

스마트 캠퍼스 정보제공 서비스에 관한 연구

최신형*

강원대학교 제어계측공학과

A Study on Smart Campus Information Services

Shin-Hyeong Choi*

Dept. of Control & Instrumentation Engineering, Kangwon National University

요 약 본 연구의 목적은 대학교 캠퍼스 내에서 공부하고 생활하는 학생들에게 개인별 맞춤형 정보를 제공하는 것을 목적으로 한다. 연구방법은 캠퍼스 내부에서 수집되는 데이터와 인터넷 상의 블로그나 SNS 등의 외부 데이터를 수집하여, 이를 저장하여 처리한 다음, 이들 데이터를 세부적으로 분석하여 학생 개인별로 일대일 마케팅을 제공하는 시스템을 제안한다. 학생 한 사람 한사람의 구매 이력과 해당 건물의 출결상황을 상세히 분석하여 패턴에 맞춰 학생개인별로 내용이 다른 쿠폰 및 정보를 학생의 휴대전화로 전송한다. 이를 통해 학생 개인별로 맞춤형 정보를 제공함으로써 보다 유익한 대학생활을 할 수 있을 것이다.

키워드 : 대학 캠퍼스, SNS, 마케팅, 스마트폰

Abstract The purpose of this study is to provide customized information to student which study and live in a university campus. In this study, we collect internal data of campus and external data on the Internet such as the blog or SNS and, then store and process them. After that, we propose a system for providing individual students one-to-one marketing by analyzing these data in detail. This system analyzes purchase history information and the attendance of the building, and then transmits the coupon and information individually according to the pattern to the student's mobile phone.

Key Words : University Campus, SNS, Marketing, Smartphones

1. 서론

90년대 들어서 인터넷이 확산되면서 사람들은 오프라인에서 보다 더욱 쉽게 원하는 정보를 검색가능하게 되었다. 더불어서 특정한 분야에 관심을 가지는 사람들은 본인들이 수집하고 알고 있는 정보를 인터넷 상에 작성하고 저장하게 되었다. 즉, 사람들은 특정한 일을 수행할 때 필요한 정보를 인터넷 상의 홈페이지에 저장된 내용을 검색하여 보다 쉽게 작업을 처리할 수 있게 되었다. 이 후 휴대폰 보급의 확산과 스마트폰의 폭발적인 성장으로 인해 도래한 모바일시대에는 개인화 서비스와 소셜

네트워크 서비스(Social Network Service)가 확산되었고, 이를 통해 다양한 종류의 데이터 즉, 정형화된 데이터 이외에 비정형화된 데이터가 대량으로 생산되고 있다. 이렇게 증가되는 데이터의 양은 엄청나서 기존의 데이터를 저장하고 관리하며, 분석하는 기법으로는 처리하는데 한계가 있어서 이를 위한 새로운 기술이 필요하다. 야후, 구글, 페이스북과 같은 정보검색 및 정보제공업체에서는 이를 대비하기 위해 이전부터 수많은 인력과 연구자금을 투입하여 왔고, 이를 통해 제안되고 개발된 기술들이 관련기업과 기관들에서 도입하여 사용 중이다. 빅 데이터

Received 2016-08-22 Revised 2016-09-05 Accepted 2016-09-09 Published 2016-09-30

*Corresponding author : Shin-Hyeong Choi (cshinh@kangwon.ac.kr)

(Big Data)라는 용어는 이와 같은 정보홍수의 시대를 대표하는 것으로서 이를 관리하는 기법과 더불어서 수많은 사례연구가 진행 중이다.

2. 관련 연구

2.1 빅 데이터의 개념

빅 데이터는 기술적 관점에 따라 다양하게 정의될 수 있으나, 데이터 형식이 다양하고 생성속도가 매우 빨라서 기존의 방식으로는 저장 및 관리 등을 할 수 없으므로 이를 처리하기 위해 새롭게 관리하고 분석할 필요가 있는 대용량 데이터를 말한다[1-4]. 다시 말하면, 빅 데이터는 기존에 일반적으로 사용되는 데이터 수집 및 관리를 위한 소프트웨어의 처리능력을 넘어서는 용량의 데이터를 말하며, 그 크기는 계속해서 변화하는 것이 특징이다. 빅 데이터의 주요 속성으로서는 규모(Volume), 다양성(Variety), 속도(Velocity)를 비롯하여 정확성(Veracity), 시각화(Visualization)와 가치(Value)를 포함한다[5-8].

2.2 빅 데이터의 종류

앞서 언급하였듯이 빅 데이터 측면에서 볼 때, 과거에는 형식이 정해져있는 텍스트 위주의 데이터가 많이 생성되었으나, 최근에는 개인화서비스와 SNS의 확산으로 스마트폰을 포함한 스마트기기를 통해 다양한 종류의 데이터 즉, 이미지, 동영상, 음성 등의 데이터가 폭발적으로 증가한다. 이와 같이 빅 데이터의 종류는 크게 정형데이터, 반정형데이터, 비정형데이터로 나눌 수 있다. 정형데이터는 형식이 정해진 것으로서 일반적으로 관계형데이터베이스에 저장되며, 반정형데이터는 XML과 HTML처

럼 메타데이터나 스키마 등을 포함하는 데이터를 말한다. 또한, 비정형데이터는 고정된 필드에 저장하기 힘든 데이터로서 그림, 동영상과 음성데이터를 말한다[9,10].

2.3 빅 데이터 기술

빅 데이터를 효과적으로 활용하기 위한 기술로서는 분석기술과 표현기술이 있다. 빅 데이터 분석기술은 새로운 것이 아니라 기존에 통계학, 전산학과 기계학습/데이터 마이닝 분야에서 사용되던 기법들로서 최근에 폭발적으로 증가하고 있는 비정형데이터를 효과적으로 분석하기 위해 대량의 데이터 처리에 맞게 수정하여 적용하고 있다. 특히, 최근에 SNS를 통해 생산되는 데이터들은 형식이 갖추어진 것이 아니라 텍스트와 이미지로 구성되어 있으므로 비정형데이터를 분석하기 위해서는 텍스트 마이닝(Text Mining)이나 오피니언 마이닝(Opinion Mining)과 같은 분석기술이 필수적이다. 또한 전 세계인들이 SNS를 통해 묶여지고, 이를 통해 데이터를 주고받으므로 이들 연결 관계를 분석하기 위해 수학의 그래프 이론을 바탕으로 한 분석도 필요하다. 방대한 양의 빅 데이터를 분석한 결과는 상당히 복잡하므로, 이를 문장으로 표현하는 것은 매우 어렵다. 그래서 빅 데이터 분석결과를 쉽게 이해할 수 있도록 표현하는 기술이 필요하다. 이를 위해 도표나 시각화된 이미지 등으로 표현하는 기술이 개발되고 있다[11-14].

2.4 빅 데이터 활용분야

기존의 컴퓨터를 활용한 데이터 관리방식과 달리 빅 데이터는 등장이후 비정형데이터에서 유용한 정보를 찾아내는 첨단 기술로 관심을 받았다. 최근에 있었던 인공이세돌 9단과 ‘알파고’라는 인공지능과의 세기의 바둑대

Table 1. The Case Study of Big Data Application Fields

Field	Domestic	Foreign	Effect
Public	Integrated Disaster Safety Information Network (Ministry of the Interior), Offshore tax evasion information sharing (Customs Administration, National Tax Service)	Intelligent Transportation Systems(Japan), Country Risk Management System(Singapore), Public Participatory Safety Management Platform(England)	Cost-effectiveness and public safety improvement
Finance	Off-site Surveillance Monitoring System (Financial Supervisory Service), Consumption Behavior of Card Users(Credit Card Company)	Money Laundering Tracking System(USA), Car Insurance Estimates(USA)	Improvement of prior detection of illegal and poor situation and the accuracy of loans inspection
Medical	Special Disease and Disorder Prevention and Control(Ministry of Health and Welfare), Drug Safety Early Warning Service(Drug Safety and Risk Management)	Disease Care System through Genes Data Sharing(USA), Epidemic Response Program (England)	Effectiveness for disease prediction and providing measures

결에서 보았듯이, 데이터가 축적되면 이를 통해 학습한 시스템이 엄청난 위력을 발휘한다는 것을 볼 수 있었다. 빅 데이터의 시작은 야후, 구글과 같은 인터넷 기업을 중심으로 시작되었지만, 최근에는 공공, 금융, 의료 등 다양한 분야에서 빅 데이터를 활용하기 위한 시도를 하고 있다[15]. Table 1에서는 세 가지 분야별 국내외 사례를 보여준다[16-20].

Table 1에서 제시된 사례들의 공통점은 기존의 정형화된 데이터뿐만 아니라 다양한 종류의 비정형데이터를 분석하여 보다 신속하고 정확하게 해당분야에 대해 대응할 수 있다는 것이다. 기존에는 공공과 민간이 각각의 서비스 제공을 위해 개별적으로 데이터를 수집하여 분류, 가공, 분석 등을 통해 정보를 제공하여 왔다. 서울특별시의 심야버스 노선 최적화와 부산광역시의 서비스 인구통계개발 등에서 볼 수 있듯이 지방자치단체들은 국민의 수요는 알고 있으나, 상당수의 지방자치단체들은 예산 및 전문 인력의 부족으로 이를 활용하는 시도가 많지는 않다. 이를 해결하기 위해 최근에는 공공기관을 중심으로 축적되는 공공데이터와 기업 등의 민간을 통해 축적되는 민간데이터를 통합해서 서비스를 제공하려는 시도가 많으며, 이를 위해 기술 및 제도적 인프라 구축 등의 준비가 필요한 시점이다.

3. 스마트 캠퍼스 정보

빅 데이터는 기존의 데이터 수집, 분석과 표현과 달리 새로운 방법에 따른 절차가 필요하다. 빅 데이터를 처리하는 과정은 데이터 수집에서부터 저장, 처리, 분석과 표현으로 구분할 수 있다. Fig. 1은 빅 데이터 처리과정을 보여준다.

보다 정확하고 양질의 데이터를 수집하기 위해서는 내부에서 생성되는 데이터 이외에 외부의 데이터를 확보할 방안을 모색해야 한다. 내부데이터로는 자체적으로 축적된 파일시스템 또는 데이터베이스 관리시스템과 센서 및 CCTV 등의 정보가 있으며, 외부데이터로는 인터넷으로 연결된 외부의 공개된 데이터를 비롯하여 휴대폰 이용자 기반의 활동인구 데이터와 신용카드 사용량 기반의 지역상권 매출 데이터 등이 있다.

본 연구에서는 대학교 캠퍼스 내에서 공부하고 생활하는 학생들에게 개인별 맞춤형 정보를 제공하는 것을 목적

으로 한다. 이를 위해서는 캠퍼스 내부에서 수집되는 데이터와 블로그 및 SNS 등의 외부 데이터를 수집한다.

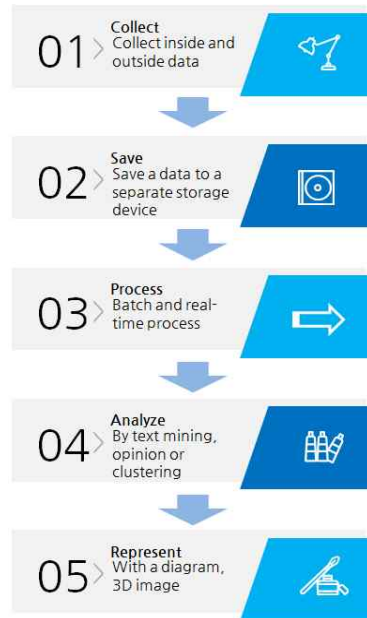


Fig. 1. Big Data Processing

학생들은 안드로이드 기반의 모바일 학생증을 지원하는 스마트폰을 각종 자판기 및 스마트출결시스템 등의 카드 판독기에 터치하면 주문내용과 출석 데이터가 수집된다. 이들 데이터를 분석하여 학생 개인별 일대일 마케팅을 제공하도록 시스템을 구현한다. Fig. 2는 빅 데이터를 이용한 분산처리시스템의 전체 구성도를 나타낸다.

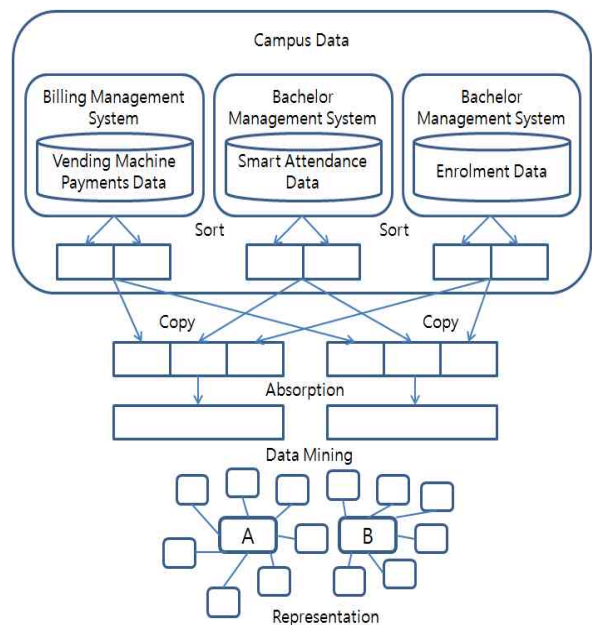


Fig. 2. System Structure

Fig. 1과 Fig. 2의 절차를 기반으로 하둡(Hadoop)을 이용하여 데이터를 처리할 경우에는 다음의 절차에 따라 수행된다.

첫 번째로, 자판기구매이력을 비롯하여 스마트 출결데이터와 학사정보 각각의 DBMS에 있는 데이터를 하둡 기반의 HDFS(Hadoop Distributed File System)로 모은다. HDFS는 하나의 서버에서만 동작하는 게 아니라 여러 개의 서버에 설치되어서 서비스가 되며, 리눅스의 local disk를 이용해 여러 대 이상으로 확장 가능한 구조이다. HDFS는 하나의 NameNode 서버와 Secondary NameNode, 그리고 다수의 DataNode 서버들로 구성된다. 이와 같이 여러 DataNode에 다수의 사본을 유지함으로써 이중화가 가능하다. NameNode는 마스터로서, DataNode는 슬레이브로서의 역할을 수행하는 구조이다. 다음으로는 맵 리듀스(Map Reduce)를 통해 데이터를 분류, 매핑, 재배치 후 추출할 수 있다. 이 과정은 크게 Mapper와 Reducer로 구분할 수 있는데, Mapper는 여러 군데 흩어져있는 데이터를 Key-Value형태로 연관성 있는 데이터로 분류하는 작업이며, Reducer는 Mapper를 통해 나온 결과물 중에서 중복된 데이터를 제거하고 원하는 데이터만을 추출하는 작업이다. Fig. 3은 HDFS로부터 적재된 파일들이 처리되는 과정을 나타낸다.

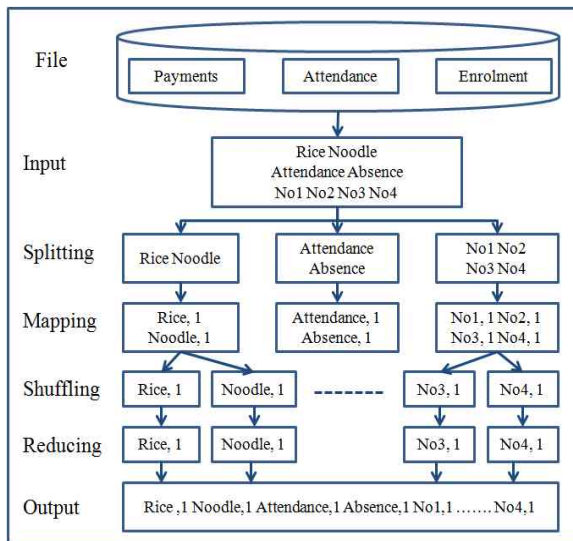


Fig. 3. Data Input and Output Process

이렇게 학생 한 사람 한사람의 구매 이력과 해당 건물의 출결상황을 상세히 분석하여 구매 패턴에 맞춰 재구성한 다음에 학생개인별로 내용이 다른 쿠폰 및 정보를 학생의 휴대폰으로 전송한다. 다음은 학생의 상황에 따

라 제공하는 맞춤형 쿠폰 예시를 나타낸다.

-주중 강의가 있는 요일에 해당 학생이 방문하면 각종 학내 행사, 시험정보와 더불어서 식단제공 및 소비조합 등의 행사쿠폰을 제공함

기존의 서비스는 방문하는 모든 사람들에게 동일한 내용의 쿠폰을 휴대폰에 전송함으로써 해당 학생에게 불필요한 쿠폰으로 인식되었지만 수집, 분석한 빅 데이터를 이용하면 보다 유의한 정보를 학생들에게 제공할 수 있다.

빅 데이터를 성공적으로 활용하기 위해서는 무엇보다도 가치 있는 양질의 대량 데이터를 획득하는 것이 매우 중요하다. 이를 위해 접근 가능한 데이터 소스를 확대할 필요가 있다. 또한 학생들에게 양질의 정보를 제공하기 위해 필요한 데이터 소스를 확보할 수 있는 구체적인 계획을 수립해야 한다.

4. 결론

과거에는 데이터를 저장하기 위해서는 테이블이라는 정형화된 틀을 생성한 다음 표준형식에 맞춰 입력하는 시스템이 일반적이었으며, 대부분 관계형 데이터베이스 관리시스템을 통해 관리하였다. 이와 달리 최근의 스마트 혁명과 함께 일반인들을 통해 생산되는 데이터양은 폭발적으로 증가하고 있으며, 데이터 종류 또한 다양하다. 이와 같은 데이터를 빅 데이터라고 부르며, 이를 저장 및 관리하기 위한 방법이 많이 연구 개발되고 있다. 이런 공공 및 민간의 빅 데이터를 효과적으로 활용하면 보다 양질의 정보를 제공가능하다. 본 연구에서는 캠퍼스 내부에서 수집되는 데이터와 인터넷상의 블로그나 SNS 등의 외부 데이터를 수집하여, 이를 저장하여 처리한 다음, 이들 데이터를 세부적으로 분석하여 학생 개인별로 일대일 마케팅을 제공하는 시스템을 제안한다. 향후 시스템으로 구현하여 이를 통해 학생 개인별로 맞춤형 정보를 제공한다면 보다 유의한 대학생활을 할 수 있을 것이다.

REFERENCES

[1] Ministry of Science, ICT and Future Planning, *Ministry of Public Administration, Security and Ministry of Knowledge Economy, Korea Communications*

- Commission, Executive Committee of the National Ministry of Science and Technology, Big Data Master Plan for the Implementation of Smart Country*, Ministry of Science, ICT and Future Planning 2012.
- [2] Y. S. Jeong, "U-healthcare Service Management SAcheme for Big Data of Patient Information," *Journal of Convergence Society for SMB*, Vol. 5, No. 1, pp. 1-6, Mar. 2015.
- [3] NIA, "Big Data Trends and Policy Implications, Policy of Broadcasting and Telecommunication," <http://www.nia.or.kr>, Vol. 20, 2016. 4.
- [4] J. H. Kim, J. Y. Go and K. H. Lee, "A Scheme of Social Engineering Attacks and Countermeasures Using Big Data based Conversion Voice Phishing," *Journal of the Korea Convergence Society*, Vol. 6, No. 1, pp. 85-90, Feb, 2015.
- [5] E. J. Kim, "A Study on Characteristics of Female Consumers Using Big Data," *Journal of Digital Convergence*, Vol. 13, No. 10, pp. 185-194, Oct. 2015.
- [6] Y. S. Jeong, "Business Process Model for Efficient SMB using Big Data," *Journal of Convergence Society for SMB*, Vol. 5, No. 4, pp. 11-16, Dec. 2015.
- [7] Y. A. Hur and K. H. Lee, "A Study on Countermeasures of Convergence for Big Data and Security Threats to Attack DRDoS in U-Healthcare Device," *Journal of the Korea Convergence Society*, Vol. 6, No. 4, pp. 243-248, Aug, 2015.
- [8] Y. S. Jeong and K. H. Han, "Big Data Processing Scheme of Distribution Environment", *Journal of Digital Convergence*, Vol. 12, No. 6, pp. 311-316, Jun, 2014.
- [9] J. I. Bum and S. J. Choi, *Big Data case of application and Implications*, Nonghyup Economic Research Institute, No. 312, 2013.
- [10] J. S. Han, "Utilization Outlook of Medical Big Data in the Cloud Environment," *Journal of Digital Convergence*, Vol. 12, No. 6, pp. 341-347, Jun. 2014.
- [11] K. H. Choi and J. A. Yoo, "A reviews on the social network analysis using R," *Journal of the Korea Convergence Society*, Vol. 6, No. 1, pp. 77-83, Feb, 2015.
- [12] H. S. Yang and S. B. Kim, "Evaluation Method of Big Data Efficiency", *Journal of Digital Convergence*, Vol. 11, No. 8, pp. 31-39, Aug. 2013.
- [13] J. S. Kim, "A Study on Application and Related Technology of Big Data", *Journal of the Korea Contents Association*, Vol. 10, No. 1, pp. 34-40, Jan. 2012.
- [14] S. H. Lee and D. W. Lee, "Current Status of Big Data Utilization", *Journal of Digital Convergence*, Vol. 11, No. 2, pp. 229-233, Feb, 2013.
- [15] M. Y. Kim and D. J. Seo, "An Analysis of the Public Data for Making the Ambient Intelligent Service", *Journal of Digital Convergence*, Vol. 12, No. 12, pp. 313-321, Dec, 2014.
- [16] Y. B. Cho, S. H. Woo and S. H. Lee, "In Small and Medium Business the Government 3.0-based Big Data Utilization Policy", *Journal of Convergence Society for SMB*, Vol. 3, No. 1, pp. 15-22, May. 2013.
- [17] S. N. Shin, "Current status and challenges on the big data of public sector in Korea," *Journal of the Korean Medical Association*, Vol 57, No. 5, pp. 398-404, May. 2014.
- [18] B. C. Kim, "Big Data Security Technology and Response Study", *Journal of Digital Convergence*, Vol. 11, No. 10, pp. 445-451, Oct, 2013.
- [19] Korea Information Research and Development Institute, A Study on Application Strategy of Private and Public Convergence Big Data through the Analysis of Big Data Application, pp. 7-22, 2014.
- [20] NIA, *National Information Society Agency Future Strategy Center*, 2015 Big Data Global Casebook, 2015.

저 자 소 개

최 신 형(Shin-Hyeong Choi)

[정회원]



- 1993년 2월: 울산대학교 전자계산학과 학사
 - 1995년 2월 : 경남대학교 전자계산학과 석사
 - 2002년 8월 : 경남대학교 컴퓨터공학과 박사
 - 2003년 9월 ~ 현재 : 강원대학교 제어계측공학과 교수
- <관심분야> : 임베디드시스템, 모바일프로그래밍, 정보보안