

이것만은 알아야 할 네트워크 보안 이야기

Part 10. APT 3. Zero Trust 보안모델, AI(인공지능)를 이용한 공격 탐지

글. 이선웅 아이크래프트 수석
ka3211a@gmail.com

이번 호에서는 기존의 보안모델로는 급증하는 APT 공격에 대응하기에는 한계점이 드러나 새롭게 주목받는 Zero Trust 보안모델에 대해 소개하고, 인공지능(AI) 기술을 공격 탐지에 활용할 수 있는 방안에 대해 알아보도록 하겠다.

연재 목차

- 1회. 방화벽 1 _ 방화벽의 역할, 보안 정책
- 2회. 방화벽 2 _ 차세대 방화벽, 웹 방화벽의 등장
- 3회. 방화벽 3 _ 가상화 기술을 통한 방화벽의 진화
- 4회. VPN _ 암호와 검증과 인증의 결정판
- 5회. DDoS 1 _ DDoS 공격의 방식과 유형, DDoS 방어 장비
- 6회. DDoS 2 _ DDoS 공격의 탐지 방안
- 7회. DDoS 3 _ DDoS 공격의 차단 방안
- 8회. APT 1 _ APT 공격의 방식과 사례
- 9회. APT 2 _ APT 공격 그룹과 악성코드(Malware)
- 10회. APT 3 _ Zero Trust 보안 모델, AI를 이용한 공격 탐지
- 11회. APT 4 _ APT 공격 가상시나리오, APT 공격 방어 장비

APT 공격이 등장하기 전에는 외부에서 내부로의 공격방식이 비교적 단순하였다. 공격자는 외부에 공개되어 있는 서비스를 검색해서 접속 가능한 서비스 포트를 확인하고 다양한 방식의 공격을 시도해 보는 방식이 일반적이었다. 즉 공격대상이 공개 서비스를 제공하는 서버로 한정되어 있었다. 그래서 보안관리자의 입장에서는 외부에서 공개된 서버로 연결되는 네트워크 구간만 모니터링하면 대부분의 공격 탐지가 가능하였다. 그러나 최근의 APT 공격의 주요 대상이 서버가 아니라 사용자가 이용하는 PC, 스마트폰으로 이동하였다. 공개된 서버는 전통적으로 보안에 대한 대비가 잘 되어 있는 반면, 개인이 사용하는 PC나 스마트폰의 경우 서버에 비해 공격에 대한 대비가 소홀할 수밖에 없다. 더욱이 서버는 나름대로 훈련된 엔지니어가 관리하는 반면, 개인 단말의 경우 IT 기술이 상대적으로 낮고, 보안에 대한 의식이 부족한 일반 사용자가 이용하기에 공격대상으로 더욱 주목받을 수밖에 없게 되었다.

이러한 개인 단말이 침투하기 쉬운 공격대상으로 인식되면서, 일반사용자를 대상으로 하는 다양한 공격수단이 개발되었다. 이런 개인을 대상으로 하는 공격에는 사회공학(social engineering)이라고 기법이 활용되고 있다. 사회공학이란 기술적인 방법이 아니라 사람 간의 기본적인 신뢰를 기반으로 사람을 속여 비밀정보를 획득하는 기법으로, 인간 상호 간의 기본적인 신뢰를 바탕으로 사람들을 속여 정상적인 보안 수단을 우회하여 공격에 활용되는 비기술적인 침입 수단이라고 할 수 있다. 한마디로 사람을 속이는 사기 수법이라는 말이다.

“기업 정보보안에 있어서 가장 큰 위협은 컴퓨터 바이러스, 패치가 적용되지 않은 중요한 프로그램이나 잘못 설정된 방화벽이 아니다. 가장 큰 위협은 바로 당신이다.”

전설적인 해커인 케빈 미트닉은 이렇게 언급했다. 이 말은 아무리 기술적으로 완벽한 보안 수단이 있다고 하더라도 이를 이용하는 개인은 사회공학적인 기법으로 쉽게 속을 수 있어 이를 통한 공격이 가장 큰 위험 요소라는 말이다. 사실 기술적인 보안을 완벽히 하더라도, 물리적이거나 인간적인 부분까지 완벽히 막기는 어렵다. 아무리 서버에 고도로 암호화되어 있는 중요 데이터라고 해도 담당자를 속여서 직원의 계정과 암호를 빼돌리면 막을 방법이 없기 때문이다.



그림 1. 사회공학적 공격기법

APT 공격을 통해 단말에 침투하는 수단 중 가장 대표적인 워터링홀(Watering hole)과 스피어피싱(Spear Phishing)이 바로 사회공학을 이용하는 공격기법이다. 이메일로 오는 ‘무료쿠폰 제공, 돌잔치 초대장, 택배 조회, 할인이벤트 정보’와 같이 사람들이 기본적으로 관심을 가질 만한 내용이나 ‘각종 보고서, 납품 대금 계좌변경, 품의서, 기안서 등 각종 업무 관련 내용’으로 위장해서 클릭이나 첨부파일을 열어보게 유도하는 것이 가장 전형적인 단말 침투 수단이다. 아무리 직원에 대한 보안 교육을 한다고 하더라도 100명 중 1명만 클릭을 할 수 있게 한다면 공격에 성공하는 것이다. 이미 공격대상자의 이메일을 알고 있는 상황에서 100명에게 한번 보내서 성공하지 않으면 두 번, 세 번 반복해서 보내 100명 중 누구든지 1명만 클릭한다면 침투에 성공할 수 있는, 성공률이 꽤 높은 공격수단이라고 할 수 있다.

이렇게 공격대상이 개인 단말이 되고 공격 성공률이 높아 지면서, 기존의 인터넷과 내부 네트워크가 만나는 구간에서만 집중적으로 공격을 탐지하는 방식인 경계 보안은 무용지물이 되었다. 이메일을 보내는 사람을 임의로 차단할 수도 없고, 첨부파일에 알려지지 않은 악성코드가 포함될 경우 탐지하기가 쉽지 않기 때문이다. ‘울타리만 집중적으로 방어하는 방식은 공격을 100% 차단할 수 있다’라는 믿음이 있기 때문에 가능했지만, 개인 단말을 통한 공격이 쉽게 성공하는 상황에서는 새로운 보안모델이 필요하게 되었다.

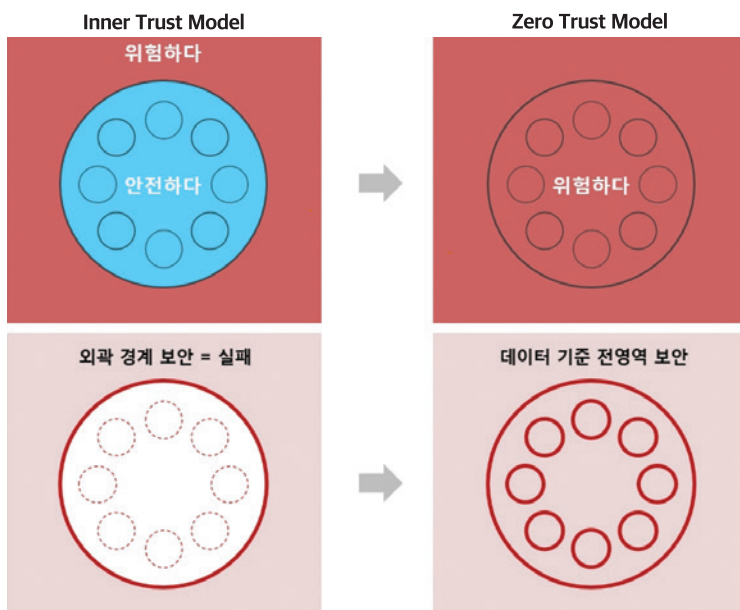


그림 2. Zero Trust 보안모델 / 출처 : 펜타시큐리티

그래서 최근 주목받는 모델이 바로 ‘Zero Trust’라는 보안모델이다. 즉, [그림 2]와 같이 ‘신뢰할 수 있는 것은 아무것도 없다’라는 전제에서 보안 대책을 수립해야 한다는 사상이다. 공격을 100% 차단하는 것은 불가능하니 공격을 당했다는 전제에서 보안 대책을 수립해야 한다는 것이다. 시스템 외부와 내부를 따로 나누지 않고 모든 곳이 위험하니 적절한 인증 절차 없이는 그 누구도 믿어서는 안 되며, 누구든 중요 데이터에 접근하려면 권한을 부여받기 전에 재차 신원을 확인해야 한다는 것을 전제로 가져가는 모델이다. 더욱이 클라우드, 모바일 사용자의 증가로 인해 내/외부망을 구분하기가 모호해지는 상황에서 이러한 보안모델이 현재 상황과도 일치하는 사상이라고 할 수 있다.

최근의 코로나19 사태에 대응하는 방식도 이런 보안모델을 동일하게 적용할 수 있다. 2020년 2~3월의 초기에는 중국에서 발원한 코로나를 국경에서만 막으면 된다고 생각하고 공항과 항만 등을 폐쇄한 사우디아라비아, 이스라엘의 확진자가 늘어나는 것도 같은 맥락이라고 할 수 있다. 아무리 국경에서 철저하게 확진자를 차단한다고 하더라도 방역 당국이 미처 파악할 수 없는 통로는 존재할 수밖에 없기 때문에 100% 차단은 불가능한 것이다. 국경 폐쇄도 중요한 방식이지만, 100% 차단이 불가능하다는 전제하에서 내부에서 증상이 발생하는 환자에 대해 추적, 테스트, 치료하는 적극적인 대응에 집중한 대한민국의 방식이 좀 더 효율적인 방역 대책이었던 것이다. 코로나19 방역에서도 확인된 것처럼 APT 공격에 대한 대응도 외부경계뿐만 아니라 내부에 존재하는 서버와 단말 등에 대해 좀 더 집중해서 모니터링하는 방식이 APT 방어에 대한 효과적인 모델이라고 할 수 있다.

이러한 내부 보안에 집중하기 위해서는 무엇보다 개인용 PC나 스마트폰과 같은 단말에서 공격을 탐지하고 차단하기 위한 노력이 필요한데, 이를 통상적으로 엔드포인트(EndPoint) 보안을 강화한다고 말한다. 기존에 사용되고 있는 바이러스 백신(Anti-Virus)의 경우 기존에 알려져 있는 바이러스, 악성코드의 패턴을 가지고 있다가 단말로 침투할 경우 탐지하여 차단하거나 이미 단말 내부에 존재하는 바이러스 등을 탐지하여 삭제하거나 격리시키는 역할을 수행하였다. 그러나 악성코드를 손쉽게 수정하여 변종을 만들기가 간편해지면서 더 이상 기존에 알려진 악성코드의 패턴을 통해서 이러한 변종 악성코드의 탐지가 불가능하게 되었다. 즉 알려지지 않은 악성코드에 대한 탐지기능이 필요하게 되었고 동적 분석을 통해 악성코드를 가상 PC에서 직접 다운로드하여 동작하는 것을 모니터링하여 악성코드 여부를 판단하는 방법이 개발되었다. 또한 악성코드로 감염시킨 개인용 PC를 발판으로 하여 다양한 공격행위가 탐지되었을 때, 이 감염 PC를 조사하여 공격행위에 대한 조사를 통해 공격이 지속된 시간과 사내 다른 PC에 대한 공격 범위, 피해 정도, 공격그룹의 추적하기 위한 단서 등의 자료를 확보하는 기능도 필요하게 되었다. 그래서 기존의 백신(Anti-Virus) 기능에 알려지지 않은 변종 악성코드 탐지능력과 침해사고조사를 위한 포렌식 기능 등을 추가하여 엔드포인트 위협탐지대응(EDR : Endpoint Detection & Response) 솔루션으로 기존 백신이 업그레이드되거나 신규벤더 솔루션이 시장에 많이 출시되었다. 아래 [그림 3]은 EDR 솔루션이 순차적으로 알려진 공격과 알려지지 않은 공격행위에 대해 어떤 단계를 통해 공격을 탐지하는지 설명한 그림이다.

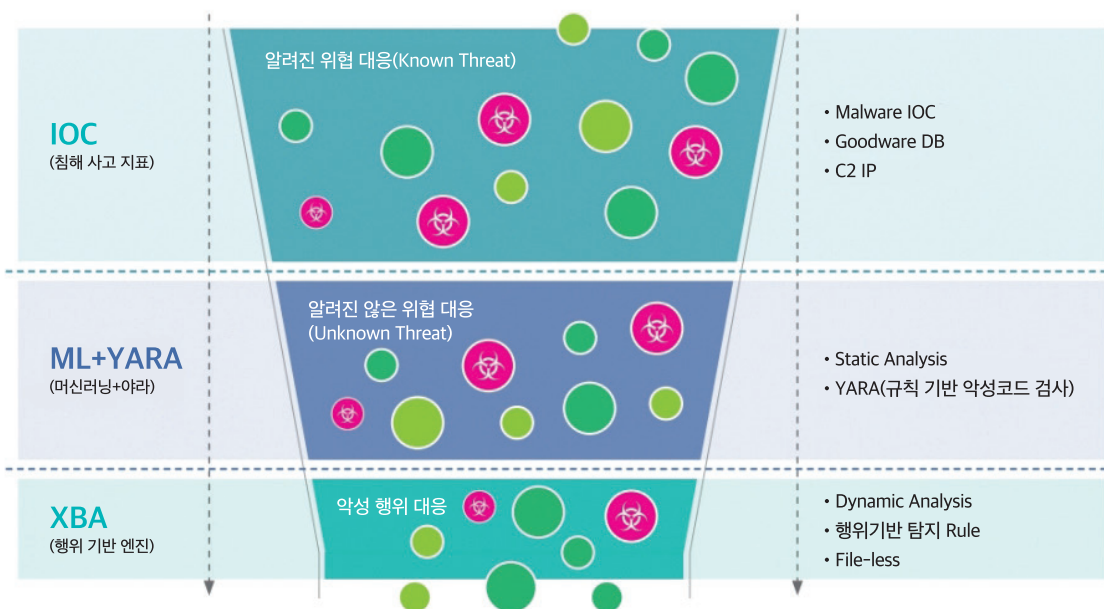


그림 3. EDR(Endpoint Detection & Response) 솔루션 동작 방식 / 출처 : 지니언스

최근 빅데이터를 기반으로 하여 인공지능에 대한 관심이 높아지고 있다. 인공지능을 이용하여 코로나 사태에서 중요한 의료분야에 적용되어 감염진단, 진단시약개발, 치료제, 백신 후보물질 개발 등에도 활용되고 있다고 한다. 정보보안 분야에도 당연히 인공지능을 활용하기 위해 많은 연구와 시도가 이루어지고 있다. 정보보안 분야에서 가장 중요한 이슈는 얼마나 정확하게 공격을 탐지할 수 있는지에 대한 공격 탐지정확도를 최대한 100%에 가깝게 유지하는 것이다. 지금까지는 보안담당자의 배경지식과 경험을 바탕으로 공격 탐지 조건을 설정하고, 시행착오를 반복해서 적용해 보며 조직의 환경에 최적화된 조건을 찾는 것이 일반적인 방법이었다. 그러나 이런 방식은 공격 탐지의 정확도를 높은 수준으로 유지할 수 없을 뿐만 아니라, 새로운 공격방식이 등장하면 정확도가 급격히 떨어지는 문제점이 존재하였다. 그리고 가장 큰 문제는 경험이 풍부한 보안 담당자의 높은 급여 수준과 인력 부족으로 채용하기가 쉽지 않다는 점이다. 이런 문제점을 해결하기 위해 공격을 높은 정확도로 탐지하는 것을 목표로 인공지능을 적용하기 시작하였다.

우선 인공지능에 대해서 간략하게 알아보자. 아래 [그림 4]와 같이 인공지능의 개념은 1950년대부터 시작되었다. 1980년대부터는 컴퓨터를 활용한 머신러닝이라는 용어가 등장하였고, 2010년 사람과 동물의 뇌를 모방하여 딥러닝이라는 용어가 등장하면서, 향상된 처리성과 디지털화를 통해 축적된 많은 학습데이터의 활용이 가능해졌고, 비약적으로 발전하기 시작하였다. 최근에는 고성능의 딥러닝머신을 구매하지 않더라도 클라우드 서비스를 통해 간편하게 컴퓨팅능력을 일정 시간 빌려서 이용할 수 있고, 공개된 공공 빅데이터를 활용하여 다양한 테스트가 가능해지면서 사회 여러 분야에 급속도로 인공지능이 적용되기 시작하고 있다.

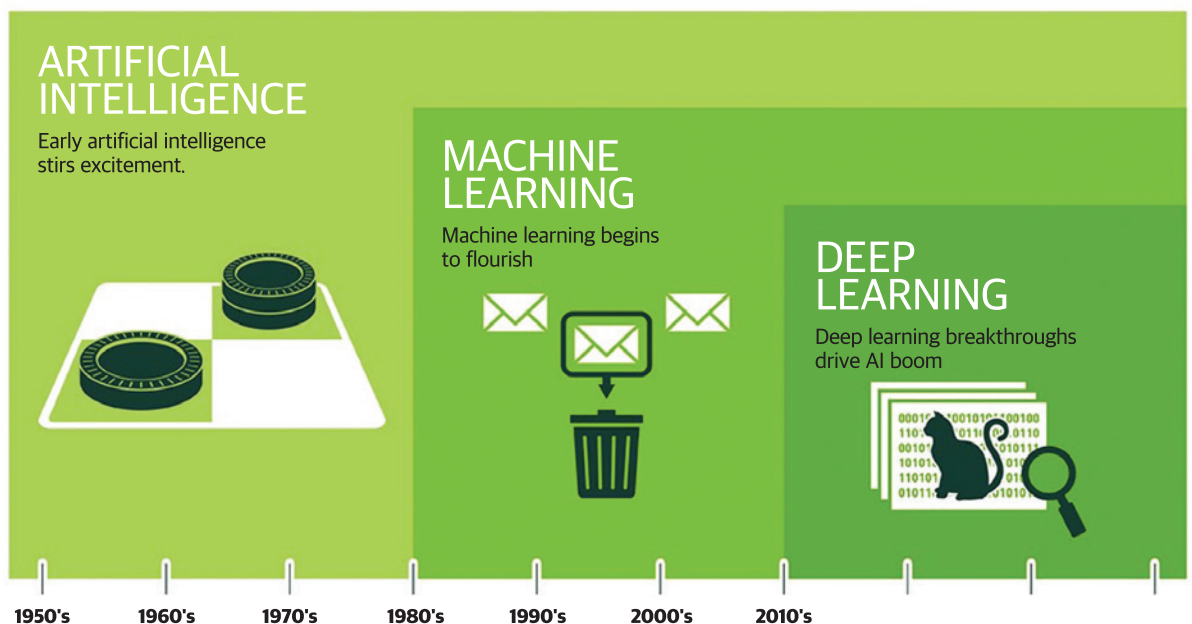


그림 4. 인공지능의 발전 역사

이제 인공지능을 사용한다는 개념이 기존의 관리자가 수동으로 탐지 장비를 통해 공격을 탐지하는 개념과 어떻게 다른지 알아보자. 일반적으로 보안장비가 공격을 탐지하는 방식은 관리자가 임의로 공격 탐지 조건을 설정하고 탐지 장비로 트래픽, 악성코드를 모니터링하다가 공격인지 정상인지를 판단해서 알려주는 방식으로 작동한다. 이에 비해서 인공지능을 활용한다는 것은 사전에 많은 양의 정보 로그, 악성코드와 각각의 로그가 정상인지 공격인지에 대한 결과가 포함된 빅데이터 정보를 입력해주면 인공지능 장비가 학습을 통해서 공격 탐지조건을 생성해서 공격 탐지에 이용하는 것을 의미한다. 이렇게 생성된 조건을 공격 탐지 장비에 입력하면 높은 정확도의 결과값을 기대하는 방식으로 인

공지능이 활용되는 것이다.

[그림 5]는 기존의 일반적인 보안장비를 운영할 때 발생하는 케이스를 설명한 것이다. 운영자가 공격으로 판단할 수 있는 조건을 입력해 주어야만 공격을 탐지할 수 있다. 그래서 운영자의 경험과 숙련도에 따라 공격 탐지의 정확도가 널뛰기할 수밖에 없다. 공격 탐지 조건을 조직의 환경에 최적화해서 적용하는 것이 탐지정확도를 높이는 관건이다.



장비 : 공격 트래픽은 어떤 기준으로 판단하나요?
 운영자 : 출발지 IP가 111.111.111.0/24 이면 모두 공격으로 판단하죠.
 장비 : 그럼 111.111.111.0/24 에서 들어오는 모든 트래픽은 공격으로 판단할게요.
 운영자 : 잠깐, 그건 아니고 포트 넘버가 8888인 것이 공격인 것 같은데.
 장비 : 그럼 111.111.111.0/24이고 포트 넘버가 8888인 것은 공격으로 판단할게요.
 운영자 : 아 그런데, 그 조건에도 정상 트래픽이 많이 포함되어 있는 것 같아.
 패킷의 데이터레벨까지 확인해서 admin-attack이라는 문자열이 포함되면 공격인 것 같아.
 장비 : 그럼 데이터레벨까지 확인해서 그 조건인 경우에 공격으로 판단하겠습니다.
 운영자 : 아, 그렇게 해도 공격이 들어왔는데 공격을 탐지하지 못하네.
 이하 무한 반복.....

그림 5. 일반적인 보안장비가 공격을 탐지하는 방법

[그림 6]의 경우 인공지능을 활용하여 공격을 탐지하는 시나리오를 설명하였다. 운영자는 먼저 양질의 학습데이터를 확보하여야 한다. 얼마나 많은 양의 질 좋은 데이터로 학습하는가에 따라 공격 탐지정확도가 좌우된다. 이렇게 제공된 학습데이터를 바탕으로 인공지능 장비는 공격과 정상을 구분할 수 있는 조건을 학습하게 된다. 이렇게 학습된 조건을 가지고, 학습용으로 제공하지 않고 테스트용으로 남겨둔 데이터로 테스트를 진행해 보고 정확도를 가늠해 볼 수 있다. 정확도가 대략 90% 이상을 상회한다면 실제 운영망에 적용해 볼 가치가 있으므로, 실망에 적용해서 일정 기간 테스트를 해 볼 수 있다. 테스트 결과 특정 영역에서는 정확도가 높으나 기타 부분에서는 탐지정확도가 낮다면 정확도가 낮은 부분에 대한 학습데이터를 확보하여 추가 학습이 필요하다.

탐지정확도를 높이기 위해 가장 중요한 것은 ‘어떻게 양질의 학습데이터(공격 로그, 악성코드와 악성/정상 여부 결과 값)를 확보할 수 있는가’라는 점과 ‘인공지능 장비가 보안담당자가 의도하는 방향으로 정상적으로 학습할 수 있도록 학습데이터를 사전에 잘 가공(전처리)할 수 있는가’라는 점이다. 인공지능이 사람이 수행하는 업무정확도를 같은 수준이나 혹은 더 높은 수준을 유지하기 위해서는 인공지능이 학습할 때 사용하는 데이터의 양이 많아야 하고 품질이 우수



장비 : 수집된 데이터를 주시면 제가 학습해서 공격을 판단하는 조건을 계산할게요.
 운영자 : 그럼 내가 기존에 저장해 놓은 데이터를 입력해 줄게, 공격을 판단하는 조건을 알려줘.
 장비 : 제공한 데이터를 통해 공격 탐지할 수 있는 조건을 계산(학습)하였으니 테스트해 보세요.
 운영자 : 그럼 별도로 저장된 수집데이터로 테스트해 보니 정확도가 92%쯤 나오니 실망에 적용해 볼게.



장비 : 인공지능파트에서 학습한 공격 탐지 조건을 입력해 주시면 공격 탐지를 시작하겠습니다.
 운영자 : 공격 탐지 결과를 보니 웹 서버에 대한 공격 탐지는 정확도가 높는데, 다른 서버에 대한 정확도는 낮아.
 장비 : 인공지능파트에서 웹 서버 공격 데이터 이외 다른 서버에 대한 공격데이터를 입력해서 학습해 보세요.
 운영자 : 음. 구할 수 있는 데이터가 없는데, 어디서 구해야 하나..

그림 6. 인공지능을 활용한 공격 탐지의 운영시나리오

해야 한다. 그런데 질 좋은 데이터를 대량으로 확보하기가 쉽지 않다. 구글, 애플, 아마존 등의 IT 기업에서는 서비스를 무료 혹은 저렴한 가격으로 제공하면서 얻으려는 가치는 이런 서비스를 이용하는 사용자의 사용패턴이 포함된 다양한 데이터이다. 이렇게 수집된 데이터를 바탕으로 더 다양하고 정확한 서비스를 개발하여 매출을 증대시키기 위한 인공지능의 학습데이터로 활용되는 것이다. 다음 [그림 7]과 같이 인공지능을 학습시키는데 사용되는 데이터의 질에 따른 학습 결과값을 표시하였다. 양질의 데이터를 입력하면 낮은 품질의 결과가 나올 수도 있고 혹은 양질의 학습 결과가 나올 수도 있지만, 질 낮은 쓰레기 데이터를 넣으면 아무리 노력해도 쓰레기 결과값밖에 나올 수 없다. 그래서 인공지능을 활용하기 원하는 조직은 바로 이 양질의 빅데이터를 수집하는데 혈안이 될 수밖에 없는 것이다.

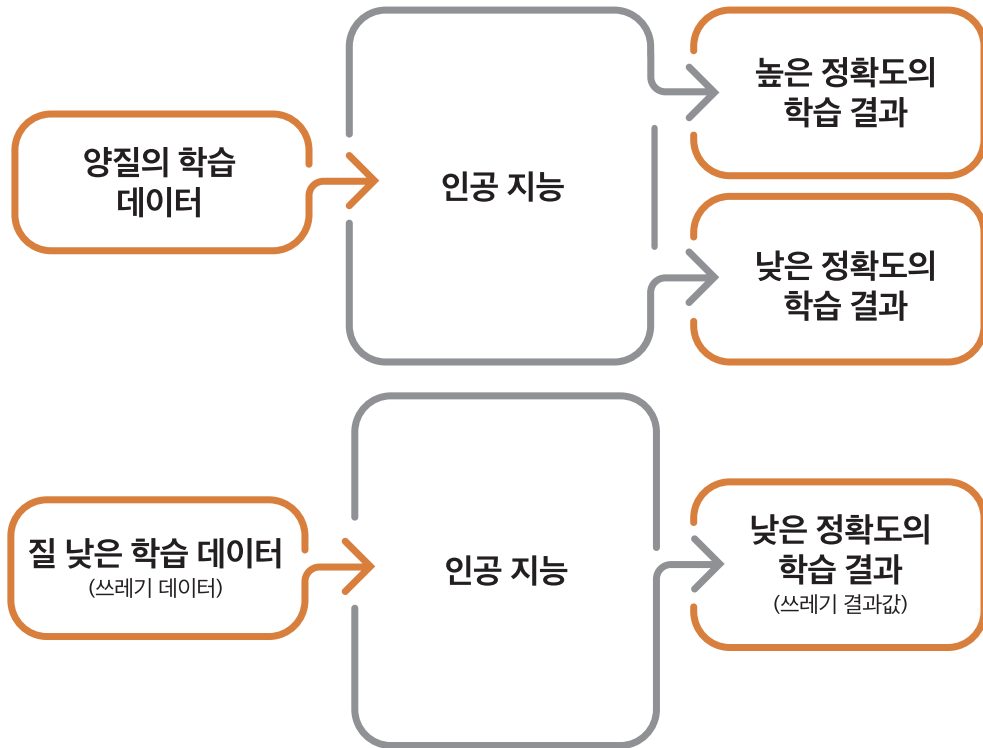


그림 7. 인공지능의 학습데이터의 질에 따른 학습 결과값의 품질

행정안전부 NIA 한국정보화진흥원

2020 공공데이터 청년 인턴십 모집

접수기간 7. 22.(수) ~ 7. 31.(금)

1 데이터 수집

공공·민간 데이터 수집

- 바이오·의료, 제조, 에너지, 금융, 행정 등

SOC 디지털화

- 도로·철도, 하천, 스마트시티 등

디지털 인프라 확충

2 데이터 축적·가공

- AI 학습용 데이터 150종 구축 (20. 380억원 ~ 2,925억원 중)
- 5대 분야(추가) 빅데이터 플랫폼 센터 구축 (20. 448억원 ~ 400억원 중)
- 공공데이터 개방 확대·품질개선 (20. 363억원 ~ 1,275억원 중)

데이터 경제 기반 강화

3 데이터 활용(AI)

- AI가공 데이터 바꾸처 (20. 249억원 ~ 489억원 중)
- AI 바꾸처 (20. 39억원 ~ 500억원 중)
- AI융합(AI+X) (20. 추경 신규 211억원)

민간 AI활용 비즈니스 창출

클라우드

그림 8. 데이터댐 (인공지능 활용을 위한 양질의 빅데이터 수집사업) 공고문

정부에서도 국민의 안전과 생활편의를 목적으로 공공분야에 대한 인공지능 활용을 활성화하기 위해 빅데이터의 중요성을 인식하고 양질의 데이터 수집을 위해 [그림 8]과 같이 청년인턴십을 활용하여 수집된 데이터에 결과값(Name Tag)을 입력하는 사업을 시작하고 있다. 최근 정부에서 디지털뉴딜 중 하나로 발표한 데이터댐 사업의 일환으로 다양한 공공데이터를 인공지능의 학습자료로 활용하기 위해 부족한 데이터를 생성하고 수집된 데이터에 대한 정확한 설명을 입력하는 작업을 추진하고 있다. 이런 작업을 넓게 보면 데이터에 대한 전처리 작업이라고 할 수 있는데, 학습 목적에 부합되는 자료를 생성한 후 선별하고 정확한 설명을 입력하는 단계로 인공지능의 학습정확도를 좌우하는 매우 중요한 단계이다. 학습하는 데이터에 대한 설명(Name Tag)이 정확하지 않으면 학습한 결과는 활용할 수 없는 낮은 수준의 결과값밖에 나올 수 없기 때문이다. 그래서 이런 작업이 단순

R&D 데이터셋 목록

데이터셋 그룹	데이터셋	제공자	소개
악성코드 데이터셋	PC악성코드	한국인터넷진흥원 세인트시큐리티	악성코드 변종탐지, 그룹분류 기술, 성능평가에 활용된 5,045개의 악성코드 샘플과 분석결과
	지능형 악성코드	안랩, 허우리, 세인트시큐리티	2017 정보보호 R&D 데이터 탐인지 대회의 '악성코드 선파' 트랙에 활용된 300여개의 지능형 악성코드
	대용량 정상, 악성파일 1 (2017 해킹)	한국인터넷진흥원, 허우리, 세인트시큐리티	2017 정보보호 R&D 데이터 탐인지 대회의 '악성코드 탐지' 트랙 예선에 활용된 5000개의 대용량 정상, 악성파일
	대용량 정상, 악성파일 2 (2017 분석)	한국인터넷진흥원, 허우리, 세인트시큐리티	2017 정보보호 R&D 데이터 탐인지 대회의 '악성코드 탐지' 트랙 본선에 활용된 5000개의 대용량 정상, 악성파일
	대용량 정상, 악성파일 3 (2018)	한국인터넷진흥원, 안랩, 이스트시큐리티, 허우리, 세인트시큐리티	2018 정보보호 R&D 데이터 탐인지 대회의 'AP'전 악성코드 탐지에 활용된 5000개의 악성코드
	대용량 정상, 악성파일 4 (2018)	한국인터넷진흥원, 안랩, 이스트시큐리티, 허우리, 세인트시큐리티	2019 정보보호 R&D 데이터 탐인지 대회의 'AP'전 악성코드 탐지에 활용된 4000개의 악성코드
	VX Heaven 악성코드	호서대학교	VX Heaven에서 배포하는 15개의 악성코드 그룹으로 구성된 236,754개의 악성코드
악성코드 탐 데이터셋	메모리 상주 악성코드	한성대학교	실행파일 없이 메모리 상에 상주하는 악성코드 564개 및 분석 보고서
	정상악성집	고려대학교	다양한 악성집 repository에서 수집한 9,990개의 악성집 샘플과 109,933개의 정상집
	오실/제로데이 악성집	고려대학교	Andro-Profil에서 오실란 업로드 정상집, 악성집 샘플, Zero-day 업 샘플을 포함하는 데이터셋
	대용량 악성, 정상 파일 1 (2018)	고려대학교	2018 정보보호 R&D 데이터 탐인지 대회의 'AP'전 악성코드 탐지 및 분석 보고서
차량 데이터셋	Car Hacking 데이터셋	고려대학교	Dx5 attack, fuzzy attack, spoofing the drive gear, spoofing the RPM gauge를 포함하는 자동차 해킹 데이터셋
	차량 이상징후 탐지	고려대학교	2017 정보보호 R&D 데이터 탐인지 대회의 '차량 이상징후 탐지' 트랙 본선에 활용된 정상 및 3종류의 차량 공격 시도 해킹 데이터
	차량 주행 데이터 (2018)	고려대학교	2018 정보보호 R&D 데이터 탐인지 대회의 '차량 주행 데이터 기반 도난 탐지'에 활용된 700km 차량 주행 데이터
	자동차용 네트워크 데이터 (2019)	고려대학교	'자동차용 침입탐지 2019'에 활용된 3개 차량의 정상 및 공격 데이터
기타	스크립트 난독화 도구	한국인터넷진흥원	난독화된 자바 스크립트 해제 기술 성능평가에 활용된 원본 스크립트, 난독화된 스크립트, 난독화 도구 4종 분석자료
	취약한 바이트코드 1 (2018)	한국인터넷진흥원	'AP'전 취약점 자동탐지 2018'에 활용된 취약점이 포함된 바이트코드 95개
	취약한 바이트코드 2 (2019)	한국인터넷진흥원	'AP'전 취약점 자동탐지 2019'에 활용된 취약점이 포함된 바이트코드 60개
	네트워크 패킷 데이터	한국인터넷진흥원, 안랩, 한국인터넷공사	'AP'전 네트워크 위협탐지 2019'에 활용된 약 60GB의 패킷을 포함한 파일
	게임유저 액션로그 데이터	고려대학교	'게임넷 탐지 2019'에 활용된 약 3주치의 AION 게임유저 액션 로그

그림 9. 인공지능을 학습시키기 위한 학습용 빅데이터 / 출처 : 한국인터넷진흥원

반복적인 작업이지만 인공지능의 학습 결과를 실생활에 활용할 정도가 되기 위한 아주 중요한 단계라고 할 수 있다.

정보보호 분야에서 인공지능을 활용하려는 분야도 다양하다. 악성코드 여부를 판독하는 작업, 온라인게임에서 오토봇을 사용하여 유저를 찾는 작업, 네트워크 패킷을 분석하여 공격 여부를 판별하는 작업, 자율주행차량의 이상 징후를 탐지하는 작업, 다양한 보안경고 로그를 분석하여 실제 공격을 판별하는 작업 등에 인공지능을 적용하려는 시도가 진행되고 있다. [그림 9]는 정보보호 분야의 인공지능을 학습시키는 데 필요한 빅데이터를 표시한 것이다. 한국인터넷진흥원(KISA)에서는 '정보보호 R&D 데이터 챌린지'라는 대회를 매년 개최하고 있는데, 대학의 정보보호동아리나 개인들이 아래와 같이 제공된 빅데이터를 이용하여, 얼마나 정확도가 높은 인공지능 학습 결과를 뽑아내는지 경쟁하는 대회이다. 이를 통해 인공지능에 대한 관심을 높이고 인재를 육성하여 국내 정보보호 산업 발전에 이바지하기 위한 취지로 개최되고 있다.

인공지능에 양질의 빅데이터를 입력시켜 실제 업무에 적용할 수 있는 레벨로 만드는 작업은 쉬운 작업이 아니다. 데이터를 기준에 맞게 선별하고, 부족한 데이터는 추가로 만들거나 수집하여야 하며, 수집된 데이터에 대한 설명 내용을 추가하는 가공 작업은 모두 사람의 수작업이 필요한 업무이다. 그뿐만 아니라 이렇게 학습된 결과에 대해 적절한 해석과 실생활에 적용하기 위한 적용 노하우를 축적하는 일은 IT 전문가가 아닌 적용하려는 분야 전문가들의 도움이 절실한 영역이다.

최근부터 업무에 선별적으로 적용되기 시작한 인공지능이 시간이 갈수록 정확도가 높아지고 좀 더 많은 업무에 적용될 수 있겠지만 아직은 초기 단계라 많은 시행착오가 필요할 것으로 예상된다. 언젠가 보안전문가의 도움 없이도 정확도 높은 공격 탐지와 차단이 가능한 보안장비가 나오기를 기대한다.

그럼 다음 호에서는 가상으로 방송시스템에 대한 APT 공격시나리오를 설명하면서 APT 공격을 통해 어떤 결과를 초래할 수 있는지 알아보고, APT 공격에 대한 효과적인 대응을 위한 APT 방어 장비의 종류와 효과적인 공격 차단을 위한 전략에 대해 소개하도록 하겠다. 📖