

# 빅데이터 표준분석모델을 활용한 CCTV우선 설치지역 도출 사례연구

성창수\*, 박주연\*\*, 가희광\*\*\*

동국대학교 기술창업학과(대학원)\*, 연세대학교 경영학과\*\*, 행정자치부 공공정보정책과\*\*\*

## The Case Study of CCTV Priority Installation Using BigData Standard Analysis Model

Chang Soo Sung\*, Joo Y. Park\*\*, Hoi Kwang Ka\*\*\*

Dept. of Technology Entrepreneurship, Dongguk University\*

School of Business, Yonsei University\*\*

Public Information Policy, Ministry of the Interior\*\*\*

요 약 본 논문에서는 공공기관의 빅데이터를 활용한 표준분석모델을 살펴보고, 실제 사례현장의 분석결과를 통해 빅데이터 표준분석모델의 적합성과 효과성을 확인하고자 한다. 특히 행정효율성을 향상시킬 수 있는 분야인 민원, CCTV 등의 빅데이터에 대한 표준분석모델을 대상으로 하였다. 이를 위해 빅데이터 표준분석 지표를 산출하고 K시의 CCTV우선설치지역 사례에 적용하여 빅데이터 표준분석모델의 정확성을 조사하였다. 빅데이터 표준분석 모델을 활용한 K시의 사례분석 결과, 우선 설치리스트 상위 지점 모두 범죄취약지수 중 환경지수 값이 전반적으로 낮게 나온 반면, CCTV미설치 지역에 따른 감시취약지수와 야간 및 심야 시간대의 유동인구 지수가 높게 나타났다. 이는 실제로 CCTV설치에 대한 민원이 높고 그 필요성을 인지하고 있는 지역으로 빅 데이터 분석결과가 높은 정확성을 나타내고 있음을 확인할 수 있다. 이러한 연구결과는 빅데이터 활용과 분석에 의한 행정 효율성의 제고 및 디지털 융합 환경에서 빅데이터 표준분석 모델의 의미있는 시사점을 제시하고 있다.

주제어 : 디지털 융합 사회, 빅데이터, 표준분석모델, CCTV우선설치, 행정 효율성

**Abstract** This study aims to investigate the public big data standard analysis model developed by Ministry of the Interior and examine its accuracy and reliability of prediction. To do this, big data standard analysis index were calculated to apply them to the real world case of CCTV monitoring system prior installation in K city. The result of this case study revealed that the areas to be installed CCTV consisted with the area where residences requested and complained to install CCTV monitoring systems, which indicated that the result of big data standard analysis model provided accurate and reliable outcomes. The result of this study suggested implications on effective exploitation of big data analysis.

**Key Words** : Digital Convergent society, Bigdata, Standard Analysis Model, CCTV prior installation index, effectiveness.

\*This study is based on the report of 'A due diligence of CCTV installation region' written in part of the 2016 Public Big Data Standard Analysis Mode project by National Information Society Agency(NIA).

Received 24 March 2017, Revised 29 April 2017

Accepted 20 May 2017, Published 28 May 2017

Corresponding Author: Joo Y, Park (Yonsei University)

Email: park3500@naver.com

© The Society of Digital Policy & Management. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1. 서론

디지털 융합시대로 접어들면서 사물과 인터넷 및 정보통신의 접목을 통해 인간의 삶과 환경에 변화를 가져오고 있다. 이로 인해 서로 다른 분야의 사물과 기기들이 정보통신기술로 융합하는 디지털 융합의 분야는 다양하게 넓어지고 있다. 정보기술과 인터넷이 기반이 된 디지털 융합사회에서는 다양한 기술들로부터 획득한 여러 형태의 데이터를 얼마나 수집하고, 어떻게 분석·활용하는가가 향후 미래사회 발전에 중요한 역할을 한다[1]. 국민의 안전과 편리한 생활 등 공공정책에 관한 행정지원 및 정책적 의사결정에 있어서도 지역주민과 사회구성원의 의견을 수렴하기 위해 빅데이터의 효율적인 활용은 중요하다. 이는 빅데이터 분석 결과가 지역주민 또는 모든 사회 구성원이 필요로 하는 정책을 수립하고 사회 발전에 기여할 수 있는 매우 중요한 역할을 할 수 있다는 것이다[2, 3]. 이에 정부와 민간 기업들은 행정시스템 및 제조업 등에 정보통신기술을 융합해서 업무효율성과 기업경쟁력을 제고하는데 노력하고 있다.

이처럼 디지털 융합사회 경제의 핵심으로 부각되고 있는 빅데이터는 디지털 환경에서 생성되는 모든 데이터를 말한다. 인터넷과 소셜미디어 등의 다양한 디지털 채널에서 생성된 데이터는 디지털융합으로 데이터의 규모가 더욱 방대해지고 지속적으로 재생산되며, 생성주기가 짧고 다양한 데이터형태를 가지고 있는 것이 특징이다. 이는 전통적인 단일 정보기술로는 빅데이터를 쉽게 처리하고 분석할 수 없는 이유이기도 하다[4, 5]. 이러한 데이터 분석을 통해 생성된 정보와 결과는 사용자, 소비자 등의 행동을 미리 예측할 수 있도록 도와주고 그들의 편익과 삶의 질 향상을 위해 정부부처와 유관기관 및 민간 기업에서 다양한 방법으로 적용되어 활용되고 있다. 컨설팅 전문업체인 맥킨지는 빅데이터가 제공하는 가치를 연구한 결과, 빅데이터는 혁신, 경쟁우위, 생산성 등에서 가치를 창출할 수 있으며, 이는 빅데이터가 창의적이고, 효율적으로 품질과 효율성을 향상시킬 수 있도록 활용되었을 때 가능하다고 한다[6,7]. 다시 말해, 방대한 데이터 자체가 혁신적이고 효과적인 결과물을 창출하는 것이 아니라, 다양한 형태의 데이터를 어떻게 모으고 분석하여 의미 있는 결과를 도출하는가에 따라 빅데이터의 가치가 달라진다는 것이다. 이처럼 빅데이터 분석 결과는 의사결

정에 중요한 영향을 미치기 때문에 이를 어떻게 처리하고 관리하는 지에 대한 빅데이터 관리가 중요한 과제로 부각되고 있다[8].

그럼에도 불구하고 현실적으로 급속히 증가하는 빅데이터를 효과적으로 처리하고 활용할 수 있는 역량의 부족과 여러 채널을 통해 얻은 데이터가 각 기관별로 중복 저장 및 분석 도구 및 방법이 상이하여 분석결과가 일관되지 못하는 문제점이 있다[9]. 따라서, 동일 데이터를 통합하고 분석모델을 표준화하여 빅데이터의 분석과 활용을 효과적으로 모색하는 것이 필요하다. 통합적인 빅데이터의 활용과 표준화된 분석결과는 데이터의 품질을 높이고 정확하고 신뢰성 있는 정보를 제공하여 업무 및 비용의 효율성과 성과향상으로 연계될 수 있다. 이를 통해 현상에 대한 예측을 정확하게 할 수 있으며 환경 및 업무 개선에 대한 의사결정을 신속하게 할 수 있게 된다.

최근 우리 정부에서도 서로 다른 공공기관에 산재되어 사용하고 있는 빅데이터를 통합한 표준분석모델을 구축하여 행정활성화 및 분석사업의 업무효율성 향상을 위해 노력하고 있다[10, 11]. 이를 통해 중앙부처 및 지자체에서 개별적으로 분석하고 활용했던 데이터를 통합함으로써 데이터수집 및 분석기간의 단축과 분석을 위한 제반시설 및 분석시스템 구축비용을 절감할 수 있는 장점을 극대화할 수 있다. 또한, 표준분석모델을 통해 수집데이터 목록, 데이터 형식, 수급절차 등을 표준화하여 자치단체별 서로 다른 데이터와 분석모델 및 분석결과와의 차이점을 극복하고 분석의 효율성과 정확성을 향상시킬 수 있다[12]. 빅데이터의 분석결과는 중소기업과 대기업 등의 의사결정 및 전략적 포지셔닝을 분석하기 위해서 사용될 뿐만 아니라 국가의 행정, 치안, 보안 등 우리 생활과 삶의 안정과 향상을 위해서도 활용되고 있다. 특히, 빅데이터 분석모델을 활용한 범죄 분석과 치안수요 분석 등 예측기술 향상과 대응능력 극대화에 대한 연구가 활발하게 진행되고 있다[13]. 하지만, 빅데이터의 데이터베이스 구축이 기존 행정업무의 흐름에 따라 이루어져 있어 데이터의 분산과 중복 및 상호 호환성이 낮으며 데이터 통합이 어려운 실정이다. 또한 여러 선행연구에서 산재되어 있는 데이터의 분석결과를 즉시적으로 상호 전달, 연계할 수 있는 시스템이 부족하기 때문에 치안수요의 예측과 범죄예방의 대응에 대한 즉시성 및 정확성이 낮다고 지적하고, 장기적인 관점에서 공공데이터의 생산과

통합적인 관리의 필요성을 제언하고 있다[14, 15, 16].

이에 본 논문에서는 공공기관의 빅데이터를 활용한 표준분석모델을 살펴보고, 실제 사례의 분석결과를 통해 빅데이터 표준분석모델의 적합성과 효과성을 확인하고자 한다. 이를 위해 행정효율성을 향상시킬 수 있는 분야인 민원, CCTV 등의 빅데이터에 대한 표준분석모델을 대상으로 하였으며, 특히 K시의 CCTV우선설치지역 도출사례를 토대로 공공 빅데이터 표준분석모델의 정확성과 타당성을 조사한다. 또한 본 연구 결과를 통해 빅데이터 활용과 분석의 효율성을 제고하고 국민안전과 생활의 편의성 향상에 대한 방향성을 모색하고자 한다.

## 2. 빅데이터 표준분석 모델

### 2.1 데이터 수집

데이터는 교육부, 행정자치부, 사이버경찰청 등 분산되어 있던 개방되어 공개된 공공데이터를 사용하였다. 데이터는 도로명주소, 거주인구, 주택정보, 상가정보, 격자데이터 등이 포함된다. CCTV 설치요청에 관한 민원데이터 및 등록외국인 데이터는 해당 지자체의 데이터를 활용하였다. 또한, 유동인구 데이터는 이동통신사의 유료로 제공되는 데이터를 확보하여 사용하였다.

<Table 1> The List of Data

Organizations	Collected Data
Local governments	Installation Status
	Installation demands by residents
	Operation management
Disaster Management Research Institute	Status of streetlight/security light
National Police Agency	Status of crime prevention/safe facilities
Ministry of Education	Facilities for children
Ministry of the Interior	Name of street
Ministry of Land, Infrastructure & Transport	Grid CELL
Ministry of the Interior	Population
Ministry of Land, Infrastructure and Transport	Housing information

이는 범죄취약지수를 산출하기 위해 공공데이터에 포함되지 않은 소셜네트워크링, 지도, 이미지 등의 비정형 데이터를 활용하여 지역주민의 이동 및 유동인구의 움직임을 파악하기 위함이다. 또한 빅데이터 분석은 다양한 채널을 통해 획득한 데이터, 즉 정형화된 데이터와 함께 전체 데이터의 80%이상을 차지하는 비정형 데이터를 모두 포함했을 때 신뢰성을 확보할 수 있다는 기존 연구에 따라 지수산출에 필요한 데이터를 이동통신사로부터 확보하였다[17]. 본 연구의 데이터 분석모델에 사용하기 위해 수집된 데이터와 제공기관은 <Table 1>과 같다.

### 2.2 데이터 표준화

빅데이터 표준분석모델을 구축하기 위해 먼저 유사 데이터에 대한 표준화가 요구된다. 따라서 CCTV설치에 대한 데이터의 표준화는 다음과 같은 과정으로 진행되었다. 첫째, 기존의 선행분석 사례를 검토하고 분석에 활용된 데이터목록을 정리한 후 데이터에 대한 필수항목, 소유자, 수급문제 또는 다른 데이터와의 대체가능성 등을 점검하였다. 둘째, 선행연구 전문가와의 인터뷰 등을 통해 분석에 사용할 표준 데이터 목록을 정리하였다. 셋째, 분석을 위한 상세한 데이터 포맷을 표준화하여 정의하였다. 이는 분석에 있어 해당 데이터를 표준화된 형태로 가공하여 표준분석모델의 입력 데이터로 활용하기 위함이다. 유동인구 데이터 등의 민간수급 데이터도 표준화 작업을 진행하였다.

이에 따라, CCTV우선설치지역을 도출하기 위한 표준 분석모델 데이터로 감시취약지수, 범죄취약지수와 가중치를 도출하였다. 표준화된 데이터의 분석방법은 Geo-coding, Classification, Natural Breaks, Cell Generation 등의 분석 방법을 사용하여 분석하였으며, 분석도구는 데이터마이닝이나 텍스트마이닝을 위해 오픈소스인 R을 기본으로 분석하였다. 지도 등을 이용한 GIS데이터 분석작업은 Q-GIS를 기본 분석도구를 사용하였다[18].

### 2.3 분석에 사용할 지수 정립

빅데이터 표준분석모델의 구축을 위해 해당 지자체 및 개방된 공공데이터를 획득하여 <Table 2>와 같은 지수산출 방식을 적용하였다. 이는 CCTV 우선설치 지역분석에 활용하기 위한 지수산출 기법이다. CCTV설치지역란 CCTV우선설치 지역을 도출하기 위한 값이다. 즉

<Table 2> Index Algorithm for Analysis

Item	Index Algorithm
Vulnerability of Monitoring (VOM)	$VOM = [(1-a)] * 10$ where a = monitoring areas by CCTV (m <sup>2</sup> )/total areas(m <sup>2</sup> )
Vulnerability of Crime (VOC)	$VOC = grade1_{crime} + grade2_{env} + grade3_{flow}$
Weight*	Rate for female population vulnerable for crime (1.095) Rate for areas of independent housing compared to total housing areas (1.041) Rate for registered foreigners in a local government. (1.014) * This types of weight can be omitted up to the local government's option
Installation index	Installation index = (Index VOM + Index VOC) * weights Installation scoring point = f (the number of safe facilities in buffer areas)

CCTV산출지수가 높은 지역일수록 CCTV설치가 우선적으로 요구되는 지역임을 알 수 있다. CCTV의 설치지수를 구하기 위해 감시취약지수, 범죄취약지수와 가중치를 각각 구하여 인용한다. 감시취약지수는 격자내의 CCTV 점유면적에 대한 해당 감시면적의 비율이며, 범죄취약지수는 범죄요인 변수, 환경변수 및 유동인구 요인 변수의 총 합산 값으로 구한다. CCTV 우선 설치지수에 대한 가중치는 범죄에 취약한 여성 인구수 비율, 단독주택 면적비율과 지자체 등록 외국인 비율로 구분하여 조사하였다. 하지만 이러한 가중치는 지자체의 의견 및 선택에 따라 부여하지 않을 수도 있다. 이러한 분석지수 산출에 따르면, 최대 CCTV 설치지수는 29점으로 도출될 수 있다. 분석지수를 적용한 자세한 내용은 다음의 '3. K시의 CCTV우선설치지역 도출사례'에서 구체적으로 설명한다.

### 3. K시의 CCTV우선설치지역 도출사례

#### 3.1 감시취약지수 (VCM)

감시취약지수를 구하기 위해서 먼저 격자 내에 CCTV가 점유하는 면적을 구한 후 각 격자에 해당 감시면적이 차지하는 비율로 구한다. 이는 최대 10점까지 부여된다. CCTV 기설치 감시면적에서 50m버퍼영역을 설정하고, 총 면적인 10,000m<sup>2</sup> 중 CCTV감시면적이 차지하는 면적을 구하여 감시취약지수를 산출하였다. 감시취약지수 =  $[(1-a)] * 10$  ( $a = \text{CCTV감시면적}(m^2) / \text{총 면적}(m^2)$ ). CCTV감시면적은 반경 50m, 360도를 부여하였고, 격자에 주거인구가 0인 지역은 감시취약지수를 0으로 처리하였다. 그 결과, 총 46,882개 격자 중 CCTV 기설치 면적

이 격자에 100% 차지하거나, 해당 주거인구가 0명으로 예외 처리하여 산출된 값이 39,841개로 많았다. 그 외에 CCTV감시 취약지수는 10점이 압도적으로 많은 분포를 보였다.

#### 3.2 범죄취약지수 (VOC)

범죄취약지수는 범죄요인변수( $grade1_{crime}$ ), 환경요인 변수( $grade2_{env}$ ), 유동인구요인변수( $grade3_{flow}$ )을 합산하여 구할 수 있으며, 최대 15점까지 도출된다.

$$VOC = grade1_{crime} + grade2_{env} + grade3_{flow}$$

##### 3.2.1 범죄요인 변수

CCTV 설치요청 내용들이 담긴 지자체의 설치요청민원 데이터를 범죄발생 측면에서 <Table 3>과 같이 5단계로 분류하고, 1~5점 등급화 하여 해당 격자의 Grade\_Crime변수에 점수를 적용하였다.

$$grade1_{crime} = f(\text{Crime installation requests})$$

<Table 3> Classification of Crime Type

Grade	Crime Type
5 grades	Kidnap murder, repeated robber or sexual crime
4 grades	Robber, sexual crime, repeated theft, violence
3 grades	theft, violence, a crime-ridden district, derailment
2 grades	requests for preventing crime
1 grades	no demand request

##### 3.2.2 환경요인 변수

범죄취약지수에 필요한 환경요인은 유흥업소밀도, 숙

박업소밀도, 금융업소 밀도를 격자별로 계산하였고, 해당 격자의  $Grade_{2\_env}$  변수에 그 최종 집계 값을 입력하였다. 범죄환경요인에 대한 요인들은 주류판매업소의 밀도, 주류취급음식점 밀도, 주점밀도가 높을수록 범죄율, 강간률, 폭행률 등 모든 범죄발생률이 높아진다는 기존 연구를 바탕으로 도출하였다. 또한, 공간적, 환경적 요인이 범죄발생에 밀접한 영향이 있다는 연구와 범죄와 밀접한 관련이 있는 요인은 유동인구수와 유흥업소수임을 증명한 연구 등의 결과를 토대로 구분하였다[19].

$$grade_{2\_env} = (\text{density of bars} + \text{density of hotels} + \text{density of banks}) * 5$$

위의 식에서 유흥업소밀도, 숙박업소밀도, 금융업소밀도는 격자 안의 총 사업체 수에서 각각의 사업체 수를 나눈 값이다. 예를 들어, 유흥업소밀도 = 유흥업소 사업체 수/총 사업체 수이다. 분석결과 K시의 범죄취약지수는 총 46,891개 격자값 중 평균이 0.007로, 표준편차는 0.095로 나타났으며, 총 격자내 총 사업체 10곳중 유흥업소와 숙박업소가 밀집해 있는 지역이 범죄취약지수 최고 점인 5점 만점 중 4점으로 가장 높은 범죄취약지수를 나타냈다.

### 3.2.3 유동인구 요인 변수

범죄취약 계층인 10~30대 여성의 심야 시간대 보행 및 야간 귀가시간대 유동인구를 전체 5단계로 등급화하고 해당 격자의  $Grade_{3\_flow}$  변수에 1~5점에 해당하는 점수를 부여하였다. 이는 시간대별로 다른 유형의 범죄가 집중된다는 기존 연구결과에 기인한다[20].

$$grade_{3\_flow} = (\text{flow}_{1\text{latehourreturn}} + \text{flow}_{2\text{midnight}}) / 2$$

위의 식에서  $flow_1$  은 18시에서 24시 사이에 보행 및 야간귀가를 하는 여성들로 ‘야간시간귀가 등급’으로 지정하고  $flow_2$ 는 24시부터 새벽 6시 사이 보행, 귀가하는 여

성들로 ‘심야시간보행등급’으로 지정하여 값을 부여하였다. 유동인구가 1보다 작은 구간에서는 해당 변수에 0의 값을 부여하고, 유동인구가 1보다 큰 구간에서는 20%씩 나누어 1~5점을 부여하였다. 이러한 등급구분을 통해 유동인구 밀집도가 높은 지역을 구분하여 CCTV우선설치 지역의 지수를 판단하였다.

### 3.3 가중치

가중치는 격자별로 취약여성인구, 주택유형, 외국인에 대한 가중치를 계산하여 반영한다. 이 지수값은 최대 1.2점까지 부여될 수 있다. 본 연구에서는 범죄취약여성인구, 단독주택, 외국인 밀집지역에 해당하는 지역에 반영하여 계산하였다. 이러한 가중치 항목들은 특정 공간 및 집단과 취약계층의 범죄가 증가하고 있다는 연구와 범죄 유발 지역 및 공간에 대한 위험성 평가를 연구한 내용을 토대로 도출하였다[21, 22]. 또한 CCTV분석사업 현황분석 및 지자체 담당자와의 인터뷰를 통해 항목들을 도출하였다. 하지만, 지자체별로 이러한 특정 가중치 관련 데이터의 수급이 어렵거나 측정이 불가피할 경우 해당 가중치 값을 부여하지 않아도 된다. 범죄취약여성 비율( $w_1$ ), 단독주택비율( $w_2$ ) 그리고 외국인 비율( $w_3$ )에 대해 적용 가중치 값을 <Table 4>와 같이 적용하였다. 따라서, 가중치 값은 다음과 같은 모델을 따른다.

$$\text{Weight} = w_1_{\text{female}} * w_2_{\text{indep.housing}} * w_3_{\text{foreigner}}$$

$w_1_{\text{female}}$ 은 10~30대 여자 거주인구수의 비율로 측정하고, 이 값의 평균값을 구하여 평균값 이상인 경우 여성거주 밀집지역으로 지정한다. 마찬가지로,  $w_2_{\text{indep.housing}}$  은 총 주택 면적 대비 단독주택 면적 비율과 평균값을 구하여, 평균값 이상인 경우 단독주택 밀집지역으로 지정하였다[23, 24].  $w_3_{\text{foreigner}}$ 은 외국인 격자 가중치에 대한 연구를 바탕으로 격자별로 등록 외국인에 대한 밀집비율을 구하였다[25]. K시의 경우, 가중치가 모두 부여된 격자가

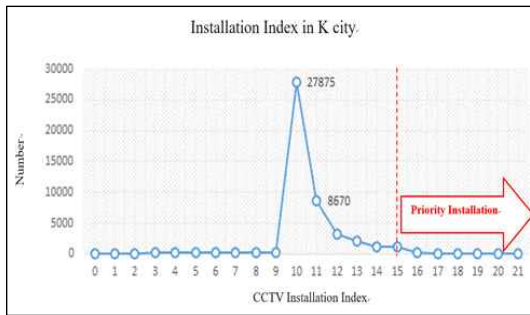
<Table 4> A Summary of Weights

상위 20개 중 2곳이었고, 평균적으로 최소 가중치 하나 이상을 적용하였다. 하지만, 전체적인 우선설치 지수값에 큰 영향을 미치지 않는 것으로 판명되었다.

### 3.4 CCTV설치지수

앞서 도출된 감시취약지수, 범죄취약지수 및 격자 가중치를 이용하여 최소 CCTV설치지수를 산출한다. 설치지수는 최대 29점까지 부여될 수 있으며, 이 값을 토대로 CCTV우선설치 지역을 확인할 수 있게 된다. 빅데이터 분석표준 모델을 기반으로 도출된 K시의 CCTV우선 설치지수는 20.805점으로 평균 설치지수 10.456, 표준편차 1.71의 결과값을 나타내고 있다.

또한 총 46,883개의 격자값 중 CCTV 최우선 설치지수는 상위 5%에 해당하는 점수구간인 15점 이상으로 2,344개가 산출되었다. [Fig. 1]과 같이 최우선 설치등급은 CCTV설치지수가 15점 이상으로 상위 5%에 해당하며, 보통 설치등급인 10점 등급이 27,875개로 10점~11점 등급의 구간이 70% 비율로 가장 높게 나타났다.



[Fig. 1] The Installation Index Scores

분석결과를 살펴보면, K시의 CCTV분포는 대체로 도심 빌딩밀집지역, 주거인구밀집지역에 분포되어 있음을 알 수 있고, 단독주택밀집지역이 아니어도 유동인구요인이나 설치요청민원이 높은 지역을 확인할 수 있었다. 또한, 절도, 폭력 등 범죄우발 지역도 우선설치리스트의 상위에 다소 분포하여 기존 민원발생지역과 빅데이터 표준 분석 모델을 활용한 우선설치 결과와 대체로 일치함을 알 수 있다. 또한 본 연구의 빅데이터 분석결과는 설치지수가 높음에도 불구하고, 아직 CCTV설치가 되지 않은 미 설치지점이 상당부분 존재하고 있고 설치요청 민원이

있음에도 불구하고 미설치된 사각지대가 있음을 확인할 수 있었다. 따라서, 빅데이터의 표준분석모델 활용은 민원처리 등의 행정절차에 대한 시간과 비용의 감소 및 지역주민과 사회구성원의 생활 안정성 및 만족도를 증대시킬 수 있을 것으로 판단된다.

## 4. 결론 및 시사점

본 연구에서는 K시의 CCTV설치우선 지역분석 사례를 바탕으로 빅데이터 표준분석 모델의 유효성과 적합성을 검증하고자 하였다. 이를 위해 민원데이터, 거주인구, 유동인구 등 관련 데이터를 해당 지자체 및 공공데이터를 통해 수집하였다. K시의 사례분석 결과, 우선 설치리스트 상위 지점 모두 범죄취약지수 중 환경지수 값이 전반적으로 낮게 나온 반면, CCTV미설치 지역에 따른 감시취약지수와 야간 및 심야 시간대의 유동인구 지수가 높게 나타났다. 빅데이터 표준분석 모델에 의한 결과 값으로 나온 우선설치 리스트 상위지역에서 범죄요인 Grade 1의 경우 대부분 5점 만점에서 3점 이상으로 절도, 폭력 등이 빈번하게 발생하는 우범지역임을 알 수 있다. 이 지역은 또한 CCTV설치요청 민원이 높은 지역으로 기존 민원 발생지역과 데이터 분석결과가 대체로 일치함을 보여주고 있다. 종합적으로, 빅데이터 표준분석모델에 따른 결과는 실제 CCTV설치대수가 미비한 사각지대에서 범죄취약지수 및 CCTV설치지수가 높게 나왔다. 이 지역은 평소 민원이 많았고, 해당 지역 담당자도 CCTV설치의 필요성을 인지하고 있는 지역으로 나타나 빅데이터 분석결과가 높은 정확성을 보여주고 있음을 확인할 수 있다[26]. 이러한 결과는 실제 정부주도하에 진행하고 있는 표준분석모델의 적합성을 사례연구를 통해 검증하였다는 점에서 연구의 독창성과 의의를 가진다.

본 연구의 결과는 다음과 같은 시사점을 제시한다. 첫째, 표준분석모델의 분석결과가 실제 민원요구 및 실사를 통해 확인된 CCTV설치의 요구지역과 상당부분 일치성을 나타내어 표준분석모델 분석의 정확성을 확인하였다. 둘째, 분석결과는 쉽게 파악되고 인식되지 않은 사각지대 및 우범지대를 명확하게 제시함으로써 미처 고려되지 못했던 지역 및 사안들을 파악할 수 있다. 이는 빅데이터 분석을 통해 신속하고 효과적인 의사결정과 업무수

행이 향상될 수 있음을 보여준다. 이러한 결과는 예측모형 개발 시 통합된 빅데이터 분석방법론을 적용하는 것이 예측모형의 오류를 줄이고 신뢰도를 높일 수 있다는 기존의 연구결과와 유사하다[27, 28]. 셋째, 빅데이터 표준분석모델을 활용함으로써 지자체는 데이터 획득 및 빅데이터 분석도구와 분석과정에서 발생하는 인력, 재정, 시간 등의 기회비용을 대폭 절감할 수 있다. 또한 지역주민들의 요구사항 및 민원을 신속하게 처리하여 주민들을 범죄위험에서 보호하고 생활의 편리함과 행정지원 및 정책을 효율적으로 운영할 수 있음을 알 수 있다[29].

본 연구는 빅데이터 표준분석모델을 실제 사례에 적용하여 분석모델의 적합성을 조사한 연구로서 빅데이터 정제 및 표준화 절차에 대한 분석과정을 자세하게 설명하지 않았다. 따라서, 향후연구에서는 구체적인 지수산출 과정 및 표준화 등 빅데이터 표준분석모델 구축과정에 초점을 맞추어 결과도출 과정을 제시할 필요가 있다. 또한, 전문가 인터뷰(AHP, pairwise, delphi) 등을 통한 분석의 타당성 검증에 대한 후속연구가 필요하다. 마지막으로, 본 연구결과는 특정 기초단체의 사례를 분석한 결과이기 때문에 일반화의 한계가 있을 수 있다. 따라서, 표준분석모델을 다양한 분야 및 지역 사례에 적용하여 관련 결과를 비교연구해 볼 필요가 있을 것이다.

## ACKNOWLEDGMENTS

This study is based on the report of 'A due diligence of CCTV installation region' written in part of the 2016 Public Big Data Standard Analysis Mode project by National Information Society Agency(NIA).

## REFERENCES

[1] Forth Industrial Revolution, Big data marketing, <http://glamstone.net/220855312522> (Nov. 2016)

[2] Tae-Young Kim, Yong-Kim, Hyo-Jung Oh, "An Analysis of the Relationship between Public Opinion on Social Bigdata and Results after Implementation of Public Policies: A Case Study in 'Welfare' Policy," *Journal of Digital Convergence*.

Vol. 15. No. 3, pp. 17-25. 2017.

[3] KeunWon Kim, DongWoo Kim, Kyoo-Sung Noh, Joo-Yeoun Lee, "An Exploratory Study on Improvement Method of the Subway Congestion Based Big Data Convergence," *Journal of Digital Convergence*, Vol. 13, No. 2, pp. 35-42, 2015.

[4] M. Chen, S. Mao, Y. Ahang, V. Leung, *Big Data: Related Technologies, Challenges and Future Prospects*, Cham: Springer, 2014.

[5] A. T. Hashem, I. Yaqoob, N. Anuar, S. Mokhtar, A. Gani, S. U. Khan, "The Rise of Big Data on Cloud Computing: Review and Open Research Issues," *Information Systems*, No. 47, pp. 98-115, 2015.

[6] J. Manyika, *Big Data: the Net Frontier for Innovation, Competition and Productivity*, McKinsey Global Institute, 2011.

[7] M. Chen, S. Mao, Y. Liu, "Big Data: A Survey," *Mobile Net Appl*, No. 19, pp. 171-209, 2014.

[8] Editorial, "Big Data Management: New Frontiers, New Paradigms," *Information Systems*, Vol. 63, pp. 63-65, 2017.

[9] B. Aaqaibeh, I. Obaidat, W. Hussien, "Big Data Analysis Techniques using Multi-GPUs Map Reduce Implementations," *International Journal of Advanced Studies in Computer Science and Engineering*, Vol. 5. No. 11, pp. 1-6, 2016.

[10] Ministry of the Interior, Development of Public Big-Data Standard Analysis Model, Risk Management, <http://www.di-focus.com/news/articleView.html?idxno=17411>(December, 2016)

[11] Seong-Hoon Lee, Dong-Woo Lee, "Current Status of Big Data Utilization," *Journal of Digital Convergence*, Vol. 11. No. 2, pp. 229-233. 2013.

[12] Boannews, Enhancing social life, security and welfare through bigdata, <http://www.boannews.com/media/view.asp?idx=53257&kind=2> (Jan. 2017)

[13] N. M., Joyce, C. H. Ramsey, and J. K. Stewart, "Commentary on Smart Policing," *Police Quarterly*, Vol. 16, No. 3, pp. 358-368, 2013.

[14] Y. S. Kim, "Information and Communications Technology in the Field of Public Security: Crime

- Prevention and Response System,” *Convergent Security Journal*, Vol. 16, No. 6, pp. 23-32, 2016.
- [15] Young-mi Kim, “The Study on Local Government’s Disaster Safety Governance using Big Data,” *Journal of Digital Convergence*, Vol. 15, No. 1, pp. 61-67, 2017.
- [16] W. L. Sung, S. S. Chen, M. E. Yun, H .R. Lee, “The Effect of Alcohol On premises Densities on Corresponding Local Criminal Rates in Seoul,” *Public Health Association*, Vol. 41, No. 4, pp. 75-86, 2015.
- [17] S. W. Lee, J. G. Cho, “The Effect of Spatial and Environmental factors on Crime,” *Victimization, Journal of Seoul City Research*. Vol. 7, No.. 2, pp. 57-76, 2006.
- [18] Sang H. Cho, Min Jae Kim, “First step for Public Big Data Analysis Explosion”, *National Information Society Agency*, 2016.
- [19] Min-Gu Song, Sun-Bae Kim, “A Study of improving reliability on prediction model by analyzing method Big data,” *Journal of Digital Convergence*, Vol. 11, No. 6, pp. 103-112, 2013.
- [20] M-J. Choi, K-S. Noh, “Exploratory Study on Crime Prevention based on Bigdata Convergence,” *Journal of Digital Convergence*, Vol. 14, No. 11, pp. 125-133, 2016.
- [21] Ministry of Trade, Industry and Energy, *Social Protection Service Design for Sexual Crime Prevention*, Report, 2016.
- [22] O. K. Shin, *The Development of Crime Risk Assessment Tool and Its Application(III) - with Focus on Sexual Violence and School Violence*, Report, *Korean Institute of Criminology*. 2014.
- [23] Y .H. Kim, J .M. Moon, D. G.J ang, “Study on the Distribution of Crime in Urban Space - Gwangju Metropolitan City from the Perspective of the Environmental Criminology,” *Journal of Korea Society of Computer Information*, Vol. 12, No. 5, pp. 235 - 241, 2007.
- [24] D. S. Lee, K .H. Lee, “A Study on the Relationship Between Crime Rates and Characteristics of Urban Areas: Focusing on Demographic and Socio-Economic Factors.” *Journal of Korean Association for Local Government Studies*, Vol. 11, No. 1, pp. 189-206, 2009.
- [25] Y. S. Choi, S. J. Kang, M. S. Kim, and I. S. Kim, “A Study on the Foreign Crimes in Ethnic Places in South Korea,” *Korea Institute of Criminology*, 12-AA-09, 2012.
- [26] Y. S. Kang, Jin Woo Park, “Report of CCTV installation region\_Due Diligence,” *National Information Society Agency(NIA)*, Nov. 25 pp. 1-46, 2016.
- [27] M. G. Song, S. B. Kim, “A Study of Improving Reliability on Prediction Model by Analyzing Method Big data,” *The Journal of Digital Policy & Management*, Vol. 11, No. 6, pp. 103-112, 2013.
- [28] J. Fan, F. Han, H. Liu, “Challenges of Big Data analysis,” *National Science Review*, Vol. 1, pp. 293-314, 2014.
- [29] Mi-Ok Kim, Jin-Soo Ko, Seung-Chul Noh, Jae-Hoon Chung, “A Study of the Vitalization Strategy for Public Sports Facility through Big-Data,” *Journal of Digital Convergence*, Vol. 15, No. 2, pp. 527-535, 2017.

성 창 수(Chang Soo Sung)



- 2011년 2월 : 중앙대학교 일반대학원 창업학과 (창업학박사)
- 2013년 12월 : 선문대학교 BT융합 제약공학과 조교수
- 2015년 5월 ~ 현재 : 동국대학교 기술창업학과 조교수
- 관심분야 : 기술창업, 기업가정신과 혁신, 창업교육

· E-Mail : redsun44@dongguk.ac.kr

박 주 연(Joo Y. Park)



- 2005년 2월 : 연세대학교 경영학과 (경영학박사)
- 2011년 ~ 2013년 : 연세대학교 정보대학원 연구교수
- 2013년 9월 ~ 현재 : 연세대학교 경영학과
- 관심분야 : 정보시스템, 빅데이터, 기술창업

· E-Mail : park3500@naver.com



가 회 광(Hoi Kwang Ka)



- 2014년 2월 : 중앙대학교 일반대학원 경영학과 (경영학박사)
- 2011년 3월 ~ 2016년 4월 : 한국창업경영연구원 책임연구원
- 2016년 5월 ~ 현재 : 행정자치부 전문관 (빅데이터)
- 관심분야 : 빅데이터, SCM, Network Biz-model