정 보 보 호 학 회 지
제 29 권 제6호, 2019. 12

사물인터넷 환경의 이상탐지를 위한 경량 인공신경망 기술 연구

오 성 택*, 고 웅*, 김 미 주*, 이 재 혁*, 김 홍 근*, 박 순 태*

요 약

최근 5G 네트워크의 발전으로 사물인터넷의 활용도가 커지며 시장이 급격히 확대되고 있다. 사물인터넷 기기가 급증하면서 이를 대상으로 하는 위협이 크게 늘며 사물인터넷 기기의 보안이 중요시 되고 있다. 그러나 이러한 사물인터넷 기기는 기존의 ICT 장비와는 다르게 리소스가 제한되어 있다. 본 논문에서는 이러한 특성을 갖는 사물인터넷 환경에 적합한 보안 기술로 네트워크 학습을 통해 사물인터넷 기기의 이상행위를 탐지하는 경량화된 인공신경망 기술을 제안한다. 기기 별 혹은 사용자 별 네트워크 행위 패턴을 분석하여 특성 연구를 진행하였으며, 사물인터넷 기기의 정상행위를 수집하고 학습데이터로 활용한다. 이러한 학습데이터를 통해 인공신경망 기반의 오토인코더 알고리즘을 활용하여 이상행위 탐지 모델을 구축하였으며, 파라미터 튜닝을 통해 모델 사이즈, 학습 시간, 복잡도 등을 경량화 하였다. 본 논문에서 제안하는 탐지 모델은 신경망 프루닝 및 양자화를 통해 경량화된 오토인코더 기반 인공신경망을 학습하였으며, 정상 행위 패턴을 벗어나는 이상행위를 식별할 수 있었다. 본 논문은 1. 서론을 통해 현재 사물인터넷 환경과 보안 기술 연구 동향을 소개하고 2. 관련 연구를 통하여 머신러닝 기술과 이상 탐지 기술에 대해 소개한다. 3. 제안기술에서는 본 논문에서 제안하는 인공신경망 알고리즘 기반의 사물인터넷 이상행위 탐지 기술에 대해 설명하고, 4. 향후연구계획을 통해 추후 활용 방안 및 고도화에 대한 내용을 작성하였다. 마지막으로 5. 결론을 통하여 제안기술의 평가와 소회에 대해 설명하였다.

Ⅰ. 서 론

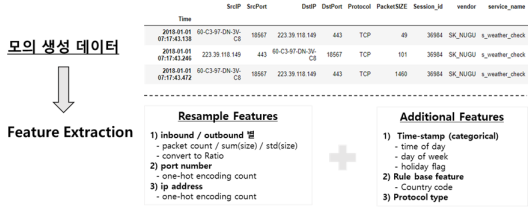
사물인터넷 기술을 이용하는 다양한 제품과 서비스가 출시되고 있다. 5G네트워크의 상용화로 사물인터넷의 활용도가 급증하며 관련 산업시장이 급격히 확대되고 있다. McAfee의 2016년 위협 보고서에 따르면 2020년까지 사물인터넷 기기의 공급수가 60억대로 증가하며, 244억 개의 기기가 인터넷에 연결되어 월 데이터 사용량이 44제타 바이트로 급증할 것으로 전망하였다. 최근 IHS마켓의 분석 결과에 따르면 전 세계 사물인터넷 기기가 2020년 400억대에서 10년 후 2030년에는 약 1,000억대가 증가한 1,400억대로 급증할 것으로 발표하였다. 하지만 이러한 사물인터넷 환경의 급격한 발전과 기기의 급증은 취약 기기의 노출 증가와 이를 악용한 사이버 범죄의 증가로 이어질 것이다. KT 경제경영연구소의 조사에서는 2030년 국내 사물인터넷 해킹 피해액만 26조 7천억에 달할 것으로 분석되고 있다. 또한 사물인터넷이 일상생활과 밀접해지면서 금전적인 피해뿐만 아니라 생명에도 피해를 입힐 수 있기 때문에 사물인터넷 보안에도 큰 관심이 집중되고 있다. 현재 국내 사물인터넷 보안 기술 연구는 기기 인증 및 암호화를 초점으로 연구개발이 진행되고 있으며, KISA에서는 사물인터넷 제품의 보안내재화를 위해 사물인터넷 하드웨어와 소프트웨어 전반에 걸쳐 준수해야 할 7개의 보안원칙을 제시한 ‘사물인터넷 공통 보안원칙‘을 발표하고 사물인터넷 공통보안 가이드를 제공하였다. 하지만 사이버 침해사고 발생에 대한 대응 기술은 미흡한 상태이다. 국외의 경우 사물인터넷 보안법을 제정하여 기기의 설계 단계부터의 보안 내재화를 법적으로 의무화 하는 등 사물인터넷 보안에 적극적으로 대응하고 있다. 하지만 기존에 사용되고 있는 다양한 사물인터넷 제품들은 여전히 취약한 상태로 인터넷

이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2018-0-00232, 클라우드 기반 IoT 위협 자율 분석 및 대응 기술개발)
* 한국인터넷진흥원 정보보호R&D기술공유센터 보안위협대응R&D팀 ({angelrick, wgo, mijoo.kim, jaehyuk, kimhg, stpark12}@kisa.or.kr)

에 연결되어 미라이 등의 악성코드의 공격 대상이 되고 있다. 또한 리소스가 제한된 저사양의 사물인터넷 제품들은 보안 내재화에 한계가 존재한다. 이러한 보안이 미흡한 사물인터넷 기기를 대상으로 하는 공격이 증가하고, 기존의 ICT 환경의 주요 침해사고가 향후 사물인터넷 환경으로 전이·확산될 것으로 예상된다. 보안에 대한 전문 지식이 부족한 사물인터넷 기기 사용자는 침해사고의 발생여부조차 인지하기 어려우며, 인지하여도 대응하기가 쉽지 않다. 본 논문에서는 이러한 사물인터넷 환경의 특성을 고려한 사물인터넷 환경에서의 이상행위 탐지 기술을 제안한다. 본 논문에서 제안하는 기술은 사용자의 사물인터넷 환경에서 스마트 게이트웨이에 위치하며 기기의 행위로 발생되는 네트워크 정보를 수집하고 기기 별 행위 패턴 분석하며 탐지 모델을 생성한다.

Ⅱ. 특성 연구

네트워크 트래픽 데이터 마이닝 관련 선행연구를 참고하여 적절성, 실현 가능성 등을 고려하여 추출할 특성 정보를 선정하였다. 특성 선정 과정에서 많은 선행연구의 경우 와이어샤크 등을 사용하여 다양한 변수들을 추출하였다. 본 논문에서는 사물인터넷 환경이 가지는 제약을 고려하여 5튜플과 시간정보를 사용하는 특성 정보를 중점적으로 연구하였다.



(그림 1) 특성 선정 과정

Ⅲ. 알고리즘 연구

모델 성능은 일정 수준 이상으로 유지하되 복잡도와 학습시간 등을 최소한으로 낮출 수 있는 머신러닝 기술에 초점을 맞추어 연구를 진행하였다. 고차원 문제에 적합하고 성능이 뛰어나 이상탐지에 가장 많이 쓰이고 있는 방법론인 One-Class SVM, Isolation Forest, Autoencoder를 표 1과 같은 평가 기준으로 비교해 대상 모델을 선정하였다[1].

〔표 1〕 모델 평가 지표

지표	세부 항목
성능	• 모델이 데이터에 얼마나 적합(Fit)되었는가 • Vanishing Gradient Problem • Noise에 강건한가
복잡도	• Neuron, Layer의 개수 • Parameter의 개수
시간	• Training Time • 빠른 Convergence Rate

3.1. Isolation Forest

트리기반으로 분류하며 모든 데이터 관측치를 고립시키는 방법이다. 특정 한 개체가 고립되는 종료 노드까지의 거리를 아웃라이어 점수로 정의하며 평균거리가 짧을수록 아웃라이어 점수가 높아진다. 비정상 데이터가 고립되려면 시작 노드와 가까운 평균거리를 가지며, 정상 데이터가 고립되려면 트리구조의 종료 노드에 가까운 평균거리를 갖는다[2]. 적은 데이터부터 많은 데이터까지 모두 높은 성능을 보이며, False Alarm의 비율이 낮다. 낮은 시간 복잡도(O(n))를 가지며 모델 학습 시 이상 데이터의 비율이 높을 경우 학습이 잘 되지 않았다.

〔표 2〕 Isolation Forest Hyperparameters

파라미터	설명
n_estimators	앙상블 base estimator의 개수
max_samples	base estimator 학습 시 사용되는 sub-sampling 데이터 개수
max_features	base estimator 학습 시 사용되는 특성 개수
bootstrap	Bootstrap 샘플링 사용

3.2. One-Class SVM

SVM 알고리즘의 비지도 학습 버전으로 커널 함수를 통해 데이터를 적절한 변수 공간에 매핑하고 원점으로 부터 마진이 최대가 되도록 분리한다. 정상 점을 둘러싸는 가장 작은 하이퍼스피어를 찾는 것과 동일하다[3]. 특징으로는 작은 데이터 셋에도 뛰어난 성능을

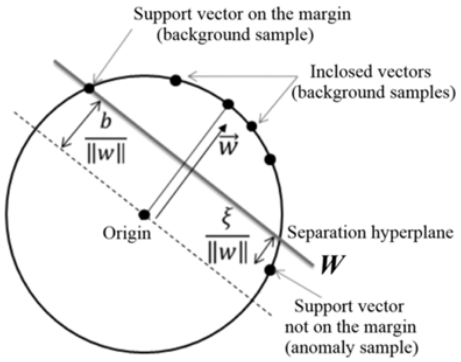
모델		One-Class SVM	Isolation Forest	Autoencoder
라이브러리		sklearn	sklearn	tensorflow
사용된 변수		• Features의 종류 • Kernel의 종류 • nu의 크기 • gamma의 크기	• Features의 종류 • n_estimators • bootstrap 여부 • max_features의 크기	• Features의 종류 • learning rate • Optimizer • activation function • Initializer • hidden_layer size
평가 지표	Anomaly Detecting Performance	Recall / Precision / Accuracy / Etc.		
	Model Performance	Model Size / Training Time / Prediction Time		

(그림 2) 모델 평가 지표

보이며, 강건한 성능을 보인다. 고차원 문제의 해결에 적합하며 시간 복잡도는 $O(n^2)$을 갖는다.

〔표 3〕 One-Class SVM Hyperparameters

파라미터	설명
Kernel	Kernel Function의 종류
Gamma	kernel Function이 Polynomial이나 rbf의 경우 설정
Nu	Slack variable의 강도



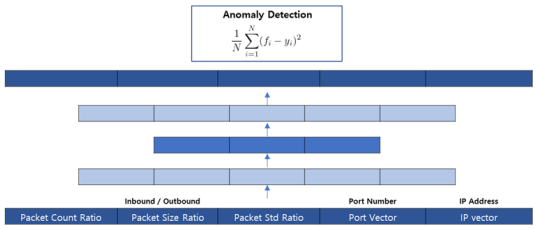
(그림 3) One-Class SVM

3.3. Autoencoder

강력한 Representation power를 바탕으로 여러 분야를 통틀어 독보적으로 뛰어난 성능을 보이는 알고리즘이다[4]. Manifold hypothesis를 바탕으로 낮은 차원의 특성을 추출하는 기법으로써 이상탐지 분야에 주로 쓰이는 방법론이다. 신경망의 경우 기본적으로 다양한 파라미터들의 조합을 튜닝하여 하나의 모델을 만들어 낸다. 해당 과정에서 굉장히 많은 조합들이 등장하며 실험을 반복하여 목적에 맞는 최적의 파라미터를 찾아나간다.

〔표 4〕 Autoencoder Hyperparameters

파라미터	설명
Hidden layer	은닉 층의 넓이 및 깊이
Optimizer	학습을 위해 사용되는 최적화 알고리즘
Activation function	은닉 노드에 비선형성을 추가하기 위한 함수



(그림 4) Autoencoder

Ⅳ. 실험 및 경량화 연구

4.1. 실험 데이터 생성 및 검증

연구 기술의 검증을 위해 모의 데이터를 생성하였다. 모의 데이터의 생성 기간은 12개월로 설정하였으며, 9개월분의 학습 데이터와 3개월분의 검증 데이터로 구성하였다. 정상 데이터만을 학습하여 이상행위를 탐지하는 것이 목적으로 모의 데이터의 생성 시나리오는 그림 6과 같다.

실험에 사용된 특성 정보들은 5초단위로 분할된 네트워크 트래픽 데이터를 리샘플하여 추출하였다. 사용

1. Autoencoder Training: P: 13,449 Test P: 742 N: 1197

Parameters						Result						
Features	Learning Rate	Optimizer	Activation	Hidden layer	Initializer	TP	FP	FN	TN	Train_time	Predict_time	Model_size
All	0.0001	GD	Sigmoid	64	Xavier	736	0	6	1197	6.1s	0.15s	99KB
	0.005	GD	Sigmoid	64	Xavier	737	0	5	1196	5.8s	0.10s	120KB

※ Auto Encoder의 경우 Feature extraction의 기능을 기대할 수 있기 때문에 feature engineering의 비중이 덜함

2. One-Class SVM

Parameters				Result						
Features	Kernel	nu	gamma	TP	FP	FN	TN	Train_time	Predict_time	Model_size
All	poly	0.0001	auto	736	128	6	1069	0.003s	0.001s	2.1KB
	rbf	0.01	auto	742	9	0	1188	0.003s	0.001s	34KB
Main_only	rbf	0.01	auto	737	7	5	1190	0.01s	0.002s	4.3KB

3. Isolation Forest

Parameters				Result						
Features	n_estimators	bootstrap	max_features	TP	FP	FN	TN	Train_time	Predict_time	Model_size
All	500	TRUE	1	737	131	5	1066	0.79s	0.26s	1,908KB
	1000	FALSE	1	736	133	6	1064	1.614s	0.52s	4,012KB
Main_only	500	FALSE	10	742	120	0	1077	0.85s	0.27s	3,578KB

※ 실험 환경 : 2.4GHz 20core(CPU), 263GB(Ram), Quadro4000(GPU)

(그림 5) 모델 평가 결과

	시간	행동	최대 시행 횟수
주중	7:05~8:00	이상행동	10
	7:15~7:25	날씨확인	5
	7:30~7:45	정보탐색	2
	18:30~20:00	음악 스트리밍	1회 10분
	19:00~19:30	냉장고 주문	1
	20:30~20:45	정보탐색	5
주말	8:05~9:00	이상행동	3
	11:30~11:45	뉴스 확인	2
	12:00~13:45	냉장고 주문	3
	19:30~19:20	냉장고 주문	1
	20:45~22:00	정보탐색	5

(그림 6) 데이터 생성 시나리오

된 변수는 아래 그림 7과 같으며, 실험에 사용된 모델 및 평가지표는 그림 2와 같다.

사용 변수	세부 정보
Main Features	• Inbound packet count / sum(size) / std(size) • outbound packet count / sum(size) / std(size) • Ratio of inbound and outbound
Source / Destination IP address(one-hot)	• Known IP • Unknown IP
Source / Destination port(one-hot)	• Well-Known port • Unknown port

(그림 7) 특성 추출 정보

4.2. 오토인코더 기반 경량 이상탐지 기술

그림 5의 실험 결과를 통해 사물인터넷 환경에서의 이상행위 탐지 기술에 적합한 알고리즘으로 오토인코더 알고리즘을 선정하였다. 선정된 알고리즘을 대상으로 사물인터넷 환경에 적합한 경량화 연구를 수행하였다. 최대한 높은 성능을 유지하면서 모델의 사이즈, 연산시간, 모델 복잡도 등을 경량화 할 수 있는 기존 방법론들을 조사하였으며, 기존 신경망의 파라미터들을 경량화 관점에서 재해석해 파라미터를 튜닝 하였다. 또한 모델의 구조 자체를 경량화 시킬 수 있는 방안에 대해서 연구하였다.

4.3. 신경망 프루닝(Neural Network Pruning)

대형 신경망은 매우 강력한 기법이지만 크기가 커서 메모리 사용량, 메모리 대역폭 및 연산 리소스를 크게 소비한다. 또한 많은 경우에 학습을 시작할 때 최적의 파라미터를 알 수 없기 때문에 충분한 수의 파라미터를 사용할 수 있도록 신경망을 구성한다. 하지만 파라미터를 확인해 보면 성능에 영향을 주지 않는 파라미터가 상당수 존재한다. 또한 사물인터넷 환경에서는 연산 성능이 매우 제한적이기 때문에 본 논문에서는 이러한 문제를 최소화 시키는 방법으로 신경망 프루닝(Pruning) 방법을 사용하였다. 학습 단계 이후 가중치 값이 낮은 파라미터를 제거하면서 밀도가 낮은 Sparse 네트워크로 변환한 뒤 재학습하는 과정을 반복하였다. 가중치 값이 낮은 파라미터라고 하여도 제거에 따른 성능의 손실이 발생할 수 있기 때문에 추가 학습을 통해 정확도 손실을 보정할 수 있다. 이를 통해 모델의 성능에 영향이 없는 쓸모없는 파라미터를 제거하며 성능은 유지하고 모델의 사이즈와 복잡도를 감소시켰다[5].

4.4. 양자화(Quantization)

사물인터넷 환경에서의 이상탐지를 위해 구성한 인공신경망의 경우 부동 소수점 수준의 정확한 연산이 요구되지는 않는다. 본 논문에서는 연산 효율성을 증가시키기 위해 양자화를 통해 부동 소수점 연산을 고정 소수점 연산으로 변환하여 비교적 적은 정확도의 손실이 있지만 실시간 탐지를 위해 연산 성능을 높이는 연구를 진행하였다.

Ⅴ. 향후연구계획

5.1. 이상탐지 결과의 악성여부 진단

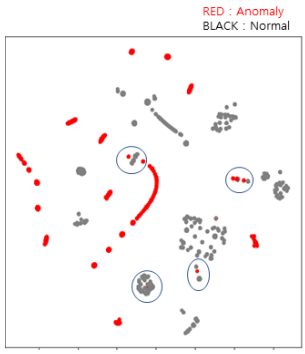
사물인터넷 환경에서의 네트워크 위협 탐지 모델을 통해 본 논문에서 제안하는 시스템에서 이상으로 탐지된 건에 대한 악성여부 진단을 연동할 계획이다. 관련하여 위협 탐지 시스템에서 활용하는 특성 정보와 인공신경망 구조의 연결 및 가중치 정보를 수집하고 본 논문에서 제안하는 모델과의 비교 분석을 통해 탐지 성능의 향상을 위한 고도화를 수행한다.

5.2. 전이학습을 통한 모델 튜닝

클라우드 플랫폼 기반의 빅데이터 분석을 통해 수집된 데이터를 통해 인공신경망을 지속적으로 학습하여 사물인터넷 환경에서 동작 중인 혹은 동작 예정인 모델에 인공신경망 구조 정보를 전달하여 초기 학습시간 절약할 수 있다. 또한 초기 정상행위 데이터를 수집하는 공백 기간을 제거하여 미탐 위협을 줄일 수 있다.

Ⅵ. 결 론

사물인터넷 환경에서의 이상행위 탐지 기술 개발을 위해 성능이 뛰어난 3개의 알고리즘에 대한 실험을 진행하였다. Rbf 커널을 사용한 SVM 알고리즘과 신경망 모델의 성능이 뛰어났으며, 검증 데이터 셋에 t-SNE를 적용하여 분석한 결과 그림 8과 같은 분포를 확인 할 수 있었다. 대개의 경우에서 이상과 정상의 데이터 분포가 달라 쉽게 구분이 가능해 보이나 특정 부분에서의 이상을 식별하기 어려웠다. 데이터의 분포 상



(그림 8) t-SNE 적용 결과

선형 분리가 어렵기 때문에 이상탐지 성능을 향상시키기 위해서는 비선형적 특성을 반영할 수 있는 모델이 필요하다. 본 논문에서는 사물인터넷 환경에서의 경량화된 이상탐지 기술 개발을 위해 3가지 모델에 대한 실험을 수행하였다. 실험결과 오토인코더 모델이 대개의 평가기준에서 월등한 성능을 보였으며, 기존 알고리즘에 비해 특성 추출 등의 필요성이 줄어든다는 장점도 확인하였다. 또한 이상행위 패턴 변화에 빠르게 대응하기 위해서는 주기적인 재학습이 필요하며, 오코인코더의 경우 재학습시 신규 데이터만 학습시키므로 패턴 변화 대응에 유리함을 갖고 있었다. 하지만 모델의 복잡도가 다른 모델에 비해 높아 과적합의 우려가 있고 추론시간이 증가하는 단점이 존재하였다.

참 고 문 헌

[1] Tuor, Aaron, "Deep learning for unsupervised insider threat detection in structured cybersecurity data streams," Workshops at the Thirty-First AAAI Conference on Artificial Intelligence. 2017

[2] Liu, Fei Tony, Kai Ming, and Zhi Hua Zhou, "Isolation forest," 2008 Eighth IEEE International Conference on Data Mining, pp. 413-422, 2008.

[3] B. Scholkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola & R. C. Williamson, "Estimating the support of a high-dimensional distribution," Neural computation, 13(7), pp. 1443-1471, 2001.

[4] Chalapathy, Raghavendra, and Sanjay Chawla, "Deep learning for anomaly detection: A survey," arXiv preprint arXiv:1901.03407, 2019.

[5] S. Han, H. Mao & W. J. Dally, "Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding," arXiv prepring arXiv:1510.00149, 2015.

〈저자소개〉

오 성 택 (Sungtaek Oh)

정회원

2013년 2월 : 아주대학교 정보컴퓨터공학부 졸업

2016년 2월 : 아주대학교 컴퓨터공학과 석사

2015년 2월~현재 : 한국인터넷진흥원 보안위협대응R&D팀 선임연구원

<관심분야> 머신러닝, 침입탐지, 정보보호

고 웅 (Woong Go)

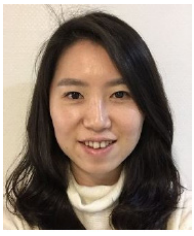
정회원

2010년 2월 : 순천향대학교 정보보호학과 석사

2013년 8월 : 순천향대학교 정보보호학과 박사

2014년 1월~현재 : 한국인터넷진흥원 보안위협대응R&D팀 책임연구원

<관심분야> IoT, 기계학습, 정보보호



김 미 주 (Mijoo Kim)

정회원

2006년 2월 : 순천향대학교 정보보호학과 졸업

2008년 2월 : 순천향대학교 정보보호학과 석사

2008년 9월~현재 : 순천향대학교 정보보호학과 박사과정

2008년 4월~현재 : 한국인터넷진흥원 책임연구원

<관심분야> 정보보호, IoT 및 모바일 보안, 융합보안

이 재 혁 (Jaehyuk Lee)

학생회원

2014년 2월 : 학점은행제 정보보호학사

2018년 2월 : 고려대학교 정보보호석사

2017년 4월~현재 : 한국인터넷진흥원 보안위협대응R&D팀 주임연구원

<관심분야> 빅데이터, AI, 정보보호

김 홍 근 (Kim Hong-Geun)

종신회원

1985년 2월 : 서울대학교 전자계산기공학과 졸업

1987년 2월 : 서울대학교 전자계산기공학과 석사

1994년 2월 : 서울대학교 컴퓨터공학과 박사

1994년 5월~현재 : 한국인터넷진흥원 연구위원

<관심분야> 병렬알고리즘, 사이버보안, 컴퓨터보안

박 순 태 (SoonTai Park)

정회원

1992년 2월 : 단국대학교 전자계산학과 졸업

1998년 8월 : 국민대학교 정보과학대학원 정보통신학과 석사

2010년 8월 : 전남대학교 대학원 정보보안협동과정 박사

2000년 4월~현재 : 한국인터넷진흥원 팀장

<관심분야> IT보안성 평가, 정보보호 인력 양성, 정보통신기반보호, 조직 정보보안/개인정보보호 실무, 정보보호 R&D