DGA 봇넷 도메인 감지 및 패밀리 분류 연구 동향

이정민[1], 강민재[2], 이연준[3]
[1]한양대학교 컴퓨터공학과 바이오인공지능융합전공 (대학원생)
[2]한양대학교 ERICA 국방정보공학과 (학부생)
[3]한양대학교 컴퓨터공학과 바이오인공지능융합전공 (교수)
lsmp12@hanyang.ac.kr, minjae0110@hanyang.ac.kr, yeonjoonlee@hanyang.ac.kr

Survey on DGA Botnet Domain Detection and Family Classification

Jungmin Lee, Minjae Kang, Yeonjoon Lee
Major in Bio Artificial Intelligence, Dept. of Computer Science & Engineering, Hanyang University (Graduate Student)
Dept. of Military Information Engineering, Hanyang University ERICA (Undergraduate Student)
Major in Bio Artificial Intelligence, Dept. of Computer Science & Engineering, Hanyang University (Professor)

요 약

봇넷은 지속적으로 사이버 범죄에 이용되고 있으며 네트워크 환경에 큰 위협이 되고 있다. 기존에는 봇들이 C&C 서버와 통신하는 것을 방지하기 위해 블랙리스트를 기반으로 DNS 서버에서 봇넷 도메인을 탐지하는 방식을 주로 사용하였다. 그러나 도메인 생성 알고리즘(DGA)을 이용하는 봇넷이 증가하면서 기존에 사용하던 블랙리스트 기반의 도메인 차단 방식으로는 더 이상 봇넷 도메인을 효율적으로 차단하기 어려워졌다. 이에 따라 봇넷 도메인 생성 알고리즘을 통해 생성되는 도메인의 특성을 분석하고 이를 토대로 봇넷 도메인을 식별하고 차단하고자 하는 시도가 계속되고 있다. 특히 연속적인 데이터 처리에 주로 사용되는 딥러닝 알고리즘을 이용하여 봇넷 도메인의 특징을 효과적으로 추출하고 정확도가 높은 탐지 모델을 구축하고자 하는 연구가 주를 이루고 있으며, 탐지뿐만 아니라 봇넷 그룹(Family) 분류까지 연구가 확장되고 있다. 이에 본 논문에서는 봇넷 도메인 생성 알고리즘에 의해 생성되는 봇넷 도메인을 식별 및 분류하기 위해 딥러닝 기술을 적용한 최근 연구 동향을 조사하고 앞으로의 연구 방향성을 논의하고자 한다.

1. 서론

봇(Bot)은 사이버 범죄자가 제작한 악성 소프트웨어에 감염된 인터넷 기반 기기를 가리키며, 봇넷(Botnet)은 이러한 기기들이 형성하는 네트워크를 의미한다. 과거에는 주로 컴퓨터만이 봇넷을 형성했지만, 최근에는 사물 인터넷(IoT) 기기들이 보급되면서 다양한 기기들이 봇으로 감염되고 봇넷의 규모가 커지고 있다. 봇넷은 C&C 혹은 C2 (Command and Control) 서버로 불리는 중앙서버와 다수의 봇으로 이루어져 있으며, 이 서버를 통해 사이버 범죄자는 봇으로부터 수집된 정보를 얻거나 봇에게 명령을 전송한다.

이러한 메커니즘을 이용하여 기존에는 봇이 C&C 서버와 통신하는 것을 막기 위해 DNS 서버에서 차단을 시도하였다. DNS 요청에 알려진 C&C 도메인 주소가 포함되면 해당 요청을 차단하는 블랙리스트 기반의 방식을 이용하였다. 그러나 도메인 생성 알고리즘을 사용하는 봇넷이 등장하면서 기존의 도메인 차단 방식과 다른 효율적인 방식의 필요성이 대두되었다.

도메인 생성 알고리즘(DGA, Domain Generation Algorithm) 기반의 봇넷은 악성 소프트웨어에 도메인 생성 알고리즘 코드가 추가되어 봇이 무수히 많은 도메인을 생성하는 봇넷을 말한다. 사이버 범죄자는 DGA 코드를 알고 있으므로 봇이 어떤 도메인을 생성할 것인지 알 수 있으며, 이 중 하나만을

C&C 서버 도메인으로 등록하면 된다. 하지만 DNS 서버는 어떤 도메인이 실제 C&C 서버의 도메인인지 알 수 없고, DGA에 의해 생성되는 모든 도메인을 블랙리스트에 추가하는 것은 비현실적인 방법이며, 알고리즘의 초깃값이 계속해서 변화하기 때문에 생성될 도메인을 유추하는 것은 쉽지 않은 일이다. 따라서 DGA 기반 도메인을 식별하기 위한 대안의 필요성이 제기되었다.

이에 봇넷 DGA에 의해 생성되는 도메인의 특성을 분석하기 위해 딥러닝 기술을 이용하는 연구들이 진행되고 있다. 일반적인 도메인과 다르게 봇넷 도메인들은 발음이 불가하거나 아무 의미를 내포하지 않는 복잡한 문자열로 구성되어있어 딥러닝 기술을 이용하여 일반적인 도메인과 구분이 쉽게 가능하다. 하지만 어떤 봇넷의 도메인인지 구분하는 것은 여전히 해결해야 할 문제점이다.

같은 공격 목표, 공격 패턴, 악성 코드를 공유하는 봇넷들을 일반적으로 하나의 봇넷 그룹(Family)으로 간주한다. 이들은 주로 같은 범죄자 혹은 범죄 그룹에 의해 관리 및 사용되며 앞서 얘기한 것처럼 많은 유사성을 가지고 있다. 봇넷 그룹마다 대응 방식이 다르므로 봇넷을 식별하는 것 외에도 어떤 봇넷 그룹에 속하는지 분류하는 것 또한 중요한 작업이다. 이에 따라 같은 봇넷 그룹은 같은 DGA를 공유한다는 특징을 기반으로, 생성된 도메인이 어떤 봇넷 그룹에 속하는지 분류하기 위한 연구가 계속되고 있다. 대부분의 봇넷 그룹이 생성하는 도메인이 서로 비슷한 형태를 띠고 있어 분류가 쉽지만은 않으나, 이를 해결하기 위해 연속적인 데이터 처리에 사용되는 딥러닝 알고리즘을 기반으로 각 봇넷 그룹 도메인의 특징 및 패턴을 찾으려는 시도가 계속되고 있다.

본 논문에서는 이와 관련된 연구들을 통해 DGA 기반의 봇넷 도메인 식별 및 분류 연구의 동향에 대해 알아보고, 앞으로의 연구 방향성에 대해 논의하고자 한다.

2. 본론

DGA 봇넷의 등장으로 기존 봇넷 탐지 방식으로는 실질적인 방어가 어려워지면서 새로운 대안의 필요성이 제기되었다. 이에 도메인 문자열 자체를 분석하여 식별 및 분류하고자 하는 연구가 진행되었다. 도메인은 알파벳, 숫자, 붙임표(-)로 이루어져 있다는 점을 기반으로, 자연어와 같은 연속적인 데이터 처리에 사용되는 딥러닝 알고리즘을 이용하고자 하는 시도가 계속되고 있다. 대표적인 알고리즘으로는 순환 신경망(RNN, Recurrent Neural Network), 장단기 메모리(LSTM, Long Short Term Memory), 어텐션 메커니즘(Attention Mechanism), 트랜스포머(Transformer)가 있으며, 이 알고리즘들은 기존 알고리즘의 단점을 완화하고자 하나의 알고리즘에서 계속해서 진화해왔다. DGA 봇넷 탐지 연구도 이러한 알고리즘들의 발전 방향과 비슷하게 흘러가고 있으며 이를 토대로 연구 흐름을 알아보고자 한다.

RNN은 텍스트와 시계열 데이터와 같이 순서가 존재하는 데이터를 처리하기 위해서 설계된 신경망이다. 내부에 순환 구조가 있어 이전 시점의 정보를 현재 시점의 입력과 함께 처리할 수 있으며 이로써 순서 정보를 학습할 수 있다는 장점이 있다. 그러나 긴 입력 데이터에 대해서는 정보 기억력이 하락하는 장기의존성 문제가 있었다.

LSTM은 RNN의 장기 의존성 문제를 해결하기 위한 알고리즘으로, 입력, 망각, 출력 게이트로 구성되는 구조 덕분에 RNN보다 긴 데이터를 더 잘 다룰 수 있게 되었다. [1]은 LSTM기반 DGA 봇넷 분류 프레임워크(LSTM.MI)를 제시하였다. 제시된 프레임워크는 먼저 이진 분류를 통해 DGA 봇넷 도메인 여부를 식별한 뒤 DGA 봇넷 도메인을 대상으로 LSTM 기반의 분류 모델을 이용하여 어떤 봇넷 그룹에 속하는지 분류하였다. 또한 데이터셋의 불균형 문제를 완화하기 위해 LSTM의 역전파 학습 메커니즘에 일종의 가중치인 비용 항목을 추가하였다. 샘플 수가 적은 클래스의 비용 가중치를 크게 하면 모델이 해당 클래스 오분류에 좀 더 민감하게 반응하게 되고, 이로써 모델의 편향성을 조절하여 더 균형잡힌 훈련 및 예측을 수행할 수 있었다. [2]는 양방향 LSTM을 이용하여 복잡한 도메인 내에서 연속적인 패턴을 찾고자 하였다. 양방향 LSTM은 한 방향으로만 학습을 진행하는 기존 LSTM과 달리 역방향으로도 학습을 진행하여 하나의 훈련 시점에서 과거의 정보와 미래의 정보를 모두 가질 수 있었다. 덕분에 좀 더 세밀한 특징을 이용하여 봇넷 그룹 분류를 수행할 수 있었고, 같은 DGA 알고리즘의 서로 다른 초깃값으로부터 생성된 도메인끼리도 클러스터링할 수 있었다. 이 연구에서 구축된 모델은 실제 ISP 중 한 곳에서 사용되었으며 실제 DNS 요청을 클러스터링함으로써 기존에 알려지지 않은 5개의 DGA 봇넷 그룹을 발견하였다. [3]은 DGA 봇넷 식

별을 하기 전 피싱 사이트와 같은 다른 악성 도메인을 걸러내기 위해 전처리 단계에 샴 신경망(Siamese Neural Network)을 도입했다. 샴 신경망은 같은 가중치를 가지는 두 개의 네트워크로 구성되며, 서로 다른 두 개의 데이터에 대한 각 네트워크 출력값의 유사성을 비교하여 데이터를 분류하는 모델이다. 이로써 정상 도메인과 유사한 형태를 띠는 악성 도메인을 유사도 측정을 통해 일차적으로 걸러낼 수 있었고 덕분에 DGA 봇넷 식별 정확도를 높일 수 있었다. DGA 봇넷 식별 및 분류 단계에서는 양방향 LSTM을 도입했다. 양방향 LSTM을 통해 과거의 정보가 먼 미래까지 전달되지 못하는 장기의존성 문제를 해결하였고 높은 정확도로 봇넷 그룹 도메인을 분류할 수 있었다.

LSTM이 긴 데이터를 처리하는 데 있어서 RNN보다 더 나은 성능을 보였으나 여전히 한계가 있었다. 더불어 학습 속도를 높이기 위해 데이터를 병렬적으로 처리하는 방법이 필요했다. 이를 해결하기 위해 Attention 기반의 Transformer 알고리즘이 등장했다. Attention은 인접한 요소 외에 다른 요소와도 상호작용하도록 하는 메커니즘으로, 이 덕분에 Transformer는 장기 의존성 문제를 크게 해소할 수 있었다. 더불어 긴 데이터 내에서의 요소 위치 정보를 병렬적으로 처리할 수 있어 빠른 학습 속도를 가질 수 있었다. [4]는 처음으로 트랜스포머 네트워크를 DGA 봇넷 탐지 모델에 적용했다. 단일 길이 임베딩에만 초점을 두었던 기존 연구들과 다르게 바이그램 단위의 임베딩을 결합하였고 트랜스포머의 인코더를 이용하여 다양한 크기에서 특징을 추출하였다. 봇넷 그룹 분류에 있어서 기존의 연구들보다 정확도가 향상되었고, 분류에 실패하는 그룹의 개수가 감소하였다.

DGA 봇넷 도메인 데이터셋을 생성하는 연구도 진행되고 있다. [5]는 기존의 봇넷 도메인 분류 연구를 서로 비교할 수 있는 정형화된 데이터셋이 없다는 문제점을 해결하고자 50개의 봇넷 그룹을 포함하는 공개 데이터셋을 구축하였다. 저자는 데이터셋의 봇넷 그룹 다양성, 도메인 양, 문서화 등의 9가지 기준을 충족하고자 하였다. 공개된 봇넷 도메인 생성 알고리즘을 이용하여 그룹마다 최대 1,000,000개의 도메인을 생성하였으며, 최대 10,000개의 도메인에 대해서 머신러닝 모델에 적용할 수 있는 특징을 추출하였다. [6]은 봇넷의 변종이 계속해서 등장하고 있음을 문제점으로 지적하고 기존 데이터셋의 업데이트가 필요함을 언급하며 새로운 데이터셋을 제안하였다. 76개의 봇넷 그룹을 포함하며, 데이터를 다양한 파일 형식으로 제공하기 때문에 특정 파일 형식을 지원하는 다양한 플랫폼에서 이용할 수 있다. 이 데이터셋 역시 머신러닝 모델에 적용할 수 있도록 도메인에서 추출한 특징값들을 제공한다.

3. 결론

연속적인 데이터 처리 기술 기반의 딥러닝 모델을 적용함으로써 복잡한 문자열 사이의 패턴을 학습할 수 있게 되었고 이는 봇넷 도메인의 식별 및 분류를 가능하게 만들었다. 하지만 여전히 일부 도메인은 식별 및 분류가 잘 이루어지지 않고 있다. 이러한 도메인들은 발음이 가능한 단어가 섞여 있거나 정상 도메인과 유사한 문자 패턴을 보여 식별이 쉽지 않다. 혹은 도메인 길이가 너무 짧거나 여러 DGA가 섞인 코드에서 생성된 도메인으로, 특징을 추출하기 어려워 분류 또한 쉽게 이루어지지 않고 있다. 이러한 도메인을 다루기 위해서는 더 세밀하고 복잡한 패턴 및 특징을 추출하기 위한 고도화된 인공지능 기술이 필요할 것으로 보인다.

더불어, 계속해서 등장하는 새로운 봇넷 그룹을 즉각적으로 모델에 학습시킬 방법을 고민해야 할 것이다. 봇넷은 지속적으로 새롭게 등장하고 있으며 이 때문에 봇넷 데이터셋 또한 주기적으로 구축되고 있다. 이러한 비효율성을 해결하기 위해 새로운 DGA 봇넷 그룹의 도메인이 발견되었을 때 적은 양으로도 해당 봇넷 그룹을 식별 및 분류할 수 있도록 사전 학습과 미세 조정을 적용하는 등 빠르게 새로운 봇넷 그룹을 분류할 수 있는 모델에 관한 연구가 필요할 것이다.

마지막으로, 딥러닝이 아닌 머신러닝을 이용한 연구도 지속될 것으로 보인다. 데이터셋을 구축하는 연구들은 공통적으로 연구에서 추출한 특징의 머신러닝 모델 적합성을 입증하고 있으며, 여전히 머신러닝을 이용한 연구를 도모하고 있다. 딥러닝이 머신러닝보다 더 좋은 성능을 가질 수 있으나 데이터셋과 관련된 문제나 컴퓨팅 자원의 한계, 모델의 복잡도 등 여전히 한계점이 존재한다. 따라서 특징 선택 기술을 이용하여 데이터셋이 제공하는 특징들을 분석하고 모델 성능 향상에 도움이 되는 특징 집합을 찾는 등 머신러닝을 이용하여 DGA 봇넷 도메인 식별 및 분류 정확도를 높이는 연구도 지속될 것으로 보인다.

ACKNOWLEDGEMENT

이 논문은 2022년도 정부(방위사업청)의 재원으로 국방기술진흥연구소의 지원을 받아 수행된 연구임 (No. KRIT-CT-22-021, 우주공간 신호정보 특화연구실)

참고문헌

[1] Tran, Duc, et al. "A LSTM based framework for handling multiclass imbalance in DGA botnet detection." Neurocomputing. 275. 1. p. 2401-2413. 2018

[2] Sidi, Lior, et al. "Helix: DGA domain embeddings for tracking and exploring botnets." Proceedings of the 29th ACM International Conference on Information & Knowledge Management. 2020.

[3] Ravi, Vinayakumar, et al. "Adversarial defense: DGA-based botnets and DNS homographs detection through integrated deep learning." IEEE transactions on engineering management. 70. 1. p. 249-266. 2023

[4] Ding, Ling, et al. "Botnet DGA Domain Name Classification Using Transformer Network with Hybrid Embedding." Big Data Research 33 (2023): 100395.

[5] Zago, Mattia, Manuel Gil Pérez, and Gregorio Martínez Pérez. "UMUDGA: A dataset for profiling DGA-based botnet." Computers & Security 92 (2020): 101719.

[6] Tuan, Tong Anh, et al. "UTL_DGA22-a dataset for DGA botnet detection and classification." Computer Networks 221 (2023): 109508.