

**TÜBİTAK–2209-A ÜNİVERSİTE ÖĞRENCİLERİ ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI**

**ARAŞTIRMA ÖNERİSİ FORMU**

2024 Yılı

1. Dönem Başvurusu

**A. GENEL BİLGİLER**

| **Başvuru Sahibinin Adı Soyadı: Alperen Ordukaya, Muhammed Hasan Medver** |
| --- |
| **Araştırma Önerisinin Başlığı:Derin öğrenme ile botnet tespiti** |
| **Danışmanın Adı Soyadı: Dr. Öğr. Üyesi Peren Jerfi CANATALAY** |
| **Araştırmanın Yürütüleceği Kurum/Kuruluş: İstinye Üniversitesi-Mühendislik ve Doğa Bilimleri Fakültesi-Bilgisayar Mühendisliği Bölümü** |

**ÖZET**

| botnet,kötü amaçlı yazılım kullanılarak ele geçiriliip büyük ve kompleks saldırılar için birbirine bağlı şekilde bekletilen sistemlerdir. çok fazla sistemin aynı anda kullanıldığı saldırılarda kullanılır. bunlar;. Bu enfekte olmuş cihazları yöneten saldırgana bot çobanı denir.  derin öğrenme ile botnet analizi sistemi;  Projenin amacı internet üzerindeki kullanıcılara veya şirketlere yapılan botnet saldırılarını derin öğrenme yöntem ve teknolojileriyle erken tespit etmek ve açtığı zararları sıfıra indirmek.Bunu saldırıların sistem üzerinde yaptığı anormallikleri tespit ederek yapar.  Sistem tarafından önlemesi beklenen senaryolar şu şekildedir, Geçici hizmet kaybı(ddos), Kişisel veri hırsızlığı, Spam e-posta saldırısı, Tıklama dolandırıcılığı VB.    Sistemin Çalışma şekli:  CNN kullanılarak ağ üzerinde yapılan hareketlerin sıcaklık haritasını çıkararak ve bu haritalar üzerinden analiz yaparak.olağanüstü durumları tespit eder.  -Ağ verilerini ön işleme  IP adresleri, port numaraları, paket boyutları, protokolleri bir renk koduna çevirir ve her işlemin bir renk kodunu oluşturur.  -Veriyi görselleştirme ve sıcaklık haritası  Gün içinde yapılan her işlem renk koduna çevrilmişti,bu renk kodlarını oluşturduğumuz matriste sırasıyla hücrelere yazıyoruz ve bu matrisi görselleştiriyoruz. elde ettiğimiz görselde Örneğin, zamanla değişen yoğunlukları veya belirli ağ davranışlarını göstermek için sıcak renkler daha yoğun hareketleri, soğuk renkler (mavi, yeşil gibi) ise daha az hareketi temsil eder.  -Olağanüstü verileri tespit etme  Oluşturduğumuz cnn modeli bu renk kodlarının dağılımını ve zaman içindeki değişimini analiz ederek olağan dışı durumları tespit edebilir. örneğin normal durumda belli portlardan gelen trafik yoğunluğu düşükse burası soğuk renk olarak gözükecek eğer bir botnet saldırısı oldugunda burası yoğun trafik olacağı için renkte bir değişme söz konusu olacak. cnn modeli bu değişimi tespit etmekle görevli |
| --- |
| **Anahtar Kelimeler: Botnet tespiti, Cnn ile derin öğrenme, Sıcaklık haritası, Ağ verileri, Görselleştirme** |
|  |

1. **ÖZGÜN DEĞER**

**1.2.Amaç ve Hedefler**

| Bir şirkete gelen botnet trafiğini farketmek için bir siber güvenlik uzmanının her dakika paketleri ve gelen uyarıları kontrol etmesi lazımdır ve kontrol insan denetiminde olduğu için en ufak hataya ve zafiyetlere sebep olabilir. Geliştirdiğimiz bu proje, uzmanın şirket ağını sürekli kontrol altında tutmasının önüne geçiyor. Derin öğrenme teknolojisiyle olağanüstü trafiği aninda işaretliyor ve uyarı geçiyor.Bu uyarı sonucu şirketin verileri korunabilir ve finansal ve itibarsal kayıp önlenebilir.  Genel hedefler:  1-Uygun Verilerin toplanıp eğitilmesi.  2-Sisteme yüklediğimiz ip adresi, protokol tipi, byte gibi ağ verilerinin sayısal değerlere dönüştürülmesi .  3-Sayısal verilerin matriste toplanarak sayısal veriyi renk koduna çeviren yazılım ile görsel bir veriye yani sıcakık haritasına dönüştürülmesi.  4-Sıcaklık haritasının anlık olarak analiz yapması ve bir sonuç çıkarması. |
| --- |
| **Sistemin geleneksel yöntemlere göre artıları,Sistemi geliştirmeye yönelik hedefler** |

**3.YÖNTEM**

| Botnet tespit projemiz 2 aşamadan oluşmaktadır.  1-Tespitin yapılması için derin öğrenme modelinin tasarımı.  2- Anomali tespiti sonrası kullanıcıyı uyaran sistemin tasarımı.    eğitilmiş sistemimiz ağ verilerini (ip, byte, protokol gibi) tek tip sayısal değere dönüştürür bu projede renk kodlarına çevrilecektir. Renk kodları ise matrislere eklenerek bir sıcaklık haritası oluşturulacaktır.  Anomali tespit mantığımızı bu sıcaklık haritalarının renk değişikliklerine göre yapmaktadır. Görsel üzerindeki değişimleri fark etmemizi sağlayan derin öğrenme modeli ise CNN modelidir.  **Veri toplama Katmanı** = Ağ üzerinde olan anlık ağ verilerini toplar.( bağlantı süreleri, IP adresleri, portlar, protokoller)  **Veri ön işleme** = Toplanan ağ verilerini sayısal değere dönüştürür(renk kodu).  **Sıcaklık haritasi oluşturma=** Oluşturulan veri matriksler yardımıyla sıcaklık haritasına çevrilir.  **Tespit analizi=** Oluşturulan harita oluşturduğumuz cnn derin öğrenme modeli ile analiz edilir ve botnet saldırısı olup olmadığı tespit edilir  **uyarı aşaması=** Anomali tespiti sonrası görevli siber güvenlik uzmanına uyarı geçilir. |
| --- |

1. **PROJE YÖNETİMİ** 
   1. **İş- Zaman Çizelgesi**

**İŞ-ZAMAN ÇİZELGESİ**

| **İP No** | **İş Paketlerinin Adı ve Hedefleri** | **Kim(ler) Tarafından Gerçekleştirileceği** | **Zaman Aralığı**  **(..-.. Ay)** | **Başarı Ölçütü ve Projenin Başarısına Katkısı** |
| --- | --- | --- | --- | --- |
|  | Literatür araştırması ve proje taslağı | Alperen Ordukaya | 2 hafta | %5 |
|  | Veri setlerinin araştırılması ve toplanması | Alperen Ordukaya  Muhammed hasan medver | 1 hafta | %5 |
|  | Verilerin ön işlenmesi | Alperen Ordukaya  Muhammed hasan medver | 2 hafta | %15 |
|  | Model miimarisinin oluşturulması | Alperen Ordukaya  Muhammed hasan medver | 4 hafta | %25 |
|  | Modelin kodlanması ve eğitimi | Alperen Ordukaya  Muhammed hasan medver | 6 hafta | %40 |
|  | Test aşaması | Alperen Ordukaya  Muhammed hasan medver | 1 hafta | %10 |

* 1. **Risk Yönetimi**

**RİSK YÖNETİMİ TABLOSU\***

| **İP No** | **En Önemli Riskler** | **Risk Yönetimi (B Planı)** |
| --- | --- | --- |
| 1 | Veri seti çok hatalı ve proje kapasitesine yetmiyorsa | Veri setinin gözden geçirilmesi, kurtarılamadığı senaryoda yedek veri setleriyle değişimi |
| 2 | Seçilen mimariden kaynaklı performans sorunları | Mimari değişimine hazır olma. |
| 3 | Kullanıcının sistemi yanlış yapılandırması | Kullanıcı dökümasyonunun daima hazır olması |

* 1. **Araştırma Olanakları**

Bu bölümde projenin yürütüleceği kurum ve kuruluşlardavar olan ve projede kullanılacak olan altyapı/ekipman (laboratuvar, araç, makine-teçhizat, vb.)olanakları belirtilir.

**ARAŞTIRMA OLANAKLARI TABLOSU (\*)**

| **Kuruluşta Bulunan Altyapı/Ekipman Türü, Modeli**  (Laboratuvar, Araç, Makine-Teçhizat, vb.) | **Projede Kullanım Amacı** |
| --- | --- |
| Kurum tarafından sağlanan bilgisayar laboratuvarları ve kurulu olan programlama dilleri. | Projede yapılması planlanan derin öğrenme için gerekli olan bilgisayar laboratuvarı ve kurumunda sağladığı lisanslı programlama dilleri kullanılacaktır. |
| Kurumun sağladığı öğrenci hesapları | Kaggle gibi sitelerden öğrenci hesabı ile veri bulma işlemi yapılacaktır. |
| Danışmanlık desteği | Projenin tıkanması durumunda akademik danışman ve uzmanlardan yardım alınacaktır |

1. **YAYGIN ETKİ**

**ARAŞTIRMA ÖNERİSİNDEN BEKLENEN YAYGIN ETKİ TABLOSU**

| **Yaygın Etki Türleri** | **Önerilen Araştırmadan Beklenen Çıktı, Sonuç ve Etkiler** |
| --- | --- |
| **Bilimsel/Akademik**  (Makale, Bildiri, Kitap Bölümü, Kitap) | Projemizin çalışması sonucu elde edilen veriler ve Yapay Zekanın getirdiğimiz boyutu ulusal sempozyum ve bildiri olarak sunulması planlanmaktadır. |
| **Ekonomik/Ticari/Sosyal**  (Ürün, Prototip, Patent, Faydalı Model, Üretim İzni, Çeşit Tescili, Spin-off/Start- up Şirket, Görsel/İşitsel Arşiv, Envanter/Veri Tabanı/Belgeleme Üretimi, Telife Konu Olan Eser, Medyada Yer Alma, Fuar, Proje Pazarı, Çalıştay, Eğitim vb. Bilimsel Etkinlik, Proje Sonuçlarını Kullanacak Kurum/Kuruluş, vb. diğer yaygın etkiler) | Yapılacak projemizin sonunda elde edilen Yapay Zeka destekli uygulamanın sonuç raporu yapılıp teslim edildikten sonra T.C. Ticaret Bakanlığı aracılığıyla projenin tüm telif haklarının alınması ve bu alanda çalışan şirketlerin değerlendirmesine gönderilerek projenin daha da gelişmesi için yatırımlar alınması amaçlanmaktadır. |
| **Araştırmacı Yetiştirilmesi ve Yeni Proje(ler) Oluşturma**  (Yüksek Lisans/Doktora Tezi, Ulusal/Uluslararası Yeni Proje) | Proje sonuçlarının yeni projeler için yol gösterici olacağı düşünülmektedir. Yüksek Lisans projesi olarak geliştirilebilmesi öngörülmektedir. |

**5. BÜTÇE TALEP ÇİZELGESİ**

| **Bütçe Türü** | **Talep Edilen Bütçe Miktarı (TL)** | **Talep Gerekçesi** |
| --- | --- | --- |
| **Sarf Malzeme** | 4000TL | Projede kullanılacak programlama dili ve derin öğrenme kütüphanelerinin lisansı (Python , Kensorflow, Keras,Matlab..) |
| **Makina/Teçhizat (Demirbaş)** | 0TL |  |
| **Hizmet Alımı** | 5000TL | Model eğitimi ve Veri depolama için yüksek gpulu sunuuclar |
| **TOPLAM** | 9000TL |  |

**6. BELİRTMEK İSTEDİĞİNİZ DİĞER KONULAR**

|  |
| --- |

**7. EKLER**

**EK-1: KAYNAKLAR**

# Gupta, H., & Singh, M. (2018). "Network Traffic Analysis and Visualization: Tools and Techniques." *International Journal of Computer Science and Information Technologies*

1. O'Leary, D. E. (2016). "Big Data in Network Management." *Journal of Network and Computer Applications*
2. *Cisco Systems. (2020). Cisco Networking Basics. Cisco.*
3. *NIST. (2018). Framework for Improving Critical Infrastructure Cybersecurity. National Institute of Standards and Technology.*
4. *Matplotlib Documentation. (2021). Matplotlib: Visualization with Python.*
5. *Pandas Documentation. (2021). Pandas: Python Data Analysis Library.*