빅데이터 보안 관리 시스템 적용 방안

전중호*, 박석천**, 김종현***
*가천대학교 일반대학원 모바일소프트웨어학과
**가천대학교 컴퓨터공학과 정교수(교신저자)
***위세아이텍 대표이사
lover030237@naver.com

Applied Method of Security Management System for Big Data

Jung-Ho Jeon*, Seok-Cheon Park**, Jung-Hyun Kim***
*Dept. of Mobile Software, Gachon University
**Dept. of Computer Engineering, Gachon University
***Representative Director, WISEITECH co., ltd

요 약

최근 스마트폰과 같은 모바일 기기의 확산과 SNS 의 성장이 결합되면서 사이버상의 데이터량이 기하급수적으로 증가됨에 따라 빅데이터가 화두로 등장하였으며 빅데이터는 활용방법에 따라 국가 기업 및 개인의 삶의 질을 향상 시킬수 있다. 그러나 빅데이터는 다양한 경로로 데이터를 생성하고 수집함으로서 보안에 대한 이슈가 대두되고 있다.
본 논문에서는 데이터를 생성하고 수집하는 구간에서의 보안관리를 통하여 잠재되어 있는 악성코드의 공격과 개인정보에 대한 안전성을 높이고 신뢰성있는 데이터로 만들어 활용할수 있는 방법에 대해 연구한다.

Ⅰ. 서론

최근 스마트폰과 같은 모바일 기기의 확산과 페이스북과 트위터 같은 SNS의 성장이 결합 되면서 사이버상의 데이터량은 기하급수적으로 증가되고 있다.

이와 같이 데이터량이 기하급수적으로 증가함에 따라 “빅데이터”가 최근 화두로 등장하였으며 세계포럼은 2012년 가장 주목해야할 기술로 빅데이터를 지목하였다.

빅데이터는 규모가 방대하고(Volume),데이터의 종류가 다양하며(Variety),데이터의 흐름이 빠르게 진행되는 속도(Velocity)특성을 가지며 새로운 가치를 창출하여 미래예측을 통한 삶에 질을 높일 것으로 예측된다.

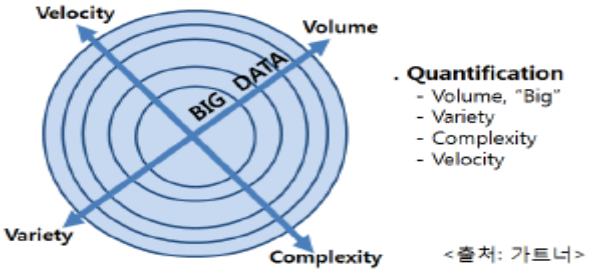
그러나 빅데이터는 인터넷 기반, 트위터, 페이스북 등의 소셜 네트워킹 ,위치기반 서비스, 스마트 폰 등을 통하여 모인 각종 개인정보가 무차별적, 실시간으로 생성되고 수집됨으로써 개인정보보안이 이슈로 떠오르고 있다.

본 연구에서는 빅데이터의 정의와 활용사례 등 빅데이터 기술들을 알아보고 빅데이터에서 개인정보 보안에 대한 연구 및 분석 하고자 한다.

Ⅱ. 빅데이터의 개요

2.1 빅데이터의 정의

빅데이터는 과거 데이터에 비해 규모가 크고 형태가 다양하여 기존의 방법으로 수집, 저장, 분석이 어려운 방대한 크기의 데이터를 의미한다. 기존의 데이터 단위를 넘어서는 방대한 양(Volume), 데이터의 생성과 흐름이 빠르게 진행되는 속도(Velocity) 사진, 동영상 등 기존의 구조화된 데이터가 아닌 다양한 (Variety) 형태의 정보 등 3가지 속성을 가진 데이터가 ‘빅 데이터’ 라는 게 대다수 전문가들의 공통된 의견이다. 가트너는 [그림1]과 같이 3V에 복잡성을 추가해 3V+C로 정의하기도 한다 [5].



[그림1] 빅데이터의 정의

2.2 빅데이터의 활용사례

미국정부에 경우 9.11 테러 이후 국토안보부를 중심으로 테러·범죄 방지를 위한 범정부적 빅 데이터 수집, 분석 및 예측체계를 도입하여 국토 보안에 빅데이터를 활용하고 있다. 또한 의료분야에서는 오바마 Health.20- 필박스 프로젝트(Pillbox)를 통해 수집된 빅 데이터를 통해 후천성 면역 결핍증(HIV) 등 관리대상 주요 질병의 분포, 연도별 증가 등에 대한 통계치 확보가 가능해졌다.

싱가포르 정부는 빈번히 발생하는 테러 및 전염병으로 인한 불확실한 미래 대비를 위하여 2004년부터 빅데이터 기반 위험 관리계획을 추진하고 있다.

공공 분야에서의 빅데이터 활용사례 뿐만 아니라 구글 검색창에 입력되는 발열 기침등의 검색 빈도로 독감을 예보하거나 자라에서는 전 세계 매장의 판매 데이터로 유행을 파악하여 재고를 줄이는 등 빅데이터를 활용, 분석하여 가치 있는 정보를 추출하고 분석된 데이터를 바탕으로 능동적으로 대응하거나 변화를 예측하여 삶의 질과 신뢰도를 높이고 있다.

아래의 [표 1]은 메킨지에서 제시한 빅데이터의 활용 방안이다.

이와 같이 빅데이터는 다양한 분야에서 사용 할 수 있다 [5].

[표1] 메킨지에서 제시한 빅데이터 활용방안

도메인	분석대상 데이터	예상효과
미국의 의료산업	제약사 연구개발 데이터 환자 치료 임상 데이터, 의료산업의 비용 데이터	연간 $3조 연간 0.7% 생산성 향상
유럽의 공공행정	정부의 행정업무에서 발생하는 데이터	연간 £2.5조 연간 0.5% 생산성 향상
소매업	고객의 거래 데이터, 구매 경향	$1조 + 서비스업자 수익 $7조 소비자 이익
제조업	고객 취향 데이터, 수요 예측 데이터, 제조과정 데이터, 센서활용 데이터	60% 마진 증가 0.5~1.0% 생산성 향상
개인 위치 데이터	개인, 차량의 위치 데이터	개발 및 조립비용 50% 감소, 운전자본 7%감소

Ⅲ. 빅데이터의 개인정보 보안

3.1 개인정보의 필요성

앞에서 언급한 것과 같이 빅데이터는 활용방법에 따라 국가, 기업, 개인에게 다양한 효과를 얻을 수 있다.

그러나 빅데이터에 많은 수가 개인 IT 단말기를 통해 생성되어 수집되기 때문에 개인정보가 노출되거나 개인 데이터가 무분별하게 상업적으로 이용될 수 있으며 다양한 보안위협이 존재한다. 지난 2011년에 네이트 사이트에서 3,500만 건의 개인정보가 유출되었으며 넥슨사 사이트에서는 1300만 건의 회원정도가 불법 유출된 사고가 발생한바 있다. 이처럼 개인정보는 유출 될 때 다량의 정보가 유출되기 때문에, 빅데이터에 개인정보 침해사고가 발생한다면 지금까지의 발생했던 침해사고 보다도 더 큰 피해를 가지고 올 것이다.

3.2 빅데이터 보안위협

트랜드 마이크로가 발표한 2011년 1분기 보안위협 보고서에 의하면 APT 위협과 모바일 기기에 대한 사이버 범죄가 증가할 것으로 예상하였다. 대표적으로는 APT 위협과 BYOD위협이 있다.

지능형 지속 위협은 단발성 공격이 아니며 공격대상네트워크에 침투하여 목적이 달성될 때까지 지속적으로 공격하는 고도화된 보안위협을 의미한다. 외부에 공개된 정보나 이전의 공격에서 얻은 데이터를 바탕으로 진행하기 때문에 공격대상에 관해 더 자세히 알수록 공격은 정교해진다. 전형적인 APT 위협의 예로서 사회 공학적 기법을 이용하여 악성코드가 내포되어 있는 Email 등을 전송하여 클릭을 유도하고 클릭 시, 악성코드를 실행하여 시스템을 점령하고 내부 시스템에 잠복한다. 악성코드는 외부 서버에서 명령을 전달받아 수행되며 장기간에 걸쳐 내부정보 등을 유출하게 된다.

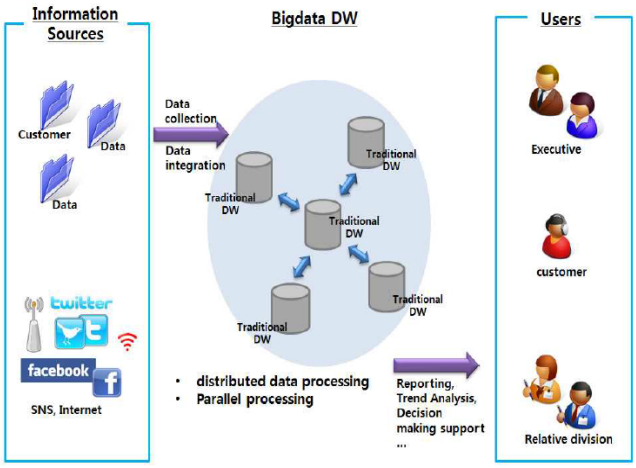
BYOD(Bring Your Own Device)는 개인소유의 IT 단말기를 업무에 활용하는 현상을 의미한다. PC 위주의 주요업무와 개인 태블릿 PC, 스마트폰을 보조적 수단으로 업무에 활용하는 최신 유형이다. 이러한 현상은 하드웨어와 소프트웨어의 발전이 더욱 가속시키고 있다. BYOD를 이용한 업무생산성, 편의성과 같은 긍정적인 요소 외에 잠재적 보안위협이 존재한다. 대표적 보안위협에는 기업의 IT 통제권 상실, 단말기의 취약점 및 악성코드로 인한 기업 내부정보 유출 위협, 악성코드에 감염된 개인용 기기의 내부 접속으로 인한 기업 IT자산위협, 단말기 도난 또는 분실로 인한 데이터 유출, 보안의식이 낮은 직원에 의한 계정 유출 등이 있다[1].

3.3 빅데이터 처리구간

[그림2]에서처럼 빅데이터의 생성부터 서비스에 이르기까지 세 단계로 나뉜다.

첫 번째 과정은 여러 소스를 통해 생산된 데이터를 수집하는 과정, 두 번째는 분산처리 및 병렬처리

를 위해 데이터를 분산 저장 및 운영 과정, 마지막으로 데이터 분석 및 2차 데이터 생성을 통해 서비스로 재사용되는 과V 정이다[2].



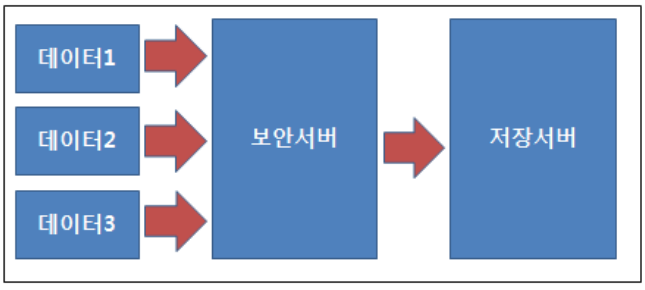
[그림 2] 빅데이터 서비스 개요

Ⅳ. 빅데이터 보안관리 시스템 적용방안

빅데이터는 다양한 경로를 통하여 데이터를 생성 및 수집하기 때문에 개인정보의 노출 및 다양한 보안위협에 노출되어 있다.

기존의 빅데이터의 데이터를 수집하는 구간에서는 전자서명, 다양한 필터링 기법. 스팸 메일방지 등의 기술들을 적용하여 보안 위협에 대처하고 있지만 빅데이터의 많은 수가 IT 단말을 통해 생성되어 수집 되면서 개인의 데이터가 유출될 수 있다.

이점을 착안하여 본 논문에서는 데이터의 수집구간에서 보안 서버를 설치함으로서 개인정보 보안에 대한 신뢰성을 높이고자 한다.



[그림 3]빅데이터 보안관리 시스템 구성

[그림 3]은 본 논문에서 제안하는 수집구간의 개인정보보안서버를 설치함 으로서 다양한 경로로 들어오는 생성되는 데이터들이 보안서버를 통과하여 저장서버에 저장 된다. 보안서버를 통과하는 동안 데이터에 잠재되어 있는 악성코드를 필터링 하거나 개인정보를 암호화 하여 신뢰성 있는 데이터를 제공한다.

Ⅴ. 결 론

최근 정부의 빅데이터 시범사업을 본격적으로서 추진됨에 따라 국내의 빅데이터 활용도도 점점 높아지고 있다.

빅데이터는 활용방안에 따라 국가, 기업, 개인의 삶의 질을 높여줄 수 있다.

그러나 다양한 경로로부터 데이터를 생성 수집함에 따라 개인정보의 유출이나 보안문제가 발생될 수 있다 .

따라서 본 논문에서는 데이터를 생성, 수집하는 과정에서 보안서버를 설치함으로서 개인정보 암호화와 잠재되는 악성코드를 제거하여 데이터의 신뢰성을 높이고 안전한 데이터를 제공해 줄 수 있도록 제안하였다.

향후, 빅데이터 보안관리 시스템을 구현하고 테스트를 진행하여 생성된 결과를 바탕으로 제안 시스템을 수정 보안 할 예정이다.

Ⅵ. 사사의 글

본 연구는 2013년도 지식경제부의 SW전문인력양성사업의 재원으로 정보통신산업진흥원의 고용계약형 SW석사과정 지원사업(HB301-13-1003)으로부터 지원받아 수행 되었습니다.

참고문헌

[1]최대수,김용민 ,“빅 데이터와 통합보안 2.0”,정보처리 과학회 ,2012

[2]정교일,박한나,정부금,장종수,정명애,“빅데이터와_정보보안”, 정보기술학회지

[3]이용수, “스마트혁명 시대 빅데이터 활용과 프라이버시 사이의 충돌에 관한 연구”,경원대학교,2011

[4]신영진, “공공분야의 빅데이터 추진과 개인정보보호에 관한 연구

[5]김정숙,“빅데이터 활용과 관련기술 고찰”, 한국콘텐츠학회,2012