컨소시엄 블록체인 네트워크 기반에서 가명처리를 활용한 안전한 기업 내부자 위협 행위 데이터 공유 시스템 연구

윤원석*, 장항배**
*중앙대학교 일반대학원 융합보안학과
**중앙대학교 산업보안학과
i3629i@cau.ac.kr, hbchang@cau.ac.kr

A study on the sharing system for insider threats behavior using pseudonymisation based on consortium blockchain network

WonSeok Yoon*, HangBae Chang**
*Dept. of Security Convergence, Graduate School, Chung-Ang University
**Dept. of Industrial Security, Chung-Ang University

요 약

본 논문에서는 지속적으로 늘어나고 있는 내부의 유출자를 탐지하기위해 컨소시엄 블록체인 기술을 활용하여 기업간 직원의 PC사용 행위 로그 데이터를 가명처리하여 블록에 기록하여 네트워크에 참여한 다른 기업들간의 안전한 공유를 통해 내부자 유출 데이터 및 시나리오의 확장하여 내부에서의 유출을 탐지할 수 있는 데이터 셋을 확보하는 연구를 제안한다. 현재 내부자 위협탐지의 한계점중 가장 큰 요소를 차지하는 부족한 실제 사례의 내부자 유출 데이터 셋의 문제점을 본 연구를 통해서 네트워크 참여 기업간의 내부자 유출 데이터를 확장하고 타기업의 유출 사례를 활용해 기업에서 발생할 수 있는 내부자 유출을 미연에 방지할 수 있다.

1. 서론

코로나 19 팬데믹 상황의 장기화로 인한 재택근무 환경 속 기업의 업무가 비교적 더뎌지는 상황 속에 기업정보자산의 내부자 유출사건이 지속적으로 늘어나고 있다. 실제로 2018년 이후 내부자에 의한 보안 사고의 수는 47%가 증가를 했고, 조직의 60%는 연간 20건 이상의 내부자 공격을 경험했다[1].

내부자의 유출이 지속적으로 일어나는 상황 속에서 여러 기업은 자신들의 중요자산을 지키기 위해 다양한 전술과 도구를 사용하는 방법을 보여준다. 대표적으로 기업 내부직원의 행위를 분석하고 모니터링하거나, 내부직원에 대해 주기적으로 교육을 진행하는 듯 다양한 방법으로 내부에서의 유출을 막기 위해 노력하고 있다[2]. 하지만 내부에서 발생하는 유출사고는 예상할 수 없는 범위에서 일어나고 완벽하게 막을 방법은 없기 때문에 다양한 방법들을 모색해야한다.

최근에는 내부자 유출을 탐지하기 위해 여러 데이터 셋을 세분화한 수준에서 기계학습 기반의 데이터 학습을 통해 모델을 개선하는 방향으로 연구가 활발히 진행되고 있다[3]. 그러나 현재 진행되고 있는 기계학습 기반의 내부자 유출 탐지연구에서는 대부분 실제 유출 데이터가 아닌 시나리오에 기반의 합성데이터 일뿐 실제 유출 데이터 셋과 차이가 존재한다. 또한 구체적인 방식이 없는 내부자 유출 행위에 대해서 새로운 방법으로 진행하는 유출 행위에 대한 유출 탐지가 불가능하다는 한계점을 꼬리표처럼 가져가게 된다[4].

따라서 본 연구는 기업 직원의 개인정보를 보호하며 기업 간 내부자 유출 탐지를 위해 컨소시움 블록체인 네트워크를 구축하여 네트워크에 참여한 기업간의 내부 유출자 행위와 일반 직원의 행위 데이터를 모두 가명처리를 통해 개인정보를 보호하고 해당 데이터를 네트워크에 참여한 기업끼리 공유하여 기업 내의 예상치 못한 방식의 유출 방법을 사전에 탐지하는 컨소시엄 블록체인 기반 내부자 행위 데이터 공유 시스템을 연구한다.

2. 선행연구

2.1 컨소시엄 블록체인

블록체인 기술은 서로를 신뢰할 수 없는 환경에서 신뢰를 보증할 수 있는 중개자 없이 참여자끼리 안전하게 거래를 할 수 있게 만든 기술이다[5]. 이러한 블록체인 기술에는 대표적으로 퍼블릭 블록체인과 컨소시엄 블록체인의 유형이 존재하는데 퍼블릭 블록체인과 컨소시엄 블록체인의 가장 큰 차이는 참여자의 허가에서 가장 큰 차이를 나타낼 수 있다. 퍼블릭 블록체인 같은 경우에는 누구든 참여할 수 있고 운영 주체가 될 수 있으므로 데이터 또한 참여자 모두가 볼 수 있다. 따라서 퍼블릭 블록체인은 민감한 정보를 다루는 기업의 관점에서는 외부로 나가기 곤란한 데이터를 모두에게 보여줘야 한다는 한계점이 존재하기 때문에 비교적으로 선호하지 않고 있다[6].

이러한 한계점을 보완하고자 컨소시엄 블록체인 기술이 등장하였고 허가된 기업이나 기관만이 참여가 가능한 신뢰 된 네트워크를 기반으로 조직은 효율적인 합의 알고리즘을 사용해 트랜잭션의 처리속도를 극대화하고 무결성을 유지하여 산업 내에서 복합적인 조직 환경과 공통의 정보를 필요로 하는 곳에 유용하게 사용되고 기업 간의 인사이트를 공유할 수 있으므로 서로의 기업조직에 이득이 된다[7].

2.2 내부자 유출 데이터 연구

기업 내부에서 유출되는 행위를 탐지하기 위해서 내부에서 의심되는 사용자의 행위를 디지털 환경에서 인식할 수 있게 만들기 위해 데이터로 만들기 위한 연구가 필요하다. 내부자 유출 탐지 데이터 셋에서 가장 많이 사용 중인 데이터는 CMU(Carnegie Mellon University) 대학에서 만든 CERT 데이터 셋이 많은 연구에 사용되고 있다[8]. CERT 데이터 셋은 가상의 기업에 다니는 사용자들의 PC 로그를 수집한 데이터로, 이 중 5개의 시나리오의 악의적인 행위자에 대한 데이터를 다루고 있다. 또한, 내부자 유출에 관한 데이터 연구에 대해서는 데이터 셋의 조작 행위에 대한 세분화가 필요하여 사용자의 유출 행위를 보다 구체적으로 탐지할 수 있게 진행한 연구를 진행하여 내부자 유출 탐지를 고도화하고 있다[9].

3. 연구방법론

3.1 공유 데이터 가명처리

기업의 행위 로그 데이터를 공유하면서 기업끼리 컨소시엄 블록체인을 통해 공유를 진행하게 된다면 직원의 개인정보가 포함된 정보를 공유받을 수 있고 직원의 PC 정보 또한 개인의 프라이버시가 담겨 있을 수 있으므로 개인정보보호에 관한 문제가 존재한다. 따라서 본 연구는 개인의 프라이버시 문제를 보호하기 위해 직원 PC의 행위 로그 데이터에서 개인정보가 들어가는 특징들을 모두 가명처리를 진행해 고유 특징을 비식별화하여 기업 간의 공유가 가능하게 설계를 진행한다. 표 1은 기업 내의 직원 행위 데이터를 컨소시엄 블록체인에 기록하기 위해 데이터를 가명 처리하여 설계한 내용이다.

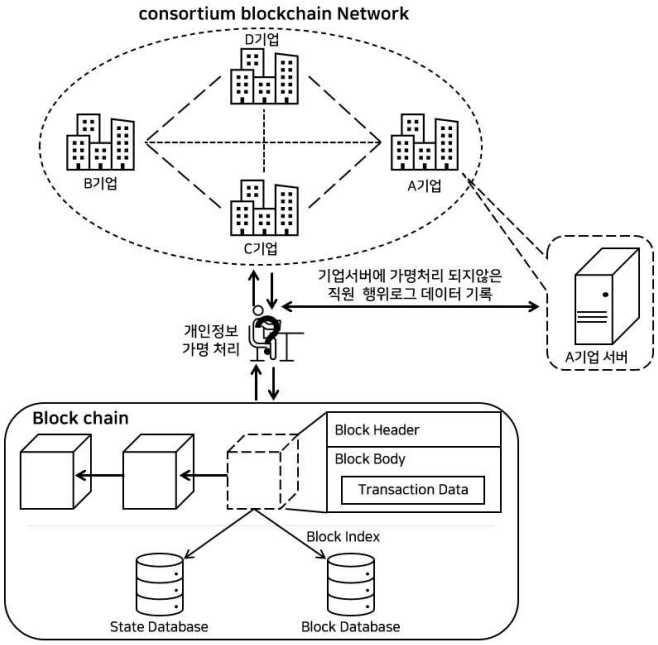
<표 1> 사용자 행위 가명처리 데이터 설계

직원 속성	실제 데이터	가명처리	기술방식
이름	임홍규	임OO	식별자 부분 삭제
사용자 ID	Hong1234	420C70543999710C352BD21099100B95A21132E4B2020D321ADD6666E775AECE	암호화
이메일	Hong11@gmail.com	H*****@gmail.com	식별자 삭제
직무	Technician	Technician	X
소속 팀	Web Software	W팀	감추기
PC ID	PC-18850	PC -1*****	식별자 부분 삭제
행위 이벤트	File Open	File Open	X
이벤트 발생일	02/21/2020 11:47:24	02/21/2020 11:47:24	X
파일이름	R:\MSW0221\대외비문서.doc	****.doc	식별자 부분 삭제(확장자만 표시)

3.2 컨소시엄 블록체인 설계

본 절에서는 가명 처리한 데이터를 기업 간의 안전한 공유를 위한 컨소시엄 블록체인을 설계한다. 블록체인 네트워크에 참여한 기업의 각 직원은 자신들의 PC 내에서 이루어지는 행위 로그에 대해서 가명처리 되어 블록에 기록하기 이전에 기업의 개인 서버에 기록하여 가명 처리되지 않은 데이터를 기록함으로 내부 유출자에 대한 추적이 가능하도록 설계한다. 개인 서버에 기록된 직원 행위 데이터는 가명처리를 통해 기업의 정보와 자신의 프라이버시를 침

해할 수 있는 개인의 정보를 숨겨 직원이 PC에서 실행한 행위에 대해서 블록에 기록하여 타 기업의 직원이 로그를 봤을 때 PC에서 일어난 행위로그와 가명화된 사용자만 구분할 수 있는 방법을 제시한다. 그림 1은 기업 간의 컨소시엄 블록체인 네트워크의 구성과 블록에 기록하기까지의 네트워크를 설계한 그림이다.



(그림 1) 컨소시엄 블록체인 설계

본 연구를 통해서 내부자 유출 연구의 가장 큰 문제점인 유출자에 의한 데이터의 부족과 새로운 유출 행위의 시나리오를 발견할 수 있는 방향성을 제시한다. 블록체인에서 통합적으로 저장하여 유출자의 행위에 대한 PC 로그 기록을 블록체인 네트워크에 속한 참여 기업이 공통으로 확인이 가능하고 여러 기업의 다양한 직원들의 행위 로그를 한곳에 모으기 때문에 유출 탐지를 분석을 위한 발판을 마련한다.

4. 결론 및 향후 연구

예상하지 못한 코로나 팬데믹 상황의 장기화로 인한 기업의 업무 마비와 재택근무 상황 속 기업 내부의 내부자 유출이 지속해서 증가하고 있다. 따라서 본 연구는 기업 간의 컨소시엄 블록체인 네트워크를 구축하여 각 기업 직원들이 PC에서 발생하는 행위 로그 데이터를 가명처리하여 블록에 기록해 다방면에서 발생하는 내부자 유출 행위 데이터를 안전하게 공유하는 방안을 제안한다. 직원들의 행위 데이터에서 발생할 수 있는 개인정보를 침해할 수 있는 데이터는 비식별화를 통해 직접적으로 해당 로그의 대상이 누구인지 구체적으로 판별하지 못하도록 가명처리하여 블록에 기록하고 가명처리를 진행하기 이전에 각 기업의 서버에 가명처리 되지 않은 직원 행위 데이터를 기록하여 기업 내의 내부자 추적성을 제공한다.

본 연구를 통해서 기업 간의 안전한 직원 행위 데이터를 공유함으로 실제 내부자 유출의 데이터 셋 부족 문제를 해결하고 다른 기업의 내부자 유출 행위를 통해 자신의 기업에 아직 발견하지 못한 내부자 유출 행위 시나리오를 제공받아 사전에 발생할 수 있는 내부자 유출을 기업간 막아주는 네트워크를 구성할 수 있다.

향후 연구로는 SIEM(Security information and event management)과 같은 기업에서 사용하는 이벤트 관리 솔루션을 활용하여 로그를 수집하여 특징에 맞게 파싱을 진행한 후 가명처리한 결과를 직접 구현함으로 기업 간의 협업을 통해 실제 가명처리를 수행한 내부자 유출 행위를 제공받아 다른 기업들의 직원 행위 데이터의 특성을 맞춰 내부의 유출자로 의심이 가는 직원을 도출해 검증을 해보는 연구와 컨소시엄 블록체인 네트워크 참여 기업의 내부자 탐지를 위한 적절한 규모 조정과 유사한 환경 속 새로운 내부자 유출 시나리오를 도출을 위한 기업 분야에 로그 데이터가 분석연구가 필요하다.

감사의글

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학 ICT 연구센터지원사업의 연구결과로 수행되었음” (IITP-2021-2018-0-01799)

참고문헌

[1] IBM. “Cost of Insider Threats: Global Report 2020”. Retrieved from IBM. 2020

[2] Cybersecurity Insiders. “Insider Threat Report”. Retrieved from Cybersecurity Insiders. 2020

[3] Al-Mhiqani, M. N., Ahmed, R., Abidin, Z. Z., & Isnin, S. N. “An integrated imbalanced learning and deep neural network model for insider threat detection”. International Journal of Advanced Computer Science and Applications, 12(1). 2021.

[4] Yuan, S., & Wu, X. “Deep learning for insider threat detection: Review, challenges and opportunities”. Computers & Security, 102221. 2021
[5] Yaga, D., Mell, P., Roby, N., & Scarfone, K. “Blockchain technology overview”. arXiv preprint arXiv:1906.11078. 2019.
[6] Syed, T. A., Alzahrani, A., Jan, S., Siddiqui, M. S., Nadeem, A., & Alghamdi, T. “A comparative analysis of blockchain architecture and its applications: Problems and recommendations”. 7, 176838-176869. 2019.
[7] Zheng, P., Xu, Q., Zheng, Z., Zhou, Z., Yan, Y., & Zhang, H. “Meepo: Sharded Consortium Blockchain”. In 2021 IEEE 37th International Conference on Data Engineering (ICDE). pp. 1847-1852. 2021.
[8] Glasser, J., Lindauer, "B.: Bridging the gap: a pragmatic approach to generating insider threat data." In: Security and Privacy Workshops (SPW), July 2013. pp.98 - 104
[9] Yoon, W. S., & Jang, H. B. “A Study on Dataset of Insider Leakage Behavior through Analysis of Document and File Manipulation Behavior”. Spring Conference. The Journal of Society for e-Business Studies. 2021. pp.120-122.