연합학습 기반 스마트팩토리 영역별 보안위협 대응방안

정인수[1], 김득훈[2], 곽진[3]
[1]아주대학교 사이버보안학과, 정보보호응용및보증연구실 석박통합과정
[2]아주대학교 소프트웨어융합연구소 박사후연구원
[3]아주대학교 사이버보안학과 교수
jis0727@ajou.ac.kr, kimdh1206@ajou.ac.kr, security@ajou.ac.kr

Countermeasures for Security Threats by Smart Factory Area based on Federated Learning

In-Su Jung[1], Deuk-Hun Kim[2], Jin Kwak[3]
[1]ISAA Lab., Dept. of Cyber Security, Ajou University
[2]Inst. for Computiong and Informatics Research, Ajou University
[3]Dept. of Cyber Security, Ajou University

요 약

스마트팩토리는 기존 제조산업에 ICT 기술이 융합된 지능형 공장이다. 이는 IT(Information Technology)영역과 OT(Operation Technology)으로 구분되고, 영역 간 연결을 통해 제조공정 자동화 및 지능화를 수행한다. IT영역은 외부 네트워크와 연결되어 스마트팩토리의 전사업무 관리를 수행하며, OT영역은 폐쇄망 네트워크로 구성되어 직접적인 제조과정을 수행한다. 이는 2개의 영역으로 구분되어 자동화 및 지능화된 제조공정 과정을 수행함에 따라 구조가 복잡해지고 있으며, 이로 인해 스마트팩토리 보안위협이 발생 가능한 공격 표면이 증가하고 있다. 이에 대응하기 위해서는 스마트팩토리 IT영역과 OT영역의 특징을 분석하고, 영역별 적합한 보안위협 대응체계를 수립해야 한다. 이에 따라, 본 논문에서는 다수의 장치에 대한 학습이 용이하고, 세부적으로 학습기법을 구분할 수 있는 연합학습을 활용하여 스마트팩토리 영역별 적합한 보안위협 대응방안을 제안한다.

1. 서론

최근 ICT 기술과 제조산업을 융합한 스마트팩토리가 발전함에 따라 ICT 기술을 기반으로 스마트팩토리의 IT영역과 OT영역이 연결되어 자동화 및 지능화를 통해 관리 비용 감소, 제조공정 최적화 등을 수행한다[1,2]. 그러나 영역 간 연결, 영역 내 장치 간 연결 등으로 인해 보안위협이 발생 가능한 공격 표면이 증가하고 있다. 이에 대응하기 위해서는 스마트팩토리 영역별 특징과 보안위협을 분석하는 과정이 필요하며, 이를 기반으로 영역별 보안위협에 대응하기 위한 대응체계 수립이 필요하다. 또한, 영역 내 장치들 간의 상호 연결성과 영역 간의 연결성 등 스마트팩토리의 구조 복합성을 고려하여 효율적인 보안위협 대응방안 연구가 필요하다. 따라서, 본 논문에서는 다수의 장치에 대한 학습이 용이하고, 세부적으로 학습기법을 구분할 수 있는 연합학습 활용하여 IT 영역에는 스마트팩토리 네트워크 데이터를 기반으로 지도 학습을 수행하고, OT 영역에서는 스마트팩토리 제조공정 데이터를 기반으로 비지도 학습을 수행하는 스마트팩토리 영역별 대응방안을 제시한다.

본 논문에서는 2장에서 스마트팩토리에 대한 정의, 스마트팩토리 영역별 보안위협 및 연합학습에 대하여 설명한다. 3장에서는 연합학습 기반 스마트팩토리 영역별 보안위협 대응방안을 제안하고, 4장에서 결론을 맺는다.

이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2021-0-01806, 스마트공장 보안 내재화 및 보안관리 기술 개발).

2. 관련 연구

2.1 스마트팩토리

스마트팩토리는 기존 제조기술에 클라우드, 빅데이터, IIoT(Industrial Internet of Things)와

같은 ICT 기술과의 융합을 통해 구축된다. 또한, 스마트팩토리는 전사업무 관리를 위한 IT 영역과 직접적인 제조공정 과정이 수행되는 OT 영역으로 구분되며, 이는 5개의 계층으로 세분화된다. 2개의 영역과 각 계층이 연결됨에 따라 제조공정 자동화 및 지능화된 인프라를 제공함으로써 생산성 향상, 에너지 절감, 안전한 생산환경 구현 등이 가능하다. 이는 국내외 스마트팩토리 산업제어시스템 표준(RAMI 4.0, ISA/IEC 62443, NIST 800-82, Purdue 모델 등)을 기반으로 구축된다[3].

❒ 스마트팩토리 IT 영역

스마트팩토리의 IT 영역은 주로 공장 내부의 생산을 관리하는 생산정보시스템과 전사 업무 관리 및 인터넷 서비스를 제공하기 위한 장치로 구성된다. 또한, Ethernet을 통해 대내·외 네트워크와 연결된 외부망으로 구축되어 있다. IT 영역은 스마트팩토리 대외로 서비스를 제공하고, 전체적인 관리 및 비즈니스 관련 활동에 필요한 기능을 수행한다[3].

❒ 스마트팩토리 OT 영역

스마트팩토리의 OT 영역은 주로 제어설비 장치로 구성된다. 또한, Fieldbus, 산업용 Eternet, Modbus, Profinet 등과 같은 산업용 네트워크를 기반으로 외부 네트워크와 구분된 폐쇄망으로 구축되어 있다. OT 영역 내 각 계층은 생산과 관련된 현장 장치로 구성된 0계층, 현장 장치들의 상태정보 수집 및 제어 명령 전달을 수행하는 1계층, 모니터링 및 공정 통제를 수행하는 2계층, 공장 또는 시설 단위로 전체 모니터링을 수행하고 최종 제품을 생산하기 위한 작업을 관리하는 3계층으로 구성되어 있다[3].

2.2 스마트팩토리 영역별 보안위협

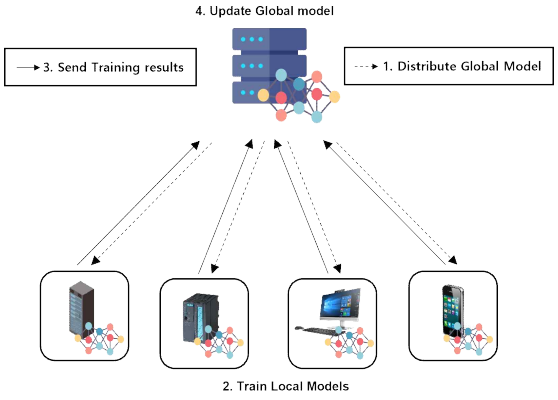
스마트팩토리 보안위협은 각 영역별 특성을 기반으로 도출되며, <표 1>을 통해 영역별 스마트팩토리의 보안위협 벡터와 보안위협을 확인할 수 있다. IT영역의 보안위협은 외부 네트워크와 연결됨에 따라 네트워크 기반 보안위협이 주로 발생하며, OT영역의 보안위협은 제조공정 과정에 직접적으로 영향을 미치는 제조공정 장치에 대한 보안위협이 주로 발생한다[3].

<표 1> 스마트팩토리 영역·계층별 보안위협

Area	Security Threat Vector	Security Threat
OT	Physical access, port, and support facility	Physical device damage, process data manipulation and leakage, etc.
OT	Industrial Control System	malfunction and service interruption, etc.
OT	Process Control Network	Seizing software permissions, malfunctioning and disrupting services, etc.
OT, IT	Factory Business Area	Ransomware infection, network failure, etc.
OT	Supply Chain	Ransomware infection, tampering with production information, etc.
OT, IT	Personnel and old facilities	Malfunction and interruption, manipulation and leakage of process data, etc.
OT, IT	External Internet	Business network failure, service interruption, information leakage, etc.

2.3 연합학습

연합학습이란 기업이나 기관 등 여러 위치에 분산된 클라이언트 데이터를 중앙서버인 글로벌 서버(Global Server)로 전달하지 않고, 글로벌 서버의 AI 모델인 글로벌 모델(Global Model)을 클라이언트로 보내 각각의 데이터로 모델을 훈련하는 분산형 머신러닝 기법이다. 이는 장치 데이터를 직접적으로 제공할 필요 없이 장치에서 학습된 모델 결과값을 기반으로 학습이 수행된다. 연합학습은 데이터가 탈중앙화된 환경에서 글로벌 모델을 학습하고 실행하는 구조로 데이터 프라이버시나 통신에 효율적이다. 연합학습의 동작 과정은 아래 5단계로 구성된다[4,5].



(그림 1) 연합학습 프로세스.

Step 1. 글로벌 서버는 보유한 글로벌 모델을 복사하여 클라이언트에 배포한다.

Step 2. 각 클라이언트는 수집하여 보유한 데이터로 배포된 로컬 모델(Local Model)을 학습한다.

Step 3. 학습을 마친 클라이언트는 학습한 로컬 모델 가중치 값을 글로벌 서버로 전송한다.

Step 4. 글로벌 서버는 전송된 결과들을 취합하여 글로벌 모델을 업데이트한다.

Step 5. Step 1.으로 돌아가 다시 반복한다.

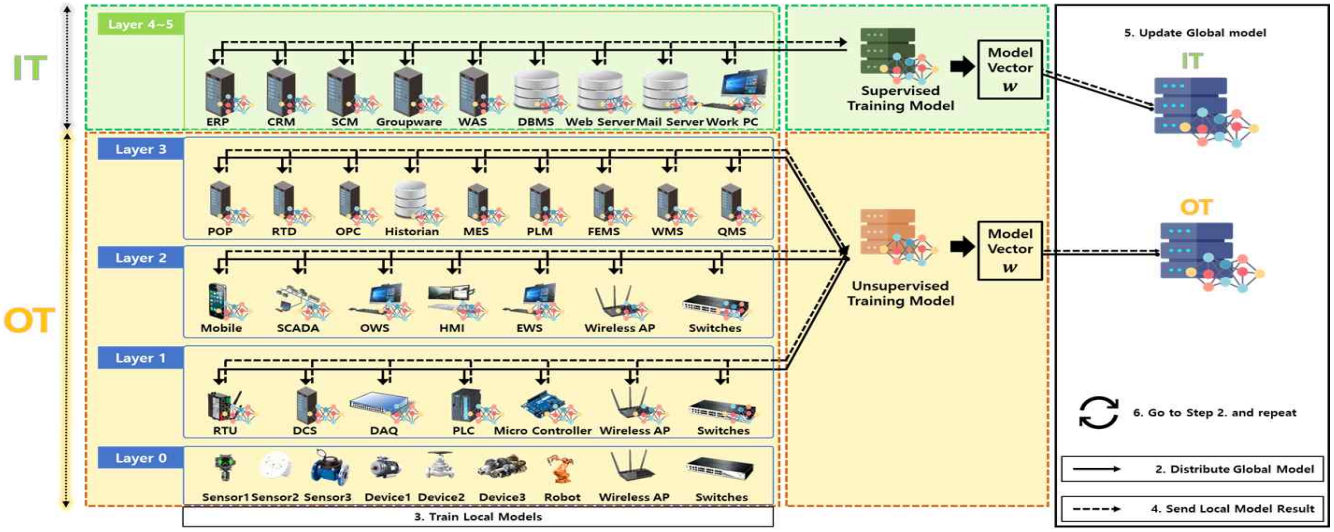
3. 제안사항

본 장에서는 연합학습을 기반으로 IT 영역과 OT 영역을 구분하여 학습을 수행하고, 이를 기반으로 한 보안위협 대응방안을 제안한다. 스마트팩토리 각 영역별 특징을 기반으로 영역별 보안위협에 대응하기 위해 활용되는 학습기법을 구분하여 학습을 수행하고, 반복적인 연합학습을 통해 최적화된 글로벌 모델은 스마트팩토리 보안위협 대응에 활용된다.

IT 영역 내 장치들은 외부 네트워크를 기반으로 연결되어 있다. 이에 따라, IT 영역에서 발생 가능한 보안위협은 네트워크를 통한 공격이 다수 존재하며, 각 장치에 대한 보안위협 분석을 위해서는 장치가 연결된 네트워크 정보들을 우선적으로 분석해야 한다. 또한, IT 영역 내 네트워크를 기반으로 한 공격은 기업 및 국가기관에서 보고서 및 데이터 형태로 다수 제공되고 있다. IT 영역에 대한 보안위협 관련 데이터 학습은 외부 네트워크와 연결되고, 이전에 발생한 공격 데이터 등을 활용할 수 있음에 따라 지도 학습을 활용하여 네트워크 데이터 기반 이상징후 탐지를 수행하는 것이 효율적이다.

OT 영역은 폐쇄망으로 구성되어 기타 환경들과 상이한 구조를 띠고 있으며, 0부터 3계층까지의 각 IIoT 장치들이 수평적 혹은 종속적인 관계를 형성하여 제조공정 과정을 수행하고 있다. 이에 따라, OT 영역의 장치들은 제조공정 과정에 직접적으로 사용되는 장치들로써, 학습을 위해 장치에서 발생하는 주파수, 진동수 등과 같은 장치 기반 데이터 분석이 필요하다. 또한, 이러한 제조공정 과정에는 IIoT 장치 정보, 제조하는 제품 정보 등과 같은 민감정보들을 포함하고 있다. 이에 따라, 스마트팩토리 내 제조공정 과정에 대한 데이터를 공개하기 어려운 실정이다. 따라서, OT 영역에 대한 보안위협 관련 데이터 학습은 비지도 학습 활용하여 장치에서 발생하는 시계열 데이터 기반 보안위협 대응을 수행하는 것이 효율적이다.

본 논문에서 제안하는 연합학습 프레임워크에서 사용하는 수식은 아래와 같으며, 사용되는 변수의 설명은 <표 2>과 같다. IT 영역 내 존재하는 장치들은 $S_I = \{1, ..., N_I\}$로 표현하고, OT 영역 내 존재하는 장치들을 $S_O = \{1, ..., N_O\}$로 표현한다. S_I은 네트워크 데이터를 기반으로 지도 학습을 수행하고, S_O는 장치 기반 시계열 데이터를 기반으로 비지도 학습을 수행한다. 영역별로 학습을 수행한 후 각 영역별 모델이 가지고 있는 $F_I(w)$, $F_O(w)$ 및 모델 벡터 $w \in \mathbb{R}^d$를 활용하여 각 영역별 연합학습을 수행한다. 아래 수식을 기반으로 연산을 수행하여 연합학습 글로벌 모델 최적화 과정이 수행된다.



(그림 2) 연합학습 기반 스마트팩토리 영역별 보안위협 대응 아키텍처

$$\min_{w\in\mathbb{R}^d} \frac{1}{N_I}\sum_{i\in S_I} F_I(w_I) \quad (1)$$

$$\min_{w\in\mathbb{R}^d} \frac{1}{N_O}\sum_{i\in S_O} F_O(w_O) \quad (2)$$

<표 2> 연합학습을 위한 변수 설명

Name	Description
t	Number of global rounds
N_I	Number of IIoT devices in IT area of smart factory
N_O	Number of IIoT devices in OT area of smart factory
S_I	Set of IIoT devices in IT area of smart factory
S_O	Set of IIoT devices in OT area of smart factory
w_I^0	Initial global model in IT area
w_O^0	Initial global model in OT area
w_I	Global model in IT area
w_O	Global model in OT area
w_{Ij}	Local model for device j in IT area
w_{Oj}	Local model for device j in OT area
$F_I(w_I)$	Global loss for IT area
$F_O(w_O)$	Global loss for OT area

Step 1. IT 영역 내 연합학습 모델의 글로벌 서버는 외부 네트워크를 통한 보안위협 벡터, 보안위협, 공격 대상 등을 기반으로 초기 데이터를 설정하고, 초기 글로벌 모델(w_I^0)을 구축한다. OT 영역 내 연합학습 모델의 글로벌 서버는 스마트팩토리 OT 영역 내 제조공정 장치의 주파수, 진동수 등과 같은 장치 기반 데이터의 정상 작동값을 초기 데이터로 설정하여 초기 글로벌 모델(w_O^0)을 구축한다. 초기 글로벌 모델은 학습에 필요한 데이터 정의, 이상 징후에 대한 지표를 제공한다.

Step 2. 글로벌 서버는 스마트팩토리의 보안위협 지표로 구축된 초기 글로벌 모델을 각 계층의 장치(Device)에 배포한다.

Step 3. IT 영역 내 장치는 네트워크 정보를 기반으로 지도 학습을 수행하며, OT 영역 내 장치는 장치 기반 시계열 데이터를 기반으로 비지도 학습을 수행한다.

Step 4. 각 영역별 장치에서 초기 글로벌 모델 기반 학습을 마친 후 이는 로컬 모델이 되며, 학습한 로컬 모델 벡터 기반 $F_I(w_I)$, $F_O(w_O)$ 값을 각 영역별 글로벌 서버로 전송한다.

Step 5. 글로벌 서버는 각 장치로부터 수신한 로컬 모델 벡터 결과값 $F_I(w_I)$, $F_O(w_O)$를 취합하여 글로벌 모델을 업데이트한다.

Step 6. 글로벌 모델 업데이트가 완료되면 Step 2.로 돌아가 다시 반복한다.

위의 과정을 반복함으로써, 각 영역별 글로벌 모델은 스마트팩토리 영역별 특성에 적합한 보안위협 대응방안을 도출한다. 또한, 각 영역별로 연합학습 최적화 연산을 반복함으로써 영역별 추가적으로 발생 가능한 보안위협에 대하여 지속적인 학습이 가능하다.

4. 결론

본 논문에서는 연합학습 기반 스마트팩토리 영역별 보안위협 대응방안을 제안하였다. 스마트팩토리 각 영역별 데이터 특징을 고려한 대응방안을 통해 정확한 보안위협 분석 및 대응에 기여할 수 있을 것이다.

참고문헌

[1] Shi, Z., Xie, Y., Xue, W., Chen, Y., Fu, L., & Xu, X., "Smart factory in Industry 4.0. Systems Research and Behavioral Science," 37, 4, pp.607-617, 2020.

[2] 이현정, 유상근, 김용운, "스마트공장 기술 및 표준화 동향", 국가기술표준원, 제32권, 제3호, pp.78-88, Jun. 2017.

[3] 한국인터넷진흥원, "스마트공장 보안 모델 해설서", Dec. 2022.

[4] Elnagar, Samaa, and Manoj A. Thomas. "Federated deep learning: A conceptual model and applied framework for industry 4.0," Americas Conference on Information Systems (AMCIS) 2020, Jul. 2020. pp. 24-34.

[5] McMahan, Brendan, et al. "Communication-efficient learning of deep networks from decentralized data," AProceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS) 2017, Florida, USA, 2017.