Domain-Based C2 Detection Using Machine Learning

1. Abstract

Bu projede, bir domain adının Command & Control (C2) altyapısına ait olup olmadığını tespit etmek amacıyla iki aşamalı bir makine öğrenmesi tabanlı sistem geliştirilmiştir. İlk aşamada domain isimleri yapısal ve metinsel özellikler kullanılarak hızlı bir şekilde sınıflandırılmış, ikinci aşamada ise DNS tabanlı bağlamsal sinyaller ile bu sonuç doğrulanmıştır. Sistem, yalnızca domain bilgisine dayalı olarak çalışmakta ve risk skoru üreterek güvenlik analizlerinde ön filtreleme amacıyla kullanılabilmektedir.

2. Introduction

Command & Control (C2) sunucuları, zararlı yazılımların uzaktan kontrol edilmesi ve veri sızdırılması gibi kritik saldırı aşamalarında merkezi bir rol oynamaktadır. Geleneksel imza tabanlı yaklaşımlar, sürekli değişen C2 altyapıları karşısında yetersiz kalabilmektedir. Bu nedenle, domain tabanlı özellikler kullanarak C2 benzeri altyapıların makine öğrenmesi ile tespit edilmesi, güncel ve etkili bir yaklaşım olarak öne çıkmaktadır.

Bu projede, domain isimlerinin yapısal özellikleri ve DNS davranışları birlikte değerlendirilerek iki aşamalı bir karar mekanizması tasarlanmıştır.

3. Dataset and Labeling

Veri seti iki farklı kaynaktan oluşturulmuştur:

- **Benign Domainler:** Popüler ve güvenilir domainleri içeren `top-1m.csv` listesi.
- **Malicious Domainler:** URLhaus tarafından sağlanan, hâlihazırda aktif olan zararlı URL'lerin domain bilgileri.

URL formatındaki verilerden yalnızca domain kısmı çıkarılmış ve IP adresleri filtrelenmiştir.

Benign domainler `0`, malicious domainler ise `1` etiketi ile işaretlenmiştir. Veri seti dengeli olacak şekilde düzenlenmiştir.

---

4. Data Preprocessing

Veri ön işleme aşamasında aşağıdaki adımlar uygulanmıştır:

- URL → domain dönüşümü
- Küçük harfe çevirme (normalization)
- Path, port ve `www` gibi gereksiz bileşenlerin kaldırılması
- IP adreslerinin ve IP-benzeri domainlerin filtrelenmesi
- Duplicate kayıtların silinmesi

Bu işlemler sonucunda model için temiz ve tutarlı bir veri seti elde edilmiştir.

5. Feature Engineering

5.1 Lexical (Yapısal) Özellikler

Her domain için aşağıdaki yapısal özellikler çıkarılmıştır:

- Domain uzunluğu
- Rakam oranı
- Tire oranı
- Alt domain sayısı
- Entropy (rastgelelik ölçüsü)
- Üst seviye alan adı (TLD)

Bu özellikler, C2 domain'lerin genellikle otomatik ve rastgele üretilmiş olmasından faydalanarak ayırt edici sinyaller sunmaktadır.

5.2 Text Mining – Character N-gram TF-IDF

Domain isimleri karakter tabanlı n-gram'lara ayrılmış ve TF-IDF yöntemi ile sayısal vektörlere dönüştürülmüştür. Bu sayede "login", "update", "secure" gibi C2 domain'lerde sık görülen pattern'ler modele yansıtılmıştır.

6. Two-Stage Architecture

Sistem iki aşamalı bir mimari ile tasarlanmıştır:

Stage 1 – Machine Learning Based Filtering

İlk aşamada, TF-IDF ve lexical özellikler kullanılarak Logistic Regression modeli eğitilmiştir. Model, domain'in C2 olma olasılığını (`stage1_prob`) üretir.

- `stage1_prob < 0.5` → Benign
- `stage1_prob ≥ 0.5` → Şüpheli (Stage-2'ye geçilir)

Stage 2 – DNS-Based Context Verification

İkinci aşamada, yalnızca şüpheli domainler için DNS tabanlı bağlamsal özellikler çıkarılmıştır:

- A kaydı sayısı
- NS kaydı sayısı
- MX kaydı varlığı
- Minimum TTL değeri

Bu bilgilerden bir doğrulama skoru (`stage2_score`) üretilmiştir. Stage-2 skoru belirli bir eşik değerin altında kalırsa sonuç "Şüpheli (Doğrulanmadı)" olarak işaretlenmiştir. Bu mekanizma, yanlış pozitifleri azaltmayı amaçlamaktadır.

7. Model Training

Stage-1 için Logistic Regression modeli kullanılmıştır. Model, karakter n-gram TF-IDF ve lexical özellikler üzerinde eğitilmiş ve sınıf dengesizliği `class_weight` yaklaşımı ile yönetilmiştir.

8. Evaluation

Model performansı aşağıdaki metrikler ile değerlendirilmiştir:

- Accuracy: %80
- ROC-AUC: 0.88
- PR-AUC: 0.91

Özellikle PR-AUC metriği, güvenlik problemlerinde dengesiz sınıflar için daha anlamlı bir performans ölçütü sunmaktadır.

9. Case Studies

Benign Örnek

Domain: `facebook.com`

Sonuç: Benign olabilir – Risk: 21/100

Şüpheli (Doğrulanmadı)

Domain: `londonairportstransfer.co.uk`

Stage-1 yüksek olmasına rağmen DNS doğrulaması zayıf olduğu için sonuç "Şüpheli (Doğrulanmadı)" olarak işaretlenmiştir.

C2 Olabilir

Domain: `crobel3.floresagapanto.cfd`

Düşük TTL, MX/NS kayıtlarının olmaması ve yüksek ML olasılığı nedeniyle sonuç "C2 olabilir" olarak değerlendirilmiştir.

10. Limitations

- Model yalnızca domain tabanlı çalışmaktadır.
- DNS sorguları ağ koşullarına ve resolver kısıtlamalarına bağlıdır.
- WHOIS verileri sınırlı ve gecikmeli olabilir.

11. Conclusion

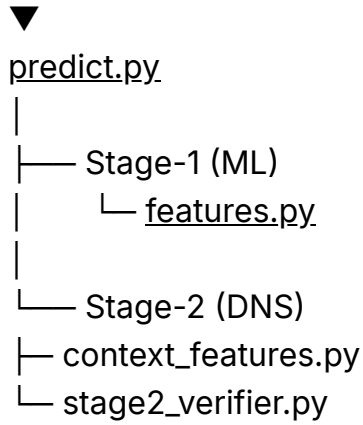
Bu projede, domain tabanlı özellikler ve DNS bağlamsal sinyaller kullanılarak C2 altyapılarının tespiti için iki aşamalı bir sistem geliştirilmiştir. Elde edilen sonuçlar,

domain isimleri üzerinden yapılan analizlerin C2 tespiti için etkili bir ön filtreleme mekanizması sağlayabileceğini göstermektedir.

Dosya Adı	Görevi	Bu Dosyada Ne Yapılıyor?	Veri Madenciliği / ML Aşaması
`dataset_builder.py`	Ham veri hazırlama	İnternetten indirilen benign ve malicious listelerden domain çıkarır, temizler, IP'leri filtreler, `benign.txt` ve `malicious.txt` üretir	Data collection, data cleaning, labeling
`prepare_data.py`	Dataset oluşturma	`benign.txt` ve `malicious.txt` dosyalarını birleştirir, etiketler ve `dataset.csv` üretir	Data preprocessing, dataset construction
`features.py`	Özellik çıkarımı (Stage-1)	Domain string'inden lexical özellikler (uzunluk, entropy, rakam oranı vb.) çıkarır	Feature engineering
`train.py`	Model eğitimi	TF-IDF + lexical özellikleri kullanarak Logistic Regression modelini eğitir	Supervised learning, classification
`evaluate.py`	Model değerlendirme	Confusion matrix, Precision, Recall, F1, ROC-AUC, PR-AUC hesaplar	Model evaluation
`context_features.py`	Bağlamsal özellikler (Stage-2)	DNS üzerinden A, NS, MX kayıtları ve TTL bilgilerini çıkarır	Context-aware feature extraction
`stage2_verifier.py`	Doğrulama mantığı	DNS özelliklerinden rule-based doğrulama skoru ve nedenler üretir	Rule-based verification
`predict.py`	Tahmin & karar	İki aşamalı karar mekanizmasını çalıştırır (ML + DNS), risk skoru ve etiket üretir	Inference, decision making
`c2_domain_model.joblib`	Eğitilmiş model	Logistic Regression + TF-IDF pipeline'ının	Model artifact

Dosya Adı	Görevi	Bu Dosyada Ne Yapılıyor?	Veri Madenciliği / ML Aşaması
		kaydedilmiş hali	
top-1m.csv	Benign veri	Popüler ve güvenilir domain listesi	Benign data source
online_urls.txt	Malicious veri	Aktif zararlı URL'leri içeren ham IOC listesi	Malicious data source
benign.txt	Temiz benign liste	Normalize edilmiş, sadece domain içeren benign veri	Clean labeled data
malicious.txt	Temiz malicious liste	Normalize edilmiş, domain tabanlı zararlı veri	Clean labeled data
dataset.csv	Eğitim verisi	Model eğitiminde kullanılan son etiketli dataset	Final dataset
requirements.txt	Bağımlılıklar	Projede kullanılan Python kütüphaneleri	Environment setup

[Ham Veri]
|
▼
dataset_builder.py
|
▼
benign.txt / malicious.txt
|
▼
prepare_data.py
|
▼
dataset.csv
|
▼
train.py
|
▼
c2_domain_model.joblib
|


predict.py
Stage-1 (ML)
features.py
Stage-2 (DNS)
context_features.py
stage2_verifier.py