

빅데이터를 이용한 IoT 활용사례와 발전방향

조영주*, 김진혁*, 소윤정*, 노창희*, 권순필⁰

⁰조선대학교 전자정보공과대학 컴퓨터공학과

e-mail: oblss@naver.com*, 6980kimjh@naver.com*, yj951006@naver.com*, ncg0382@naver.com*, ekexn_@naver.com⁰

Application cases of IoT using big data and its' direction of improvement.

Young-Ju Cho*, Jin-Hyuk Kim*, Yoon-Jeong So*, Chang-Hee No*, Soon-Pil Kwon⁰

⁰Dept. of Computer Engineering, Chosun University

● 요약 ●

빅데이터는 인터넷의 발달로 인하여 규모를 가늠할 수 없을 정도로 많은 양의 정보가 생산되어지는 데이터를 말한다. 빅데이터를 활용하는 다양한 사례 중 기업에서 활용된 사례로 사람들의 구매데이터를 분석하여 개인의 구매 취향을 분석한 뒤 구매자가 원하는 제품을 빠르게 안내하는 시스템이 있으며, 정부에서 빅데이터를 활용한 사례로는 사회복지 자금이 대상자들에게 제대로 지급되고 있는지 판단하여 부정수급자를 원천봉쇄 하는 시스템이 있다. 이처럼 빅데이터가 인간의 삶에 긍정적인 영향을 주는 등 다양한 분야에서 활용되고 있는데 본 논문에서는 빅데이터를 이용한 IoT의 활용사례를 알아보고, 긍정적 사례와 부정적인 활용사례를 분석한 뒤 그 발전방향에 대해 제시하고자 한다.

키워드: IoT(Internet of Things), 빅데이터(Big Data), 보안의식(Security consciousness)

I. Introduction

인터넷의 발달로 인하여 많은 사람들이 편리하게 정보들을 생산하고 이용하게 되었다. 이 정보들을 동시에 여러 사람들이 한꺼번에 많이 생산함으로써 데이터의 양이 폭발적으로 증가하고 있다. 빅데이터는 이처럼 디지털의 확산으로 규모를 가늠할 수 없을 정도로 많은 정보가 생산되어지는 데이터들을 말한다.

다시 말해, 빅 데이터는 디지털 환경에서 생성되는 데이터로 그 규모가 방대하고, 생성 주기도 짧고, 형태도 수치 데이터뿐 아니라 문자와 영상 데이터를 포함하는 대규모 데이터를 말한다. 빅 데이터 환경은 과거에 비해 데이터의 양이 폭증했다는 점과 함께 데이터의 종류도 다양해져 사람들의 행동은 물론 위치

정보와 SNS를 통해 생각과 의견까지 분석하고 예측할 수 있다.

빅데이터의 특성으로는 미국의 정보기술 및 자문회사인 가트너그룹에서 빅데이터의 특성을 3V(Volume, Variety, Velocity)로 정의하였으며, 학계와 산업계의 지속적인 연구와 개념정의를 해서 빅데이터에 대한 정의가 추가되고 있다. 현재는 가치와 정확성이라는 특성을 더해서 5V(Volume, Variety, Velocity, Veracity, Value)라고 정의했다.

그림1에 보듯이 2010년 한해에만 1.2ZB(제타바이트)를 생산되었고 여기서 말한 1ZB(제타바이트)는 1조GB로 변환될 수 있고 정보량은 미국 의회도서관 저장정보량의 약 4백만 배에 해당하는 것을 알 수 있다. 2012년에는 한 해 동안 생성된 데이터의 양이 2.7ZB로 이전 데이터양을 모두 합친 것 보다 많다는 것을 보여주고 있다.

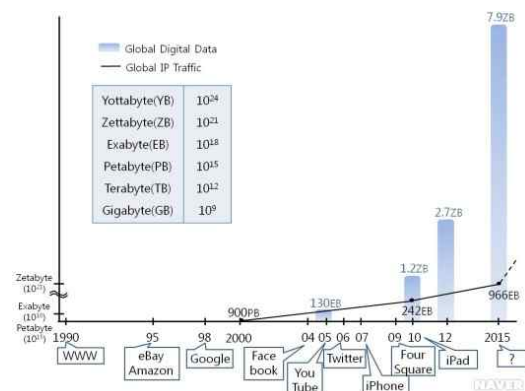


Fig. 1. Data amount by year [1]

본 논문에서는 빅데이터를 이용한 IoT의 활용사례를 알아보고, 긍정적 사례와 부정적인 활용사례를 분석한 뒤 그 발전방향에 대해 제시하고자 한다.

II. Use case

1. Positive use case

1.1 기업의 긍정적 활용사례

1) IBM의 Watson

왓슨은 인공지능으로 빅데이터를 이용해 학습한다. 왓슨은 2011년 미국에서 가장 인기 있는 퀴즈쇼에 나가서 퀴즈쇼의 챔피언과의 대결에서 승리하였다. 퀴즈들의 문제는 굉장히 은유적인 문제들도 있었던 만큼 왓슨에게도 많이 불리하였지만, 왓슨의 4테라바이트의 디스크에 저장된 2억 페이지에 달하는 콘텐츠를 통하여 퀴즈쇼의 챔피언과의 대결에서 승리하였다.

2) Amazon의 추천시스템

온라인 서적 구매 사이트였던 아마존 또한 빅데이터를 활용한 사례가 있다. 아마존은 고객의 서적구매 패턴을 빅데이터로 저장하여 고객이 서적을 구매한 것을 토대로 다음 구매할 것 같은 책을 선정해서 그 책들을 할인해주는 쿠폰들을 발급해준다. 전형적인 데이터 분석에 기반을 둔 마케팅 방법이다. 아마존은 이러한 빅데이터 분석 경험으로 현재 클라우드 서비스를 제공하고 있으며 비정형데이터 처리를 위한 데이터베이스를 개발하는 등 빅데이터 관련 기업으로써 위상을 높이고 있다.

3) Google의 자동번역시스템

구글의 자동번역 시스템은 통계적 기계 번역이라고 한다. 이 시스템은 인터넷의 많은 웹페이지들 중 번역되어 있는 빅데이터를 딥러닝을 통해 학습해서 번역해주는 시스템이다. 웹페이지에 번역되어 있는 자료가 많으면 많을수록 번역의 질은 높아진다. 번역되어 있는 자료가 절대적으로 부족한 아르메니아어, 라틴어는 번역의 질이 높지 않아 테스트 언어로 분류되어 있다.

4) Google의 Google Flue Trends

구글은 검색어 추이를 나타내는 ‘Google Flue Trends’를 통해 어느 지역에서 독감이 유행할 것인지를 예측한다. 그림2는 Google의 Flue Trends를 보여준다.

5) Netflix의 Cine-match

미디어 콘텐츠 유통기업인 넷플릭스(Netflix)는 이용자의 영화 대여 목록에 기초해서 새로운 영화를 추천해주는 시네매치(Cinematch) 시스템을 개발했다. 넷플릭스는 시네매치 시스템의 정확도를 높이기 위해 상금을 걸고 경진대회를 열기도 한다. 넷플릭스의 빅데이터 경영은 경쟁기업인 블록버스터(Blockbuster)를 파산에 이르게 한 동인으로 평가하고 있다.



Fig.2 . Google Flu Trends[2]

6) Youtube의 선호채널

하루 40억 회 이상 동영상상이 검색되는 유튜브도 이용자가 자신이 선호하는 동영상 채널을 구성할 수 있는 개별 홈페이지를 제공하고 있다. 개인별로 동영상 이용 데이터가 축적되면 이를 SNS 정보, 인적 네트워크 정보와 연계해 다양한 개인 맞춤형 서비스를 제공할 수 있다.

7) Zara의 최적 재고관리 시스템

패스트 패션의 선도자인 자라는 현재 유행하는 패션 트렌드를 즉시 반영해 단기간에 다품종 소량 생산하는 초스피드 전략을 채택하고 있다. 이러한 전략을 뒷받침하기 위해서 상품 수요를 예측하고, 실시간 데이터 분석에 의존 한다. 빅데이터 활용을 위해 자라는 MIT 연구팀과 최적 재고관리 시스템을 개발했다.

1.2 정부의 긍정적 활용사례

1) 국내

2010년 2월부터 보건복지부를 포함한 각 부처와 공공기관이 수행하는 복지사업과 수혜자 정보를 통합 관리하는 사회복지통합관리망(그림3)을 가동하고 있다. 정부는 이 시스템을 통해 사회복지 자금들이 대상자들에게 제대로 지급되는지 관리하고 있는데, 현재까지 부정 수급자, 횡령사건을 원천적으로 차단하는 성과를 거두고 있다.



Fig. 3. Social welfare integration management network

2) 싱가포르와 미국

싱가포르와 미국 정부는 보안과 위험관리 분야에 빅데이터를 활용하고 있다. 싱가포르 정부는 재난방재와 테러감지, 전염병 확산과 같은 불확실한 미래를 대비하기 위해 2004년부터 국가위험관리시스템(RAHS, Risk Assessment & Horizon Scanning)을 추진했다. 다양한 국가적 위험 데이터를 수집·분석해 사전에 예측하고 대응방안을 모색하고 있다. 미국 연방수사국(FBI)의 DNA 색인 시스템도 빅데이터 활용사례다. 빅DNA데이터를 활용해 단시간에 범인을 검거하는 시스템을 운영하고 있다. 오바마 정부가 추진한 필박스(Pillbox) 프로젝트는 국립보건원(NIH) 전용 사이트를 통해 의약품 정보 서비스를 제공하고 제조사와 사용자 간 유기적인 정보 공유를 가능하게 했다. 이를 통해 후천성면역결핍증 등 관리 대상 주요 질병의 분포와 증감 현황 데이터를 수집·분석할 수 있게 되었다.

미국 미시간 주정부는 관련 정부기관 통합 데이터웨어하우스(IDW, Integrated Data Warehouse) 구축으로 시민에 대한 보다 나은 서비스를 제공하고 비용을 절감했다. 미시간 주의 21개 정부기관은 데이터 통합을 통해 공공의료보험(Medicaid) 부정행위 발생 감지, 개인 건강관리 개선, 최적의 입양가정 선택 등 공공 서비스 품질 개선에 활용하고 있다. 오하이오 주와 오클라호마 주 정부는 국세청(IRS) 데이터와 고용데이터를 분석해 새로운 세원을 확보하고 미납세금을 확인하고 있다.

3) 영국과 호주

주요국 정부는 정부데이터를 공개하는 전용 사이트를 만들어 데이터를 활용한 새로운 지식을 만들기 위해 노력하고 있다. 영국(<http://data.gov.uk>)과 호주(<http://www.data.gov.au>)는 공공 부문의 데이터 공개를 통해 정부의 투명성을 높이고 국민의 알 권리를 향상시키며 시간과 자원을 절감하는 효과를 지향하고 있다. 이러한 정부 데이터 공개 정책(open data)은 빅데이터 시대에 소통과 공유, 협업(crowdsourcing) 전략이 무엇보다 중요하다는 것을 의미한다.

2. Negative use cases

그림4는 Insecam.org라는 사이트이다. 해킹되어 있는 CCTV들을 각 나라 별로 볼 수 있게 카테고리화 만들었다. 우리나라 카테고리를 클릭하여 CCTV를 보면 8000여개의 CCTV가 무방비로 보이고 있다. 위 사이트는 아직 비밀번호가 설정되어 있지 않음, 또는 초기비밀

번호로 설정되어 있는 CCTV의 보안의식을 강화하기 위해 만들어진 사이트라고 하여도 8000여개의 CCTV가 손쉽게 볼 수 있다는 점을 생각한다면 심각한 문제가 발생할 여지가 있다고 생각된다.

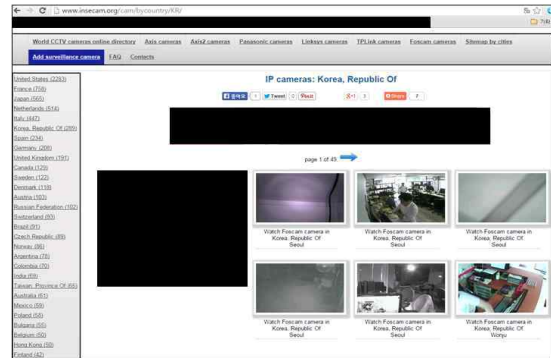


Fig. 4. Insecam.org Web Site [3]

또 다른 부정적인 활용사례로는 사생활침해가 있다. 예를 들어 페이스북이나 트위터에서 일상적으로 하던 대화들이 기업이나 다른 업체들에서 빅데이터로 판단하여 데이터들을 모아 유용한 정보로 활용한다면 개인입장에서는 자신의 구매취향이나 성격, 복용약 등을 빠르게 안내 받을 수 있지만 나의 개인적인 구매취향이나, 나의 상세한 정보들을 누가 보는지 모르는 기업들의 데이터베이스에 있다는 것이 불쾌하게 느껴질 수 있고, CCTV의 사례와 마찬가지로 보안에 관련하여 해킹을 당할 시 스팸문자, 보이스피싱등 많은 문제점을 야기할 수 있다.

III. Conclusions

지금까지 본 논문에서는 많은 사람들이 쉽게 정보를 주고받을 수 있는 현대사회에서 빅데이터라는 주요한 이슈와, 빅데이터를 활용한 긍정적인 사례와 부정적인 사례를 알아보았다. 위 내용에서 보았듯이 긍정적인 사례에서는 빅데이터를 활용하여 기업의 이익을 창출하는 기업의 활용사례와 공익을 창출하는 정부의 활용사례로 나누었다. 기업은 구매자가 원하는 상품을 빠르게 보여주는 시스템을 주요하게 활용하였고, 정부는 정부기금이 제대로 활용되고 있는지, 혹시 잘못 사용되고 있는 곳은 없는지 확인하는 방향으로 활용하고 있다. 부정적인 활용사례는 많은 개인정보보호가 큰 이슈가 되고 예방되어지지 못한 것들을 알아보았다.

이러한 많은 부정적인 사례들이 있음에도 불구하고 부정적인 측면을 제도적, 법률적으로 보완하고 긍정적인 측면을 부각시킨다면 인간의 삶의 질을 향상시키는데 이바지 할 것으로 기대된다. 따라서 본 논문에서는 부정적인 측면을 해결하는 방안으로 세 가지 내용을 제안하고자 한다.

첫째, 법의 제도적인 개편이 필요하다. 현재의 법은 개인정보보호법이 존재하지만 국외와 비교하여 볼 때 미흡한 점이 많다. 너무 포괄적으로 동의하는 부분과, 이미 시행 중인 법률을 소규모 사이트에서는 지켜지지 않고 간과되고 있다.

둘째, 보안의식과 관련된 교육제도의 확립이다.

일반적으로 사용자에게는 보안의식과 관련된 교육적 제도의 부재로

인해 기본적으로 지켜야 할 보안의식마저 결여된 상황이기때 교육제도의 정책이 확립되어야 한다.

셋째, 보안관련 문제가 발생했을 때 그것을 해결할 수 있는 방향을 제시할 필요가 있다. 방법과 절차를 배포할 기관들의 확립이 절실하다.

이와 같이 세 가지 방안이 활성화된다면 안정된 보안의식과 제도와 교육으로 보다 나은 IoT 속 빅데이터 환경이 구축될 것이라 기대된다. 향후 보안의식 분야까지 적용된 IoT 관련 빅데이터 활용방안을 연구, 구현할 계획이다.

References

- [1] Jung Yong-chan, "Big Data Revolution and Media Policy Issues", KISDI Premium Report
- [2] Cho Hyeon-hyun, "Current status of utilization of big data, problems and countermeasures", KERI Column
- [3] Yun Seong Ko, "Analysis of vulnerability through countermeasure of security threat of physical security control system and study of countermeasure", Journal of Korea Multimedia Society Vol. 19, No. 1, January 2016(pp. 51-59)
- [4] Mantika, James, et al. "Big data: The next frontier for innovation, competition, and productivity."(2011)
- [5] Jung Yong-chan (2012b). Big Data, Big Brother. KISDI expert column. 2012.6. Information and Communication Policy Institute.
- [6] Kim, Jang-hyun (2013) A Huge Change in Strategic Communication - Big Data Rumble and After: Toward the Proper Data Age. Korea PR Society 2013 Strategic Communication Seminar 'Strategic Communication for Overcoming the Upheaval'. Dec. 07, 2013.
- [7] <http://blog.naver.com/lool2389/>
- [8] Big Data Utilization Business Model (Big Data Platform Strategy, Feb. 11, 2013, Korea Electronics and Telecommunications Research Institute (ETRI), Electronic Newspaper)
- [9] Big Data Utilization Business Model (Big Data Platform Strategy, Feb. 11, 2013, Korea Electronics and Telecommunications Research Institute (ETRI), Electronic Newspaper)