Journal of The Korea Institute of Information Security & Cryptology(JKIISC) ISSN 1598-3986(Print)
VOL.23, NO.5, October 2013 ISSN 2288-2715(Online)
http://dx.doi.org/10.13089/JKIISC.2013.23.5.969

빅데이터를 이용한 보안정책 개선에 관한 연구

김 송 영,† 김 요 셉, 임 종 인, 이 경 호‡
고려대학교 정보보호대학원

A study on the security policy improvement using the big data

Song-young Kim,† Joseph Kim, Jong-in Lim, Kyung-ho Lee‡
Korea University, Graduate School of Information Security

요 약

조직이 보유한 정보보호시스템들은 모든 취약점, 침입, 자료유출 등을 탐지하는 것을 목적으로 하고 있다. 그에 따라, 기업은 조직구성원들의 모든 행동이 어떤 경로에서든지 기록되고 확인할 수 있도록 하는 시스템들을 지속적으로 도입하고 있다. 반면에 이것을 관리하고 이 시스템들에서 생성되는 보안로그들을 분석하는 것은 더욱 어려워지고 있다. 보안시스템들을 관리하고 로그를 분석하는 대부분의 인원은 현업의 정보유통 프로세스와 중요정보의 관리절차에 대해 사용자, 또는 유출자보다 알기 어렵다. 이러한 현실은 내부정보유출의 심각성을 더 키우고 있다고 할 수 있다. 최근 빅데이터에 관한 연구가 활발히 진행 되면서 다양한 분야에서 성공사례들을 발표하고 있다. 본 연구는 빅데이터 처리기술과 활용사례를 보안 분야에 적용하여, 기존에 분석할 수 없었던 대용량 정보를 좀 더 효과적으로 분석가능하도록 할 수 있었던 사례와 효율적으로 보안관리 업무를 개선할 수 있는 방안을 제시하고자 한다.

ABSTRACT

The information protection systems of company are intended to detect all weak points, intrusion, document drain. All actions of people in company are recorded and can check persistently. On the other hand, what analyze security log generated by these systems becomes more difficult. Most staff who manages the security systems, and analyze log is more incomprehensible than a user or a person of drain for an information distribution process of the work-site operations and the management procedure of the critical information. Such a reality say the serious nature of the internal information leakage that can be brought up more. While the research on the big data proceeds actively recently, the successful cases are being announced in the various areas. This research is going to present the improved big data processing technology and case of the security field.

Keywords: Big Data; Internal Information Leakage; Data Mining

I. 서 론

최근 중·대기업들의 보안사고가 연이어지면서 사회적인 이슈가 되고 있고, 이러한 조직의 정보나 데이터 유출은 외부자 침입에 의한 것보다 내부자의 침입에 의해서 발생하는 경우가 더 많은 것으로 나타나고 있다. 국가정보원 산업기밀보호센터에서 발표한 통계자료에 따르면, 최근 5년간 발생한 국내 핵심 기술 유출 사건 202건 중 80%에 해당하는 162건이 내부자(전·현직 직원)에 의해 발생되었다.[1] 이는 현재의 단위 보안시스템에 대한 분석 및 관리가 외부의 위협에는 효과적일 수 있으나 내부의 위협에 대해서는 여전히 기업의 핵심기술을 보호하기 어렵다는 사실을 시사한다.

접수일(2013년 4월 1일), 수정일(1차: 2013년 6월 3일, 2차: 2013년 7월 12일), 게재확정일(2013년 8월 29일)
† 주저자, donotcry0@gmail.com
‡ 교신저자, kevinlee@korea.ac.kr(Corresponding author)

이에 따라 기업에서는 내부자에 의한 기업정보 유출 사고 발생을 방지하기 위하여 보안 조직을 강화하고 내부구성원에 대한 보안의식교육을 실시하는 한편, 업무시스템 사용이력 보관, 단일화된 사용자 인증(Single Sign On), 시스템 및 정보에 대한 권한별 접근 통제 등의 다양한 대책을 수립하고 각종 보안 활동을 수행하고 있다. 하지만 이러한 노력을 무력화시키기 위해 나날이 지능적으로 발전하고 있는 내부자 위협에 대응하기 위해서는 정보유출 가능성이 높은 내부자들에 대한 정밀 모니터링의 필요성 또한 높아지고 있다.〔2-3〕

본 연구는 각 보안솔루션들의 편향적인 결과만을 분석하던 기존의 보안정책을 개선 하고자 빅데이터 처리를 위한 오픈소스 중 하나인 하둡(Hadoop)〔4〕 파일 시스템을 이용해서 기존의 다양한 보안 솔루션들의 데이터를 하나로 통합하고, 하이브(Nosql)〔5〕를 이용해서 통합된 데이터를 분석했다. 〔표 1〕의 통합시스템 성능지표는 구축 전의 RDB 환경에서는 분석이 불가능했던 대용량 데이터(1년, 약229TB)를 위와 같은 환경을 구축한 후에는 비교적 빠른 시간에 분석이 가능하다는 것을 보여준다.

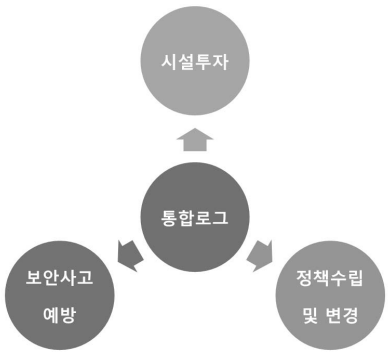
〔표 1〕 통합시스템 성능 지표

		기존 (억건)	통합 (억건)	
테스트항목		1개월 (68)	1개월 (68)	3개월 (202)
특정 URL 검출	5개	불가능	54분	166분
	10개	불가능	55분	172분
특정 목적지 IP 검출	5개	불가능	55분	176분
	10개	불가능	54분	167분
최대 접속 URL TOP 10		불가능	57분	171분

또한, 정보유출 방지를 위한 상용솔루션 도입 및 설비 투자의 의사결정에 도움이 되는 사실적 근거 자료 수집에 통합로그를 활용 할 수 있는 방안을 마련하였다. 기존 보안인프라 투자의 근거는 사고사례에 대한 후속조치로서, 사업장의 규모와 근무하는 임직원, 관리 포인트가 되는 장비들의 갯수만으로 의사결정이 진행되었다. 하지만 각 사업장의 특징을 고려하지 않은 획일적인 투자는 ROI(투자대비효과)에 대한 부담을 가지고 있으며, 실제 사고가 발생하기 전까지는 의사결정권자인 경영진을 설득하기 어렵다.

이번 보안정책 개선에 관한 연구는 단순히 대용량 로그 조회를 위해 사용되었던 통합로그 시스템을 개선해서 보안정책 수립을 위한 근거자료를 추출하고, 〔그림1〕과 같은 의사결정 시 활용가능 하도록 하는 것을 목적으로 한다.

1) 보안사고 예방을 위한 위험징후 탐지
2) 보안시설 투자 예측
3) 보안정책 수립 및 변경에 관한 근거 제시



〔그림 1〕 통합로그 활용방안

II. 관련연구

최근 보안시스템 로그관리 분야의 동향은 통합보안관리(ESM)에서 위험관리시스템(RMS), 보안정보 및 이벤트 관리(SIEM)에 이르기까지 지속적으로 진화하고 있다. 이 가운데 로그 통합관리를 위해 등장한 SIEM은 이기종 환경의 인프라 및 보안로그를 효율적으로 통합운영하고 리스크를 낮추기 위한 해결책으로 받아들여지고 있다.〔6〕 하지만 SIEM은 각각의 단위 보안장비들을 통합 관리하기 위한 수단으로 이해되는 것이 옳다. 기업 보안에 있어서 단순히 SIEM 시스템을 구매하는 관점으로 접근하여서는 적절한 성과를 얻기 힘들다. 이는 사용하는 목적이나 기대하는 기능이 각각의 기업이 처해진 상황별로 다를 수 있기 때문이다. 시나리오 및 패턴 기반의 외부위협을 탐지하는 ESM의 기술적인 계보를 이어받은 SIEM은 침해대응체계 분야에서 그 실효성이 입증되고 있으나 내부 위협자를 선별하고 사용자 행동에 따른 이상징후를 탐지하기 위한 방법으로서는 그 한계성이 지속적으로 논의되고 있다. 특히 내부 위협자의 정보유출 가능 경로를 추측해서 이를 임계치 기반의 시나리오로 적용하는 단순한 접근방법은 내부 위협자 선별에 많은 오탐이 발생된다. 또한 무분별하게 많은 시나리오가 적용될 경우, 내부 위협자를 식별하는 변별력을 저하시키게 되어 보안 활동의 효율성을 저하시키고 있다. 기술적

측면에서는, 분석 되는 보안 장비들 간의 연동 표준 미비, 대부분의 기업 업무가 정보시스템을 활용하게 됨으로써 기하급수적으로 늘어나고 있는 대용량 로그 통합분석의 어려움이 문제로 대두되고 있다.〔7〕 현재 100% 완벽한 보안장비는 없으며 과거의 기술과 사고의 연장선상에서 대응하기 보다는, 보다 창의적인 대응방법이 필요하다. 이를 위해서는 사용자의 행위를 기록하고 있는 빅데이터 분석을 중심으로 한 지능형 보안 시스템 구축이 필요하다.〔6〕 가트너 그룹에서는 빅데이터 분석을 활용한 보안 분석을 통하여 예전에는 보이지 않았던 사고패턴을 발견하고, 정보보안을 포함한 기업경영에 대한 선명한 통찰력을 제공함으로써 기업의 비즈니스 가치를 높일 수 있다고 예측하고 있다.〔8〕

III. 고도화된 보안사고 모니터링 방법론

고도화된 보안사고 모니터링의 체계를 구축하기 위해서는 현재 기업에서 사용하고 있는 네트워크, DRM, 저장매체, 서버보안, 물리보안 등과 관련된 보안 솔루션 로그 정보를 사용자 표준화 과정을 거쳐 통합한 후 내부정보 유출과 관련된 Rule 시나리오를 통해 개인별 행위 분석을 실시한다. 개인별 행위 분석은 개인별 보안 위험도를 산출하는 가장 근간이 되는 기준으로 통상 행위 분석과 이상 행위 분석으로 구분된다.

통상행위 분석은 개인의 통상 업무 패턴을 분석하는 과정으로 〔표 2〕의 『통상행위 기본정의서』에 정의된 평가 항목에 대해서 통상 업무 패턴을 수치화 하는 과정이다. 개인별 통상 행위 분석은 위험도 분석 시 통상 행위를 벗어나는 이탈 행위를 정의하기 위한 기준 데이터로 사용된다.

이상행위 분석은 보안 위협으로 판단 할 수 있는 비정상적인 행위를 분석하는 과정으로, 〔표 3〕의 『이상행위 기본 정의서』에 정의된 평가 항목에 대해서 이상 행위 건수를 수치화 하는 과정이다.

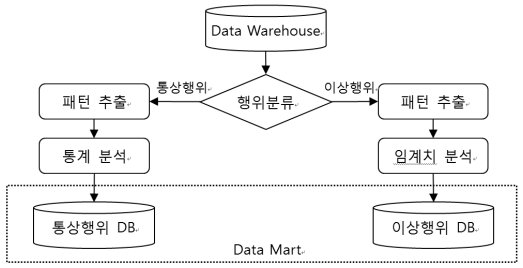
〔표 2〕 통상행위 기본 정의서 샘플 예시

통상행위 평가 항목 정의	상세 내용
근무시간	평균 출퇴근 시간 및 근무시간
인쇄 건수	평균 인쇄 건수
문서 생성 건수	DRM을 통한 문서 생성 건수
업무 서버 사용량	특정 업무 서버 평균 Upload/Download 사용량
인터넷 사용량	TCP/UDP에 대한 인터넷 평균 사용량

〔표 3〕 이상행위 기본 정의서 샘플 예시

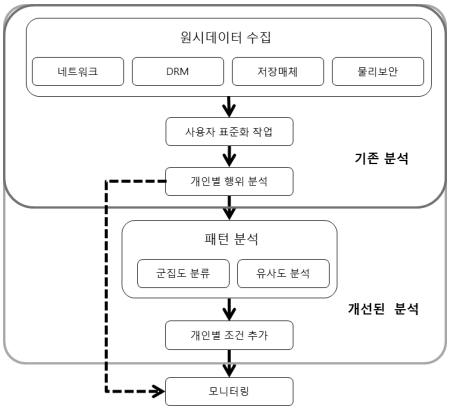
이상행위 평가 항목 정의	상세 내용
자기 자신에게 메일 발송	발송된 메일의 수신자가 발신자 Mail-ID와 동일한 경우
동일 IP에서 다중 Mail-ID 검출	동일 IP에서 2개 이상의 사내 Mail-ID가 검출된 경우
IP 변경	허가 받지 않고 IP를 변경하는 경우
OS 재설치	PC의 OS를 재설치 하는 경우
외부 저장 매체 사용	외부 저장 매체 사용 및 차단 이력

〔그림 2〕와 같이 통합된 로그들을 개인별 행위 분석 프로세스를 이용하여 통상행위와 이상행위로 구분하고 추출된 데이터를 이용하여 위험의 사전관리를 실시한다.



〔그림 2〕 개인별 행위 분석 프로세스

본 장에서는 〔그림 3〕과 같이 기존에 추출된 개인별 행위 분석 결과를 토대로, 실제 위험자와 유사한 패턴을 보이는 대상자를 재분석 하여 기존의 분석 방식보다 효율적인 사고추적 및 증적분석이 가능한 프로세스를 제시하고자 한다.



〔그림 3〕 고도화된 내부정보 유출방지 분석 프로세스

3.1. 고도화된 내부정보 유출방지 분석 프로세스

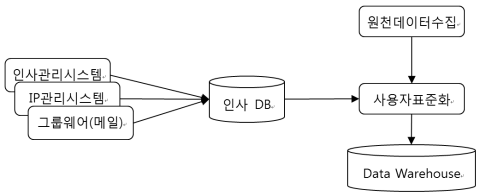
고도화된 내부정보 유출 징후분석 프로세스는 개별 보안시스템에 쌓여있는 대용량 로그를 통합 분석 한 후, 내부자의 사고패턴을 발견함[9]으로써, 전체임직원수의 일부에 불과한 관리가능영역에 내부정보 유출 징후자가 최대한 포함되도록 하는 것을 목적으로 한다.

3.1.1 원시데이터 수집

개인별 보안위험도를 분석하는 기준이 되는 행위에 대한 원천 데이터를 의미하며 대상은 기업에서 사용하고 있는 네트워크, DRM, 저장매체, 물리보안 관리 시스템 등에 대한 사용이력을 모두 포함한다.

3.1.2 사용자 표준화 작업

대부분의 보안 시스템마다 인사정보를 연동하는 기준이 되는 인사관리 시스템이 상이한 경우가 많으며, 또한 유동 IP를 사용하는 경우 대부분 실시간 연동이 아닌 주기에 의한 일괄 연동으로 동기화 하고 있으므로 연동 시점에 모든 보안 시스템의 인사정보가 정확히 동기화 된다고 보기 어렵다. 이 경우 비표준화된 서로 다른 시스템에서 추출된 인사정보는 그 신뢰성이 보장되지 않으며, 개인별 행위 분석 시 상이한 사용자에 대해서 동일인으로 취급될 경우, 그 개인별 보안 위험도는 신뢰성을 보장하기 힘들다. 따라서 [그림 4]와 같이 수집된 원시데이터에 인사관리 시스템, IP관리 시스템, 그룹웨어 시스템 등 의 DB를 연동해서 사용자 표준화 과정을 거친다.



[그림 4] 사용자 표준화 과정

3.1.3 개인별 행위 분석

내부정보 유출과 관련된 Rule시나리오를 통해 개인별 행위 분석을 실시한다. 개인별 행위 분석은 개인별 보안 위험도를 산출하는 근간이 되는 기준으로 통상행위 분석과 이상행위 분석으로 구분된다.

3.1.4 패턴분석

패턴분석은 실제 사고자의 개인별 행위 분석결과를 군집도와 유사도 분석을 이용하여 패턴을 추출하는 과정이다.

3.1.4.1 군집도(K-means Clustering) 분류

군집도 분류는 대상들이 갖고 있는 특성에 기초해서 유사한 성질을 갖고 있는 대상들을 동일한 집단으로 분류하는 기법이다. 이를 통해 개인별 행위 분석결과의 그룹을 선별해서 그룹 내의 패턴을 역산 분석함으로써, 실제 사고자가 포함되어 있는 대상 그룹의 개인별 행위 분석 결과가 공통적으로 포함되어 있는 통상/이상행위 리스트의 군집도를 확인한다.

3.1.4.2 유사도(Similarity) 분석

유사도는 두 개체가 닮은 정도에 대한 수치적인 척도이다. 군집도 분석을 통해 실제 사고자가 속해 있는 그룹의 개인별 특징을 정의한 후 다차원 변수의 유사도 측정 방식인 유클리드(Euclid) 거리 공식을 적용하여 실제 사고자와 전체 임직원의 유사도를 추출한다.

$$d(p,q)=\sqrt{\sum_{i=1}^{n}(p-q_i)^2} \quad (1)$$

p : 정보유출 행위자 (사고자)
q : 전체 임직원
n : 전체 임직원 수
i : 동일 기간 내의 비교데이터[그림 5]

이번 연구에 사용된 유사도 측정 방법은 [그림 5]와 같이 비교대상으로 선정된 실제 사고자와 전체 임직원들의 개인별 행위 분석의 비교 기간을 동일하게

[그림 5] 유사패턴 분석 방식

맞추기 위하여, 기준 데이터와 비교 데이터의 연속적인 기간 내의 유사도 값을 모두 추출하였으며, 그 중 가장 유사도가 높은 값을 각각의 임직원의 대표 값으로 선출하였다. 또한 기준 데이터보다 비교데이터가 적을 시에는 측정에서 제외하였다.

3.1.5 개인조건 추가

확보된 정보유출 위협자(사고자)의 사고발생전 행동패턴, 입사년도, 문서 사용 패턴등을 분석 후 개인조건에 대한 가중치로 활용한다.

〔표 4〕 가중치 부여 예시

정보유출 위협자와의 유사도 차이	가중치
동일행동 패턴	0.2
±1	0.1
±2	0.05
±3	0.025
기타	없음

$$P = MAX(d(p,q))$$
$$R = P + w_1 + w_2 + \ldots + w_n \quad (2)$$

P : 비교 대상자의 기간별 최대 유사도 값
w: 실제 사고자의 개인조건 (근무년수, 문서삭제량)
R : 개인조건 값을 추가한 최종 결과

3.1.6 모니터링

고도화된 내부정보 유출방지 프로세스의 마지막 단계에서는 최종적으로 추출된 실제 사고자의 패턴 분석 결과에 기반한 유사 행위자를 내림차순으로 정렬함으로써 유사도가 높은 내부 위험자들을 주요 모니터링 대상으로 선정해서 집중 관리 할 수 있다.

Ⅳ. 고도화된 보안사고 탐지 모니터링 구축사례

4.1 고도화된 모니터링 케이스 스터디 개선율

이번 케이스 스터디의 효율성 검증을 위하여 내부정보 유출 사고자 대신 실제 퇴직자를 이용한 실험을 진행하였다.

〔표 5〕 고도화된 내부정보 유출 방식의 개선율

실제 퇴직자	기존 분석 방식	개선된 분석 방식	개선율
251명	33명 (13.1%)	116명 (46.2%)	33.1%

〔표 5〕는 기존의 개인별 행위 분석방식과 고도화된 내부정보 유출 분석 방식의 적중률 차이로서, 케이스 스터디 실험을 통해 실제 퇴직자에 대한 유사 패턴의 추출 결과로 전체 임직원의 상위 20%를 각각의 방법으로 추출하였을 때 퇴직자가 포함되어 있는 수치를 비교하였으며, 분석방식에 따른 개선율은 33.1%로 나타났다.

4.2 고도화된 보안사고 모니터링 케이스 스터디

Ⅲ. 고도화된 보안사고 모니터링 방법론에서 언급하였던 통상/이상행위 정의서의 항목 중 퇴직과 관련이 있는 31개의 개인별 행위 추출 시나리오를 대상으로 1건이라도 추출된 적이 있는 조사대상자를 실험데이터와 검증데이터로 구분하여 사용하였다. 〔표 6〕의 실제 퇴직자 중 조사대상자로 선정되지 않은 인원은 퇴직 전 출산휴가, 파견, 출장, 기타이유 등으로 인해 PC를 사용하지 않아서, 시스템 로그로는 분석이 불가능한 데이터를 포함하고 있다.

〔표 6〕 검증 데이터

	실제 퇴직자	선정된 조사 대상자
실험 데이터(9,10,11월)	169명	128명
검증 데이터(12,1월)	137명	123명
합 계	306명	251명

〔표 7〕은 전체임직원 30,518명 중 9월~11월 퇴직자와 유사도가 높은 순서대로 정렬한 결과를 상위 1/10/20%로 추출한 후, 12월~1월의 데이터를 검증데이터로 사용하여 패턴분석 결과의 적중률을 확인했다. 확인결과, 패턴분석 기반으로 정렬한 전체 임직원

〔표 7〕 유사 행위 패턴 유사도 추출

	9월	10월	11월	12월	1월	합계	비율 (%)
퇴직자 (명)	28	58	42	87	36	251	-
유사도 상위1%	2	5	3	4	2	16	6.3
유사도 상위10%	4	7	6	10	7	34	13.5
유사도 상위 20%	5	11	9	15	8	48	19.1

의 상위 20%(6,104명)이내에 퇴직자 251명 중 내부임직원 48명이 결과에 속해 있었으며, 이는 전체 퇴직자의 19.1%에 해당한다.

전체 내부자의 퇴직자 추이를 분석 하였을 때, 진급과 관련된 4/8/12/18년 그리고 25년 이상에서 퇴직 확률이 크게 증가 하고 있음을 알 수 있었고, 실제 퇴직자의 DRM 문서 사용 패턴을 분석 하였을 때, 퇴직 전 3개월 기간 내에 문서 삭제량의 증가가 약 71% 로 나타났다.

〔표 8〕은 유사행위 패턴으로 추출된 유사도 값에 근무년수의 관련성을 개인조건으로 추가한 결과로, 전체 임직원의 상위 20%(6,104명)이내에 퇴직자 251명 중 내부임직원 64명이 결과에 속해 있었으며, 이는 전체 퇴직자의 25.4%에 해당한다.

〔표 8〕 유사 행위 패턴 유사도 추출 +근무년수

	9월	10월	11월	12월	1월	합계	비율(%)
퇴직자 (명)	28	58	42	87	36	251	-
유사도 상위1%	2	1	4	3	1	11	4.3
유사도 상위 10%	6	8	11	17	8	50	19.9
유사도 상위 20%	7	12	14	21	10	64	25.4

〔표 9〕는 〔표 8〕의 결과에 문서삭제량을 개인조건으로 추가한 결과로, 전체 임직원의 상위 20%(6,104명)이내에 퇴직자 251명 중 내부임직원 116명이 결과에 속해 있었으며, 이는 전체 퇴직자의 46.2%에 해당한다.

〔표 9〕 유사 행위 패턴 유사도 추출 +근무년수+문서삭제량

	9월	10월	11월	12월	1월	합계	비율(%)
퇴직자 (명)	28	58	42	87	36	251	-
유사도 상위1%	1	4	4	2	2	13	5.1
유사도 상위 10%	7	10	11	17	8	53	21.1
유사도 상위 20%	13	31	23	35	14	116	46.2

〔표 10〕은 최종적으로 계산된 유사행위 패턴 추출 적중률로써 기업 내에서 동일한 근무년수 그리고 문서삭제량의 조건을 추가 하였을 때 기존방식대비 적중률이 27.1% 향상 되었음을 보여준다.

이 결과는 위의 조건들이 실제 퇴직관련 패턴과 관련성이 높으며, 더 많은 개인조건을 반영하거나, 유사도의 범위를 넓히면 더 높은 적중률을 얻을 수 있다는 것을 보여 주고 있다. 여기서, 유사도의 범위에 따라

〔표 10〕 특이패턴 반영에 따른 위험자 포함비율의 변화

	유사도	유사도+근무년수	유사도+근무년수+문서삭제량
위험자 포함비율	19.1	25.4	46.2

정책적용 인원조정이 가능하므로, 투입 가능한 보안자원에 따른 임계치 정의가 가능하다.

4.3 케이스 스터디 결과

〔표 11〕의 실험결과와 같이 전체임직원 30,518명 중 상위 20%(6,104명)를 기준으로 볼 때, 대상자 128명 중 약 52.3%에 해당하는 67명의 실제 퇴직자를 포함시킬 수 있었으며, 검증 대상인 최근 2개월에 대해 동일한 패턴을 적용한 결과, 대상자 123명 중 39.8%에 해당하는 49명을 포함시킬 수 있었다. 전체 실험기간에서는 전체 퇴직자251명 중 46.2%에 해당하는 116명을 포함시킬 수 있었다.

실험의 데이터로 사용한 퇴직이라는 위험행위는 이직할 회사를 알아보다가 최종행동으로 옮기지 않을 가능성이 높으므로, 상위 20%안에 추출된 검토대상자는 퇴직여부를 갈등하고 있는 인원이 포함 되어있으므로, 보안정책 수립 및 인력관리에 대한 근거로 활용 가능하다.

〔표 11〕 실험 결과

	퇴직자	선별된 20%에 포함된 퇴직자	적중률
표본 데이터 (9,10,11월)	128명	67명	52.3%
검증 데이터 (12,1월)	123명	49명	39.8%
합 계	251명	116명	46.2%

V. 정책 수립

5.1 보안시설 투자 예측

통합로그 시스템의 데이터는 각각의 보안솔루션들의 데이터를 통합하여 임직원 개개인 또는 속해 있는 그룹의 데이터를 한 곳으로 집중시켜 관리 할 수 있다. 이는 제품별 투자나 수요예측 시 별도의 통계를 추출 후 반영하였던 기존방식에서 벗어나 효율적인 투자를 가능하게 하였다.

1) 임직원의 이동경로를 활용한 사내 외 출입동선 변경
2) 입출입 빈도수를 활용한 사무실레이아웃 변경
3) 같은 구역의 물리 보안장비(스피드게이트) 사용 빈도 통계를 활용한 장비 추가 및 위치변경
4) 그룹별 사용량 분석 추이를 반영한 장비 추가 및 변경

5.2 보안정책 수립 및 변경에 관한 근거 제시

기존의 보안정책 의사결정은 보안 솔루션들이 검출해 내는 결과와 보안 담당자들의 경험에 비추어 수행되었다고 해도 과언이 아니었다. 하지만 이러한 정책결정은 예방차원이 아닌 기존의 사고사례에 대한 대응차원의 결정일 뿐이었다.

앞서 서론에서 언급하였던 보안 솔루션의 증가로 인해 보안 담당 관리자들의 보안활동은 운영·관리 측면으로 편향되고 있으며, 본연의 업무인 보안사고 예방활동의 비중이 점점 작아지게 되는 위험을 유발하고 있다. 하지만 금번 연구를 통해 산개 되어 있는 관리 포인트를 하나로 모아서 데이터를 분석함으로써 사내에서 발생하고 있는 보안사고 트렌드에 대응하기위한 의사결정에 아래와 같이 도움이 될 수 있다.

1) 임직원의 90%이상이 검출되는 시나리오 수정
2) 직군별 업무 시 필요한 보안 위배 정책 예외 처리
3) 주기별 Rule 시나리오 검출 빈도 분석을 통한 보안정책 수정
4) 보안제품은 다르지만 같은 기능을 하는 보안정책 통합으로 중복처리 제거

5.3 실제 적용 사례

구글은 높은 퇴사율이 문제가 되자 직종별, 성별, 직급별 데이터를 분석해 본 결과, 특히 최근 출산한 여성의 이직률이 직원 평균의 2배 이상임을 발견 했다.[10] 이는 데이터 분석에 근거하여 출산 여직원들의 개별 니즈를 적극 반영한 지원책을 제시한 사례로써 이번 연구를 통해 내부정보 유출자에 대한 데이터를 분석하여 결과가 아닌 과정의 원인을 파악함으로써 더 큰 사고 예방을 막을 수 있는 정책 수립에 도움이 될 수 있다는 것을 보여준다.

VI. 결론

빅데이터를 이용한 보안정책 개선을 통해 기존의 정보유출 사고 대응에 한정되어 있는 솔루션 활용에서 벗어나, 각 장비에서 추출된 데이터를 통합분석해서 보안 정책결정에 보다 신뢰성 있고 효과적으로 대응할 수 있는 사실적인 근거를 마련할 수 있었다.

1) 기존의 보안 솔루션을 통합함으로써, 다양한 제품들을 연계한 Rule 기반의 시나리오를 이용한 정보유출에 행위를 추출할 수 있으며, 패턴, 친밀도 분석을 통한 사고예방 및 추적대응이 가능했다.

2) 타 사업장의 규모나 근무인원의 데이터를 기반으로 예측되고 투자하였던 보안 시설의 투자에 있어서 통합로그를 활용한 기초 데이터 추출은 기존의 예측되는 데이터보다 좀 더 정확한 데이터를 제공할 수 있었다. 부서의 성격, 주 이동경로, 직군, 부서에 따른 통계 추출 및 누적된 데이터의 추이 그래프는 자칫 잘못 반영될 수 있는 보안 시설 투자에 있어서 좀 더 정확한 사실적인 데이터를 추출해 내었다.

3) 내부자의 위협에 효과적으로 대응하기 위해서 사용중인 SIEM의 시나리오 분석 방법을 발전시켜 기존 내부 위협자의 행동 패턴을 시나리오 기반으로 분석하고 이와 유사한 행동 패턴 징후 보유자를 선별하여 내부 위협자 선별의 오탐을 최소화는 방법론을 제시함으로써 기업 내 보안 수준 향상에 기여 할 수 있는 방안을 제안할 수 있었다.

이번 연구는 협소의 의미로 내부 위협자를 퇴직 예정자로 간주하고 특정 기업(제조업)의 SIEM 시스템에서 공통적으로 운영 중인 퇴직자 식별 시나리오를 발췌하여 진행한 내부정보 유출에 관한 케이스 스터디로써 정보유출 방지체계 구축을 위한 지속적인 개선·관리를 위해 개인 행동 패턴 연구에 대한 다양한 기업에서의 케이스 스터디가 필요하다.

참고문헌

[1] 산업기밀보호센터, http://service4.nis.go.kr/servlet/page?cmd=preservation&menu=AAA00
[2] 신혜원, "기업 내 정보유출방지를 위한 내부자 위험도 분석 방법론 연구," 한국컴퓨터종합학술대회 논

문집, Vol39, No 1. pp. 295-297, 2012년 6월
〔3〕 이재용 외 1명, "기업정보 유출 방지를 위한 통합 로그분석 시스템 설계 및 검증," 디지털콘텐트학회 논문지 제 9권, pp. 491-498, 2008년 9월
〔4〕 Tom White, "Hadoop: The Definitive Guide, Second Edition," OReilly Media , 2010
〔5〕 Ashish Thusoo외 8명, "Hive: a warehousing solution over a map-reduce framework," Proceedings of the VLDB Endowment VLDB Endowment Volume 2 Issue 2, pp. 1629-1629, August 2009
〔6〕 김종현 외 4명, "빅데이터를 활용한 사이버 보안 기술 동향," ETRI 사이버 보안 기술 특집, 제28권 제3호, pp. 19-29, 2013년 6월
〔7〕 이기혁 외 1명, "내부정보 유출 징후 분석을 통한 유출방지체계 구축에 관한 연구," 정보보호학회지, 제19권 제3호, pp. 70-79, 2009년 6월
〔8〕 S. Curry et al., "Big Data Fuels Intelligence-driven Security," RSA Security Brief, Jan. 2013.
〔9〕 Han, J. and Kamber, M., "Data Mining: Concepts and Techniques," Morgan Kaufmann, 2006.
〔10〕 박주영, "직원의 마음을 읽는 창, 빅데이터," SERI 경영노트, 제177호, 2013년 6월

〈저자소개〉

김 송 영 (Song Young Kim) 정회원
2005년 2월: 중앙대학교 산업정보학과 졸업
2005년 3월~현재: ㈜삼성전자
2013년 8월: 고려대학교 정보보호대학원 수료
〈관심분야〉 융합보안, 디지털포렌식, 빅데이터 등

김 요 셉 (Joseph Kim) 정회원
2000년 3월: 육군 3사관학교 전산정보과 졸업
2007년 1월: 국방대학교 전산정보 석사
2012년 3월~현재: 고려대학교 정보보호대학원 박사과정
〈관심분야〉 무기체계 보안인증, ISMS, 개인정보보호, 상호운용성 등

임 종 인 (Jong In Lim) 종신회원
1980년 2월: 고려대학교 수학과 졸업
1982년 2월: 고려대학교 수학과 이학석사
1986년 2월: 고려대학교 수학과 이학박사
現 고려대학교 정보보호대학원 원장, 고려대학교 사이버국방학과 교수, 개인정보보호위원회 위원, 대검찰청 디지털수사자문위원회 위원장, 금융보안연구원 보안전문기술위원회 위원장, 행정안전부 정책자문위원회 위원, 국방부 정보화책임과 자문위원, 한국저작권위원회위원 등
〈관심분야〉 사이버국방, 정보법학, 디지털포렌식, 개인정보보호, 융합기술보안 등

이 경 호 (Kyung Ho Lee) 종신회원
1989년 8월: 서강대학교 수학과 학사
1997년 8월: 서강대학교 정보통신대학원 석사
2009년 8월: 고려대학교 정보보호대학원 박사
1994년 2월~현재: 삼성그룹, nhn, 시큐베이스 등 근무
2011년 9월~현재: 고려대학교 정보보호대학원 조교수
〈관심분야〉 위험관리, 정보보호컨설팅, 정보보호 및 개인정보보호정책