 서울시 인구 변화 추이

단위: 1000명

연도	상주인구	유입인구	유출인구	주간인구	주간인구지수
2000	9688	1087	586	10189	105.0
2005	9632	1174	612	10194	106.0
2010	9550	1457	638	10370	108.6
2015	9516	1500	733	10284	108.1

이러한 양상은 ‘유동인구’에서도 나타난다. 서울시와 같이 서비스업이 발달한 대도시의 경우, 통근·통학 이라는 정기적인 방문 외에도 기타 목적으로 도시를 일시적으로 방문하는 인구가 증가하고 있으며, 유동인구 자료는 이러한 특징을 포괄하는 자료이다. 유동인구는 해당 지역에 상주라는 목적으로 정기적으로 머무는 것이 아니며, 일시적으로 해당 지역에 일시적으로 머무는 것이다. 이는 곧 해당 지역을 일시적·한시적으로 도시공간을 점유한다는 ‘일시적 도시공간 점유’와 맥락을 같이한다(정혜미, 2013). 해당 연구에 의하면 단순히 일회적인 공간점유 뿐만이 아니라 장기적인 스케일에서 특정 시간/시점마다 반복적으로 나타나는 것 역시 ‘일시적 도시공간 점유’에 해당한다. 또한 이서진(2007) 역시 이와 유사한 개념으로 ‘한시적 가시경관’이라는 용어를 언급하였으며, 특정 계층이 한시적으로 특정한 시공간을 장악함으로써 생기는 경관이라고 정의하였다.

즉 이처럼 시공간적으로 달라지는 유동인구에 의해 도시 내 역동성, 즉 유의미한 ‘시공간 점유 패턴’이 형성되고 있다. 이처럼 시공간적으로 달라지는 유동인구에 따라 유의미한 ‘시공간 점유 패턴’이 나타나고 있으며, 이를 파악하는 연구는 여러 지리학분야에서 흥미를 끄는 주제이다.

이러한 측면에서 유동인구의 시공간적 분포를 계량적으로 파악하는 것 역시 중요하다. 하지만 기존의 연구는 조사원이 직접 각 거리에서 보행하는 인구를 카운팅하는 자료의 특성으로 인해 그 상시성이 부족하였고, 또한 서울시와 같은 특정 도시에서만 조사가 시행이 되었기 때문에 타 도시연구에 적용이 어려웠다. 그 결과 공간통계적인 방법이 진행되기 보다는 단순히 행정동 및 시군구 등의 행정경계별로 수치를 취합해 그 차이를 살펴보는 연구가 대부분이었다(변미리·서우석, 2011; 김기현 외, 2014; 이경민·정창무, 2014; 이향숙 외, 2014; 정윤영·문태현, 2014; 김미경 외, 2016).

한편 빅데이터 시대라고 불리는 오늘날, 인공위성, 항공기, GPS 수신기 등 다양한 센서를 통해 수집된 공간데이터는 시간이 지날수록 누적되었고, 보다 미시적이고 자세한 시공간 정보를 담게 되었다. 이는 곧 공간데이터의 정략적인 확장을 의미하며, 공간빅데이터로 정의될 수 있다(서양모·김원균, 2015). 이러한 흐름 속에서 유동인구 역시 보다 미시적이고 자세한 시공간정보를 갖게 되었다. 이동통신 빅데이터 기반의 유동인구 자료가 대표적인 공간빅데이터에 해당한다. 이는 기존의 유동인구 자료에 비해 보다 정확하고 구체적인 시공간 정보를 포함한다.

이에 따라 본 연구에서는 공간빅데이터에 해당하는 이동통신 빅데이터와 시공간클러스터링의 통계적인 방법론을 활용하여, 유동인구의 시공간적 역동성을 보다 체계적이고 미시적으로 파악하고자 한다. 이를 통해 도시 내 시공간 점유패턴을 도출할 수 있다.

제 2절 연구 자료 및 분석 도구

2016년 기준으로 국내추계인구는 약 5100만명이며, 이동통신 서비스 가입자는 6100만명 수준으로, 1인당 이동전화 가입자는 1.18로 인구를 상회하고 있을 정도로 모바일폰의 보급률은 매우 높다. 따라서 통신자료는 국민 대다수가 이용하는 전수에 가까운 자료라고 할 수 있다(김경태 외,

2015). 빅데이터시대에 이르러 모바일폰데이터와 정보기술의 융합으로 인해 정적인 국토에서 동적인 정보를 생산가능하게 되었다. 개인정보 등의 문제로 인해 이동통신서비스 가입자에 대한 직접적인 위치정보는 파악이 불가능하지만, 통신사들에서는 기지국 단위로 통신자료를 집계하고 있으며, 이를 활용하여 다양한 정보를 생산하고 있다. 이와 같이 이동통신데이터 기반의 유동인구자료는 가장 대표적인 공간빅데이터 자료에 해당한다.

이동통신데이터 기반의 유동인구는 이동통신 서비스 이용 현황을 기반으로 시간대별 유동인구(서비스인구 혹은 활동인구)를 집계한 50mX50m 간격의 점형 자료이다. 통신회사에서 제공하는 공간단위별로 특정 셀 안에 있는 사람들의 모바일폰 사용(데이터 송수신, 전화, 문자)를 시간단위로 파악하여 유동인구를 측정하였다<그림 1-1>. 본 연구에서는 이동통신기반의 유동인구 자료를 ‘이동통신 빅데이터’라고 칭한다.

해당 자료는 특정 시점 특정 장소를 보행하는 통근자, 통학자의 정기적 방문자 이외에도 관광객 등의 단기적 방문자를 포괄한다. 또한 <그림 1-1 >에서 볼 수 있듯이, 기존 유동인구 자료로 주로 활용되었던 서울시유동인구조사자료에 비해 보다 데이터의 정밀성이 높다는 장점이 있다. 이러한 점에서 이동통신 빅데이터를 활용한 분석은 기존의 서울시유동인구자료 기반의 분석에 비해 보다 정확성이 높을 것이라 기대된다.



<그림 1-1> 유동인구의 조사지점
(좌: 이동통신 빅데이터 유동인구, 우: 서울시 유동인구조사자료)

이에 따라 본 연구에서는 KT에서 제공하는 이동통신 빅데이터를 활용하였다. 자료는 서울특별시에 대한 80만개의 셀 자료이며, 성별, 시간대별 인구를 기록한 자료이다. 셀 자료는 그리드의 좌표값이 포함된 csv형태로 제공된다. 서울특별시 빅데이터캠퍼스에서 구독 가능하다. 해당 자료는 1일을 24시간으로 구분하여 각 시간대별로 해당 셀에 존재하였던 유동인구의 수를 나타낸다. 통신데이터를 기반으로 작성이 되며, 동일 시간대에서는 여러 차례 통신한 기록이 있더라도 1번으로 계산하고, 시간대가 달라질 경우에는 중복하여 산정한다. 예를 들어 3-4시간대에 총 5번의 통신기록이 있더라도 1번으로 계산하며, 한 대의 휴대전화에 대해서 하루 최대 24번까지 기록될 수 있다(김경태 외, 2015). 해당 데이터의 구성은 <표4-1> 같다.

<표 1-2> KT 이동통신 빅데이터의 구성

필드명	필드내용
셀 id	셀 별 고유 id 부여
x좌표	셀단위(50*50)의 셀 중심 X좌표 좌표계 : GRS80 TM 중부원점 (EPSG ; 5186)
y좌표	셀단위(50*50)의 셀 중심 Y좌표 좌표계 : GRS80 TM 중부원점 (EPSG ; 5186)
성별	성별 코드 1(남), 2(여)
연령대	5세단위의 연령대코드 (예: 60세~64세 ⇒ 6064)
주중유동인구수	주중(월~금)의 유동인구수 일 평균
주말유동인구수	주말(토,일)의 유동인구수 일 평균
거주지코드	유동인구의 거주지(행정동) 코드
행정동코드	유동인구 좌표가 위치한 행정동 코드
기준년월	데이터 생성 기준년월 (월단위 데이터는 생성 기준일자)

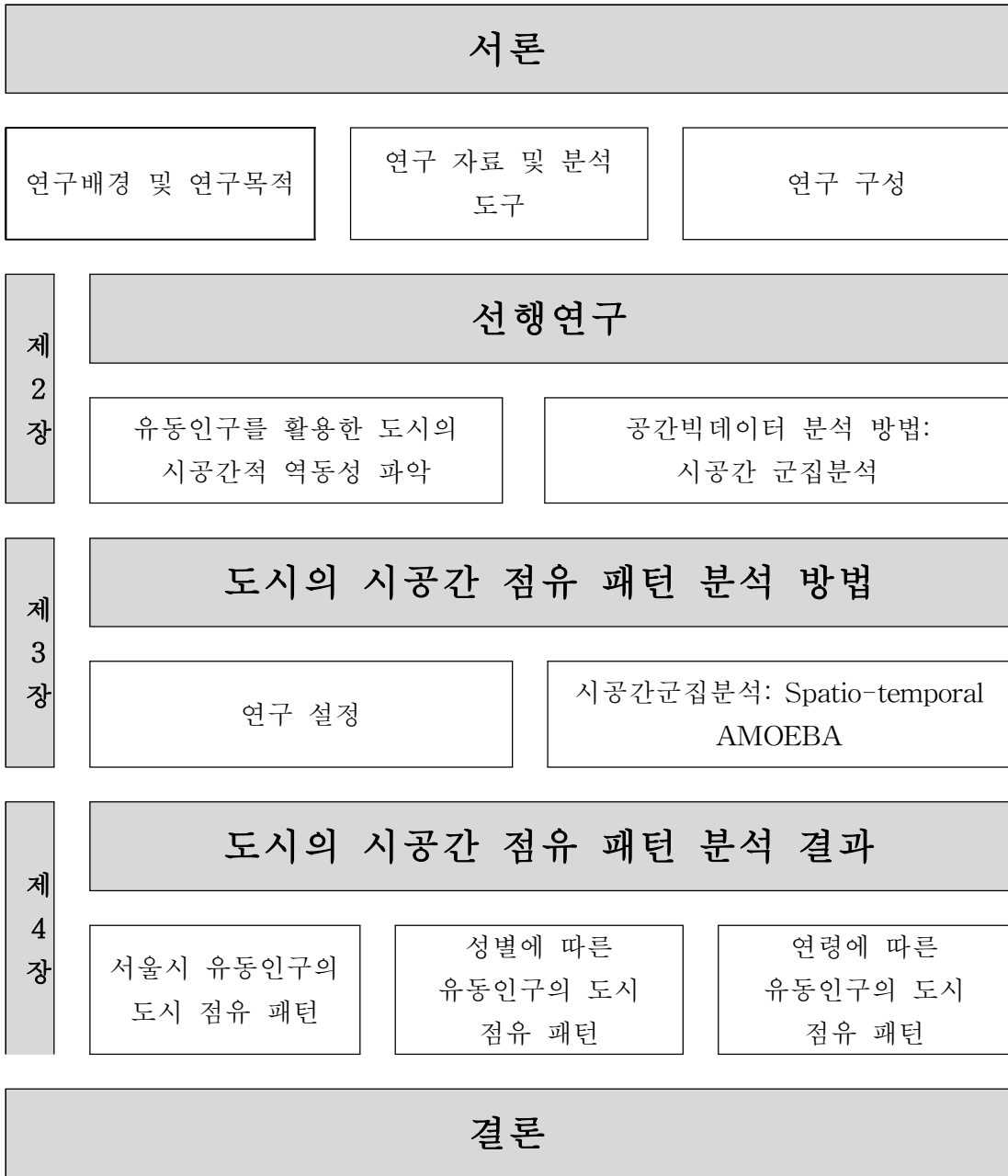
이러한 이동통신 빅데이터 기반의 유동인구 원자료를 총 4개의 시점에

서 단계구분도로 표현해보면 <그림 1-2>와 같이 나타난다. 시간대에 따라 유동인구가 특정한 공간에 분포되어 있으며, 분포가 활발하게 변화하는 것을 알 수 있다. 빅데이터처리기술인 MySQL과 통계프로그래밍언어인 R프로그래밍을 통해 2016년 6월부터 2017년 6월까지의 자료를 시간대별로 합역하였다.

연구 방법으로는 시공간군집분석 기법을 적용하였고, 이를 통해 서울시 유동인구의 군집 및 변화 양상을 체계적으로 살펴보았다. 시공간군집분석 기법으로는 공간군집기법인 AMOEBA 기법을 시간축으로 확장·적용한 Spatio-temporal AMOEBA를 이용하였으며, 이 분석에는 R프로그래밍을 이용하였다. Spatio-temporal AMOEBA에 대한 자세한 설명은 3장 2절에서 다루었다. 그리고 결과의 시각화에는 QGIS 2.18을 이용하였다. 유동인구의 시공간적군집을 성별, 연령 등의 인구적 속성에 따라 각각 분석을 시행하였다.

제 3절 연구 구성

본 연구의 구성은 다음의 <그림 1-3>과 같다. 우선 2장에서는 선행연구들을 통해 유동인구의 시공간적 분포를 포착한 연구들을 살펴봄으로써 기존 유동인구 관련 연구의 한계점을 제시한다. 또한 유동인구의 시공간적 군집을 포착하기 위해 활용하기 위하여 기존의 공간군집분석 및 시공간적 군집방법론들을 살펴보고 본 연구에 적합한 방법론을 선택한다. 3장에서는 연구 방법에 대해 설명하고자한다. 본 연구를 원활히 하기 위한 연구설정을 언급하고, 분석에 활용한 방법론인 Spatio-temporal AMOEBA 기법을 설명한다. 구체적으로는 분석의 공간단위 및 인구밀도의 과대추정을 방지하기 위한 일련의 설정 그리고 AMOEBA를 시간축으로 확장하여 적용하는 방법 등이다. 4장에서는 3장에서 설명한 연구방법을 실제 서울시 유동인구 자료에 적용하여 그 분포와 시공간적 변화를 파악한다. 마지막으로 5장에서는 분석 결과를 요약하고 정리하여 본 연구의 의의와 한계를 언급하고자 한다.



<그림 1-2> 연구 구성

제 2 장 선행연구

유동인구를 통해 도시 내 시공간적 역동성을 파악한 연구는 지리학 이외에도 도시공학 등 여러 영역에서 꾸준히 진행되어 왔다. 본 장에서는 유동인구를 활용하여 도시 내 시공간적 역동성을 살펴본 기존의 연구를 확인하며, 도시의 시공간 점유패턴을 파악하기 위한 방법론들을 살펴본다. 먼저 1절에서는 유동인구의 군집 및 그 변화 양상을 파악한 연구를 살펴본다. 이를 통해 유동인구의 분포 파악의 중요성을 확인하고, 기존 연구의 한계점을 언급하고자 한다. 2절에서는 유동인구분포의 시공간적 군집을 파악하고, 나아가 도시 내 시공간 점유패턴을 확인하기 위한 방법론적 선행연구를 살펴본다. 이는 크게 공간군집분석과 시공간군집분석으로 나눌 수 있다.

제 1 절 유동인구를 활용한 도시 내 시공간적 역동성 파악

도시계획 등 도시 전반적인 분야에서 인구관련 자료는 기초적이고 필수적인 자료에 해당하며, 대체적으로 ‘주민등록인구’가 활용된다. 하지만 주민등록인구는 상주인구에 해당하는 것으로, 낮시간에 실제 해당지역에서 활동하거나 이동하는 비상주인구와는 차이가 있으며, 도시 활동에 대한 분석에는 활용하기 어렵다. 이에 통계청에서는 인구주택총조사의 일환으로 12세 이상 통근 및 통학을 대상으로 유입인구와 유출인구를 조사하고 있으며, 주간인구 및 주간인구지수 자료를 시군구 단위로 제공한다. 주간인구는 해당지역의 상주인구에 타지역으로부터 유입된 통근·통학인구를 더하고 타지역으로 유출된 통근·통학인구를 뺀 것이다. 하지만 이 주간인구는 통근·통학 외 기타목적으로 해당 지역을 방문하는 인구는 포

함되지 않는다.

이러한 통근자와 통학자만을 기준으로 하는 주간인구를 보완하기 위해 ‘유동인구’의 개념이 제시되었다(변미리·서우석, 2011). 유동인구는 특정 지점을 기준으로 일정시간동안에 유입·유출 및 이동한 인구수로서 건물 유출입 인구와 가로보행 등을 포함하는 인구이다(서울시, 2010). 그러므로 유동인구자료는 도시의 동적활동을 파악하기에 효과적인 지표로서 이용 가능하다(장운영·문태현, 2014). 특히 서울과 같은 대도시권에서는 통근, 통학 등의 정기적인 체류 이외에도 단기적 방문이 빈번하기 때문에 유동인구의 파악은 그 중요성이 더욱 높아지고 있다. 이러한 흐름 속에서 서울시는 2010년부터 조사원이 직접 거리에서 보행하는 인구를 카운팅하는 ‘서울시 유동인구자료’를 제공하고 있다.

이러한 유동인구자료를 활용한 연구들이 다수 진행되었으나, 자료 및 방법론적인 한계로 인해 시공간적 범위 및 정확도가 미시적이거나 구체적인 스케일에서 진행되지는 못하였다. 해당 연구들은 크게 두가지의 연구 목적을 가지며 진행되어왔다. 먼저, 상권분석 등 상업적인 관점에서 인구적 속성으로 유동인구의 분포를 살펴본 연구가 있다(김기현 외, 2014; 이경민·정창무, 2014; 김미경 외, 2016; 김완희 외, 2018).

김기현 외(2014)는 상업가로를 대상으로 유동인구의 공간적 분포와 가로의 물리적 특성과의 연관성을 분석하였다. 서울시 유동인구자료를 활용하였으며, 유동인구의 밀도가 가로망의 특성에 따라 토지이용 및 개방성 등의 다양한 형태와 연관성을 보임을 밝히었다.

이경민·정창무(2014)는 유동인구가 업종별 점포입지에 미치는 영향에 대하여 평가하였다. 블록별 업종 점포 밀집도를 종속변수로, 유동인구의 시간대별 통행량 변수와 기타 통제변수(공간적인 속성)를 설명변수로 설정하여 다중회귀 기반의 연구를 수행하였다. 그 결과 시간대별 유동인구의 인구적 속성이 점포입지에 영향을 미치고 있음을 확인하였다.

김미경 외(2016)은 서울시의 대피소의 수용능력을 분석하기 위해 유동인구자료를 활용하였다. 이동통신데이터 기반의 유동인구 및 상주인구를 연구자료로 활용하였으며, 서울시 대피소의 실질적인 접근성 및 수용능

력을 분석하였다. 또한 연구지역 내 대피소의 불균등한 분포와 그로 인한 취약 지역을 제시하였다.

김완희 외(2018)는 상업집적지의 탐지를 위해 유동인구의 분포를 살펴 보았다. 이동통신 빅데이터 기반의 유동인구와 공시지가 그리고 요식업 현황 자료 등을 통해 상업 집적지를 탐지하였고, 집적지의 시공간적 변화양상을 살펴보았다. AMOEBA 등의 공간군집분석을 통해 진행하였다.

또한 유동인구의 시공간적 분포를 살펴보고, 그 변화 패턴을 분포패턴을 토지이용 등을 통해 해석하고자한 연구들이 진행되었다(변미리·서우석, 2011; 이원도·조창현, 2011; 윤정미·최돈정, 2014; 이향숙 외, 2014; 정윤영·문태헌, 2014).

변미리·서우석(2011)는 유동인구자료를 활용하여 도시의 주간활동인구를 측정하는 방법을 제안하였다. 통근·통학을 근거로 인구이동을 측정한 기존의 주간인구의 한계를 해결하기 위해 가로를 중심으로 사람들의 이동을 측정한 보행량으로서의 유동인구 개념을 도입하였고, 이는 도시관리의 수요예측 면에서 중요한 역할을 할 수 있음을 언급하였다. 이를 바탕으로 시군구별 유동인구의 공간적 분포 차이와 시간적 분포 차이를 살펴보았다.

이원도·조창현(2011)는 유동인구의 시공간적 분포의 양상을 살펴보았다. 수도권 가구통행조사를 연구자료로 활용하였으며, 행정동별 유동인구에 대하여 시간대에 따른 편차를 살펴보았다. 이후 교통카드자료와 사용기록을 비교하여 그 집중지역을 탐색하였고, 해당 지역의 지역특성을 설명하였다.

윤정미·최돈정(2014) 역시 유동인구 분포의 공간패턴에 영향을 준 요인을 탐색하였다. 유동인구분포와 토지이용특성과의 관계를 파악하기 위해 지리가중회귀모형(GWR)을 이용하였고, 각 지역에 따라 유동인구분포에 영향을 준 요인이 서로 상이함을 분석하였다.

한편 이향숙 외(2014)는 유동인구자료를 통해 보행특성을 분석하였다. 서울을 5개의 권역으로 나누어, 권역별로 인구의 유동인구, 시간대별 유동인구, 조사지점 속성에 따른 유동인구의 차이를 분석하고, 다중선형회

귀 분석을 통해 유동인구에 영향을 미치는 요인을 규명하였다. 그 결과 중앙선, 버스정류장 및 경사로 등 도로적 속성에 따라 유동인구의 수가 크게 차이가 나고 있으며, 이외에도 다양한 토지이용지표, 사회적 지표 등도 유동인구에 영향력을 발휘함을 밝혔다.

반면 정운영·문태현(2014) 서울시의 유동인구자료를 이용하여 시간대별, 유동인구 밀집지역을 분석하였다. 행정동경계와 유동인구자료를 중첩하여 행정동별 유동인구를 시간대별로 추출하여 분석에 이용하였다. 이후 통근통학 O-D데이터를 이용한 중심성지수를 함께 사용하여 2030서울플랜에서 제시한 도심 및 광역중심지와 비교하였다. 그 결과 3도심에 해당하는 지역은 유동인구 및 중심성지수가 높게 나왔으나, 7광역중심지의 경우 일부지역은 일치하지 않는 것으로 나타났다. 해당 연구에서 이용한 유동인구데이터는 집계원이 직접 도로에서 실제적으로 유동하는 수를 직접 카운팅한 데이터이며, 직접 카운팅한다는 집계 방식의 특성상 07~21시라는 시간적인 제약을 갖고 있었다.

한편 유동인구의 공간적 및 시공간적 분포를 살펴본 연구는 해외에서도 활발하게 이루어졌다(Bhaduri et al., 2007; Giradin et al., 2009; Thomas et al., 2014; Song Gao, 2015). 먼저 Bhaduri et al.(2007)는 미국을 공간적 대상으로 기존의 통계데이터로는 파악이 불가능하였던 주간 및 야간의 유동인구를 추정 및 비교하기 위한 모델을 개발하였다. 유동인구를 추정하기 위한 데이터로는 거주인구, 주야간 노동자, 관광객 등의 통계자료를 이용하였고, 방법론으로는 면적가중치(Area Weighted interpolation), 대시메트릭 기법(Dasymetric) 등 여러 공간보간법을 이용하였다. 해당 분석 결과를 주간인구와 비교를 하였으며, 두시간대의 인구가 공간적으로 상이하게 군집되어있음을 관찰 가능하였다.

Giradin et al.(2009)는 모바일폰의 통화 데이터의 통화 밀도를 통해 뉴욕 내 폭포를 방문하는 인구의 숫자 및 분포를 추정하였다. 각 지역에 대하여 방문객의 시간적 분포 차이를 살펴보았다.

Thomas et al.(2014)는 이동통신 빅데이터는 인간의 행동을 보다 잘 포착할 수 있는 데이터라고 언급을 하며, 유동인구의 공간적인 분포를

살펴보았다. 요일에 따른 도시 간 유동인구의 수를 비교하였고, 이를 스페인의 도시구조와 연결지어 해석하였다. 하지만 시간대별 데이터로의 활용은 이어지지 못하였다.

Song Gao(2015)는 유동인구를 통해 인간의 이동패턴과 도시의 역동성을 탐구하였다. O-D기반의 이동통신데이터를 이용하여 시공간적 분석을 진행하였으며, 거주지 및 출퇴근지 등 실제로 자주 방문되는 지역을 파악 가능하였다. 모바일 셀을 기준으로 시공간 커널밀도 분석 및 시공간 자기상관 분석 등을 통해 시공간 궤적을 파악하였다.

이처럼 유동인구의 시공간적 분포를 살펴본 연구가 다수 진행되었다. 하지만 국내연구의 경우 그 분석 대상이 서울시에 국한되어 있으며, 단순히 행정동 별 유동인구의 수치를 파악하는 것에 그쳤다. Getis-ord G_i^* 등의 공간통계기법을 통해 군집을 파악하는 연구가 일부 존재하였지만 이는 인구 자체의 군집을 바라보기 보다는 유동인구패턴과 토지이용과의 상관성을 파악하기 위한 분석이었으며, 군집분석에 사용된 사건의 개수 역시 서울시 전역을 포괄하지 못하였다. 또한 유동인구의 시공간적 군집과 그 변화 양상, 즉 ‘시공간적 점유패턴’을 체계적으로 파악하고자하는 연구는 미흡하였다.

이러한 한계는 유동인구에 이용하는 자료 자체에서 기인한다. 유동인구의 시공간적 분포를 파악한 기존의 연구에서는 대체적으로 인구주택총조사의 일환으로 이루어지는 가구별 이동조사 혹은 조사원이 직접 카운팅한 자료를 활용하였다. 이러한 기존의 자료는 그 구축에 있어서 많은 노동력과 비용이 들기 때문에 상시성과 정밀성이 부족하다는 단점을 갖고 있다. 하지만 최근 빅데이터 시대에 이르러 이동통신데이터의 가용성이 증대되었고, 이를 활용한 연구가 다수 시도되었다. 하지만 대부분 유동인구보다는 특정 시점, 특정 위치에 존재하는 모든 인구를 의미하는 ‘현주인구’의 패턴을 살펴하는데에 국한되었다.

한편 이동통신데이터 기반의 유동인구를 인 이동통신 빅데이터를 활용하여 현주인구를 추정하는 연구는 다음과 같다(김종학 외, 2014; 이건학·김감영, 2016; Deville *et al.*, 2014; Douglas *et al.*, 2015). 김종학 외

(2014)는 모바일폰 사용자의 실시간 공간분포를 공간정보와 연동하여, 요일별, 시간대별 수도권 현주인구의 동적변화를 파악하고자 하였다. 이를 위해 커널밀도분석, 히트맵분석, 무게중심점의 개념을 적용하였다.

이건학·김감영(2016)은 이동통신 빅데이터를 활용하여 대구광역시의 서비스인구의 시간대별 분포와, 성별, 연령별 인구의 공간적 분포 특성을 살펴보았으며, 이를 공간통계기법을 이용해 현주인구를 추정하는 데에 비교자료로 이용하였다. 속성별로 인구밀집지역의 분포 및 변화 양상을 파악할 수 있었다.

Deville *et al.*,(2014)은 포르투갈과 프랑스를 공간적 대상으로 요일별, 주중과 주말, 휴가철과 비휴가철의 유동인구의 변화양상을 비교 분석하기 위한 모델을 작성하였다. 데이터로는 모바일 폰 위치데이터를 이용하였고, 방법론으로는 공간보간법(Areal interpolation), 대시메트릭 기법(Dasymetric) 등을 이용하였다. 해당 연구에서 이용한 방법론들을 통한 인구밀도의 추정은 기존의 downscaling 방법만큼 정확성을 갖는다는 것을 밝히었다.

Douglas *et al.*,(2015)역시 이동통신데이터의 통화량을 바탕으로 인구 밀도를 추정하는 연구를 진행하였다. 해당 연구에서도 회귀식을 직접 작성하여 이용하였으며, Deville *et al.*의 연구와 유사하게 토지이용데이터와 결합하여 그 설명력을 높이고자 시도하였다. 이탈리아의 밀라노를 사례지역으로 삼았으며, 연구 결과를 통해 이동통신 빅데이터를 통해 실시간에 가까운 인구추정이 가능함을 밝혔다.

이를 종합하면, 유동인구의 시공간적 분포를 살펴본 연구가 다수 존재하였으나, 통계적 방법론을 통해 체계적으로 살펴본 연구는 미흡하였으며, 공간적 대상 역시 서울시라는 특정 도시에 한정되어 있었다. 이러한 한계점들은 연구에 활용한 자료의 한계에서 기인하였다. 이러한 자료의 한계성을 해결하기 위하여 이동통신 빅데이터가 활용될 수 있다. 이동통신 빅데이터를 활용한 연구 역시 활발히 진행되고 있으나, 유동인구의 측면에서 이루어진 연구는 미흡하였다. 이러한 기존의 한계점들에서 착안하여 본 연구는 이동통신 빅데이터 기반의 유동인구를 통계적인 분

석을 통해 그 군집을 도출하고자 한다.

제 2 절 공간빅데이터 분석 방법: 시공간군집분석

본 절에서는 유동인구의 시공간적 분포와 그 변화양상을 체계적으로 탐색 위한 방법론들을 살펴보고자 한다. 공간빅데이터에 해당하는 이동통신 빅데이터를 분석하기 위해 시공간군집분석을 활용하고자 한다. 이를 위해 먼저 공간군집기법의 개념과 예시를 살펴보고자 한다. 이후 공간군집기법을 시간축으로 응용한 시공간군집기법에 대한 방법론적 기존의 연구들을 살펴보고자 한다. 이는 본 연구에 적합한 시공간군집기법을 선택하기 위함이다.

1. 공간군집분석

공간군집분석 방법은 공간빅데이터를 분석하는 대표적인 방법에 해당하며, 이는 공간적자기상관성을 기반으로 한다. 많은 사건들은 공간상에서 특정한 위치를 중심으로 군집되는 성향을 갖고 있다. 이러한 성향은 일정한 형태의 공간적 패턴으로 나타난다. 공간적 패턴이라는 말은 공간적 이질성과 동질성 및 거리조락현상에 의거한다. 예를 들어, 어떤 지역의 유동인구 밀도는 다른 지역에 있는 유동인구의 밀도와 차이가 난다(공간적 이질성). 그러나 서로 인접한 지역들의 유동인구밀도는 그 수치가 비슷하다(공간적 동질성). 만일 어느 지역에서의 수치가 10,000명인데 바로 인접한 지역의 수치가 100명으로 급격하게 떨어지는 경우는 특별한 이유가 없는 한 흔치 않다. 또한 이러한 차이는 지역간 거리가 멀어질수록 커질 것이다(거리조락현상). 서울시 종로구의 유동인구밀도는 지방 농촌지역의 유동인구밀도보다 높을 것이다. 이는 해당 지역 토지의 물리적인 특성보다는 각 지역의 사회적 환경이 다르기 때문이다. 특히 도시 공간구조의 특성은 산업체·거주지 등의 환경적 차이를 가져오고, 이는

다시 해당 공간 위에 존재하는 인구의 차이를 야기한다. 이는 즉 ‘공간적 자기상관성(spatial autocorrelation)’을 의미한다. 공간적 자기상관성은 공간적으로 가까운 거리에 있는 것들은 비슷한 특성을 갖는다는 것을 의미하며, 이는 토블러의 ‘지리학 제1법칙’과도 통한다. 즉 이러한 공간적 자기상관성을 통해 유사한 집단을 하나의 군집으로 설정함으로써, 데이터의 양을 축소시키면서도 본래의 의미를 변화시키지 않고, 그 정보를 전달하고 시각화 할 수 있는 것이다. 이러한 점에서 공간군집분석은 공간 빅데이터를 효과적으로 분석하고 다룰 수 있는 방법론에 해당한다.

이처럼 공간데이터 및 공간빅데이터의 군집을 파악하는 것은 지리학을 비롯하여 범죄나 보건 등 다양한 연구 분야의 기본적인 관심사이다. 특히 어떠한 객체나 사건이 특정 공간상에 군집되어 분포하는지(클러스터링 분석), 만일 그러하다면 그 영역은 구체적으로 어디인지(클러스터 탐색 분석)를 살펴보는 것은 매우 중요한 관심 주제이다. 이러한 객체나 사건과 관련된 정보들은 특정 위치에서의 정보 유무로 기록되거나 행정 구역과 같은 영역단위로 합산된 카운트로 기록된다(조대현, 2013). 카운트데이터에 대한 공간군집분석은 1990년대 들어 급성장한 공간통계학 혹은 공간데이터분석과 밀접하게 관련되어있다. 이러한 공간현상의 군집 분석방법은 그 데이터의 유형 및 분석 목적과 관련하여 상이하게 발전되었다. 이는 Besag and Newell(1991)에 의해 처음 제시되었으며, Waller and Gotway(2004), Cromley and McLafferty(2012), 조대현(2013)등에 의해 정리되고 발전하였다.

공간 군집 분석 방법은 크게 ‘클러스터링 테스트’와 ‘클러스터 테스트’로 분류 가능하다. 클러스터링 테스트는 전역적 공간 패턴을 나타내는 것으로, 한 지역 혹은 그 이상의 지역에서 군집이 나타나는지를 판단하는 것이며, 클러스터 테스트는 국지적 공간패턴을 나타내는 것으로, 전체 지역에서 클러스터로 불릴 국지적 범위를 설정하는 것과 관련이 있다(김영호, 2012; 조대현. 2013).

클러스터링테스트의 대표적인 방법은 Moran’s I, Geary’s C, K-function, Score-test 등이 있으며, 클러스터 테스트의 방법론으로는

Geographical Analysis Machine, Spatial Scan statistic, Local Indicators of Spatial Autocorrelation 등이 있다.

<표 2-1> 공간군집분석 분류

전역적 방법(클러스터링 테스트)	국지적 방법(클러스터 테스트)
<ul style="list-style-type: none"> ▪ Moran's I (Moran, 1948) ▪ Geary's C (Geary, 1954) ▪ K-function (Ripley, 1976) ▪ Socre test (Lawson, 1993) 	<ul style="list-style-type: none"> ▪ Geographical Analysis Machine (Openshaw and Charlton, 1987) ▪ Local Indicators of Spatial Association (Anselin, 1995) ▪ Spatial Scan Statistics (Kulldorff and Negarwalla, 1995)

전역적방법에 해당하는 클러스터링 테스트는 모든 사건들의 전반적인 패턴을 나타내는 것으로, 군집의 여부 및 그 정도를 판단할 수 있는 방법이다. 대표적인 방법으로는 Moran's *I*(Moran, 1950), Geary's *C*(Geary, 1954), K-function(Ripley, 1976), Score-test(Lawson, 1993)등이 있다. Moran's *I* 및 Geary's *C*는 각 사건들이 서로 이웃한지 아닌지를 판단하는 공간가중행렬을 만들며 이를 바탕으로 군집의 여부 및 정도를 평가하는 방법이다. K-function은 연구지역 내에서 사건들에 대하여 사전에 반경을 설정한 뒤, 해당 반경 안에 위치하는 다른 사건들의 개수의 평균값을 구하고, 이를 무작위 패턴에서의 값과 비교를 진행한다. 이를 통해 군집 여부 및 정도를 파악한다. 한편 Score test(Lawson, 1993)은 미리 선정된 지역들을 기준으로, 해당 지역에서 군집이 형성되는지 아닌지를 판단하는 방법에 해당한다. 이는 클러스터링 테스트 중 집중테스트에 해당한다.

반면 국지적방법에 해당하는 클러스터 테스트는 어떠한 공간에서 군집이 발생하는지를 찾아내는 방법이다. 전역적 군집을 형성하는 데에 기여를 한 지역 및 시점이 구체적으로 어디인지를 탐색한다. Openshaw *et*

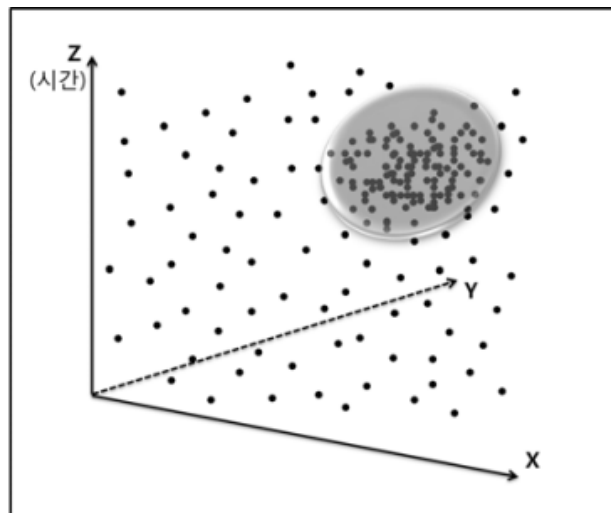
al.(1987)에서 비롯된 것으로, 어린이 백혈병 환자가 집중적으로 발생한 지역을 찾고 통계적 유의성을 검증하기 위해 GAM(Geographical Analysis Machine)이라는 분석방법을 제시하였다. 하지만 이 방법론은 분석과정의 특성으로 인해 도출된 군집들이 중첩되어 있는 경우, 해당 군집들의 통계적 유의성 기준을 동시에 적용하기 위한 다중가설검정문제에 취약하다(김대중·구형수, 2011). 두 가설을 동시에 검정하기 위해서는 서로 독립적인 자료를 활용해야하는데, 이 경우에는 공간적 자기상관성으로 인해 자료들의 독립성이 보장되지 않기 때문이다.

Spatial Scan Statistics(공간스캔통계, Kulldorff, 1997)은 이러한 한계점을 해결하기 위해 제시된 기법이다. 군집의 지역은 ‘우도비’로 정의되는 공간스캔통계량을 통해 탐색된다. 특정 사건의 지점으로부터 검색반경을 넓혀가며 공간스캔통계량을 계산하며, 모든 사건들에 대하여 이러한 분석을 시행한다. 계산 결과 중 그 값이 가장 큰 지역이 군집에 해당한다.

Local Indicators of Spatial Association(Anselin, 1995)는 국지적 자기상관 통계량을 의미하는 것으로, 줄여서 LISA라고 부른다. LISA에는 국지적 Moran's I_i (Anselin, 1995)와 Geary's C_i (Anselin, 1995), 그리고 Getis-ord's G_i 와 G_i^* (Getis and ord, 1992; Ord and Getis, 1995)가 있으며, 일정한 그리드 데이터 이외에도, 일정하지 않은 공간경계를 가진 역형 객체로 구성된 데이터에도 적용하기 용이하다는 장점이 있다. 이 방법론은 사건이 발생한 위치(혹은 지역)과 그 이웃사건의 위치(지역)과의 관계에 초점을 두고 있으며, 사건과 그 이웃의 값의 유사성이나 집중도를 측정한다. 이때 이웃내의 값들과 전반적인 평균 혹은 다른 이웃의 값과 비교를 함으로써 유사성과 집중도를 평가한다. 이러한 LISA는 가장 널리 이용되는 공간클러스터 탐색 기법에 해당한다.

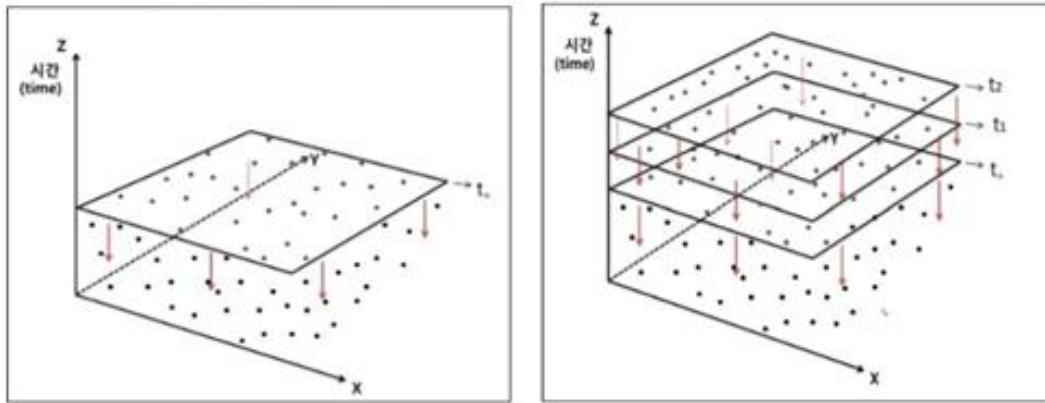
2. 시공간군집 분석

앞서 언급하였듯 공간군집분석은 특정 사건의 공간적 밀도가 높은 것을 탐색하는 것으로, 사건들 간의 거리가 상대적으로 가까운 것을 말한다(Bailey and Gatrell, 1995). 최근에는 공간군집분석에 ‘시간’이라는 변수를 함께 고려한 시공간군집분석이 많이 활용되고 있다. 이는 공간 군집에서의 공간적 자기상관성과 같은 맥락에서, 시간적인 측면을 고려한 시공간자기상관성을 도입한 것이다. 시공간적 자기상관성은 비슷한 공간 및 시점에 발생하는 사건들은 유사한 성격을 가진다는 개념이다. 즉, 시공간군집은 공간적 측면만을 강조하는 공간적인 분포 패턴 분석의 한계를 인식하고 공간상에서 나타나는 사건 간 인과관계를 시간으로 확대한 것으로(박선영·김영호, 2013), 토블러 지리학 제1법칙을 시간축으로 확대한 것으로 볼 수 있다(Arbia *et al.*, 2010). <그림 2-1>은 시공간군집을 그림으로 표현한 것이다. 공간은 X, Y좌표를 기반으로 하는 좌표 상에서 표현되며, 시간은 Z로 표현된다. 시공간군집은 사건들 사이에서 공간적으로나 시간적으로 거리가 모두 가까운 값들로 구성이 된다. 그 결과 시공간군집분석은 결과로서 특정 사건들이 시공간적으로 군집되는 경향성을 갖는지, 그리고 그러한 구체적인 군집이 어떠한 범위에서 형성되었는지를 보여준다.



<그림 2-1> 시공간군집 개념

이러한 시공간군집역시 공간클러스터탐색방법과 마찬가지로 크게 시공간클러스터링 테스트와 시공간클러스터 탐지테스트 두가지로 유형 가능하다. 그리고 이러한 시공간군집분석은 데이터의 시간적 범위와 연속성의 여부에 따라 소급적 분석방법(retrospective analysis)과 전망적 분석방법(prospective analysis)으로도 구분이 가능하다(Sossen and Bock, 2003). 소급적분석은 연구의 데이터는 고정되어 있고, 전체의 데이터세트에 대해 군집분석이 동시에 수행된다. <그림 2-2>의 좌측에서 제시하듯 시공간적으로 제한된 범위에 있는 데이터를 대상으로 단 한번의 시공간적 군집분석을 진행한다. 특정 지점과 시간에서 핫스팟을 발견한다(Sossen and Bock, 2003; 김영호, 2012). 반면 전망적분석의 경우 연속적인 시간데이터에 대해 데이터셋을 업데이트하며 시공간 군집분석을 진행하는 것으로. <그림 2-2>의 우측에서 보여주고 있으며, 데이터가 업데이트됨에 따라 t_0 , t_1 , t_2 의 개별시간에서 연속적으로 분석이 진행된다. 이때 연속성의 개념은 연구대상의 시공간적 속성에 따라, 년, 월, 주, 일 등의 단위로 설정이 된다(Sossen and Bock, 2003; 김영호, 2012). 이러한 전망적 분석은 시공간군집분석은 데이터들을 대상으로 해당 군집의 구체적인 시공간적 위치를 포착하고 그 변화 양상을 파악하는 데에도 이용 가능하다(김영호, 2012). 나아가 연속적으로 데이터를 업데이트한다면, 지속적인 분석을 통해 미래에 대한 예방과 예측도 가능하다. 또한 시공간 군집분석은 단순히 군집의 확인에서 그치는 것이 아니라 군집들의 시간대별 성장(Growth), 이동(Movement), 분할(Splits) 및 결합(Joins)을 관측 가능하다. 각 방법들 예시는 <표 2-2>로 정리하였다.



<그림 2-2> 시공간 군집분석 분류(소급적 분석방법, 전망적 분석방법)

<표 2-2> 시공간 군집기법의 분류

	전역적 방법(클러스터링 테스트)	국지적 방법(클러스터 테스트)
소급적 연구방법	<ul style="list-style-type: none"> ▪ Knox test ▪ K-Nearest neighbor test ▪ Space-time K-function 	
전망적 연구방법	<ul style="list-style-type: none"> ▪ CUSUM test 	<ul style="list-style-type: none"> ▪ Space-time Scan ▪ Permutation Scan ▪ Bootstrap-based surveillance model

‘시공간 클러스터링 테스트’는 분석하고자하는 현상(객체 및 사건)이 연구지역 내에서 어떠한 시공간 군집경향을 보이는지 파악하기 위한 분석 방법이다. 분석 결과를 통해 해당 현상이 시공간적으로 군집되어있는지, 무작위하게 분포되어있는지, 혹은 분산되어있는지를 알 수 있다. 즉 이 유형에 해당하는 방법들은 군집도여부를 평가하기 위한 상대적인 기준에 해당한다. Knox test, Space-time K-function와 K-Nearest Neighbors가 이에 해당하는 가장 대표적인 시공간군집방법이다.

Knox(1964)는 사건의 거리와 빈도를 통해 시공간적 군집을 포착할 수 있는 Knox test방법을 제시하였다. Knox test는 개별 사건의 시공간적

거리와 빈도를 조사하여 각 사건들의 시공간적 거리가 한계수준보다 가까울 경우, 시공간군집으로 설정하는 방법이다. 그러므로 해당 방법은 미리 설정된 시공간 거리의 한계수준이 필요하다. 공간거리와 시간거리를 기준으로 멀고 가까움을 나누며, 이를 통해 공간군집, 시간군집, 시공간군집을 분류한다. 수식은 다음과 같다(식4-1).

$$X = \sum_{i=1}^N \sum_{j=1}^{i-1} S_{ij} t_{ij} \quad [\text{식 2-1}]$$

이때, S_{ij} 는 공간적 근접성을 나타내는 변수로, 사건 i, j 가 기준에 설정한 한계거리 내에 있으면 1, 그렇지 않으면 0으로 표현된다. t_{ij} 는 시간적 근접성을 의미하며, 공간적 근접성과 동일하게 i, j 가 기준에 설정한 한계거리 내에 있으면 1, 그렇지 않으면 0의 값을 갖는다. Knox test의 귀무가설은 X 가 공간적으로 독립되어 있는 포아송 분포를 따르며 그 기댓값은 $ST/n(n-1)$ 로 정의된다. 이때의 ST 는 시공간거리가 한계치보다 작은 값들의 경우들을 의미한다. 해당 방법론은 시공간 군집의 탐색과 관련하여 여러 연구에 응용되었다.

Space-time K-function(시공간 K-함수) 역시 어떠한 시공간 군집경향을 보이는지 파악하기 위한 분석 방법이다. 이것은 K-function을 시공간으로 확장한 것이며, K-function은 다음과 같이 정의된다(Bailey and Gatrell, 1995).

$$K(h) = \frac{[(\sum_{i=1}^n \sum_{j=1}^n) / n]}{(n/R)}, i \neq j \quad [\text{식 2-2}]$$

이때 n 과 R 은 각각 연구지역 내에서 발생한 사건의 개수와 연구지역의 면적을 나타낸다. i 와 j 는 연구지역 내 서로 다른 두 사건을 의미한다. I_h 는 사건 i 를 중심으로 또 다른 사건 j 가 반경 h 내에 있으면 1의 값

을, 그렇지 않으면 0의 값을 갖는 이진함수이다. 반경 h 는 사건들 간 공간적 인접성을 나타내는 매개변수이다. h 값에 해당하는 공간적 범위에 대하여 군집경향을 측정한다. 식에서의 분자는 각 사건의 발생지점을 중심으로 일정한 반경 내에 위치하는 다른 사건들 개수의 평균값이다. 이 평균값은 실제 발생한 사건들에 대하여 계산한 값이다. 식의 분모는 연구지역의 단위면적당 사건의 발생횟수, 즉 밀도를 의미한다. 결국 [식 4-2]은 사건의 발생지점(i)에서 일정한 반경(h)내에 들어오는 사건들의 개수를 연구지역 내 사건의 발생밀도로 나눈 것이다. 실제로 관찰된 자료들과 무작위한 패턴에서의 자료들을 대상으로 $K(h)$ 값을 구하고 비교를 통해 해당 사건들이 군집되어있는지, 무작위하게 분포되어있는지, 혹은 분산되어있는지를 파악할 수 있다. 이때, 일정한 공간적인 반경(h)과 함께 일정한 시간거리(t)를 동시에 고려함으로써, 해당 시공간 범위 내에서 발생할 것으로 예상되는 사건의 수를 계산한 것이 Space-time K-function이다.

Space-time K-function 이용한 연구는 다음과 같다. 먼저 박선영·김영호(2009)는 서울시를 대상으로 Space-time K-function(시공간 k함수)와 Space-time Permutation Scan Statistics(시공간순열스캔통계)를 통해 섬유 및 의복제조업의 시공간클러스터를 분석하였다. 이때 시공간에 따른 집중현상을 확인하고, 그 차이를 분석하기 위해 시공간 K-함수를 활용하였고, 섬유 및 의복제조업 산업에서는 모두 시공간적 군집현상이 나타나고 있음을 확인하였다.

Arbia et. al.,(2014)은 이탈리아의 Veneto 지역을 대상으로 제약 및 의료 기기 회사의 시공간적 군집을 살펴보았다. 이를 위해 기업들의 개폐일 정보를 활용하였고, 시공간 K-함수를 통해 군집 시기를 살펴보았으며, 전체 산업체 정보와 비교를 통해 해당 산업이 시공간적으로 군집해 있던 시기를 파악하였다.

한편 K-nearest Neighbor(Jacquez, 1996)은 공간적으로도 가까운 거리에 있는 사건들이 시간적으로도 가깝게 위치해있는지를 판단하는 방법이다. 각 사건들의 시공간 거리에 대하여, 연구자가 사전에 설정한 K개만

클 시공간적으로 가까운 사건들이 분석의 대상이 된다. 예를 들어 k 를 5로 설정을 하였다면, 각 각의 시간거리와 공간거리가 가장 가까운 5개씩의 사건이 분석의 대상이 된다.

$$J_k = \sum_{i=1}^N \sum_{j=1}^N s_{ijk} t_{ijk} \quad [\text{식 2-3}]$$

이때 i 와 j 는 사건을 의미하고, s_{ijk} 는 공간적인 k 근린지표를 t_{ijk} 는 시간적인 k 근린 지표에 해당한다. 각각 i 와 j 가 공간적 및 시간적으로 k 근린 지표 안에 해당하는 경우에 따라 s_{ijk} 와 t_{ijk} 는 1의 값을 가지며, 그렇지 않는 경우 0의 값을 갖는다. 시공간군집이 존재할 때, 공간적으로 k 근린에 해당하는 사건들은 시간적으로도 k 근린에 해당하기 때문에 J_k 는 큰 값을 가지게 된다. 통계적 유의도 검사는 몬테카를로시뮬레이션을 통해 검정이 된다.

Jacque(2010)은 원격탐사영상의 자동화된 분류 기법을 증대시키고자 연구를 진행하였고, 이를 통해 파푸아뉴기니의 열대우림을 탐지하였다. 1992~1993년까지의 계절별 영상을 이용하였으며, K-nearest Neighbors 기법을 통해 열대우림 및 비열대우림의 범위를 설정하였고, 시공간적 군집을 포착하였다.

Cumulative Sum(Page, 1954) 기법 역시 시공간클러스터링 탐색 방법에 해당한다. 이 기법은 일련의 과정에 양적인 변화가 발생하였을 때, 이를 발견하는 모델이다. 이 모델은 새로운데이터를 추가할때마다 반복적으로 군집분석을 다시 시행하고, 개별시간에서 연속적으로 분석이 진행된다는 점에서 전망적 연구에 해당한다. 그 수식은 다음과 같다.

$$y_t = \max\left(0, y_{t-1} + \left(\frac{x_t - \mu_t}{\delta_t} - k\right)\right) \quad [\text{식 2-4}]$$

$$y_0 = 0$$

이 모델의 귀무가설은 $x_t \sim N(\mu_t, \delta^2)$, 독립정규분포를 가정한다. k는 변화량을 탐색하는 데의 기준이 되는 파라미터 값으로, 사전에 설정된다. 이 모델은 평균값에서 표준편차의 k배 이상으로 차이가 나는 값들을 축적한다. 이때, y_t 값이 한계치보다 커질 때, 유의미한 변화가 발생한 것으로 해석한다. Grobler *et al.*,(2012)는 남아프리카의 Gauteng과 Limpopo 지역을 대상으로 토지피복 변화를 시계열적으로 살펴보았다. 원격탐사 영상을 기반으로 Cumulative Sum test를 적용하여 토지피복별 분류 및 시공간적 변화를 탐지하였다.

두 번째 유형인 ‘시공간클러스터 탐지테스트’는 현상이 어떠한 공간과 어떠한 시점에서 군집되어 발생하는지를 찾아내는 방법이다. 앞선 시공간클러스터링 테스트가 전역적인 분석이라면, 시공간클러스터 탐지 테스트는 전역적 군집을 형성하는데에 기여를 한 지역 및 시점이 구체적으로 어디인지를 찾아내는 것이다. Scan Statics(Kulldorff *et al.*, 2001.)을 시간축으로 확대한 Space-time Scan statistics이 가장 대표적인 예시에 해당한다. Space-time Scan statistics는 Scan Statics와 비슷한 방식으로 군집의 탐색의 기준이 되는 모양(탐색 윈도우)를 사전에 설정하고 크기가 다른 원통을 중복하여 사용한다. 원통 윈도우는 행정구역의 센트로이드를 중심으로 공간을 탐색하는 원통의 바닥과, 시간을 나타내는 원통의 높이로 구성되어있다. 원통 윈도우는 0에서 공간과 시간탐색의 최대 크기까지 증가시키며 이웃한 행정구역의 센트로이드로 이동한다(Kulldorff *et al.*, 2005). 다음 일련의 과정([식 2-5] ~ [식 2-8])을 통해 계산한다.

$$\mu_{zt} = \frac{1}{C} \left(\sum_z C_{zt} \right) \left(\sum_t C_{zt} \right) \quad [\text{식 2-5}]$$

$$C = \sum_z \sum_t C_{zt}$$

μ_{zt} 는 각 지역,각 시간에서 발생할 사건의 기대값을 의미한다. 지역z에서 관측된 모든 사건의 개수와 시간 t동안 모든 지역에서 관측된 사건의 개수의 곱을 관측된 총사건의 개수인 C로 나눈 것이다. 이때 C_{zt} 는 시간

t 동안 지역 z 에서 관측된 사건의 값이다

[식 2-5]를 기본으로 원통 A 속에서 발생된 사건의 기댓값인 μ_A 를 구한다[식 2-6]. 이때 특정 원통 A 속의 사건의 기대값 μ_A 는 원통 내의 모든 z 와 t 에서 발생할 사건의 기댓값의 합이다.

$$\mu_A = \sum_{(z,t) \in A} \mu_{zt} \quad [\text{식 2-6}]$$

$\sum_{z \in A} C_{zt}$ 와 $\sum_{t \in A} C_{zt}$ 가 모두 C 보다 작을 때, 원통 속 사건의 관측 개수인 C_A 는 포아송분포에 근사한다는 가정을 하고 있기 때문에, 원통별 사건 발생의 측정에 포아송 우도비를 사용할 수 있다. 이는 다음의 [식 4-7]과 같이 구한다.

$$\left(\frac{C_A}{\mu_A}\right)^{C_A} \left(\frac{C - C_A}{C - \mu_A}\right)^{C - C_A} \quad [\text{식 2-7}]$$

[식 4-7]은 원통 내부의 관측치를 원통 내부의 기댓값으로 나누는 것과 원통 외부의 관측치를 원통 외부의 기댓값으로 나누는 것의 곱이다. 모든 원통의 우도비 중 최대우도비는 실제사건이 발생할 때에 시공간적으로 가장 개연성 있다는 것을 의미하며, 이를 통해 시공간균집이 어떠한 시공간적 위치에서 발생하는 지를 알 수 있다(Kulldorff *et al.*, 2005; 박선영·김영호, 2009).

시공간스캔통계량을 활용한 연구는 다음과 같다. 정경석 외(2010)는 시공간스캔통계량과 Ripley-K함수를 이용하여 도시범죄의 핫스팟분석을 진행하였다. 연구결과 범취위험도가 유의미하게 높은 지역들이 나타났으며, 시공간적인 분석 결과 기존 공간분포만을 고려하였을 때와는 다소 차이가 나타남을 알 수 있었다.

Abarmovich *et al.*,(2014)는 감상선암 질병에 대하여 시공간적 균집을 탐색하였다. 러시아의 민스크 지역에 대하여 행정구역별 질병감염자의

수를 카운팅하였으며 1991~2005년을 대상으로 진행을 하였다. 방법론으로는 Scan Statistics 방법을 이용하였다. 그 결과 1991-1995년, 1996-2000년, 2000-2003년 갑상선암 환자의 시공간 군집이 공간적으로 다르게 분포함을 파악할 수 있었다.

한편 시공간스캔통계량을 응용한 여러 방법들도 존재한다. ‘시공간순열스캔통계량(Space-Time Permutation Scan Statistics)’과 Bootstrap-based Space-time surveillance model이 대표적인 예시이다. 시공간순열스캔방법은 그 방법론 자체는 시공간스캔통계량과 유사하다. 유의성 검증을 위해 위험인구 데이터를 생성함으로써 각 사건의 시공간적 위치정보에 대하여 임의적인 순열을 발생시켜 많은 양의 개수를 생성하고 이를 위험인구로 사용하는 것이다(Kulldorff *et al.*, 2005).

시공간순열통계를 활용한 연구는 다음과 같다. 먼저 앞서 언급한 박선영·김영호(2013)역시 시공간순열스캔통계량을 활용하였다. 섬유 및 의복제조업의 시공간클러스터의 패턴을 파악하기 위해 활용하였다.

배선학 외(2013)는 구제역 전과양상의 시공간클러스터를 일단위로 분석하였다. 시공간스캔통계량의 확장모델인 Permutation model을 이용하여 분석을 진행하였다. 가축의 종별 분석 결과, 발생 시기 및 군집도가 상이하게 발생하였으며 이를 바탕으로 차단방역의 우선순위를 제안하였다.

Kim and O’Kelly(2008)은 리샘플링 과정에 해당하는 Bootstrap 방법을 통해 기존의 Space-time Scan Statistics의 관측값의 정확도를 개선시킨 Bootstrap-based Space-time surveillance model을 개발하였다. Bootstrap 방법은 통계적 추론의 단계에서 리샘플링 과정을 거쳐 새로운 통계적 분포를 만들어내는 것인데, 이 분포를 통해 특정 신뢰수준 하에서 통계적 검증이 가능하다. 과정을 기존의 방법론에 비해 보다

범죄데이터에 대하여 Bootstrap-based space-time surveillance model을 적용하였고, 그 시공간군집을 포착하였으며 군집의 변화 양상을 탐색하였다.

이처럼 시공간군집의 변화양상을 탐색할 수 있는 시공간클러스터 탐색

기법으로를 탐색하기 위해 Space-time Scan statistics은 시공간군집을 측정하기 위해 여러 연구에서 활용이 되었으며(박선영·김영호, 2013; 배선학외, 2013; Arbia *et al.*,2014; Abarmovich *et al.*, 2014), 다양하게 응용되었다.(Kulldorff *et al.*, 2005; Kim and O'kelly, 2008).

이처럼 Space-time Scan statistics에 기반하여 시공간군집을 파악한 연구가 다수 존재하였다. 시공간스캔통계량은 시공간클러스터를 탐지하는데에 가장 대표적인 방법론이지만, 기본적으로 시공간스캔통계량은 기본적으로 스캔통계량을 응용한 것이기 때문에 이 역시 스캔통계량의 한계점을 계승한다. 시공간스캔통계량에서는 군집의 모양을 원통형이나 원뿔형으로 미리 가정을 한다. 그렇게 때문에 시공간데이터의 불규칙적인 클러스터를 파악하지 못하고, 클러스터를 과대 추정하는 경향이 있다는 한계점을 갖는다.

이러한 한계점을 극복하기 위한 시도가 이루어졌다. Kim(2009)과 김대중·구형수(2012)는 Space-time Scan statistics에 AMOEBA 기법을 적용한 Space-time Chain statistics(시공간체인통계량) 기법을 제시하여 불규칙적인 시공간클러스터를 확인하고자 하였다. 이 기법에서는 원통형 탐색 위도우를 체인과 같은 형태로 이동시켜 불규칙한 시공간 클러스터의 탐색이 가능하다. 그들은 시공간스캔통계량을 이용하여 어떠한 지점에서 토지거래가 집중적으로 발생하는지를 탐색하였다. 하지만 이 방법론은 유사한 셀이 인접해 있는 '면 형태'의 카운트자료에 적용하기는 힘들다는 한계점을 갖고 있다.

이에 대한 대안적 접근으로 Kroon, J.,(2010)은 Aldstadt and Getis(2006)의 AMOEBA(A Multidirectional Optimum Ecotope-Based Algorithm)기법을 응용한 Spatio-temporal AMOEBA를 제안하였다. 하지만 해당 연구는 학술발표 외 논문 등의 후속연구가 진행되지 않았다. Kim(2009) 역시 사상들을 단순히 공간적으로 이웃한 것이 아닌, 시공간적으로 이웃한 것으로 고려한다면 AMOEBA기법을 적용하여 시공간군집분석을 진행할 수 있을 것이라고 언급하였다.

지금까지 시공간군집기법에 대해 살펴보았다. 본 연구는 이동통신 빅

데이터를 통해 도시의 시공간 점유패턴을 살펴보고자 하며, 공간적 단위는 ‘기초구역단위’로 분석을 진행하고자 한다. 기초구역단위는 비교적 상세한 공간단위를 가진 면데이터에 해당한다. 그렇기 때문에 군집 역시 불규칙한 모양으로 나올 가능성이 높다. 이러한 이유에서 불규칙한 클러스터 탐색에 유리하고, 보다 상세한 공간단위에 적용하기 용이한 Spatio-temporal AMOEBA기법을 연구 방법론으로 활용하고자 한다.

제 3 장 연구 방법

본 장에서는 앞서 언급한 유동인구의 시공간 군집을 탐색하는 방법에 대하여 구체적으로 살펴보고자한다. 이를 위해 1절에서는 연구자료인 이동통신 빅데이터에 대해 보다 상세하게 알아보고, 본 연구를 위해 세가지의 사전설정을 설명하고자 한다. 그리고 2절에서는 방법론인 AMOEBA와 이를 시간축으로 응용하는 Spatio-temporal AMOEBA기법에 대해 설명하고자 한다.

제 1 절 연구 설정

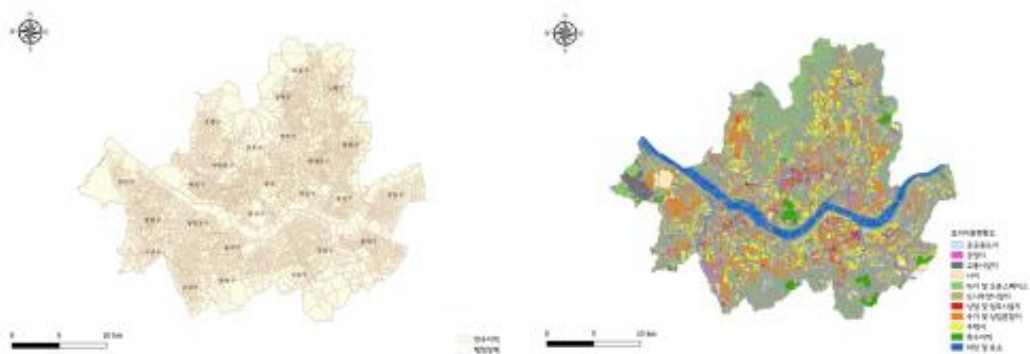
본 연구를 원활히 수행하기 위한 공간적, 시간적 사전설정을 진행하였다. 공간적 설정으로는 분석의 공간단위 설정을 진행하고 자료의 특성상 과소추정을 방지하기 위하여 유동인구의 분포 불가능지역 설정한다. 시간적 사전설정으로는 시간대별 합역 및 시간축의 시드의 설정이 이루어진다.

1. 공간적 사전 설정

먼저 분석의 공간단위를 기초구역으로 설정하였다. 본 연구에서 활용한 자료는 50m*50m의 포인트 데이터로 매우 상세한 공간단위로 제공이 된다. 하지만 산지, 수변과 같이 통신이 불가능한 지역에서는 그 값이 누락되어있는 등, 불규칙하게 데이터가 구성되어 있다<그림 3-1>. 또한 도시 내 유동인구의 시공간적 분포를 살펴보는 만큼 기존에 설정되어있는 공간단위를 바탕으로 분석을 진행하고자 하였다. 본 연구에서는 기초구역단위를 공간단위로 활용하였고, 이를 위해 포인트데이터를 면형의 공간데이터로 변환하였다. 기초구역 단위는 행정동 단위보다는 보다 상

제한 공간단위로, 현재 국가 행정서비스 등의 기준이 되는 공간단위 중 비교적 상세한 편에 속한다. 각 격자들의 중심점의 좌표에 대하여 포인트데이터 형태로 제공되는 이동통신 빅데이터를 spatial join을 통해 기초구역단위로 합역을 하였다.

둘째, 과소추정을 방지하고자 토지이용도를 보조자료로 활용하여 인구의 분포가능지 및 불가능지를 구분하였으며, 분포 불가능지의 면적을 제외한 기초구역 면적을 기준으로 인구밀도를 계산하였다. 앞서 언급하였듯이 산지, 수변 등과 같이 통신이 불가능한 지역에서는 그 값이 누락되어있다. 만일 유동인구의 분포를 탐색하는 과정에서 특정 기초구역단위가 통신불가지역을 포함하고 있으면, 그 지역의 유동인구밀도는 과소추정이 될 것이다. 보조자료를 이용하여 통신불가지역을 제외한 작업은 이를 방지하기 위한 사전 설정이다. 유동인구의 분포 불가능의 지역으로는 수면, 군사시설, 산지 등을 선택하였다<그림 3-3>.



<그림 3-1> 연구지역 및 토지이용도 지도

2. 시간적 사전 설정

시공간 AMOEBA에서는 시공간군집분석인 만큼 시간축에 대한 설정 또한 중요하다. 시간대의 경우, 원활한 분석을 위해 데이터의 크기를 축소하고자 3시간단위로 시간대를 합역하였다. ‘03-06시’, ‘06-09시’, ‘09시-12시’, ‘12시-15시’, ‘15시-18시’, ‘18시-21시’, ‘21시-24시’, ‘24시-03시’까지 총 8개의 시간대로 설정을 하였다. 예를 들어, 03-06시는 수면활동이

이루어지는 시간으로 고려하고, 06-09시는 출근 및 등교가 이루어지는 시간으로, 18시-21시는 퇴근이 이루어지는 등 일반적인 활동패턴에 따라 분류하였고, 그들을 하나의 시간대로 병합하였다.

시간대별 편차를 고려하기 위하여 유동인구의 밀도를 각 시간대별로 표준화하였다. <표3-2>는 시간대별 유동인구 수에 대한 단순 통계치이다. 표에서 알 수 있듯이 시간대별 인구분포의 차이가 크게 나타나고 있으며, 인구수가 가장 높게 집계되는 15-18시의 인구와 가장 낮게 집계되는 03-06시 사이의 인구차이가 약 2.6배의 차이가 발생함을 알 수 있다.

본 연구의 관심사는 유동인구의 시공간적 군집을 알아보는 것이다. 이를 위해 각 시간대별로 인구가 군집하는 지역을 살펴보고, 이를 바탕으로 시공간적 이웃과 그 군집 정도를 비교함으로써 군집을 설정한다. 이러한 과정에서 시간대에 따른 클러스터의 과대추정 및 과소추정을 방지할 필요성이 있으며, 인구밀도를 각 시간대별로 표준화함으로써 이를 해소하고자 하였다.

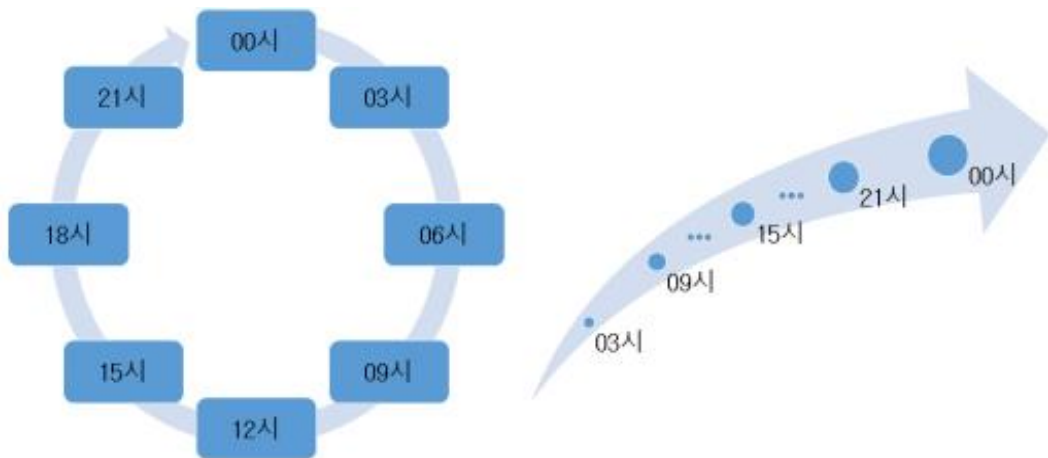
<표 3-1> 시간대별 유동인구

단위: 천명

시간대	03-06	06-09	09-12	12-15	15-18	18-21	21-24	00-03
합계	3600.21	4186.47	6844.65	8644.35	8305.63	7838.34	5567.15	3870.11
최대값	3.2	3.9	8.7	14.1	15.3	13.5	8.3	3.3
평균	3.0	3.1	4.8	7.2	8.0	7.5	5.8	3.7
표준편차	3.1	3.3	5.5	9.1	11.5	10.8	7.7	4.0

시공간군집분석을 위한 시간적 시드의 설정 또한 이루어져야 한다. 시공간군집분석은 시간 축을 일방적인 방향성을 갖고 있는 것으로 가정하며, 시드선정에 있어서 시간축을 따로 고려하지 않는다. 하지만 본 연구는 ‘시간대’라는 순환하는 성격의 시간적 데이터를 이용하기 때문에 시드를 선정하는데 있어서 시간대를 고려하여야 한다. 이를 위해 시간대라는 순환하는 시간 축을 순환하지 않는 일방향성을 갖는 반직선형태로 조정

을 해야한다. 본 연구에서는 시드의 시간축 설정에 있어서 일반적인 인간의 하루 활동시간을 고려하였으며, 03시-06시를 클러스터의 시작으로 설정하였다. 00-03시의 시간대의 경우 그 전날의 활동에서 이어지는 것으로 판단하였다. 예를 들어 월요일의 00-03시에 해당하는 시간대는 일요일에서의 활동의 연장으로 가정하였다.



<그림 3-2> 시간축 설정

제 2 절 시공간 군집 방법: Spatio-temporal AMOEBA

본 절에서는 연구에서 이용한 시공간군집탐색기법인 Spatio-temporal AMOEBA기법에 대해 살펴보려고 한다. 시공간아메바 기법은 공간군집 방법인 아메바 기법을 시공간 축으로 확장 및 적용한 방법이다. 이를 위해 우선 시공간아메바기법의 바탕이 되는 아메바 기법에 대해 설명한다. 이후 이를 시공간으로 적용한 Spatio-temporal AMOEBA 기법을 살펴본다.

1. AMOEBA 기법

AMOEBA(A Multidirectional Optimum Ecotope Based Algorithm)기

법은 공간가중행렬을 구성하는 동시에 공간클러스터의 기하학적 형태를 탐색하기 위한 알고리즘이다. 이 알고리즘은 시드(Seed, 클러스터 탐색에 사용되는 시작위치)로부터 여러 방향으로 공간적 연관성을 탐색하여, 가장 높은 값들이 모여 있는 최적의 클러스터를 도출한다(이상일 외, 2010; 김영호, 2012). 시드라는 하나의 공간단위에서 탐색을 시작하기 때문에 미세하고 로컬한 스케일에서의 공간적 연관성도 밝혀낼 수 있다는 특징을 갖고 있다.

이때, 전체 공간 데이터셋 내에서 공간적 연관성의 소구역을 찾으며, 이를 에코톱이라고 부른다. 아메바는 이러한 에코톱을 찾기 위한 알고리즘이며, 이러한 에코톱은 대부분 공간적으로 불규칙하거나 아메바와 같은 모양을 가지고 있다(Aldstadt and Getis, 2006).

AMOEBAs는 국지적으로 공간적 자기상관성을 밝히는 통계량인 G_i^* 를 기반으로 한다. G_i^* 에서 i 는 특정 공간단위를 뜻한다. 즉 G_i^* 는 특정 공간단위에 초점을 두고 있는 로컬한 통계량이다(Getis and Aldstadt, 2004).

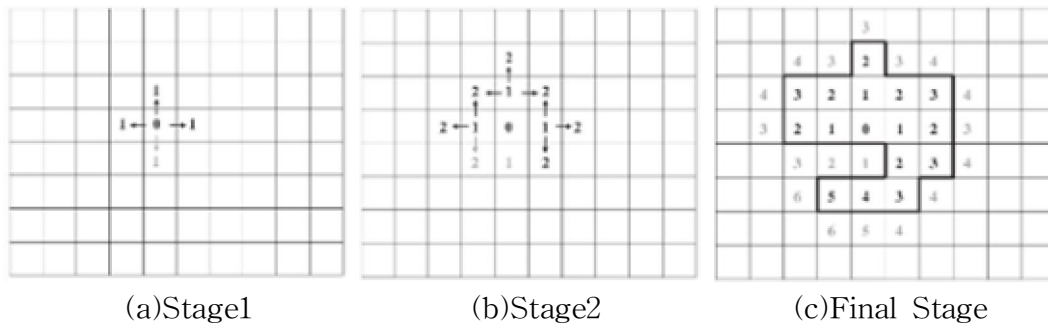
앞에서 언급하였듯이 AMOEBAs 기법은 LISA를 이용하여 공간클러스터의 범역을 설정하는 기법이다. AMOEBAs에서 활용하는 LISA는 Getis-ord의 G_i^* 이며, 이는 다음과 같은 수식을 통해 구해진다.([식3-1])

$$G_i^* = \frac{\sum_{j=1} w_{ij} x_j - \bar{x} \sum_{j=1} w_{ij}}{\sqrt{\frac{n \sum_{j=1} w_{ij}^2 - (\sum_{j=1} w_{ij})^2}{n-1}}} \quad [\text{식 3-1}]$$

$$s = \sqrt{\frac{\sum_{j=1}^N x_j^2}{N} - (\bar{x})^2}$$

이때 s 는 표준편차를, w_{ij} 는 공간가중행렬의 요소 값을, n 은 전체 케이스 수를 의미한다. i 와 j 는 개별 공간단위(혹은 위치)를 의미하는 것으로

두 공간단위가 이웃으로 정의되면 $w_{ij} = 1$ 이, 그렇지 않으면 $w_{ij} = 0$ 이 되며, 자신을 이웃으로 간주하기 때문에 $w_{ii} = 1$ 로 주어진다. 이 통계량의 기댓값은 0이고, 분산은 거의 1이다. 따라서 이 통계량의 유의성 검정은 정규분포를 상정한 표준화 점수에 대한 것과 거의 동일하게 이루어진다 (Aldstadt and Getis, 2006; 이상일 외, 2010). Geits-rod의 G_i^* 가 갖는 가장 장점은 통계치로부터 직관적으로 핫스팟인지, 콜드스팟인지 확인할 수 있다는 점에 있다. 즉 통계치가 양수이면, 핫스팟과 연관되어있고, 음수면 콜드스팟과 관련된다.



<그림 3-3> AMOEBA 알고리즘(Aldstadt and Geits, 2006)

<그림 3-3> 은 AMOEBA 알고리즘을 단계별로 보여주고 있다. 0단계에서, 우선적으로 한 셀에 대한 그 셀만을 대상으로 G_i^* 를 계산하고, 그 값을 $G_i^*(0)$ 이라고 설정한다. 여기서 $G_i^*(0)$ 값이 0보다 크다면 이는 해당 셀의 값이 다른 셀의 평균값보다 높다는 의미로, 잠재적인 핫스팟에 해당하며, $G_i^*(0)$ 값이 0보다 작으면 그 반대를 의미한다. 그 셀을 중심으로 상하좌우에 위치한 네 개의 그리드 셀을 이웃으로 규정하고, 그 네 셀로 가능한 모든 조합 각각에 중심 셀을 합하여 규정되는 ‘구역’에 대해 G_i^* 를 계산하는 것이 1단계의 핵심이다(이상일 외, 2010). 이 네 개의 셀로 만들 수 있는 조합은, 네 개에서 하나를 뽑는 방법, 네 개에서 두 개를 뽑는 방법, 네 개에서 세 개를 뽑는 방법, 네 개에서 네 개를 뽑는 방법을 모두 합한 것이다. 이 경우 15가지(4+6+4+1)조합이 가능하다. 이 각각의

조합에 중심 셀을 포함시켜 서로 다른 15개의 구역이 1차적인 클러스터의 후보가 된다. 이때, 구역에 대해 계산된 값을 $G_i^*(1)$ 이다. $G_i^*(0)$ 가 0보다 큰 경우, 이 $G_i^*(1)$ 들 중 최대값이 $G_i^*(0)$ 보다 크다면, 그 최대값을 갖는 지역을 1차적인 핫스팟으로 간주한다. $G_i^*(0)$ 가 0보다 작은 경우에는, $G_i^*(1)$ 들 중 최소값이 $G_i^*(0)$ 보다 작으면, 그 최소값을 갖는 지역을 1차적인 콜드스팟으로 간주한다. 그림에서의 (a)를 보면 중심 셀과 상, 좌, 우에 위치한 셀을 합한 것이 1차적인 클러스터로 설정된다. 이때 누락된 하면의 셀은 다음의 모든 과정에서 배제된다. 즉 절대로 어떤 클러스터의 구성이 될 수 없다.

2단계에서는 1단계에서 포함된 이웃 셀을 중심으로 상하좌우의 가능한 위치에 두 번째 이웃 셀의 세트를 설정한다. 그림의 (b)는 (a)를 전제할 때 7개의 이웃 셀이 클러스터 후보로 설정될 수 있음을 보여준다. 앞의 단계와 마찬가지로 모든 가능한 조합에 대해 G_i^* 를 계산한 뒤, $G_i^*(2)$ 최대값이 $G_i^*(1)$ 보다 크면, 최대값을 산출한 조합이 2차적인 클러스터가 된다. 이러한 과정을 G_i^* 값의 더 이상 커지지 않을 때 까지 진행하게 된다. 최종적으로 도출된 세트들의 조합이 클러스터의 범역이 된다. 같은 방식으로 3~6단계를 진행하면, 더 이상의 G_i^* 값의 증가가 없는 그림(c)와 같은 클러스터 범역이 도출된다.

위의 과정은 하나의 셀을 중심으로 공간클러스터의 범역을 설정하는 방법이다. 이러한 과정이 모든 셀에 대하여 진행된다면, 범역들이 서로 겹치는 문제가 발생할 것이므로, 공간클러스터를 선별하는 이차적인 기준이 필요하다. 위와 같은 그리드 셀 개수만큼의 공간클러스터가 형성되고, 범역들이 서로 겹치는 문제가 발생한다. 이를 해결하기 위해 Aldstadt and Getis(2006)은 G_i^* 를 내림차순으로 정리한 다음, 가장 큰 G_i^* 값을 보이는 클러스터를 우선적으로 선택하고 해당 클러스터와 공간적 범위를 공유하고 있는 클러스터들을 제거한다. 그리고 두 번째로 높은 G_i^* 값을 갖는 클러스터를 선택하고, 마찬가지로 해당클러스터와 공간적 범위를 공유하는 클러스터들을 모두 제거한다. 이와 같은 과정을 반복하면, 최종

적으로 공간클러스터들의 세트가 도출된다. AMOEBA 기법은 AMOEBA 기법은 다른 클러스터 기법과는 달리 클러스터의 영역을 미리 가정하지 않는다. 그렇기 때문에 불규칙적인 클러스터를 탐색가능하며, 클러스터를 과대추정 할 가능성이 적다.

2. Spatio-temporal AMOEBA 기법

본 연구에서 활용하는 Spatio-temporal AMOEBA 기법은 공간군집분석인 AMOEBA 기법을 시간 축으로도 확대한 것이다. 이를 위해서는 우선 각 시간대별로 G_{it}^* 의 값을 계산한다. G_{it}^* 는 t 시점에서의 G_i^* 값을 의미한다. 이를 위해 우선 시공간이웃(Spatio-temporal Neighbors)를 측정하고, 이를 바탕으로 시공간가중행렬인 w_{ijt} 를 생성한다. 이때, t 시점을 기준으로 시간 축에서의 이웃은 $t+1$ 인 경우만을 고려한다. 본 연구에서의 i 와 j 는 연구지역인 서울시 기초구역의 개별 공간단위에 해당하고, t 는 연구의 시간단위에 해당한다.

이후 앞선 절에서 시공간가중행렬을 기반으로 시공간 자기상관성을 측정하는 G_{it}^* 를 계산한다. 그 식은 다음과 같다([식 3-2]).

$$G_{it}^* = \frac{\sum_{j=1} w_{ijt} x_j - \bar{x} \sum_{j=1} w_{ijt}}{\sqrt{\frac{n \sum_{j=1} w_{ijt}^2 - (\sum_{j=1} w_{ijt})^2}{n-1}}} \quad [\text{식 3-2}]$$

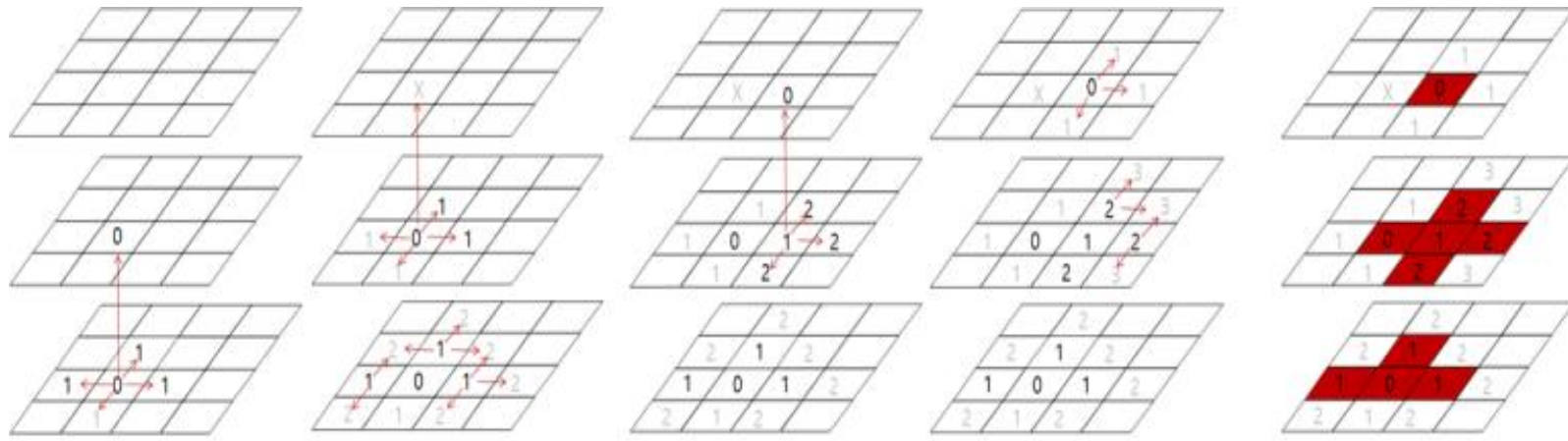
$$s = \sqrt{\frac{\sum_{j=1}^N x_j^2}{N} - (\bar{x})^2}$$

모든 시공간지역에 대하여 G_{it}^* 를 계산하고, 그 값을 기준으로 AMOEBA 분석을 진행한다. 이때 AMOEBA 분석을 진행함에 있어 이웃(neighborhood) 역시 상하좌우, 그리고 $t+1$ 시점에 위치한 다섯 개의

그리드 셀을 이웃으로 규정한다. 이는 결국 공간적인 관계와 시간적인 관계를 모두 포괄한 것이다. <그림3-4>은 Spatio-Temporal AMOEBA 알고리즘을 단계별로 보여주고 있다. 0단계에서, 우선적으로 한 셀에 대한 그 셀만을 대상으로 G_{it}^* 를 계산하고, 그 값을 $G_{it}^*(0)$ 이라고 설정한다. 여기서 $G_{it}^*(0)$ 값이 0보다 크다면 이는 해당 셀의 값이 다른 셀의 평균값보다 높다는 의미로, 잠재적인 핫스팟에 해당하며, $G_{it}^*(0)$ 값이 0보다 작으면 그 반대를 의미한다. 그 셀을 중심으로 상하좌우, 그리고 $t+1$ 시점에 위치한 셀($G_{it+1}^*(0)$)까지 총 다섯 개의 그리드 셀을 이웃으로 규정하고, 그 다섯개의 셀로 가능한 모든 조합 각각에 중심 셀을 합하여 규정되는 ‘구역’에 대해 G_{it}^* 를 계산하는 것이 1단계의 핵심이다. 이 각각의 조합에 중심 셀을 포함시킨 구역이 1차적인 클러스터의 후보가 된다. 이때, 구역에 대해 계산된 값을 $G_{it}^*(1)$ 이다. $G_{it}^*(0)$ 가 0보다 큰 경우, 이 $G_{it}^*(1)$ 들 중 최대값이 $G_{it}^*(0)$ 보다 크다면, 그 최대값을 갖는 지역을 1차적인 핫스팟으로 간주한다. $G_{it}^*(0)$ 가 0보다 작은 경우에는, $G_{it}^*(1)$ 들 중 최소값이 $G_{it}^*(0)$ 보다 작으면, 그 최소값을 갖는 지역을 1차적인 콜드스팟으로 간주한다. 그림에서의 (a)를 보면 중심 셀과 상, 좌, 우, 시간축 상에 위치한 셀을 합한 것이 1차적인 클러스터로 설정되었다.

2단계에서는 1단계에서 포함된 이웃 셀을 중심으로 상하좌우, 그리고 $t+1$ 시점에 위치한 셀까지 총 다섯 개의 셀들을 두 번째 이웃 셀의 세트를 설정한다. 그림의 (b)는 (a)를 전제할 때 2차적인 클러스터를 보여준다. 앞의 단계와 마찬가지로 모든 가능한 조합에 대해 G_{it}^* 를 계산한 뒤, $G_{it}^*(2)$ 최대값이 $G_{it}^*(1)$ 보다 크면, 최대값을 산출한 조합이 2차적인 클러스터가 된다. 이러한 과정을 G_{it}^* 값의 더 이상 커지지 않을 때 까지 진행하게 된다. 최종적으로 도출된 세트들의 조합이 클러스터의 범역이 된다. 이러한 과정을 반복하여 더 이상의 G_{it}^* 값의 증가가 없는 그림(e)와 같은 클러스터 범역이 도출된다. 이후, 겹쳐지는 클러스터의 영역을 제거하기 위하여 G_{it}^* 를 내림차순으로 정리한 다음 겹쳐지는 영역의 G_{it}^* 를 제거해나간다. 해당 과정을 통해 최종적인 시공간클러스터가 도출된

다. 이때 분석 결과의 정확성은 몬테카를로시뮬레이션 시행과 p-value를 통해 보정한다. 서울시 전역을 대상으로 Spatio-temporal AMOEBA 기법을 통해 시공간 클러스터 탐색을 진행하였으며, 총 99번의 몬테카를로 시뮬레이션 시행과 0.05이하의 p-value 값을 유의수준으로 설정하였다.



(a) Stage 1

(b) Stage 2

(c) Stage 3

(d) Stage 4

(e) Stage 5

<그림 3-4> Spatio-temporal AMOEBA 모식도

제 4 장 유동인구에 따른 도시의 시공간 점유 패턴 분석 결과

이번 장에서는 유동인구의 시공간 군집을 살펴봄으로써 도시의 시공간 점유 패턴을 파악하고자 한다. 이를 위해 앞선 장에서 설명한 이동통신 빅데이터와 Spatio-temporal AMOEBA 기법을 통해 유동인구의 시공간군집 분석을 진행함으로써, 군집의 공간적인 분포 및 시계열적 변화를 탐지한다. 1절에서는 서울시를 대상으로 유동인구의 시공간 군집 탐색을 실시하고 그 결과, 즉 점유 패턴을 포착한다. 2절과 3절에서는 성별 및 연령에 따른 시공간 군집 탐색을 실행함으로써 각 인구적 속성에 따른 도시 점유 패턴 변화를 파악한다.

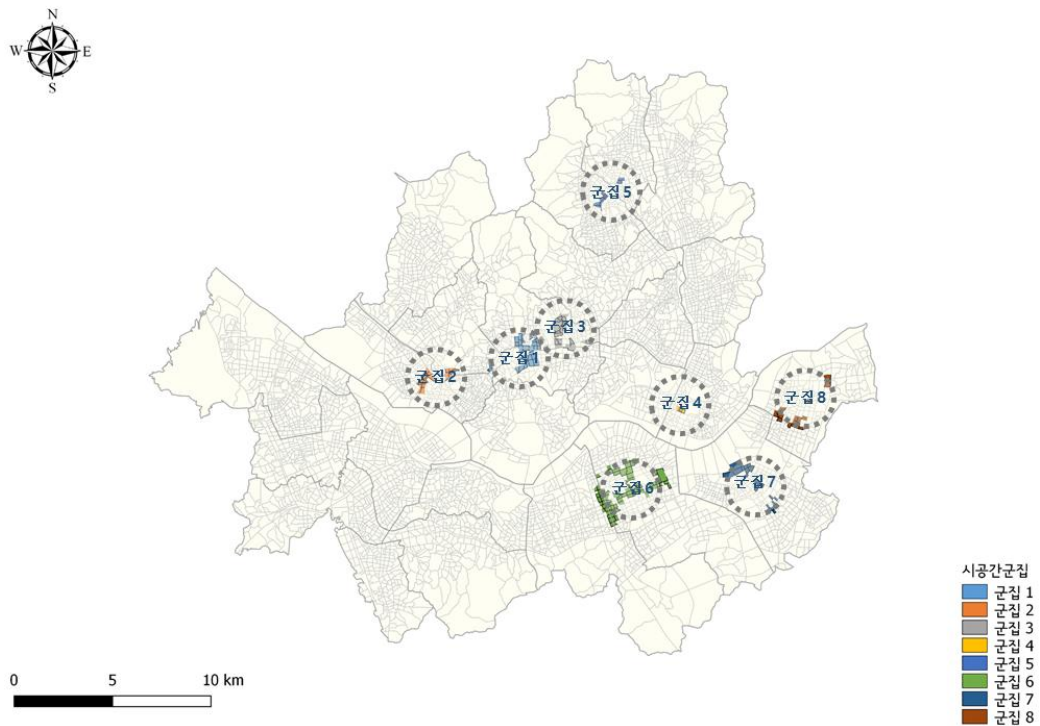
제 1 절 서울시 유동인구의 도시 점유 패턴

1. 서울시 유동인구의 시공간군집 탐색 결과

서울시를 대상으로 유동인구의 시공간군집분석을 실시하였으며, 그 결과 총 8개의 군집이 탐색되었다. 각 군집은 종로구-중구(군집 1), 홍대-신촌(군집 2), 동대문-혜화(군집 3), 건국대학교 인근(군집 4), 수유-쌍문(군집 5), 강남대로-테헤란로(군집 6), 잠실-수서(군집 7) 그리고 강동구(군집 8)<그림 4-1>에 위치한다. 유의수준은 모두 0.05이하로 유의미한 시공간군집에 해당한다<표 4-1>. 탐지된 군집의 시공간적 변화 양상을 통해 도시의 점유 패턴을 파악할 수 있다.

<표 4-1> 유동인구의 서울시 점유 공간

분류	위치	공간적 변화
군집 1	종로구-중구	확장 후 축소
군집 2	홍대-신촌	확장 후 축소
군집 3	동대문-혜화	확장 후 축소
군집 4	건국대학교 인근	확장 후 축소
군집 5	수유-쌍문	축소 후 확장
군집 6	강남대로-테헤란로	확장 후 축소
군집 7	잠실-수서	축소 후 확장
군집 8	강동구	축소 후 확장

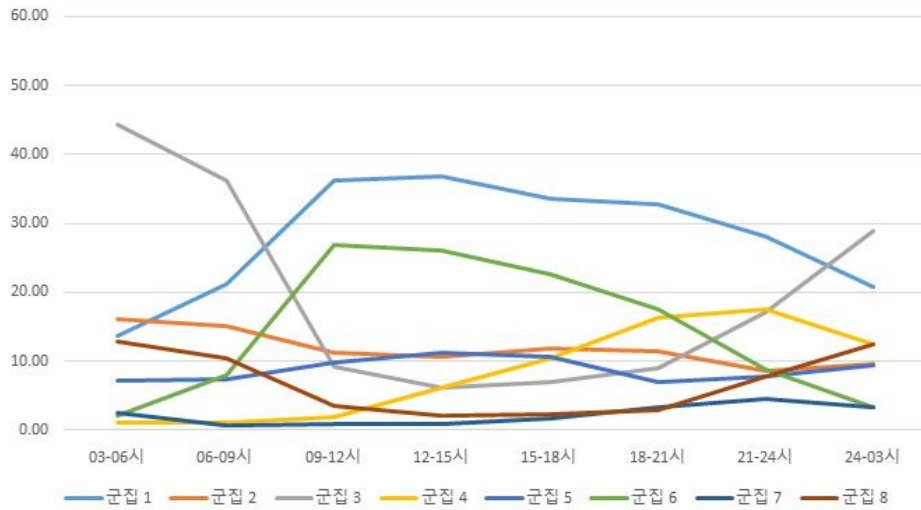


<그림 4-1> 도시의 시공간 점유 지역 탐색 결과(시공간 군집)

탐지된 군집들의 총 인구는 각 시간대별로 295.52천명(03-06시), 321.59천명(06-09시), 862.67천명(09-12시), 1363.65천명(12-15시), 1273.62천명(15-18시), 1150.96천명(18-21시), 713.32천명(21-24시), 368.30천명(24-03시)이다. 한편 탐지된 군집의 총 인구를 전체 유동인구와 비교해보면, 각각 8.21%(03-06시), 7.26%(06-09시), 12.60%(09-12시), 15.78%(12-15시), 15.33%(15-18시), 14.68%(18-21시), 12.81%(21-24시), 9.52%(24-03시)로 증가 후 감소하는 양상을 보인다. 오전과 새벽시간대에 비해 낮과 오후 시간에 인구의 유동이 많이 이루어지고 있으며, 공간적으로도 보다 군집해있다.

분석 결과를 각 군집 및 시간대별로 정리하면 <표 4-2>와 같다. 표의 좌측은 각 군집의 유동인구 수를, 표의 우측은 전체 군집 대비 각 군집의 유동인구 비율을 표현한 것이다. 예를 들어, 03-06시의 군집 1에는 총 40.41천명의 유동인구가 분포해 있으며, 이는 해당 시점의 모든 군집에 속하는 유동인구 중 13.61%의 인구에 해당한다. 각 수치들을 통해 군집 내 유동인구의 절대적인 수치를 알 수 있으며, 타 군집들과의 상대적인 비교를 할 수 있다.

우선 탐지된 군집의 전반적인 양상을 시간대 별로 살펴보면 03-06시에는 군집 3(131.08천명, 44.36%), 군집 2(47.49천명, 16.17%), 군집 1(40.41천명, 13.67%), 군집 8(38.21천명, 13.97%) 군집 6(21.31천명, 7.21%), 군집 4(6.32천명, 2.14%), 군집 5(3.05천명, 1.03%)의 순으로 분포한다. 09-12시에는 군집 1(3012.38천명, 36.21%) 및 군집 4(232.60천명, 26.96%)의 성장 및 군집 3(79.33천명, 9.20%), 군집 2(97.36천명, 11.29%), 군집 8(31.24천명, 3.62%)의 축소가 두드러지게 드러난다. 12-15시에는 이와 같은 변동성이 심화되고 있으며, 특히 군집 5의 증가가 크게 나타난다. 15-18시를 기점으로 군집 1과 군집 4는 감소하는 반면 군집 3과 군집 8은 증가하는 추세를 보인다. 한편 24-03시에는 03-06시에서의 분포와 비슷하게 다시 군집 3이 가장 높게 나타나고 있으며, 군집 1(20.87%) 역시 그 군집 양상은 21-24시에 비해 감소한다<그림 4-2>.



<그림 4-2> 점유 지역의 유동인구 변화 (전체군집 대비 유동인구 수)

2. 군집별 탐색 결과

분석 결과에 대하여 보다 구체적으로 살펴보면 다음과 같다. 먼저 군집 1(종로구-중구)은 행정구역상 종로구의 종로1·2·3·4가동, 사직동과 중구의 명동, 소공동, 회현동 및 중림동에 걸쳐있으며, 지하철역을 기준으로 광화문역, 시청역, 서울역, 을지로입구역, 명동역, 충정로역 등 인근에 위치한다. 03-06시 중림동의 충정로역과 06-09시에는 종로1·2·3·4가동의 광화문 및 회현동의 시청 부근에서 탐색되기 시작하며, 점차 명동·소공동 등의 인접한 지역으로 확장되어 15-18시까지 유지된다. 18-21시 이후 군집은 급속도로 축소되며, 24-03시에는 종로1·2·3·4가동의 종각 일대와 중림동의 충정로역 부근에서만 탐색된다. 이때 03-06시에서 군집에 속하는 유동인구 수는 40.41천명이고 이는 전체 군집 대비 13.67%에 해당한다. 06-09시에는 68.11천명, 해당 시점의 21.18%에 해당하는 인구가 분포한다. 지속적인 증가 후, 18-21시를 기점으로 376.11천명(4.80%), 24-03시에는 76.87천명(1.99%)으로 전체군집 대비 유동인구수의 비율은 감소한다. 이 지역은 중심업무지구에 속하는 지역이며, 오전 및 야간 시간대에 비해 오후시간대에 군집의 공간적 범위와 해당지역을 방문하는

인구가 많은 것으로 보인다. 공간적 범위는 도심의 외곽에서 시작되어 점차 중심지역으로 이동 및 확장을 한 뒤 다시 후퇴하는 양상으로 나타나며, 인구 역시 유사하게 변화한다. 특히 도심공동화현상이 나타나는 중심업무지구의 지역적 특성이 반영되어 있는 것으로 해석할 수 있다.

군집2는 홍대-신촌 지역에서 포착되며 행정구역상으로는 마포구의 서교동, 서강동 및 서대문구 신촌동에 속한다. 지하철역을 기준으로 신촌역, 홍대입구역, 상수역 일대에 위치해있다. 군집의 외연적인 변화추이를 살펴보면 다음과 같다. 03-06시와 09-12시 신촌과 홍대입구 인근에서 인구의 군집지역이 발생한다. 이후 두 군집은 각각 남쪽 및 북쪽으로 확장이 진행되며 18-21시에는 하나로 병합된다. 병합되었던 군집지역은 21-24시를 기점으로 다시 분할 및 축소된다. 한편 군집에 속하는 인구의 변화는 다음과 같다. 03-06시에는 해당시간대의 모든 군집에 속하는 인구 중 16.17%(47.79천명)가 군집2에 속해있지만 이후 12-15시에는 10.54%(143.77천명) 24-03시에는 9.58%(35.29천명)로 점차 감소하는 추세를 보인다. 결국 공간적인 범위는 증가 후 감소하고 있지만, 해당 군집에 분포하는 인구의 비중은 감소 후 증가하고 있다. 해당 지역은 대표적인 상업집적지 중 하나이며, 분석을 통해 탐색된 시공간 군집 역시 상업집적지에 위치한다. 특히 대로변 중심의 상업지역에서 형성된 군집이 점차 확장되어가고 있는 양상을 보인다. 인구 역시 오전에 비해 야간시간대에 보다 군집되어 있는 것으로 나타났다.

군집 3은 동대문-혜화 지역에 위치해있다. 행정구역상 종로구의 종로5·6가동 및 이화동, 창신1동을 중심으로 신당동(중구) 및 광희동(중구) 등에 걸쳐 군집지역이 형성되어있다, 지하철역을 기준으로 동대문역, 동대문역사공원역, 종로5가역, 혜화역 주변 지역에 해당한다. 군집의 공간적 분포는 다음과 같이 변화한다. 03-06시 혜화역 부근의 이화동, 혜화동과 동대문역, 동대문역사공원역 인근의 창신1동, 신당동, 숭인2동에서 군집지역이 각각 발생한다. 점차 공간적인 확장을 이루며, 12-15시에 하나로 병합이 되어 확대된다. 이후 공간적인 범위는 내부에서의 활발한 변동 이후 축소되는 모습을 갖는다. 그 결과 24-03시에는 다시 혜화역

부근과 동대문-동대문역사공원역 인근의 군집지역의 군집만이 존재한다. 특히 이는 주변지역에 비해 상대적으로 새벽 및 오전시간대의 유동인구가 많은 동대문시장, 평화시장, 동문시장 등이 위치한 동대문역 주변에서 점차 북쪽으로 이동해가는 모습이 포착된다. 반면 군집에 속하는 유동인구의 분포양상은 감소 후 증가하는 형태를 보인다. 03-06시에는 해당 시간대의 군집으로 탐색된 인구 대비 44%(131.08천명)정도가 위치해있으며, 이는 다른 군집지역에 비해 상대적으로 많은 인구에 해당한다. 하지만 이후 점차 감소를 하며, 12-15시에는 6.21%(84.73천명)정도만이 분포한다. 15-18시를 기점으로 다시 군집정도는 증가를 하며, 24-03시에는 28.83%(106.19천명)로 군집들 중 가장 많은 유동인구가 탐색된다. 공간적인 범위는 증가 후 감소를 하고 있지만, 해당 시점에서 해당 군집에 분포하는 인구의 비중은 감소 후 증가하고 있다.

군집 4는 건국대학교 인근에서 형성되어있으며 건대입구역을 기준으로 주변에 위치한 광진구의 성수 2·3가동, 화양동 및 자양동에 걸쳐있다. 시간대에 따른 군집의 공간적인 범위의 변동은 소폭 확장 후 축소된다. 특히 해당 군집은 시간의 흐름에 따라 서쪽 및 북쪽으로 확장되었다 다시 쇠퇴하는 양상을 갖는다. 군집 4는 03-06시에는 총 3.05천명의 유동인구가 탐색이 되었으며, 이는 해당 시점의 총 군집으로 탐색된 인구 중 1.03%에 해당한다. 이후 계속 증가하여 21-24시에는 124.56천명(17.46%)의 인구가 해당 지역에서 탐색이 되지만, 24-03시에는 45.60천명(3.26%)로 축소된다. 용도지역을 통해 군집으로 탐색된 지역을 살펴보면, 군집은 일반상업지역과 제2종일반주거지역 등에 걸쳐 있으며, 점차 ‘건대맛의거리’가 위치한 제2종일반주거지역으로 군집이 확장 및 이동해나감을 확인할 수 있다. ‘건대맛의거리’의 경우 용도지역상으로는 제2종일반주거지역으로 설정되어있으나, 요식업 등 상업시설이 밀집되어있는 지역이다.

군집 5는 수유-쌍문 지역에 위치해있다. 행정구역상으로는 미아동, 수유3동, 번1동, 쌍문3동, 창2동에 해당하고 지하철역을 기준으로는 수유역-미아역에 위치해 있다. 군집의 시공간적 변화는 축소 및 분할 이후 다시 확장 및 병합되는 형태로 나타났다. 03-09시와 21-03시에는 해당 지

역의 전역에 걸쳐 형성되어 있다. 하지만 09-21시에서는 수유역과 미아역 인근에서만 군집이 형성되어 있다. 이때 03-06시에는 군집에 해당하는 유동인구 중 1.03%(21.31천명)의 인구가 해당 지역에 분포한다. 하지만 15-18시에는 10.61%(133.22천명)로 점차 증가한 뒤, 18-21시를 기점으로 다시 감소하여 21-24시에는 7.87%(56.12천명)의 인구가 분포한다.

군집 6은 테헤란로 및 강남대로 지역에 위치한다. 행정구역상으로는 강남구의 삼성1·2동, 대치1·2동, 역삼 1·2동, 논현1·2동 및 서초구의 반포 1동과 서초 2동, 서초 4동에 속해있다. 지하철역으로는 삼성역-선릉역-역삼역-강남역 및 논현역-신논현역-강남역-양재역 등의 인근에 위치해 있다. 03-06시 강남역, 역삼역, 삼성역 인근에서 군집이 탐색이 시작되며, 총 6.32천의 유동인구가 분포한다. 이는 해당 시점의 군집 접체와 비교해보았을 때, 2.14%의 수치에 해당한다. 탐색된 3개의 군집은 공간적으로 점차 크게 확장을 하며 12-15시에 하나로 병합된다. 이때의 유동인구는 354.03천명이며, 이는 해당시점의 군집 전체 대비 28.96%를 차지한다. 하지만 15-18시를 기점으로 점차 그 공간적 범위가 축소 및 분할된다. 24-03시에는 12.01천명(12.38%)만이 해당 지역에 밀집되어 있다. 군집 6 역시 중심업무지구해 해당하며 오전 및 야간 시간대에 비해 오후시간대에 군집의 범위 및 해당지역을 방문하는 인구가 많다. 해당 결과 역시 도심공동화현상이 나타나는 중심업무지구의 지역적 특성이 반영된 것으로 보인다.

군집 7은 잠실-수서 지역에서 탐색되었다. 행정 구역상으로는 잠실 및 수서 지역 전반에 걸쳐 있으나, 구체적인 위치와 그 규모는 지속적으로 변화한다. 03-06시에는 방이2동, 가락동, 수서에 걸쳐 군집이 넓게 분포해있으나, 이후 잠실, 가락동, 수서 세 지역으로 분할이 된다. 또한 잠실 지역에 위치한 군집의 경우 점차 서쪽으로 이동을 한다. 주거지가 밀집해있는 방이2동에서 부도심이자 산업 및 상업시설 등 다양한 기능일 밀집한 잠실로 유동인구의 군집이 이동해가고 있는 것으로 판단된다. 한편 인구의 측면에서 해당 지역을 살펴보면 다음과 같다. 03-06시에는 7.33천명이 군집 7에 위치해있으며, 이는 전체군집의 2.48%에 해당하는 수치

다. 12-15시에는 13.58천명(1.00%), 18-21시에는 33.27천명(3.27%)가 위치하며 이후 감소를 해 24-03시에는 11.98천명(3.25천명)으로 감소한다.

군집 8은 강동구 지역 전반에 걸쳐서 포착이 되며, 그 공간적인 범위는 축소된 후 15-18시를 기점으로 성장한다. 03-09시와 21-03시에는 해당 지역의 전역에 걸쳐 형성되어있으나 12-21시에는 해당 군집들은 축소되며 일부 지역에서만 군집이 형성된다. 각 군집의 인구적 변화 또한 공간적인 범위의 양상과 궤를 같이하여 감소 후 증가하는 추세를 보인다. 03-06시에는 12.93%의 군집이 군집8에 분포해있으나, 12-15시에는 2.02%로, 지속적인 감소를 한다. 이후 다시 점차 증가를 하여 24-03시에는 3.25%의 인구가 해당 지역에 분포한다.

시공간군집분석을 통해 살펴본 결과 총 8개의 군집이 탐색되었으며, 종로구-중구, 홍대-신촌, 동대문-혜화, 건국대학교 인근, 수유-쌍문, 강남대로-테헤란로, 잠실-수서 그리고 강동구에 위치하였다. 각 군집의 시공간적인 변화 양상을 통해 유동인구의 도시 점유 패턴을 파악할 수 있었다. 종로구-중구, 홍대-신촌, 동대문-혜화, 건국대학교 인근, 강남대로-테헤란로에서는 증가 후 감소하는 점유 패턴을 보였으며, 수유-쌍문과 잠실-수서, 그리고 강동구에서는 감소 후 증가하는 형태로 나타났다. 또한 이러한 점유 패턴의 차이는 군집이 위치한 지역의 공간적 특성이 반영된 것으로 보인다.

<표 4-2> 서울시 시공간 점유 패턴의 시간대별 변화 양상

단위: 1000명(%)

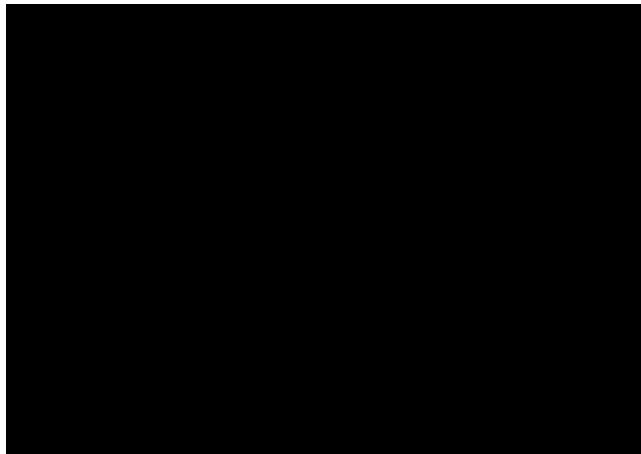
분류	03-06시		06-09시		09-12시		12-15시		15-18시		18-21시		21-24시		24-03시	
군집 1	40.41	13.67	68.11	21.18	312.38	36.21	500.91	36.73	426.18	33.46	376.11	32.68	199.98	28.04	76.87	20.87
군집 2	47.79	16.17	48.71	15.15	97.36	11.29	143.77	10.54	149.82	11.76	130.64	11.35	60.67	8.51	35.29	9.58
군집 3	131.08	44.36	116.18	36.13	79.33	9.20	84.73	6.21	89.45	7.02	103.04	8.95	121.73	17.07	106.19	28.83
군집 4	3.05	1.03	3.38	1.05	16.80	1.95	84.65	6.21	133.22	10.46	188.05	16.34	124.56	17.46	45.60	12.38
군집 5	21.31	7.21	24.09	7.49	85.40	9.90	154.50	11.33	135.17	10.61	81.01	7.04	56.12	7.87	34.49	9.37
군집 6	6.32	2.14	25.46	7.92	232.60	26.96	354.03	25.96	287.90	22.60	201.22	17.48	63.04	8.84	12.01	3.26
군집 7	7.33	2.48	2.30	0.72	7.56	0.88	13.58	1.00	21.23	1.67	37.62	3.27	32.31	4.53	11.98	3.25
군집 8	38.21	12.93	33.36	10.37	31.24	3.62	27.48	2.02	30.65	2.41	33.27	2.89	54.90	7.70	45.87	12.45
합계	295.52	100	321.59	100	862.67	100	1,363.65	100	1,273.62	100	1,150.96	100	713.32	100	368.30	100



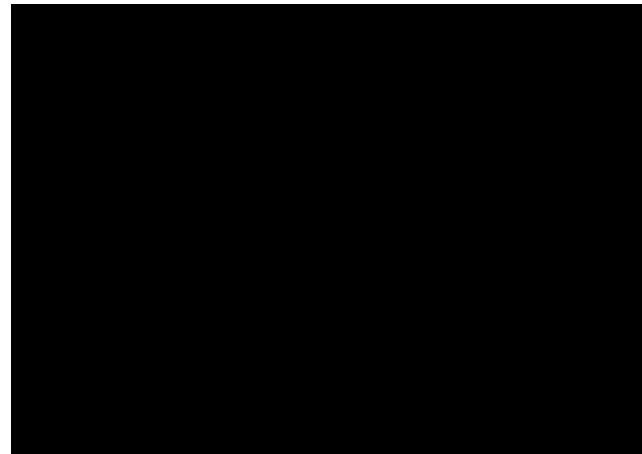
a) 03시-06시 점유 지역



b) 06-09시 점유 지역



c) 09-12시 점유 지역



d) 12-15시 점유 지역

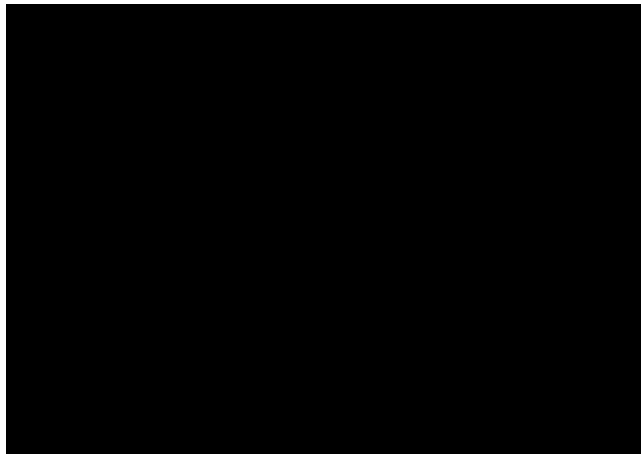
<그림 4-3> 유동인구에 기반한 서울시의 시공간 점유 패턴 1



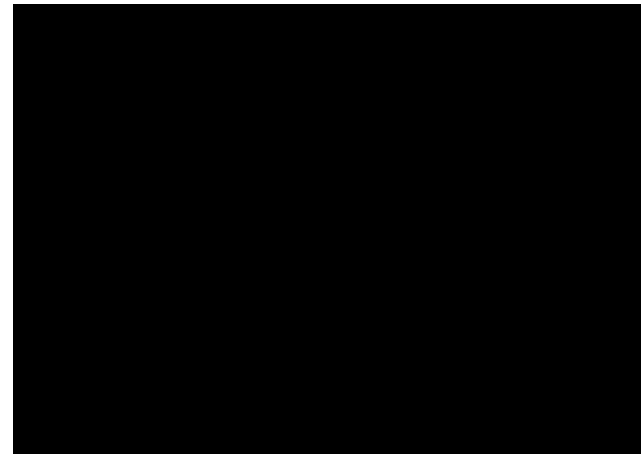
e) 15-18시 점유 지역



f) 18-21시 점유 지역



g) 21-24시 점유 지역



h) 24-03시 점유 지역

<그림 4-4> 유동인구에 기반한 서울시의 시공간 점유 패턴 2

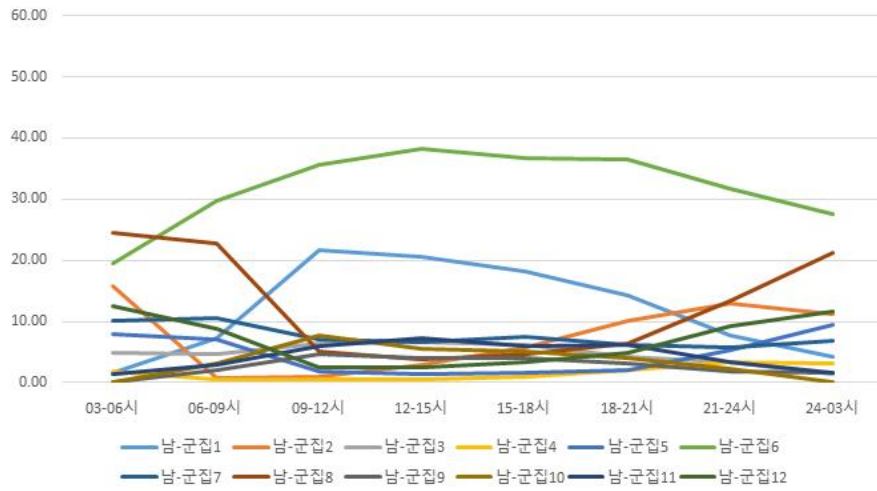
제 2 절 성별에 따른 유동인구의 도시 점유 패턴

1. 서울시 남성 및 여성 유동인구의 시공간군집 탐색 결과

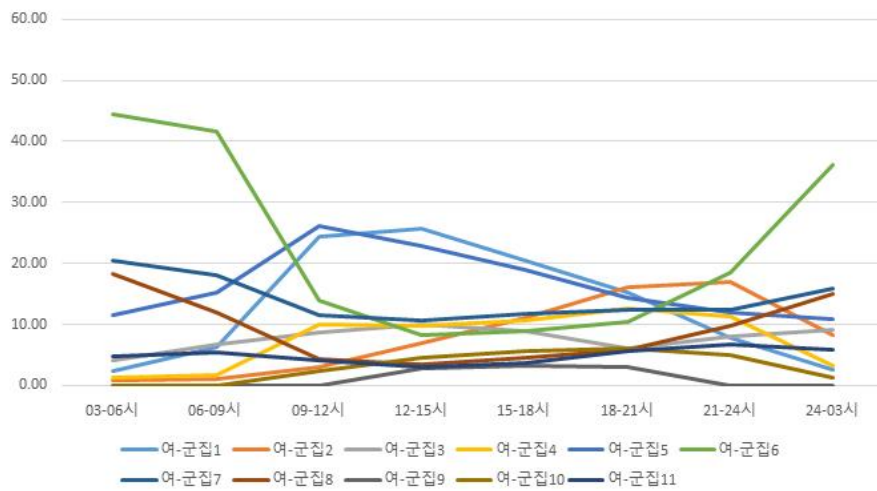
본 절에서는 성별에 따른 유동인구의 시공간군집 분석을 진행하고, 그 결과를 통해 남성 및 여성의 도시 시공간 점유 패턴을 파악하고자 한다. 앞서 언급하였듯이 남성 및 여성에게서 포착된 시공간 군집의 개수는 각 12개와 11개이다. 해당 군집의 구체적인 위치는 표<4-3>와 같다. 한편 탐지된 군집의 총 인구는 남성과 여성 각각 03-06시에는 230.31천명(남성), 146.58천명(여성), 15-18시에는 859.91천명(남성), 730.26천명(여성)이며, 21-24시에는 485.60천명(남성), 357.74천명(여성)이다. 시간대별로 살펴보면 남성과 여성 모두 03-06시에는 강동구(남-군집 8, 여-군집 6)에 가장 많은 인구가 분포한다. 09-12, 강동구에 군집한 인구는 그 규모가 축소되며, 종로구-중구 및 강남대로, 테헤란로에 가장 많은 인구가 분포한다. 반면 21-24시, 남성은 여전히 강남대로-테헤란로에서 가장 많은 수가 위치해 있으나, 여성은 강동구에 가장 많이 분포한다.<그림 4-5>. 즉 해당 군집들에서 남성과 여성의 공간 점유 현상이 나타난다.

<표 4-3> 성별 점유 공간

분류	위치	공간적 변화	분류	위치	공간적 변화
남-군집 1	종로구-중구	확장 후 축소	여-군집 1	종로구-중구	확장 후 축소
남-군집 2	홍대-신촌	확장 후 축소	여-군집 2	홍대-신촌	확장 후 축소
남-군집 3	동대문-혜화	확장 후 축소	여-군집 3	동대문-혜화	확장 후 축소
남-군집 4	건국대학교 인근	확장 후 축소	여-군집 4	강남대로	확장 후 축소
남-군집 5	수유-쌍문	축소 후 확장	여-군집 5	테헤란로	확장 후 축소
남-군집 6	강남대로-테헤란로	확장 후 축소	여-군집 6	강동구	축소 후 확장
남-군집 7	잠실-수서	축소 후 확장	여-군집 7	잠실-수서	축소 후 확장
남-군집 8	강동구	축소 후 확장	여-군집 8	신림동	축소 후 확장
남-군집 9	구로디지털단지	확장 후 축소	여-군집 9	고속터미널	변화 없음
남-군집 10	여의도	확장 후 축소	여-군집 10	영등포동	확장 후 축소
남-군집 11	고속터미널-교대- 남부터미널	확장 후 축소	여-군집 11	사당-이수	확장 후 축소
남-군집 12	신림동	축소 후 확장	.	.	



a) 남성



b) 여성

<그림 4-5> 점유 지역의 남성 및 여성의 인구변화 (전체군집 대비 유동인구 수)

2. 군집별 탐색 결과

탐색 결과 중 총 7개의 지역이 남성 및 여성이 공통적으로 점유하는 지역으로 탐색되었다. 이 지역들에서 탐색된 군집의 시공간적 분포 및 인구의 변화 양상 역시 서로 유사하게 나타났다. 해당 지역들의 구체적인 위치는 한강을 기준으로 북쪽에는 종로구-중구, 동대문-혜화, 홍대-신촌에, 남쪽에는 강남대로-테헤란로, 잠실-수서, 강동구, 그리고 신림동

에 분포한다. <표 4-4>.

<표 4-4> 남성 및 여성의 점유 공간 1(남 여 공통)

분류(남성/여성)		위치	공간적 변화
남-군집 1	여-군집 1	종로구-중구	확장 후 축소
남-군집 2	여-군집 2	홍대-신촌	확장 후 축소
남-군집 3	여-군집 3	동대문-혜화	확장 후 축소
남-군집 6	여-군집 4	강남대로-테헤란로	확장 후 축소
	여-군집 5		
남-군집 7	여-군집 7	잠실-수서	축소 후 확장
남-군집 8	여-군집 6	강동구	축소 후 확장
남-군집 12	여-군집 8	신림동	축소 후 확장

남성 및 여성에서 공통적으로 점유되는 공간, 즉 시공간적 군집의 변화 양상은 대체적으로 1절에서 살펴본 결과와 유사하게 나타난다(종로구-중구, 동대문-혜화, 홍대-신촌, 잠실-수서, 강동구).

강남대로-테헤란로(남-군집 6, 여-군집 4, 여-군집 5)에서 탐지된 군집의 공간적 변화 및 인구의 변화 양상은 모두 확장 후 축소(증가 후 감소)하는 공통적인 형태를 갖지만, 남성과 여성에게서 다소 상이한 패턴의 결과가 포착되었다. 남성의 경우 강남대로와 테헤란로가 하나의 군집으로 형성이 되어있는 반면, 여성의 경우 두 지역이 서로 다른 군집으로 포착된다. 즉 여성에게서는 두 지역은 분리되어있는 군집으로 나타났으며, 시공간적 연결성이 뚜렷하게 드러나지는 않았다<그림 4-6>, <그림

4-7>.

신림동(남-군집12, 여-군집8)은 행정구역상으로는 신림동, 신원동, 서원동, 청룡동, 중앙동, 행운동, 낙성대동에 해당하며, 서울대입구역·봉천역·신림역을 잇는 길을 기준으로 위치해있다. 03-06시를 기준으로 이 길의 주변으로 형성되어 있던 군집은 06-09시와 09-12시 이후 분할이 되어 각 역을 중심으로 군집을 형성되고 이후 다시 성장한다. 인구의 변화 추이 역시 감소 후 증가하는 양상으로 나타난다.

한편 남성 혹은 여성에서만 탐색된 군집 역시 다수 존재하며 해당 군집의 구체적인 위치는 <표 4-5>와 같다.

<표 4-5> 남성 및 여성의 점유 공간 2(남성 혹은 여성)

분류	위치	공간적 변화
남-군집 4	건대입구	확장 후 축소
남-군집 5	수유-쌍문	축소 후 확장
남-군집 9	구로디지털단지	확장 후 축소
남-군집 10	여의도	확장 후 축소
남-군집 11	고속터미널-교대-남부터미널	확장 후 축소
여-군집 9	고속터미널	변화 없음
여-군집 10	영등포동	확장 후 축소
여-군집 11	사당-이수	확장 후 축소

한편 남성이 점유하는 지역은 건대입구(남-군집 4), 그리고 수유-쌍문(남-군집 5)과 구로디지털단지(남-군집 9), 여의도(남-군집 10), 고속터미널-교대-남부터미널(남-군집 11)에 해당한다. 점유 패턴 역시 각 군집별

로 상이하게 나타나고 있다. 먼저, 남-군집 4(건대입구) 및 남-군집 5(수유-쌍문)에서 포착된 군집은 각각 공간적으로 확장 후 축소 혹은 축소 후 확장하는 형태로 나타나고 있으며, 구체적인 변화 양상은 1절에서 탐색된 결과와 유사하다.

남-군집 9는 구로디지털단지(구로구)에 위치해있으며 군집은 행정구역 상으로는 구로구 구로3동에 해당한다. 지하철역으로는 구로디지털단지역 인근에 위치해있으며, 06-09시부터 24-03시까지 탐색된다. 군집의 공간적인 변화는 동쪽으로 확장한 후 축소는 되는 형태로 나타난다. 09-12시 군집의 서쪽지역에서 군집이 형성되기 시작되며, 곧 구로디지털단지 전역으로 확장된다. 이후 18-21시에 이르러 다시 축소가 되며, 21-24시, 24-03시에는 구로디지털단지역 인근의 상업시설 밀집지역에서 군집이 포착된다<그림 4-6>, <그림 4-7>. 인구의 변화를 살펴보면 06-09시에는 1.97%(4.39천명)의 군집이 해당 지역에서 탐색되었으며, 09-12시 이후 그 비율은 4%대를 유지한다. 이후 18-21시(3.12%, 24.45천명)와 21-24시(1.78%, 8.54천명)을 거쳐 24-03시에는 1.52%(3.70천명)로 감소한다<표 4-6>. 변화 양상을 용도지역의 공간적 분포를 중첩하여 분석한 결과, 06-21시까지 탐색되는 시공간 군집은 준공업지역에 위치하며, 21-03시에는 일반상업지역을 중심으로 일부 준공업지역이 함께 군집으로 포착되었다. 남성의 근무비율이 높은 구로디지털단지의 특성상 해당 지역을 방문하는 인구에 의해 점유 패턴이 형성된 것으로 보인다. 특히 업무시간인 09-21시 사이에는 전역에 걸쳐 군집이 형성되고 있으나 21시 이후에는 ‘깎깎거리’라고 불리는 인근의 상업시설로 남성의 군집이 이동해간다.

영등포구 여의도동(남-군집 10)에서도 남성인구의 군집이 포착된다. 이 지역에서 포착되는 군집은 06-09시부터 21-24시 기간 동안 탐색이 된다. 06-09시 여의도의 동쪽에서 군집이 형성되어, 점차 서쪽으로 이동하며 확장하지만 18-21시 이후 축소된다<그림 4-6>, <그림 4-7>. 인구의 측면에서 살펴보면, 06-09시에는 6.95천명의 유동인구가 군집해있으며 이는 전체 군집 대비 3.12%에 해당한다<표 4-6>. 이후 군집 정도의 증가 및 변화를 거쳐 축소된 후 24-03시에는 소멸된다.

고속터미널-교대-남부터미널 지역에 위치한 군집은 행정구역상으로는 서초1·3·4동과 반포 1·4동에 해당하며, 고속터미널역-교대역-남부터미널역을 잇는 대로를 따라 위치한다. 03-06시와 06-09시에는 각각 고속터미널과 남부터미널 인근에서 군집이 형성된다. 두 군집들은 각각 북쪽, 남쪽의 방향으로 공간적인 확장을 이루며, 12-15시에는 하나로 병합이 된다<그림 4-6>, <그림 4-7>. 이때 군집에는 52.79천명이 분포하며, 전체 군집에 해당하는 유동인구 대비 7.37%에 해당한다. 이후 공간적인 분포는 축소 및 분리가 되며 그에 따른 인구의 분포 비율 역시 감소한다<표 4-6>.

한편 여성에서만 점유되는 공간은 총 3개이다(여-군집 9, 여-군집 10, 여-군집 11). 각각 고속터미널(여-군집 9), 영등포동(여-군집 10), 사당-이수지역(여-군집 11)에 위치한다. 먼저 고속터미널(여-군집 9)에서 탐색된 군집은 서초구 고속터미널에 위치하며, 12-21시까지 유지된다. 해당 군집은 공간적인 변화가 발생하지 않았다<그림 4-8>, <그림 4-9>. 고속터미널을 이용하기 위해 방문한 인구에 의해 해당 지역이 점유된 것으로 볼 수 있다.

영등포동에 위치한 여-군집 10은 09-12시에 탐지되기 시작하며 24-03시까지 지속된다. 영등포동의 타임스퀘어와 롯데백화점을 중심으로 군집이 형성되기 시작하며 인근의 상업지역들로 확장된다. 이후 24-03시를 기점으로 군집은 다시 축소된다<그림 4-8>, <그림 4-9>. 전체 군집 대비 해당 군집의 유동인구 역시 증가 후 감소하는 변화를 보인다. 구체적으로 살펴보면 12-15시에는 9.55천명이 군집으로 포착이 되며, 이는 전체 군집 중 2.92%정도이다. 18-21시에는 40.86천명(26.11%)으로 증가지만 24-03시에는 다시 25.65천명(4.93%)로 감소한다<표 4-6>.

한편 사당-이수 지역(여-군집 11) 역시 여성에 의해 점유되는 지역으로 나타났으며, 행정구역상 동작구 사당1·2·동과 서초구 방배2·4동에 해당한다. 이 지역은 사당역과 이수역을 중심으로 군집이 형성되며 점차 남북방면으로 확장된 후 다시 축소되는 형태를 보인다. 한편 해당 지역에서의 인구 변화 역시 증가 후 감소되는 형태로 나타난다<표 4-7>.

이상으로 성별에 따른 도시 점유 패턴을 살펴보았다. 분석 결과 종로구-중구 홍대-신촌, 동대문-혜화, 강남대로-테헤란로, 잠실-수서, 강동구, 신림동은 남성 및 여성 모두에게서 점유되는 지역으로 나타났으며, 그 공간적 변화 및 인구 변화 양상은 1절에서의 결과와 유사하게 나타났다. 한편 건대입구(남-군집 4), 수유-쌍문(남-군집 5), 구로디지털단지(남-군집 9), 여의도(남-군집 10) 및 고속터미널-교대-남부터미널(남-군집 11)은 남성에게서만 점유되는 지역으로, 고속터미널(여-군집 9), 영등포동(여-군집 10), 사당-이수(여-군집 11) 지역은 여성에서만 점유되는 지역으로 탐색되었다. 이처럼 성별에 따라 점유되는 공간과, 그 점유 패턴이 다소 상이하게 나타났다.

이건학 외(2013)에 의하면 젠더에 따라 점유되고 사용되는 도시공간은 서로 다르게 나타난다. 그는 토지이용을 통해 남성 및 여성에게서 상이하게 소비되는 도시공간을 실증적으로 분류하였다. 연구에 의하면 중구와 종로구 같은 전통적 도심과 강남-서초와 같은 부도심(현재는 도심), 그리고 신촌-홍대 지역, 강동구 일대는 중성적 공간으로 탐지가 되었다. 한편 영등포구 여의도동은 타임스퀘어와 영등포 시장을 중심으로 여성에 의해 점유되고 있으며, 업무나 비즈니스의 중심기능을 수행하는 여의도와 첨단산업단지가 입지해있는 구로구지역은 남성의 공간으로 점유되고 있다고 밝혔다. 이러한 연구 결과는 본 절에서 탐지한 결과와 유사하게 나타나고 있음을 알 수 있다. 즉 성별에 따라 유동인구가 분포하는 지역, 즉 점유되는 공간이 다소 차이가 나고 있다.

<표 4-6> 서울시 시공간 점유 패턴의 시간대별 변화 양상(남성)

단위: 1000명(%)

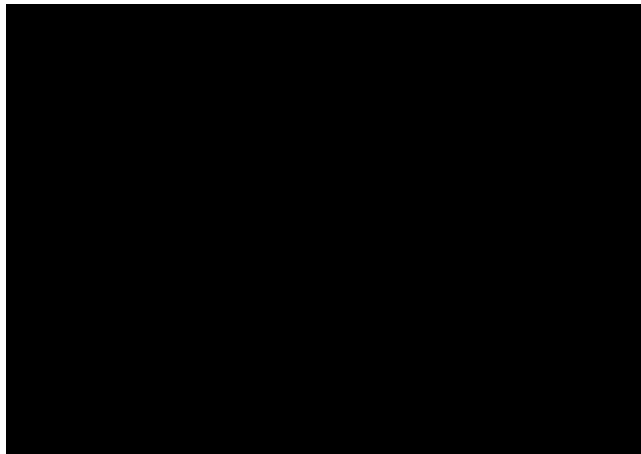
분류	03-06시		06-09시		09-12시		12-15시		15-18시		18-21시		21-24시		24-03시	
군집 1	3.51	1.53	16.43	7.37	143.66	21.63	196.62	20.50	155.69	18.11	112.68	14.36	37.06	7.71	10.03	4.13
군집 2	36.18	15.71	1.64	0.74	6.93	1.04	27.30	2.85	48.98	5.70	78.55	10.01	62.24	12.95	27.31	11.25
군집 3	11.01	4.78	10.43	4.68	42.74	6.43	63.36	6.61	52.37	6.09	33.04	4.21	16.19	3.37	3.63	1.50
군집 4	4.10	1.78	1.14	0.51	3.30	0.50	5.54	0.58	8.82	1.03	16.73	2.13	15.96	3.32	7.77	3.20
군집 5	18.34	7.97	15.89	7.13	12.59	1.90	12.39	1.29	14.09	1.64	15.64	1.99	26.06	5.42	22.79	9.39
군집 6	45.04	19.55	66.38	29.77	236.29	35.58	367.01	38.27	315.48	36.69	287.06	36.59	152.28	31.68	66.72	27.49
군집 7	23.48	10.19	23.32	10.46	46.41	6.99	63.10	6.58	64.07	7.45	47.83	6.10	27.52	5.72	16.79	6.92
군집 8	56.56	24.56	50.57	22.68	33.52	5.05	35.72	3.72	39.44	4.59	50.83	6.48	64.27	13.37	51.61	21.26
군집 9	0.00	0.00	4.40	1.97	31.02	4.67	39.06	4.07	35.16	4.09	24.46	3.12	8.54	1.78	3.70	1.52
군집 10	0.00	0.00	6.95	3.12	51.64	7.77	53.76	5.61	44.14	5.13	31.14	3.97	10.44	2.17	0.00	0.00
군집 11	3.35	1.45	6.36	2.85	40.19	6.05	70.72	7.37	52.29	6.08	48.45	6.17	15.68	3.26	3.88	1.60
군집 12	28.74	12.48	19.46	8.73	15.90	2.39	24.34	2.54	29.39	3.42	38.22	4.87	44.41	9.24	28.51	11.74
합계	230.31	100	222.98	100	664.20	100	958.92	100	859.91	100	784.62	100	480.65	100	242.75	100



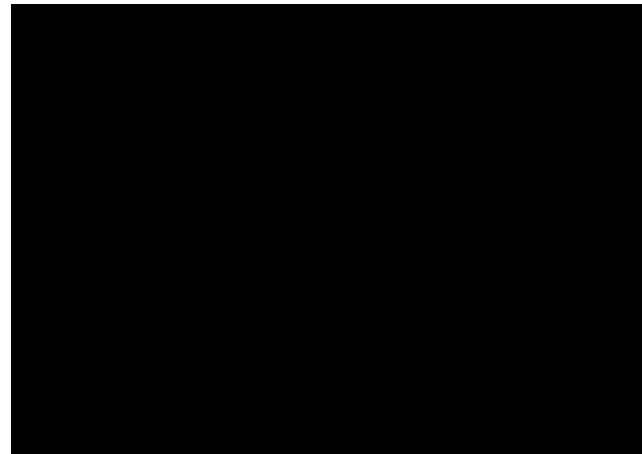
a) 03시-06시 점유 지역



b) 06-09시 점유 지역



c) 09-12시 점유 지역



d) 12-15시 점유 지역

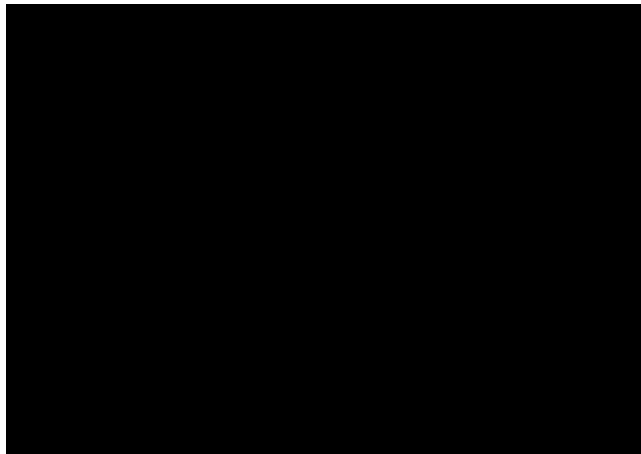
<그림 4-6> 남성 유동인구에 기반한 서울시의 시공간 점유 패턴 1



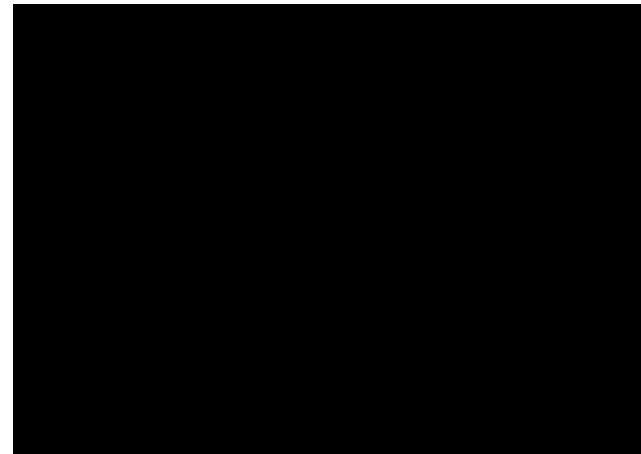
a) 15-18시 점유 지역



b) 18-21시 점유 지역



c) 21-24시 점유 지역



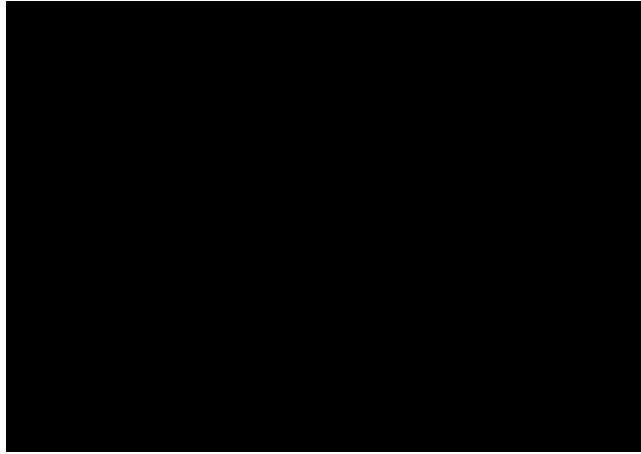
d) 24-03시 점유 지역

<그림 4-7> 남성 유동인구에 기반한 서울시의 시공간 점유 패턴 2

<표 4-7> 서울시 시공간 점유 패턴의 시간대별 변화 양상(여성)

단위: 1000명(%)

분류	03-06시		06-09시		09-12시		12-15시		15-18시		18-21시		21-24시		24-03시	
군집 1	3.44	2.34	9.95	6.41	101.22	24.41	195.08	25.80	149.98	20.54	102.58	15.35	28.12	7.86	4.58	2.67
군집 2	1.39	0.95	1.60	1.03	12.20	2.94	52.71	6.97	80.79	11.06	107.40	16.07	60.70	16.97	14.22	8.28
군집 3	6.11	4.17	10.49	6.76	36.32	8.76	76.48	10.11	65.91	9.03	41.02	6.14	28.76	8.04	15.71	9.15
군집 4	1.81	1.24	2.71	1.74	41.57	10.02	74.47	9.85	78.66	10.77	85.11	12.73	40.13	11.22	5.68	3.31
군집 5	17.06	11.64	23.77	15.31	108.22	26.09	173.13	22.90	138.31	18.94	96.38	14.42	42.72	11.94	18.51	10.78
군집 6	65.11	44.42	64.70	41.68	57.94	13.97	62.69	8.29	65.19	8.93	70.02	10.48	66.07	18.47	62.02	36.13
군집 7	29.97	20.45	27.98	18.03	47.48	11.45	79.84	10.56	86.38	11.83	82.93	12.41	44.20	12.36	27.30	15.90
군집 8	26.78	18.27	18.68	12.03	17.71	4.27	26.52	3.51	32.73	4.48	39.33	5.88	35.35	9.88	25.66	14.94
군집 9	0.00	0.00	0.00	0.00	0.00	0.00	22.06	2.92	24.52	3.36	20.02	2.99	0.00	0.00	0.00	0.00
군집 10	0.00	0.00	0.00	0.00	9.56	2.30	34.02	4.50	41.53	5.69	40.86	6.11	17.63	4.93	2.35	1.37
군집 11	7.12	4.86	8.27	5.33	17.09	4.12	22.18	2.93	27.11	3.71	38.47	5.76	23.87	6.67	9.97	5.81
합계	146.58	100	155.2 2	100	414.74	100	756.18	100	730.26	100	668.42	100	357.74	100	171.69	100



a) 03시-06시 점유 지역



b) 06-09시 점유 지역



c) 09-12시 점유 지역



d) 12-15시 점유 지역

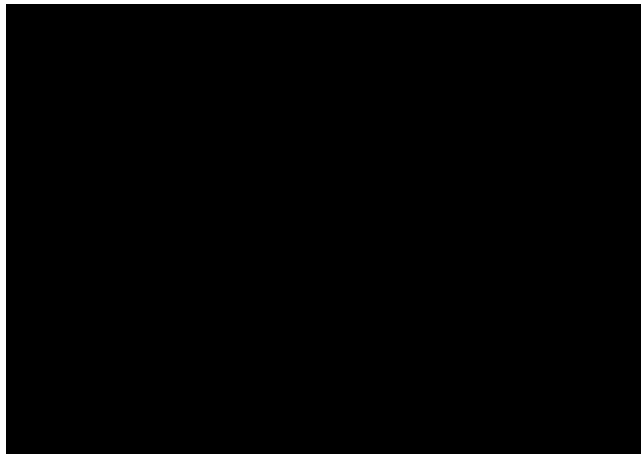
<그림 4-8> 여성 유동인구에 기반한 서울시의 시공간 점유 패턴 1



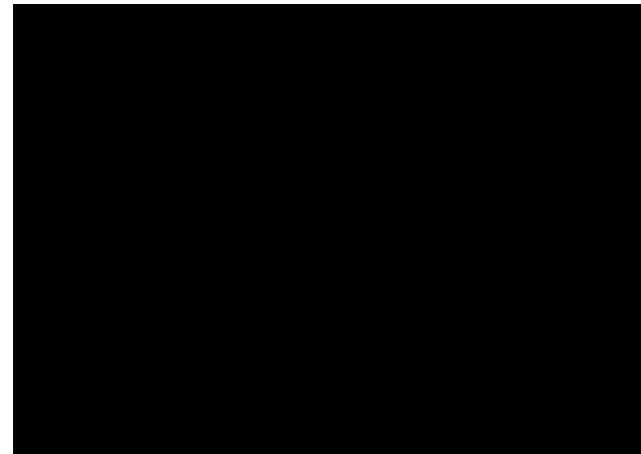
a) 15-18시 점유 지역



b) 18-21시 점유 지역



c) 21-24시 점유 지역



d) 24-03시 점유 지역

<그림 4-9> 여성 유동인구에 기반한 서울시의 시공간 점유 패턴 2

제 3 절 연령에 따른 유동인구의 도시 점유 패턴

1. 서울시 연령별 유동인구의 시공간군집 탐색 결과

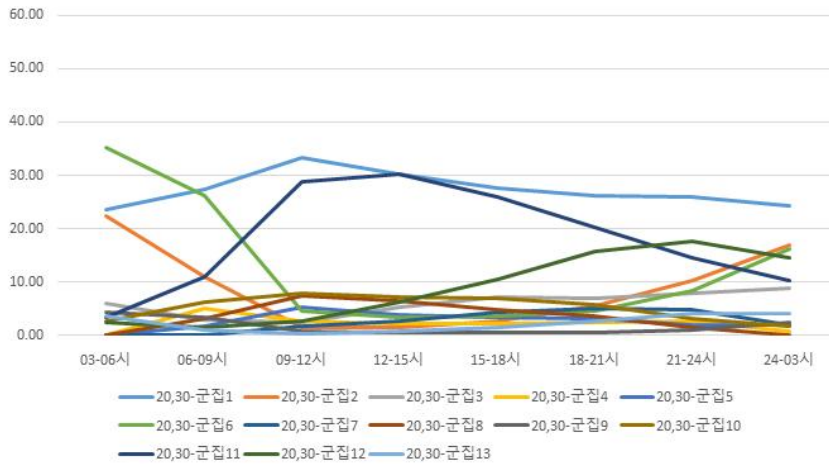
연령별 유동인구의 시공간 점유 패턴 탐색 결과는 <표 4-8>과 같다. 20,30대에서는 총 13개이며, 40,50대에서는 11개, 그리고 60,70대에서는 총 8개의 지역이 점유되는 지역으로 탐색되었다. 각 연령에 따라 점유되는 지역 역시 다소 차이를 보였다.

한편 점유 지역에서 시간대별로 탐지된 총 인구는 각 연령대별로 03-06시에는 110.65천명(20,30대), 146.74천명(40,50대), 33.53천명(60,70대), 15-18시에는 711.45천명(20,30대), 546.23천명(40,50대), 165.4천명(60,70대)이며, 21-24시에는 417.05천명(20,30대) 282.65천명(40,50대) 48.00천명(60,70대)이다. 60,70대에 비해 20,30대와 40,50대의 유동인구 수가 많이 탐지 되었다. 또한 03-12시까지 30,40대의 유동인구 수가 가장 높게 나타나며, 12-03시까지 20,30대의 유동인구가 가장 높게 나타난다.

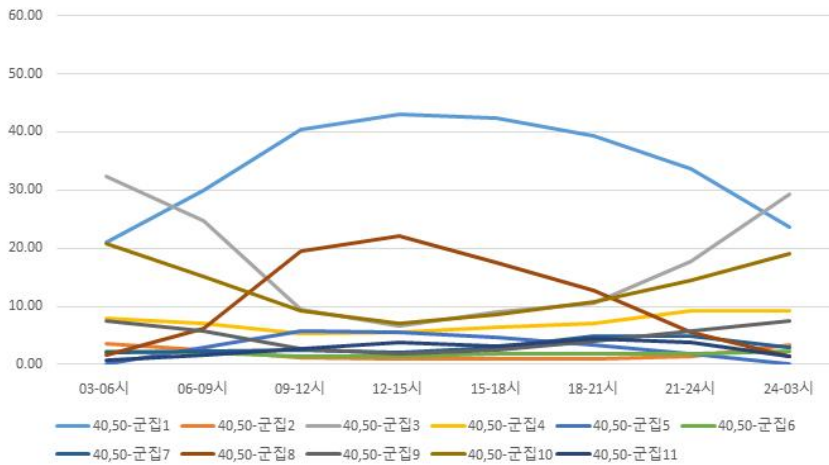
시간대별로 살펴보면 모든 연령대에서 03-06시에는 강동구와 강남대로-테헤란로에서 가장 많은 유동인구에 의해 점유되고 있는 지역으로 탐지되며, 그 다음으로 20,30대에서는 신림동이, 40,50대에서는 노원-중계-공릉지역이, 60,70대에서는 쌍문-수유이 높은 점유 형태를 보였다. 09-12시에서는 강남대로-테헤란로에서 가장 많은 인구 군집이 분포하고, 강동구에 군집한 인구는 그 규모가 축소되고 있다. 반면 21-24시에서는 다시 20,30대에서는 강남대로-테헤란로에, 40,50대와 60,70대에서는 강동구에 가장 많은 인구가 군집한다<그림 4-9>. 공간적 변화 양상 및 인구의 변화는 <그림 4-11> ~ <그림 4-16>과 <표 4-9> ~ <표 4-11>에 자세하게 나타나있다,

<표 4-8> 연령별 점유 공간

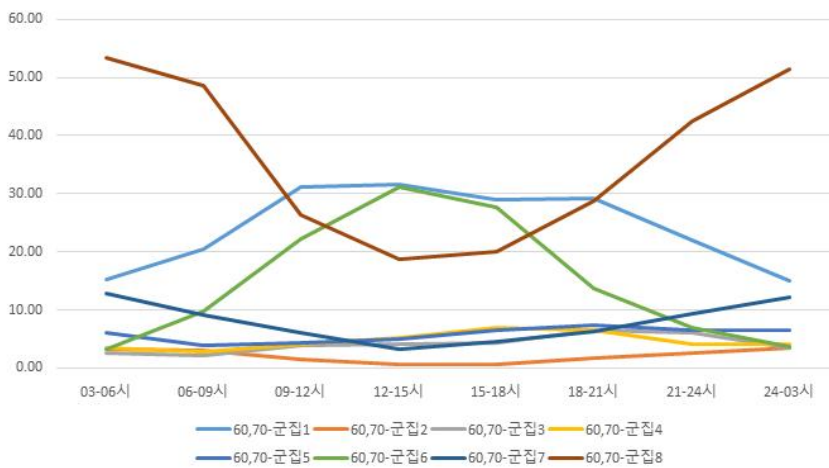
분류	위치	공간적 변화	분류	위치	공간적 변화	분류	위치	공간적 변화
2030-군집 1	강남대로-테헤란로	확장 후 축소	4050-군집 1	강남대로-테헤란로+고속터미널-교대-남부터미널	확장 후 축소	6070-군집 1	잠실-강동구	축소 후 확장
2030-군집 2	신림동	축소 후 확장	4050-군집 2	가락동	축소 후 확장	6070-군집 2	강남대로-테헤란로+고속터미널-교대-남부터미널	확장 후 축소
2030-군집 3	잠실	축소 후 확장	4050-군집 3	강동구	축소 후 확장	6070-군집 3	가락동	축소 후 확장
2030-군집 4	노량진	확장 후 축소	4050-군집 4	잠실	축소 후 확장	6070-군집 4	사당-이수	확장 후 축소
2030-군집 5	구로디지털단지	확장 후 축소	4050-군집 5	여의도	확장 후 축소	6070-군집 5	영등포동	확장 후 축소
2030-군집 6	강동구	축소 후 확장	4050-군집 6	목동	변화 없음	6070-군집 6	신림동	확장 후 축소
2030-군집 7	영등포동	확장 후 축소	4050-군집 7	사당-이수	확장 후 축소	6070-군집 7	종로구-중구	확장 후 축소
2030-군집 8	여의도	확장 후 축소	4050-군집 8	종로구-중구	확장 후 축소	6070-군집 8	수유-쌍문	축소 후 확장
2030-군집 9	가락동	축소 후 확장	4050-군집 9	수유-쌍문	축소 후 확장	.	.	.
2030-군집 10	고속터미널-교대-남부터미널	확장 후 축소	4050-군집 10	노원구-중랑구	축소 후 확장	.	.	.
2030-군집 11	종로구-중구	확장 후 축소	4050-군집 11	공덕-신촌	확장 후 축소	.	.	.
2030-군집 12	홍대-신촌	확장 후 축소
2030-군집 13	건국대학교 인근	확장 후 축소



a) 20,30대



b) 40,50대



c) 60,70대

<그림 4-10> 점유 지역의 연령별 인구변화 (전체군집 대비 유동인구 수)

2. 군집별 탐색 결과

종로구-중구, 강남대로-테헤란로와 고속터미널-교대-남부터미널, 잠실과 강동구, 가락동은 모든 연령대에서 공통적으로 점유되고 있는 지역이다. 하지만 해당 위치, 즉 시공간군집이 형성된 위치 및 그 공간적 변화에 있어서는 일부 다르게 나타났다.

종로구-중구지역(20,30-군집 11, 40,50-군집 8, 60,70-군집 7)은 모든 연령대에서 공간적으로 확대된 후 축소되는 형태를 보이며, 그에 따른 유동인구의 변화 역시 증가 후 감소한다. 하지만 군집의 범위는 연령에 따라 다소 상이하게 나타난다. 20,30대 및 60,70대에서 점유되는 지역은 40,50대에 비해 비교적 넓게 탐색이 된다. 종로구-중구로부터 혜화-동대문 지역까지 걸쳐 탐색되고 있으며, 이는 특히 09-12시 이후 현저하게 나타난다. 하지만 40,50대에게 점유되는 지역은 종로구-중구 지역의 CBD를 중심으로 그 범위가 넓게 나타나지는 않는다.

강남대로-테헤란로(2030-군집 1, 40,50-군집 1, 60,70-군집 2)와 고속터미널-교대-남부터미널 지역(2030-군집 10, 40,50-군집 1, 60,70-군집 2)은 20,30대에서는 독립된 두 개의 군집이 점유 되는 것으로 탐지되나, 40,50대와 60,70대에게서는 해당 지역이 하나의 점유 공간으로 탐지되었다. 또한 잠실과 강동구 지역 역시 60,70대에서는 다른 연령대와는 달리 하나의 점유 지역으로 포착되고 있다

가락동(20,30-군집 9, 40,50-군집 2, 60,70-군집 3)에서 탐색되는 군집은 공간적으로 축소 후 확대되는 모습을 갖는다. 03-06시에는 1.85천명으로 전체 군집의 4.26%의 인구가 지역에서 포착되지만 15-18시에는 1.63천명(0.43%)이 분포한다. 이후 다시 21-24시에는 15.06천명이 탐지되며 이는 전체 군집의 8.35%를 차지한다.

두 연령대에서 공통적으로 점유되는 지역은 다음과 같다. 먼저 20,30대와 40,50대는 여의도를 공통적으로 점유한다. 시공간 군집 탐색 결과 20,30대(20,30-군집 8)에서는 09-12시에, 40,50대(40,50-군집 5)에서는 06-09시에 군집이 형성되며 모두 21-24시까지 군집이 지속된 후 소멸한

다. 두 연령대에서 모두 군집은 확장 후 축소되는 형태로 나타나며, 군집에 해당하는 인구 역시 유사한 형태로 변화한다.

40,50대와 60,70에서 공통적으로 점유되는 지역은 사당-이수와 수유-쌍문 지역에 위치한다. 시공간군집 분석 결과, 해당 군집들의 공간적 분포 및 변화 양상은 앞선 절에서 기술한 결과와 유사하다. 사당-이수지역(40,50-군집 7, 60,70-군집 4)에 위치한 군집은 확장 후 축소(증가 후 축소)되는 형태로 나타나고 있다. 반면 수유-쌍문 지역(40,50-군집 9, 60,70-군집 8)에 위치한 군집의 경우 공간적인 변화와 인구의 변화추이는 축소 후 확장(감소 후 증가)하는 형태로 나타나고 있다.

20,30대와 60,70대는 영등포동을 공통적으로 점유하고 있다. 분석결과, 20,30대(20,30-군집 7)의 경우 영등포역 인근에 위치한 영등포동에서, 탐 반면 60,70대(60,70-군집 5)에서는 영등포동과 당산동에서 시공간 군집이 탐색된다. 시간적인 범위 또한 20,30대에서는 09-12시 이후로 유동인구가 군집하지만, 60,70대에서는 모든 시점에서 군집으로 탐지된다. 즉 영등포동이라는 같은 지역을 점유하고 있지만, 그 공간적인 범위나 시간적인 범위에 있어서 점유 패턴이 다소 상이하게 나타나고 있다. 점유 지역 내 유동인구의 군집정도 역시 증가 후 감소하는 형태로 나타난다.

반면 신림동에서는 20,30대와 60,70대의 점유 패턴이 다소 다르게 나타난다. 20,30대(20,30-군집 2)의 경우 공간적인 분포 및 전체 군집 대비 해당 군집의 유동인구 비율이 축소 후 확장되는 형태로 나타나지만, 60,70대(60,70-군집 6)의 경우 그 반대의 변화를 보인다. 즉 해당 공간은 낮과 오후에는 주로 60,70대에 의해 점유되는 반면 오전과 새벽시간대에는 20,30대에 의해 주로 점유된다. 신림동은 학원가와 고시원이 밀집해있는 곳으로, 20,30대에서는 학원가를 방문하는 인구에 의해 군집이 형성된 것으로 보인다. 동시에 해당 지역은 노년인구가 관악노인종합복지관, 동명노인복지센터 등이 인근에 위치하므로, 60,70대에서는 노인복지시설을 방문하는 인구에 의해 해당 지역이 점유되는 것으로 추측할 수 있다.

한편 특정 연령대에서만 점유되는 지역도 다수 존재한다. 시공간군집 분석 결과, 구로디지털단지, 홍대-신촌, 건국대학교 인근, 그리고 노량진

은 20,30대에서만 점유된다. 구로디지털단지(20,30-군집 5), 건국대학교 인근(20,30-군집 13), 홍대-신촌(20,30-군집 12)은 앞선 절에서 기술한 결과와 유사한 시공간적 분포와 그 편화패턴을 갖는다. 구로디지털단지는 산업단지를 중심으로 군집이 형성되나, 야간시간대에는 상업시설이 밀집한 지역으로 군집이 이동하고 있으며, 홍대-신촌과 건국대학교 인근 지역에서는 상업시설이 밀집된 지역을 방문하는 인구에 의해 군집이 확장 후 축소되는 것으로 보인다. 한편 노량진(20,30-군집 4)은 학원가가 위치한 대표적인 지역이다. 학원가를 방문하는 인구에 의해 군집지역이 형성된 것으로 보인다.

반면 목동, 공덕-신촌, 노원구-중랑구 지역은 40,50대에서만 점유되는 지역에 해당한다. 40,50-군집 6은 양천구 목동의 오목교역을 중심으로 남북쪽으로 형성되어 있으며, 공간적인 변화는 크게 나타나지 않는다. 각 시점별 인구 역시 06-09시에는 12.74천명(1.99%), 15-18시 84.48천명(1.83%), 24-03시에는 12.05천명(2.16%)으로 큰 변화가 나타나지 않는다.

노원구-중랑구 지역에 위치한 군집(40,50-군집 6)은 노원구의 상·중·하계동 및 공릉동과 중랑구의 목동, 중화동에 걸쳐 남북방면으로 걸쳐있다. 해당 군집은 상계역-공릉역-중화역까지 7호선의 지하철역을 중심으로 형성된다. 군집의 공간적 변화와 인구 변화 모두 축소 후 확장(감소 후 증가)되는 형태로 나타난다.

마지막으로 공덕-신촌 지역(40,50-군집 11) 역시 40,50대에서 점유되고 있다. 분석 결과 03-06시 공덕역을 중심으로 군집이 형성되기 시작하여 신촌지역으로 확장한 뒤 다시 축소되는 변화를 보인다. 이때, 군집에 속해있는 인구는 시간에 따라 03-06시에는 3.61천명(0.73%) 12-15시 82.11천명(3.90%)을 거쳐 24-03시에는 7.97천명(1.43%)이며, 증가 후 감소하는 것으로 나타났다.

이상으로 연령대에 따른 도시의 시공간적 점유 패턴을 살펴보았다. 시공간군집 분석 결과 종로구-중구과 강남대로-테헤란로, 고속터미널-교대-남부터미널 지역, 잠실, 가락동 및 강동구 지역은 모든 연령에게서 공통적으로 점유되는 지역으로 탐색되었다. 또한 20,30대와 40,50대는 여의

도에서, 40,50대와 60,70대는 수유-쌍문 그리고 사당-이수 지역에서, 마지막으로 20,30대와 60,70대는 신림동에서 공통적인 점유 패턴이 나타났다. 한편 홍대-신촌, 신림, 노량진, 구로디지털단지, 건대입구 지역은 20,30대에서 그리고 목동, 공덕-신촌, 노원-중랑 지역은 30,40대에서만 점유되는 지역으로 탐색되었다. 이처럼 유동인구는 연령대에 따라서도 시공간적인 분포가 다르며, 그 결과 도시의 시공간 점유 패턴 역시 다르게 나타나고 있었다.

또한 일부 지역은 특정한 시간대에서만 군집이 위치하는 것으로 탐지되었다. 노량진(20,30대)과 구로디지털단지(20,30대)는 06-09시에서 24-03시까지, 영등포동(20,30대)은 09-12시에서 24-03시에, 여의도(20,30대 및 40,50대)는 06-09시에서 21-24시까지만 유동인구의 군집이 이루어졌다. 특정 시간대에서만 해당 공간이 각 연령대의 인구에 의해 점유되고 있는 것이다.

<표 4-9> 서울시 시공간 점유 패턴의 시간대별 변화 양상(20,30대)

단위: 1000명(%)

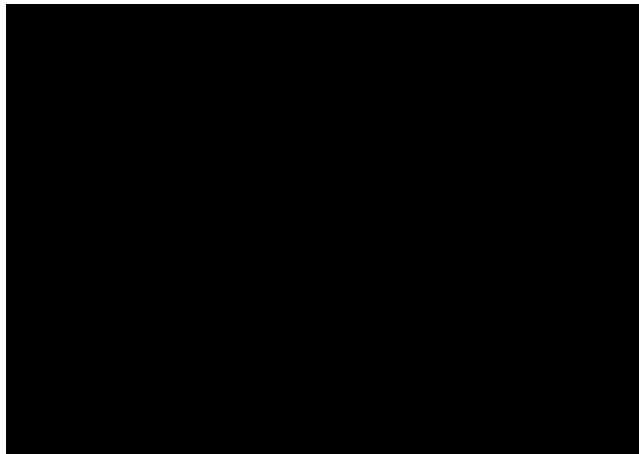
분류	03-06시		06-09시		09-12시		12-15시		15-18시		18-21시		21-24시		24-03시	
1	26.09	23.57	33.71	27.29	144.22	33.44	227.24	30.29	195.87	27.53	189.82	26.18	108.30	25.97	47.57	24.22
2	24.83	22.44	13.67	11.07	7.08	1.64	10.62	1.42	19.15	2.69	40.50	5.59	43.45	10.42	33.49	17.05
3	6.70	6.05	3.65	2.96	10.94	2.54	40.54	5.40	51.65	7.26	50.09	6.91	32.65	7.83	17.55	8.94
4	0.00	0.00	6.31	5.11	12.17	2.82	16.20	2.16	15.15	2.13	18.56	2.56	11.78	2.83	1.45	0.74
5	0.00	0.00	2.04	1.65	23.27	5.40	28.93	3.86	24.66	3.47	23.69	3.27	8.24	1.98	3.74	1.90
6	38.93	35.18	32.40	26.23	20.05	4.65	24.93	3.32	25.37	3.57	33.20	4.58	34.58	8.29	31.84	16.21
7	0.00	0.00	0.00	0.00	7.38	1.71	19.54	2.60	31.01	4.36	36.37	5.02	20.30	4.87	3.84	1.96
8	0.00	0.00	3.84	3.11	31.79	7.37	49.46	6.59	34.78	4.89	25.83	3.56	6.46	1.55	0.00	0.00
9	4.71	4.26	4.37	3.54	3.94	0.91	4.83	0.64	4.09	0.58	4.43	0.61	3.91	0.94	4.77	2.43
10	2.91	2.63	7.85	6.36	34.67	8.04	53.72	7.16	49.74	6.99	42.28	5.83	12.92	3.10	3.34	1.70
11	3.68	3.33	13.71	11.10	124.43	28.85	227.26	30.29	185.12	26.02	146.15	20.16	60.59	14.53	20.08	10.22
12	2.80	2.53	1.95	1.58	11.37	2.64	47.01	6.27	74.85	10.52	114.00	15.73	73.86	17.71	28.71	14.62
13	4.15	3.75	1.24	1.01	1.86	0.43	6.28	0.84	10.62	1.49	19.05	2.63	17.42	4.18	8.29	4.22
합계	110.65	100	123.50	100	431.30	100	750.29	100	711.45	100	724.93	100	417.05	100	196.38	100



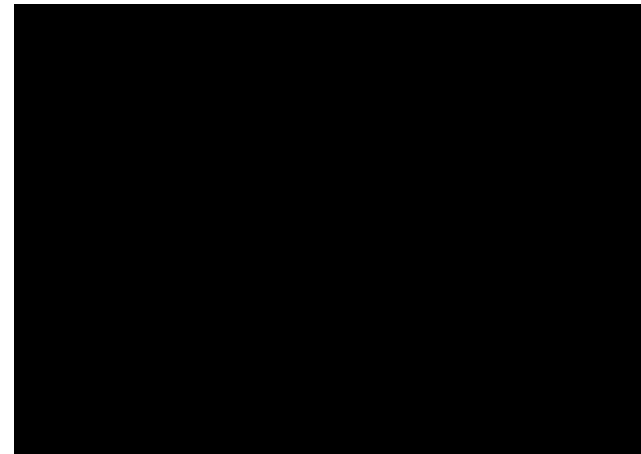
a) 03시-06시 점유 지역



b) 06-09시 점유 지역



c) 09-12시 점유 지역



d) 12-15시 점유 지역

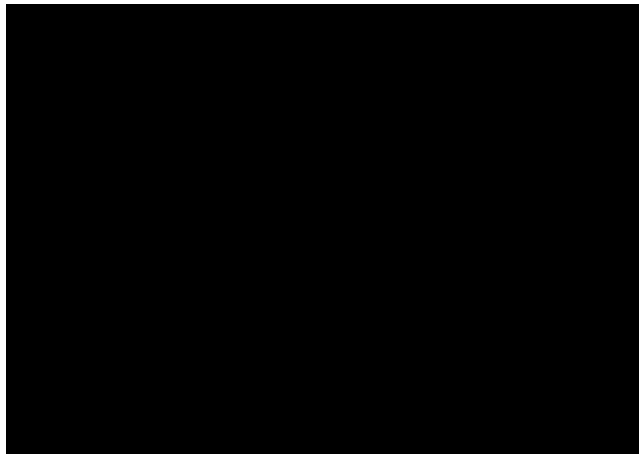
<그림 4-11> 20,30대 유동인구에 기반한 서울시의 시공간 점유 패턴 1



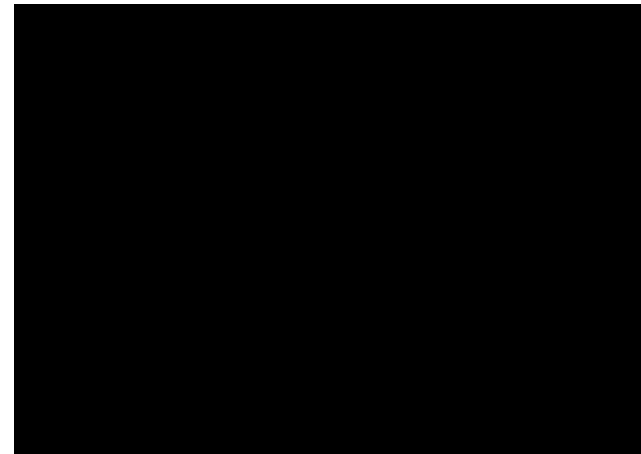
a) 15-18시 점유 지역



b) 18-21시 점유 지역



c) 21-24시 점유 지역



d) 24-03시 점유 지역

<그림 4-12> 20,30대 유동인구에 기반한 서울시의 시공간 점유 패턴 2

<표 4-10> 서울시 시공간 점유 패턴의 시간대별 변화 양상(40,50대)

단위: 1000명(%)

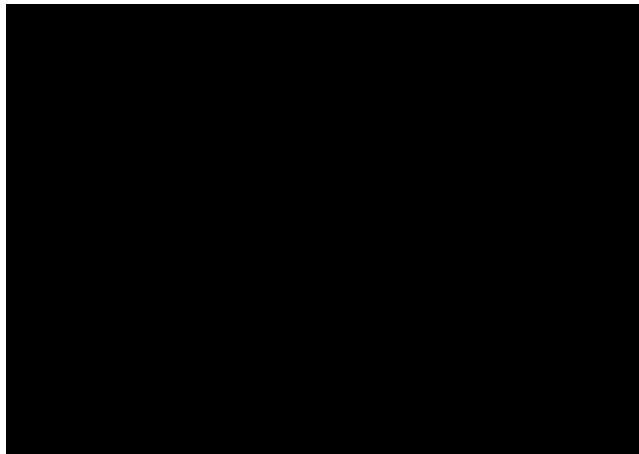
분류	03-06시		06-09시		09-12시		12-15시		15-18시		18-21시		21-24시		24-03시	
1	30.93	21.08	57.49	29.93	180.76	40.32	271.72	43.04	231.49	42.38	191.23	39.36	95.00	33.61	39.47	23.67
2	5.39	3.67	4.71	2.45	4.96	1.11	5.70	0.90	5.15	0.94	4.81	0.99	3.67	1.30	5.49	3.29
3	47.37	32.28	47.67	24.82	42.35	9.45	41.98	6.65	49.28	9.02	50.93	10.48	50.40	17.83	48.73	29.22
4	11.57	7.88	13.49	7.02	23.74	5.30	34.64	5.49	34.65	6.34	34.81	7.17	25.96	9.18	15.51	9.30
5	0.00	0.00	5.73	2.98	26.13	5.83	35.30	5.59	25.35	4.64	16.15	3.33	5.26	1.86	0.00	0.00
6	3.36	2.29	3.82	1.99	6.19	1.38	9.29	1.47	10.05	1.84	9.08	1.87	5.29	1.87	3.62	2.17
7	3.05	2.08	4.42	2.30	10.73	2.39	13.61	2.16	16.25	2.97	23.66	4.87	13.69	4.84	4.91	2.95
8	2.52	1.72	11.93	6.21	87.43	19.50	139.12	22.03	96.06	17.59	61.95	12.75	15.69	5.55	2.19	1.31
9	10.88	7.42	10.92	5.68	11.89	2.65	11.36	1.80	14.03	2.57	19.34	3.98	16.15	5.71	12.58	7.54
10	30.58	20.84	28.97	15.08	41.57	9.27	44.03	6.97	46.87	8.58	51.88	10.68	40.85	14.45	31.88	19.11
11	1.08	0.74	2.93	1.53	12.54	2.80	24.63	3.90	17.05	3.12	21.95	4.52	10.69	3.78	2.39	1.44
합계	146.74	100	192.10	100	448.29	100	631.37	100	546.23	100	485.79	100	282.65	100	166.79	100



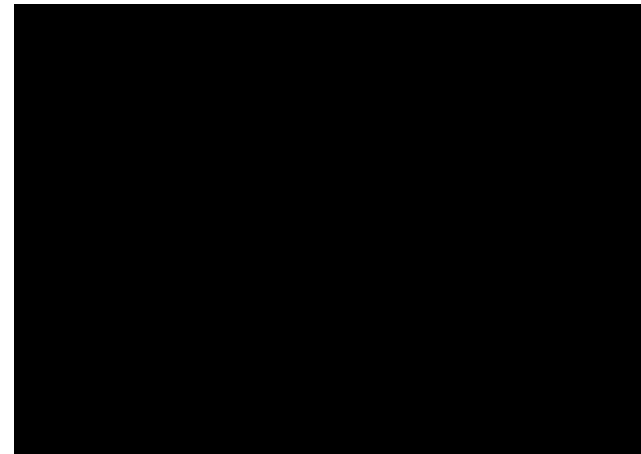
a) 03시-06시 점유 지역



b) 06-09시 점유 지역



c) 09-12시 점유 지역



d) 12-15시 점유 지역

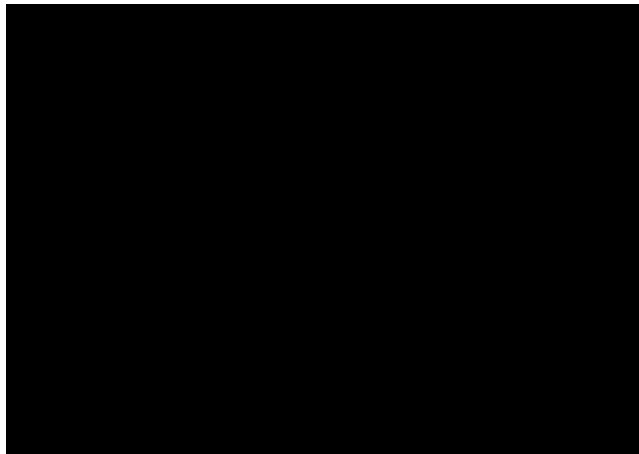
<그림 4-13> 40,50대 유동인구에 기반한 서울시의 시공간 점유 패턴1



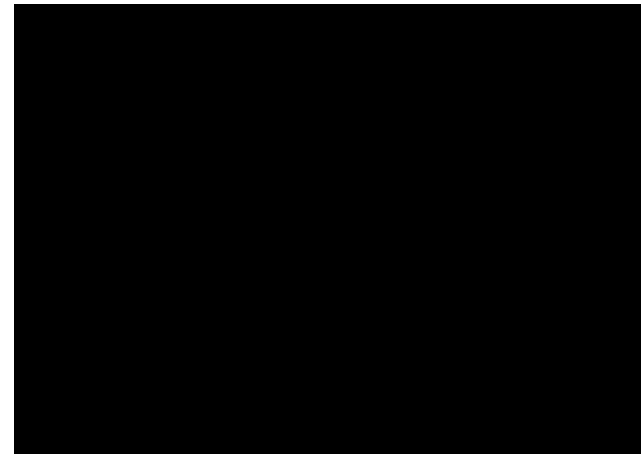
a) 15-18시 점유 지역



b) 18-21시 점유 지역



c) 21-24시 점유 지역



d) 24-03시 점유 지역

<그림 4-14> 40,50대 유동인구에 기반한 서울시의 시공간 점유 패턴 2

<표 4-11> 서울시 시공간 점유 패턴의 시간대별 변화 양상(60,70대)

단위: 1000명(%)

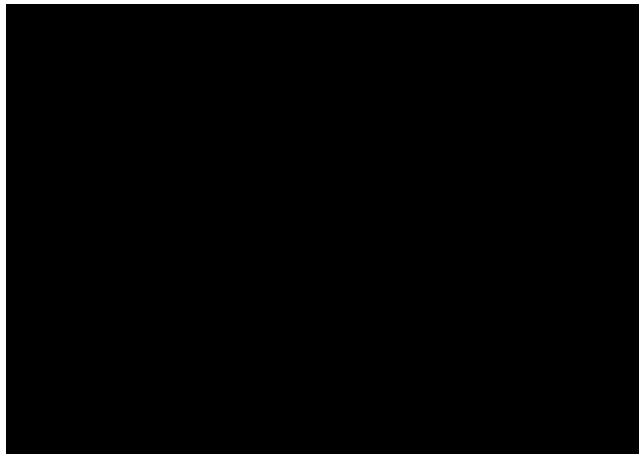
분류	03-06시		06-09시		09-12시		12-15시		15-18시		18-21시		21-24시		24-03시	
1	17.91	53.40	23.24	48.63	28.77	26.40	30.96	18.76	27.19	20.08	25.76	28.81	20.37	42.45	17.20	51.47
2	5.14	15.32	9.75	20.40	34.04	31.23	52.12	31.57	39.27	29.00	26.06	29.14	10.51	21.90	5.00	14.96
3	1.09	3.26	1.49	3.11	1.76	1.61	1.12	0.68	1.01	0.75	1.58	1.77	1.26	2.63	1.16	3.46
4	0.84	2.51	1.06	2.22	4.33	3.97	6.99	4.23	6.03	4.45	5.74	6.42	2.89	6.01	1.19	3.58
5	1.13	3.37	1.34	2.80	4.49	4.12	8.73	5.29	9.51	7.03	5.86	6.55	1.96	4.08	1.41	4.22
6	2.04	6.10	1.87	3.91	4.71	4.32	8.22	4.98	8.80	6.50	6.60	7.38	3.10	6.46	2.17	6.49
7	1.09	3.25	4.69	9.81	24.22	22.22	51.59	31.25	37.39	27.61	12.25	13.70	3.38	7.04	1.20	3.58
8	4.30	12.81	4.37	9.14	6.67	6.12	5.34	3.24	6.21	4.58	5.57	6.23	4.52	9.42	4.09	12.25
합계	33.53	100	47.79	100	108.98	100	165.07	100	135.40	100	89.41	100	48.00	100	33.41	100



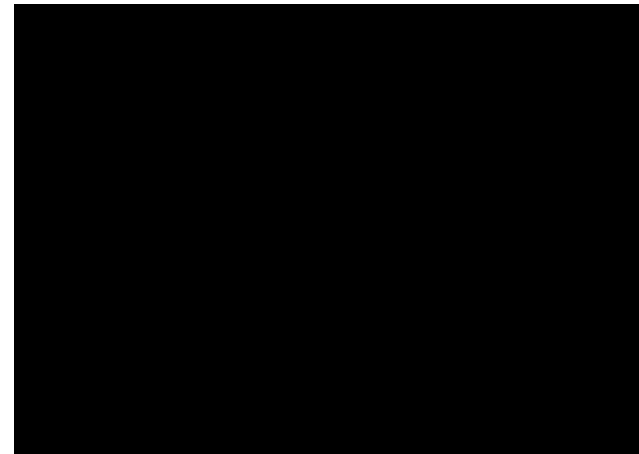
a) 03시-06시 점유 지역



b) 06-09시 점유 지역



c) 09-12시 점유 지역

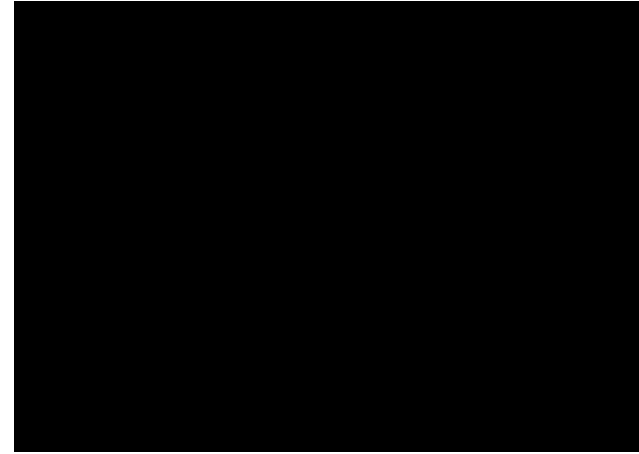


d) 12-15시 점유 지역

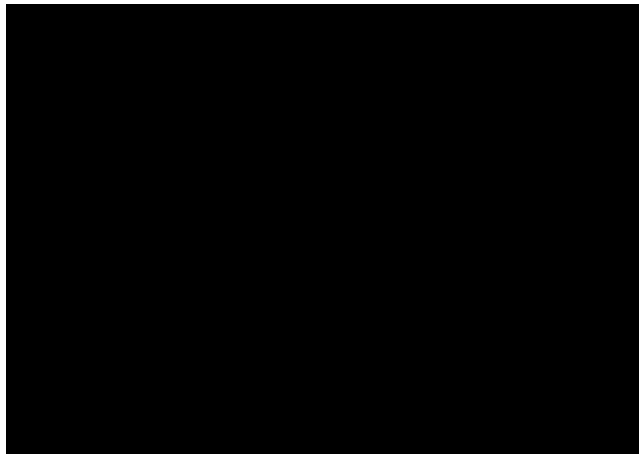
<그림 4-15> 60,70대 유동인구에 기반한 서울시의 시공간 점유 패턴 1



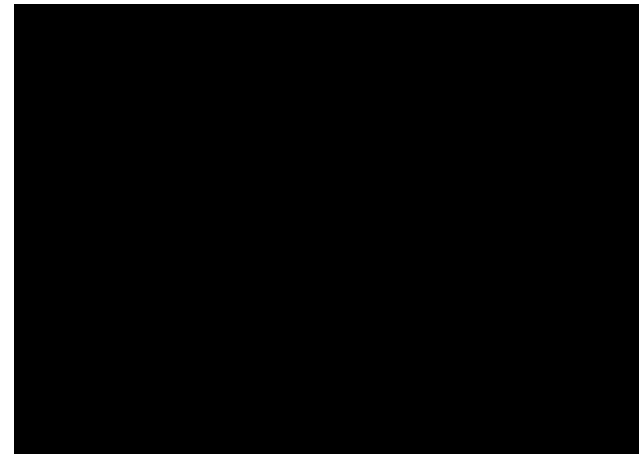
a) 15-18시 점유 지역



b) 18-21시 점유 지역



c) 21-24시 점유 지역



d) 24-03시 점유 지역

<그림 4-16> 60,70대 유동인구에 기반한 서울시의 시공간 점유 패턴 2

제 5 장 결론

본 연구는 시공간적으로 달라지는 도시 유동인구에 따라 유의미한 ‘시공간 점유패턴’이 형성되고 있는 것을 통계적인 기법을 통해 파악하였다. 유동인구의 시공간적 분포를 살펴본 연구가 다수 존재하였으나, 통계적 기법을 통해 살펴본 연구는 미흡하였으며, 공간적 대상 역시 서울시라는 특정 도시에 한정되어 있었기 때문이다. 그 결과 도시 유동인구의 역동적인 변화 양상을 파악하지 못하였다. 이러한 한계점들은 연구에 활용한 자료에서 기인하였다. 자료의 한계성을 해결하기 위하여 본 연구에서는 이동통신 빅데이터를 활용 하였으며, 시공간군집기법을 활용해 유동인구의 군집을 보다 체계적으로 도출하였다.

본 연구에서 활용한 시공간군집분석 기법은 Spatio-temporal AMOEBA이다. 시공간군집분석은 시공간자료의 군집을 포착할 수 있고, 군집의 시간대별 성장, 이동, 분할 및 결합 등 변화추이를 관측할 수 있다. 하지만 기존에 주로 사용되어온 시공간군집분석은 Space-time Scan statistics에 기반하고 있는데, 이는 불규칙적인 클러스터를 파악하지 못하고, 클러스터를 과대 추정하는 경향이 있다는 한계점을 갖는다. 이러한 한계점을 극복하기 위해 불규칙한 클러스터를 탐지하는 데에 유리한 공간군집기법인 AMOEBA 기법을 응용한 방법이 일부 제시되었다. 그 중 AMOEBA기법을 시공간축으로 응용한 Spatio-temporal AMOEBA를 활용하여 유동인구의 시공간 군집을 탐지하였다.

또한 보다 원활한 탐지를 위해 4가지의 사전설정을 진행하였다. 먼저 연구의 분석단위를 기초구역단위로 설정하였다. 기초구역단위는 행정동보다는 현재 국가 행정서비스 등의 기준이 되는 공간단위 중 비교적 상세한 편에 속하기 때문에, 군집의 위치 및 변화를 보다 상세하게 살펴볼 수 있다. 둘째, 산지 및 수변과 같이 인구의 유동이 불가능한 지역에서는 자료가 누락되어있다. 그로 인해 점형으로 제공되는 원자료를 면단위의 공간단위로 합역하는 과정에서 일부 지역의 유동인구 값은 과소추정된다. 이를 방지하기 위해 토지이용도를 활용하여 인구의 분포가 가능한

지역만을 추출하여 연구에 활용하였다. 셋째, 시공간군집분석의 과정에서 원활한 분석을 위해 데이터를 3시간 단위로 합역 하였으며, 시간대별 편차를 고려하기 위해 각 시간대별 인구밀도를 표준화하여 분석에 이용하였다. 넷째, 시공간군집분석의 특성상 시간적 시드를 설정할 필요성이 있으며, 본 연구에서는 03-06시를 인간의 활동시간대의 시드로 설정하였다.

연구의 결과 및 의의는 다음과 같다.

첫째, 시공간군집기법의 통계적 방법론을 통해 기존의 연구에서는 볼 수 없었던 유동인구의 시공간적 분포를 체계적으로 탐지하였으며, 이를 통해 도시의 유의미한 시공간적 점유 패턴을 파악할 수 있었다. 기존에 주로 이용되어온 서울시유동인구자료는 조사지점의 수가 매우 적어 통계적인 방법론을 통해 군집을 탐지하기에는 어려웠다. 하지만 본 연구에서는 이동통신 빅데이터와 시공간군집기법의 통계적 방법론을 활용하였으며, 서울시를 대상으로 유동인구의 군집과 그 시계열적 변화를 분석하였다. 이러한 시공간적 점유 패턴은 해당지역의 공간적 특성과 연관이 있는 것으로 보인다. 더불어 기존에 주로 이용되어 온 서울시유동인구자료는 조사지점의 수와 자료의 한시성이라는 문제로 인해 서울시와 같은 특정 지역에서만 자료가 구축되어 타 도시 연구에 적용하기 어려웠다. 하지만 본 연구에서 활용한 이동통신 빅데이터의 경우 상시적인 특성을 갖고 있으므로, 타 지역에서도 활용 가능할 것이다.

둘째, 성별 및 연령에 따라 유동인구의 군집 및 도시 점유 패턴이 다소 상이하게 나타남을 확인할 수 있었다. 즉 특정 계층에 따라 공간의 점유 및 소비하는 경향이 상이한 패턴을 보이는 것이다. 그동안 계층에 의한 도시공간의 점유 및 소비에 대한 담론은 여러 연구를 통해 논의되어 왔다. 하지만 이견학 외(2013)를 제외하고는 그러한 논의는 규범적이고 추상적인 면에서 논의가 이루어져왔으며(McDowell, 1993; Scraton and Watson, 1998), 실세계에서 어떻게 재현되고, 실제적 공간적 분포가 어떻게 이루어지는지에 대해서는 실증적으로 보여주지 못하였다. 하지만 본 연구에서는 이동통신 빅데이터를 통해 현실에 대한 상황을 분석하고 입증하였다는 점에서 관련 연구 분야에 실증적인 측면에서 기여를 할

수 있을 것이다.

셋째, 일부 지역은 특정 시간대에서만 점유되고 있음을 포착하였다. 시공간군집분석의 특성상 이러한 현상이 포착 가능하였다. 즉 도시공간 상에서 일시적인 공간점유가 나타난 것이다. 이서진(2007)과 정혜미(2013)에 의하면 공간은 특정 계층에 의해 장악되며, 이를 ‘한시적 경관’, 혹은 ‘일시적 도시공간 점유’라고 명명하였다. 이처럼 담론으로만 제기되었던 일시적 도시공간의 점유를 실증적으로 포착할 수 있었다.

한편 이러한 의의에도 다음과 같은 한계점이 존재한다.

첫째, MTUP(Modifiable Temporal Unit Problem)를 고려하지 못하였다. 본 연구에서는 시간대를 3시간단위로 임의적으로 나누어 총 8개의 시간단위에 대하여 분석을 진행하였다. 이처럼 임의적으로 시간단위를 정의하여 진행하였기 때문에 일부 지역을 제외하고는 대부분의 군집들이 모든 시간대에서 군집을 형성하고 있는 것으로 탐지되었다. MTUP에 대한 고려가 이루어지면, 보다 정교한 시공간군집의 포착이 가능해질 것이다.

둘째, 본 연구에서는 유동인구의 시공간적 군집, 및 성별, 연령에 따른 군집의 분포 차이를 확인하였지만, 그 구체적인 원인을 파악하지는 못하였다. 이러한 한계점들은 후속 연구에서 보완될 수 있을 것이다.

참 고 문 헌

- 김경태, 이인묵, 곽호찬, 민재홍, 2015, 유동인구 추정 시 통신자료의 활용에 관한 연구, 서울도시연구, 16(3), 177-187.
- 김기현, 손동욱, 이동훈, 2014, 상업가로 물리적 특성과 보행밀도와의 연관성 분석, 한국도시계획학회지, 15(5), 161-171.
- 김대중, 구형수, 2011, 시공간패턴분석을 통한 토지이용 변화 예측 및 활용방안 연구, 국토연구원 .
- 김미경, 강신혜, 김상필, 손홍규, 2016, 유동인구를 활용한 대피소 수용능력 분석, 한국측량학회지, 34(1), 1-10.
- 김영호, 2012, AMOEBA 기법을 응용한 서울시 자전거 네트워크의 클러스터 분석, 한국도시지리학회지, 15(3), 77-87.
- 김오석, 2014, 시계열 원격탐사 기법을 활용한 도시 식별 연구, 한국지도학회지, 14(2), 119-126.
- 김완희, 임종서, 송아현, 2018, 인구유동 자료로 살펴본 상업 집적지의 시공간적 변화: 2016년~2017년 사이의 홍대지역을 사례로, 한국지도학회지, 18(1), 49-63.
- 박병호, 인병철, 김태영, 2009, 전국 산업단지 쇠퇴 분석, 지역연구, 25(3), 61-73.
- 박선영·김영호, 2013, 서울시 R&D 산업체의 시공간 클러스터 분석. 한국경제지리학회지, 16(3), 492-511.
- 변미리, 서우석,, 2011. 연구논문, 도시 거리의 주간활동인구 측정과 해석: 서울시 유동인구 조사 사례 조사연구 12(2), 27-50.
- 신정엽, 2015, 시애틀의 도시 공간 특성: 역동성과 다양성, 한국지리학회지, 4(1), 119-137.
- 정수열, 임석희, 2012, 도시 내 이민자 자영업의 시공간적 역동성 -시카고 거주 한국인 이민자를 사례로 한국경제지리학회지, 15(3), 376-389.
- 윤정미, 최돈정 2015. 서울시 유동인구 분포의 공간 패턴과 토지이용 특성에 관한 지리가중 회귀분석, 한국지형공간정보학회지 23(3), 77

- 이건학, 김감영, 2016, 공간통계 기법을 이용한 현주인구 추정 모델링, 한국지도학회지 16(2), 71-93.
- 이건학, 신정엽, 홍유진, 2013, 젠더화된 도시 공간 탐색 : 도시 건물의 용도별 공간 분포를 중심으로. 한국지도학회지, 13(3), 57-73.
- 이경민, 정창무, 2014, 시간대별 유동인구가 업종별 점포 입지에 미치는 영향에 관한 연구: 수원시 소매업 및 음식업 점포를 중심으로, 대한건축학회 논문집 - 계획계, 30(8), 47 - 55.
- 이광국·양위주, 2007, LQ지수를 통한 관광산업의 도시공간적 분포의 시간적 변화특성에 관한 연구, 관광·레저연구, 19(1), 93-108.
- 이서진, 2007, 게이남성의 장소 형성 - 종로구 낙원동을 사례로- , 서울대학교 지리학과 석사학위 논문
- 이경주. (2013). 공간통계지수를 이용한 외연적 도시 확장 과정의 시공간 패턴고찰에 관한 연구: 미국 버펄로 지역을 중심으로. 한국지도학회지, 13(1), 155-165.
- 이상일, 조대현, 손학기, 채미옥, 2010, 공간 클러스터의 범역 설정을 위한 GIS-기반 방법론 연구 -수정 AMOEBA 기법-, 대한지리학회지, 45(4), 502-520.
- 이향숙, 추상호, 김지윤, 2014. 수도권 고령자의 통행사슬 특성에 관한 연구, 한국ITS학회논문지, 13(2), 68-79.
- 정윤영·문태현, 2014, 유동인구 자료를 이용한 서울시 도시공간구조 분석 연구 - 2030 서울플랜과 비교연구, 한국지역개발학회지, 26(3), 139-158.
- 정혜미, 2013, 일시적 도시공간 점유의 건축적 특징에 관한 연구, 서울대학교 건축학과 석사학위 논문.
- 조대현, 2013, 카운트 데이터 기반 공간 군집 분석 연구의 동향과 방법론적 이슈, 대한지리학회지, 48(5), 768-785.
- Aldstadt, J. and Getis, A., 2006, Using AMOEBA to Create a Spatial Weights Matrix and Identify Spatial Clusters, *Geographical Analysis*, 38(4), 327-343.
- Anselin, L., 1995, Local Indicators of Spatial Association - LISA,

- Geographical Analysis*, 27(2), 93–115.
- Arbia, G., Espa, G., Giuliani, D., and Dickson, M. M., 2014, Spatio-temporal clustering in the pharmaceutical and medical device manufacturing industry: A geographical micro-level analysis, *Regional Science and Urban Economics*, 49, 298 - 304.
- Clark, 1965
- Besag and Newell, 1991, The Detection of Clusters in Rare Diseases, *Journal of the Royal Statistical Society. Series A, Statistics in Society*, 154(1), 143–155.
- Bhaduri, B., Bright, E., Coleman, P., and Urban, M.L., 2007, LandScan USA: a high-resolution geospatial and temporal modeling approach for population distribution and dynamics, *GeoJournal*, 69, 103–117.
- Cromley, R. G. and McLafferty, 2012, GIS and Public Health, *Guilford Press*, New York, NY.
- Demattei, C., and Cucala, C., 2010, Multiple Spatio-Temporal Cluster Detection for Case Event Data: An Ordering-Based Approach, *Communications in Statistics - Theory and Methods*, 40(2) 358–372.
- Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F. R., Gaughan, A. E., Blondela, V. D., and Tatem, A. J., 2014, Dynamic population mapping using mobile phone data, *Proceedings of the National Academy of Sciences*, 111(45), 15888–15893.
- Douglass R. W., Meyer, D. A., Ram, M., Rideout, D., and Song, D., 2015, High resolution population estimates form telecommunications data, *EPJ Data Science*, 4(1), 1–13.
- Fotheringham, A. S., Crespo, R., and Yao, J., 2015, Geographical and temporal weighted regression(GTWR), *Geographical Analysis*, 47(4), 431–452.

- Geary, R. C., 1954, The contiguity ratio and statistical mapping, *The incorporated statistician*, 5, 115-145.
- Getis, A. and Aldstadt, J., 2004. Constructing the Spatial Weights Matrix Using a Local Statistic. *Geographical Analysis*, 36(2), 90 - 104.
- Getis, A., and Ord, J. K., The Analysis of Spatial Association by Use of Distance Statistics, *Geographical Analysis*, 24(3), 1992, 189 - 206.
- Girardin F, Vaccari A, Gerber A, Biderman A, Ratti C., 2009, Quantifying urban attractiveness from the distribution and density of digital footprints, *International Journal of Spatial Data Infrastructures Research*, 4, 175-200.
- Jacquez, G. M. and Greiling, D. A., 2003, Local clustering in breast, lung and colorectal cancer in Long Island, New York, *International Journal of Health Geographics*, 2(1), 3
- Jacque,C, 2010, Evaluating ERS-1 SAR data for the discrimination of tropical forest from other tropical vegetation types in Papua New Guinea, *International Journal of Remote Sensing*, 18(14), 2967-2984.
- Kim, D. J., 2009 Prospective Monitoring for Land Use Change Prediction: Applications of spatiotemporal pattern analyses to land transaction data, *the University at Buffalo, State University of New York*.
- Kroon. J., 2010 Spatio-temporal cluster detection using AMOEBA, *the Pennsylvania State University*.
- Kulldorff, M., 1997, A spatial scan statistic, *Communications in Statistics: Theory and Methods*, 26(6), 1487-1496.
- Lawson, A. B., 1993, On the analysis of mortality events associated with a prespecified fixed point, *Journal of the Royal Statistical Society: Series A*, 156(3), 363-377.
- Moran, P.A.P., 1948, The interpretation of statistical maps., *Journal of*

- the Royal Statistical Society: Series B*, 10, 245–251.
- Openshaw, S. and Charlton, M., 1987, A mark 1 Geographical Analysis Machine for the automated analysis of point data-sets, *International Journal of Geographical Information Systems*, 1(4), 335–358.
- Ord, J. K., and Getis, A., 1995, Local spatial autocorrelation statistics: Distributional issues and an application, *Geographical Analysis*, 27(4), 286 - 306.
- Simmons, J. and Yeates, M., 1998, Toronto: commercial structure and change, *Progress in Planning*, 50(4), 253–272.
- Song Gao, 2014, Spatio–Temporal Analytics for Exploring Human Mobility Patterns and Urban Dynamics in the Mobile Age. *Spatial Cognition & Computation*, 15(2). 86–114.
- Tsai, Y.H., 2005, Quantifying urban form: compactness versus 'sprawl', *Urban studies*, 42(1), 141–161.
- Thomas, L., Maxime, L., Oliva, G., Cantu, R., Miguel ,P., Ricardo, H., Enrique, F. M., Jose, J.R., Marc, B., 2014, From mobile phone data to the spatial structure of cities, *Scientific Reports*, 4(1), 5276–5187.
- Waller and Gotway, C. A., 2004. Applied spatial statistics for public health data Lance A. Waller, Carol A. Gotway., Hoboken, N.J., John Wiley & Sons.

Abstract

A study on Spatio-temporal occupation patterns in urban area, through the Mobile Big data

Wanhee Kim

Department of Geography

The Graduate School

Seoul National University

This study aims to identify meaningful ‘spatio-temporal occupation patterns in urban areas’ through the employment of spatial statistical techniques. Due to problems of raw data, prior research failed to detect systematic spatio-temporal distribution of floating populations. As a result, it was difficult to detect the dynamically changing aspects of cities. This study utilizes mobile big data - relatively free from the standards of accuracy and consistency -, and the spatio-temporal AMOEBA method, which is effective in identifying irregular cluster as a spatial-temporal clustered method.

The results and implications are summarized as follows:

First, the spatio-temporal distribution of the floating population was detected systematically through the statistical application of the spatio-temporal cluster method, which was not covered in previous studies. Meaningful occupation pattern was also detected through the change of clusters in time series. Such spatio-temporal occupation patterns reflect spatial characteristics of the area. In addition, the consistency feature of the mobile big data used in this study, allows for further application of this data to analyze in other regions.

Second, this research has verified the distribution of floating population and occupancy patterns in urban areas are different according to gender and age. Therefore, this study offers an empirical analysis of discourse on the occupancy and consumption of urban areas by a particular class. It also contributes quantitatively to prior research, which have mostly been normative and abstract overall.

Third, some clusters were distributed only at specific temporal scales. The concept of 'temporary landscape' or 'temporary occupancy of urban areas' dominated by a certain class are demonstrated through this quantitative research.

However, 'MTUP(Modified Temporal Unit Problem)' was not considered because time units were arbitrarily combined to simplify the operation. Once these limitations are addressed, additional research with higher temporal precision may be possible.

keywords : floating population, spatio-temporal occupation, spatio-temporal cluster, mobile big data, Spatio-temporal AMOEBA

Student Number : 2016-20205