머신러닝 기반 네트워크 침입탐지 및 상황 분석 솔루션 프레임워크 구현

Implementation of Framework for Machine Learning-Based Network Intrusion Detection and Situational Analysis Solutions

DO-Hee Kang, Yu-Jin Kim, Ye-Rim Yeon and Byung-Il Kwak Hallym Univ, Data-Driven Cybersecurity Research LAB



CONTENT

머신러닝 기반 네트워크 침입탐지 및 상황 분석 솔루션 프레임워크 구현 Implementation of Framework for Machine Learning-Based Network Intrusion Detection and Situational Analysis Solution

01. 요약

본 논문에서는 네트워크 보안 상황을 파악할 수 있는 네트 워크 트래픽 분석 솔루션을 웹 프레임워크로 구현하였음.

02. 서론

네트워크 규모의 지속적인 확장은 정보의 양과 복잡성을 증가시켜 네트워크 보안 상황 인지를 어렵게 만듦.

03. 침입 탐지 시스템

CICIDS2017 데이터셋과 직접 수집한 데이터셋으로 12가 지 공격 유형으로 식별하는 침입 탐지 시스템을 개발함.

04. 데이터 시각화

Circle Network 및 Parallel Coordinates Attack Visualization(PCAV) 등 11가지 그래프로 시각화하였음.

05. 구현

프레임워크는 입력한 네트워크 트래픽을 Pandas로 전처 리하고, 오픈소스 EChart를 통해 대시보드를 구성하였음.

06. 결론

구현한 웹 프레임워크는 보안 이벤트 시각화 기술로 네트 워크 보안 상황을 인지하고 대응할 수 있도록 구성함.

요약 (Abstract)

- <u>네트워크 보안 상황 파악이 가능한 네트워크 트래픽 분석 솔루션</u>을 구현함.
- 웹 프레임워크를 통해 제공하는 대시보드는 데이터 시각화, 데이터 분석, 보안 솔루션 및 백신 추천으로 이루어짐.
 - 입력한 네트워크 트래픽을 RandomForest 를 통해 네트워크 침입을 12가지 공격 유형으로 식별 및 분류함.
 - Circle Network 및 PCAV 등 11가지 시각화 그래프를 통해 공격 패턴에 대한 가시성을 높임.
 - 공격 분석표를 통해 구체적인 정보를 제공하고, 분석 데이터를 기반으로 보안 솔루션 및 백신을 추천하여 컴 퓨터 네트워크를 보호하도록 구성하였음.

I. 서론 (Introduction)

- 네트워크 규모의 지속적인 확장은 정보의 양과 복잡성을 증가시켜 네트워크 보안 상황 인지를 어렵게 만듦.
 - 네트워크 보안 상황을 분석하기 위한 방법으로 **보안 이벤트 시각화 기술**이 연구되고 있음.

• 보안 이벤트 시각화란?

- 네트워크상에서 발생하는 방대한 양의 이벤트를 시각화하는 기술로,
 - 보안과 관련된 많은 정보를 신속하고 정확하게 전달함.
- 본 논문에서는 네트워크 보안 상황을 파악할 수 있는 웹 기반 네트워크 트래픽 분석 솔루션을 구현함.
 - 본 프레임워크는 보안 이벤트 시각화 및 침입 탐지 시스템으로 네트워크 트래픽을 시각화 및 분석하고,
 - 분석 데이터를 기반으로 보안 솔루션을 제공하여 컴퓨터 네트워크를 보호하도록 구성함.

II. 침입 탐지 시스템 (Intrusion Detection System: IDS)

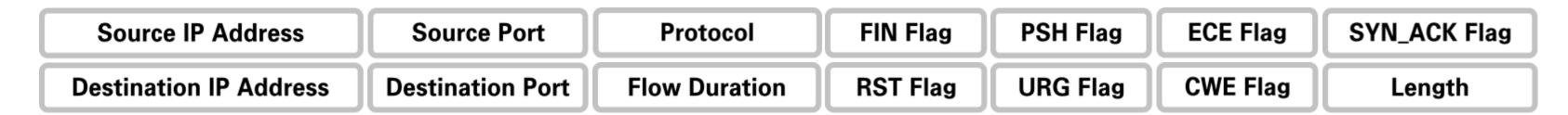
2.1 학습 데이터셋

- 여러 공격 데이터가 포함된 CICIDS2017 데이터셋 + 직접 수집한 트래픽 데이터셋을 사용함.
 - 학습에 사용된 데이터셋은 총 13개의 유형으로 분류됨.
 - Normal, ICMP Flooding, DDoS, DoS Hulk, DoS slowloris, DoS slowhttptest, DoS GoldenEye,
 Heartbleed, Infiltration, PortScan, Botnet, FTP-Patator, SSH-Patator

II. 침입 탐지 시스템 (Intrusion Detection System: IDS)

2.2 피처 엔지니어링

- 침입 탐지 시스템(IDS)을 서버에서 운영하기 위해 경량화시킴.
 - 피처의 수를 최소화하고, 데이터 시각화 및 모델 학습에 사용되는 피처를 동일하게 구성함.
 - 데이터 시각화에서 네트워크 트래픽을 쉽게 파악할 수 있도록,
 - 송수신 IP 주소와 같은 기본적인 정보로만 피처 엔지니어링을 수행함.
- CICIDS2017 데이터셋에서 13개의 피처를 선택하였으며,
 - 수신 상태를 확인하기 위해 SYN Flag와 ACK Flag를 동시에 카운트하는 SYN_ACK Flag를 재구성하여 추가함.



II. 침입 탐지 시스템 (Intrusion Detection System: IDS)

2.3 모델 성능

- 데이터셋을 8:2 비율로 랜덤하게 분할하여, Random Forest 알고리즘으로 학습 및 테스트를 수행함.
- 성능 지표 : 정확도(Accuracy), 정밀도(Precision), 재현율(Recall), F1-Score
 - 공격을 공격으로 식별했을 때 : True Positive(TP), 정상을 정상으로 식별했을 때 : True Negative(TN)
- 실험 결과: Accuracy 99.99%, Precision 99.82%, Recall 98.27%, F1-Score 98.95%

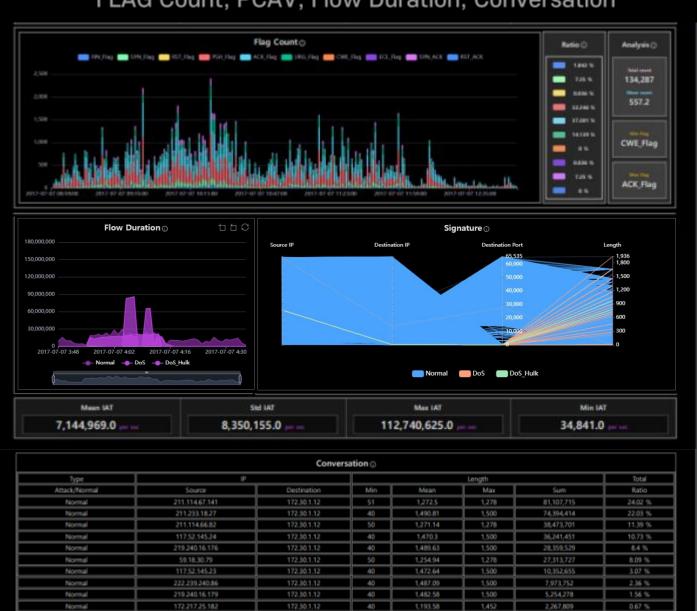
데이터 시각화의 상위 영역은 네트워크 트래픽의 전체적인 통계와 흐름을 파악할 수 있으며, 하위 영역은 데이터를 세분화하여 공격유형을 직관적으로 식별할 수 있음. 이를 통해 네트워크 보안상황을 종합적으로 파악할 수 있음.

네트워크, 정상/이상 통계, Length 빈도, BPS, PPS, Protocol, QPS, RPS

[상위 영역]

Data analysis was successful. Check out the Dashboard Date: 2023-11-24 00:32:38 TCP 21 0% --- Normal --- DoS --- DoS_Hulk TCP 465 0% TCP 389 0% MEAN VALUES Type: TCP Ratio: 90.51% Type: UDP Ratio: 9.489 QPS: 1.63 kg/s RPS: 8.24 kr/s QPS: 0.22 g/s RPS: 2.14 ler/s

[하위 영역] FLAG Count, PCAV, Flow Duration, Conversation



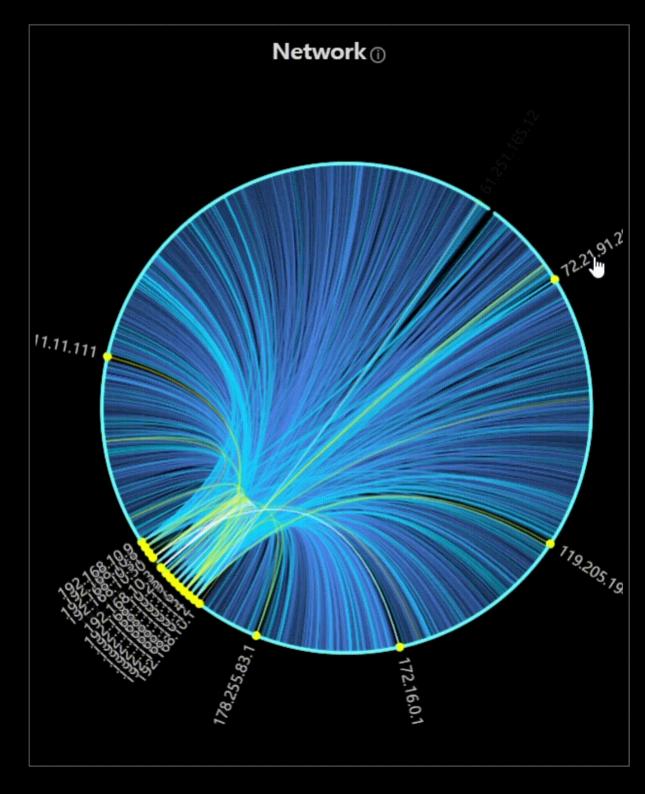
3.1 Line Chart

- 네트워크 트래픽 흐름을 파악하기 위해 시간에 따른 BPS, PPS, QPS, RPS를 Line Chart로 나타냄.
 - 투명도 차이를 활용한 정상 및 공격 데이터 구분
 - BPS : 초당 바이트 수, PPS : 초당 패킷 수
 - QPS: 초당 DNS 쿼리, RPS: HTTP(S) 요청 수
- 보안 이벤트가 언제 발생했는지 직관적으로 검출할 수 있음.



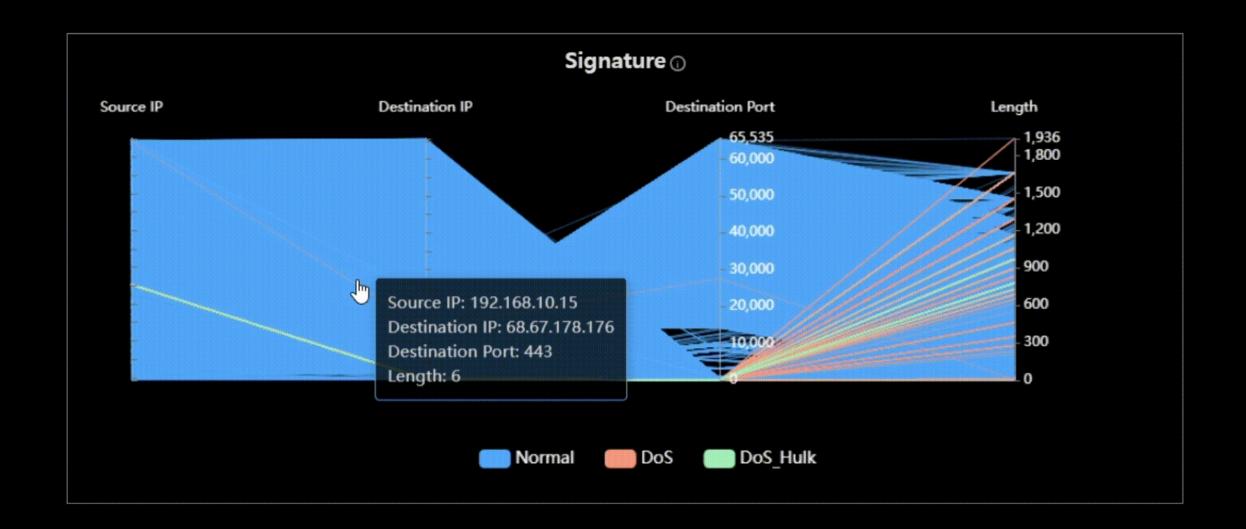
3.2 Circle Network

- 전체적인 IP 주소 간의 관계를 Circle Network 형태로 나타냄.
 - IP 주소 : 노드, 교류량 : 노드와 노드 사이 연결선
 - 교류량이 일정 임계치를 초과하면 다음과 같이 표시하였음.
 - 상위 노드 : 노란색, 연결선 : 파랑 -> 노랑 -> 흰색
- 특정 노드를 선택할 경우,
 - 해당 IP 주소와 트래픽 교류량을 구체적으로 확인할 수 있음.



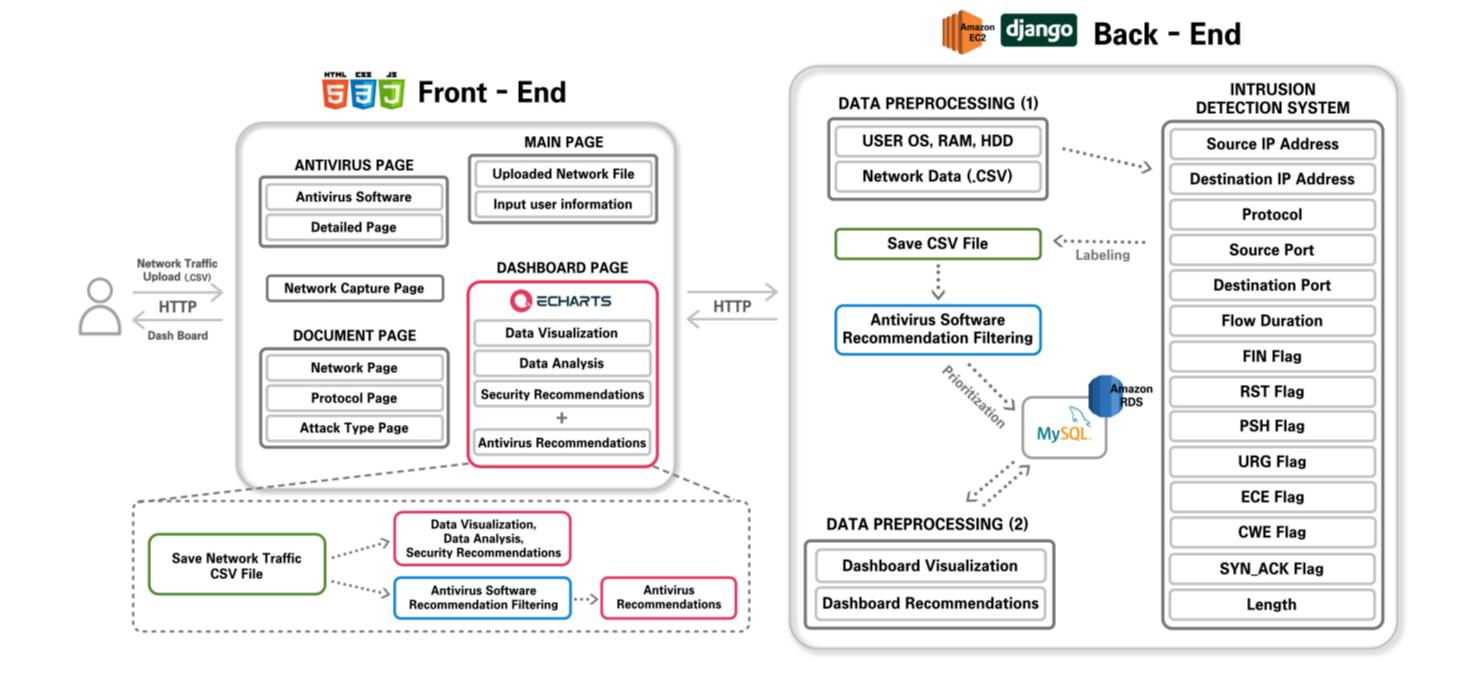
3.3 Parallel Coordinates Attack Visualization

- Parallel Coordinates Attack Visualization(PCAV)을 활용하여 복잡한 네트워크 데이터를 간단하게 시각화함.
 - PCAV는 Source IP, Destination IP, Destination Port, Packet Length를 x축에 두며,
 - 공격 데이터의 경우 특정 패턴을 형성함.



IV. 구현 (Implement)

본 프레임워크는 입력한 네트워크 트래픽을 Pandas로 전처리하고, 오픈소스 EChart를 통해 대시보드를 구성함.



IV. 구현 (Implement)

4.1 Dashboard

- 대시보드는 데이터 시각화, 데이터 분석, 보안 솔루션 및 백신 추천으로 이루어짐.
 - 해당 내용은 보고서 형태로 PDF 로 다운로드 할 수 있음.

4.1.1 데이터 분석

- 전체 데이터(Data Analysis)를 통계 분석한 후,
 - 공격 유형이 식별되면 해당 공격에 대한 분석표를 제공함.

종류		설명			
IP 주소	Source	송신 IP 주소			
IP 72	Destination	수신 IP 주소			
Length	Mean	트래픽 길이의 평균값			
	Min	트래픽 길이의 최솟값			
	Max	트래픽 길이의 최댓값			
	Sum	트래픽 길이의 총합			
Time		트래픽이 들어온 시간			

Data Analysis

대시 보드를 통한 데이터 분석을 해드러겠습니다

전체 데이터 중 정상 데이터는 34.3 %, 이상 데이터는 65.7 % 앱니다. 이상데이터는 DDoS 65.616 %, DoS Hulk 0.037 %으로 이루어져 있습니다. 전체 패킷 필어의 평균 값은 577.2 bytes 이때, 표 한편차, 희닷값, 희솟값은 작곡 571.73 bytes, 1,936 bytes, 0 bytes입니다. 패킷 끝이 벤도는 0~10 구간이 68,420 개로 가장 높은 비율을 차지하였습니다.

8PS의 최고값은 7,043,119 bytes/s 이며, 해당 시간은 2017-07-07 03:57:00 입니다. PPS의 최고값은 8,243 packets/s 이며, 해당 시간은 2017-07-07 03:57:00 입니다. CPS의 최고값은 1,627 queries/s 이며, 해당 시간은 2017-07-07 04:13:00 입니다. RPS의 최고값은 8,243 rquests/s 이며, 해당 시간은 2017-07-07 03:57:00 입니다. 따라서 해당 시간대의 네트워크 트랙틱을 확인해보세 이

FLAG는 중 48 개 이때, 조당 평균 개수는 4,549.5 입니다. 가장 높은 비율을 자지하는 FLAG는 ACK_Flag. 가장 낮은 비율을 자지하는 FLAG는 CWE_Flag 입니다. 전체 비율은 작각 FIN_Flag : 0.153 %, SYN_Flag : 2.73 %, RST_Flag : 0.006 %, PSH_Flag : 32.735 %, ACK_Flag : 48.334 %, URG_Flag : 13.306 %, CWE_Flag : 0.% FCE_Flag : 0.006 %, SYN_ACK : 2.73 %, RST_ACK : 0 % 입니다.

네트워크 트레틱에서 교류량이 가장 높은 Source IP 와 Declination IP는 각자 172.16.0.1. 192.168.10.50 이며, 횟수는 127.951 메달합니다. 해당 때첫 집에의 출함은 94.240.927 bytes 이며, 평균은 736.54 bytes, 취대는 1,936 bytes, 취소는 2 bytes, 비율은 83.73 % 입니다. Protocol은 TCP 가 90.51%의 비율로 가장 많이 차지합니다. IAN도착 자연 시간)의 평균은 17,342,566.4, 표준 변차는 17,032.892.6, 화닷값은 85,934,092.9, 최댓값의 해당 시간은 2017-07-07 04:07:50 입니다.

OR DORSE WHEN SHEET WAS ARRESTED FOR

현재 이상 데이터 ODBS 는 2017-07-07 03-56:00 - 2017-07-07 04:16:00 에서 발견 되었으며, 현재 데이터의 최고감 8PS는 7,043,119 bytes/s, PPS는 8,243 packets/s, QPS는 0 queries/s, RPS는 8,243 rquests/s입니다. 이상 데이터로 찍히는 Source IP는 172.16.0.1 이미, Destination IP는 192.168.10.50 입니다.

아래 표는 (Block 공격이 탐지된 데이터의 정보입니다. 이래 표를 확인하여, 해당 IP를 차단하세요

IP.		Port		Length				
Source	Destination	Soutte	Destination	Mean	Min	Max	Sum	Time
172.16.0.1	192.168.10.50	50499	80	279	930.33	1,453	2,791	2017-07-07 04:05:00
172.16.0.1	192.168.10.50	59111	80	345	923.67	1,453	2,771	2017-07-07-04:10:00
172.16.0.1	192,168,10.50	533662 2	80	833	1,384.5	1,936	2,769	2017-07-07 03:57:00
172.16.0.1	192.168.10.50	62176	80	462	915.67	1,453	2,747	2017-07-07 04:12:00
172.16.0.1	192.168.10.50	62763	80	462	883.0	1,291	2,649	2017-07-07 04:03:00
172.16.0.1	192.168.10.50	59095	80	279	876.33	1,291	2,629	2017-07-07 04:10:00
172.16.0.1	192(168.10.50	627623	80	279	876.0	1,291	2,628	2017-07-07 04:03:00
172.16.0.1	192.168.10.50	59857	80	279	875.67	1,453	2,627	2017-07-07 04:01:00
172.16.0.1	192,168,10.50	54135	80	279	875.67	1,452	2,627	2017-07-07 04:08:00
172.16.0.1	192.168.10.50	53442	80	1,164	1,309.0	1,454	2,618	2017-07-07 03:57:00

현재 이상 데이터 (bit. bial) 는 2017-07-07 03:56:00 - 2017-07-07 04:15:00 에서 발견 되었으며, 현재 데이터의 최고값 BFS는 24,281 bytes/s, PPS는 31 packets/s, QPS는 0 queries/s, RPS는 31 rquests/s입니다. 이상 데이터로 찍히는 Source IP는 172.16.0.1 이에, Destination IP는 192.168.10.50 입니다.

아래 표는 'Shill Halls' 공격이 합지된 데이타의 정보입니다. 아래 표를 확인하여, 해당 IP를 차단하세요

IP		Port		Length				
Source	Destination	Source	Destination	Mean	Min	Max	Sum	Time
172.16.0.1	192.168.10.50	54783	80	1,660	1,660.0	1,660	1,660	2017-07-07 03:58:00
172.16.0.1	192,168,10.50	42235	80	1,162	1,162,0	1,162	1,162	2017-07-07 04:15:00
172.16.0.1	192.168.10.50	41430	80	1,162	1,162.0	1,162	1,162	2017-07-07-04:15:00
172.16.0.1	192,168,10.50	51981	80	1,058	1,058.0	1,05H	1,058	2017-07-07 03:56:00
172.16.0.1	192,168,10.50	58765	80	971	971.0	971	971	2017-07-07-04:10:00
172.16.0.1	192,168,10.50	49771	80	971	971.0	971	971	2017-07-07 04:05:00
172.16.0.1	192.168.10.50	50445	80	971	971.0	971	971	2017-07-07 04:05:00
172.16.0.1	192.168,10.50	52261	80	971	971.0	971	971	2017-07-07-04:06:00
172.16.0.1	192.168.10.50	52781	80	971	971.0	921	971	2017-07-07 04:07:00
172.16.0.1	192.168.10.50	53927	80	968	968.0	968	968	2017-07-07 03:58:00

보선 경기에 되로 ID 전성을 함께하시라면 다듬가 같은 사이트 이용을 존재하니?

1. Criminal IP ()는 개인이나 기업의 사이비 자산과 관련된 취약점을 실시간으로 참지하고, 선제적인 대용을 가능하게 하는 종합 위험 인텔리전스 검색 엔진입니다. 2. Shodan ()은 사용자가 다양한 필타를 사용하여 안타넷에 연관된 다양한 유형의 서비를 검색할 수 있는 검색 엔진입니다. 해당 ip 검색을 통해 보안을 강화해보세요.

IV. 구현 (Implement)

4.1.2 보안 솔루션 및 백신 추천

- 공격이 식별되면 해당 공격에 대한 보안솔루션을 제시함.
 - 공격 데이터가 없는 경우, 백신 업데이트 등 가장 기본적인 보 안 솔루션을 제공함.
- 백신은 분석 데이터를 기반으로 연관성을 계산하여 추천함.
 - 분석 데이터는 네트워크 트래픽정보, 사용 유형(개인용, 가정용 등), 사용 환경(OS, RAM, HDD), 가격 등으로 구성됨.

Security Recommendations

ON TROUGH PRICES ME SHORTERS A WALNOT AN USE SAMERICE MAALE

DDoS 공격 유형이 팀지되었습니다. DDoS 공격에 대한 보안 권장 사항은 다음과 같습니다.

1. KISA 사이버 대피소를 이용하세요. DDoS 공격이 발생하면, KISA에서 제공하는 사이버 대피소를 이용하여 웹사이트로 향하는 DDoS 트레픽을 대피소로 우회하여 분석 및 자단하여 웹사이트를 정성 적으로 운영할 수 있도록 합니다. 이 서비스는 중소기업 무료지원 서비스로 제공되며, 에서 신청할 수 있습니다.

백업 서비를 구축하세요. DDcS 공격에 의한 서비 장애를 대비하여 중요서비들은 서로 다른 회전에 이중화 구성을 해야 합니다. 서로 다른 회전을 이용해서 이중화를 구성하면 Master 서비의 회전에 DDcS 공격으로 인한 장애가 발생해도, 다른 회전에 있는 Slave 서비가 동작하여 서비스 장애를 최소화 할 수 있습니다.

3. 서배를 중성하세요. 서배 중성을 통해. 서배스 용량을 확보하고 부하를 분산시킴으로써 DDos 공격에 대응할 수 있습니다. 이를 통해 웹사이트나 서비스의 가용성을 유지할 수 있습니다.

DOCK CONTROL THE THEFT HER WITH DAZING MAY THE

4. 공격 대상을 최소화하세요. 외부에 노출된 웹서비스 외 기업 내부용 서비는 외부에 노출되지 않도록 내부랑으로 망 분리 조치 후 운영하세요. 내부용 서비는 DDoS 공격 및 기타 사이비 공격을 예방 하기 위해 기업의 내부 서비 IP 및 서비스가 외부에 OPEN 되어 있는지 주기적인 스케닝을 시행하고, shodan 또는 다크 웹 노출 여부를 확인하세요. 외부에서 부득이 하게 내부망으로 접속이 필요할 경우에는 DDoS 대용 및 방어 설정이 가능한 가상사설망 (VPN) 등의 별도 서비스를 이용하세요.

홍막 : 2023 달기사 학보보인가서 찢지 CE 마운?

PortScan 공격 유형이 탐지되었습니다. 포트 스캔 자체는 단순한 조사이미 이것만으로 실제 손해가 발생하진 않습니다. 하지만 시스템에 취약점이 있다면 포트 스캔으로 유출된 정보를 통해 공격을 받아 피해를 볼 우리가 존재합니다. PortScan 공격에 대한 보안 권장 사랑은 다음과 같습니다.

1. 포트스캔은 보통 자동화된 프로그램을 이용하여 수행되므로, 백신 프로그램을 통해 포트스캔을 수행한 프로그램을 찾아 삭제해야 합니다.

포트벤호는 65535까지의 청수로 표현되며, 컴퓨터 통신 시 IP와 포트벤호를 알아야 상호 통신이 가능합니다. 포트벤호는 특정 프로그램을 사용하기 위해 정해진 포트가 있으며 대표적으로 80(홈페이 지L 21(FTP), 25(SMTP) 등입니다.

Antivirus Sotfware Recommendations

A를 사용한 맞충형 백신 제품을 추천해드리도록 하겠습니다







V. 결론 (Conclusions)

- 본 논문은 네트워크 보안 상황을 인지할 수 있는 웹 기반 네트워크 트래픽 분석 솔루션을 구현함.
- 본 프레임워크는 머신러닝 분류 알고리즘인 Random Forest를 활용하여 네트워크 침입을 12가지 공격을 다양으로 식별 및 분류하고, 다양한 보안 이벤트 시각화 기술을 사용하여 네트워크상의 여러 공격을 직관적으로 파악할 수 있음.
- 또한, 공격 분석표 및 보안 솔루션을 제공하여 네트워크 보안 상황을 인지하고 대응할 수 있도록 구성함.

참고 문헌

[1] 박재범, 김휘강, 김은진.(2014).대규모 네트워크의 효과적 보안상황 인지를 위한 벌집 구조 시각화 시스템의 설계 및 구현.정보보호학회논 문지,24(6),1197-1213.

- [2] H. Shiravi, A. Shiravi and A. A.Ghorbani, "A Survey of Visualization Systems for Network Security," in IEEE Transactions on Visualization and Computer Graphics, vol. 18, no. 8, pp.1313-1329, Aug. 2012, doi: 10.1109/TVCG.2011.144.
- [3] 정치윤, 손선경, 장범환 and 나중찬. (2009). 시각화 기반의 효율적인 네트워크 보안상황 분석 방법. 정보보호학회논문지, 19(3), 107-117.
- [4] Iman Sharafaldin, Arash Habibi Lashkari, and Ali A. Ghorbani, "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization", 4th International Conference on Information Systems Security and Privacy (ICISSP), Portugal, January 2018
- [5] Hyunsang Choi, Heejo Lee, Hyogon Kim, Fast detection and visualization of network attacks on parallel coordinates, Computers & Security, Volume 28, Issue 5, 2009, Pages 276-288, ISSN 0167-4048, https://doi.org/10.1016/j.cose.2008.12.003
- [6] K. hun Han and H. K. Kim, "FDANT-PCSV: Fast Detection of Abnormal Network Traffic Using Parallel Coordinates and Sankey Visualization," Journal of the Korea Institute of Information Security & Cryptology, vol. 30, no. 4, pp. 693–704, Aug. 2020.

END

DO-Hee Kang, Yu-Jin Kim, Ye-Rim Yeon and Byung-Il Kwak Hallym Univ, Data-Driven Cybersecurity Research LAB

