**딥러닝을 사용한 네트워크 이상 탐지에 대한 최신 문헌 검토**

1. 소개

최근 딥러닝을 통한 네트워크 이상 감지를 강조함.

APT, DDOS, Zero-day 와 같은 다양한 공격 유형이 있다.(각 고유한 패턴을 효율적으로 감지할 수 있는 것을 찾는게 중요함)

-의도된 공격 유형 탐지, 기능적 차이, 탐지 정확도의 차이와 같은 측면을 고려함

이 논문에 초점은 DDos, Zero-day,웹 공격, 딥러닝을 사용한 APT탐지

두 번째 섹션에서는 이상징후를 감지하기 위해 딥 러닝을 사용한 다양한 이상탐지기술을 설명

세 번째 섹셔은 개선 사항을 요약

네 번째 섹션은 제시된 연구의 장단점에 대해 논의

1. 딥러닝을 사용한 네트워크 이상 탐지

Chuanlong Yin et al.

딥 러닝 기반 침입 탐지(RNN-IDS) 방법을 제안

순환 신경망,시스템 성능으로 IDS를 개선

이진 분류 및 다중 클래스 분류의 두 가지 유형의 분류 문제를 해결할 수 있는 가능성을 검토

(정상, DoS, User to root(U2R), Probe(probing), Root to Local(R2L)) 이 5가지로 테스트

-Theano를 딥러닝 프레임워크로 작성한 파이썬을 사용

-데이터 차원이 많기에 RNN을 사용

-NSL-KDD데이터 셋 사용(41개의 변수, 그 중 38개는 숫자 3개는 숫자X)

-RNN을 훈련 후 테스트를 위해 데이터 전처리(차원 수를 41개에서 122개로 늘림)

-이진 분류의 경우 100Epoch, 80 히든레이어, 0.1 학습률이 가장 정확도가 높다

-다중 클래스 분류의 경우 80Epoch, 80 히든 레이어, 0.5 학습률이 가장 정확도가 높다

-GPU 가속, 그라디언트가 폭발 및 사라지는 것을 방지하는 방법에 중점

-침입 탐지 목적으로 LSTM,양방향 RNN알고리즘의 분류성능을 향상시키는 방법을 연구

-APT 공격을 탐지할 때 다차원 처리는 매우 중요함

Xiaoyong Yuan et al

Bidirectional Recurrent Neural Network 기반 DDoS 탐지를 제안

DDos 공격 트래픽은 일반 트래픽과 매우 유사하기 때문에 자동으로 탐지 하기 어렵고

공격자는 Flash-crowd방식을 사용하여 정상적인 높은 사용량 피크를 모방하려 함

통계 기반 탐지는 특정 DDoS에서 잘 수행되며 다양한 공격에 대한 사전 메트릭이 필요

DeepDefence 모델 제안

ㄴRNN,LSTM, GRU와 같은 딥러닝 방법으로 테스트 후 Random Forest ML과 비교테스트

제안된 방법에서 양뱡향 순환 신경망 설계

입력노드 : 정방향, 역방향 순환 계층이라는 두개의 분리된 병렬 히든 레이어가 있다.

데이터는 히든레이어를 통과 후 결과는 이상 징후를 예측하기 전 후자 계층에 연결됨

RNN은 과거 정보 문제(LSTM, GRU)를 잘 처리하기에 선택되며 과거 데이터 패턴을 기반으로 성능을 향상시킨다. 단일 패킷 감지는 수행X

제안된 방법을 테스트하기 위해 ISCX2012데이터 세트의 7일 기록을 선택, 두 개의 개별 데이터 세트(data14, data15)로 추출.

첫째.두 데이터 세트로 서로 다른 RNN방법을 테스트 후 비교

LSTM은 data14(97.996%)에서 가장 높은 정확도

3LSTM은 data15(98.410%)

둘째. 동일한 데이터 세트로 랜덤 포레스트 정확도를 테스트

정확도는 낮았지만 데이터 세트 간에 약간의 차이가 있었음(data14는 97.1% data15는 92.5%)

다양한 환경에서 성능테스트를 위해 다양한 DDoS 벡터 및 시스템 설정을 늘림

이 방법을 기반으로 Deep DDos 방어 시스템을 구축하고 실제 환경에서 테스트

그 결과 정확도는 0.099%차이로 거의 같았음. 랜덤 포레스트에서는 4.599%결과

LTSM은 APT 공격을 탐지할 때 중요할 수 있는 과거 정보를 불러올 수 있기 때문에 APT 공격 탐지에 대한 추가 연구의 가치가 있다.

Sergey Andropov et al.

백스페이스 전파 알고리즘 훈련을 통한 다층 퍼셉트론 ANN을 제안

-시그니쳐 기반 탐지 방법으로 불가능하거나 알려지지 않은 공격을 탐지하는데 목적.

-네트워크 공격, Malware, 하드웨어 소프트웨어 오작동으로인해 이상징후 발생->위협으로 처리

-새로운 네트워크 장치 또는 소프트웨어 설치 또는 새 사용자 장치가 추가되는 네트워크 토폴리지 변경이 수시로 발생하고 네트워크에서 새로운 패턴이나 동작이 발생

하지만 신경망은 이러한 변화를 허용하고 가중치를 조정함으로써 적응할 수 있다.

제안된 방법

1.오프라인 트래픽 분석

2.온라인 트래픽 분석

-데이터 집계를 사용하여 다른 방식의 가변적인 트래픽 데이터에서 패턴을 감지

-네트워크 장치에서 정보 수집을 위해 Netflow 프로토콜을 구현 후 ANN의 이상 탐지 프로세스를 위해 필터링 및 집계

집계 : 1. 시간당 패킷 수

2. 평균 패킷 크기

3. 신경망에 대한 입력으로 사용된 포트 사용.

ANN : 1. 입력 계층

2. 히든 레이어

3. 출력 레이어(모든 이상에 대한 뉴런과 2개의 추가뉴런(정상용, 알수없음))

6개의 입력 레이어, 10개의 히든 레이어, 7개의 출력레이어 / sigmoid함수 사용

테스트: 로컬 ISP가 수백개의 L2노드에서 한달동안 데이터 수집 (정상 트래픽으로 간주)

테스트목적으로 (Dos, DDos 공격 ,포트 스캔,이메일 스팸,라우터)와 같은 소규모 이상현상을 만듦

신경망이 알수없는 이상클래스를 감지할 수 있는지 테스트하기 위해 사용자 지정 이상현상을 만듦

-생성된 이상 현상은 탐지 테스트를 실행하기 전 일반 트래픽 데이터 세트에 삽입

결과 : 제안된 ANN은 높은 정확도로 알려진 / 알려지지 않은 이상 징후를 모두 감지 가능

그러나 idle scan 은 감지하기 여려움

사용자 지정 이상 현상은 150,000회의 분류 반복으로 정확도가 가장 낮았음

ARP 스푸핑은 두 번째로 낮았음

Idle scan은 세 번째로 낮았음

300,000회의 분류 반복 후 모든 이상 현상은 80%이상 정확도, idle scan은 80% 미만

오탐율 감소 및 집계 알고리즘 최적화에 초점을 둠

동시 공격 벡터가 여러개 있어 ART 공격을 탐지하는 데 유용함

R. Can Aygun et al

-Zero-day 공격 탐지를 위한 딥 러닝 기반 IDS를 제안

-향상된Autoencoder(AE) 기반 모델로 Zero-day 공격 탐지를 개선하는 것이 목적.

-현재 IDS는 이전 공격의 서명DB를 기반으로 하기에 최신 상태로 유지되어도 잘 알려지지 않은 새로운 공격을 탐지하지 못함. -> 머신러닝활용

-확률적 이상 임계값 결정 알고리즘으로 AE를 제안

이 알고리즘과 확률적/결정적 AE의 성능을 테스트하고 비교

AE는 인코딩, 디코딩의 기본 기능이 있다. 입력 레이어와 출력 레이어의 노드 수는 동일

인코딩을 사용하여 히든레이어에서 더 적은 양의 단위로 입력을 표현 후 다음 디코딩 단계에서 인코딩 된 입력을 출력으로 재구성.

AE는 훈련데이터로 인해 항등함수가 될수있어 정확도가 낮을 수 있다->확률적 노이즈 제거 방법

AE훈련에 반지도훈련 방법 사용(제로데이 공격에 대한 탐지 정확도가 더 높음)

훈련과 모델 분류를 위해 일반 트래픽 데이터만 사용

정상/비정상으로 레이블 지정, 임계값에 따라 탐지 결과를 해석

반지도 훈련에는 비정상인 경우 낮고 정상인 경우 높은 재구성 오류가 있기에 감지 결과는 적절한 임계값으로 올바르게 해석해야됨.

테스트 : NSL-KDD 데이터 세트 선택. (테스트 셋과 학습 셋은 완전히 다른 공격 유형을 가지고 있다)

데이터는 사전처리되고 특징 수는 41개->121개로 증가, 데이터 셋은 [0,1]범위로 확장

결과 : 결정적AE는 88.28% 확률적AE는 88.65%

-APT탐지에서도 가능하지만 탐지 결과에 대한 자동 해석이 권장된다.

Nguyen T. Van et al.

딥러닝 기반 NIDS를 제안

NIDS에 DL 방법을 구현하여 이상 기반 탐지를 얻고 알려진/잘안알려진 공격을 탐지하는것이 목적

인터프라이즈 네트워크는 매우 중요하며 보안 솔루션이 유연해야 하는 상황이 발생->딥러닝 활용

이상현상 : 1. 포인트 이상

2. 문맥적 이상

3. 집단적 이상

Dos, Probe, U2R, R2L과 관련,

-multi-layer 신경망 구성을 위해 스택 방법 사용

(첫번째 신경망의 히든레이어 출력이 두번째 신경망에 대한 입력으로 총 세 개 결합된 신경망)

사전 훈련을 위해 Restricted Boltzmann Machines (RBM) 과Autoencoder (AE) 선택

테스트로 KDDCup99 데이터셋 선택, 위 두가지 방법으로 테스트

결과 : 두 방법 모두 탐지 정확도가 높음. 데이터 분류는 AE가 성능이 더 우수함.

학습 시간은 AE가 더 김

-빠른 계산 속도를 위해 시스템을 병렬 플랫폼에 결합하는 방법을 연구해야함

-훈련 오류 감소와 탐지 정확도 향상을 위해 진폭을 최적화하고 줄이기 위한 사전 훈련 기술을 재연구해야함.

APT감지에 대해 연구가능

Sunhee Baek et al

탐지 정확도 향상을 위해 지도/비지도 학습 방법의 기능 조합을 제안.

개별 훈련 방법의 문제점 :

1. 지도학습은 레이블이 지정된 데이터를 사용해 빠른 데이터 포인트 테스트를 수해앟며 높은 정확성을 갖지만 모든 데이터에 레이블을 지정하는 것이 어려움
2. 비지도 학습은 레이블 대신 학습 과정 중 데이터를 분류하며 탐지에 적은 양의 데이터를 요구하지만 탐지 정확도가 낮고 런타임 복잡성이 높다.

-빠른 탐지가 중요한 기업, ISP 네트워크에서 이상 탐지를 개선하는 것이 목적.

이상징후를 탐지하는 데 더 관심

기초 연구를 통해 전통적인 이상탐지의 정의: 가정 : 정상 데이터 인스턴스는 크고 밀집한 군집에 속하며 이상은 작거나 희박한 군집에 속한다

-(size := {small or not})와 (density := {sparse or not})에 대한 매개변수가 정의됨

이상 탐지의 가정을 재정의 : 1. 매우 밀집한 군집은 이상으로 레이블됨

2. 작거나 희박한 군집은 이상으로 레이블됨

3. 그렇지 않은 경우 군집은 정상으로 레이블됨

-이 방법으로 테스트,최적화

i) Naive Bayes, ii) Adaboosting, iii) SVM and iv) Random Forest 로 레이블된 데이터로 훈련하고 테스트

테스트 : NSL-KDD데이터 셋 선택

-기존 이상 탐지 방법의 성능을 테스트->이상 징후 탐지를 제대로 못함

군집 크기(16,32,64,128,256)으로 테스트 이 중 16이 가장 높은 성능, 정확도는 62%미만

-재정의된 이상 탐지 방법으로 테스트->16이 62%에서 88%로 향상.

-레이블된 데이터와 다른 군집 크기로 네가지 지도 학습 방법으로 테스트-> 군집 크기와 데이터가 달라도 비지도학습으로 레이블링된 데이터로 지도 학습으로 학습하면 높은 정확도로 수행됨

S. Baek et al.

K-Means 클러스터링의 임의성을 최소화하는 방법을 개발할 것이라고 언급함

방법 : 여러번의 실행 결과를 포함하는 것

분류를 개선하기 위해 데이터 차원을 줄이는 것

테스트 결과 : 서로 다른 방법들의 높은 성능의 기능을 결합하여 전반적인 성능을 향상

이상을 감지하는 것만으로 사고 대응의 우선 순위를 충족할 수 없고 가장 위험한 공격에 먼저 대응하도록 하는 공격 유형 분류가 필요하다.

Vrizlynn L. L.

IEEE 802.11 표준 무선 네트워크에 대한 네트워크 이상 탐지 , 공격 분류 방법 제안

딥러닝을 통해 무선 네트워크에서 다양한 유형의 공격을 분류하는 것이 목적

합법적인 트래픽 외의 공격 : : i) flooding, ii) injection and iii) impersonation type attacks

-Wi-Fi는 다양한 취약점이 있다.->소프트웨어 취약점 패치를 포함하여 나중에 분석 및 가능한 복구 작업을 위해 분류가 필요-> SAE(Stacked Auto-encoder) 제안

-2개 및 3개의 히든레이어가 있는 두가지 SAE 프레임워크를 제안

첫 번째 히든 레이어에 256개의 뉴런, 두번째는 128개의 뉴런, 세번째는 64개의 뉴런

ReLu, LReLU, Sigmoid, PReLU 활성화 함수 사용

Softmax회귀 사용(다중 클래스 분류를 지원 로지스틱 회귀는 이진분류만 지원)

테스트 환경 : 데스크톱, 노트북2대, 스마트폰2대, 태블릿, 스마트TV를 갖춘 실험실 환경(WEP 암호화 액세스 포인트(AP))를 만들어 무선 트래픽 생성 후 직접 캡처)

Kali Linux는 15가지 공격 유형을 실행, 여러 침투 테스트 도구로 사용

공격은 injection, flooding, impresonation 세 가지로 분류되고 정의

테스트는 2개의 SAE 프레임워크와 4개의 작업 기능으로 실행,알려지지 않은 15가지 공격 유형을 포함하는 데이터 세트를 사용

결과 : PReLU함수가 있는 두개의 히든 레이어 프레임워크가 98.6%로 가장 높은 정확도

사칭공격이 가장 제거하기 어려움

이상탐지를 위해 표준 프로토콜 트래픽 사용을 제안

Bluetooth, LoRaWan, ZigBee 및 IoT 시스템에 사용되는 센서 네트워크에 사용되는 기타 무선 프로토콜로 해결할 수 있다면 연구할 가치가 있음

히든레이어 수와 활성화 함수의 양을 변경하면 다른 예측을 제공할 수 있어 신경망을 최적화하는 것이 정말 중요함

Wei Wang et al

트래픽 데이터를 이미지로 변환하여 CNN을 사용하여 Malware 트래픽 분류를 제안

원시 데이터를 입력으로 사용하고 이미지로 변환하여 분류

트래픽 세분성에는 TCP연결, 흐름, 세션, 서비스, 호스트, 세션정보가 포함

플로우 패킷에는 소스, 소스포트, 대상 , 목적지 포트, 전송 레벨 프로토콜로 정의

세션은 양방향 플로우로 정의

이미지 크기를 줄이고 IP, MAC주소의 세션 문제를 제거하기 위해 4가지 유형의 데이터 표현이 정의.( Flow + All, Flow + L7, Session + All, Session + L7 (L7 is OSI layer 7).)

테스트 : USTC-TFC2016데이터셋을 만든후 USTC-TK2016을 개발

공개 웹 사이트에서 10개의 서로 다른 malware트래픽 데이터 셋과 IXIA BreakingPoint 시뮬레이터로 일반 트래픽 수집

4가지 표현 방법에 대해 8개의 테스트 실행

전처리 단계를 거쳐 752,040개의 레코드를 생성

시각적 테스트 결과 각 트래픽 클래스들이 명백한 차별 정도를 가지고 있으며 높은 일관성을 가졌다

세션 + 전체 표현이 가장 높은 정확도를 보여줌

평균 정확도는 99.41%

변환에 시간이 많이 걸리는 경우 이 방법이 적합하지 않음

Youbiao He et al.

딥러닝을 통해 스마트 그리드에서 실시간FDI(False Data Injection)감지를 제안

스마트 전력 그리드의 모니터링과 통신은 품질과 지능적 기능을 얻기 위해 IP 네트워크로 진화

이러한 발전은 FDI 공격의 가능성을 증가

주요 초점은 FDI에 의해 실행되는 전기 절도 탐지를 개선하는 것

-스마트 전력 그리드는 페이저 측정 장치(PMUs), 스마트 미터, 원격 단말 장치(RTUs) 및 감독 및 제어(SCADA) 시스템을 포함한 복잡한 연결된 시스템

시스템 상태 정보는 전력망의 안정성과 효율성에 매우 중요

이상탐지를 위해 확장된 Deep Belief Network (DBN)인 "Conditional Deep Belief Network (CDBN)"을 제안

CDBN은 첫 히든레이어에서만 CGBRBM을 실행 나머지 히든레이어에서는 RBM수행

시스템 성능은 패턴 감지, 민감도, 정상 데이터 패턴 차이, FDI침해 테이터에 좌우함

이 방법은 전력 시스템의 정적 물리적 토폴로지에 크게 의존함

테스트 : CGBRBM, RBM 훈련

Sigmoid활성화 함수 사용

시뮬레이션 목적으로 실제 소스에서 데이터 수집(소량의 손상데이터 포함)

수집된 실제 위협 데이터와 유사한 패턴을 가진 인공적으로 레이블이 지정된 위협 데이터를 생성

TCP 덤프 또는 원시데이터 대신 IEEE118버스 및 IEEE300 버스시스템과 특정 부하 프로필 데이터를 사용

ANN, SVM같은 다양한 방법으로 테스트

결과 : 전반적으로 매우 낮은 위 양성 및 위음성률로 탐지 정확도가 높았음

FDI 공격 행동보다 실용적인 방법으로 모델링하기 위한 연구를 할 것

효율적 감지를 위해 필요한 감지 장치의 최소 수가 무엇인지 연구할 것

1. 추가 개선 사항 요약

-보다 정확한 이상 탐지를 위해 인식된 개선 사항을 제시

-KDDCup99 데이터가 구식, 부족하다고 언급

ㄴ 새로운 방법을 벤치마킹하는 데 널리 사용

-현재의 데이터 세트를 찾거나 KDDCup99를 대체할 데이터 세트를 만들어야 함

-이상 탐지를 위해 IP 및 TCP헤더의 특정 부분만 선택하는 것은 누락 데이터로 인해 탐지 기능을 제한할 수 있다

-신경망은 알려지지 않은 공격에 대해 더 높은 탐지율을 보인다.

-NN은 탐지 시스템 비용 구조에 부정적인 영향을 미친다.

-탐지율은 훈련 데이터의 양에 크기 좌우함

-이진분류, 정상 또는 비정상 트래픽은 실제 공격 벡터에 충분한 정보 제공X.

-이진분류는 거짓 부정 비율이 매우 낮아야함.

1. 결론

-하드웨어가 충분히 강력하지 않은 경우 고속 네트워크에서 실시간 감지에도 문제를 일으킬 수 있는 더 많은 계산 능력이 필요.

-더 많은 훈련 데이터가 필요.

-자가 학습을 통해 급변하는 네트워크 환경에 적응하고, 다차원 데이터를 다루며, 이전에 알려지지 않은 공격을 탐지하는 신경망 기능은 오랫동안 탐지되지 않은 채 행동하려고 하기 때문에 APT와 같은 정교한 공격을 탐지하는 데 큰 이점

-탐지 시간을 줄여 탐지

-딥러닝 이상 탐지 시스템 및 방법을 개발할 때 이러한 장단점을 고려해야함