Journal of the Korea Convergence Society
Vol. 10. No. 5, pp. 9-16, 2019

ISSN 2233-4890
https://doi.org/10.15207/JKCS.2019.10.5.009

딥러닝 기술이 가지는 보안 문제점에 대한 분석

최희식[1], 조양현[2*]
[1]삼육대학교 컴퓨터·메카트로닉스공학부 외래교수, [2]삼육대학교 컴퓨터·메카트로닉스공학부 교수

Analysis of Security Problems of Deep Learning Technology

Hee-Sik Choi[1], Yang-Hyun Cho[2*]
[1]Assistant processor, Division of Computer & Mechatronics Engineering, Sahmyook University
[2]Processor, Division of Computer & Mechatronics Engineering, Sahmyook University

요 약 본 논문에서는 딥러닝 기술이 인터넷과 연결된 다양한 비즈니스 분야에 새로운 형태의 비즈니스 업무에 활용할 수 있도록 보안에 관한 문제점을 분석하고자 한다. 우선 딥러닝이 비즈니스 영역에 보안 업무를 충분히 수행하기 위해서는 많은 데이터를 가지고 반복적인 학습을 필요하게 된다. 본 논문에서 딥러닝이 안정적인 비즈니스 보안 업무를 완벽하게 수행할 수 있는 학습적 능력을 얻기 위해서는 비정상 IP패킷에 대한 탐지 능력과 정상적인 소프트웨어와 악성코드를 탑재하여 감염 의도를 가지고 접근하는 공격을 탐지해낼 수 있는 인지 능력을 갖추고 있는지를 분석하였다. 이에 본 논문에서는 인공지능의 딥러닝 기술이 시스템에 접근하여 문제의 비즈니스 모델을 안정적으로 수행할 수 있게 하기 위해서는 시스템내의 비정상 데이터를 추출해 내고 시스템 데이터 침해를 구분해 낼 수 있는 수학적 역할의 문제점을 보완하기 위해 새로운 IP에 대한 세션 및 로그 분석을 수행할 수 있도록 보안 엔진이 탑재된 딥러닝 기술을 개발하여 비즈니스 모델에 적용시켜서 취약점을 제거하여 비즈니스 업무 능력을 향상시키도록 문제적 방안을 비교 분석하였다.

주제어 : 인공지능, 기계학습, 딥러닝, 보안, 비즈니스모델

Abstract In this paper, it will analyze security problems, so technology's potential can apply to business security area. First, in order to deep learning do security tasks sufficiently in the business area, deep learning requires repetitive learning with large amounts of data. In this paper, to acquire learning ability to do stable business tasks, it must detect abnormal IP packets and attack such as normal software with malicious code. Therefore, this paper will analyze whether deep learning has the cognitive ability to detect various attack. In this paper, to deep learning to reach the system and reliably execute the business model which has problem, this paper will develop deep learning technology which is equipped with security engine to analyze new IP about Session and do log analysis and solve the problem of mathematical role which can extract abnormal data and distinguish infringement of system data. Then it will apply to business model to drop the vulnerability and improve the business performance.

Key Words : Convergence, AI, Machine Learning, Deep Learning, Security, Business Model

*This study is supported by Basic Science Research Program through the Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2017R1D1A1B03030759)
*Corresponding Author : Yang-Hyun Cho(yhcho@syu.ac.kr)
Received February 25, 2019 Revised April 02, 2019
Accepted May 20, 2019 Published May 30, 2019

1. 서론

최근 4차 산업혁명 시대가 도래되면서 인공지능과 빅데이터 처리에 대한 활용이 높아지고 있다. 인공지능 인식은 최근 전 세계적 관심을 모았던 이세돌 선수와 인공지능 알파고의 바둑 대결을 기억하면 쉽게 이해할 수 있다. 인공지능의 활용은 인공지능의 핵심기술의 하나인 딥러닝이라 하는 인공신경망 기반의 기술로 어떠한 데이터가 있을 때 이를 컴퓨터가 알아들을 수 있는 명령 중심으로 학습을 통해 적용하도록 되어있으며 학습 후, 목적에 부합된 특수한 영역에 활용될 수 있는 알고리즘 형태의 기계학습이다. 본 논문에서는 인공지능 기술의 하나인 딥러닝 기술이 인터넷과 연결된 다양한 비즈니스 분야에 새로운 형태의 친화적 서비스를 업무적으로 잘 활용할 것이라는 기대를 가지고 있다. 특히, 딥러닝과 같은 인공지능의 핵심 기술은 사용자가 손쉽게 원하는 지식을 대화하면서 원하는 학습 방향으로 유도하며 정보를 획득하고 의사소통할 수 있도록 되어있기 때문에 그 가능성을 비즈니스 업무의 특수한 영역인 보안 업무에 활용할 수 있도록 타당성을 검토하여 문제점을 도출해 내고자 한다. 이러한 역할의 딥러닝이 경쟁력과 충분한 가능성을 갖추고 있다면 이를 잘 훈련시키어 IT 업계 보안 시스템에 적용시킬 경우에는 기존 보안시스템에 문제점과 비용적인 측면을 고려해 볼 때에 충분한 경쟁력이 확보될 것으로 예상하고 있다[1].

본 논문의 구성은 다음과 같다. 1장 서론에서는 인공신경망(Artificial Neural Network) 기반의 기계 학습법의 기본적인 내용에 대해서 알아보고 2장 관련 연구에서는 딥러닝 작동방식 및 특징에 대해서 알아보고 국내외 전반적인 기술적 동향에 대해서 살펴본다. 3장에서는 딥러닝의 활용할 수 있는 기술적 특징에 대해서 알아보고, 4장에서는 딥러닝이 안고 있는 문제점을 검토한 후 학습 방향으로 유도하여 비즈니스 업무에 활용할 수 있도록 비인가자의 IP 및 세션정보에 대한 문제점을 분석하여, 5장에서 향후 기술적 기대 방향과 함께 결론으로 마무리한다.

2. 관련연구

딥러닝은 어떠한 문제 처리를 사람이 직접 지시하지 않아도 데이터를 통해 컴퓨터가 패턴 인식 문제 또는 특징적 학습을 하여 그것을 스스로 처리하고 해결할 수 있도록 하는 기계학습 기술이다. 이는 실제 인간의 뇌가 뉴런들 간의 연결이 매우 깊은(deep) 구조를 가지고 있다는 점에서 보다 진보된 학습과 추론에 대한 인공지능 기술이라 정의할 수 있다[2].

2.1 딥러닝 작동 방식

딥러닝의 학습 작동 방식 중 가장 중요한 부분은 분석 기술의 정보 추출 분야로 구조적 분류의 관점에서 어떻게 접근해야 하는지 이다. 이를 위해서는 해당 도메인에서의 적정량의 학습데이터와, 해당 데이터로부터 최적의 분류 함수를 얻기 위한 다음과 같은 두 가지 기계 학습(machine learning)이 필요하다.

① 첫 번째 기계 학습에 대한 부분은 경험적 역할로 기계 학습을 통해 스스로가 특정한 분야에 반복적으로 적응하면서 지적 영역과 경험적 영역을 넓혀 나가는 방식이다.
② 두 번째 기계 학습에 대한 부분은 규칙과 패턴을 익히는 역할로 기계 학습을 통해 문제의 규칙을 익히고 학습을 통해 필요한 과정을 습득하여, 같은 유형을 습득한 후, 비슷한 다른 문제의 영역까지도 해결 수 있는 강력한 응용력을 적응하는 방식이다.

2.2 딥러닝과 패턴

딥러닝 기술을 통해 훈련시켜야 하는 가중치들이 모인 곳을 은닉계층 (Hidden Layer)이라고 한다. 딥러닝의 최종 목적은, 테스트 데이터(Test Set)를 인공신경망에 통과시켜서 원하는 결과(Output)를 얻는 것이다. 딥러닝

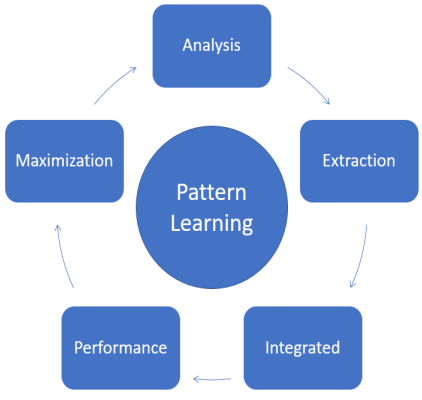

Fig. 1. Pattern Learning Process

에서는 은닉계층에 가중치를 훈련용 데이터(Training Set)를 이용하여 훈련(Training) 시키게 된다. 이 때 신경망 훈련을 통해 얻어진 데이터를 첫 번째에 레이어에 넣어서 데이터의 정확도를 산출해 낸다. 아울러 Fig. 1. 과 같이 이러한 과정을 반복으로 패턴화하여 학습과정을 규칙으로 습득하도록 한다[3].

① 첫 번째는 해당 분야의 전문가의 지식 없이도 데이터로부터 자동적으로 필요한 특징들을 추출해 낼 수 있는 강력한 분석 능력이다[4].
② 두 번째는 특징 추출과 분류기가 하나의 모델로 통합됨으로써 패턴인식의 성능이 극대화 되면 가중치를 바탕으로 얻어진 결과 값을 빠르게 다음 수행 작업 층으로 전달하는 능력이다[5].

2.3 딥러닝 보안 적용 가능성

기본적으로 일반적인 보안 업무 중 하나는 새로 다운로드 하거나 설치한 애플리케이션이 악성인지를 판단하는 것인데, 전통적인 접근방식은 기본 전문가 시스템으로 애플리케이션의 서명과 알려진 악성코드의 서명이 일치하는지 판단하게 된다. 딥러닝이 가지고 있는 인공지능의 학습적 응용 기술을 다양한 업무에 대한 보안 기능을 탑재하여 적용시킬 수 있는 가능성 타당여부를 살펴본다.

① 인공지능의 딥러닝 기술로 학습된 보안시스템은 약 1GB 크기로 대부분 애플리케이션에는 너무 크지만 팁 인스팅트는 이를 약 20MB 데이터양의 크기로 충분히 정리할 수 있다.
② 모바일을 포함하여 범용 엔드포인트 기기에나 설치할 수 있으며, 가장 느린 기기에서도 수 밀리초 만에 유입되는 위협을 빠르게 분석하여 침입에 대한 판단을 결정할 수 있다[6].
③ 딥러닝 학습을 통해 보안 영역에 적용시키기 위해서 인스팅트의 작동 방식이나 이미 잘 알려진 모든 악성코드 샘플을 딥러닝 시스템으로 학습하도록 제공한다. 딥러닝 프로세스가 학습된 악성코드를 찾고 분석하는 패턴을 익히게 되면, 이 프로세스는 같은 유형과 비슷한 유형의 악성코드를 훈련된 습관으로 빠르고 강력하게 처리를 수행하게 된다[7].

2.4 딥러닝 활용

현재 딥러닝에 대한 활용은 국내뿐만 아니라 해외에서도 다양한 비즈니스 측면에 널리 활용되고 있다. 특히 스타트업 기업에서는 딥러닝과 인공지능을 활용한 기술적 특성을 비즈니스 업무에 적용시키고 있어 기업의 빠른 성장을 돕는데 크게 기여하고 있는 게 사실이다. 딥러닝의 활용적 두각은 이미지 분류 뿐 아니라 음성 인식, 영상 인식, 자연어 처리, 빅데이터와 같이 데이터양이 많고 데이터 처리가 빠르고 정확한 업무적 성과가 필요한 부분에서 딥러닝의 활용은 성과적으로 매우 우수하다고 평가되고 있다.

딥러닝을 활용한 대표적 기업의 사례를 살펴본다.

① 국내 : 국내 대표적 2개의 IT 포털업체인 네이버(Naver)는 음성 검색에 딥러닝 알고리즘을 적용해 성능 향상을 경험했고 다음(Daum)도 뒤질세라 꾸준히 투자와 연구에 나서고 있다. 또한, 국내의 스타트업과 같은 신생 기업에서도 딥러닝 알고리즘을 잘 활용하고 있는 성공적인 사례로 비추어 볼 때 앞으로도 꾸준한 성장 가능성이 기대되고 있다[8].
② 해외 : 해외 글로벌 ICT 대표적인 업체인 구글, 마이크로소프트(MS), 페이스북, IBM, 바이두 등 IT 업체에서도 인공지능의 활용을 위해 꾸준한 기술적 투자를 아끼지 않고 있으며 딥러닝을 탑재한 비즈니스 업무적 역할에 점차 의존성이 높여가고 있다.

국내. 외 인공지능 기술이 적용된 서비스는 Table 1. 과 같다.

Table 1. Artificial Intelligence Service[9]

Company	Service
Google	Neural Machine Translation
Facebook	Deep Face
Microsoft	Cotana
CNN	Utilizing Extraction Techniques
Naver	Naver Cloud, Knowledge in

2.5 딥러닝이 보안 적용에 필요한 이유

기존 보안 관련 백신의 특징은 새로운 악성코드가 등장하면 감지가 어렵고 많은 피해를 경험한 후, 이를 토대로 대처하는 유형이 백신이었다. 뿐만 아니라 많은 기능적 수정과 새로운 악성 코드의 등장에 따라 지속적인 업

데이트가 필요하여 완전함 보다는 일시적인 대처 방법으로 안정보다는 불안정적인 요소가 더 많았다. 하지만 딥러닝을 활용한 보안 탑재는 학습에 사용할 수 있는 알려진 기존 악성코드 샘플이 현재 약 10억 개에 달하고 있다. 딥러닝의 가장 큰 장점인 실제 데이터를 보유하고 있는 대량의 악성 코드의 데이터를 통하여 악성코드의 유형을 분석하고 특징을 파악하여 침입 탐지와 공격에 대비할 수 있다는 패턴적인 습관을 빠르게 익히어 적용시킬 수 있다는 것이다[10].

3. 딥러닝 기술

딥러닝은 인공신경망의 한계를 극복하기 위해 제안된 기계학습의 특징 등을 2장에서 살펴보았다. 많이 알려진 의사결정나무, 베이지안망, 서포트벡터머신(SVM), 인공신경망과 같은 알고리즘은 데이터를 어떤 분야에 어떻게 활용될 것인지는 매우 중요하다. 3장에서는 딥러닝의 기술적 분류 역할로 딥러닝이 신경망을 활용하여 사물이나 데이터를 군집화하거나 분류(Classification)하는 데 사용되는 일종의 기술적 특징 및 학습에 대한 수행 역할도 함께 알아본다[11].

3.1 딥러닝 데이터 마이닝

딥러닝이 다양한 비즈니스 업무에 효율성 있게 이행하고 자료를 분석, 도출하기 위해서는 아래와 같은 4가지 수행 역할을 이행한다.

① 지각 : 개체를 시각적으로 인지하여 데이터를 수집하는 역할
② 추론 : 기억하고 있는 비슷하고 다양한 표현들을 사례로 삼아서 미래에 어떤 일이 일어날 것인지를 예측하는 역할
③ 학습 : 입력 데이터와 출력 데이터를 통해서 데이터 값에 의해 알려주는 대로 인식하여 정확도를 높이는 역할
④ 실행 : 지각과 추론에 대한 데이터를 기본으로 하여 복잡한 계획을 구조적으로 세워서 실행에 옮기는 역할

3.2 딥러닝 학습 모델

우선적으로 딥러닝이 학습을 수행하기 위한 습득 능력은 데이터의 속성, 형태에 따라 서로 다른 알고리즘이 적용된다. 하지만 대부분의 딥러닝 기반 발견학습은 데이터를 수집하고, 수집한 방대한 양의 데이터를 딥러닝을 통해 각 데이터의 관계를 설명하는 모델을 생성하고 이를 분석하고 추론하여 일반화에 적용하는 행동적 플로우 학습 모델이 Fig. 2.와 같이 적용된다[12].

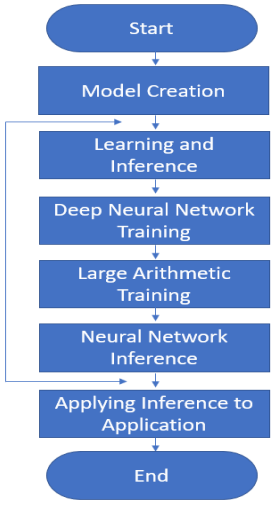

Fig. 2. Deep Learning Model

3.3 딥러닝 트레이닝

딥러닝 트레이닝(training)은 입력 데이터를 통해 모델을 학습하는 과정으로 기본적으로 피드포워드(feed forward) 과정과 백 프로파게이션(back propagation) 과정의 반복이다. 즉, 여러 은닉 계층을 거쳐 출력 계층까지 특징 값과 목적함수를 계산해 나가는 피드포워드 과정과 오류를 반영하여 출력 계층으로부터 은닉 계층을 거쳐 입력 계층까지 가중치를 수정하는 백 프로파게이션의 과정이 트레닝 과정으로 반복된다[13].

3.4 딥러닝 병렬처리

딥러닝 병렬처리는 GPU 컴퓨팅에도 일종의 컴파일러 역할을 수행하는 도구 역할을 한다. 일반적인 GPU는 계층적으로 구성된 수백 개의 연산 코어를 가지고 있으며, Fig. 3과 같이 CUDA 아키텍처 덕분에 GPU 가속화 애플리케이션에서 애플리케이션의 연산 집약적인 부분은 수천 개의 GPU 코어에서 동시에 실행하여 단일 스레드 성능에 최적화된 CPU에서 처리할 수 있게 된다[13].

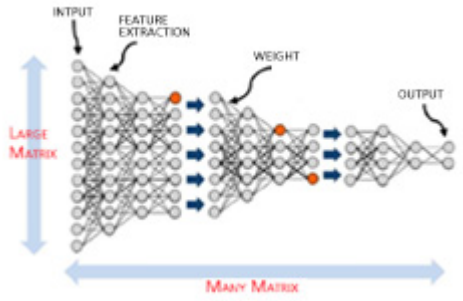

Fig. 3. Data Processing of Large Deep Learning Model[13]

3.5 딥러닝 추론 엔진(TensorRT)

딥러닝 추론 엔진(TensorRT)은 TensorRT는 딥러닝 애플리케이션의 배포를 위한 고성능 뉴럴 네트워크 추론 엔진이므로 주로 하이퍼스케일 데이터센터, 임베디드, 자율주행 플랫폼의 추론에 쓰이는 훈련된 뉴럴 네트워크를 빠르게 최적화하고, 검증, 배포하는데 사용된다[14].

3.6 딥스트림 SDK(DeepStream SDK)

딥스트림은 딥러닝의 GPU 가속 트랜스코딩(Transcoding) 과 딥너링 추론(Inference) 기능을 빠르게 통합해서 보다 반응이 빠른 AI 기반 서비스를 제공할 수 있도록 한다[14].

3.7 cuDNN 가속화 프레임워크

cuDNN 가속화 프레임워크는 cuDNN의 최적화된 루틴은 딥러닝 개발자들이 뉴럴 네트워크 모델의 디자인과 훈련에 집중할 수 있도록 돕는다. cuDNN은 TensorFlow, Theano 및 PyTorch를 포함해서 대중적으로 널리 사용되는 딥러닝 프레임워크를 집중적으로 가속화한다[14].

3.8 비동기식 방식 처리

파라미터 서버의 경우 H2O, DeepSpark과 같은 빅데이터 처리 계열 프레임워크와 Petuum, CNTK, MXNet과 같은 기계학습/딥러닝 전용 프레임워크들은 키-밸류 저장소 형태의 파라미터 서버를 지원하고 있으며, H2O와 DeepSpark의 경우 비동기 방식을 도입하여 속도를 개선한다[15].

4. 딥러닝 보안 기술 문제점 분석

4장에서는 딥러닝 기술이 신경망 기술을 적용해 의사결정을 내릴 수 있고 다양한 비즈니스 모델에 보안 업무를 적용시키고 활용하기 위해서는 우선적으로 딥러닝 기술이 안고 있는 보안 문제점이 무엇인지를 분석해야 한다. 최근 딥러닝은 인터넷과 연결된 다양한 환경에서의 취약한 인터넷 환경에서의 비정상 행위를 탐지하고 분석하여 위협적인 공격으로부터 공격을 차단하는 학습을 스스로 익히어 공격에 대한 인지패턴을 만들 수 있도록 발전하였다. 즉, 인공지능을 기반으로 하는 새로운 분야의 딥러닝과 같은 기술은 이미 선진국의 구글, MS, 아마존과 같은 기업에서 클라우드 기반과 빅데이터 기반으로 이를 접목한 연구와 적용이 활발히 진행되고 있다. 앞으로 딥러닝을 활용한 비즈니스 영역의 보안 서비스는 새로운 경쟁력으로 4차 산업의 핵심에서 여러 비즈니스 모델에 활용될 것으로 전망하고 있다. 본 논문에서는 지금까지 살펴본 딥러닝 학습 프로세스 모델이 보안이라는 특수한 임무를 제대로 수행하고 효과적으로 실효를 얻기 위해서는 기존 자료의 정확한 데이터 분석과 보안의 문제점이 무엇이었나에 대한 부분과 보안의 수행 역할에 있어서 오류 탐지율에 대한 분석이 제대로 이루어졌는지에 대한 부분이다. 또한 비즈니스 모델에 딥러닝 기술의 보안 알고리즘을 좀 더 효과적으로 적용시키기 위해서는 기존 자료의 정확한 데이터 제공과 분석, 그리고 적용할 데이터의 양이 방대할수록 정확한 예측 대안을 찾을 수 있는지에 대한 부분이 중요한 핵심 요소가 될 것이다.

4.1 딥러닝 보안 탐지 모델

딥러닝은 관찰 가능한 모든 위협 환경을 학습할 수 있는 확장형 탐지 모델을 제공하게 된다. 또한 기존에 처리되어 수집된 수억 개의 샘플데이터를 활용하여 학습 데이터로 사용할 수 있어 기존 전통적인 머신러닝이 가지고 있는 오류 탐지율과 속도가 느린 것에 비하면 빠른 속도로 더욱 정확한 예측이 가능한 경쟁력을 가질 수 있다. 물론, 제공되고 있는 기존 데이터가 얼마만큼의 신뢰하고 정확한 제공한 데이터가 딥러닝 보안 모델에 제공되는가에 따라 위험 탐지에 대한 정확성이 고려되고 신뢰할 수 있는 규칙과 패턴을 정할 수 있는 품질의 척도가 달라지는 변수가 있음을 고려해야 한다. 아래 Table 2.는 기존 보안 솔루션의 문제점을 도출하여 표로 제시하였다.

Table 2. Problems of Existing Security Model

Classifi-cation	Problem
Problem Analysis	• As the amount of data increases, computational complexity becomes complicated and computation speed decreases • Security detection ability drops much when the model size is at the gigabit level – False alarm rate appears – Lack of malicious code and discernment of legitimate software – IT Productivity and efficiency are poor • If there is insufficient amount of training data sample that are used for algorithm learning or if implementation is invalid, it could led to bad result.

4.2 다양한 비즈니스 업무 보안 검토

딥러닝이 인터넷과 연결된 다양한 각종 비즈니스 업무와 관련된 서비스 영역에서 보안 업무를 충분히 수행하기 위해서는 많은 데이터를 가지고 반복적인 학습이 필요하다. 특히 딥러닝이 보안 업무를 수행할 수 있는 학습적 능력을 얻기 위해서는 비정상 IP패킷에 대한 탐지와 정상적인 소프트웨어와 악성코드를 탑재하여 감염의 의도를 가지고 접근하는 공격자가 의도가 확실한지에 따른 빠르고 정확한 판단이 가능해야 한다. 이러한 보안 수행 업무를 완벽히 해내기 위해서는 반복 학습을 통해 정확하고 올바른 판단과 결론을 도출해내야 한다. 그 과정을 분석하고 학습하여 습관적인 패턴을 수행하기 위해서는 기존 적용된 보안 문제점이 무엇인가를 우선적으로 파악할 수 있는 근거 데이터를 제공해야 하는데 알고리즘의 학습에 사용되는 트레이닝 데이터 샘플에 대한 정확성이 떨어지거나 미흡할 경우, 또는 구현이 잘못된 경우 그 적용성이 떨어지게 된다. 또한, 머신러닝 알고리즘 설계가 미흡할 경우, 그 결과를 유용하게 활용할 수 없다. 즉, 다양한 비즈니스 환경에서 비정상 탐지에 대한 추론을 스스로 익힐 수 있도록 충분한 실험적 데이터 접근을 통한 학습 환경이 제공되어야만 한다.

4.2.1 비정상 탐지 학습

딥러닝이 다양한 인터넷 환경에서의 비정상 접근 및 악성코드에 대한 비정상 소프트웨어를 구분해 내기 위해서는 시스템에 접근하기 위한 다양한 IP를 분석하고 대처하기 위한 무수한 학습적 검토가 필요하다. 딥러닝 알고리즘은 비정상 탐지 및 악성 행위를 더 빨리 탐지하고, 공격 시작 전에 이를 저지할 수 있도록 훈련을 통해 스스로 익힐 수 있도록 충분한 능력을 갖추게 된다. 그렇다면 딥러닝 기술이 공격자에 대한 비정상 IP를 어떻게 탐지하고 공격이 의심되는 패킷을 어떻게 구분하여 추론할 수 있는지에 대한 올바른 학습 이행 여부를 알아본다.

① 특정 시스템 내 네트워크로 진입하는 비인가 침입자를 탐지해 낼 수 있는가?
② 공격자로부터 데이터 유출 공격을 인지하여 탐지할 수 있는가?
③ 랜섬웨어와 같은 악성코드를 포함한 의도된 공격자의 시스템 접근 시 이를 감지하고 이와 유사한 공격도 탐지할 수 있는가?
④ 침투 목적을 가진 위협적인 분산서비스 공격으로부터 조기에 이를 감지하고 예방하여 공격으로 벗어날 수 있는 추론 능력을 보유하고 있는가?
⑤ 로그인 정보가 5회 이상 잘못 입력된 경우 비정상 패킷을 분석해 내어 이를 관리자에게 통지할 수 있는가?

위 5가지 중 하나만이라도 비정상 추론 동작을 이행하지 못하거나 오류율이 보일 경우, 학습을 통해 문제의 원인을 파악하여 오류율을 재 탐지하고 정상 동작할 수 있도록 학습 추론, 반복 연산, 즉, 신경망 학습에 대한 규칙 패턴 인식에 대한 수정 보강이 절대적으로 필요하다. 지금까지 정상으로 이용되고 있는 세션 및 로그 분석 패턴 트래픽 정보를 토대로 머신러닝이 5회 이상 로그인 정보가 오류 및 비정상 탐지를 인지할 때에는 일정 시간 동안 시스템 접근을 못하도록 사용을 중지하거나 비밀번호 변경에 대한 요청을 하여 이메일로 관련 사항을 보내는 경우가 대부분이었고, 비정상 공격 탐지로 분류 시키지는 못하였다. 하지만 이러한 부분에서 공격 탐지로 시도되고 있는 비정상 데이터 트래픽 탐지 IP가 기존에서 사용된 부분의 오류와 엄격히 구분하여 새로운 위치의 IP와 기존에 사용된 IP 위치를 분석하여 문제가 될 수 있는 IP 부분에 대해서는 비정상 탐지로 구분해 낼 수 있도록 학습적 능력을 새롭게 부여하도록 해야만 한다.

4.2.2 위협 분석 능력

그렇다면 하루에도 수백만 개의 로그인이 접속되고 있는 시스템에 비정상 탐지에 대한 위협을 어떻게 완벽히 탐지하여 분석해낼 수 있을까? 이 또한 시스템 로그인을 분석하여 기존의 시스템을 이용하는 사용자의 로그인 정보를 매일매일 업데이트한 후, 확인하여 로그인 정보를 분류한 후, 시스템에 정상적으로 접근하는 인가자 정보인지 또는 시스템 침입을 목적으로 접근하는 비인가자의 침투 목적을 분석 구분하여 데이터를 추출해 낼 수 있어야만 한다. 대부분 비인가자의 경우 시스템에 침투하는 목적이 악성 코드를 포함하여 시스템에 접근하여 비즈니

스 업무를 방해하거나 데이터 유출을 목적으로 접근하는 공격자가 대부분이다. 딥러닝 기술의 업무적 보안 능력은 바로 허가되지 않은 IP와 비인가자 IP를 빠르게 분석하고 인지하여 사전에 비인가자의 공격 의도를 탐지하고, 네트워크와 엔드포인트 보안 상태를 분석하여 위협으로부터 보호하도록 해야 한다. 그럼 IP 탐지에 대한 오류를 줄이고 인가자와 비인가자를 구별하는 정확도를 높이기 위해서는 위협 분석 테스트를 수행하여 비인가자의 공격 위협으로부터 공격 탐지를 90% 이상 탐지할 수 있도록 하는 것이다. 그래야만이 딥러닝의 비인가자의 공격적 의도를 파악할 수 있는 1차 분석 능력을 갖추었다고 능력을 인정할 수 있다. 그렇지 않을 경우 반복적인 학습 추론을 통해 트레이닝 훈련 보강을 수정하여 탐지에 대한 정확도를 높이고 해킹 위협을 탐지할 수 있는 위협 분석 능력 학습 추론을 강화해야만 한다.

4.3 서버 운영에 따른 취약점 인지

다양한 비즈니스 모델 시스템 운용 시 네트워크, 시스템 서버 등에 다양한 위협이 존재할 수 있다. 인공지능의 보안 기능이 적용된 딥러닝은 최신 유행하고 있는 제로데이 공격 위협, 익스플로잇, 랜섬웨어 공격, 안전하지 않은 IoT 장치를 표적으로 삼는 공격을 중심으로부터 취약점을 사전에 인지하여 시스템 운용자, 서버 관리자에게 문제가 될 수 있는 부분을 통지하여 위협을 보안, 강화할 수 있는 수행 업무가 마련되어야 한다. 현재 많은 기업들이 상업용 보안 솔루션을 설치하여 자사 보안 시스템 의 침투를 예방하고 시스템 상의 고객 정보 데이터의 유출을 막는데 신경을 쓰고 있다. 이에 딥러닝 기술은 취약점의 존재를 파악하고 이를 꾸준히 규칙적인 시스템 운용에 따른 학습을 통해 트래픽 모니터링에 기계학습을 통해, 제로데이와 익스플로잇과 같은 시스템의 취약한 부분을 노리는 공격을 인식할 수 있는지에 대한 학습 능력 파악이 우선적으로 되어야만 한다. 대부분의 시스템은 성능적 개선과 해킹에 대한 데이터 유출 방지, 사전 보안사고와 유사한 공격 사례를 통해 보안을 강화시키어 대비하고는 있다. 하지만 공격자는 시스템의 허점과 보안 시스템의 허점을 노리고 있으므로 딥러닝 학습 추론의 다양한 시뮬레이션 유추를 통해 보안 능력을 강화하고 시스템 오류와 보안 탐지율이 저조한 시간대를 분석하여 서버 시스템을 강화토록 해야 한다. 하지만 보안 시스템의 취약한 부분과 소프트웨어의 특정한 부분에 오류를 파악하기란 쉽지 않다. 하지만 정상적인 데이터에 오류 코드 삽입 및 사용 시간대가 저조한 심야, 대량 데이터의 연산 및 추출, 로컬 IP, 해외 IP 접근, 의심되는 멀웨어 악성 코드 삽입 후 시스템 접근 등 위협을 통한 다양한 방식으로 시스템에 접근하고 이에 대한 유형을 분석하고 대처할 수 있는 패턴 학습을 익혀 나가게 된다면 시스템 운용에 따른 보안 사고는 현저히 줄어들 것으로 예측한다. 또한, 시뮬레이션을 통해 시스템 운용에 따른 위협적 취약한 부분이 발견 된다면 관리자에게 메시지로 통지하고, 취약한 부분을 제거하고, 보안 수행 기능을 더욱 강화시키는 사고 대응 인지 능력도 필요하다고 본다.

Table 3. Improvement of Existing Security Solution

Classifi-cation	Solution to the Problem
Suggested Solution	• Modify training drills for repetitive learning for accuracy and detection rate, when attack detection rate is below than 90% • To prepare for the recognition of attacks which targets vulnerable parts of systems such as zero-day and exploits. - Insert error code - low time period of usage such as midnight - Computation and extraction of large data - Access to local IP and overseas IP - System access after insertion of suspicious malware

V. 결론

4차 산업 시대가 도래된 현재, 딥러닝 기술 등장으로 인해 다양한 비즈니스 영역에 그 활용 가치가 더 광범위해졌으며 보안의 중요성과 함께 그 능력의 기대 또한 커지게 되었다. 본 논문을 통해 인공지능이 가지고 있는 추론 능력을 보안에 적용하여 기존 프로그래밍 방식을 벗어나서 스스로가 비인가 된 접근자의 위협과 보안 취약점, 침입 의도를 가진 네트워크 IP 패킷을 딥러닝 기술 스스로가 분석하여 구별해 낼 수 있는 추론 능력을 보유하여 보안 업무를 수행할 수 있도록 하는 것이 핵심적 과제였다.

본 논문에서 딥러닝 기술의 보안 문제점을 통해 다양한 비즈니스 모델에서 적용 가능도록 IP 탐지 능력 및 시스템의 취약점을 인지하거나 시스템 소프트웨어의 오류 등의 문제점을 찾아내 이 후 진행되는 프로세스에 대처할 수 있는 능력을 보유하고 있는지에 대한 다양한 사례 검토를 통해 기술적 문제점에 대한 구체적 방안을 제시하였다.

끝으로, 우리가 기대하는 딥러닝 기술은 앞으로 컴퓨터 시스템을 통해 여러 다양한 학습 가능 기대 효과를 넘어서 금융, 경영, 교육, 의료, IT 산업 분야 등 어느 비즈니스 분야를 총망라하여 활용할 수 있는 범위가 무궁무진하리고 예측되고 있다. 현재 적용 가능한 모든 산업 분야에 딥러닝 기술이 그 진가를 발휘하여 실효성을 거두기 위해서는 더욱더 정확하고 분석된 대량의 데이터를 도출하여 이를 적용할 수 있는 보안 관련 학습 모듈의 체계 구축과 대용량 범용 비즈니스 모델 기반의 보안 솔루션을 구축하는 것이 과제이다. 인공지능에 탑재된 딥러닝 기술은 잠재력 있는 기술력으로 재평가 되고 있는 만큼 앞으로도 끊임없는 노력과 연구가 절실히 필요하다.

REFERENCES

[1] M. R. Choi. (2017), Artifical Intelligence Technology based on Natural Language Processing, *Telecommunications Technology Association, (36),* 33-37

[2] S. E. Moon, S. B. Jang, J. H. Lee & J. S. Lee. (2016), *Machine Learning and Deep Learning Technology Trends,* Information and Telecommunication, 49-56

[3] Y. H. Shin, J. S. Yun, S. H. Seo & J. M. Chung. (2017). Deployment of Network Resources for Enhancement of Disaster Response Capabilities with Deep Learning and Augmented Reality, *JICS, 18(5),* 69-77 DOI 10.7472/jksii.2017.18.5.69.

[4] J. S. Yun, K. Y. Kim, Y. C. Jung, H. S. Oh & D. J. Seo. (2017). Development of Deep Learning Technologies and Applications for the Information Extraction of S&T Open Texts, *Korea Institute of Science and Technology Information*, 1711042891, 1-43

[5] H. Y. Choi & Y. H. Min. (2015), Intorduction to Deep Learning and Major Issues, *Korea Information Proccessing Society, 22(1),* 7-21

[6] K. Maria. (2016. 11. 10). *Security Solution with Machine Learning.* CIO Center Article, http://www.ciokorea.com/t/21990/%EC%95%85%EC%84%B1%EC%BD%94%EB%93%9C/31931#csidx734eda637286190876b5fbcd63a16cd

[7] K. Maria. (2016. 11. 10). *Security Solution with Machine Learning.* CIO Center Article, http://www.ciokorea.com/t/21990/%EC%95%85%EC%84%B1%EC%BD%94%EB%93%9C/31931#csidx68a1b7c636c6fd6b63e1e9482b31a7d

[8] P. S. Kang & J. H. Kim. (2014. 7. 31). What is Deeep Learning, http://www.bloter.net/archives/201445

[9] J. Y. Kim and T. W. Lee, (2016). A Study on the Development of Smart Education Using Deep Learning Algorithm, *Korea Computer Information Association, 24(2),* 169-171

[10] K. Maria. (2016. 11. 10). CIO Center Article, http://www.ciokorea.com/t/21990/%EC%95%85%EC%84%B1%EC%BD%94%EB%93%9C/31931#csidxbd14f7f63ec4ddfa75a2c452de87a2e

[11] P. S. Kang & J. H. Kim. (2014. 7. 31). What is Deeep Learning, http://www.bloter.net/archives/201445

[12] J. W. Kim, H. A. Pyo, J. W. Ha, C. K. Lee & J. H. Kim. (2015). *Utilizing various Deep Learning Algorithms,* Electronics and Telecommunications Research Institute, 25-31

[13] S. Y. Ahn & Y. M. Park & E. J. Lim & W. Choi. (2017), Trends on Distributed Frameworks for Deep Learning, *Electronics and Telecommunications Trends, 31(3),* 131-141

[14] NVIDIA KOREA Center. (2018. 1. 16). *Computered Unified Deviced Architecture.http://blogs.nvidia.co.kr/*2018/01/16/cuda-toolkit/

[15] ETRI. (2016). Trends on Distributed Frameworks for Deep Learning, ETRI Article, https://ettrends.etri.re.kr/ettrends/159/0905002137/0905002137.html

최 희 식(Choi, Hee Sik) [정회원]

· 2006년 2월 : 숭실대학교 컴퓨터공학과(공학석사)
· 2012년 2월 : 숭실대학교 컴퓨터학과(공학박사)
· 2008년 3월 ~ 현재 : 삼육대학교 컴터학부 외래교수
· 관심분야 : 정보보안, 클라우딩컴퓨터, 핀테크, 금융보안
· E-Mail : dali3054@ssu.ac.kr

조 양 현(Cho, Yang-Hyun) [정회원]

· 1982년 2월 : 광운대학교 전자통신공학과(공학사)
· 1985년 2월 : 광운대학교 전자통신공학과(공학석사)
· 2012년 2월 : 광운대학교 전자통신공학과(공학박사)
· 1987년 9월 ~ 1997년 8월 : LG정보통신 전송기술개발실 과장
· 1997년 9월 ~ 현재 : 삼육대학교 컴퓨터•메카트로닉스공학부 교수
· 2014년 3월 ~ 2016년 2월 : 삼육대학교 산학협력단장/연구처장
· 관심분야 : 컴퓨터네트워크, 통신망(BcN), GMPLS, IoT
· E-Mail : yhcho@syu.ac.kr