https://doi.org/10.33778/kcsa.2022.22.5.145

N-gram을 활용한 DGA 기반의 봇넷 탐지 방안★

정 일 옥*, 신 덕 하**, 김 수 철***, 이 록 석****

요 약

최근 봇넷의 광범위한 확산과 고도의 정교함은 기업과 사용자뿐만 아니라 국가 간 사이버전에도 심각한 결과를 초래하고 있다. 이 때문에 봇넷을 탐지하고자 하는 연구는 꾸준히 되고 있다. 하지만, DGA 기반의 봇넷은 기존의 시그니처 및 통계 기반의 기술로는 탐지율은 높지만, 오탐율 또한 높은 한계가 있다. 이에 본 논문에서는 DGA 기반의 봇넷을 탐지하고자 문자 기반의 n-gram을 활용한 탐지모델을 제안한다. 제안한 모델을 통해 기존의 탐지 기술의 한계인 탐지율을 높이고 오탐율을 최소화할 수 있다. 다양한 DGA 봇넷에서 사용하는 대규모의 도메인 데이터셋과 정상 도메인에 대한 실험을 통해 기존의 모델보다 성능이 우수함을 확인하였다. 제안된 모델의 오탐율은 2~4% 미만이며 전체 탐지 정확도와 F1 점수는 모두 97.5%임을 확인하였다. 이처럼 본 논문에서 제안한 모델을 통해 DGA 기반의 봇넷에 대한 탐지 및 대응 능력이 향상될 것을 기대한다.

DGA-based Botnet Detection Technology using N-gram

Jung Il Ok*, Shin Deok Ha**, Kim Su Chul***, Lee Rock Seok****

ABSTRACT

Recently, the widespread proliferation and high sophistication of botnets are having serious consequences not only for enterprises and users, but also for cyber warfare between countries. Therefore, research to detect botnets is steadily progressing. However, the DGA-based botnet has a high detection rate with the existing signature and statistics-based technology, but also has a high limit in the false positive rate. Therefore, in this paper, we propose a detection model using text-based n-gram to detect DGA-based botnets. Through the proposed model, the detection rate, which is the limit of the existing detection technology, can be increased and the false positive rate can also be minimized. Through experiments on large-scale domain datasets and normal domains used in various DGA botnets, it was confirmed that the performance was superior to that of the existing model. It was confirmed that the false positive rate of the proposed model is less than 2 to 4%, and the overall detection accuracy and F1 score are both 97.5%. As such, it is expected that the detection and response capabilities of DGA-based botnets will be improved through the model proposed in this paper.

Key words : DGA, Botnet, intrusion detection

접수일(2022년 09월 30일), 수정일(2022년 11월 30일), 게재확정일(2022년 12월 22일)

★ 본 논문은 2022년 정부(국토교통부)의 재원으로 국토교통과학기술진흥원(KAIA)의 지원을 받아 연구가 수행된 연구임(22TLRP-B152767-04, 자율협력주행 도로교통체계 통합보안시스템 운영을 위한 기술 및 제도개발)

* 고려대학교/정보보호학과 (주저자, 교신저자)
** 경희대학교/응용수학과(공동저자)
*** 숭실대학교/IT정책경영학과(공동저자)
**** 전남대학교/정보보호협동과정(공동저자)

1. 서 론

최근 봇넷은 지속적으로 증가하고 있으며, 이는 오늘날 주요 보안 사고의 주요 원인이 되고 있다. 이는 봇넷이 대규모 DDoS 공격, 악성 이메일, 스팸, 악성코드 전송 및 제어, 민감한 데이터 유출 등과 같은 다양한 공격과 관련이 있기 때문이다. 솔라윈즈(Solarwinds) 해킹 사건(2020.12)[1]에서 침투에 성공한 악성코드는 보안시스템 탐지를 우회하기 위해 DGA를 통해 C&C 서버로 접근을 시도하였으며, 러시아-우크라이나 전쟁에서도 수년 전부터 구축된 봇넷을 통해 가짜 뉴스와 허위 사실을 유포하거나 적군의 인프라를 교란하는 데 사용되었다고 한다[2][3]. 또한 다른 위험한 유형의 봇넷 지원 공격은 웹 삽입, URL 스푸핑, DNS 스푸핑 및 민감한 데이터 수집에 사용될 수 있으며, 봇넷을 활용한 공격의 주요 대상은 일반적으로 정부 기관이나 금융 등이 되고 있다[4].

봇넷에 관한 기술은 지속적으로 정교해지고 있다[5]. 일반적으로 봇넷은 봇이라고 하는 특수한 유형의 악성코드에 감염되어 인터넷으로 연결된 네트워크라 할 수 있다[6]. 봇은 일반적으로 봇마스터라고 하는 해킹 그룹에 의해 생성된다. 인터넷에 연결된 장치에서 실행되는 봇을 사용하면 봇마스터가 장치를 원격으로 제어할 수 있다. 봇에 감염된 장치는 컴퓨터, 스마트폰 또는 IoT 장치일 수도 있다. 봇은 고도로 자율적이며 C&C 등의 채널을 사용하여 제어 시스템에서 명령 및 코드를 업데이트하고 주기적으로 작업 상태를 제어 시스템으로 보낸다는 점에서 다른 유형의 악성코드와 매우 다르다[7]. 이러한 봇넷에 대한 피해를 최소화하기 위해 C&C 서버 IP 탐지하고 차단하는 다양한 솔루션과 연구가 진행되었으며, 이를 회피하기 위해서 봇넷에 DGA 기술을 적용하게 되었다[8].

이렇게 DGA 기반의 봇넷이 증가하면서 보안 담당자들은 봇넷을 탐지하고 차단하는 데 어려움을 겪고 있다. 이에 본 논문에서는 DGA 기반의 봇넷에 대한 탐지방안을 제안한다. 기존의 DGA에 대한 특성(피처) 추출과 통계적인 방법은 탐지율에 대해서 좋은 성능을 내었지만, 오탐율에 대한 이슈가 존재하였으며, 특성(피처)을 추출하는 데 전문가의 노력과 시간이 필요하였다. 본 논문에서는 이러한 부분을 개선한 Character level bigram을 적용한 DGA 기반의 봇넷 탐지모델을 제안한다[7].

제안된 탐지모델은 특성(피처) 추출에 Character level bigram을 사용하였으며, 이렇게 추출된 피처로 생성된 학습 데이터에 RF(Random Forest), XGBoost[9], NN(Neural Network) 등의 알고리즘을 적용하여 모델을 생성한다. 이렇게 생성된 모델은 기존의 모델보다 F1-score에서 좋은 성능을 도출할 수 있으며, 특히, 오탐율에 대한 부분을 최소화하는데 기여할 것이다.

본 논문의 구성은 다음과 같다. 2장에서는 봇넷 탐지, DGA에 대해서 알아보고, DGA 기반의 봇넷 탐지 관련 연구를 알아보고 3장에서는 본 논문에서 제안하는 기법에 관해 설명한다. 4장에서는 제안된 기법에 대한 실험 및 평가 결과에 대해 분석하였다. 5장에서는 결론 및 향후 연구 방향을 제시한다.

2. 관련 연구

본 장에서는 봇넷 탐지 및 DGA 기술과 DGA 기반의 봇넷 탐지에 관한 연구를 살펴본다.

2.1 봇넷 탐지와 DGA에 관한 연구

최근 몇 년 동안 봇넷 탐지에 관한 다양한 연구가 진행되고 있다[10]. 봇넷 탐지는 크게 호스트 기반과 네트워크 기반의 탐지로 나눌 수 있으며, 특히, 네트워크 기반 탐지는 시그니처 기반, DNS 기반, 트래픽 기반, 이상 기반 및 마이닝 기반 방법으로 나눈다. 이 가운데 DNS 기반 봇넷 탐지 기술[11]은 흐름 기반, 이상 기반, 플럭스 기반, DGA 기반 및 봇 감염 기반의 5가지 범주로 분류할 수 있다.

이 가운데 많은 봇넷이 DGA(Domain Generation Algorithm) 이라는 기술을 사용하여 C&C 서버에 서로 다른 도메인 이름을 자동으로 생성하고 등록하여 보안체계를 벗어나려 한다. 봇넷이 DGA 기술을 이용하는 주된 이유는 등록된 도메인 이름

의 제어 및 관리가 어렵게 하기 위한 것도 있다. 이러한 DGA 기술을 이용한 봇넷은 DGA 기반의 봇넷이라고 한다.

일반적으로 DGA 기술은 연, 월, 일과같이 지속적으로 변화하는 값을 갖는 변수에 연산자를 이용하여 임의의 도메인 이름을 생성한다. 다음은 DGA 기반의 봇넷 중 하나인 GameoverZeus 에 대한 설명이다. 해당 봇넷은 날짜 등의 값을 이용하여 해시 형식의 시드를 생성하고, 32글자의 시드를 8글자씩 쪼개어 Hex 값으로 변환 후 연산을 통해 출력된 수를 아스키코드에 대응(Mapping)하여 도메인을 구성한다. 기본적으로 알파벳과 숫자 형의 합으로 구성된다. 또한, 최상위 도메인으로는 '.com', '.net', 'biz', 'org'를 가지게 된다.

2.2 DGA 기반의 봇넷 탐지에 관한 연구

위에서 언급했듯이 봇넷의 봇은 로컬 DNS 서버를 활용하여 일상 활동에서 봇넷 C&C 서버의 IP 주소를 찾는다. 따라서 DNS 트래픽이나 추적을 모니터링하고 분석하면 봇넷의 활동을 감지하는데 도움이 될 수 있다. 이러한 원리를 기반으로 다양한 연구자들이 연구를 수행하였으며, 특히, Hoang and Nguyen (2018)[7], Qiao et al. (2019)[13], Hostiadi et al. (2020)[15] 등은 DNS 트래픽 기능을 사용하여 도메인 플럭스 봇넷을 탐지하는 방법을 제안한다. 그들은 도메인 길이와 예상값을 포함한 DNS 도메인 기능을 사용하여 일부 봇넷에서 생성된 합법적인 도메인 이름과 의사 무작위 도메인 이름(PDN)을 구별하였다. 도메인 이름의 예상값은 Alexa(DN Pedia, n.d)에서 순위를 매긴 가장 인기 있는 합법적인 도메인 이름 100,000개의 문자 분포를 기반으로 계산하였다. 실험 데이터 세트는 Alexa(DN Pedia, n.d)가 순위를 매긴 가장 인기 있는 합법적인 도메인 이름 100,000개와 Conficker 및 Zeus 봇넷에서 생성한 약 20,000개의 도메인 이름으로 구성하였다. Naive Bayes, KNN, SVN, DT, RF를 포함한 다양한 알고리즘을 사용하여 지도학습 기반의 봇넷 탐지모델을 구성하고 검증하였다. 실험 결과는 의사 결정 트리가 전체에서 가장 큰 값(92.30%의 탐지 정확도와 4.80%의 오탐율)을 제공하는 알고리즘임을 보여준다. 제안된 모델의 전체 탐지 정확도는 상대적으로 높지만, 오탐율이 높아 최적의 경우 총 약 7.70%를 보여준다. Hoang and Nguyen(2018)은 지도학습 기반의 머신러닝 기술을 사용하여 합법적인 도메인 이름과 봇넷 생성 도메인 이름의 분류를 기반으로 하는 DGA 봇넷 탐지모델을 제안하였다. 그들은 제안된 모델을 구성하고 검증하기 위해 16개의 n-gram 특징과 2개의 모음 분포 특징을 포함하여 18개의 도메인 특징을 사용할 것을 제안하였다. 16개의 n-gram 특성 중 8개의 특성은 각 도메인의 2-gram 부분 문자열을 기반으로 계산되고 나머지 8개의 특성은 도메인의 3-gram 부분 문자열을 기반으로 계산하였다. 실험 데이터 세트는 Alexa(DN Pedia, n.d)[13]에 의해 순위가 매겨진 30,000개의 정상 도메인 이름과 DGA 봇넷(Netlab 360, n.d)[12]에서 사용하는 30,000개의 악성 도메인 이름으로 구성된다. 나이브 베이즈(naive Bayes), kNN, 의사결정 트리 및 랜덤 포레스트와 같은 전통적인 지도학습 기반의 머신러닝 알고리즘이 제안된 모델을 구축하고 검증

<표 1> DGA 기반의 봇넷 탐지에 관한 다른 연구

저자	제안기법	ACC(%)	F1(%)	특징
Truong and Cheng (2016)[12]	J48 Decision Tree	92.30	-	간단하지만, 오탐율이 높음(7.70%)
Hoang and Nguyen (2018)[7]	Various Learning Methods	90.90	90.90	간단하지만, 오탐율이 높음(9.30%)
Qiao et al. (2019)[13]	LSTM Deep Learning		94.58	자원소모가 큼, 오탐 비율(5%)
Zhao et al. (2019)[14]	n-gram Statistics	94.04	-	오탐율이 높음(7.42%)
Hostiadi et al. (2020)[15]	Statistics	89.16	-	오탐율이 높음(10%)

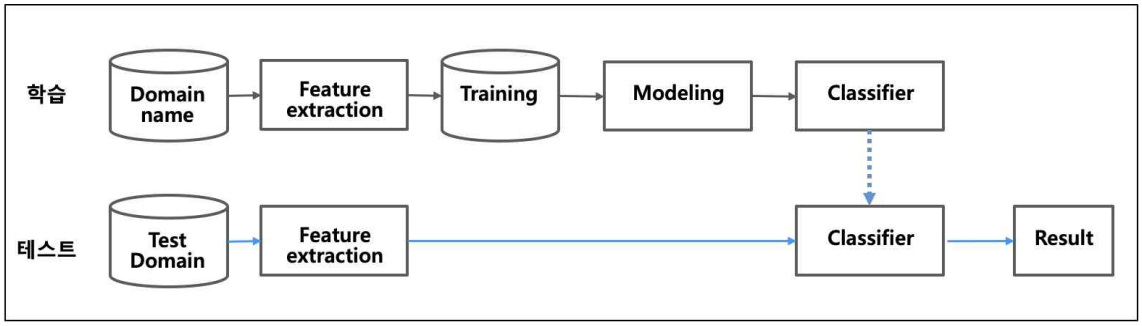
하는 데 사용하였다. 실험 결과는 머신러닝 기술이 봇넷이 사용하는 합법적인 도메인 이름과 알고리즘 생성 도메인 이름의 분류를 기반으로 봇넷을 탐지하는 데 효과적으로 사용될 수 있음을 확인하였다. 실험 결과 또한 랜덤 포레스트 알고리즘이 90% 이상의 가장 높은 전체 탐지율을 생성하였다. 그러나 제안된 모델의 주요 문제는 각 테스트 시나리오에 대한 실험 데이터 세트가 다른 접근 방식에 비해 매우 작고 오탐율이 9.30%로 비교적 높다는 점이다. 작은 실험 데이터 세트는 결과 신뢰도를 감소시키고, 높은 오탐율은 제안된 모델의 실제 적용 가능성을 제한하였다.

Qiao et al. (2019)는 LSTM(Long Short-Term Memory)을 기반으로 DGA 도메인 이름을 분류하는 방법을 제안하였다. LSTM은 지도학습 기반의 딥러닝 방법으로 보안 분야에서 비교적 새로운 접근 방식이다. 제안된 방법에서 각 도메인 이름은 DGA 문자열 추출, 임베딩 기반의 전처리 단계를 통해 생성된다. 그런 다음 훈련 및 테스트를 위해 54×128 행렬로 변환된다. 실험 데이터셋은 Alexa(DN Pedia, n.d)에 의해 순위가 매겨진 상위 100만 개의 합법적인 도메인 이름과 다양한 DGA 봇넷(Netlab 360, n.d)에 의해 생성된 1,675,404개의 악성 도메인 이름으로 구성된다. 실험 결과 제안한 방법이 평균 F1-score 94.58%로 성능이 우수함을 보여주었다. LSTM 학습 방법을 사용하여 제안된 모델은 피처 추출 프로세스에 드는 시간을 최소화 할 수 있다. 그러나 이 논문은 성능에 필요한 자원이 크고 결과에 대한 설명이 간단하지 않으며, 오탐율이 제시되어 있지 않지만 약 5% 정도로 상대적으로 높아서 정확도와 재현율 모두 약 95%에서 유추할 수 있다.

Zhao et al. (2019)는 n-gram 기법을 기반으로 악성 도메인 이름을 탐지하는 통계 기반 방법을 제안하였다. 합법적인 도메인 훈련 데이터는 각 도메인 이름은 먼저 3, 4, 5, 6 및 7-gram 기술을 사용하여 하위 문자열 시퀀스로 나누었다. 그런 다음 모든 훈련 영역의 부분 문자열의 통계 및 가중치 값을 계산하여 'Profile'을 형성한다. 입력 도메인 이름이 정상이거나 악의적인지 확인하기 위해 도메인 이름도 먼저 3, 4, 5, 6 및 7-gram 기술을 사용하여 하위 문자열 시퀀스로 나눈다. 그런 다음 도메인 네임 부분의 문자열에 대한 통계를 계산한 다음 Profile을 기반으로 도메인 네임의 평판 값을 계산하는 데 사용한다. Profile을 사용하여 악성 도메인 이름의 각 범주에 대해 도메인 평판 임계값이 생성된다. 도메인 이름의 평판 값이 임계값보다 큰 경우 정상이고, 그렇지 않으면 악성이다. 실험 결과 제안된 접근 방식이 94.04%의 탐지 정확도를 달성함을 보여주었다. 그러나 제안된 접근 방식의 탐지 성능은 현재 수동으로 생성 및 선택되는 도메인 평판 임계값의 선택에 크게 의존한다. 또한 오탐율은 각각 6.14% 및 7.42%로 상당히 높다.

Hostiadi et al. (2020)은 네트워크 흐름 트래픽 분석을 기반으로 봇 그룹 활동 감지를 위한 B-Corr 모델을 제안한다. B-Corr 모델은 단일 봇의 활동을 감지하는 대신 봇 그룹의 활동을 감지하는 데 중점을 두었다. 봇 그룹 활동 탐지는 네트워크 관리자가 봇 그룹 공격의 활동 또는 액세스를 격리하고 봇 간의 관계를 확인하고 상관관계를 측정하는 데 도움이 될 수 있다. B-Corr 모델은 봇 활동 흐름에서 특징 추출, 봇 간의 교차 측정, 유사성 값 생성을 포함한 세 단계로 구성된다. B-Corr 모델은 봇 그룹의 활동을 지정하기 위해 유사한 대상을 가진 유사한 봇을 분류한다. 이보다 포괄적인 관점을 얻기 위해 B-Corr는 유사한 봇 그래프의 형태로 봇 간의 유사도 값을 시각화한다. 또한 실제 봇넷 데이터 세트를 사용하여 다양한 시나리오에서 수행된 광범위한 실험을 통해 높은 탐지 정확도를 확인했다. 모델의 봇 그룹 활동 IP 주소 탐지 정확도는 89.16%이다. 제안된 접근 방식의 장점은 봇 그룹의 활동을 높은 정확도로 감지할 수 있다는 것이다. 그러나 실험 데이터 세트에서 발견된 봇 그룹 및 그룹 활동의 수가 적기 때문에 탐지 정확도의 신뢰성에 의문이 있다.

<표 1>에서 DGA 기반의 봇넷 탐지에 관한 다른 연구들에 대해서 제안기법과 정확도, F1-score 장단점 등의 특징을 정리하였다. 나타난 바와 같이 지금까지 DGA 기반의 봇넷 탐지에 관한 연구는



(그림 1) 제안된 DGA 기반의 봇넷 탐지 프로세스

비교적 탐지율은 높지만, 오탐율에 대한 이슈가 남아 있음을 알 수 있다.

이에 본 논문에서는 기존의 탐지율을 유지하면서 오탐율을 최소화하는 방안을 제시하고자 한다.

3. 제안된 기법

3.1 제안된 DGA 기반의 봇넷 탐지 프로세스

본 연구의 목적은 DGA 기반의 봇넷 탐지를 위한 모델을 제안한다.

(그림 1)에서 보는 바와 같이 정상 도메인과 DNS 기반의 봇넷 도메인을 통해 Domain name 데이터셋을 구축한다. 구축된 데이터셋에 대해서 n-gram 기반의 피처 추출을 통해 학습 데이터셋을 생성한다. 생성된 학습 데이터셋에 대해서 알고리즘을 통해 분류 모델을 생성한다. 이렇게 생성 모델을 통해 테스트 도메인을 입력하여 도메인의 봇넷 유무를 분류하거나, 봇넷의 유형을 도출하게 된다. 이러한 봇넷 탐지 프로세스는 기존의 시그니처 기반의 탐지보다 탐지율이 높지만, 오탐율이 높다. 이에 본 논문에서는 문자 단위 bigram 기반의 피처 추출을 통한 모델링을 통해 탐지율을 향상하면서도 오탐율을 최소화한 프로세스를 제안한다.

3.2 문자 단위 bigram 기반의 피처 추출

DGA 기반의 봇넷을 탐지하기 위해 DGA 기반으로 생성된 도메인을 탐지하기 위해서 문자 단위(Character level)의 bigram을 적용한다. bigram은 n-gram 사용 시 n=2인 경우를 말한다.

n-gram은 n개의 연속적인 단어 나열을 의미하며, 코퍼스에서 n개의 단어 뭉치 단위로 끊어서 이를 하나의 토큰으로 간주한다. n-gram 점수는 n-gram 데이터셋의 도메인에서 n-gram 순서를 나타낸다. 도메인이 DGA 봇넷에 의해 생성되었으면 n-gram 점수는 정상보다 작다. 도메인 d에서 n-gram을 추출하고 단어사전의 순서에 따라 점수를 준다.

$$S(d) = \sum_{t \in p} Count(t) \times n(t) \ /|p| \qquad (1)$$

(1)에서 n(t)은 도메인 d에서 n-gram의 빈도를 나타내고, |p|는 도메인 d 안에서 n-gram 수를 나타낸다.

이러한 방법은 일반적으로 도메인에 사용되는 단어들을 구성하는 데 사용되지 않은 알파벳의 조합이 존재하므로 단어들을 조합하여 만든 일반적인 도메인보다 다양한 문자 조합으로 생성되는 DGA 도메인을 탐지하는 데 유리하다. 하지만, 모든 일반적인 도메인이 단어를 조합한 것은 아니기 때문에 DGA 알고리즘 중에서 단어를 조합하여 도메인을 생성하는 알고리즘이 있다는 것은 고려할 필요가 있다.

3.3 DGA 알고리즘

다음은 본 논문에서 사용된 DGA 기반의 봇넷

<표 2> DGA 알고리즘 특징 (일부)[16]

DGA	특징
banjori	tld: tk, com, pages.dev sld: A fix length of 9, a-z chars; 27 domains per month e.g: nerjyzkup.com knjpeuzyr.tk zrkyenupj.pages.dev
rovnix	tld: ru, com, net, biz, cn sld: A fix length of 18, mix a-z and 1-8; Generated 10000 domains, actually infinity in theory e.g: c7thuhy8agn43zzgi.biz aby71fqwc3ai12wseh.com lryja5lrm835m7byr8.ru
tinva	tld: variant, depend on seed sld: Fix length of 12, a - y;100 | 200 | 1000 domains in total, depend on seed e.g: nvfowikhevmy.com oykjietwrmlw.ru oqxvkgnpxhyi.in
pykspa	tld: [biz, com, net, org, info, cc] sld: A length of 6-15, a-z; 5000 domains per two days e.g: agadss.biz ynrvwgfqbex.org ssegsguiwcymao.biz
bazardoor	tld: bazar sld: A fix length of 12; 2160 domains per month e.g: aceijlahgijp.bazar acfiimahhiiq.bazar efhijkekjijo.bazar

중 DGA 알고리즘 일부를 나타내고 있다. <표 2>에서 tld는 최상위 도메인을 말하며, sld는 도메인 구성에 필요한 사전을 말한다. eg. 도메인 예제를 나타내고 있다. 이처럼 각 DGA 알고리즘은 각자의 특징을 가지고 있으며, 해당 특징에 맞게 도메인이 생성된다. 이렇게 생성된 도메인은 쉽게 구별되기도 하지만, 정상적인 도메인과 유사하게 생성되기도 한다.

4. 실험 및 평가

이 장에서는 DGA 기반의 봇넷(Botnet) 데이터셋에 대해서 머신러닝 알고리즘을 통해 분류 모델을 생성하고, 생성된 모델을 통해 정상적인 도메인과 봇넷 도메인을 분류하고자 한다.

우리는 본 실험을 통해 다음과 같은 질문을 해결하고자 한다. 정상적인 도메인과 봇넷 유형의 도메인을 분류할 수 있는가? 봇넷이라 분류된 도메인에 대해서 봇넷의 유형을 분류할 수 있는가? 이를 위해 본 실험에서 Netlab 360 DGA 프로젝트에서 사용된 데이터셋으로 61개의 DGA로 구성된 도메인 데이터셋을 사용한다[16].

4.1 데이터셋

본 실험에 사용된 데이터셋은 Netlab 360에서 제공된 데이터셋으로 2016년부터 지속해서 BotNet에 대한 데이터를 수집, 추가하고 있다. 해당 데이터셋에서 포함된 DGA Families는 현재, 61개 구성되어 있다[14].

<표 3> DAG Families 데이터셋

NO	Class	DGA Families	데이터 수
1	legit	alexa	1,000,000
2	Botnet	banjori	483,072
3	Botnet	rovnix	179,995
4	Botnet	tinba	102,115
5	Botnet	pykspa_v1	44,598
6	Botnet	flubot	30,000
7	Botnet	bazardoor	28,410
8	Botnet	simda	30,280
9	Botnet	ramnit	20,058
10	Botnet	ranbyus	13,640
11	Botnet	gameoverzeus	12,000

<표 3>에서 보는 바와 같이 본 논문에서는 사용하는 데이터셋은 정상 도메인으로 Alexa(DN Pedia, n.d)[17] 에서 순위를 매긴 가장 인기 있는 합법적인 도메인 이름 1,000,000개의 문자 분포를 기반으로 하였으며, 61개 봇넷 유형(banjori, rov

nix, tinba, pykspa_v1 등) 상위 10개의 봇넷 유형에서 생성한 921,360개의 악성 도메인으로 구성하였다.

4.2 데이터 전처리

데이터 전처리 단계는 데이터셋을 머신러닝에서 사용할 수 있는 데이터의 형태로 처리하는 단계이다. 먼저, 데이터셋에서 중복 데이터를 제거한다. 둘째, 정상 도메인과 상위 10개의 DGA 알고리즘에서 생성한 도메인을 선정한다. 선정된 도메인별로 10만 건의 데이터를 추출한다. 셋째, 길이가 2인 문자열 생성, "."을 제외한 특수문자의 빈도와 특수문자의 빈도를 각각 특성(피처)화한다. 길이가 2인 문자열의 빈도를 특성(피처)화 (36*36=1296) 하고, 수행, 같은 문자열이 3글자 이상 포함된 경우 연속적 카운터 적용한다.

4.3 실험 환경 및 평가 방법

본 실험 환경은 Ubuntu 18.04.2 LTS에서 Python 3.7을 사용하여 구현되었다. 사용된 고전적인 기계 학습 알고리즘은 Scikit-learn 0.20.4 을 사용하였다. 하드웨어 사양은 GPU는 NVidia Geforce RTX 2060이었으며 256GB RAM, 6TB 하드디스크, AMD Ryzen Threadripper 1900X 8-Core Processor 환경이다.

본 실험의 목적은 머신러닝을 통해 DGA로 구현된 봇넷 도메인과 정상적인 도메인을 분류하는 모델을 구현하고 검증하고자 한다. 이에 대한 평가지표로는 일반적으로 사용되는 Confusion Matrix[18]를 기반으로 기본적으로 사용되는 Accuracy(ACC), Precision, Recall, F1-score, FPR 사용하였다. Accuracy는 모든 샘플 중에서 정상과 봇넷 도메인을 올바르게 분류된 항목의 비율로 정의된다. Precision은 봇넷이라고 예측한 도메인 중 실제 봇넷이라고 분류한 비율을 말한다. Recall은 실제 봇넷 도메인 중 봇넷이라고 예측한 비율을 말한다. F1-score는 Precision과 Recall 간의 조화평균(harmonic mean)을 의미한다. FPR(False Positive Rate)은 정상 트래픽을 공격이라고 잘못 예측한 비율로 오탐에 대한 탐지 결과로 위양성률 또는 오탐율이라고도 한다. 평가지표별 공식은 <표 4>와 같다.

<표 4> 평가지표별 공식

평가지표	공식
Accuracy	$(TP + TN) / (TP + TN + FP + FN)$
Precision	$TP / (TP + FP)$
Recall	$TP / (TP + FN)$
F1-score	$2 \times (Recall \times Precision) / (Recall + Precisio$
FPR	$FP / (FP + TN)$

4.4 실험 결과

본 실험은 크게 두 가지 시나리오를 가지고 수행하였다. 먼저, 정상 도메인과 DGA를 통해 구현된 도메인을 모두 봇넷으로 구분한 이진 분류(Binary Classification), 다음으로 DGA 알고리즘에 따라 봇넷 유형을 나눈 다중 분류(Multi Classification)로 나누어 실험하여 평가하였다.

4.4.1. 이진 분류(Binary Classification)

Binary Classification은 정상 도메인과 DGA를 통해 구현된 봇넷 도메인을 분류하는 모델이며, 이에 대한 성능평가표는 <표 5>와 같다.

<표 5> Binary Classification 성능평가표

	RF	XGBoost	NN
Accuracy	0.9493	0.9551	0.9761
Precision	0.9515	0.9654	0.9791
Recall	0.9465	0.9438	0.9728
F1-score	0.9490	0.9545	0.9759
FPR	0.0480	0.0336	0.0207

해당 <표 5>에서 보는 바와 F1-score 기준으로 NN(Neural Network) > XGBoost > RF(Random Forest) 순으로 성능이 나타나고 있다. 또한, ACC(Accuracy) 기준으로는 NN(Neural Network) > XGBoost > RF(Random Forest) 순으로 나타낸다. 보는 바와 같이 신경망(MLP) 기반의 알고리즘이나 Boosting 기반의 알고리즘(XGBoos

<표 6> Multi Classification 성능평가표

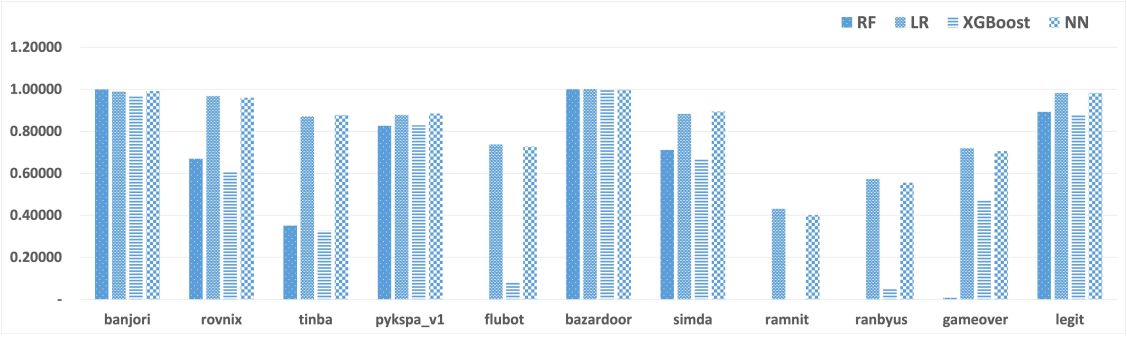
	Precision	Recall	Accuracy	F1-score
banjori	0.98774	0.99502	0.99565	0.99137
rovnix	0.96307	0.95517	0.99205	0.95910
tinba	0.86952	0.88495	0.98708	0.87717
pykspa_v1	0.89847	0.86998	0.99460	0.88400
flubot	0.72190	0.73032	0.99055	0.72609
bazardoor	0.99667	0.99667	0.99990	0.99667
simda	0.87852	0.90951	0.99648	0.89375
ramnit	0.43085	0.37587	0.98793	0.40149
ranbyus	0.57639	0.53548	0.99335	0.55518
gameover	0.74510	0.66901	0.99603	0.70501
legit	0.98128	0.98255	0.98185	0.98191

t)이 좋은 성능을 내고 있음을 알 수 있다. 특히, 본 실험 결과에서는 F1-score가 높을 뿐 아니라 오탐율을 나타내는 FPR 부분에서도 성능 향상을 나타내었다.

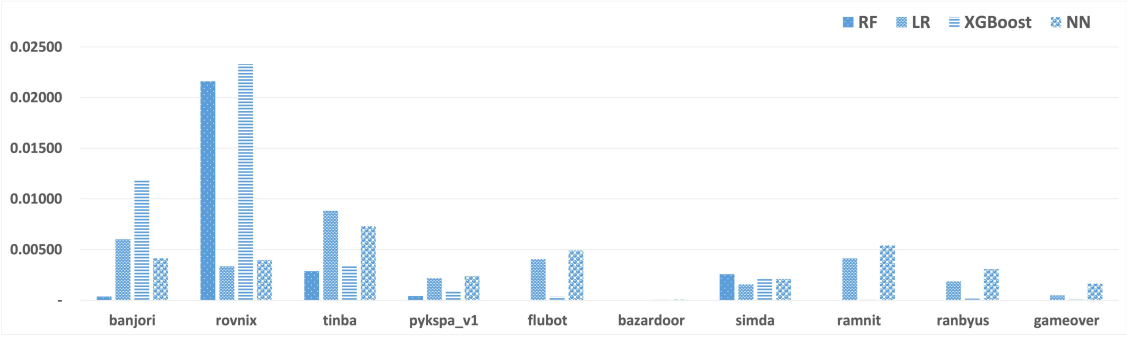
4.4.2 다중 분류(Multi Classification)

Multi Classification은 정상 도메인(legit 기반)과 10개의 DGA를 통해 구현된 봇넷 도메인을 유형별로 분류하는 모델이며, 이에 대한 성능평가표는 <표 6>과 같다. <표 6>은 알고리즘 NN(Neural Network)으로 만들어진 모델로 평가지표는 전체적으로 0.9567의 정확도(Accuracy)와 F1-score 0.9572이다. 평가표에서 ramnit, ranbyus 봇넷은 탐지 결과가 낮게 나오고 있다.

(그림 2)는 DGA 기반의 유형별 봇넷에 대한 알고리즘별 F1-score 값을 나타내고 있다. 전체적으로는 LR > NN(Neural Network) > RF > XGBoost 순으로 다수의 유형에서 봇넷 탐지 결과가 잘



(그림 2) DGA 기반의 유형별 봇넷에 대한 알고리즘별 F1-score



(그림 3) DGA 기반의 유형별 봇넷에 대한 알고리즘별 FPR(오탐율)

도출된 것을 알 수 있다. 이중 ramnit, ranbyus의 경우 전체적으로 점수가 낮으며, 데이터 확인 결과 봇넷 도메인의 형태가 정상 도메인의 형태와 많은 유사도를 하고 있음을 확인할 수 있다. 이처럼 DGA를 통해 생성된 봇넷 도메인의 결과가 Alexa에서 생성된 legit 도메인의 형태와 유사한 경우 분류가 어렵지만, 대다수의 DGA 기반의 봇넷 도메인에 대한 탐지율 성능이 좋음을 알 수 있다. 또한, (그림 3)은 DGA 기반의 유형별 봇넷에 대한 알고리즘별 FPR을 나타내고 있다. 전체적으로 기존 연구보다 오탐율이 상대적으로 적은 성능을 보이고 있으며, 다른 알고리즘보다 XGBoost 경우 오탐율이 높은 것을 볼 수 있다.

5. 결론

본 연구를 통해 DGA 기반의 봇넷을 탐지하는 분류 모델을 제안하였다. 제안된 모델은 문자 기반(Character level)의 bigram을 적용하여 보안 전문가들의 특성(피처)을 추출하기 위한 수고를 줄일 수 있으며, 기존의 모델에 비해 제안된 모델은 95% 이상의 정탐율로 정확도로 테스트 데이터 세트에서 대부분의 DGA 봇넷을 효과적으로 탐지할 수 있었으며, 오탐율에 대해서도 좋은 성과를 도출하였다.

향후에는 본 결과에서도 도출되었지만, 합법적인 도메인 이름과 유사한 도메인 이름을 생성하는 DGA 기반의 봇넷을 탐지할 수 있도록 모델을 계속 개선할 것이다.

참고문헌

[1] M. Willett, “Lessons of the SolarWinds hack. Survival”, 63(2), 7-26, 2021.

[2] S. T. Eun, “Cyber Warfare in the Russo-Ukrainian War: Assessment and Implications”. IFANS FOCUS, 2022(16), 1-4, 2022.

[3] 손현우, 이승진, 허원석. “러시아 우크라이나 간 사이버 전장 내 공격 유형 분석”. 한국정보과학회 학술발표논문집, 2160-2162, 2022.

[4] Y. Zhou, Q. S. Li, Q. Miao, & K. Yim, “DGA-Based Botnet Detection Using DNS Traffic”. J. Internet Serv. Inf. Secur, 3(3/4), 116-123, 2013.

[5] M. Feily, A. Shahrestani, & S. Ramadass, “A survey of botnet and botnet detection”. In 2009 Third International Conference on Emerging Security Information, Systems and Technologies (pp. 268-273). IEEE, 2009.

[6] M. Singh, M. Singh, and S. Kaur, "Issues and challenges in DNS based botnet detection: a survey," Computers & Security, vol. 86, pp. 28-52, 2019.

[7] X. D. Hoang, & X. H. Vu, “An improved model for detecting DGA botnets using random forest algorithm”. Information Security Journal: A Global Perspective, 31(4), 441-450, 2022.

[8] D. Tran, H. Mac, V. Tong, H. A. Tran, & L. G. Nguyen, “A LSTM based framework for handling multiclass imbalance in DGA botnet detection.” Neurocomputing, 275, 2401-2413, 2018.

[9] H. Gohiya, H .Lohiya, & K. Patidar, “A Survey of Xgboost system”. Int. J. Adv. Technol. Eng. Res, 8, 25-30, 2018.

[10] I. Ali, A. I. A. Ahmed, A. Almogren et al., "Systematic literature review on IoT-based botnet attack", IEEE Access, vol. 8, pp. 212220-212232, 2020.

[11] M. Singh, M. Singh, and S. Kaur, "Issues and challenges in DNS based botnet detection: a survey", Computers & Security, vol. 86, pp. 28-52, 2019.

[12] D. T. Truong, & G. Cheng, “Detecting domain-flux botnet based on DNS traffic features in managed network”. Security Comm. Networks 2016 (Vol. 9, pp. 2338-2347). John Wiley & Sons, 2016.

[13] Y. Qiao, B. Zhang, W. Zhang, A. K. Sangaiah, & H. Wu, “DGA domain name classification method based on long short-term memory with attention mechanism”, Applied Science, (2019(9), 4205. https://doi.org/10.3390/ app9204205, 2019.

[14] H. Zhao, Z. Chang, G. Bao & X. Zeng, “Malicious domain names detection algorithm based on N-Gram”, Journal of Computer Networks and Communications 2019, 9. Hindawi. https://do-

i.org/10.1155/2019/4612474, 2019.
[15] D. P. Hostiadi, W. Wibisono & T. Ahmad, “B-corr model for bot group activity detection based on network flows traffic analysis”. KSII Transactions on Internet and Information Systems, 10(2020), 4176-4197. https://doi.org/10.3837/tiis.2020.10.014 14, 2020.
[16] Netlab 360. (n.d.). DGA Families. Available online: https:// data.netlab.360.com/dga/(accessed on 10 August 2022).
[17] DN Pedia. (n.d.). Top Alexa one million domains. CodePunch Solutions. https://dnpedia.com/tlds/topm.php (accessed on 10 August 2022).
[18] C. Yin, Y. Zhu, S. Liu, J. Fei & H. Zhang, “An enhancing framework for botnet detection using generative adversarial networks”. In 2018 International Conference on Artificial Intelligence and Big Data (ICAIBD) (pp. 228-234). IEEE, 2018.

〔저자소개〕

정 일 옥 (Il-ok Jung)
2001년 2월 전남대학교 물리학과 학사
2008년 8월 고려대학교 컴퓨터공학과 석사
2021년 8월 고려대학교 정보보호학과 박사

email : okkida@korea.ac.kr

신 덕 하 (Deok-ha Shin)
2013년 2월 청주대학교 통계학과 학사

email : deokha.shin@igloo.co.kr

김 수 철 (Soo-chul Kim)
2008년 8월 고려대학교 컴퓨터공학과 석사
2022년 8월 숭실대학교 IT정책경영학과 박사 수료
email : kscfuture@naver.com

이 록 석 (Rock-seok Lee)
2002년 8월 광주대학교 전자공학 학사
2011년 2년 전남대학교 정보보호협동과정 석사 수료
email : rockseok.lee@igloo.co.kr