논문 2023-3-12 http://dx.doi.org/10.29056/jsav.2023.09.12

운영 기술 환경에서 수동적 스캐닝 기반 장치 식별의 비교 분석

박민수*, 안석현**, 박세연**, 조성제**† , 김홍근***

A Study of Passive Scanning-based Device Identification in Operational Technology Environments

Minsu Park*, Seokhyun Ahn**, Seyeon Park**, Seong-je Cho**† , HongGeun Kim***

요 약

보안 위험을 완화하기 위해서는 조직의 자산을 식별하고 보안 취약점을 제거해야 한다. 자산 식별을 위해 기존 IT 환경에서는 능동적 스캐닝과 수동적 스캐닝 기법을 적용해왔다. 본 논문에서는 먼저 기존 네트워크 스캐닝 방식을 OT 시스템에 적용할 때의 문제점을 살펴본다. 수동적 스캐닝 기법을 적용하여 OT 장치를 식별할 때 필요한 네트워크 트래픽의 특징 정보를 선택하기 위해, OT 장치와 유사한 특성을 가진 IoT 장치 식별 연구에서 사용된 네트워크 트래픽 특징(TTL, TCP Window Size, IAT 등)을 분석하였다. 또한, 네트워크 트래픽 특징 정보의 유형을 식별하고, 이를 기반으로 효과적인 OT 장치 식별을 위한 앞으로 해결해야 할 연구 이슈들을 제시한다.

Abstract

To mitigate security risks, organizational assets must be identified and security vulnerabilities eliminated. active scanning and passive scanning techniques have been applied in existing IT environments to identify assets.. In this paper, we first examine the problems of applying the existing network scanning methods to OT systems. Then, to select network traffic features required when identifying OT devices by applying a passive scanning, we analyze network traffic features (TTL, TCP Window Size, IAT, etc.) used in prior studies on identifying IoT devices which have similar properties to OT devices. In addition, we recognize the types of network traffic features, and present research issues to be addressed for effective identification of OT devices.

한글키워드 : 운영 기술, IoT, 장치 식별, 네트워크 스캐닝, 네트워크 트래픽 특징 정보

keywords : Operational Technology, Internet of Things, Device Identification, Network Scanning, Network Traffic Features

* 단국대학교 인공지능융합학과
** 단국대학교 소프트웨어학과
*** 동국대학교 국제정보대학원 정보보호학과
† 교신저자: 조성제(email: sjcho@dku.edu)
접수일자: 2023.09.01. 심사완료: 2023.09.09.
게재확정: 2023.09.20.

1. 서 론

디지털 전환으로 스마트 팩토리, 스마트 홈, 스마트 빌딩, 스마트 그리드 등 다양한 스마트+'X'기술이 등장하고 있다[1]. 기존에 인터넷 및 클라우

드와 연결되지 않은 OT(Operational Technology) 시스템의 구성요소들이 연결되면서 사이버 보안 위험이 증가하고 있다[2][3]. OT 시스템을 대상으로 한 사이버 공격은 시스템에 인가받지 않은 사용자가 시작 및 중지 명령으로 지속적인 동작을 수행하는 공정을 멈추거나, 물리적 현장의 인력을 위험에 빠뜨릴 수 있다[4]. 즉, OT 시스템이 사이버 공격을 받아 동작하지 않을 경우, 경제적인 손실뿐만 아니라 사람의 안전까지 위협받게 된다. 따라서 OT 시스템에 대한 보안 위험을 제거하거나 완화하는 것이 매우 중요하다. OT 시스템에 대한 보안 위험을 완화하기 위해서는 네트워크에 연결된 자산(장치 등의 시스템 구성요소들)을 식별하고 이에 대한 보안 취약점 관리가 이루어져야 한다.

즉, 사이버 보안 위협을 제거 또는 완화하기 위해서는, OT 시스템을 구성하는 장치를 식별하는 연구가 우선 선행되어야 한다. 이를 위해, 전통적인 IT 환경에서 각 장치를 식별하는 방법, 즉 네트워크 스캐닝(network scanning) 기법들[7][8]을 고려할 수 있다. 네트워크 스캐닝 기법은 네트워크에서 기본적으로 전송되는 패킷들 혹은 특정 도구에 의해 생성된 패킷들을 관찰하여, 네트워크에 연결된 각 장치의 유형, 하드웨어 정보, 운영체제 등을 식별하는 기술이다.

그러나, 전통적인 네트워크 스캐닝 방식을 OT 시스템에 그대로 적용하기는 어렵다. IT 시스템에서는 기밀성(confidentiality) 유지에 대한 요구가 높고, 각 구성요소가 보유한 컴퓨팅 자원이 풍부한 편이다. 반면에, OT 시스템은 물리적 세계와 상호작용하면서 실시간성을 만족해야 하므로 가용성(availability)에 대한 요구가 높고, 각 구성요소가 보유한 컴퓨팅 자원이 부족한 편이다[1]. OT 시스템의 높은 가용성 요구사항으로 인해, 네트워크 패킷을 생성하여 장치를 식별하는 방법을 적용하기 어렵게 만든다[5][6].

또한, OT 시스템은 다양한 제조업체가 개발한 구성요소와 특정 상황에서만 동작하는 장치들을 포함하고 있어, 전통적인 네트워크 스캐닝을 적용하기 어렵다. 예로, 건물 자동화 시스템(Building Automation System)과 같은 OT 시스템은 난방. 환기, 공기 조절, 화재 경보, 조명 제어 등과 관련된 다수의 장치로 구성되는데, 화재 경보 기기처럼 특정 상황에서만 동작하는 장치들은 네트워크 패킷을 관찰하기 어렵다는 문제가 있다. 이에, OT 환경에 적용할 수 있는 장치를 식별하는 연구가 필요하다.

본 논문에서는 먼저 네트워크 스캐닝과 관련된 기존 방식들의 장단점을 조사하여, OT 장치 식별 기법의 이슈를 파악한다. 이를 통해, 패킷을 생성하여 장치를 식별하는 능동적 스캐닝(active scanning)은 OT 장치 식별에 한계점이 있음을 제시한다. 수동적인 스캐닝(passive scanning)의 경우, 네트워크에서 기본적으로 전송되는 패킷만으로 장치를 식별하는 방식으로 OT 시스템의 가용성에 영향을 주지 않는다. 본 논문에서는, OT 시스템과 유사한 특징을 가진 IoT 시스템을 대상으로 한 수동적 스캐닝 연구들과, 그리고 OT 장치를 식별하는 네트워크 스캐닝 연구들을 비교 분석한다. 이를 통해, OT 장치를 효과적으로 식별하기 위한 연구 방향을 제시한다.

논문의 구성은 다음과 같다. 2장에서는 전통적인 IT 환경에서의 네트워크 스캐닝 방식을 정리하고, 3장에서는 OT 환경에 네트워크 스캐닝 기법을 적용할 경우, 발생하는 문제점을 살펴본다. 4장에서는 수동적 스캐닝에 의한 IoT 자산 식별 연구를 조사하여 사용된 네트워크 트래픽 특징 정보를 그룹화하고, 이를 비탕으로 OT 자산 식별 연구에서 사용된 네트워크 특징 정보들을 조사한다. 5장에서는 OT 자산 식별을 위해 해결해야 하는 연구 과제들을 제시하고, 6장에서 결론을 맺는다.

2. IT 환경에서의 네트워크 스캐닝

IT 환경에서의 네트워크 스캐닝 기법은 능동적 스캐닝과 수동적 스캐닝으로 구분된다.

능동적 스캐닝은 네트워크에 직접적인 쿼리(Query)를 발생시켜 연결된 장치를 파악하는 기법이다[9,10]. 능동적 스캐닝은 수신한 전송 제어 프로토콜 패킷 헤더의 필드를 특징 정보로 활용하여 알려진 운영체제의 특징 데이터베이스와 비교하여 동작 중인 장치의 운영체제 정보를 예측한다. 능동적 스캐닝 기법을 사용하는 대표적인 도구인 Nmap[11]은 nmap-os-db에 있는 운영체제에 대한 정보를 토대로 네트워크에서 활동 중인 장치에 쿼리를 보내고 응답 패킷의 헤더 정보들을 통해 장치의 운영체제를 추측한다. 즉, Nmap은 최대 16개의 TCP, UDP, ICMP 패킷을 대상 장치의 알려진 포트로 전송하여 돌아오는 응답 패킷의 헤더 정보를 통해 스캔 대상의 운영체제를 식별할 수 있다.

수동적 스캐닝은 네트워크에 직접적인 쿼리를 발생시키지 않고 네트워크에서 발생하는 트래픽을 관찰하여 얻은 정보를 토대로 자산을 파악하는 방법이다[9]. 대표적인 수동적 스캐닝 도구인 p0f[12]는 장치가 생성하는 TCP 트래픽(SYN, SYN+ACK 및 RST, RST+ACK 등)을 기반으로 특징 데이터베이스를 참조하여 장치의 운영체제를 예측한다. p0f는 장치가 주고받는 TCP SYN, TCP SYN+ACK, HTTP Request, HTTP Response 등의 패킷을 확인한다. 관찰된 패킷의 헤더에서 최대 세그먼트 크기(Maximum Segment Size), 윈도우 크기(Window Size), TCP 헤더 옵션 등의 정보를 서명 레이아웃으로 만들어 장치별 운영체제를 식별한다.

3. OT 환경에서의 네트워크 스캐닝 한계점 분석

이번 장에서는 2장의 능동적 스캐닝과 수동적 스캐닝을 OT 환경에 적용했을 때, 발생할 수 있는 문제들을 식별한다.

능동적 스캐닝은 식별 대상에 직접적인 쿼리를 발생시키기 때문에 식별 대상 장치에 부하를 준다. OT 네트워크에서 수행하는 침투 테스팅(Penetration Testing) 혹은 네트워크 스캐닝 작업은 처리 능력이 부족한 OT 장치에 상당한 부담을 줄 수 있다. 2005년 SANDIA[24]는 능동적 스캐닝이 산업 제어 시스템(Industrial Control System) 환경에 악영향을 줄 수 있다고 보고하였다. 예로, 한 가스 회사에서 SCADA(Supervisory Control and Data Acquisition) 시스템에 대한 침투 테스트로 인해 4시간 동안 파이프라인이 잠겨 서비스 손실이 발생하였다. M. Caselli[25] 등은 OT 장치의 낮은 연산 능력이 많은 양의 네트워크 트래픽을 처리하기 어려워 장치를 식별하기에 문제가 될 수 있음을 지적하였다. 반면, 수동적 스캐닝은 능동적 스캐닝보다 식별 정확도가 낮을 수 있다. 또한, 특정 상황에만 동작하는 장치 또는 네트워크 트래픽이 드문드문 생성되는 장치를 탐지하지 못할 수 있다. 예로, 특정 상황에서만 울리는 경보 센서는 탐지하기 어렵다.

또한, OT 장치를 대상으로 능동적 스캐닝 도구가 영향을 미치는지 평가한 연구가 있다. [13]은 38개 장치를 포함하는 3개의 테스트베드를 대상으로 능동적 스캐닝을 수행하였는데, 38개 장치 중 1개의 장치에서 오작동을 발생하였다. [14]는 능동적 스캐닝 도구(Nmap, Zmap)가 산업 제어 시스템 테스트베드에 미치는 영향을 조사하였다. Nmap은 네트워크 지연 혹은 통신 장애를 발생시켰으며, Zmap은 통신 오류로 인해 동작 중이던 모터의 작동을 정지시켰다. [15]는 5가지 능동적

스캐닝 도구가 7개의 PLC 장치에 영향을 미치는지 분석하였는데, 스캐닝 도중 대부분의 장치는 몇 밀리초에서 수백 밀리초 가량의 지연이 발생하였다.

표 1은 OT 장치를 대상으로 능동적 스캐닝과 수동적 스캐닝을 활용해 자산 식별을 수행한 연구 목록으로 자산 식별 수행의 한계점을 분석한 것이다.

[16,17,18]는 OT 장치를 식별하기 위해 산업 환경에서 사용되고 있는 프로토콜의 정보를 토대로 장치 식별을 수행하였다. Modbus와 BACnet 프로토콜을 대상으로 장치 식별을 수행하였으며, 제조사 별 레지스터 주소, 프로토콜 내 속성값 등이 다르므로 장치의 제조사 및 모델을 식별함에 있어 유용한 정보로 활용되었다. 이외에 KNX, DNP3, OPC UA 등 다른 프로토콜을 활용한 연구는 확인되지 않았다. 산업 제어 프로토콜에서 제공하는 정보는 장치 식별에 유용하지만, 장치가 해당 프로토콜을 사용하지 않으면 식별할 수 없다는 문제가 있다.

[19]는 포트 미러링(Port Mirroring) 기술이 지원되지 않는 환경에서 수동적 스캐닝 도구의 한계점을 파악하기 위한 연구를 수행하였다. 수동적 스캐닝 도구인 PVS(Passive Vulnerability Scanner)를 사용해 1개의 HMI 서버와 2개의 PLC를 대상으로 장치 식별을 수행하였다. 결과로 브로드캐스트 패킷을 생성한 HMI 서버는 식별할 수 있었지만, 해당 패킷을 생성하지 않은 PLC 장치는 식별되지 않았다.

[20]은 CLRT(Cross-Layer Response Times)와 FF-ANN (Feed Foward Artificial Neural Network) 알고리즘을 활용하여, 자산 식별을 수행하였다. [21]은 특징 정보들(TTL, IP ID, MAC ID)을 활용하고 장치가 동작하는 계층을 식별하는 알고리즘으로 자산 식별을 수행하였다. [20,21]에서는 장치 식별을 위해 각각 약 5개월과 11일 동안 네트워크 트래픽을 수집하였다. OT 장치의 긴 세션 지속 시간과 특수한 상황에서 트래픽이 생성되는 장치로 인해 트래픽 수집 기간이 길어질 수 있다. 이러한 OT 장치의 특성으로 수집 기간 내에 장치 식별에 유용한 정보가 확인되지 않거나, 특수한 장비를 필요로 할 수 있다.

[22,23]은 오픈소스 자산 식별 도구를 사용해 OT 장치를 식별할 수 있는지 조사하였다. 능동적 스캐닝 도구인 Nmap과 수동적 스캐닝 도구인 Grassmarlin을 사용하였으며, Nmap은 지멘스

표 1. OT 장치에 대한 능동적 및 수동적 스캐닝 기법 비교
Table 1. Active and Passive scanning techniques for OT devices

Technology	Ref No.	Limitation	year
Active Scanning	[16]	산업 제어 프로토콜(Modbus, BACnet 등)을 사용하는 장치만 식별 가능	2016
	[17]		
	[18]		2021
Passive Scanning	[19]	포트 미러링(Port Mirroring) 기술이 지원되지 않는 환경에서 브로드캐스트 패킷을 생성하지 않는 장치 식별 불가	2015
	[20]	긴 네트워크 트래픽 수집 기간	2016
	[21]		2019
Hybrid	[22]	자산 식별 도구(Nmap, Grassmarlin)의 장치 식별 능력 부족	2018
	[23]		2022

표 2. 수동적 스캐닝에 의한 IoT 장치 식별에 기계학습 모델을 적용한 연구 비교
Table 2. Comparison of studies applying machine learning models to IoT device identification by passive scanning

Year	Ref No.	Models	Packet Features	Datasets	Evaluation
2017	[27]	RF	ARL, LLC, IP, etc.	Public [27]	27개 장치 중 17개에 대해 약 95% 이상, 10개에 대해 약 50%의 정확도
2018	[28]	NB, RF	DNS Queries, NTP Queries, Flow volume, etc.	Public [28]	28개 장치와 non-IoT 장치에서 99.88%의 정확도
2018	[29]	GB, KNN, DT	IP, ICMP, ICMPv6, etc.	Private	14개 장치에 대해 평균 99% 이상의 정확도
2018	[30]	RF, DT, SVM, etc.	Packet Length, IAT(Inter Arrival Time), etc.	Private	4개 장치에 대해 99.9%의 정확도
2018	[31]	RF, KNN, NB, etc.	DNS Queries, Packet Length, Flow Duration etc.	Private	16개의 장치에 대해 평균 70.55%의 정확도
2018	[32]	LSTM-CNN	MAC Address, Packet Length, Packet Count, etc.	Public [28]	15개 장치에 대해 평균 74.8%, 최고 99.7%의 정확도
2019	[33]	KNN, SVM, RF, ADABoost, etc.	ARP, LLC, IP, etc.	Public [28]	21개 장치에 대해 최대 95.5% 정확도와 F1-Score
2019	[34]	LDA, KNN, CART, etc.	TTL, IP Header Length, IAT, etc.	Public [27]	27개 장치에 대해 평균 90.3%의 정확도와 91% F1-score
2019	[35]	KNN	Periodic Flows, Period Accuracy, Period Duration etc.	Private	33개 장치에 대해 98.2% 정확도
2019	[36]	J48, PART, DT, etc.	TTL, TCP Window Size, Packet Length, etc.	Public [27]	23개 장치에 대해 최소 95%의 정확도
2020	[37]	J48, PART	IP Header Length, TTL, TCP Window Size, etc.	Public [27,28]	두 가지 데이터셋에서 각각 최대 99.77%, 99.93% 정확도
2020	[38]	HMM, LSTM	ARP, LLC, IP, ICMP, etc.	Public [27,28]	두 가지 데이터셋에서 각각 최대 92%, 99% 정확도
2020	[39]	RF, Bayes Net	IAT	Private	39개의 ZigBee, Z-Wave 프로토콜을 사용하는 장치에 대해 평균 91% 이상의 정밀도와 재현율
2020	[40]	KNN, RF, DT	Packet Length, IAT, Packet Count, etc.	Public[28], Private	5개 장치와 공개 데이터셋에서 평균 90% 이상의 정확도
2021	[41]	SVM, DT, RF, GB	ARP, EAPoL, IGMPv2, etc.	Public[27], Private	18개 장치에 대해 평균 98.8% 정확도
2021	[42]	CNN+BiLSTM	source IP, source Port, Destination IP, etc.	Public [27,45]	두 가지 데이터셋에서 각각 99.91%, 99.68% 정확도
2022	[43]	RF, DT, RNN, etc.	Packet Length, IAT, Flow Duration etc.	Private	7개의 장치에 대해 최대 99.97% 정확도
2022	[44]	RF, DT, GB	TTL, TCP Window Size Packet Length, etc.	Public [27,28]	두 가지 데이터셋에서 각각 93.3%, 94.3% 정확도와 96.1%, 93.7% F1-Score

(Siemens)사의 PLC를 대상으로 하드웨어 정보를 파악할 수 있었지만, 소프트웨어 정보는 알 수 없었다. Grassmarlin은 IP, MAC, 제조업체 정보는 확인할 수 있었지만, 이외 하드웨어와 소프트웨어 정보 모두 알 수 없는 한계를 보였다. 즉, 대표적으로 알려진 자산 식별 도구가 OT 장치를 완벽히 식별하는 것은 부족함을 알 수 있다.

IT 환경에서는 능동적 스캐닝에 의한 일부 가용성 훼손이 있더라도 장치의 자원이 이를 견딜 만큼의 용량을 갖추므로, 능동적 스캐닝의 적용을 제한하지 않는다[1]. 반면, OT 환경은 24시간 연속적으로 운영되는 환경의 특성으로 인해 OT 장치의 가용성 요구사항이 높으며[26], 능동적 스캐닝으로 인해 가용성에 영향을 받으면 심각한 문제로 이어질 수 있어, 가용성 침해의 발생 가능성을 정확히 분석한 후 제한적으로 사용해야 한다. 이상의 문제들을 고려하여 다음 장에서는 수동적 스캐닝 기법을 중심으로 관련 연구를 비교 분석한다.

4. 수동적 스캐닝 적용을 위한 네트워크 특징 정보 식별

네트워크 트래픽을 수집하고 기계학습 모델을 개발하여 장치를 식별하려는 연구가 최근 많이 이루어졌다. 기계학습 모델을 개발하기 위해서는 수집한 네트워크 트래픽으로부터 특징 정보를 선정해야 한다. 본 장에서는 OT 장치를 네트워크 스캐닝을 통해 식별하기 위해 수동적 스캐닝에 필요한 네트워크 트래픽의 특징 정보를 식별한 연구들을 조사분석 하였다. 먼저, OT 장치와 유사한 특성을 가진 IoT 장치에 대한 네트워크 트래픽의 특징 정보를 다룬 연구를 조사하고, 3가지 유형으로 특징 정보들을 분류하였다. 이를 바탕으로 OT 장치에 대한 네트워크 트래픽의 특징 정보에 대한 연구들을 조사하였다.

4.1 IoT 장치 식별을 위한 네트워크 특징 정보 조사

IoT 장치를 대상으로 수동적 스캐닝에 의한 자산 식별에 기계학습 모델을 적용하는 18건의 연구를 연도순으로 표 2로 정리하였다. 표 2의 연구들에서 장치 식별을 위해 다양한 머신러닝 모델과 딥러닝 모델을 사용하였다. 사용한 데이터셋은 공개 데이터셋과 개별 수집 데이터셋으로 나뉜다. 공개 데이터셋은 IoT Sentinel, UNSW, YourThings[27][28][45]을 사용하였다. 연구 대상 장치 수는 최소 4개이며, 두 가지 데이터셋[27,28]을 사용한 연구에서 최대 55개로 확인되었다. 대부분의 연구에서 90% 이상의 높은 식별 정확도를 보이며, 활용한 네트워크 특징 정보들이 장치를 식별하는데 유용하다는 것을 알 수 있다.

다음으로, 표 2의 IoT 장치 식별 연구에서 사용한 특징 정보들은 프로토콜 기반, 패킷 속성 기반, 통신 행위 기반 3가지로 그룹화하여 표 3, 표 4, 표 5로 정리하였다.

표 3. 프로토콜 기반 특징 정보
Table 3. Protocol based feature information

Ref No.	Layer	Features
[27,29,30, 32,33,38]	Datalink	ARP/LLC
	Network	IP/ICMP/ICMPv6 /IGMP/EAPoL
	Transport	TCP/UDP
	Application	HTTP/HTTPS/D HCP/BOOTP/SS DP/DNS/NTP

표 3은 네트워크 계층별 프로토콜 기반 특징 정보들이다. 프로토콜 기반 특징 정보들은 OSI 7계층 모델을 기준으로 데이터링크 계층, 네트워크 계층, 전송 계층, 응용 계층으로 나뉜다. 이 정보

들은 프로토콜들의 사용 유무를 통해 장치를 식별하는 정보로 활용된다. 프로토콜 기반 특징 정보들은 다양한 IoT 장치의 네트워크 통신 특성을 이해하는데 활용될 수 있다.

표 4. 패킷 속성 기반 특징 정보
Table 4. Feature Information Based on Packet Properties

Ref No.	Features
[29,34,35,36,37,38,41,44]	Packet Header related features (TTL, TCP Windows size, etc.)
[27,29,32,34,44]	Packet Payload related features (Payload Entropy, Raw Data, etc.)
[27,29,31,32,34,36,37,38,39,40,41,43,44]	Packet Length related features (Packet Length, TCP Segment Length, etc.)

표 4는 통신 중 발생하는 패킷으로부터 패킷 헤더(header), 패킷 페이로드(Payload), 패킷 길이(Length) 등의 패킷 속성 기반의 특징 정보의 목록이다.

패킷 헤더 특징 정보는 패킷의 헤더에서 추출할 수 있는 정보로서, 대표적으로 TTL과 TCP 윈도우 크기(Window Size)가 있다. TTL은 패킷이 생존할 수 있는 시간을 의미하며, 하나의 패킷이 하나의 라우터를 지날 때마다 값이 1 감소한다. TTL 값은 소프트웨어를 개발자에 따라 정해진 값이 다르다[46]. 이는 IT 시스템에서 PC 운영체제 식별에도 사용되어왔다[47]. 또한, TCP 윈도우 크기는 TCP 기반에 통신을 수행할 때, 송신자에게 응답을 보내기 전, 자신이 처리할 수 있는 데이터의 양을 알려주기 위해 사용된다. 이 특징 정보는 장치가 수행하는 기능마다 차이가 있다. 예를 들어, 스마트 전구와 같은 장치는 윈도우 크기가 작으며, 스마트 카메라와 같은 장치는 더 가변적이고 큰 윈도우 크기를 가질 수 있다[29].

페이로드 관련 특징 정보는 패킷의 페이로드로부터 추출한 정보로서, 대표적으로 페이로드 엔트로피(Payload Entropy)가 있다. 페이로드 엔트로피는 패킷 내부 정보인 페이로드 메시지의 유형 및 크기와 관련이 있다[29]. 페이로드를 평문으로 전송할 경우 엔트로피는 가장 낮으며, 압축 또는 암호화된 데이터를 전송할 경우 엔트로피는 증가한다[48]. 이는 패킷을 암호화하는 장치를 구분하는데 특징 정보로 사용될 수 있다.

패킷 길이 관련 특징 정보는 특정 통신 프로토콜에서 보내는 패킷의 헤더 길이 혹은 페이로드의 길이를 말한다. 관련 특징 정보로 IP 패킷 헤더 길이, UDP 데이터 길이(Data Length), TCP 세그먼트 길이(Segment Length) 등이 있다. 패킷 길이는 특정 프로토콜을 사용하는 패킷 길이의 최소, 최대, 평균 등의 통계적인 값을 사용한다. 이 정보들은 사용하는 프로토콜, 제어 신호의 종류, 제조사별 장치의 기능 등에 따라 달라질 수 있다. [37,44]에서는 패킷 길이 관련 특징 정보의 중요도가 높게 평가하여, 장치의 고유한 식별 정보로 활용될 수 있음을 보였다.

표 5. 통신 행위 기반 특징 정보
Table 5. Communication Behavior Based Feature Information

Ref No.	Features
[28,30,31,32,33,34,35,39,40,42,43]	IAT, Flow Volume, Flow Duration, Packet Count, etc.

표 5는 장치별 네트워크 행위로부터 추출할 수 있는 특징 정보이다. 관련된 특징 정보는 네트워크 양방향 흐름으로부터 추출할 수 있다. 대표적

인 특징 정보로 패킷 간 도착 시간을 나타내는 IAT(Inter Arrival Time)가 있다. IAT는 다음 패킷이 도착하기까지 걸리는 지연 시간을 측정한 값을 말한다. IAT는 일반적으로 회로 설계의 작은 차이와 회로 모듈의 시간적 결함으로 인해서 발생한다[39]. 이러한 차이는 장치 유형을 분류할 수 있는 특징 정보로 활용되고 있으며, 여러 연구에서 장치의 고유한 특징 정보로 활용될 수 있음을 볼 수 있다[49,50]. 추가적인 특징 정보로 흐름 크기(Flow Volume), 흐름 기간(Flow Duration) 등을 사용할 수 있다.

4.2 OT 장치 식별 연구 현황

본 장에서는 OT 장치를 대상으로 한 조사된 수동적 스캐닝 연구를 표 1을 통해 비교한다. 또한, 앞 절에서 그룹화한 3가지 특징 정보(프로토콜 기반, 패킷 속성 기반, 통신 행위 기반) 관점에서 OT 장치 식별에 사용된 특징 정보들을 조사한다.

표 1의 연구들에서 [19,22,23]은 자산 식별 도구의 한계점 파악을 목표로 연구를 수행하였으며, [20,21]은 특징 정보를 추출하여 자산 식별을 수행하였다. 기계학습 모델을 활용한 연구는 1건으로 FF-ANN(Feed Forward Artificial Neural Network)를 사용했다. 사용된 도구로 PVS, TShark, Grassmarlin이 네트워크 트래픽 수집과 장치 식별을 위해 사용되었다. 공개 데이터셋은 단 1건에서만 사용하였으며, 사용한 공개 데이터셋은 ICS pcaps Github, NETRESEC, SWAT(Secure Water Treatment)[51,52,53]이다. 장치 식별 대상은 OT 장치인 PLC, HMI, RTU(Remote Terminal Unit) 등 이다. [19,22,23]은 현재 알려진 자산 식별 도구가 OT 장치를 완벽하게 식별하지 못한다는 한계점을 보였다. [20,21]에서는 패킷으로부터 몇 가지 특징 정보(CLRT, TTL 등)를 사용하여 장치 모델을 식별할 수 있음을 보였다. 다음으로, 표 1에서 특징 정보를 추출한 연구[20,21]와 앞 장에서 그룹화한 3가지 특징 정보 관점에서 OT 장치 식별에 사용된 특징 정보들을 조사한 내용을 보인다.

프로토콜 기반 특징 정보를 활용한 연구는 능동적 스캐닝 연구 외에 확인되지 않았다. IoT 장치는 기존 IT 시스템에서 사용하던 프로토콜을 그대로 사용하고 있는 것을 파악했다. 반면, OT 장치는 제어 시스템을 위한 프로토콜을 주로 사용한다. 예를 들어, 응용 계층 프로토콜에서 IoT 장치는 HTTP, HTTPS, DHCP, BOOTP, SSDP 등 IT 시스템에서 사용하던 프로토콜을 주로 사용한다. 반면, OT 장치는 Modbus, Profinet, OPC UA, DNP3, BACnet[50] 등을 주로 사용한다.

패킷 속성 기반 특징 정보를 활용한 연구[21]는 2개의 공개 데이터[51,52]에서 특징 정보를 분석하고 11일간 수집된 싱가포르 iTrust의 수처리 SCADA 테스트베드 (SWAT)[53]에서 PLC, HMI 등 9개 장치를 대상으로 특징 정보와 장치가 동작하고 있는 계층을 인식하는 알고리즘을 제안하여 장치를 식별했다. 특징 정보는 패킷으로부터 추출할 수 있는 3가지(TTL, IP ID, MAC ID)를 사용했다. 하지만, 일부 장치에서 특징 정보 값이 중복될 수 있음을 보이며, 사용한 특징 정보가 장치별로 고유한 값을 가지지 않았다. 이를 해결하기 위해 장치가 동작하는 계층을 식별하여 장치를 식별할 수 있음을 보였다.

D. Formby[20] 등은 OT 시스템들이 주기적으로 통신하는 특성을 활용한 연구를 수행하였다. 실시간 전력 변전소에서 약 130개의 장치를 대상으로 약 5개월 동안 네트워크 트래픽을 수집한 데이터와 약 80개의 장치를 가진 다른 변전소에서 데이터를 수집했다. 사용한 특징 정보는 통신 행위로부터 추출할 수 있는 CLRT를 사용하였다. CLRT는 요청에 대한 응답(Response) 시간과 전송 계층에서 확인 응답(ACK) 시간 간의 차이를

말한다. 해당 연구는 장치 하드웨어와 소프트웨어별로 CLRT의 평균과 분산을 확인했을 때, 고유한 분포를 형성하고 있어 특징 정보로 활용될 수 있음을 보였다. 하지만, CLRT를 관찰하기 위해 장치에서 사용하는 프로토콜이 TCP 기반으로 통신해야 하며, TCP ACK의 지연이 발생하지 않도록 장치의“quick ACKs”를 설정해야 하는 조건이 있다.

상기에서 살펴본 바와 같이 OT 장치 식별을 위한 특징 정보에 대한 연구는 IoT 장치 식별을 위한 특징 정보에 대한 연구에 비하여 상대적으로 미흡함을 알 수 있었다.

5. 논의 및 연구 방향

4.2장에서 살펴본 바와 같이 OT 장치에 대한 네트워크 트래픽 특징 정보 식별 연구는 미흡한 수준이다. 이를 해결하기 위해 IoT 장치 식별에서 사용된 특징 정보가 OT 장치에도 적용할 수 있는지 검토하고, OT 장치에 특화된 3가지 유형의 특징 정보를 식별하기 위한 추가적인 연구가 필요하다.

본 연구는 네트워크 스캐닝 기법과 OT 장치를 식별하는 문제를 중심으로 관련 연구들을 조사 분석하였다. 수동적 스캐닝을 위한 네트워크 트래픽 특징 정보 식별 연구 외에도 OT 장치 식별을 위해 해결해야 하는 연구주제들을 다음과 같이 제시한다.

첫째, OT 장치 식별에 기계학습을 적용할 경우 네트워크 트래픽 특징 정보와 기계학습 알고리즘 간의 최적 조합을 찾아내는 연구가 필요하다. 수동적 스캐닝 연구들은 대부분 기계학습을 활용해 장치를 식별했다. IoT 장치 식별 연구는 다양한 네트워크 특징 정보 세트와 기계학습 알고리즘을 통해 높은 식별 정확도를 보였다. 반면, OT 장치 식별 연구는 확인된 특징 정보가 부족하며, 적용한 모델의 수도 적다. 따라서, OT 장치의 네트워크 트래픽 특징 정보에 여러 기계학습 알고리즘을 적용하는 실험이 이루어져야 한다.

둘째, 수동적 스캐닝으로 OT 장치를 식별할 때 최적의 스캐닝 소요시간을 측정하는 연구가 필요하다. OT 장치는 긴 세션 지속 시간으로 인해 세션을 연결하고 끊는 패킷을 확인하기 어렵다[25]. 또한, 2장에서 언급한 바와 같이 휴먼 장치 또는 네트워크 트래픽이 드문드문 생성되는 장치가 존재할 수도 있다. 보안 취약점 관리를 위한 자산 식별에 지나치게 긴 시간을 할당하는 것은 보안 위험을 키울 수 있다. 따라서, OT 네트워크에서 허용 가능한 장치 식별 소요 시간을 산정할 필요가 있다.

셋째, 능동적 스캐닝이 OT 장치에 미치는 가용성 침해 정도를 산출할 수 있으면 가능한 적용 범위를 판단할 수 있다. 장치 식별 정확도가 높고 수행 시간이 짧은 능동적 스캐닝을 선별적으로 적용하기 위한 관련 연구가 필요하다. 역으로 능동적 스캐닝을 수행하기 위해 장치에 필요한 자원 요구사항을 분석할 수도 있다.

넷째, OT 네트워크에서 적정한 시간 범위 내에서 능동적, 수동적 스캐닝을 둘 다 사용해도 식별할 수 없는 OT 장치가 있다면, 제 3의 자동화된 식별 방법이 제안되어야 한다. 능동적 스캐닝과 수동적 스캐닝 두 기법 모두 한계점을 가지고 있어서, 주어진 OT 환경에서 네트워크 스캐닝 기반의 자산 식별 기법이 제공하는 커버리지를 확인할 필요가 있다.

6. 결 론

본 논문에서는 IT 환경에서 자산 식별을 위해 사용하는 네트워크 스캐닝 방식을 OT 네트워크

에 적용할 때의 문제점을 살펴보고, OT 장치의 특성을 반영한 스캐닝 기법을 조사하였다. 또한 수동적 스캐닝 기법을 적용하여 OT 장치를 식별할 때 필요한 네트워크 트래픽의 특징 정보를 선택하기 위해 OT 장치와 유사한 성질을 가진 IoT 장치 식별 연구에서 사용된 네트워크 트래픽 특징 정보를 조사하였다. 이를 통해 네트워크 트래픽 특징 정보의 유형을 분류하고, OT 장치 식별 연구를 조사 분석하였으며, OT 장치 식별을 위한 특징 정보 연구가 필요함을 파악했다. 향후 OT 장치 식별을 위해 수동적 스캐닝을 사용할 때, 3가지 유형의 네트워크 트래픽 특징 정보를 식별하는 연구를 수행할 예정이다.

본 연구는 산업통상자원부(MOTIE)와 한국에너지기술평가원(KETEP)의 지원을 받아 수행한 연구 과제입니다. (No. 20212020800120)

참 고 문 헌

[1] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pilltteri & S. Lightman. (2022). Guide to Operational Technology (OT) Security. National Institute of Standards and Technology. https://csrc.nist.gov/pubs/sp/800/82/r3/ipd

[2] Fortinet Inc. (2022). State of Operational Technology and Cybersecurity Report. https://www.fortinet.com/content/dam/fortinet/assets/analyst-reports

[3] M. Bristow. (2021). A SANS 2021 Survey: OT/ICS Cybersecurity. SANS Institution. https://www.sans.org/white-papers/SANS-2021-Survey-OTICS-Cybersecurity/

[4] Dean Parsons. (2022). The State of ICS/OT Cybersecurity in 2022 and Beyond. SANS Institution. https://www.sans.org/white-papers/state-ics-ot-cybersecurity-2022-beyond/

[5] RISI. (2015). RISI Online Incident Database. https://www.risidata.com/Database

[6] H. Pulkkinen. (2022). Safe security scanning of a production state automation system. Tampere University.

[7] P. M. S. Sanchez, J. M. J. Valero, A. H. Celdrán, G. Bovet, M. G. Pérez & G. M. Pérez. (2021) A Survey on device behavior fingerprinting: Data sources, techniques, application scenarios, and datasets. IEEE Communications Surveys & Tutorials, 23(2), 1048-1077. DOI : 10.1109/COMST.2021.3064259

[8] J. M. Banoczi. (2018). IT Asset Management: Financial Services. National Institute of Standards and Technology, No. NIST Special Publication (SP) 1800-5.

[9] O. Giorgio. (2022). Asset Discovery Tools Supporting Cybersecurity Inventory. POLITECNICO DI TORINO.

[10] E. B. Harb, M. Debbabi & C. Assi. (2013). Cyber scanning: a comprehensive survey. IEEE Communications Surveys & Tutorials,16(3), 1496-1519. DOI : 10.1109/SURV.2013.102913.00020.

[11] G. F. Lyon. (2009). Nmap network scanning: The official Nmap project guide to network discovery and security scanning. Nmap official website. https://nmap.org/

[12] M. Zalewski. (2014). p0f v3. p0f official website. https://lcamtuf.coredump.cx/p0f3/

[13] J, Helms, B. Salazar, P. Scheibel, M. Engels & C. Reiger. (2017). Safe Active Scanning for Energy Delivery Systems Final Report. California : Lawrence Livemore National Lab. https://doi.org/10.2172/1409972

[14] O. Pospisil, P. Blazek, R. Fujdiak & J. Misurec. (2021). Active Scanning in the industrial control systems. IEEE

International Symposium on Computer Science and Intelligent Controls (ISCSIC). (pp.227-232). Rome : IEEE.

[15] T. Hanka, M. Niedermaier, F. Fisher, S. KieBling, P. Knauer & D. Merli. (2020). Impact of Active Scanning Tools for device discovery in industrial network. Security, Privacy, and Anonymity in Computation, Communication, and Storage (SpaCCS), Proceedings 13. (pp.557-572). Nanjing : Springer.

[16] A. Keliris & M. Maniatakos. (2016). Remote Field Device Fingerprinting Using Device-Specific Modbus Information. IEEE 59th International Midwest Symposium on Circuits And Systems (MWSCAS). (pp.1-4). Abu Dhabi : IEEE.

[17] X. Feng, Q.Li, H. Wang & L. Sun. (2016). Characterizing Industrial Control System Devices on the Internet. IEEE 24th International Conference on Network Protocols (ICNP). (pp.1-10). Singapore : IEEE

[18] M. Cash, S. Wang, B. Pearson, Q. Zhou & X. Fu. (2021) On Automating BACnet Device Discovery and Property Identification. ICC 2021 - IEEE International Conference on Communications. (pp.1-6). Montreal : IEEE.

[19] A. Wedgbury & K. Jones. (2015). Automated Asset Discovery in Industrial Control systems-exploring the problem. 3rd International Symposium for ICS & SCADA Cyber Security Research (ICS-CSR). (pp.73-83). Germany : ICS-CSR.

[20] D. Formby, P. Srinivasan, A. Leonard, J. Rogers & R. Beyah. (2016). Who's in Control of Your Control System? Device Fingerprinting for Cyber-Physical Systems. Network and Distributed System Security Symposium. San Diego : NDSS.

[21] AT. AI. Ghazo & R. Kumar. (2019). ICS/SCADA Device Recongnition: A Hybrid Communication-Patterns and Passive-fingerprinting Approach. IFIP/IEEE Symposium on Integrated Network and Service Management (IM). (pp.19-24). Arlington : IEEE.

[22] M. Abdularazzaq & Y. Wei. (2018). Industrial Control System (ICS) Network Asset Identification and Risk Management. HALMSTAD University.

[23] Haein Kang, Minsu Park. Seongje Cho & Jihun Jung. (2022). Limitation of Scanning Tools for Asset Discovery in Operational Technology Networks. Korea Software Congress. (pp.920-922). Jeju : KIISE

[24] D. Duggan, M. Berg, J. Dillinger & J. Stamp. (2005) Penetration Testing of Industrial Control Systems. Sandia National Laboratories.

[25] M. Caselli, D. Hadžiosmanović, E. Zambon & F. Kargl. (2013). on feasibility of device fingerprinting in industrial control systems. Critical Information Infrastructures Security: 8th International Workshop (CRITIS), Revised Selected Papers 8. (pp.155-166). Amsterdam : Springer.

[26] W. A. Conklin. (2016). IT vs OT Security: A Time to Consider a Change in CIA to Include Resilience. 49th Hanwaii International Conference on System Sciences (HICSS). (pp.2642-2647). Koloa : IEEE.

[27] M. Miettinen, S. Marchal, I. Hafeez, N. Asokan, A. R. Sadeghi & S. Tarkoma. (2017). IoT Sentinel: Automated device-type identification for security enforcement in IoT. IEEE 37th International Conference on Distributed Computing Systems (ICDCS). (pp.2177-2184). Atlanta : IEEE.

[28] A. Sivanathan, H. H. Gharakheili, F. Loi, A. Radford, C. Wijenayake, A. Vishwanath & V. Sivaraman. Classifying IoT Devices in Smart Environments Using Network Traffic Characteristics. IEEE Transactions

on Mobile Computing, 18(8), 1745-1759. DOI : 10.1109/TMC.2018.2866249

[29] B. Bezawada, M. Bachani, J. Peterson, H. Shirazi, I. Ray & I. Ray. (2018). Behavioral Fingerprinting of IoT devices. Proceedings of the 2018 workshop on attacks and solutions in hardware security. (pp.41-50). Toronto : ACM.

[30] M. R. Shahid, G. Blanc, Z. Zhang & H. Debar. (2018). IoT Devices Recognition Through Network Traffic Analysis. IEEE International Conference on big data (big data). (pp.5187-5192). Seattle : IEEE.

[31] V. Thangavelu, D. M. Divakaran, R. Sairam, S. S. Bhunia & M. Gurusamy. (2018) DEFT: A Distributed IoT Fingerprinting Technique. IEEE Internet of Things Journal, 6(1). 940-952. DOI : 10.1109/JIOT.2018.2865604

[32] L. Bai, L. Yao, S. S. Kanhere, X. Wang & Z. Yang. (2018). Automatic Device Classification from Network Traffic Streams of Internet of Things. IEEE 43rd Conference on Local Computer Networks (LCN), (pp.1-9). Chicago : IEEE.

[33] N. Msadek, R. Soua & T. Engel. (2019). IoT device fingerprinting: Machine learning based encrypted traffic analysis. IEEE Wireless Communications and Networking Conference (WCNC). (pp.1-8). Marrakesh : IEEE.

[34] S. A. Hamad, W. E. Zhang, Q. Z. Sheng & S. Nepal. (2019). IoT Device Identification via network-Flow Based Fingerprinting and Learning. 18th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/13th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE). (pp.103-111). Rotorua : IEEE.

[35] S. Marchal, M. Miettinen, T. D. Nguyen, A. R. Sadeghi & N. Asokan. (2019). AuDI: Toward Autonomous IoT Device-Type Identification Using Periodic Communication. IEEE Journal on Selected Areas in Communications, 37(6), 1402-1412. DOI : 10.1109/JSAC.2019.2904364

[36] A. Aksoy & M. H. Gunes. (2019). Automated IoT Device Identification using Network Traffic. IEEE International Conference on Comunications (ICC). (pp.1-7). Shanghai : IEEE.

[37] R. R. Chowdhury, S. Aneja, N. Aneja & E. Abas. (2020). Network Traffic Analysis based IoT Device Identification. Proceedings of the 2020 4th International Conference on Big Data and Internet of Things. (pp.79-89). Singapore : ACM.

[38] N. Najari, S. Berlemont, G. Lefebvre, S. Duffner & C. Garcia. (2020). Network Traffic Modeling For IoT-device Re-identification. International Conference on Omni-layer Intelligent Systems (COINS). (pp.1-6). Barcelona : IEEE.

[39] L. Babun, H. Aksu, L. Ryan, K. Akkaya, E. S. Bentley & A. S. Uluagac. (2020). Z-IoT: Passive Device-class Fingerprinting of ZigBee and Z-Wave IoT Devices. IEEE International Conference on Communications (ICC). (pp.1-7). Dublin : IEEE.

[40] M. Skowron, A. Janicki & W. Mazurczyk. (2020). Traffic Fingerprinting Attacks on Internet of Things Using Machine Learning. IEEE Access, 8, 20386-20400. DOI : 10.1109/ACCESS.2020.2969015

[41] V. A. Ferman & M. A. Tawfeeq. (2021). Machine Learning Challenges for IoT Device Fingerprints Identification. Journal of Physics: Conference Series, 1963(1), p012046. DOI : 10.1088/1742-6596/1963/1/012046

[42] F. Yin, L. Yang, Y. Wang & J. Dai. (2021). IoT ETEI: End to-End IoT Device Identification Method. IEEE Conference on Dependable and Secure Computing (DSC). (pp.1-8). Aizuwakamatsu : IEEE.

[43] O. Salman, I. H. Elhajj, A. Chehab & A.

Kayssi. (2022). A machine learning based framework for IoT device identification and abnormal traffic detection. Transactions on Emerging Telecommunications Technologies, 33(3). e.3743. DOI : https://doi.org/10.1002/ett.3743

[44] K. Kostas, M. Just & M. A. Lones. (2022). IoTDevID: A Behavior-Based Device Identification Method for the IoT. IEEE Internet of Things Journal, 9(23), 23741-23749. DOI : 10.1109/JIOT.2022.3191951

[45] YourThings dataset Alrawi Omar, Lever Chaz, Antonakakis Manos & Monrose Fabian. (2019). SoK: Security Evaluation of Home-Based IoT Deployments. IEEE Symposium on Security and Privacy (SP). (pp.1362-1380). San Francisco : IEEE.

[46] Y. Meidan, M. Bohadana, A. Shabtai, M. Ochoa, N. O. Tippenhauer, J. D. Guarnizo & Y. Elovici. (2017). Detection of Unauthorized IoT Devices Using Machine Learning Techniques. arXiv preprint arXiv:1709, 04647. DOI : https://doi.org/10.48550/arXiv.1709.04647

[47] E. Hjelmvik. (2011). Passive OS Fingerprinting. Netresec. https://www.netresec.com/?page=Blog&month=2011-11&post=Passive-OS-Fingerprinting

[48] AR. Khakpour & A.X Liu. (2012). An Information-Theoretical Approach to High-Speed Flow Nature Identification. IEEE/ACM Transactions on networking, 24(4), 1076-1089. DOI : 10.1109/TNET.2012.2219591

[49] L Wang, L Peng, M Su, B Yang, & X Zhou. (2016). On the impact of packet inter arrival time for early stage traffic identification. IEEE International Conference on Internet of Things (iThings) and Green Computing and Communications (greenCom) and Cyber, Physical and Social Computing (CPSCom) and Smart Data (SmartData). (pp.510-515). Chengdu : IEEE.

[50] SV Radhakrishnan, AS Uluagac, & R Beyah. (2014). GTID: A technique for physical device and device type fingerprinting. IEEE Transactions on Dependable and Secure Computing, 12(5), 519-532. DOI : 10.1109/TDSC.2014.2369033

[51] T. Yardley. (2018). ICS-Security Tool Github. https://github.com/ITI/ICS-Security-Tools/tree/master/pcaps.

[52] L. Hansson. (2018). Capture files from 4sics geek lounge. https://www.netresec.com/index.ashx?page=PCAP4SICS

[53] iTrust Lab. (2018). Secure water treatment tesetbed. https://itrust.sutd.edu.sg/testbeds/secure-water-treatment-swat/

[54] Digital bond. (2011). Control System Port List. https://web.archive.org/web/20111203052654/http://www.digitalbond.com/tools/the-rack/control-system-port-list

저 자 소 개

박민수(Minsu Park)

2020.2 단국대학교 소프트웨어학과 졸업
2022.3-현재 단국대학교 인공지능융합학과 석사과정
<주관심분야> 시스템 보안, 네트워크 보안

안석현(Seokhyun Ahn)

2020.3-현재 단국대학교 소프트웨어학과 학사과정
<주관심분야> 시스템 보안, 네트워크 보안

박세연(Seyeon Park)

2020.3-현재 단국대학교 소프트웨어학과 학사과정
<주관심분야> 머신러닝, 인공지능, 딥러닝, 정보보안, 안드로이드 등

조성제(Seong-je Cho)

1989.2 서울대학교 컴퓨터공학과 공학사
1991.2 서울대학교 컴퓨터공학과 공학석사
1996.8 서울대학교 컴퓨터공학과 공학박사
1997년 3월~현재 단국대학교 소프트웨어학과/컴퓨터학과 교수
<주관심분야> 시스템 보안 및 악성코드 분석, 소프트웨어 보증, 시스템 소프트웨어 임베디드 소프트웨어 등

김홍근(Honggeun Kim)

1994.5 한국전산원
1996.5 한국인터넷진흥원
<주관심분야> 컴퓨터보안, 정보보호