



A Proposal of a Defense Model for the Abnormal Data Collection using Trace Back Information in Big Data Environments

Yan-Ting Mu¹, Hyun-Chul Baek², Jae-Yeong Choi¹, Won-Chang Jeong³, Sang-Bok Kim^{*1}

¹Department of Computer Science, Gyeongsang National University

²Department of Internet Information Technology, Gyeongnam Provincial Namhae College

³Department of Medicinal & Welfare Information, Jinju Health College

ABSTRACT

There are mass productions and collections of information in different fields with real time in our present society. Big data is not limited as a specific single system in its information collection process and produces useful information through collecting different data from multiple systems. Thus, there are rapid increases in users who abuse such information through accessing specific distributed systems by collecting and combining the information. Therefore, it is considered that damages caused by the information abuse are very serious. In this study, trace back information has primarily used in a process for collecting data in the different information environments that manage the data as a distributed storage manner in order to react currently increased illegal data collections. In addition, the data is to be classified according to values of the data and the values are also evaluated and managed through applying proper security classes. Then, the data that causes significant damages from combinations of it due to leaks of the data is to be managed. Based on the management, a way of performing stable information provides and collections under the current situation that represents rapidly and differently changed environments for collecting information is proposed.

© 2015 KKITS All rights reserved

KEYWORDS : Big data, Big data security, Cloud Computing, Computer networks, Trace router.

ARTICLE INFO: Received 23 February 2015, Revised 3 April 2015, Accepted 14 April 2015.

*Corresponding author is with the Department of Computer Science, Gyeongsang National University, 501, Jinju-Gajwa-Dong, Jinju-si, Gyeongsangnam-do, 660-701,

KOREA.

E-mail address: sbkim@gnu.ac.kr

1. 서론

오늘 날 컴퓨터 네트워킹 환경은 초기의 단순한 송/수신자간의 연결 환경에서 인터넷 환경을 거쳐 클라우드 컴퓨팅 환경으로 접어들었다. 빅 데이터란 이러한 클라우드 컴퓨팅 기술을 기반으로 하는 정보 수집환경을 의미한다. 클라우드 컴퓨팅 기술에 기반한 빅 데이터 환경에서의 정보 수집은 기존 인터넷 환경에서의 정보 서비스 차원을 넘어 다양한 정보 시스템으로부터 필요 정보 수집과 이를 통한 새로운 정보들을 만들어 내는 과정을 보이고 있다. 그러므로 빅 데이터를 위한 정보 제공 시스템들은 불법적인 공격자들의 집중적인 표적이 되고 있는 실정이다.[4][6][10] 본 논문에서는 이러한 다수의 정보 제공 시스템과 이들이 보유하고 있는 중요 정보 유출을 막기 위하여 다음과 같은 방법을 적용하였다. 먼저 공격자들의 불법적인 접근에 대비하여 트래이스 백 정보를 이용하여 정상적인 사용자 여부를 검증할 수 있도록 하였다.[11] 그 다음 각 정보 제공 시스템에서 보유하고 있는 자료들에 대하여 가중치를 두어 해당 정보들의 조합에 대한 보안 레벨을 설정하였다. 이를 통하여 일차적으로 각 정보 시스템에 대한 접근 허용 여부를 통제 가능하도록 하였다. 그 다음 각 사용자들이 정보 시스템으로 접근했던 자료들을 분석한 후 정상적인 정보 생성을 위하여 상호 협조 하고 있는 각 시스템으로 통보한다. 이렇게 각 시스템으로 통보된 자료들은 각각의 정보 시스템에서 보유하고 있는 자료들의 보안 레벨과 비교 분석하여 이상 징후로 판단되면 마지막 자료에 대한 연결을 단절하여 중요 자료에 대한 유출을 막을 수 있도록 하였다.[7][8][9] 본 논문의 구성은 다음과 같다. 2장에서 빅 데이터와 관련된 연구를 살펴보고, 3장에서는 빅 데이터 환경에서의 보안 모델을 제안하고 그 동작 과정을 설명하였다. 4장에서는 이에 대한 시뮬레이션과 제안 모델의 보

안 레벨에 대한 단계별 보안 정책을 수행하였다. 그리고 결론에서 향후 본 논문의 응용 가능성에 대한 언급을 하였다.

2. 관련연구

2.1 빅 데이터의 개념

빅 데이터에 대한 정의는 다양하게 나타나고 있다. 최근 여러 연구 보고서를 보게 되면 빅 데이터의 정의를 다음과 같이 분석하고 있다. 먼저 기존 분석도구나 시스템 체계에서 처리 가능한 범위를 넘어선 데이터 환경을 빅 데이터로 정의하고 있다. 빅 데이터에 대한 또 다른 시각은 빅 데이터를 다양한 종류의 대규모 데이터로부터 저렴한 비용으로 가치 추출이 가능하고, 필요 데이터에 대한 빠른 수집과 발굴 및 분석을 할 수 있는 차세대 기술 및 아키텍처를 의미한다.

이와 같이 빅 데이터에 대한 개념 정의는 보는 시각에 따라 그 차이를 보일 수 있다. 그렇지만 일반적으로 빅 데이터가 의미하는 내용은 다음과 같다. 빅 데이터란 대용량의 데이터를 저장, 수집, 발굴, 분석, 비즈니스화 하는 일련의 과정이며, 기존 데이터 처리에 비해 복잡도와 비정형 데이터의 비중이 높으며 정형화된 분석 모델이 없기 때문에 그 유연성이 높다고 할 수 있다. 그러므로 빅 데이터 환경에서의 정보 수집 과정은 다양한 네트워크 및 시스템 환경을 이용할 수밖에 없다. [2][3][5]

2.2 빅 데이터 환경에서의 보안 이슈

빅 데이터 서비스 환경은 앞에서 언급 했듯이 단일 시스템 및 네트워크 환경이 아닌 클라우드 컴퓨팅 환경을 기반으로 하고 있다. 그러므로 지금까지 일반적으로 적용되고 있는 단일 시스템 및

네트워크 환경에서 구축 사용되고 있는 보안 시스템의 환경을 넘어서 각 시스템과 네트워크 상호간 협력 방어 체계를 필요로 하고 있다. 그렇지만 이런 다양한 문제점을 해결하기 위한 보안 시스템 구축은 현실적으로 아주 어렵다고 할 수 있다.

2.3 클라우드 컴퓨팅의 개념

클라우드 컴퓨팅의 정의는 각 기관이나 학자마다 그 견해가 다르며, 새로운 컨셉이 자주 등장하기 때문에 기존의 정의가 새로운 정의로 대체되기도 한다. 다만, 공통된 점을 분석하여 정의해 보면 ‘네트워크 환경에서 이용자의 요구에 따라 실시간으로 소프트웨어, 플랫폼, 인프라 등 IT 자원이 필요한 만큼 공급되고 그에 따른 비용을 지불하는 서비스’ 라고 정의할 수 있다. 이러한 클라우드 컴퓨팅 환경은 개인이나 기업, 공공기관 등의 보안 시스템 환경의 변화도 요구하고 있다. 아울러 공격자들의 공격기법도 빠르게 변하고 있다. 그러므로 클라우드 컴퓨팅 환경 하에서 발생 가능한 다양하고 불법적인 공격으로부터 중요한 정보 자산을 보호하기 위해 클라우드 컴퓨팅 환경에 대한 많은 이해가 요구되고 있는 실정이다[1].

2.4 트래이스 백 정보의 개념

현재 일반적인 인터넷 망은 출발지에서 목적지까지 여러 개의 경로를 거쳐 연결되는 과정을 가지고 있다. 트래이스 백 정보는 이러한 인터넷 경로 분석을 위하여 사용하는 프로그램이다. 즉, 특정 컴퓨터가 인터넷 망을 통하여 최종 목적지에 도착할 때 까지 지나가는 각 구간 경로 정보를 기록하는 프로그램이다. 그러므로 본 논문에서는 트래이스 백 정보를 이용하여 각 구간 경로 정보를 분석한 후 원격 접속을 시도하는 사용자들에 대하

여 인증 과정을 수행하고 있다.

3. 제안 모델 설계

3.1 제안 모델

본 논문에서 제안하는 불법적인 정보수집에 대한 방어 모델은 다음 식과 같은 구조를 가지며, 이를 식 1, 식 2, <그림 1>로 나타내었다.

$$\text{시스템 A} = \{a_1, a_2, \dots, a_n\} \quad (1)$$

$$\text{시스템 B} = \{b_1, b_2, \dots, b_n\}$$

$$\text{시스템 Z} = \{z_1, z_2, \dots, z_n\}$$

$$\text{자료 조합} \quad (2)$$

$$A_i = \{a_1, a_2\}, \{a_1, b_1\}, \dots, \{a_m, b_n, z_k\}$$

$$B_i = \{b_1, b_2\}, \{b_1, a_2\}, \dots, \{b_n, a_m, z_k\}$$

$$Z_i = \{z_1, z_2\}, \{z_1, a_2\}, \dots, \{z_k, a_m, z_k\}$$

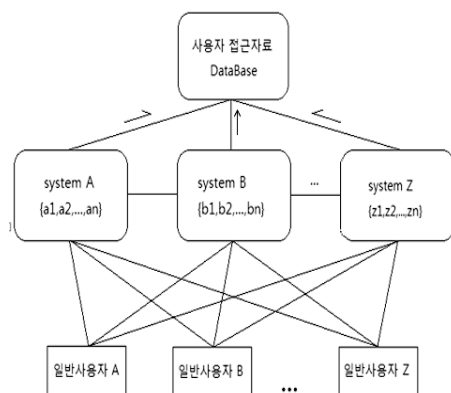


그림 1. 사용자 접근자료 정보 구축
Figure. 1 Establishing user access information

다음 <그림 2>와 <그림 3>은 본 논문에서 구축 사용한 각 대학의 학생정보 자료구조와 해당 자료 구조에 대한 항목별 보안 레벨에 대한 자료구조의 예를 각각 나타낸 것으로 본 논문에서는 일부 항목만 사용하여 시뮬레이션 하였다.

Univ_name	char [20]
stu_name	char [20]
stu_ju_num	char [13]
stu_p_name	char [20]
stu_address	char [20]
stu_hobby	char [20]
stu_special	char [20]

그림 2. 대학 학생정보 자료구조

Figure. 2 Student Information Data structure in University

Univ_name	char [20]
stu_name_Lev	char [1]
stu_ju_num_Lev	char [1]
stu_p_name_Lev	char [1]
stu_address_Lev	char [1]
stu_hobby_Lev	char [1]
stu_special_Lev	char [1]

그림 3. 항목별 보안 레벨 자료구조

Figure. 3 Itemized security level Data Structure

빅 데이터 환경에서는 일반적으로 여러 시스템이 보유하고 있는 자료들을 조합하여 유용한 정보 생성이 가능하다. 예를 들어 대학생인 홍길동과 홍길서, 홍길남이는 홍길부라는 동일 아버지를 두고 있다. 그리고 홍길동은 지역 소재 대학에 재학중이고, 홍길서는 다른 지역, 홍길남 또 다른 지역의 대학에 재학 중이라고 가정 했을 때, 세 학생의 기초 정보는 해당 학교 학사 정보에 등록하게 된다. 본 논문에서는 세 학생의 개인 정보에 각각 등급

을 부여한 후 이를 이용하여 불법적인 정보 수집에 대한 방어 모델을 제안하고자 한다. 먼저 세 학생의 학사 정보 중 1등급 분류 정보에는 주민등록번호가 있다. 그 다음 2등급 정보에는 가족관계가 있다. 즉 홍길동, 홍길서, 홍길남의 아버지는 홍길부로 각각의 학사 정보에 등록되어 있음을 의미한다. 마지막으로 일반정보에는 세 학생의 취미나 특기 같은 정보가 있다. 현재 이들 정보 중 2등급 가족관계는 자식과 부모 관계 이외에는 다른 정보를 제공할 수 없는 상태이다. 그렇지만 이를 불법적인 목적으로 이용하기 위해 각 대학의 학사 정보에 접근하여 해당 가족관계 정보를 획득하게 되면, 이를 통하여 세 학생은 홍길부라는 동일 아버지를 두고 있는 형제 사이라는 정보가 도출될 수 있는 것이다. 이와 같이 심각한 정보 조합이 언제든 발생할 수 있기 때문에 본 논문에서는 이들 자료들에 등급을 부여한 후 이들 2등급 자료 접근시 사용자 접근자료 데이터베이스에 등록 관리하도록 하였다. 그리고 이들 내용을 좀 더 일반화 시킨 것을 앞에서 식 1, 식 2, <그림 1>을 통하여 나타내고 이를 위한 자료구조는 <그림 2>와 <그림 3>에 나타내었다. 즉, 각 대학 시스템 A, B, C에서 관리하는 자료들에 대하여 일반 사용자들이 접근했던 자료 목록 중 2등급 정보는 다른 2등급 정보와 조합이 발생할 경우 1등급 정보로 전환 가능한 자료들이다. 그러므로 2등급 자료에 대한 접근이 발생하면 이들 정보를 사용자 접근자료 데이터베이스에 생성해 두도록 하였다. 그 다음 다른 서버에 접근하여 추가로 2등급에 대한 서비스 요청이 발생하면 사용자 접근자료 데이터베이스를 참조하여 해당 2등급 자료에 대한 조합 발생 여부를 분석할 수 있도록 하였다. 본 논문에서는 해당 정보의 일치 여부 자료를 세 학생의 가족 주소를 이용하였다. 아울러 각 시스템 자료들의 보안등급은 보유하고 있는 자료들에 대하여 일련번호가 앞선 것을

보안 등급이 높은 것으로 가정하였다.

예를 들어 {a1}, {b1}...{z1}의 경우 시스템 A와 시스템 B, 시스템 C에서 가장 중요하게 관리하는 자료들이므로 해당 자료를 서비스 할 경우에는 원타임 패스워드를 발생시켜 이에 대한 인증 과정을 바로 거치도록 한다. 그 다음 각 시스템에서 관리하는 자료 중 2등급 자료에 대한 서비스 요청이 발생하게 되면 첫 번째는 바로 서비스를 해주지만 이에 대한 접근 정보를 <그림 1>의 사용자 접근 자료 데이터베이스에 해당 내용을 등록을 해 둔다. 그리고 이를 이용하여 추가적인 2등급에 대한 서비스 요구가 발생하면 원타임 패스워드를 발생시켜 새로운 인증 과정을 거치도록 한다. 이런 과정을 수행하기 위하여 각 시스템에서 관리하는 자료들에 대한 보안 등급을 적절하게 조정하여 각 시스템에서 보유하도록 한다. 이에 대한 처리 과정은 3.2 제안 모델 동작 과정에 나타내었다.

3.2 제안 모델 동작과정

본 논문에서 제안하고 있는 모델의 사용자 접근 처리 과정은 <그림 4>와 같다. 먼저 사용자 접근이 발생하면 트레이스 백 정보를 이용하여 정상 사용자 유무를 검사한다. 정상 사용자란 트레이스 백 정보를 이용하여 구축해 놓은 정상 사용자들의 접근 경로 정보와 일치하는 사용자들의 접근을 의미한다. 이때 정상 사용자 접근으로 판정되면 서비스 접근을 허용하고, 일치하지 않으면 원타임 패스워드를 발생시켜 접근 여부를 결정한다. 사용자 접근 정보가 상이한 경우는 불법적인 사용자가 IP 스푸핑을 이용하여 접근을 시도하는 경우와, 정상적인 사용자 접근이지만 접근 위치가 일반적인 사용자 접근 위치를 이탈하여 접근하는 경우가 있다. 그러므로 이러한 경우에는 서비스 가용성을 향상시키기 위해 원타임 패스워드를 이용하여 접근 경로

정보가 바뀌었다 하더라도, 서비스를 지속적으로 유지할 수 있도록 한 것이다.

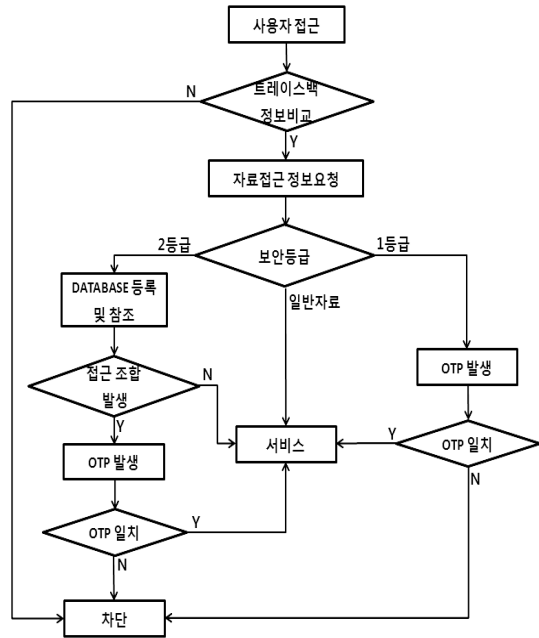


그림 4. 제안모델 동작과정 플로우차트
Figure. 4 Operational process flow chart in the proposed model

접근 경로 정보에 대한 인증 과정을 통과하면 보유하고 있는 자료의 서비스를 실시한다. 빅 데이터 환경에서 사용자들은 2대 이상의 서버를 통하여 유용한 정보를 가공해 낸다. 그러므로 자신이 보유하고 있는 자료뿐만 아니라 다른 협력 서버에 저장하고 있는 자료들을 고려한 자료 서비스 조합 정보를 구축하여야 한다. 본 논문에서는 이들 정보를 일반적인 서비스 자료, 자신이 관리하는 중요 자료, 다른 자료와 조합이 되는 경우 중요 정보로 사용 가능한 자료로 분류하여 이들을 각각 1등급 자료, 2등급 자료, 일반 자료로 구축하였으며, 이는 다음 <그림 5> 각각 구축하였다. <그림 5>에 정의해 놓은 자료는 사용자의 자료 접근 요청이 발생하면 트레이스 백 정보 비교를 통하여 1차적인 사

용자 인증 과정을 거쳤기 때문에 등급별 서비스를 실시할 자료들이다.

A서버 자료	
이름 : 장길수	등록번호 : 9508041945123
부도 : 장호명	
주소 : 서울시	
전화 : 중동	
특기 : 축구	
이름 : 이민주	등록번호 : 9607072925123
부도 : 이정일	
주소 : 대구시	
전화 : 등산	
특기 : 수영	
이름 : 홍길동	등록번호 : 9509041925342
부도 : 홍길부	
주소 : 진주시	
전화 : 운동	
특기 : 서예	
B서버 자료	
이름 : 강정실	등록번호 : 9312121925134
부도 : 강미구	
주소 : 진주시	
전화 : 남서	
특기 : 불향	
이름 : 홍길서	등록번호 : 9312121923415
부도 : 홍길부	
주소 : 진주시	
전화 : 등산	
특기 : 축구	
이름 : 박미정	등록번호 : 9508052912345
부도 : 박정수	
주소 : 부산시	
전화 : 독서	
특기 : 피아노	
C서버 자료	
이름 : 이정남	등록번호 : 9612191924134
부도 : 이우성	
주소 : 대구시	
전화 : 등산	
특기 : 피아노	
이름 : 홍길남	등록번호 : 9806241925345
부도 : 홍길부	
주소 : 진주시	
전화 : 만화보기	
특기 : 달리기	
이름 : 박미주	등록번호 : 9402132912345
부도 : 박정수	
주소 : 창원시	
전화 : 여행	
특기 : 여행	

그림 5. 분산 자료 등급 정의
Figure. 5 Definition of distributed information classes

1등급에 대한 자료 요청이 발생하면 중요 자료이므로 OTP를 발생시켜 검증을 한 후 서비스 여부를 결정한다. 다음 2등급 자료에 대한 자료 요청이 발생하면 이들 접근 정보를 상호 협력체계에 있는

시스템들이 공동으로 구축해 놓은 접근 자료 데이터베이스에 등록을 하고, 이들 사용자에게 대한 모니터링을 실시한다. 아울러 원타임 패스워드를 발생시켜 인증 과정을 거친 후 서비스 여부를 결정한다. 그러므로 다른 협력 시스템에서는 자료 서비스 요청이 발생하면, <그림 1>의 사용자 접근 자료 데이터베이스 정보를 확인하여 추가적인 서비스 유무에 대한 대응이 가능하다. 이렇게 모든 접근 경로 정보와 원타임 패스워드 인증 과정을 거친 후 자료들에 대한 서비스가 이루어지고 있기 때문에, 현재 빅 데이터 환경에서 중요한 자료들에 대한 관리 문제와 일반적인 자료에 대한 관리 및 서비스 가용성 문제를 개선 할 수 있다.

4. 실험 및 평가

4.1 시뮬레이션 환경

본 논문에서 제안하고 있는 불법적인 정보 수집에 대한 방어 시스템 모델과 관련한 시뮬레이션 환경은 다음과 같다. 먼저 사용된 응용 소프트웨어는 VisualStudio 6.0, 구현언어는 Visual C++을 사용하였다. 시뮬레이션을 위한 운영 체제는 WindowsXP 이고, 시스템 사양은 8GB 메모리를 채택한 Xeon E5506 2.13 Ghz Dual System으로 구성하였다.

DB서버 등록 자료	
ID : leejae5764	
이름 : 이민주	
부도 : 이정일	
주소 : 대구시	
ID : gun2015	
이름 : 홍길동	
부도 : 홍길부	
주소 : 진주시	
ID : kin2356	
이름 : 강정실	
부도 : 강미구	
주소 : 진주시	

그림 6. 사용자 접근 자료 데이터베이스 예
Figure. 6 Example of user access information database

<그림 6>은 빅데이터 환경에서 정상적인 사용자들이 상호 필요에 의해서 2등급 자료들에 대한 접근을 시도 했을 경우 이에 대한 자료를 사용자 접근 자료 데이터베이스에 등록해 놓은 자료구조의 예이다.

다음 <그림 7>은 정상적인 사용자 ‘gun2015’의 접근이 정상적으로 성공을 한 경우를 의미한다.

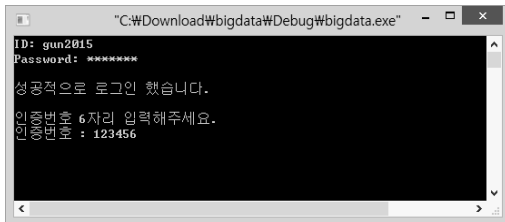


그림 7. 사용자 일반 접근시 동작 과정
Figure. 7 Operational process for general accesses of users

<그림 8>은 ‘gun2015’가 정상적으로 A서버에 접근 후 <그림 5>에 정의되어 있는 자료를 기반으로 일반등급 자료에 대한 서비스 요청이 정상적으로 이루어졌음을 보여주고 있다.

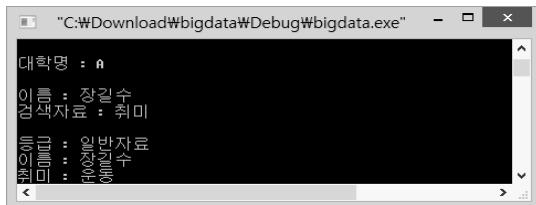


그림 8. 일반자료 접근시 동작 과정
Figure. 8 Operational process for general data information accesses

<그림 9>는 ‘gun2015’가 추가로 C서버의 2등급 자료를 요구하는 경우, 원타임 패스워드를 발생시킨 후 일치하면 자료 접근을 허용하고 서비스를 실시하는 것을 보이고 있다. 이 때 2등급 자료 요청이 발생했기 때문에 <그림 1>의 사용자 접근자료 데이터베이스에 등록이 발생한다.

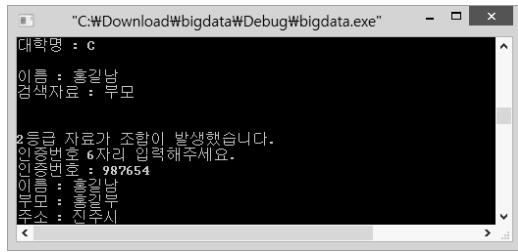


그림 9. 2등급 자료 접근 및 조합시 동작 과정-1
Figure. 9 Operational process for class II information accesses and combinations-1

<그림 10>은 ‘gun2015’가 추가로 B 서버로 2등급 자료를 요구하는 경우 원타임 패스워드를 발생시킨 후 불일치하는 경우 자료 접근을 차단하는 과정을 보이고 있다.

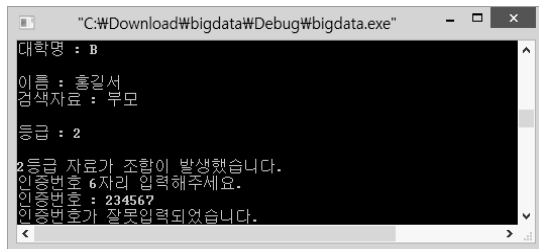


그림 10. 2등급 자료 접근 및 조합시 동작 과정-2
Figure. 10 Operational process for class II information accesses and combinations-2

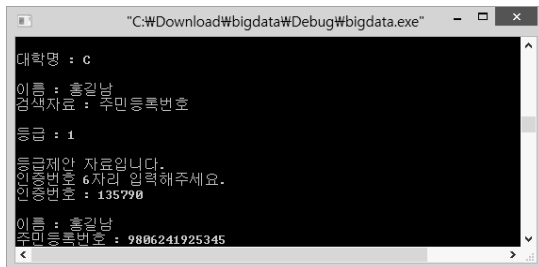


그림 11. 1등급 자료 접근시 동작 과정
Figure. 11 Operational process for class I information accesses

마지막으로 <그림 11>은 1등급 자료이기 때문에 <그림 1>의 사용자 접근자료 데이터베이스에 등록

없이 바로 원타임 패스워드를 발생시킨 후 이에 대한 처리 결과를 보이고 있다.

5. 결 론

본 논문은 현재 우리사회에 대두하고 있는 빅 데이터 환경에서 여러 시스템에 분산 저장되어 관리되는 자료들 각각에 대하여 보안 수준을 정의하였다. 아울러 분산 저장 관리되는 자료들에 대한 조합이 발생하여 보안 수준을 상향시켜야 하는 경우, 이에 대한 보안 정책까지 고려하였다. 그러므로 현재 넘쳐나는 다양한 정보들이 여러 시스템에 분산 저장 운영 될 경우에 대비한 보안 정책 적용도 가능하도록 하였다. 이는 오늘 날 다양한 빅 데이터 환경에 응용이 가능하리라고 본다. 향후 연구 과제로는 분산 저장되는 자료들의 보안 레벨 정의를 위하여 이를 자동화 할 수 있는 연구가 함께 이루어져야 할 것으로 본다.

References

- [1] O. Chen, and O-n.Deng, *Cloud computing and its key techniques*, Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China, Vol. 29, No. 9, pp. 2562~2567, 2009.
- [2] Q.Miao, *When intelligence meeting wity big data : Review and perceptions of big Data'S hotspot intelligence tracking*, Institute of Scientific&Technical Information of Shanghai, Shanghai 200031, No. 5, Serial No. 187, 2013.
- [3] J-K. Park, *A study on measures to active cultural contents service in big data age*, Vol. 20, No. 1, pp. 324~334, Mar. 2014.
- [4] X-F. Meng, and X-B. Ci, *Data management: Concepts, techniques and challenges*, School of Information, Renmin University of China, Beijing 100872, pp. 146~169, 2013.
- [5] S-Y Kim, J-I Lim, and K-h Lee, *A study on the security policy improvement using the big data*, Korea University, Graduate School of Information Security, Vol. 23, No. 5, pp. 969~976, 2013, <http://dx.doi.org/10.13089/JKIISC.2013.23.5.96>
- [6] J.z. Li, and X.M. Liu *An important aspect of big data : Data usability*, School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, pp. 1147~1162, 2013.
- [7] Y-Z. Wang, X-L. Jin, and X-Q. Cheng, *Network big data : Present and future*, Key Laboratory of Web Data Science&Technology. Institute of Computing Technology, Chinese Academ of Sciences, Beijing 100190, Vol. 36, No. 60, 2013.
- [8] Y.Zhang, M.Chen, and X.F.Liao, *Big data applications : A surVey*, School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, Vol. 50, pp. 216~233, 2013.
- [9] D-G. Feng, M. Zhang, and L. Hao, *Big data security and privacy protection*, Trusted Computing and Information Assurance Laboratory, Institute of Soitware, Chinese Academy of Sciences, Beijing 100190, Vol. 37, No. 1, pp. 246~258, 2014.
- [10] S. Curry, E. Kirda, E. Schwartz, and W. H. Stewart, A. Yoran, *Big data fuels intelligence-driven security*, RSA Security Brief, Jan. 2013.

- [11] H-D. Lee, H-T. Ha, H-C Baek, C-G. Kim, and S-B Kim, *Efficient detection and defence model against IP spoofing attack through cooperation of trusted hosts*, Journal of the Korea Institute of Information and Communication Engineering, Vol. 24, No. 12, pp. 2649~2656, 2012.

빅 데이터 환경에서의 트래이스 백 정보를 이용한 불법적인 정보 수집에 대한 방어 모델 제안

무엔팅¹, 백현철², 최재영¹, 정원창³, 김상복¹

¹경상대학교 컴퓨터과학과

²경남도립남해대학 인터넷정보기술과

³진주보건대학 의약복지정보계열,

요 약

오늘 날 우리사회는 다양한 분야에서 실시간으로 정보의 대량 생산과 수집이 이루어지고 있다. 빅 데이터란 정보의 수집과정이 특정 하나의 시스템으로 한정되지 않고, 다수의 시스템으로 부터 다양한 자료를 수집하여 유용한 정보를 만들어 내는 것을 의미한다. 이에 따라 해당 분산 시스템으로 불법적으로 접근한 후 자료를 수집하고 이들을 조합하여 악용하는 사용자들이 계속하여 빠르게 증가하고 있는 실정이다. 그러므로 이로 인한 피해는 아주 심각하다고 할 수 있다. 본 논문에서는 급격하게 증가하고 있는 불법적인 자료 수집에 대응하기 위하여, 분산저장 관리되는 다양한 정보 환경에서의 자료 수집을 위한 접근 과정에 1차적으로 트래이스 백 정보를 이용하였다. 이는 접근 경로 정보를 이용하여 정상적인 사용자 접근 여부를 판정할 수 있도록 하기 위함이다. 아울러 해당 자료들의 가치 정도에 따라 이들을 각각 분류한 후, 각 자료의 가치를 평가하고 적절한 수준의 보안 등급을 적용하여 관리 할 수 있도록 하였다. 그 다음 이들 자료들에 대한 조합이 발생하여 유출시 큰 피해가 예상되는 자료들에 대한 관리도 가능하도록 하였다. 이를 통하여 오늘 날 빠르고 다양하게 변하고 있는 정보 수집 환경에서 안정적인 정보 제공과 수집이 이루어 질 수 있도록 제안하였다.



Yan Ting Mu received the Master's degree in the Department of Computer Science from Gyeongsang National University in 2013. His current research interests include network architecture, bigdata security, network security.

E-mail address: muyanting@naver.com



Hyun Chul Baek received the Ph.D. degree in the Department of Computer Science from Gyeongsang National University in 2003. He was a chairman in the

Committee of Computer System technology at The Korea Association of Regional Public Hospital in 2007. He has been a professor in the Department of Internet Information Technology, Gyeongnam Provincial Namhae College since 2013. His current research interests include network, network security, encryption, bigdata security, cloud computing. He is a member of the KKITS.

E-mail address: dosi_gas@lycos.co.kr



Jae Yeong Choi received the Master's degree in the Department of Computer Science from Gyeongsang National University in 2014. His current research

interests include network architecture, bigdata security, network security.

E-mail address: jyoungc67@naver.com



Won Chang Jeong

received the Ph.D. degree in the Department of Computer Science from Gyeongsang National University in 2009. He has been a professor in the Department of Medicinal & Welfare Information, Jinju Health College since 2001. His current research interests include network, network security, Internet of Things, Automatic Meter Reading.

E-mail address: jwcbblue@hanmail.net



Sang Bok Kim

received the Ph.D. degree in the Department of Electronics Engineering from Chung-ang University in 1989. He was a director in the Department of Education Information Computer Center at The Gyeongsang National University from 2007 to 2010. He has been a professor in the Department of Computer Science at Gyeongsang National University since 1984. He has been a researcher in the Computer Data Communication Research Institute at The Gyeongsang National University since 1984. His current research interests include computer network and security, computer system architecture. He is a member of the KKITS.

E-mail address: sbkim@gnu.ac.kr