

# 공간정보기반 클러스터링을 이용한 초고속인터넷 결합유형별 해지의 지역별 특성연구

박장혁

연세대학교 공과대학 기술정책협동  
(parkjanghyuk@kt.com)

박상언

경기대학교 경영정보학과  
(supark@kgu.ac.kr)

김우주

연세대학교 공과대학 정보산업공학과  
(wkim@yonsei.ac.kr)

『2016 인터넷이용실태조사』에 따르면 인터넷 이용자수 및 이용률은 점점 증가하고 있으며 접속방법에 있어서는 컴퓨터보다 스마트폰을 통한 접속이 많아지고 있다. 스마트기기의 증가에 따라 초고속인터넷의 수요가 감소할 것이라는 전망도 있다. 하지만, 스마트기기의 증가에도 불구하고 기가인터넷을 통한 속도 향상과 IoT 시장의 성장으로 인해 초고속인터넷 시장은 당분간 유지될 것으로 전망된다. 시장의 포화로 인해 통신사업자들이 신규고객 확보를 위해 과도한 경쟁을 하고 있지만, 고객이탈의 원인을 알 수 있다면 보다 효과적인 마케팅을 통해 과도한 마케팅비용을 절감할 수 있을 것으로 기대된다. 본 연구에서는 통신사업자 A사가 보유하고 있는 안양시, 군포시, 의왕시 3개 도시의 결합유형별 해지 데이터와, 통계청으로부터 구한 지역별 데이터를 결합하여, 지역별 해지율과 이에 영향을 미치는 지역특성간의 관계를 분석하고자 하였다. 특히 인접지역에 따라 결합유형별 해지율의 분포에 차이가 있을 것으로 보고, 클러스터링을 이용하여 해지유형이 유사한 지역을 도출 및 분석하고자 하였다. 공간검색통계도구인 SatScan은 기존의 클러스터링 방법에 공간정보를 추가하여 인접지역을 중심으로 군집이 형성되도록 한다. 따라서 본 연구에서는 SatScan을 이용해 지역의 공간정보를 기반으로 유사지역을 군집화하고, 군집별 해지율과 지역별 데이터와의 연관성을 분석하였다. 분석 단계에서는 먼저 공간정보와 해지데이터를 결합하여 도출된 군집들의 특성을 정리하였으며, 다음으로 군집분석 결과를 바탕으로 하여 각 동의 초고속 인터넷 해지율과 지역별 데이터와의 연관성을 분산분석, 상관분석, 회귀분석을 이용하여 분석하였다. 그리고, 분석결과를 기반으로 하여 지역에 따른 적절한 마케팅 방안을 제안하였다.

**주제어** : 공간검색통계, 초고속인터넷, 인터넷해지, 타겟마케팅, SatScan, 클러스터링

논문접수일 : 2017년 7월 21일    논문수정일 : 2017년 9월 19일    게재확정일 : 2017년 9월 20일  
원고유형 : 일반논문    교신저자 : 김우주

## 1. 서론

'16년 12월 미래창조과학부와 한국인터넷진흥원이 발표한 '2016 인터넷이용실태조사'에 따르면 인터넷 접속이 가능한 가구는 전체 99.2%이고 만3세 이상 인구의 인터넷이용률은 약 88.3%에 달한다(KI&SA, 2016). 인터넷을 이용하는 목적은 이메일, SNS, 채팅 등 커뮤니케이션이 제일

높았으며 <Table 1>과 같이 인스턴트메시지, SNS, 인터넷 쇼핑은 무선기기를 선호 하고, 이메일은 아직까지는 유선기기를 좀 더 선호하는 것으로 나타났다(Digital Media, 2016).

이와 같이 무선기기 이용이 확대됨에 따라 초고속인터넷(유선인터넷)의 수요가 감소할 것으로 전망하는 사람들이 적지 않다. 1998년 도입 이후 높은 성장률을 보인 초고속인터넷 시장은

〈Table 1〉 Internet Device Usage

Device	Instant messages	SNS	Internet shopping	e-mail
Mobile	<b>92.3%</b>	<b>64.7%</b>	<b>50.8%</b>	74.1%
Wire	21.2%	23.4%	40.7%	<b>81.7%</b>

Source : “2016 Internet Usage Survey”, Korea Internet & Security Agency

〈Table 2〉 Subscriber Growth Trend (unit : Thousand)

Type		2013E	2014E	2015E	2016E
Mobile	Subscriber	54,681	57,290	58,935	61,295
	<b>Increase rate</b>	<b>2.0%</b>	<b>4.8%</b>	<b>2.9%</b>	<b>4.0%</b>
wire internet	Subscriber	18,738	19,199	20,024	20,556
	<b>Increase rate</b>	<b>2.7%</b>	<b>2.5%</b>	<b>4.3%</b>	<b>2.7%</b>

Source : Ministry of Science, ICT and Future Planning (<http://www.msip.go.kr>)

2002년 이후 둔화되고 있다. 그러나, 인터넷 이용자 중 94.8%가 주로 이용하는 장소가 ‘가정’이라고 답했으며(KI&SA, 2016), 미래창조과학부에서 발표하는 가입자현황 자료에 따르면 <Table 2>과 같이 휴대폰뿐만 아니라 초고속인터넷 가입자도 매년 증가하고 있는 것으로 나타났다(Ministry of Science, 2017).

이와 같이 인터넷 이용자들의 행태와 이용자 수 변화 등을 볼 때 향후에도 무선인터넷으로의 완벽한 대체가 아닌 상호 보완적인 관계가 유지될 것으로 보여지며, 기가인터넷을 통한 속도 업그레이드, IoT서비스 발전 등으로 인해 초고속인터넷에 대한 수요는 당분간 증가할 것으로 판단된다(KISDI, 2015). 초고속인터넷 시장의 마케팅 비용은 2010년 ‘통신사업자 마케팅비 가이드라인’ 발표이후 감소세였으나 ’13년에 다시 증가추세로 전환되어 ’14년에는 전년대비 27.5% 증가한 6,489억원을 기록하였다(AoTMC, 2015). 통신

사업자들은 기존고객 유지 및 신규고객 확보를 위해 상품결합할인, 요금할인, 위약금 대납, 현금 지급 등의 다양한 마케팅 수단을 사용하고 있지만, 신규 확보는 어렵고 해지위협에는 쉽게 노출되어 있는 편이다(Kim, 2011; Park and Min, 2002). 이러한 상황에서 고객들의 해지원인을 알 수 있다면 보다 효과적인 마케팅이 가능할 것으로 기대된다.

본 연구는 해지원인을 파악하고 이에 대처하기 위한 방법의 하나로, 초고속인터넷 해지데이터를 지역별로 결합상품의 유형에 따라 정리하여, 지역별 특성 데이터와의 연관성을 분석하는 방안을 제안하고자 한다. 특히 지역별 특성을 마케팅에 보다 효과적으로 활용하기 위해서는 기존 연구의 주 분석단위인 시군구 단위보다 더 세분화된 동 단위의 연구가 더 적합할 것으로 판단하고, 안양시, 군포시, 의왕시 3개 도시 지역의 행정동 단위 29개동을 대상으로 하여 동 단위로

데이터를 수집 및 정리하였다. 그리고, 이러한 동 단위의 공간적 특성을 분석에 반영하기 위해, 인접한 동일수록 결합상품별 해지유형이 더 유사한 경향이 있을 것이라는 가정 아래, 각 동의 공간정보를 결합상품에 따른 해지유형 데이터와 결합하여 유사한 군집을 도출하는 공간정보 기반의 군집분석을 실시하였다. 이와 같이 공간정보를 이용하여 지역에 따른 특성을 분석하는 기법은, 구글 독감예측을 비롯하여 다양한 분야에서 활용되고 있으며 그 효과가 입증되고 있다 (Arvind, 2012; Hwang, 2012; Park and Choi, 2016).

분석 단계에서는 먼저 공간정보와 해지데이터를 결합하여 도출된 군집들의 특성을 정리하였으며, 다음으로 군집분석 결과를 바탕으로 하여 각 동의 초고속 인터넷 서비스의 결합유형별 해지율과 연령대별 비중, 자동차 등록비율, 아파트 거래내역 데이터와 같은 지역별 데이터와의 연관성을 분산분석, 상관분석, 회귀분석을 이용하여 분석하였다. 그리고 분석 결과를 기반으로 하여 지역에 맞게 적절한 마케팅방안을 제시하였다. 본 논문의 구성은 다음과 같다. 2장에서는 본 연구와 관련된 기존 연구의 내용을 정리하고 3장에서는 연구 설계의 내용을 기술하였다. 4장에서는 각 동의 해지유형과 지역특성 데이터와의 연관성에 대한 분석 결과를 정리하였다. 마지막으로 5장에는 결론과 연구의 한계 및 향후 연구 방안을 기술하였다.

## 2. 관련 연구

### 2.1 공간정보 및 공간검색통계량 활용 연구

공간정보는 위치, 분포 등 지도 위에 표현이

가능한 모든 정보로 지형, 지질, 토지, 가옥, 상하수도 등이 포함되며 공간정보가 포함된 다양한 지도를 이용한 분석을 의사결정에 활용할 수 있다(KNSDIP, 2017). 대표적인 공간정보 활용 연구의 사례로 구글 독감예측 시스템을 들 수 있다. 2008년 구글은 독감과 관련된 검색을 하는 건수와 실제 독감 증세가 있는 사람 수 간에 상관관계가 있음을 발견하고, 검색에 입력된 단어들 중 ‘감기’와 연관된 단어의 빈도수를 분석하여 특정 지역의 독감을 예측하는 시스템을 개발하였다. 당초 기대와는 달리 실제와 예측 값이 다르다는 연구 결과가 있으나, 구글 독감예측 시스템이 감기 등 질병 감지, 경향 분석 등으로 충분한 활용 가치가 있다는 주장 또한 많다 (Andrea, 2013; Pervaiz and Abdur, 2012). 구글 외에도 스마트기기의 발전에 따라 지도정보, 교통정보, 시설정보, 날씨정보 등 공간정보를 이용한 서비스 및 관련 연구가 최근 증가하고 있다 (KNSDIP, 2017). 특히 공간검색통계량은 발생된 데이터를 기반으로 지역별 통계를 내거나 발생 위험이 높은 영역을 탐지하는 등 다양한 분야에 활용되고 있다(Arvind, 2012). 예를 들어, 범죄, 질병 등이 불특정 지역에서 무작위로 발생하는지 혹은 가까운 지역 별로 일정한 군집을 형성하는지의 여부에 따라 대처방법이 달라지므로, 그와 같은 상황에서 공간검색통계량이 유용하게 활용될 수 있다. 공간검색통계량은 Martin Kulldorff(1997)에 의해 제안된 방법으로, 분석대상지역 내에서 발생한 사건 데이터를 이용하여 향후 발생 확률이 다른 지역에 비하여 높거나 낮은 유사지역을 도출하는 방법이다. 공간검색통계량의 수학적 정의는 베르누이 모형(Bernoulli model)과 포아송모형(poisson model)로 정의되며 (Lee and Hwang, 2015) 군집은 우도비(likelihood

ratio)의 크기로 결정된다(Han and Lee, 2016; Ji, 2015; Jung, 2010).

박진옥(2016)은 2005년부터 2013년까지의 한국의 대기오염수치와 순환기질환 표준화 사망자수를 이용하여 대기오염이 순환기계 건강에 미치는 영향을 분석하였다. 공간정보를 시군구 251개 구역으로 세분화하여 분석하였는데 시공간 스캔통계량은 SatScan 포아송모델을 활용하였으며, 이를 이용해 대기오염과 질환의 발생이 높은 위험지역을 추정하여 질환 예방 및 관리의 근거를 제시하였다(Park and Choi, 2016). 황승식(2012)은 시군구별 자살률을 지도화하여 공간검색통계량을 이용해 자살률이 높은 군집을 도출하고, 지역의 사회환경적 요인과 자살률의 관련성을 분석하고자 하였다. 또한, 지역의 사회환경적 요인과 공간정보에 대한 회귀분석 결과, 물질적 박탈지수, 기초생활보장가구비율, 초졸이하 가구 비율, 정신적 활동제한 인구비율, 65세 이상 인구비율이 유의한 요인으로 분석되었으며, 특히 초졸 이하 가구 비율이 높은 지역이 상대적으로 자살률이 높은 것으로 추정되었다(Hwang, 2012). 이 밖에도 구제역, 고병원성 조류인플루엔자, Dengue 등 전염성이 높은 질병이 발생한 경우, 감염 확산 및 재발을 방지하는 차원에서 기 발생 데이터 근간으로 공간검색통계를 활용한 선행연구들이 있다. Arvind B. Bambhroliya (2012)은 공간검색통계량을 활용하여 텍사스주 카운티 단위의 유방암 사망률을 분석하였다. 유방암 사망률 고위험군을 탐지하고, 인종 및 혈통이 유방암 사망률의 지리적 분포에 미치는 영향을 알아내기 위해, SatScan 포아송모델을 이용하여 유방암 사망률이 높은 5개 군집과 인종, 혈통에 따라 유방암 사망 위험이 다른 3개의 군집을 도출하였다. 그 결과, 텍사스 남부지역 군집히스

패닉의 위험이 4.77배로 가장 높게 나타난 반면 비히스패닉 흑인의 위험이 가장 낮았으며, 텍사스 북동부 지역의 군집에서는 비히스패닉 흑인의 위험이 3.43배로 가장 높은 반면 히스패닉의 위험이 가장 낮았다(Arvind, 2012).

## 2.2 지역 특성 데이터 활용 연구

김리영, 서원석(2016)은 저 출산으로 인해 학생수가 감소하고 있음에도 불구하고 증가하고 있는 사교육시장을 학생밀도 관점에서 바라보았다. 학생밀도가 높은 지역이 어떻게 분포되어 있는지, 학생밀도에 영향을 주는 변수가 무엇인지 연구하는데 있어서 대학진학률, 사설학원수, 교통사고율, 대학원이상비율, 도시공원, 통학시간, 자가/전세, 지방세 등의 지역특성 변수를 사용하였다. 지역특성 변수 중 사설 학원수, 교통사고율, 자가/전세비율 등이 학생층 가구가 모여드는데 영향을 주는 요인으로 나타났다(Kim and Seo, 2016). 이지은, 김갑성(2016)은 지역의 주거환경 특성이 자살률에 미치는 영향요인을 분석하는 연구에서 물리환경적 특성변수로 빈집 비율, 노후주택비율, 1인당 공원면적, 농업/어업비율을 사용하였고 경제적 특성변수로 1인당 GRDP, 자동차 등록대수, 사회적 특성 변수로 고령인구비율, 이혼율, 음주율 등을 사용하였다. 그 결과, 공원면적은 5%수준에서 유의했고 빈집비율, 고령비율, 흡연율, 이혼율, 스트레스인지율은 1%수준에서 유의미하게 나타났다. 개인단위에서 일어나는 자살현상을 지역단위로 분석함으로써 향후 정책적 고려사항을 제시했다는 것에 의의를 들 수 있다(Lee and Kim, 2016). 이의용(2014)은 경기도 각 시군의 특성을 경제, 인적, 환경, 공공, 문화, 복지 6개 요소로 나누고 GRDP, 1인가구비

율, 쓰레기배출량, 자동차 보급률, 1인당 경찰인원수 등 총 28개 지표를 활용 하여 6개 군집으로 유형화하여 분석하였다(Lee, 2014).

### 2.3 초고속인터넷 서비스 가입결정요인 관련 연구

초고속인터넷 서비스에 대한 기존 연구 유형은 결합상품에 대한 효과에 대한 연구들과 서비스만족도, 충성도 등이 전환에 미치는 연구들로 분류할 수 있는데, 사실상 초고속인터넷을 선택할 때 미치는 결정요인에 관한 연구라고 볼 수 있다. 먼저 결합상품에 대한 연구를 살펴보면, 초고속인터넷의 결합상품은 구매결정에 크게 작용하는 요인으로 2007년 등장 이후 지속적으로 증가하였고 휴대폰과 결합된 결합상품이 출시되면서 경쟁이 본격화되었다(AoTMC, 2015). 초기에는 휴대폰과 결합하면 공짜라는 식의 공짜마케팅이 유효했으나 과장광고라는 논란으로 현재는 규제되고 있지만(Lee and Kim, 2016) 결합상품 가입률은 세대기준 62.4%(14년기준)로 여전히 경쟁에 미치는 영향이 크다고 할 수 있다(Kim, 2015). 결합상품에 대해서는 정책, lock-in 효과, 전환비용 등 다양한 분야에서 연구되고 있다(Kim, 2011; Kim, 2015; Kim and Park, 2010). 한편, 고객의 구매 및 전환행동 요인을 파악하고자 하는 선행연구들을 보면, 고객이 기존 통신사업자를 유지할 것인지 타사로 전환할 것인지를 결정하는 과정에 어떤 요인들이 영향을 미쳤는지에 대한 연구들이다. 세부적으로 보면 브랜드, 기술력, 요금, A/S 등 전반적인 만족도 평가를 통한 충성도(Kim and Park, 2008; Kim, 2011; Lee, 2011)를 분석하는 연구와 전환의도(Jae and Jeon, 2010; Park and Min, 2012) 또는 해지이유(Park

and Lee, 2002)를 분석하는 연구, 소비자 불만, 보상체계 등 사후서비스가 재계약에 미치는 영향에 관한 연구(Lee and Kim, 2011; Lee, 2009)로 분류된다.

### 2.4 기존 연구와의 차이

초고속인터넷에 대한 기존 연구는 대체적으로 설문조사를 통한 분석방법이지만 본 연구는 A 통신사에서 실제 수집한 동 별 데이터와 이에 대응하여 통계청 등의 기관으로부터 수집한 지역 특성 데이터를 최대한 활용하고자 하였다. 본 연구를 진행하는 데 있어 가장 크게 어려웠던 점은, 통계청을 비롯한 국가기관에서 제공되는 데이터들은 대부분 시군구 단위였고 동 단위의 데이터들은 잘 제공되지 않고 있다는 점이었다. 시군구 단위의 데이터는 보유한 해지 데이터의 특성과 본래의 동 단위 연구 목표로 인해 분석에 활용할 수 없었기 때문에, 제한된 데이터를 이용하여 최대한의 분석을 수행하고자 하였다.

지역특성 분석과 관련한 기존 연구들이 주어진 행정구역 단위로 지역특성을 분석한 것과는 달리, 본 연구는 인접지역을 중심으로 자연스럽게 해지유형이 유사한 지역 즉 동들이 군집으로 묶일 수 있도록 공간정보 기반의 클러스터링을 수행한다는 점에서 학술적 차별점을 갖는다고 할 수 있다. 또한 초고속인터넷 서비스의 가입결정요인에 대한 기존 연구에서는 지역적인 특성을 고려한 연구가 거의 없는 반면, 본 연구에서는 지역에 따라 서로 다른 요인이 있을 것이라 가정하고 도출된 군집들과 지역적 특성 데이터와의 연관성을 분석하고자 하였다.

### 3. 연구 방법

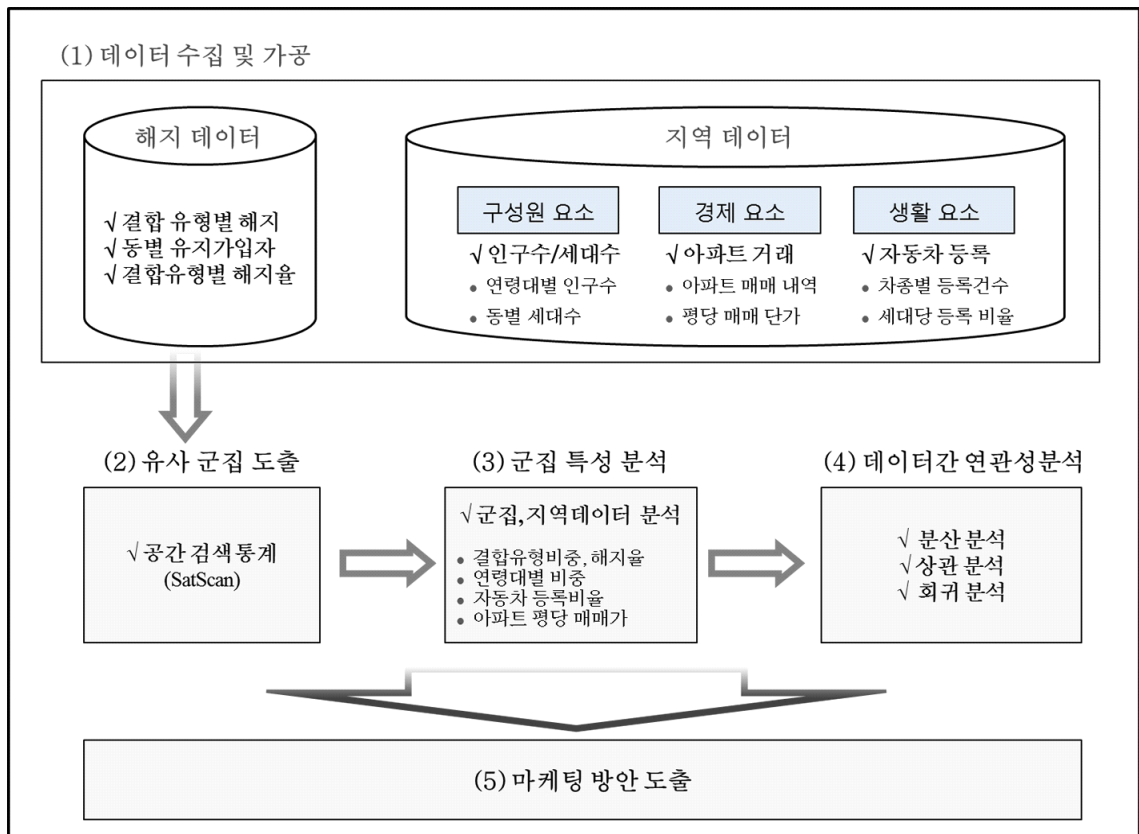
#### 3.1 연구 설계

연구는 <Figure 1>와 같이 5단계로 진행된다. 먼저 (1) 데이터를 수집 및 가공하고, (2) 공간검색통계를 이용해 유사 군집을 도출한 후, (3) 도출된 군집의 특성을 지역별 데이터를 기반으로 분석하고, (4) 분산분석, 상관관계, 회귀분석을 통해 데이터 간의 연관성을 파악하며, 마지막으로 (5) 분석결과를 기반으로 마케팅 방안을 도출하고자 한다.

#### 3.2 단계별 수행 내용

##### 3.2.1 데이터 수집 및 가공

초고속 인터넷 해지데이터는 안양시, 군포시, 의왕시 3개 도시의 행정동 단위 29개동(총17,177건, '15년 1월~12월)의 데이터를 상품 결합유형별로 수집하였다. 각 지역별 세대수, 연령대별 인구수, 자동차 등록대수는 통계청 자료를 사용하였고, 아파트 거래내역은 국토교통부에서 실거래 내역(총18,400건)을 수집하였다. 데이터들의 기준은 2015년 12월말이다. 이밖에 1인당 GRDP, 인구변동, 지역변수, 교육 정도, 주거형



<Figure 1> Research Procedure

태, 업종 등의 지역특성 데이터들이 있으나 대부분 시군구까지만 공개되어 동 단위의 분석을 필요로 하는 본 연구에서는 활용할 수 없었다. 인구데이터는 동별 연령대별로 구분하였고, 자동차 등록비율은 동별로 승용차, 승합차, 트럭, 특수차로 나누었다. 자동차 등록비율은 동별 구성원들의 소득, 직업 등을 표현하는 데이터로 가정하였고, 아파트 거래내역은 각 아파트별 평당 거래가격을 통해 해당 지역의 자산 수준을 나타내는 경제적인 지표로 활용하고자 하였다.

### 3.2.2 공간정보를 이용한 유사 군집 도출

통신사에서 제공하는 결합유형은 <Table 3>와 같이 인터넷을 기준으로 휴대폰이 포함된 결합, 휴대폰이 포함되지 않은 유선상품결합, 미결합으로 분류할 수 있다. 인접지역에 따라 결합유형별 해지율의 분포에 차이가 있을 것으로 보고 공간검색통계 도구인 SatScan을 이용해, 동을 대상으로 각 동의 위치정보와 결합유형별 해지율로 군집분석을 실시하였다. 군집분석은 사건발생 데이터가 3가지 이상의 유형들(categories)로 나뉘어 있을 때 사용하는 다항모형을 활용하여 (Jung, 2010) 휴대폰결합, 유선결합, 미결합 3가지의 결합유형으로 분류하여 수행하였다. 공간검색통계를 위해 필요한 위치정보는 각 동의 주민센터 좌표 값을 기준으로 하였다.

SatScan은 각 지역의 대표 좌표 값을 중심으로 반지름을 0에서부터 분석대상지역의 50%가 될

때까지 등근 원(window)을 확장시키면서 탐색하면서, 원 내부의 우도비율(likelihood ratio)과 원 외부의 우도비율 간의 차이를 기준으로 군집을 찾는다(Han and Lee, 2016; Hong, 2003; Jung, 2010). 우도비율 차이가 크다는 것 즉, 우도비가 높다는 것은 해당 군집에서의 발생 빈도가 이의 지역보다 높으며 하나의 군집으로 볼 수 있는 가능성이 크다는 것을 의미한다(Lee and Hwang, 2015).

### 3.2.3 군집 특성 분석

SatScan을 이용한 군집분석은 발생유형이 유사한 동들의 군집을 도출한다. 이 때 각 군집은 위치 정보와 서비스 결합유형별 해지율에 따라 생성되었지만, 다른 지역적 특성 관점에서도 유의한 차이가 있는 지를 확인할 필요가 있다 (Jung, 2010). 이를 위해 군집들에 대해 각 동별 결합비중, 해지율, 세대수 및 연령대별 비중, 세대당 자동차 등록비율, 아파트 거래내역을 기반으로 군집들의 특성을 분석하였다.

### 3.2.4 결합유형별 해지율과 지역적 특성 데이터의 연관성 분석

연령대별 비중, 세대당 자동차 등록비율, 동별 아파트 평당 매매가격 등의 지역 특성 값들과 도출된 군집들의 결합유형별 해지율과의 연관성을 분석하기 위해, 분산분석, 상관분석, 그리고 회귀 분석을 실시하였다. 그리고, 분석결과를 기반으로

<Table3> Type of Combination

Category	Combination Type	Combination example
1	Cellphone	Wire Internet + Cellphone + $\alpha$
2	Wired goods	Wire Internet + IPTV or telephone + $\alpha$
3	None	Wire Internet only

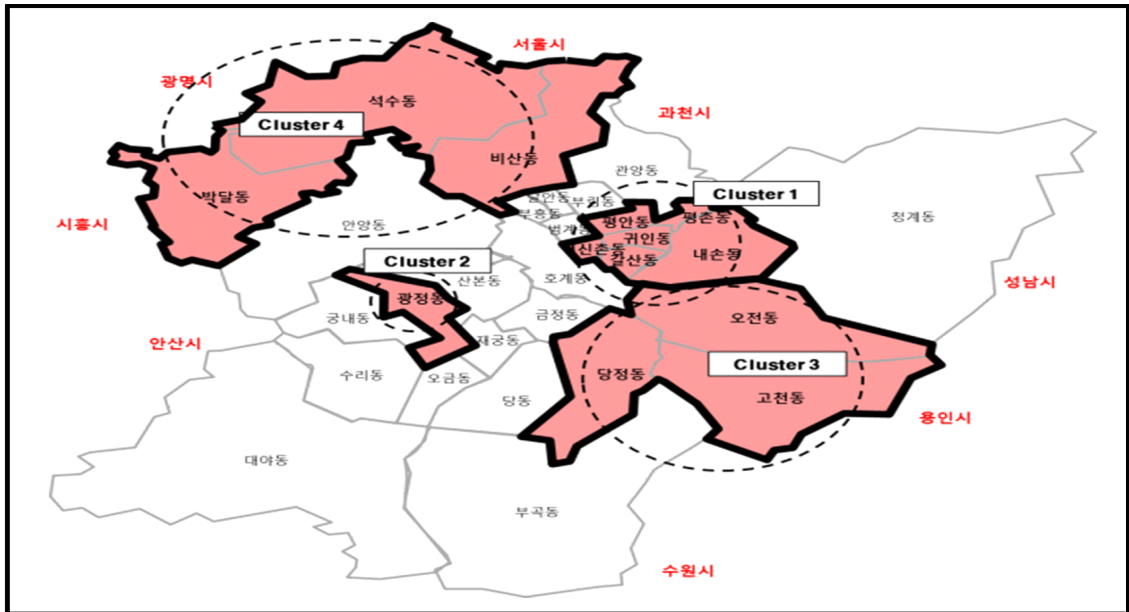
로 하여 지역에 따른 적절한 마케팅 방안을 제안 하였다.

#### 4. 결합유형별 해지와 지역특성 간의 연관성 분석 결과

##### 4.1 공간검색통계량을 이용한 군집 도출

SatScan을 이용해 초고속인터넷 결합유형별 해지율과 위치정보로 군집분석을 수행한 결과

<Figure 2>와 같이 총 13개동, 4개의 군집이 도출 되었다. <Table 4>에서 보는 바와 같이 군집1은 반경 약1.6km의 6개동, 군집2는 1개동, 군집3은 반경 2.04km의 3개동, 군집4는 반경 2.95km의 3개동으로 형성 되었으며, 4개 군집 모두 p-value 가 0.001로 유의한 것으로 나타났다. 이는 16개동에 대해서는 위치정보와 결합유형별 해지율의 유의한 차이가 나타나지 않았다고 해석할 수 있으며, 본 연구에서는 군집이 도출된 13개 동을 중심으로 분석을 진행하고자 한다.



<Figure 2> Clustering Result on the Map

<Table 4> Clustering Result

Cluster	Region	Radius	Log Likelihood	p-value
1	Galsan, naeson, kwiin, sinchon, pyeongchon, pyeongan (6ea)	1.61Km	36,183.956	0.001
2	Kwangjung (1ea)	0 Km	30,299,248	0.001
3	Kocheon, ojeon, dangjung (3ea)	2.04 Km	8,919,909	0.001
4	Seoksu, bakdal, bisan (3ea)	2.95 Km	2,364,047	0.001



## 4.2 군집 특성 분석

군집별 특성을 파악하기 위해 <Table 5>와 같이 각 군집별로 결합유형 비중과 해지율, 연령대별 비중, 자동차 등록비율, 아파트 평당 매매가를 비교하였다. 평균대비 7%포인트 높으면 높은 수치로 보았고, 평균대비 7%포인트 낮으면 낮은 수치로 보았다.

### 4.2.1 결합 비율 및 해지율

군집별 결합유형 비중과 해지율을 비교해본 결과 군집1과 군집4는 미결합 비중은 낮으나 해지율이 높은 반면, 군집2와 군집3은 미결합 비중은 높고 해지율은 낮았다. 군집1,4는 미결합고객 비중은 낮지만 미결합고객 중에서 해지고객이 많았고, 군집2,3은 미결합고객 비중은 높지만 미

<Table 5> Analysis of Cluster Feature

		Cluster1	Cluster2	Cluster3	Cluster4	Average
Combination rate	Cellphone	45.40%	49.70% *	36.90% **	45.90% *	42.90%
	Wire	43.00%	35.30% **	48.40% *	42.20%	43.80%
	None	11.60% **	15.00% *	14.70% *	11.80% **	13.30%
Termination rate	Cellphone	0.46%	0.56% *	0.49%	0.49%	0.46%
	Wire	1.22%	1.06% **	1.12% **	1.37% *	1.22%
	None	4.42% *	1.76% **	3.05% **	4.76% *	3.72%
	All	1.25%	0.92% **	1.17%	1.37% *	1.23%
Ratio of age group	1. Preschooler (0~7)	7.60%	7.80% *	7.10%	7.00%	7.10%
	2. Student (8~19)	16.80% *	15.10% *	13.20%	12.90%	13.40%
	3. Youth (20~35)	21.60% **	22.10% **	22.90%	23.50%	23.70%
	4. Middle(36~55)	36.90% *	34.90%	34.20%	34.30%	34.50%
	5. elderly (56~69)	12.00% **	12.10% **	16.30% *	15.60%	14.60%
	6. Old age (70~)	5.20% **	7.90% *	6.30%	6.80%	6.60%
	Sum	100%	100%	100%	100%	100%
Vehicle registration rate per household	1. Sedan	88.50% *	78.90%	87.50% *	75.10%	76.60%
	2. Van	3.10% **	3.30% **	4.80% *	4.60% *	4.10%
	3. Truck	7.80% **	5.70% **	17.80% *	11.50%	11.90%
	4. Special	0.10% **	0.10% **	0.40% *	0.10% **	0.30%
	Sum	99.50% *	78.90%	110.40% *	91.30%	92.80%
Apartment transaction price per PY	1. 20PY under	1,907 *	1,499	1,229 **	1,464	1,508
	2. 20PY ~ 29PY	1,725 *	1,419	1,213 **	1,366	1,439
	3. 30 PY ~39PY	1,671 *	1,205 **	1,007 **	1,222 **	1,337
	4. 40 PY ~	1,412 *	982 **	925 **	1,174	1,238
	All	1,742 *	1,348	1,181 **	1,361	1,440

Apartment price unit : 10,000won, PY = 3.3m<sup>2</sup>

\*: higher than 7% points of the average, \*\*: lower than 7% points of the average

결합고객 중 해지고객이 많지 않았기 때문이다. 이는 결합 비중과 상관 없이 지역적 특성에 따라 기존 가입자 결합 및 유지 마케팅이 필요하다는 의미를 준다.

#### 4.2.2 연령대 비중으로 본 지역 특성

연령대별 비중으로 비교한 결과로는 군집1은 학생층과 중년층 비율이 타 지역 대비 높게 나타났고 군집3은 장년층 비율이 비교적 높게 나타났다. 군집1 지역은 주로 학생자녀를 둔 중년층 세대들로 구성되어 있음을 추정할 수 있다.

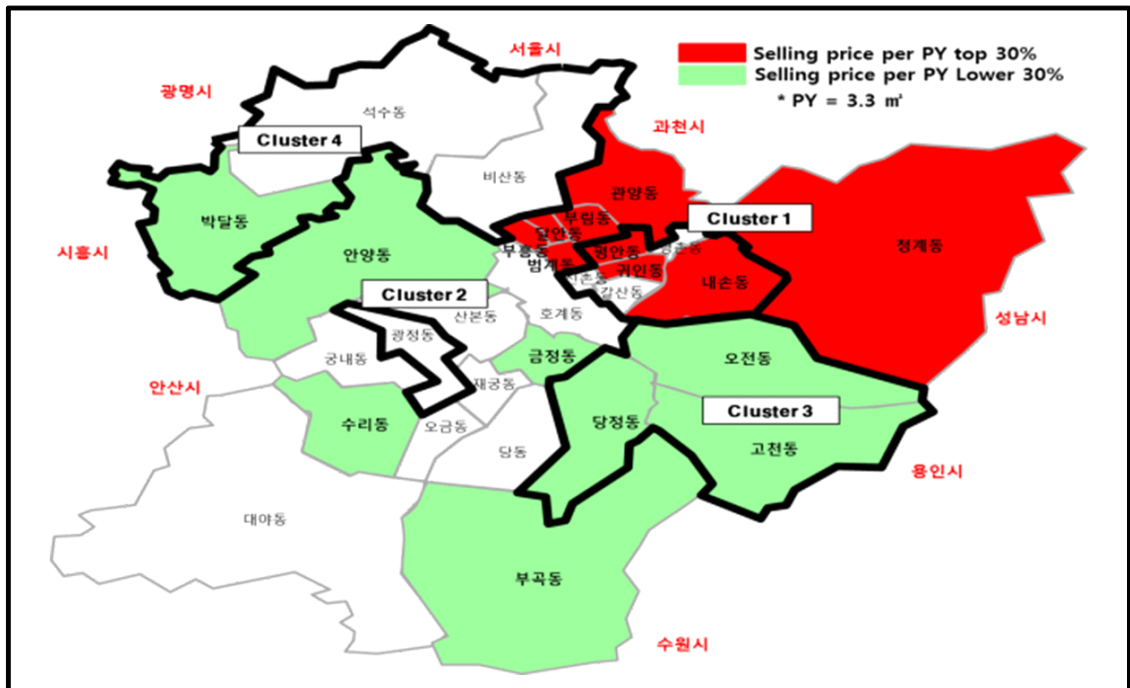
#### 4.2.3 자동차 등록비율로 본 지역 특성

세대 수 대비 자동차 등록비율과 비교한 결과, 군집1은 승용차 비중이 상대적으로 높고, 군집3

은 차량 유형에 상관없이 등록비율이 높게 나타났는데 특히 승합차와 화물차의 등록비율이 높다. 이는 자영업자 비중이 높은 것으로 추정할 수 있다.

#### 4.2.4 아파트 평당 매매가로 본 지역 특성

4개 군집지역의 경제력을 유추하기 위해 아파트 평당 매매가격을 비교하였다. 군집1 지역의 아파트시세가 평형에 상관없이 전반적으로 높았고, 군집3은 전반적으로 거래가격이 낮았다. <Figure 3>과 같이 동별 평당 매매가격 상, 하위 30%를 표시해 본 자료에도 나타나듯이 군집1에 해당하는 동은 대부분 상위에 포함되었고 군집3에 해당되는 동들은 전부 하위에 포함되었다.



<Figure 3> Apartment Transaction Price Top and Lower 30%

### 4.3 결합유형별 해지율과 지역적 특성 변수의 연관성 분석

#### 4.3.1 군집간 분산분석 결과

분산분석의 목적은 군집분석에서 도출된 군집들 간에 지역적 특성을 나타내는 데이터들의 유의한 차이가 있는지 확인하는 것이다. 만일 위치 정보와 결합유형별 해지율을 이용해 도출한 군집들 간에 지역적 특성을 나타내는 값들의 유의한 차이가 없다면, 군집과 지역적 특성과는 연관성이 없다고 할 수 있을 것이다.

먼저 각 군집의 연령대별 비중이 유의한 차이가 있는지 확인하기 위해 각 연령층 비중과 군집간 분산분석을 하였다. 미취학층의 귀무가설( $H_0$ )을 ‘미취학층 비중이 군집별로 차이가 없다’로, 대립가설( $H_1$ )은 ‘미취학층 비중이 군집별로 차이가 있다’로 설정하고 분산분석을 실시한 결과, p-value가 0.10으로 미취학층 비중은 군집별로 차이가 없는 것으로 나타났다. 그러나 학생층, 청년층, 중년층, 장년층, 노년층의 경우 각각 가설을 설정하고 분산분석을 실시한 결과, <Table 6>과 같이 미취학층을 제외한 모든 연령층에서 연령대 비중이 군집별로 유의한 차이가 있는 것으로 나타났다.

<Table 6> Analysis of Variance (Age Group Rate)

Age group	F test value	p-value
1. Preschooler (0~7)	2.11320	0.10101
2. Student (8~19)	15.52414 ***	0.00000
3. Youth (20~35)	8.29273 ***	0.00003
4. Middle(36~55)	2.74191 *	0.04305
5. elderly (56~69)	20.35403 ***	0.00000
6. Old age (70~)	16.54918 ***	0.00000

+:p<0.1, \*:p<0.05, \*\*:p<0.01, \*\*\*:p<0.001

다음으로 동단위별 평형대별 평당매매단가가 유의미한 데이터인지 확인하기 위해 각 평형대별로 군집간 분산분석을 실시하였다. 귀무가설과 대립가설은 연령대별 분석과 동일한 방식으로 설정하였으며, 분석한 결과, <Table 7>에 보이는 바와 같이 평형대별 구분 없이 전 평형대에서 아파트별 평당 매매단가가 군집에 따라 유의한 차이가 있는 것으로 나타났다. 본 연구에서 지역적 특성을 나타내는 변수 중 하나로 선택한 자동차 등록비율은 동별 데이터의 수가 부족하여 군집간 분산분석을 실시할 수 없었다. 예를 들어 연령층의 경우, 각 동별 연령대 비중이 안양동에 대해 안양1동, 안양2동, 안양3동까지 세분화되어 있는 반면, 자동차 등록비율은 안양동에 대해서만 있어 전체 데이터 수가 분석의 요건을 만족하지 못했다.

<Table 7> Analysis of Variance (Apartment Transaction Price)

Area Size	F test value	p-value
1. 20PY under	711.04567 ***	0.00000
2. 20PY ~ 29PY	847.40987 ***	0.00000
3. 29 PY ~39PY	436.01472 ***	0.00000
4. 40 PY ~.....	75.52477 ***	0.00000

PY = 3.3m<sup>2</sup>, +:p<0.1, \*:p<0.05, \*\*:p<0.01, \*\*\*:p<0.001

#### 4.3.2 결합유형별 해지율과 지역적 특성 변수 간 상관분석 결과

상관분석의 목적은 각 동의 결합유형별 해지율과 지역적 특성변수 간에 상관관계가 있는지를 세부적으로 살펴보는 데 있다. 군집분석의 결과가 상관관계에 영향을 미치는지 보기 위해, 전체 동을 대상으로 한 상관관계 분석결과와 군집으로 묶여진 13개 동을 대상으로 한 분석결과

를 비교하였다. <Table 8>은 전체 동을 대상으로 한 상관분석 결과이다. 일반적으로 상관계수가 0.2이상 0.4미만이면 약한 상관관계가 있다고 볼 수 있고, 0.4이상 0.7미만은 상관관계가 상당히 있다고 볼 수 있다(Nam and Kim, 2011). 이에 따르면 전체 해지율은 지역 특성 변수와 유의미한 상관관계가 없는 것으로 나타났다. 반면 해지율을 결합유형별로 세분화하여 분석한 결과에서는 유선결합, 미결합에서 상관관계들이 나타났다. 특히, 미결합 해지율은 전체 17개 변수 중에서 14개 변수와 0.2 이상의 상관관계를 보였다. 이는 해지율 분석을 세분화해서 보는 것이 의미가 있음을 보여준다. 군집으로 묶여진 13개 동에 대

해 상관관계를 분석한 <Table 9>를 보면 전체 해지율, 휴대폰결합해지율, 유선결합해지율은 자동차 등록비율 변수와 상관관계가 있고 다른 변수와는 상관 관계가 크게 나타나지 않은 반면 미결합 해지율은 3가지 대표 독립변수와 전부 상관관계가 있는 것으로 나타났다. 이 결과를 <table 8>의 결과와 비교해보면 군집으로 묶인 동에서의 상관관계가 더욱 높은 것을 알 수 있으며, 이는 군집화된 동들에서 결합유형별 해지율과 지역특성 변수간의 상관관계가 더 높음을 의미한다. 이 결과를 분석함으로써 보다 효과적인 마케팅 방안을 도출할 수 있을 것이다.

<Table 8> Correlation Coefficients on Every Region

			Termination Rate			
			All	Cellphone	Wire	None
A G E	Ratio of age group	1. Preschooler (0~7)	-0.067	0.046	0.305 *	0.117
		2. Student (8~19)	0.106	-0.178	0.162	0.401 **
		3. Youth (20~35)	-0.026	0.058	-0.079	-0.233 +
		4. Middle(36~55)	0.077	-0.224+	0.151	0.383 *
		5. elderly (56~69)	-0.037	0.149	-0.335 *	-0.359 *
		6. Old age (70~)	-0.193	0.324 *	-0.102	-0.506 **
	<b>Average Age</b>		<b>-0.130</b>	<b>0.164</b>	<b>-0.334 *</b>	<b>-0.490 **</b>
Vehicle registration rate per household	1. Sedan	-0.057	-0.237 +	-0.090	0.110	
	2. Van	0.005	0.168	-0.287 +	-0.319 *	
	3. Truck	-0.251 +	-0.143	-0.405 **	-0.424 **	
	4. Special	-0.232 +	-0.254 +	-0.193	-0.252 +	
	<b>All</b>	<b>-0.183</b>	<b>-0.267 +</b>	<b>-0.316 *</b>	<b>-0.163</b>	
Apartment transaction price per 3.3m <sup>2</sup>	1. 20PY under	0.086	-0.079	-0.108	0.319 *	
	2. 20PY ~ 29PY	0.154	-0.066	0.034	0.366 *	
	3. 29PY ~ 39PY	0.217 +	-0.076	0.030	0.338 *	
	4. 40PY ~	0.281 +	-0.063	-0.007	0.407 **	
	<b>All</b>	<b>0.212 +</b>	<b>-0.006</b>	<b>0.099</b>	<b>0.410 **</b>	

+ :|r|>0.2, \* :|r| >0.3, \*\* :|r|>0.4, \*\*\* :|r|>0.7

〈Table 9〉 Correlation Coefficients on Clustered Regions

			Termination Rate			
			All	Cellphone	Wire	None
A G E	Ratio of age group	1. Preschooler (0~7)	0.187	0.487 **	0.540 **	0.180
		2. Student (8~19)	0.061	-0.187	-0.016	0.457 **
		3. Youth (20~35)	0.024	0.082	0.015	-0.296 +
		4. Middle(36~55)	0.184	0.009	0.149	0.545 **
		5. elderly (56~69)	-0.119	-0.146	-0.189	-0.470 **
		6. Old age (70~)	-0.413 **	0.277 +	-0.286 +	-0.748 ***
	<b>Average Age</b>		<b>-0.241 +</b>	<b>-0.024</b>	<b>-0.280 +</b>	<b>-0.601 **</b>
Vehicle registration rate per household	1. Sedan	-0.294 +	-0.509 **	-0.422 **	-0.020	
	2. Van	-0.112	0.129	-0.228 +	-0.406 **	
	3. Truck	-0.321 *	-0.129	-0.346 *	-0.506 **	
	4. Special	-0.344 *	-0.154	-0.261 +	-0.455 **	
	<b>All</b>	<b>-0.419 **</b>	<b>-0.453 **</b>	<b>-0.540 **</b>	<b>-0.336 *</b>	
Apartment transaction price per 3.3m <sup>2</sup>	1. 20PY under	-0.084	-0.266 +	-0.263 +	0.324 *	
	2. 20PY ~ 29PY	0.013	-0.242 +	-0.060	0.458 **	
	3. 29PY ~ 39PY	0.001	-0.360 *	-0.138	0.368 *	
	4. 40PY ~	0.123	-0.314 *	-0.098	0.552 **	
	<b>All</b>	<b>0.098</b>	<b>-0.195</b>	<b>0.066</b>	<b>0.528 **</b>	

+ : |r| &gt; 0.2, \* : |r| &gt; 0.3, \*\* : |r| &gt; 0.4, \*\*\* : |r| &gt; 0.7

#### 4.3.3 결합유형별 해지율과 지역적 특성 변수 간 회귀분석 결과

회귀분석의 목적은 변수들 간의 인과관계를 분석하고, 모형을 통해 독립변수로 종속변수를 예측하는 것이다. 지역적 특성 변수의 값으로 해지율을 설명하고 예측할 수 있다면, 마케팅 계획을 세우는데 더 많은 도움을 줄 수 있을 것이다. 여기서는 지역적 특성 변수가 결합유형별 해지율에 미치는 영향을 분석하고자 하였다. 상관관계 분석은 군집으로 분류된 동들과 전체동을 구분하여 수행하고 둘을 비교하였으나, 회귀분석에서는 군집으로 분류된 동의 수가 너무 적어 유

의한 결과를 얻기 어렵다고 판단되어, 우선 전체동에 대해 지역적 특성변수와 해지율 간의 단순회귀분석을 실시하였다. 또한 상관분석 결과, 결합유형 전체와 무선결합의 경우, 해지율과 지역 특성 변수 간의 상관관계가 너무 낮아 회귀분석을 실시하지 않았다. <Table 10>은 결합유형 중, 유선결합과 미결합에 대한 회귀분석 결과를 보여주며, 미결합 해지율만 전체 평균연령, 아파트 평당 매매가와 관계가 있었다. 다음은 단순회귀분석과 마찬가지로 유선결합과 미결합에 대해 중회귀분석을 실시하였다. 먼저 세 지역특성 변수만으로 회귀분석을 실시한 결과, <Table 11>과

〈Table 10〉 Simple Regression Analysis Results

Independent Variable	Wire		None	
	F	(Adjusted R <sup>2</sup> )	F	(Adjusted R <sup>2</sup> )
Average Age	3.395 +	0.079	8.5 **	0.212
Vehicle	2.991 +	0.066	0.739	-0.009
Apartment	0.267	-0.027	5.451 *	0.137

+ : p &lt; 0.1, \* : p &lt; 0.05, \*\* : p &lt; 0.01, \*\*\* : p &lt; 0.001

〈Table 11〉 Multiple Regression Analysis Results

Regression Model	Variable	Wire (B)	None (B)
Model1 : 3 variables	Average Age	-0.0533	-0.2443
	Vehicle	-0.4936	-0.9744
	Apartment	-0.0002	0.0005
	F	2.528 +	3.079 *
	Adjusted R <sup>2</sup>	0.1407	0.1822
Model2 : 3 variables with dummy (clustered=1, none=0)	Average Age	-0.0542	-0.2311
	Vehicle	-0.4162	-2.1469
	Apartment	-0.0002	0.0005
	dummy	-0.0395	0.5976
	F	1.91	3.033 *
	Adjusted R <sup>2</sup>	0.1150	0.2251

+ : p &lt; 0.1, \* : p &lt; 0.05, \*\* : p &lt; 0.01, \*\*\* : p &lt; 0.001

같이 유선결합은 p-value가 0.08로 0.05 기준으로 유의하지 않았으나, 미결합에 대해서는 0.046으로 모형이 유의하게 나왔다. 다음으로 군집의 효과를 보기 위해, 각 동에 대해 군집으로 분류된 경우는 1, 그렇지 않은 경우는 값을 0으로 할당한 더미변수를 추가하여 회귀분석을 실시하였다. 그 결과, 유선결합은 여전히 유의하지 않게 나왔으나, 미결합의 경우는 p-value 0.037, 설명력 22.5%로 군집여부가 없는 모형에 비해 더 나은 결과를 보였다.

#### 4.4 분석 결과 토의 및 마케팅 방안 도출

##### 4.4.1 분석 결과 요약

분석을 위해 먼저 공간검색통계 도구인 SatScan을 활용하여 4개의 유사군집을 도출하였고 군집별 특성을 분석하였다. <Table 12>는 각 군집별로 지역특성을 정리한 결과이다. 각 군집의 특성을 요약하면, 군집1로 도출된 지역은 평촌신도시 지역으로 미결합비율이 낮지만 미결합 고객의 해지율이 높아서 전체 해지율이 평균보

〈Table 12〉 Cluster Feature

		cluster1	cluster2	cluster3	cluster4
Combination rate	Cellphone		High	Low	
	Wired goods		Low	High	
	None	Low	High	High	Low
Terminations rate	Cellphone		High		
	Wired goods		Low	Low	High
	None	High	Low	Low	High
	All		Low		High
Ratio of age group	Preschooler (0~7)		High		
	Student (08~19)	High	High		
	Youth (20~35)	Low			
	Middle (36~55)	High			
	elderly (56~69)	Low	Low	High	
	Old age (70~ )	Low	High		
Vehicle registration rate per household	Sedan	High		High	
	Van	Low	Low	High	High
	Truck	Low	Low	High	
	Special	Low	Low	High	Low
	Apt. transaction price per 3.3m <sup>2</sup>	High		Low	

High : higher than 7% points of the average, Low: lower than 7% points of the average

다 높다. 지역 특성은 타 지역 대비 아파트 매매 가격, 승용차 등록 비율이 높은 비교적 소득수준이 높고 학생자녀가 있는 중년층 가정으로 유추할 수 있다. 군집2는 1개동으로 특별한 특징은 찾기 어렵고, 군집3은 의왕시에 위치한 3개동으로 미결합율은 높지만 해지율은 평균보다 낮다. 지역 특성은 장년층(55~65세) 비율이 제일 높고 승합차, 화물차 등록 비율이 상대적으로 높다. 아파트 매매가격이 가장 낮은 지역으로 생계형 자영업자 비율이 높을 것으로 추정된다. 군집4는 안양시 만안구에 위치한 3개동으로 미결합비율이 낮지만 미결합 고객의 해지율이 높아 전체 해지율이 제일 높게 나타났다. 이는 미결합 비율이 낮더라도 지역적 특성에 따라 결합 및 유지마케

팅이 필요하다는 의미이다.

다음으로, 결합유형별 해지율과 지역적 특성 데이터의 연관성을 분석하기 위해 분산분석, 상관분석, 회귀분석을 하였다. 먼저, 군집들 간에 지역적 특성을 나타내는 데이터들의 유의한 차이가 있는지 확인하기 위해 분산분석을 하였다. 그 결과 미취학층을 제외한 전 연령대별 비중 및 모든 평형대의 아파트 평당 매매가가 유의한 차이가 있는 것으로 나타났다. 이는 결합유형별 해지율과 위치정보를 통해 구한 군집들 간에 지역적 특성의 차이가 있는 것으로 해석할 수 있으며, 해지율과 지역적 특성 간의 연관성이 있는 것으로 볼 수 있다.

둘 사이의 연관성을 보다 상세히 확인하기 위

해 각 동의 결합유형별 해지율과 지역적 특성변수 간의 상관관계를 분석하였다. 군집화의 효과를 확인하기 위해, 전체 동과 군집으로 묶여진 13개 동의 두 모형으로 나누어 상관관계를 비교하였다. 그 결과, 군집으로 묶인 동에서의 상관관계가 더욱 높게 나타났는데 이는 군집화된 동들에서 결합유형별 해지율과 지역특성 변수간의 상관관계가 더 높음을 의미한다고 볼 수 있다. 또한 지역적 특성 변수는 전체 해지율, 휴대폰결합 해지율과는 상관관계가 나타나지 않았으나, 유선결합 해지율, 미결합 해지율과는 비교적 높은 상관관계를 보임으로써, 결합유형에 따른 분석이 효과가 있음을 알 수 있었다.

마지막으로 지역적 특성 변수가 결합유형별 해지율에 미치는 영향에 대한 회귀분석의 결과, 미결합 해지율에서 지역특성 변수의 영향을 확인할 수 있었다. 더미변수를 추가하여 분석한 결과에서는 군집화 여부에 대한 더미변수가 있는 모형이 없는 모형에 비해 더 나은 결과를 보여 군집의 효과가 있는 것으로 분석되었다.

#### 4.4.2 마케팅 적용 방안

여기서는 이상의 분석결과를 바탕으로 마케팅 방안을 제시하고자 한다. <Table 13>은 군집특성에 따른 다양한 마케팅 방안의 예시이다. 예를 들어, 군집1과 군집4의 경우 해지로 연결되는 미결합 비중이 낮지만 미결합 고객의 이탈률이 높기 때문에 휴대폰 또는 IPTV와 결합을 최대한 유도하는 마케팅이 필요하다. 특히, 군집1은 학생층이 많으므로 입학/졸업시즌 판매이벤트, 학원과 연계한 특별수강 이벤트 등 학생대상의 마케팅이 효과적일 것이며, 비교적 고가의 아파트 단지로 형성되어 있으므로 아파트 단지 내 시연을 통해 구매를 유도할 수도 있다. 군집2, 군집3의 경우 미결합 비중이 높음에도 해지율은 낮다. 하지만 언제 이탈 될지 모르기 때문에 결합 마케팅은 필요하다. 군집2의 경우 미취학 아동 비율이 높으므로 유아용 콘텐츠로 마케팅하면 효과적일 수 있다. 군집3은 아파트 매매가격이 낮은 단지로 형성되어 있고 특히 승합차, 화물차 등록 비율이 높은 생계형 자영업자가 많을 것으로 추

<Table 13> Examples of Regional Marketing

Cluster	Feature	example
1	· 높음 : 미결합 해지율, 학생층/중년층, 승용차등록비중, 아파트평당매매가 · 낮음 : 미결합비중, 청년/장년/노년층	- 결합유도 마케팅 - 입학/졸업 시즌 이벤트 - 학원 연계 마케팅 - 아파트 현장 시연회 등
2	· 높음 : 휴대폰결합비중, 미결합비중, 휴대폰결합분 해지율, 미취학/학생/노년층 · 낮음 : 유선결합비중, 해지율	- 유아용/초등생용 콘텐츠 마케팅 - 경로당 휴대폰 마케팅 등
3	· 높음 : 유선결합비중, 미결합비중, 장년층, 승용차/승합차/화물차 비중 · 낮음 : 휴대폰결합비중, 아파트 평당매매가	- 차량관련 업체 연계 (ex. 인터넷가입시 타이어교환 등) - 아파트 현장 시연회 (현금마케팅, 휴대폰 연계 판매) 등
4	· 높음 : 해지율, 승합차 비율 · 낮음 : 미결합비중, 특수차 등록비중	- 리텐션 및 결합유도 마케팅 - 신규 가입자 유치 마케팅 등



정되므로 현금 마케팅이 및 차량에 관련된 사은품 마케팅이 유효할 것으로 보여진다. 군집4는 지역 데이터로 봐서는 별다른 특징은 없으나 결합유형에 상관없이 전반적으로 해지율이 높기 때문에 기존 가입자 관리 및 및 신규 고객유치 마케팅이 시급하다.

## 5. 결론

본 연구는 초고속인터넷 신규가입 및 고객 관리 시에 단편적인 마케팅에서 벗어나 지역별 특성을 고려한 좀더 효율적인 마케팅 방안을 마련하기 위하여, 결합유형별 해지데이터와 지역특성 데이터 사이의 연관성을 분석하였다. 지역특성을 보다 세부적 및 구체적으로 살펴보기 위하여 기존 연구의 주 분석단위인 시군구 단위보다 세분화된 동단위로 연관성을 분석하였다. 먼저 검색통계도구인 SatScan을 이용하여 결합유형별 해지율과 위치정보에 대해 유사한 유형을 보이는 군집을 도출한 뒤, 군집화 결과를 반영하여 결합유형별 해지율과 연령대별 비중, 차량 등록 비율, 아파트 평당 매매가의 지역별 특성 데이터의 연관성을 분석하였다. 분산분석 결과 군집들 간에 지역별 특성 변수가 유의한 차이가 있음을 확인하였으며, 상관분석을 통해 전체 동에 비해, 군집화된 동들에 대해 결합유형별 해지율과 지역특성 변수 사이에 더 강한 상관관계가 있음을 확인하였다. 또한 회귀분석을 통해 미결합 해지율과 지역특성변수간의 연관성을 분석하였다. 결과적으로 지역적 특성에 따라 해지율의 차이가 있고 이를 이용해 지역별로 차별화된 타겟 마케팅이 가능하다는 결과를 얻었으며, 마지막으로 분석결과를 기반으로 적절한 마케팅방안을

제시하고자 하였다.

본 연구는 지역적 특성을 분석하기 위해서 공간정보 기반의 클러스터링을 실시했다는 점에서 학술적 기여를 갖는다고 할 수 있다. 주로 질병이나 범죄 등과 같이 지역적인 연관성을 갖는 요인 등을 대상으로 사용한 공간검색통계량을 마케팅에서의 고객분석에 적용하였으며, 단순히 군집을 나누어 비교하는 것으로 그치지 않고, 해지율에 영향을 미치는 요인을 찾기 위하여 상관관계분석과 회귀분석과 같은 다양한 통계적 분석을 적용하고자 하였다. 실무적인 기여로는 A 통신사가 보유한 실제 데이터와 통계청 등으로부터 수집한 지역 특성 데이터를 결합하여 최대한 의미 있는 결과를 끌어내고자 했다는 점을 들 수 있다. 또한 분석결과를 이용하여, 실제 환경에서 활용 가능한 마케팅 방안을 제시하고자 하였다.

본 연구의 가장 큰 한계로, 연구를 수행하기에 충분한 데이터를 확보하기가 어려웠다는 점을 들 수 있다. 특히 지역특성을 나타낼 수 있는 변수를 동단위로 구하기가 어려웠다. 통계청이나 기타 국가기관에서 제공하는 자료도 동별로 제시된 데이터가 거의 없었다. 즉, 대부분의 데이터들이 동 단위가 아닌 구 단위까지만 공개되어 세부적으로 분석하는 것에 한계가 있었으며, 지역의 생활수준을 알 수 있는 소득, 카드사용정보, 외제차등록건수 등의 자료는 공개되지 않았고 해지원인에 영향을 줄 수도 있는 통신사의 정책이나 특성도 당시 자료를 구하지 못했다. 향후 연구에서 가장 시급한 부분은 이러한 지역별 특성 변수에 대한 동단위 데이터를 확보하는 것으로 생각된다.

향후 연구 방향은 본 연구의 분석결과를 기반으로 연령대, 소득 수준 등 지역별 특성에 맞는

마케팅을 실시하는 것이다. 타겟 마케팅 전후의 결과 차이를 비교 분석함으로써 마케팅의 효과를 분석하는 것도 유의미한 작업으로 판단된다. 또한, 본 연구에서는 해지데이터를 결합유형으로만 분류했지만 계약상태(약정만료, 약정미도래, 미약정 등)로 분류 할 수도 있고 동 단위에서 좀 더 세분화하여 주택, 아파트, 상가 등으로 분석할 수 있을 것이다. 해지데이터뿐 아니라 기존 고객 분포 및 가입데이터를 기반으로 군집을 구성하여 마케팅에 활용하는 것도 효과적인 것으로 보여진다.

## 참고문헌(References)

- Andrea Freyer Dugas, Mehdi Jalalpour, Yulia Gel, Scott Levin, Fred Torcaso, Takeru Igusa, Richard E. Rothman, “Influenza Forecasting with Google Flu Trends”, PLoS One, Vol. 8, Issue 2, (2013)
- Arvind B.Bambhroliya, Keith D.Burau, and Ken sexton, “Spatial Analysis of County-level Breast Cancer Mortality in Texas”, Journal of Environmental and Public Health, Vol.2012, A ID959343, 8.
- Assessment of Telecommunication Market Competition 2015, Korea Information Society Development Institute, (2015).
- Digital Media Convergence Company, “Analysis of 2016 Internet Usage Survey and Advertising Effect”, 2016
- Han,J.H., Lee,M.J., “Cancer cluster detection using scan statistic” , Korean Data Information Science Society, Vol.27, No.5,(2016), 1193-1201.
- Hong,H.P., “Detection of hotspot areas using spatial clustering methods and echelon analysis”, Statistical Research Institute, Vol.8, No.2,(2003).
- Hwang.S.S, “Spatiotemporal analysis of the association between socio-environmental factors and suicide using open source software”, FOSS4G(Free Open Source SW for GeoSpatial) Korea 2012 conference.
- Jae,M.K., Jeon,H.R., “The Intention to switch Broadband Internet Service Providers”, Korea Association of Human Ecology, Vol.19, No.5, (2010).
- Ji,B.G., “Explore Hazardous Area of Varicella Incidence in Korea”, Chungbuk National University, (2015)
- Jung I, Kulldorff M, Klassen A., A spatial scan statistic for ordinal data. Statistics in Medicine, 26: , (2007),1594-1607
- Jung I, Kulldorff M, Richard OJ., A spatial scan statistic for multinomial data. Statistics in Medicine, epub, (2010)
- Kim,H.K., Kim,D.M., “The Study on the Bundle Service and Competitive Restraints of Broadcasting & Telecommunication”, Korean Association for Broadcasting & Telecommunication, Vol.25, No.6, (2011), 248-289.
- Kim,M.G.,Park,J.H.,Baek,J.H., “Characteristics and Influencing Factors of High-speed Internet Customer churn Behavior”, Korean Institute of Industrial Engineers, 2008.05, (2008), 655-661.
- Kim,S.H., “Recent Issues in the Regulation of Bundled Communications Services”, Law & Economic Regulation, Vol.8, No.2, (2015), 122-136.

- Kim,S.H., Switching Costs of the Consumers Using Bundled Telecommunication Services, Korea Academic Society of Industrial Organization, Vol.23, No.4, (2015).
- Kim,T.I., Kim,M.S., “Sources of the customer loyalty in the high-speed Internet service”, Korea Academia Industrial cooperation Society, Vol.12, No.2, (2011), 728-735
- Kim,W.S., Park,M.S., “Analysis of Competition in the Market for Bundling of Broadband Internet and Pay-TV Services”, Korean Association for Broadcasting & Telecommunication Studies, No. winter, (2010), 119-146.
- Kim. L.Y., Seo. W.S., “Investigating Spatial Patterns and Urban Influential Factors of the School-Age Population using Spatial Econometric Analysis”, Korea Regional Development Association, Vol.28, No2 (2016), 113~130
- Korea Information Society Development Institute, “A Study on 2016 Communications Market Forecast and Policy”, KCC-2015-43, (2015).
- Korea Internet & Security Agency, “2016 Internet Usage Survey”, 2016
- Korea National Spatial Data Infrastructure Portal, “What is spatial information?”, Available at <http://www.nsdi.go.kr/?menuno=2759> (Accessed 23 June, 2017)
- Lee,C.S., “Analysis of the Relative Importance for Factors to Select Provider in High-Speed Internet Service”, e-business studies, Vol.12, No.2, (2011), 185-198.
- Lee,D.J., Kim,M.S., “Relationship between customer satisfaction and re-contract intention in the ISP industry: The role of the level of responses on customer complaints”, International Trade & Commerce, Vol.7, No.4, (2011), 385-400.
- Lee,K.J., Hwang,M.H., Han,S.H., Yang,E.J, “The first step to understanding and utilizing spatial statistical analysis”, Korea Research Institute of Human Settlements, (2015), 121~P131.
- Lee,S.H., Kim,H.C., Park,M.S.,” The Effects of Free Offers in Bundled Products: A Case of Korean Broadcasting and Telecommunication Market”, Korea Academic Society of Industrial Organization, Vol.24, No.1, (2016).
- Lee,Y.R., ”Improvement direction of High speed internet consumer compensation system”, Korea Information Society Development Institute, Vol.21, No.1, (2009), 1-22.
- Lee. J. E., Kim. K. S., “The Effect of Residential Environment Characteristics on Suicide Rate in Korea”, Korea Association For Housing Policy Studies, First half seminar presentation, (2016), 21~31
- Lee. Y. Y., “A Study on Regional Classification through Regional Characteristics Analysis (Focused on 31 Cities in Gyeonggi Province )”, Residential Environment institute of Korea, Vol12, No1(2014), 129~143
- Ministry of Science, ICT and Future Planning, “Number of Mobile and High Speed Internet subscribers”, Available at <http://www.msip.go.kr/web/msipContents/contents.do?mId=MTQ2> (Downloaded 3 June, 2017).
- Nam,C.H., Kim,H.J., “Fundamentals of Statistics”, ShinYoungSa, 2011
- Park,J.H., Lee,J.H., “A study on the churn intention and customer’s retention strategy of High-speed Internet Service”, Korea Institute of Communication Sciences, 2002.11, (2002),

1224-1227.

Park,T.J., Min,B.G., “Determinants of Customer Loyalty and Switching Behavior in High-speed Internet Service: A Qualitative Approach”, Korean Internet E-Commerce Association, Vol.12, No.12, (2012), 189-204.

Park.J.O, I.S.Choi and M.H.Na, “A Study on the

effects of air pollution on circulatory health using spatial data”, J Korea Soc Qual Manag, Vol.44, No.3, (2016), 677~688.

Pervaiz F, Pervaiz M, Abdur Rehman N, Saif U, “early epidemic detection from Google flu trends”, Journal Of Medical Internet Research [J Med Internet Res], Vol.14 (5), (2012), e125

## Abstract

# A Study on the Regional Characteristics of Broadband Internet Termination by Coupling Type using Spatial Information based Clustering

Janghyuk Park\* · Sangun Park\*\* · Wooju Kim\*\*\*

According to the Internet Usage Research performed in 2016, the number of internet users and the internet usage have been increasing. Smartphone, compared to the computer, is taking a more dominant role as an internet access device. As the number of smart devices have been increasing, some views that the demand on high-speed internet will decrease; however, Despite the increase in smart devices, the high-speed Internet market is expected to slightly increase for a while due to the speedup of Giga Internet and the growth of the IoT market. As the broadband Internet market saturates, telecom operators are over-competing to win new customers, but if they know the cause of customer exit, it is expected to reduce marketing costs by more effective marketing. In this study, we analyzed the relationship between the cancellation rates of telecommunication products and the factors affecting them by combining the data of 3 cities, Anyang, Gunpo, and Uiwang owned by a telecommunication company with the regional data from KOSIS(Korean Statistical Information Service). Especially, we focused on the assumption that the neighboring areas affect the distribution of the cancellation rates by coupling type, so we conducted spatial cluster analysis on the 3 types of cancellation rates of each region using the spatial analysis tool, SatScan, and analyzed the various relationships between the cancellation rates and the regional data. In the analysis phase, we first summarized the characteristics of the clusters derived by combining spatial information and the cancellation data. Next, based on the results of the cluster analysis, Variance analysis, Correlation analysis, and regression analysis were used to analyze the relationship between the cancellation rates data and regional data. Based on the results of analysis, we proposed appropriate marketing methods according to the region. Unlike previous studies on regional characteristics analysis, In this study has academic

---

\* Graduate Program in Technology Policy, yonsei University

\*\* Department of Management Information Systems, Kyonggi University

\*\*\* Corresponding Author: Wooju Kim

Department of Information and Industrial Engineering, Yonsei University

50 Yonsei-ro Seodaemun-gu, Seoul 03722, Korea

Tel: +82-2-2123-4010, Fax: +82-2-364-7807, E-mail: wkim@yonsei.ac.kr

differentiation in that it performs clustering based on spatial information so that the regions with similar cancellation types on adjacent regions. In addition, there have been few studies considering the regional characteristics in the previous study on the determinants of subscription to high-speed Internet services, In this study, we tried to analyze the relationship between the clusters and the regional characteristics data, assuming that there are different factors depending on the region. In this study, we tried to get more efficient marketing method considering the characteristics of each region in the new subscription and customer management in high-speed internet. As a result of analysis of variance, it was confirmed that there were significant differences in regional characteristics among the clusters, Correlation analysis shows that there is a stronger correlation the clusters than all region. and Regression analysis was used to analyze the relationship between the cancellation rate and the regional characteristics. As a result, we found that there is a difference in the cancellation rate depending on the regional characteristics, and it is possible to target differentiated marketing each region. As the biggest limitation of this study and it was difficult to obtain enough data to carry out the analyze. In particular, it is difficult to find the variables that represent the regional characteristics in the Dong unit. In other words, most of the data was disclosed to the city rather than the Dong unit, so it was limited to analyze it in detail. The data such as income, card usage information and telecommunications company policies or characteristics that could affect its cause are not available at that time. The most urgent part for a more sophisticated analysis is to obtain the Dong unit data for the regional characteristics. Direction of the next studies be target marketing based on the results. It is also meaningful to analyze the effect of marketing by comparing and analyzing the difference of results before and after target marketing. It is also effective to use clusters based on new subscription data as well as cancellation data.

**Key Words** : Spatial Scan Statistic, High Speed Internet, Internet Termination, Target Marketing, SatScan, Clustering

Received : July 21, 2017 Revised : September 19, 2017 Accepted : September 20, 2017

Publication Type : Regular Paper Corresponding Author : Wooju Kim

## 저 자 소개



### 박장혁

연세대학교에서 기술정책 박사과정을 수료하였으며, 동 대학원 산업정보경영 석사학위를 취득한 바 있다. 현재는 kt에서 지사장으로 재직 중이다. 주요 연구 관심 분야는 모바일 금융, NFC, IoT, 지능형 전자상거래 등이다.



### 박상언

KAIST에서 전산학 학사 및 경영공학 석사, 박사 학위를 취득하였으며 현재는 경기대학교 경영정보학과 교수로 재직 중이다. 주요 연구 관심 분야는 머신러닝, 딥러닝, 텍스트 마이닝, 빅데이터 분석, 지능정보시스템, 시맨틱 웹, 의미기반검색 등이다.



### 김우주

연세대학교 BBA과정 학사 학위를 취득하였고 KAIST에서 경영과학 석사, 박사 학위를 취득하였다. 현재 연세대학교 정보산업공학과 교수로 재직 중이며, 주요 연구 관심 분야는 지능형 시맨틱웹, 지식관리 및 인공지능, 빅데이터 등이다.