

https://doi.org/10.7236/JIIBC.2024.24.5.211
JIIBC 2024-5-30

빅 데이터를 활용한 연도별 범죄 분석을 통한 CPTED 활용방안 제시

Suggestion of CPTED Application through Annual Crime Analysis using Big Data

황승연*, 문진용**, 김정준***

Seung-Yeon Hwang*, Jin-Yong Moon**, Jeong-Joon Kim***

요약 우리나라의 세계 치안율은 상위권이며, 새벽에 술을 먹고 마음껏 돌아다닐 수 있을 정도로 안전한 나라이다. 하지만 이러한 우리나라에서도 검찰청 자료에 따르면 매년 약 200만 건 전후의 범죄가 발생하고 있다. 범죄는 폭력범죄, 지능범죄, 기타범죄 등 여러 범죄로 분류될 수 있다. 또한, 범죄별 연령, 성별, 장소가 모두 다르다. 대부분의 범죄는 사람들의 신뢰를 무너뜨리고 사회에 크고 작은 피해를 입히며, 범죄로 인해 무너진 신뢰는 회복하기 어렵다. 이러한 범죄의 종류에 따라 연령, 성별, 장소와 같은 조건을 기준으로 본 논문에서 빅 데이터를 이용해 분석하여 범죄 발생에 대해 접근 및 증/감폭을 시각화 할 것이다. 이 논문을 통해 범죄의 전체적인 흐름을 파악하고, 분석한 자료를 바탕으로 범죄 예방 디자인(CPTED)의 활용 방안을 제시할 것이다.

Abstract Our world security is at the top of the list, and it is a safe country to be able to drink and move around freely at dawn. However, according to the data by the prosecution's office, about 2 million crimes are committed each year. Crime can be classified as a violent crime, intelligent crime, or other crime. Also age, gender and location vary by crime. Generally, crime breaks people's trust, inflicts large and small damages on society, and it is difficult to restore trust, which is crushed by crime. In the paper based on the conditions such as age, gender, place, and area, we will use Big Data to visualize the access and increase / decrease of crime occurrence according to the types of these crimes. Through this paper, we will identify the overall trend of crime and present a method of utilizing the crime prevention design (CPTED) based on the analyzed data.

Key Words : Big Data, CPTED, Crime, Condition

1. 서론

범죄는 과거부터 지금까지 끊임없이 발생하고 있다. 고조선 시대에서부터 현재까지 법과 이념으로써 범죄

에 대해 규명하고 있다. 고조선 시대에는 '사람을 죽이면 죽음으로 배상한다. 다른 사람에게 상처를 입히면 곡물로 배상한다. 남의 물건을 훔친 사람은 노비로, 또는 50만을 내야 한다.'와 같이 범죄에 대해 물질적인

*준회원, 안양대학교 창의융합대학 컴퓨터공학과

**정회원, 강동대학교 온라인영상홍보전공

***정회원, 안양대학교 창의융합대학 소프트웨어학과

접수일자 2024년 6월 11일, 수정완료 2024년 9월 5일
게재확정일자 2024년 10월 4일

Received: 11 June, 2024 / Revised: 5 September, 2024 /

Accepted: 4 October, 2024

***Corresponding Author: jkim@anyang.ac.kr

Dept. ICT Convergence Engineering at Anyang University, Korea

보상으로 규정되어 있으며, 현대에는 헌법으로써 범죄자와 피해자에 대한 법률이 명시되어 있다. 헌법 제 30조 '타인의 범죄행위로 인하여 생명·신체에 대한 피해를 받은 국민은 법률이 정하는 바에 의하여 국가로부터 구조를 받을 수 있다.'와 제 13조 1항 '모든 국민은 행위 시의 법률에 의하여 범죄를 구성하지 아니하는 행위로 소추되지 아니하며, 동일한 범죄에 대하여 거듭 처벌받지 아니한다.'와 같이 범죄자, 피해자에 대한 명시가 되어있다. 헌법은 대한민국 최고 기본법으로 모든 법 중에서 가장 상위에 위치하며, 대한민국 모든 국민에게 적용되며 대한민국의 영역 내에서 적용된다. 또한 분야별로 법률을 가지고 있다.

이렇듯 범죄에 대해 과거부터 현재까지 법으로, 예방책으로써 방지하려는 노력을 하고 있다. 하지만 대검찰청이 발간하는 범죄 분석에 따르면, 2010년 우리나라에서 교통범죄를 제외한 전체 범죄 발생 건수는 1,332,153건으로, 이는 국민 100명당 3명 가량이 매년 범죄 피해를 겪고 있음을 의미한다. 우리나라의 범죄 발생 추이는 지난 30여 년간 연평균 3.4% 증가해 왔는데, 2000년 이후 최근 10년을 놓고 보면 2.1%로 다소 낮아졌지만 증가 추세는 계속 이어지고 있다. 아울러 폭력사건들이 감소 양상을 보이고 있음에도 불구하고 살인, 강도, 강간, 방화와 같은 흉악범죄가 높은 증가추세를 보이고 있다는 점도 주목해야 할 부분이다¹¹.

범죄에 따라 강력범죄, 폭력범죄, 지능범죄 등과 같이 여러 분류와 범죄별 나이, 성별 등 여러 조건이 있다. 범죄는 단순히 일어나는 것으로 끝이 아닌 여러 인과관계가 존재한다. 본 논문에서는 연도별, 범죄별, 조건별 경찰청 조사 자료를 바탕으로 빅 데이터를 활용하여 범죄의 여러 방면에 대해 조사와 분석을 통해 범죄 흐름을 파악하고 분석한 자료를 바탕으로 범죄 예방 디자인(CPTED) 활용 방안을 제시할 것이다.

II. 관련 기술

1. 빅데이터(Big Data)

빅데이터는 방대한 양의 정형 및 비정형 데이터를 수집, 저장, 분석하여 유용한 통찰을 도출하는 기술을 말한다. 이는 전통적인 데이터베이스 관리 도구로는 처리할 수 없는 규모의 데이터를 다루기 위해 발전되었다. 빅데이터의 핵심 개념은 세 가지로 나눌 수 있다. 첫째, 볼륨(Volume)은 데이터의 방대한 양을 나타내며, 둘째, 다양

성(Variety)은 데이터의 다양한 형태와 출처를 의미하고, 셋째, 속도(Velocity)는 데이터 생성과 처리 속도를 나타낸다. 이러한 특성을 통해 빅데이터는 기존 데이터 분석 방식으로는 얻을 수 없는 새로운 가치를 창출할 수 있다¹²⁻³¹.

2. Apache Hadoop

하둡은 대규모 데이터를 저장하고 처리하기 위한 오픈 소스 프레임워크로 분산 파일 시스템(HDFS)과 맵리듀스(MapReduce) 프로그래밍 모델로 구성되며 수백 대의 컴퓨터 클러스터에서 데이터를 효율적으로 분산 저장하고 병렬 처리할 수 있다. 빅데이터 분석, 데이터 웨어하우징, 머신러닝 등의 분야에서 널리 사용되며 스케일 아웃 아키텍처 덕분에 하드웨어 확장에 유연하다. 데이터 손실을 방지하기 위해 복제 기능을 제공하며 주요 컴포넌트로는 하둡 커먼, HDFS, 맵리듀스, 안(YARN)이 있다¹⁴⁻⁵¹.

3. Apache Pig

아파치 피그(Apache Pig)는 대규모 데이터 세트를 분석하고 처리하기 위한 고수준 플랫폼이다. 피그는 스크립트 언어인 피그 라틴(Pig Latin)을 사용하여 데이터 흐름을 정의하며, 복잡한 맵리듀스 작업을 단순화한다. 이를 통해 개발자는 코드 작성 시간을 줄이고 데이터 처리 작업을 효율적으로 수행할 수 있다. 피그는 하둡과 통합되어 HDFS에 저장된 데이터를 처리하며, 데이터 필터링, 집계, 조인 등의 작업을 쉽게 수행할 수 있다. 데이터 엔지니어와 분석가들이 대량의 데이터를 다루기 위해 주로 사용하며, 유연성과 확장성이 뛰어나며 피그의 주요 구성 요소로는 피그 인터프리터, 피그 컴파일러, 피그 런타임이 있다¹⁶⁻⁷¹.

III. 분석 자료

1. 범죄별 분류(2012-2016)

표 1. 연도별 전체 범죄 횟수(2012-2016)
Table 1. Total Number of Crimes by year

연도	발생 횟수
2012	1,934,410
2013	1,996,389
2014	1,933,835
2015	2,020,731
2016	2,008,290

범죄 구분으로는 강력범죄, 절도, 폭력범죄, 지능범죄, 풍속범죄, 특별경제범죄, 마약범죄, 보건범죄, 환경범죄, 교통범죄, 노동범죄, 안보범죄, 선거범죄, 병역범죄, 기타로 분류가 되며 2012년에서 2016년의 연도별 전체 범죄 횟수는 위 표 1과 같다.

2. 성별/나이별 분류(2012-2016)

표 2. 성별/나이별 범죄 전체 횟수(2012-2016)
 Table 2. Total Number of Crimes by Sex / Age

나이	남자	여자
20세 이하	515,498	104,443
21-30	1,169,258	225,374
31-40	1,510,227	316,545
41-50	1,838,210	453,426
51-60	1,501,444	378,784
61-70	485,089	109,826
71세 이상	146,434	32,952
미상	4,410	4,627

범죄별 나이와 성별에 따른 2012년에서 2016년의 연도별 전체 범죄 횟수는 위 표 2와 같다.

3. 장소별 분류(2012-2016)

범죄별 장소에 따른 2012년에서 2016년의 연도별 전체 범죄 횟수는 위 표 3-4와 같다.

표 3. 장소별 범죄 전체 횟수(2012-2016)
 Table 3. Total Number of Crimes by Place

장소	발생 횟수
아파트, 연립다세대	316,206
단독주택	419,622
고속도로	21,810
노상	3,958,941
백화점	12,761
슈퍼마켓	39,654
편의점	51,944
대형할인매장	14,568
상점	154,753
시장, 노점	56,566
숙박업소, 목욕탕	106,549
유흥접객업소	454,821
사무실	632,563
공장	56,477
공사장, 광산	41,314
창고	15,213
역, 대합실	34,760
지하철	12,854
기타교통수단내	74,966
홍행장	8,686
유원지	29,098

표 4. 장소별 범죄 전체 횟수(2012-2016) (계속)
 Table 4. Total Number of Crimes by Place (Continue)

장소	발생 횟수
학교	33,169
금융기관	125,322
의료기관	50,806
종교기관	15,479
산야	43,051
해상	64,572
부대	21,691
구금장소	362
공지	2,225
주차장	126,001
공중화장실	10,759
PC방	47,050
기타	2,086,136

IV. 본 론



그림 1. 2012-2016 원형 그래프 지역 라벨
 Fig. 1. Pie Graph Region Label for 2012-2016

본 논문의 목적은 경찰청 범죄 자료를 바탕으로 범죄를 분석하고, 분석 한 자료를 바탕으로 CPTED 기법 활용 방안을 제시 하고자 한다. 경찰청 범죄 자료에는 범죄에 따른 분류와 종류가 많다. 범죄의 분류로는 강력범죄, 절도범죄, 폭력범죄를 포함하여 15가지 분류가 있다. 본 논문에서는 가장 발생 빈도가 높은 폭력범죄, 지능범죄, 절도범죄에 대해 다룬다. 다음 원형 그래프들은 폭력범죄, 지능범죄, 절도범죄 중 가장 범죄가 많이 일어난 지역을 파악하기 위해 전체 범죄 횟수를 원형 그래프로 표현하였다.

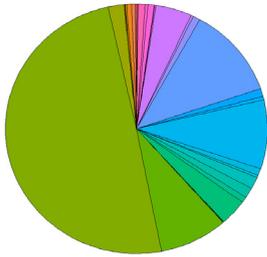


그림 2. 2012-2016 폭력 범죄 발생 지역 비교 원형 그래프
Fig. 2. Comparison Pie Graph of Violent Crime Region for 2012-2016

폭력범죄의 경우 노상이 가장 범죄율이 높았으며 그 뒤를 이어 유흥접객업소, 단독주택, 아파트-연립다세대, 주차장 등의 순으로 범죄율 순위가 높은 모습을 보이고 있다.

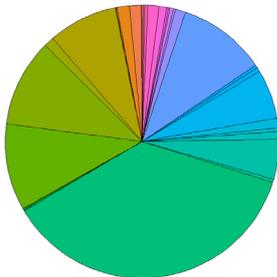


그림 3. 2012-2016 지능 범죄 발생 지역 비교 원형 그래프
Fig. 3. Comparison Pie Graph of Intelligent Crime Region for 2012-2016

지능범죄의 경우 사무실이 가장 범죄율이 높았으며 그 뒤를 이어 유흥접객업소, 금융기관, 단독주택, 노상 등의 순으로 범죄율 순위가 높은 모습을 보이고 있다.

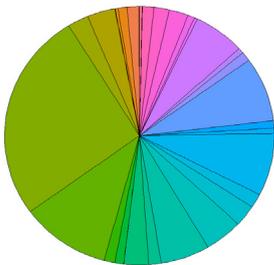


그림 4. 2012-2016 절도 범죄 발생 지역 비교 원형 그래프
Fig. 4. Comparison Pie Graph of Theft Crime Region for 2012-2016

절도범죄의 경우 폭력범죄와 마찬가지로 노상에서 가장 높은 범죄율을 보였으며 단독주택, 아파트-연립다세대, 유흥접객업소, 숙박업소-목욕탕 등의 순으로 범죄가

많이 발생하였다.

그림 2-4의 원형 그래프를 통해 범죄의 종류마다 발생하는 지역 범죄율의 분포가 상이한 결과를 보였다. 본 논문은 가장 많은 범죄가 발생하는 폭력범죄, 지능범죄, 절도범죄에 중점을 두어 가장 범죄가 많이 발생하는 5가지 지역에 대하여 연도별 범죄율 증감 비교 그래프를 제시한다.

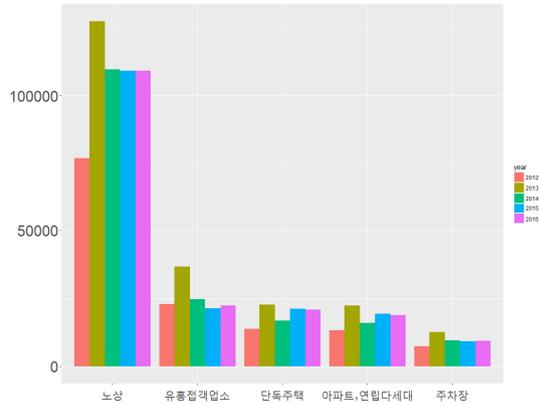


그림 5. 2012-2016 폭력범죄 발생시 발생장소 빈도 비교 그래프

Fig. 5. Comparative Graph of Violence Crime Occurrence Location by Year for 2012-2016

폭력범죄의 경우 노상에서 발생한 범죄의 비율이 다른 범죄 발생률에 비해 압도적으로 높았다. 또한 모든 범죄가 2013년도에 급증 후 감소하였으며 이후 범죄율은 큰 차이를 보이지 않았다.

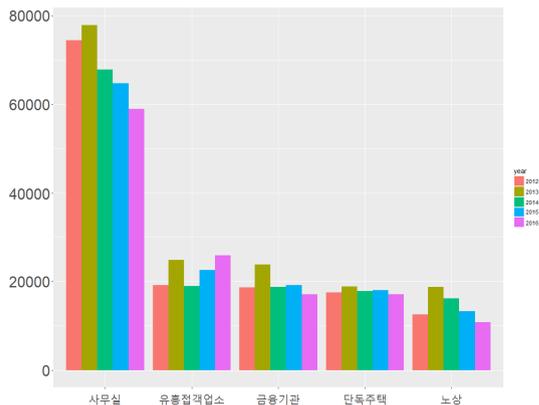


그림 6. 2012-2016 지능범죄 발생시 발생장소 빈도 비교 그래프

Fig. 6. Comparative Graph of Intelligent Crime Occurrence Location by Year for 2012-2016

지능범죄의 경우 사무실에서 발생한 범죄의 비율이 다른 범죄 발생률에 비해 압도적으로 높았다. 또한 전체적으로 범죄가 2013년도에 소폭 증가 후 감소하는 추세를 보였지만 유흥접객업소에서 발생하는 범죄는 예외적으로 범죄율이 상승하는 추세를 보였다.

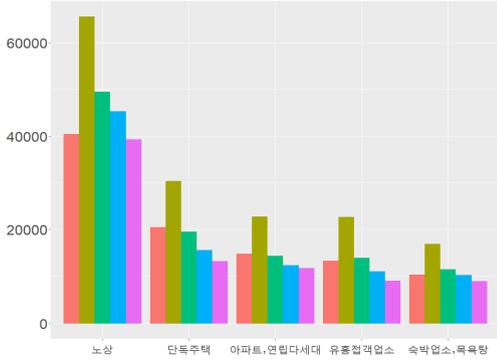


그림 7. 2012~2016 절도범죄 발생 시 연도별 발생장소 빈도 비교 그래프

Fig. 7. Comparative Graph of Theft Crime Occurrence Location by Year for 2012-2016

절도범죄의 경우 노상에서 발생한 범죄의 비율이 다른 범죄 발생률에 비해 압도적으로 높았다. 또한 모든 범죄가 2013년도에 급증한 이후 점차 범죄율이 감소하는 추세를 보였다.

다음 그래프들은 범죄별 성별 및 나이에 따른 범죄 횟수를 분석한 그래프이다. 전체 연령 중 가장 범죄 비율이 높은 20세 이하 부터 51-60세 까지의 구간을 기준으로 분석하였다. 전체적으로 남성이 여성에 비해 높은 범죄율을 나타내었다.

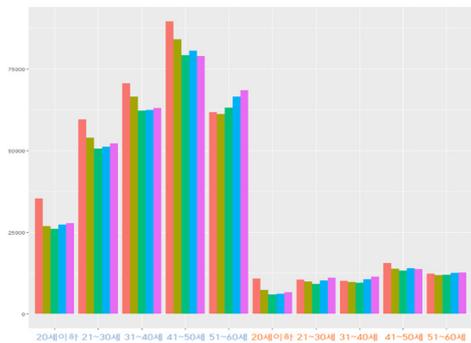


그림 8. 2012~2016 폭력범죄 발생 시 연도별 범죄자의 성별 및 나이 비교 그래프

Fig. 8. Graph of Gender and Age Comparison of Offenders by the Year When Violence Crime Occurs for 2012-2016

폭력범죄의 경우 남성은 51세~60세를 제외하고 2012년 이후 크게 감소하는 추세를 보였다. 51세~60세는 점점 연도에 따라 범죄율이 증가하였고 여성은 전체적으로 연도에 따른 범죄율 증감 변화를 보이지 않았다.

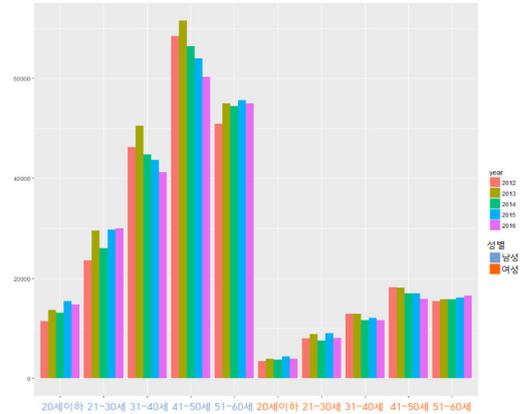


그림 9. 2012~2016 지능범죄 발생 시 연도별 범죄자의 성별 및 나이 비교 그래프

Fig. 9. Graph of Gender and Age Comparison of Offenders by the Year When Intelligent Crime Occurs for 2012-2016

지능범죄의 경우 남녀 공통으로 나이를 먹을수록 범죄율이 증가하는 추세를 보였고, 특히 41-50세 구간에서 가장 많이 발생하였다. 남성은 전체적으로 2013년 급증 후 감소하는 추세를 보였으며 여성은 전체적으로 연도에 따른 범죄율 증감 변화를 보이지 않았다.

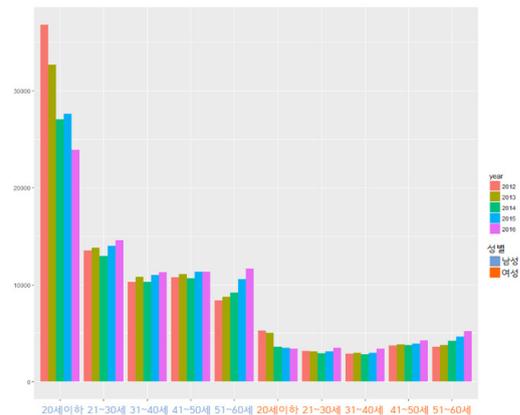


그림 10. 2012~2016 절도범죄 발생 시 연도별 범죄자의 성별 및 나이 비교 그래프

Fig. 10. Graph of Gender and Age Comparison of Offenders by the Year When Theft Crime Occurs for 2012-2016

절도범죄의 경우 남성, 특히 20세 이하의 남성의 비율이 압도적으로 높았다. 그 외의 연령에 따른 범죄율은 큰 차이를 보이지 않았다. 특이사항으로는 전체적으로 절도범죄율은 연도에 따라 비슷하거나 소량 증가하는 추세를 보였지만 20세 이하의 범죄율은 연도에 따라 크게 감소하는 추세를 보였다.

V. 결 론

본 논문에서는 공공데이터 포털 및 검찰청에서 제공하는 2012~2016년 범죄데이터를 중심으로 5년간 범죄 현황을 분석하였다. 위 그래프들이 나타내는 결과를 추론해보면 각 범죄별로 가장 범죄가 많이 발생하는 지역이 뚜렷하고, 연도별 범죄율도 특정 연령대의 범죄율이 타 연령대보다 상대적으로 높게 나타났다. 특히 남성의 범죄율이 여성의 범죄율보다 월등히 높게 나타난 걸로 확인되었다. 범죄율을 최대한 효율적으로 줄이기 위해선 가장 범죄율이 높은 분야를 공략해야 한다.

절도와 폭력범죄의 경우 노상에서의 범죄가 가장 많이 나타났다. 이를 예방하기 위해선 가장 많이 범죄가 발생한 노상에 CPTED를 적용시켜 좀 더 많은 사람들의 범죄를 예방할 수 있도록 하는 것이 최선의 방법이다. 지능범죄의 경우 사무실에서 발생하는 범죄가 압도적으로 많았다. 이는 사회 차원보다는 회사 차원에서 보안강화 및 내부금융 투명성 보장 등으로 예방하는 경우가 더 효율적일 것이다.

References

- [1] Kim, D. U., "Public Security, Welfare, Economic Growth : Evaluating Crime Incidence in Korea", National Knowledge Information System(NKIS), Vol. 1 No. 1, pp. 1-2, 2012.
- [2] Jae-Young Chang, "Bigdata Prediction Support Service for Citizen Data Scientists", The Journal of The Institute of Internet, Broadcasting and Communication, Vol. 19, No. 2, pp. 151-159, 2019.
DOI: <https://doi.org/10.7236/IIBC.2019.19.2.151>
- [3] In-Kyu Park, Gyoo-Seok Choi, "Regularized Optimization of Collaborative Filtering for Recommender System based on Big Data", The Journal of The Institute of Internet, Broadcasting and Communication, Vol. 21, No. 1, pp. 87-92, 2021.
DOI: <https://doi.org/10.7236/IIBC.2021.21.1.87>

- [4] Hadoop, <https://hadoop.apache.org/>
- [5] Tae-Hoon Keum, Won-Joo Lee, Chang-Ho Jeon, "Design and Implementation of a Monitor for Hadoop Cluster", Journal of the Institute of Electronics Engineers of Korea CI, Vol. 49, No. 1, pp. 8-15, 2012.
- [6] Apache Pig, <https://pig.apache.org/>
- [7] Hong-Geun Ji, Tae-Ki Lee, "Recommender System Development Based on Wine Review Big Data Analysis and Deep Learning", Proceedings of the Korea Information Processing Society Conference, pp. 769-766, 2019.
DOI: <https://doi.org/10.3745/PKIPS.y2019m10a.763>

저 자 소 개

황 승 연(준회원)

• Seung-Yeon Hwang received his BS in Department of Computer Science at Korea Polytechnic University in 2019. He is currently studying MS in Department of Computer Science at Anyang University. His research interests include Database System, Big Data, Data Analysis, Machine Learning, etc.

문 진 용(정회원)

• Jin Yong Moon received his MS in Computer Science from Konkuk University in 1998 and his PhD from Suwon University in 2001. He is currently a professor of Online Video Journalism Major at Gangdong University. His research interests include Information Technology, Graphic Design, Internet, Database and Multimedia, etc.

김 정 준(정회원)

• Jeong Joon Kim received his BS and MS in Computer Science at Konkuk University in 2003 and 2005, respectively. In 2010, he received his PhD in at Konkuk University. He is currently a professor at the department of Computer Science at Anyang University. His research interests include Database Systems, Big Data, Semantic Web, Geographic Information Systems (GIS) and Ubiquitous Sensor Network (USN), etc.