International Journal of Multidisciplinary Studies and Innovative Technologies

e-ISSN: 2602-4888
dergipark.org.tr/en/pub/ijmsit
Research Article

2022, 6 (1), 23-28
DOI: 10.36287/ijmsit.6.1.23
Received: May 12, 2022; Accepted: June 12, 2022

Nesnelerin İnterneti Cihazlarına Karşı Yapılan Makine Öğrenmesi Saldırıları

Ahmet Emre ERGÜN[1*] ve Özgü CAN [2]

[1*]*Bartın Üniversitesi, Mühendislik, Mimarlık ve Tasarım Fakültesi, Bilgisayar Mühendisliği Bölümü, Bartın, Türkiye (aergun@bartin.edu.tr) (ORCID: 0000-0002-3025-5640)*
[2] *Ege Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, İzmir, Türkiye (ozgu.can@ege.edu.tr) (ORCID: 0000-0002-8064-2905)*

Türkçe Özet – Nesnelerin İnterneti (IoT) cihazlarının sayısının günden güne artmasıyla birlikte bu cihazlara yönelik yapılan saldırılar da artmaktadır. Bu çalışmada, IoT cihazlarında güvenliği sağlama yöntemleri ve IoT cihazlarına yönelik saldırılar ele alınmış, sıfır-güven mimarisinin IoT güvenliğini sağlamadaki önemi açıklanmıştır. Ayrıca, dolgu yöntemlerinin saldırganın kullandığı makine öğrenmesine karşı savunma oranları gösterilmiş ve makine öğrenmesi teknikleriyle kullanılan savunma yöntemleri anlatılmıştır. Bu amaçla, makine öğrenmesi yöntemlerinin etkili olduğu saldırılar, makine öğrenmesi teknikleriyle yapılan saldırılar ve oluşturulan ihlaller belirtilmiştir. Ek olarak, makine öğrenmesi tekniklerinin şifreli trafikte IoT cihazlarını sınıflandırmadaki etkinliği incelenmiştir. Rastgele Orman ve Karar Ağacı sınıflandırma algoritmalarının IoT cihazlarını sınıflandırmadaki etkinliği değerlendirilmiştir. Son olarak, yaygın kullanılan saldırı ve savunma yöntemleri için deneyler gerçekleştirilmiştir. Bu amaçla, IoT cihaz trafiği analiz edilerek dolgulu ve dolgusuz olarak yapılan deneylerin doğruluk oranları karşılaştırılmıştır. Dolgusuz verilere sınıflandırma yapıldığında IoT cihazlarının %84 doğruluk oranı elde edilirken, saldırganın doğru bilgiye erişme oranını düşürmeyi hedefleyen rastgele dolgu yöntemi ile bu doğruluk oranı %19'a düşürülmüştür.

Anahtar Kelimeler – Nesnelerin İnterneti; Makine Öğrenmesi; Dolgu; Şekillendirme; Sıfır-Güven Mimarisi; Sınıflandırma

Atıf: Ergün, A., Can, Ö. (2022). Nesnelerin İnterneti Cihazlarına Karşı Yapılan Makine Öğrenmesi Saldırıları. International Journal of Multidisciplinary Studies and Innovative Technologies, 6(1): 23-28.

Machine Learning Attacks Against Internet of Things Devices

Extended Abstract

As the number of Internet of Things (IoT) devices increases day by day, attacks against these devices are also increasing. In this study, methods of ensuring security in IoT devices and attacks on IoT devices are discussed, and the importance of zero-trust architecture in ensuring IoT security is explained. In addition, the defense rates of padding methods against machine learning used by the attacker are shown and the defense methods used with machine learning techniques are explained. For this purpose, machine learning methods that are effective on attacks, attacks and violations that are achieved by machine learning techniques are specified. In addition, the effectiveness of machine learning techniques in classifying IoT devices in encrypted traffic is examined. The effectiveness of Random Forest and Decision Tree classification algorithms in classifying IoT devices are evaluated. Finally, experiments are carried out for commonly used attack and defense methods. For this purpose, the accuracy rates of the padded and unpadded experiments are compared by analyzing the IoT device traffic. When classifying unpadded data, 84% accuracy rate of IoT devices is achieved, while this accuracy rate has been reduced to 19% with the random padding method that aims to reduce the attacker's rate of accessing correct information.

Keywords – Internet of Things; Machine Learning; Padding; Shaping; Zero-Trust Architecture; Classification

Citation: Ergün, A., Can, Ö. (2022). Machine Learning Attacks Against Internet of Things Devices. International Journal of Multidisciplinary Studies and Innovative Technologies, 6(1): 23-28.

I. GİRİŞ

Nesnelerin İnterneti (*Internet of Things*, IoT), İnternet aracılığıyla diğer cihaz ve sistemlere bağlanarak veri alışverişi yapmak üzere sensörler, yazılımlar ve diğer teknolojilerle oluşturulan fiziksel nesnelerin ağını tanımlar. IoT cihazlarının günümüzde popülerliğinin ve kullanım talebinin artması nedeniyle, bağlantılı IoT cihazlarının toplam sayısının 2030 yılına kadar yaklaşık 80 milyara ulaşacağı tahmin edilmektedir [1].

Teknolojik gelişmelerin ilerlemesi ve IoT cihazlarının gerekliliğinin artmasıyla birlikte ağ ortamına bağlı bu cihazların kullanımının getirdiği güvenlik sorunları mevcuttur.

Verilerin gizliliğini, bütünlüğünü ve kullanılabilirliğini hedef alan saldırıların önlenmesi için güvenlik önlemleri kullanmak önemlidir. Makine öğrenmesi yöntemleri sayesinde önlemler daha etkin bir hale getirilebilmektedir. Makine öğrenmesi yöntemleri güvenliği sağlamakta kullanılabildiği gibi güvenlik sorunlarına yol açan saldırılarda da kullanılabilmektedir.

Bu çalışmada, IoT cihazlarının güvenliğini sağlamaya yönelik önlemler ve IoT cihazlarına karşı yapılan saldırılar değerlendirilmiştir. Makine öğrenmesinin yalnızca güvenliği sağlamada değil saldırılarda da kullanılabileceği belirtilmiştir. Bu çalışmanın hedefi doğrultusunda hangi makine öğrenmesi tekniğinin hangi güvenlik önlemini sağlamada ve hangi makine öğrenmesi tekniğinin hangi saldırıda daha etkin olduğu açıklanmıştır.

Öncelikle, IoT saldırılarına karşı yaygın kullanılan güvenlik önlemleri anlatılmıştır. Bir sonraki bölümde, makine öğrenmesi temelinde kullanılan güvenlik önlemleri açıklanmıştır. Daha sonra, Makine öğrenmesi ile IoT'ye yönelik yapılan saldırılar ve hedef aldığı ihlaller ele alınmıştır. Son olarak, makine öğrenmesi algoritmalarının saldırı ve güvenlik yönünden etkinliği değerlendirilmiştir.

Bu çalışmanın organizasyonu şu şekildedir: ikinci bölümde sıfır-güven (*zero-trust*) mimarisi, dolgu (*padding*) ve şekillendirme (*shaping*) yöntemleri, güvenliği sağlamaya yönelik olarak kullanılan makine öğrenmesi yöntemleri ve makine öğrenmesi tabanlı IoT saldırıları açıklanmış, üçüncü bölümde Rastgele Orman ve Karar Ağacı sınıflandırıcı algoritmalarının dolgulu ve dolgusuz doğruluk performansları değerlendirilmiş ve dördüncü bölümde tartışma ve sonuç sunulmuştur.

II. MATERYAL VE METOTLAR

A. *Yaygın Kullanılan Güvenlik Önlemleri*

IoT cihazlarında güvenliği artırmak için sıfır-güven ilkesini temel alan sıfır-güven mimarisi ve paket boyutlarını değiştirerek makine öğrenmesi modellerini yanıltmayı amaçlayan dolgu ve şekillendirme yöntemleri kullanılabilmektedir.

1. Sıfır-Güven Mimarisi

IoT güvenliğini sağlamak için tüm ağ altyapısı boyunca görünürlük, segmentasyon ve kesintisiz koruma sağlayabilen entegre çözümler gereklidir. Her zaman doğrulamayı ve hiçbir zaman güvenmemeyi esas alan sıfır-güven modelinde, içerideki veya dışarıdaki ağdan gelebilecek verilerin hiçbirine güvenilmemektedir [2]. Ağlara uzaktan erişimin artmasıyla birlikte, IoT cihazlarını korumak için sıfır-güven yaklaşımı önem kazanmaktadır. Sıfır-güven mimarisinde, sıfır-güven erişim ilkesini kullanan Rol Tabanlı Erişim Denetimi, kullanıcılara görevleri için gereken minimum ağ erişimi düzeyini sunarken ağın diğer bölümlerine erişmelerini veya bunları görmelerini engellemek için en az erişim ilkesini kullanan sıfır-güven erişimi, ağ erişim yönetiminin kritik bir bileşenidir. Sıfır-güven mimarisi ayrıca tüm ağ bileşenlerinin kapsamlı yönetim kontrolünü ve görünürlüğünü geliştirmek ve sürdürmek için birbirine bağlı akıllı cihazların kimlik doğrulamasını sağlayabilmektedir [3].

2. Dolgu

IoT trafiği şifreli olsa da ağdaki pasif trafik izleyicilerinin ağı gözetleyerek ağdan elde ettiği verilerle makine öğrenmesi yöntemleri kullanarak ağ trafiğinin gizliliğini ihlal etmesi mümkündür. Saldırganın, IoT cihazları tarafından oluşturulan trafiği makine öğrenmesi yöntemleriyle sınıflandırmasını mümkün olduğunca yanıltmak ve aynı zamanda trafiği sekteye uğratmamak için bantgenişliğini ayarlamak önemlidir. Bu nedenle, mahremiyet ve fayda arasındaki dengeyi sağlamak için saldırganı yanıltırken, mümkün olan en düşük dolgu yöntemini uygulamak gerekmektedir. Dolgu yöntemi, trafikteki paketlerin boyutuna ekstra boyut ekleyerek, saldırganın cihazlar hakkında bilgi sahibi olmak üzere makine öğrenme modeline girmek üzere ele geçirdiği trafiğin, öğrenmeyi yanıltmak için değiştirilmesidir. Rastgele Orman (*Random Forest,* RF) makine öğrenmesi tekniği kullanılarak ağın içindeki ve dışındaki saldırgan tarafından IoT cihazının doğru bir şekilde tespit edilmesine yönelik gerçekleştirilen çalışmanın [4] sonuçları Tablo 1'de sunulmaktadır.

TABLO 1. Rastgele Orman modeli için doğruluk oranları [4]

	Ağ dışındaki saldırgan	**Ağ içindeki saldırgan**
Dolgusuz	%96	-
Seviye-100 Dolgu	%32,77	%66,03
Seviye-500 Dolgu	%14,28	%52,18
Seviye-700 Dolgu	%5,94	%50,38
Seviye-900 Dolgu	%4,96	%49,83

Tablo 1'de sunulan sonuçlar için 21 cihazın 20 günlük trafikten oluşturduğu veri seti [5] kullanılmıştır. İlgili çalışmada [4], ağ dışındaki saldırgan bir ağ içindeki etkinlikleri izlemekte, ağ içindeki saldırgan dolgu yönteminin sağlayıcısı olup, sahip olduğu ağın kullanıcılarını izlemektedir. Bir saniyelik zaman pencerelerinde gruplanmış şifreli trafiğin standart sapma, ortalama ve toplam büyüklüklerini kullanan saldırganın bir IoT cihazını tanımlama doğruluk oranı dolgudan önce %96 iken seviye-900 dolgudan sonra %4,96'ya düşmektedir. Böylelikle, gerçek paket uzunluğuna dolgu eklenerek oluşturulan sahte paket boyutu ile saldırganın öğrenme modeli yanıltılabilmektedir.

3. Şekillendirme

IoT trafiğinin trafik analizi saldırısına karşı korunması için trafiği şekillendirmek etkili bir yöntemdir. Trafik şekillendirmesi, trafiğe sahte paketler ekleyerek kullanıcı etkinliğinin çıkarımının yapılmasını önlemeyi sağlamaktadır. [6] çalışmasında Stokastik Trafik Dolgusu (*Stochastic Traffic Padding, STP*) algoritmasına dayalı trafik şekillendirme yöntemi önerilmektedir. Böylelikle, kullanıcı aktivitesi olmadığında bile trafik periyotlarını belirli aralıklarla dolgulayarak saldırganın hangi zaman periyotlarında gerçek trafik etkinliğinin olduğunu algılaması önlenmektedir. STP, farklı cihazlar ve kullanıcı etkinliği frekansları için trafik bant genişliği yükü ve saldırganın kendinden eminliği arasında ayarlanabilir bir denge sağlamaktadır. Bant genişliği yükü artırılarak saldırganın IoT cihaz tespit doğruluk oranı %50'den %10'a düşürülmüştür.

B. *Makine Öğrenmesi Temelinde Kullanılan Güvenlik Yöntemleri*

Teknolojik gelişmeler sayesinde IoT cihazlarının ağ güvenliğini sağlama olanakları artmaktadır. Yaygın kullanılan IoT saldırılarına karşı makine öğrenmesi yöntemlerinin

kullanılması IoT güvenliğini sağlamak için önemlidir. Doğrulama, erişim kontrolü ve kötü amaçlı yazılım algılama gibi işlemlerin öğrenme tabanlı yapılması ile IoT güvenliği sağlanabilmektedir [7,8]. Bu kapsamda, [7] çalışmasında öğrenme temelli kimlik doğrulama, güvenli yük boşaltma (*secure offloading*), kötü amaçlı yazılım tespiti ve erişim kontrolü yöntemlerinin kullanımına yönelik detaylı bir karşılaştırma sunulmaktadır. Saldırganın makine öğrenmesi olanağına sahip olduğu durum dikkate alındığında, saldırgan makine öğrenmesine yönelik alınabilecek çeşitli önlemler bulunmaktadır [9].

Makine öğrenmesi temelli güvenlik önlemleri aşağıda belirtilmiştir:

- *Öğrenme Tabanlı Kimlik Doğrulama:* Sahtecilik (*spoofing*) ve Dinleme (*eavesdropping*) gibi saldırılara karşı etkilidir. Q-öğrenme tabanlı kimlik doğrulama, önceden bir eğitim veri seti gerektirmeden buluttaki ortamdan öğrenerek IoT cihazlarının kimlik doğrulama başarısını geliştirmesini sağlamaktadır [10].
- *Öğrenme Tabanlı Güvenli Yük Boşaltma:* Hizmet Reddi (*Denial of Service*, DoS) ve Karıştırma (*Jamming*) gibi saldırılara karşı etkilidir. Yük boşaltma, verilerin cihaz veya bulut gibi farklı platforma aktarılması işlemidir. IoT cihazları, Karıştırma ve Sahtecilik saldırılarına karşı yük boşaltma veri değerlerini seçmek için Q-öğrenme tabanlı güvenli yük boşaltmayı kullanmaktadır [11].
- *Öğrenme Tabanlı Kötü Amaçlı Yazılım Tespiti:* Virüs ve Trojan gibi kötü amaçlı yazılım saldırılarına karşı etkilidir. K-en yakın komşu algoritması (*K-Nearest Neighbors*, K-NN) ve Rastgele Orman algoritmaları kullanılarak kötü amaçlı yazılımların tespit edilmesine yönelik bir değerlendirmenin sunulduğu [12] çalışmasında, makine öğrenmesi sınıflandırıcılarının güncel en son kötü amaçlı yazılımları tespit edebildiği kanıtlanmaktadır.
- *Öğrenme Tabanlı Erişim Kontrolü:* DoS, Mahremiyet Sızıntısı ve Kötü amaçlı yazılım saldırılarına karşı etkilidir. Liteartürde, sızma tespiti için Destek Vektör Makinesi (*Support Vector Machine*, SVM) ve K-NN gibi makine öğrenmesi tekniklerinin kullanıldığı çalışmalar yer almaktadır [13].
- *Çekişmeli Eğitim Gradyanı (Adversarial Training Gradient):* Amaç, modelin dayanıklılığını artırmak için eğitim setine yanıltıcı veriler eklemektir. Ancak, kara-kutu (*black-box*) saldırısına karşı bu savunma stratejisi etkisizdir [9].
- *Aktarılabilirliği Engelleme (Blocking the Transferability):* Bu strateji, aktarılabilirlik özelliğini ortadan kaldırmayı ve bir saldırganın yanıltıcı eğitim verileri oluşturmasını önlemeyi amaçlamaktadır. Bu amaçla, eğitim kümesinin girişine bir "NULL" etiket sınıfı eklenmektedir. Sonuç olarak, sınıflandırıcı orijinal etikete daha az güvenir ve yanıltıcı verileri NULL olarak kategorize ederek reddeder [9].
- *Savunma-Üretici Çekişmeli Ağ (Defense-GAN):* Bu mekanizma, kara-kutu (*black-box)* ve beyaz-kutu (*white-box*) saldırılarında çalışmaktadır ve derin sinir ağlarını bozulmalara karşı korumaktadır Amaç, bir üretici çekişmeli ağ yardımıyla yanıltıcı örnekleri engellemektir [9].

C. Makine Öğrenmesi ile Yapılan IoT'ye Yönelik Saldırılar

Makine öğrenmesi teknikleri kullanarak IoT cihazlarındaki trafiğin manipüle edilmesi veya trafik şifreli olsada trafikteki bilgilerin ele geçirilip mahremiyete tehdit olarak kullanılması mümkündür. Saldırganların, makine öğrenmesi tekniklerini IoT ortamlarında yaygın olarak kullanmalarında mahremiyeti ihlal etmeye, sistemi aksatmaya veya yanıltmaya yönelik amaçları bulunmaktadır. Makine öğrenmesi algoritmaları üç kategoride incelenmektedir. Bu kategorilerden, IoT saldırılarında yaygın kullanılan makine öğrenmesi algoritmaları aşağıda belirtilmiştir:

- *K-En Yakın Komşu (K-NN):* Gözetimli öğrenme (*supervised learning*) algoritmasıdır. Bir veri noktasının, kendisine en yakın veri noktalarının hangi gruba ait olduğuna bağlı olarak bir grubun veya diğerinin üyesi olma olasılığını tahmin etmeye yarayan bir veri sınıflandırma yöntemidir.
- *K-Ortalama (K-Means):* Gözetimsiz öğrenme (*unsupervised learning*) algoritmasıdır. Temel amacı, gözlemler ile küme ağırlık merkezi arasındaki mesafelerin toplamını minimize etmek olan bir kümeleme algoritmasıdır.
- *Q-Öğrenmesi (Q-Learning):* Pekiştirmeli öğrenme (*reinforcement learning*) algoritmasıdır. Mevcut durum göz önüne alındığında yapılacak en iyi eylemi bulmaya çalışan bir algoritmadır.

Makine öğrenmesi kategorilerine ait saldırıların oluşturduğu ihlaller ve saldırı türleri Tablo 2'de gösterilmiştir [9]. Gizlilik, Bütünlük ve Kullanılabilirlik ihlallerine yol açan saldırılar ve bu saldırıların hangi makine öğrenmesi kategorisi yoluyla yapıldığı Tablo 2'de sunulmaktadır.

TABLO 2. Makine Öğrenmesi Kategorilerine Göre Saldırı Örnekleri

	Gözetimli Öğrenme	**Gözetimsiz Öğrenme**	**Pekiştirmeli Öğrenme**
İhlaller	Gizlilik ve Mahremiyet İhlali	Bütünlük İhlali	Kullanılabilirlik İhlali
Saldırılar	Trafik Analizi Kriptanaliz Yan Kanal Sosyal Ağ	Kaçınma (*Evasion*) Nedensel (*Causative*) Keşif (*Exploratory*) Sahte Enjeksiyon Kod Enjeksiyonu	Sahtecilik (*Spoofing*) Karıştırma (*Jamming*) Kara Delik (*Blackhole*)

1. Mahremiyet İhlali

Kullanıcıların hassas bilgilerini kötüye kullanım ile ciddi sonuçlara yol açabilecek mahremiyete yönelik saldırılarda mümkündür [14]. Şifrelenmiş IoT trafiğindeki şifrelenmiş paketlerin boyutları makine öğrenmesi teknikleri kullanılarak IoT cihazları ve trafiği hakkında bilgilere erişilebilmektedir [4,15,16].

Makine öğrenmesi teknikleriyle şifrelenmiş verileri sınıflandırmaya karşı dolgu ve şekillendirme yöntemleri kullanılsa da saldırıya uğrayan bir ağdaki cihazlar ve etkinlikler tanımlanabilmektedir. [17]'de sunulan çalışmada,

saldırganın dolgulu ve şekillendirilmiş trafiğe eriştiğinde K-NN algoritmasıyla %81 doğruluk oranında IoT cihazlarını bir saniyelik zaman penceresi düzeyinde ayırt edebildiği gösterilmektedir. Tablo 3'de [17] çalışmasında cihaz sayılarına göre gerçekleştirilen sınıflandırmaların doğruluk oranları gösterilmektedir. Elde edilen sonuçlar doğrultusunda, cihaz sayısı arttıkça doğruluk oranının azaldığı görülmektedir.

TABLO 3. Cihaz sayılarına göre doğruluk oranları [17]

Cihaz Sayısı	Doğruluk Oranı
5	%81
10	%77
14	%75

2. *Bütünlük ve Kullanılabilirlik İhlali*

IoT cihazlarının kullanımı yaygınlaştıkça yeni güvenlik riskleri ortaya çıkmaktadır. Saldırgan, IoT trafiğini izleyerek elde ettiği veriler doğrultusunda cihazların taklit edilmesini veya ağın akışını bozacak şekilde manipüle edilmesini makine öğrenmesi yöntemleri ile sağlayabilmektedir. Böylece saldırgan IoT cihazlarındaki trafiğin bütünlüğünü ve kullanılabilirliğini ihlal etmiş olmaktadır. Sosyal mühendislik saldırıları, keşif saldırısı, Hizmet Reddi (*Denial of Service, DoS*) ve Ortadaki Adam Saldırısı (*Man in the Middle Attack*) gibi saldırıların yanında makine öğrenmesini kullanarak bir saldırgan, bir kullanıcıyı veya sistemi taklit ederek IoT cihazlarını yanıltabilir, sekteye uğratabilir veya hassas bilgileri ele geçirebilir [3,19-22].

Saldırganın IoT cihazlarının yalnızca küçük bir bölümünü kontrol ederek veri füzyonunda karar vermeyi etkilediği çekişmeli makine öğrenmesi (*adversarial machine learning)* tabanlı kısmi model (*partial-model*) saldırısı [21] çalışmasında sunulmaktadır. [22] çalışmasında sunulan çekişmeli makine öğrenmesi üzerine oluşturulan yeni teknikler, karıştırma (*jamming*), spektrum zehirlemesi (*spectrum poisoning*), ve öncelik ihlali (*priority violation*) saldırılarına uygulanmıştır. Saldırgan derin yapay sinir ağı sınıflandırıcı kullanarak bir IoT veri aktarıcısının kanal erişim algoritmasını çıkarmak için keşif (*exploratory*) saldırısını ve bu tahmin sonuçlarına dayanarak veri aktarıcısını test aşamasında yanıltmak için kaçınma (*evasion*) saldırısını kullanmaktadır. IoT aktarıcısı kanal erişim algoritmasını yeniden öğrenme işlemi için eğittiği sırada saldırgan nedensel (*causative*) saldırısıyla aktarıcıya giden verileri manipüle etmektedir.

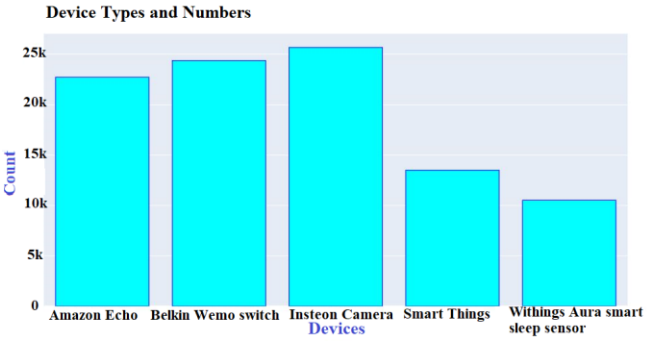
III. BULGULAR

Bu bölümde, makine öğrenmesi yöntemleri kullanarak mahremiyet ihlaline yol açan saldırgan ve bu ihlali dolgu yöntemi ile önlemeye çalışan kurban ele alınmaktadır. Bu doğrultuda, [5] çalışmasında yer alan veri seti kullanılarak Rastgele Orman ve Karar Ağacı algoritmalarına karşı dolgu yöntemlerinden biri olan rastgele dolgunun etkinliği gösterilmiştir. Rastgele Orman ve Karar Ağacı sınıflandırıcı algoritmalarının dolgu olmadan performansı ve rastgele dolgu olduktan sonraki doğruluk oranları Tablo 4'de verilmiştir. Test verilerinin her bir paketinin uzunluğu kendi uzunluğu ve 1600 bayt arasında rastgele bir değere atandığı rastgele dolgu yöntemine karşı alınan doğruluk oranları ve eğitim(*train*)-test verilerine dolgu yapılmayan dolgusuz yönteme karşı alınan doğruluk oranları gösterilmiştir [23].

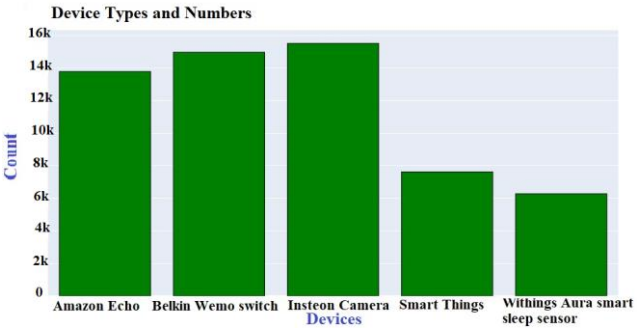
TABLO 4. Algoritmalara göre doğruluk oranları.

	Dolgusuz	Rastgele Dolgulu
Rastgele Orman	%83,3	%23
Karar Ağacı	%84,2	%19,8

Bu çalışmada gerçekleştirilen deneylerde, [5] çalışmasında oluşturulan veri seti kullanılmıştır ve bu veri setlerinden beş cihazın trafik analizi yapılmıştır. Eğitim ve test dosyalarındaki cihazların sayısal dağılımı Şekil 1 ve Şekil 2'deki grafiklerde gösterilmektedir.

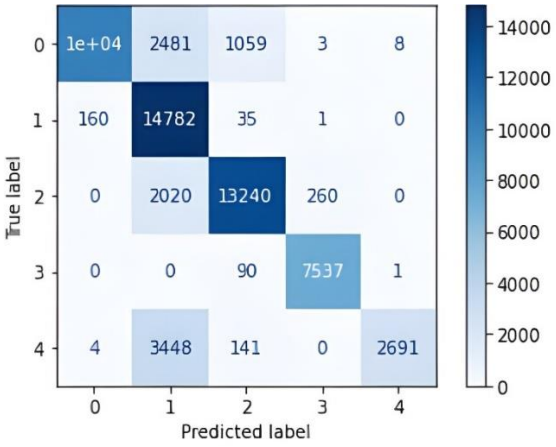


Şekil 1.Train Cihazları ve Sayıları

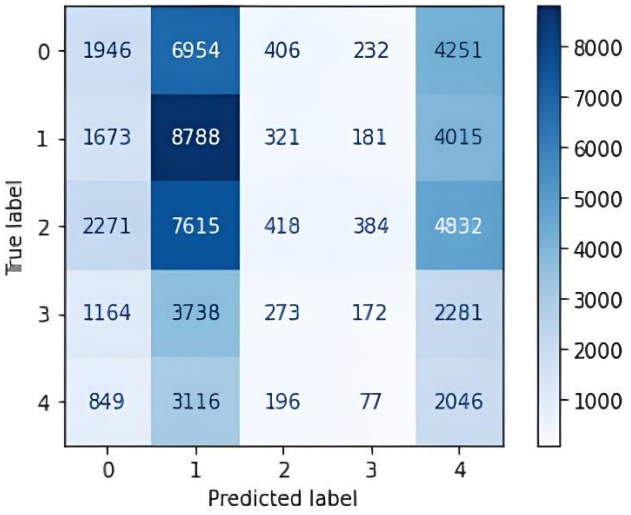


Şekil 2. Test Cihazları ve Sayıları

Rastgele Orman algoritmasıyla yapılan analizdeki karışıklık matrisleri (*confusion matrix*) Şekil 3 ve Şekil 4'de gösterilmektedir.

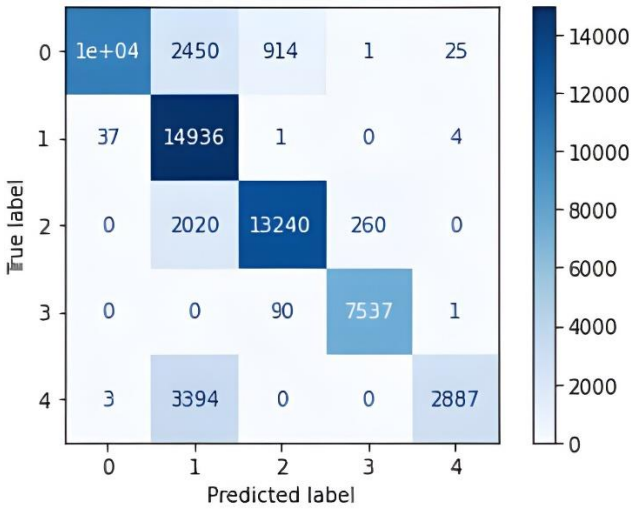


Şekil 3. Dolgusuz Rastgele Orman Karışıklık Matrisi

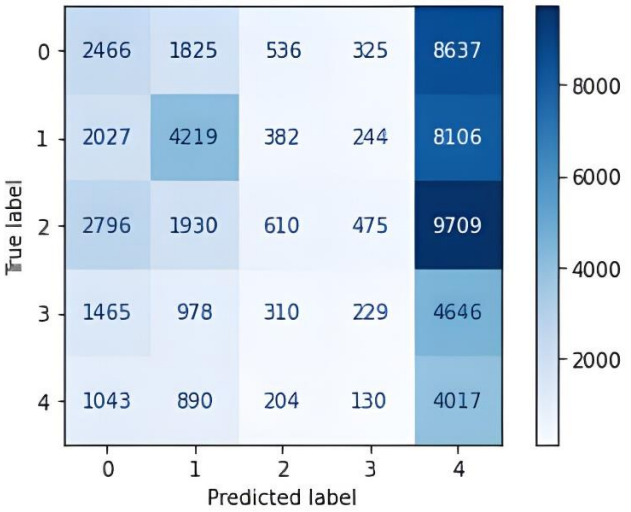


Şekil 4. Rastgele Dolgulu Rastgele Orman Karışıklık Matrisi

Karar Ağacı algoritmasıyla yapılan analizdeki karışıklık matrisleri Şekil 5 ve Şekil 6'da gösterilmektedir.



Şekil 5. Dolgusuz Karar Ağacı Karışıklık Matrisi



Şekil 6. Rastgele Dolgulu Karar Ağacı Karışıklık Matrisi

Rastgele dolgu ile saldırganın doğruluk oranının her iki sınıflandırıcı için azaldığı gözlemlenmiştir. Dolgusuz yapılan analizlerde Karar Ağacı sınıflandırma algoritması %84,2 doğruluk oranına sahipken Rastgele Orman sınıflandırma algoritması %83,3 doğruluk oranına sahiptir. Rastgele Orman algoritması rastgele dolgu %23 doğruluk oranı ile Karar Ağacı'nın %19,8 doğruluk oranından daha etkili olduğu görülmüştür.

IV. SONUÇ

Makine öğrenmesi yöntemleri saldırılara karşı savunma amacı ile de kullanılmaktadır. Bu çalışmada, IoT'ye yönelik yapılan saldırıların makine öğrenmesiyle daha etkili hale geleceği ve uygulanabilecek saldırıların hangi amaçla kullanılabileceği belirtilmiş ve bu saldırılara karşı alınabilecek önlemler açıklanmıştır. Ek olarak, hangi makine öğrenmesi tekniğinin hangi saldırılarda daha etkili olduğu ve hangisinin saldırılara karşı savunmada daha etkili olduğuna yönelik bir değerlendirme sunulmuştur.

Bu çalışmada ayrıca, IoT cihazlarında mahremiyete yönelik makine öğrenmesi ile yapılan sınıflandırma ve buna karşı alınabilecek dolgu yöntemlerinden biri olan rastgele dolgu yöntemi kullanılarak bir deney gerçekleştirilmiştir. Yaygın kullanılan bir sınıflandırıcı algoritması olan ve içinde çokça Karar Ağacı bulunduran Rastgele Orman algoritmasının dolgulu verilerdeki etkinliği gözlemlenmiştir. Rastgele dolgu yöntemine karşı Rastgele Orman algoritmasının %23 doğruluk oranı ile Karar Ağacı'nın %19,8 doğruluk oranından daha etkili olduğu görülmüştür. Rastgele Orman algoritması eğitim aşamasında çok sayıda Karar Ağacı kullanması nedeni ile Karar Ağacı algoritmasına göre daha iyi sonuç vermiştir.

Gelecek çalışmalar kapsamında, IoT cihaz trafiğinde mahremiyet, bütünlük ve kullanılabilirlik ihlallerine karşı optimal ve güvenilir bir savunma yöntemi geliştirilmesi hedeflenmektedir.

KAYNAKLAR

[1] Chettri, L., & Bera, R. (2019). A comprehensive survey on Internet of Things (IoT) toward 5G wireless systems. *IEEE Internet of Things Journal*, *7*(1), 16-32.

[2] Samaniego, M., & Deters, R. (2018, July). Zero-trust hierarchical management in IoT. In *2018 IEEE international congress on Internet of Things (ICIOT)* (pp. 88-95). IEEE.

[3] Lakhani, A. (2019, June 17). *Examining Top IoT Security Threats and Attack Vectors.* Fortinet. https://www.fortinet.com/blog/industry-trends/examining-top-iot-security-threats-and-attack-vectors

[4] Pinheiro, A. J., de Araujo-Filho, P. F., Bezerra, J. D. M., & Campelo, D. R. (2020). Adaptive Packet Padding Approach for Smart Home Networks: A Tradeoff Between Privacy and Performance. *IEEE Internet of Things Journal*, *8*(5), 3930-3938.

[5] Sivanathan, A., Gharakheili, H. H., Loi, F., Radford, A., Wijenayake, C., Vishwanath, A., & Sivaraman, V. (2018). Classifying IoT devices in smart environments using network traffic characteristics. *IEEE Transactions on Mobile Computing*, *18*(8), 1745-1759.

[6] Apthorpe, N., Huang, D. Y., Reisman, D., Narayanan, A., & Feamster, N. (2019). Keeping the smart home private with smart (er) iot traffic shaping. *Proceedings on Privacy Enhancing Technologies*, *2019*(3), 128-148.

[7] Xiao, L., Wan, X., Lu, X., Zhang, Y., & Wu, D. (2018). IoT security techniques based on machine learning: How do IoT devices use AI to enhance security?. *IEEE Signal Processing Magazine*, *35*(5), 41-49.

[8] Mohamed Shakeel, P., Baskar, S., Sarma Dhulipala, V. R., Mishra, S., & Jaber, M. M. (2018). Maintaining security and privacy in health care system using learning based deep-Q-networks. *Journal of medical systems*, *42*(10), 1-10.

[9] Bout, E., Loscri, V., & Gallais, A. (2021). How Machine Learning changes the nature of cyberattacks on IoT networks: A survey. *IEEE Communications Surveys & Tutorials*.

[10] Xiao, L., Li, Y., Han, G., Liu, G., & Zhuang, W. (2016). PHY-layer system using learning spoofing detection with reinforcement learning in wireless networks. *IEEE Transactions on Vehicular Technology*, *65*(12), 10037-10047.

[11] Xiao, L., Xie, C., Chen, T., Dai, H., & Poor, H. V. (2016). A mobile offloading game against smart attacks. *IEEE Access*, *4*, 2281-2291.

[12] Narudin, F. A., Feizollah, A., Anuar, N. B., & Gani, A. (2016). Evaluation of machine learning classifiers for mobile malware detection. *Soft Computing*, *20*(1), 343-357.

[13] Buczak, A. L., & Guven, E. (2015). A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications surveys & tutorials*, *18*(2), 1153-1176.

[14] Zhang, Z. K., Cho, M. C. Y., Wang, C. W., Hsu, C. W., Chen, C. K., & Shieh, S. (2014, November). IoT security: ongoing challenges and research opportunities. In *2014 IEEE 7th international conference on service-oriented computing and applications* (pp. 230-234). IEEE.

[15] Sivanathan, A., Sherratt, D., Gharakheili, H. H., Radford, A., Wijenayake, C., Vishwanath, A., & Sivaraman, V. (2017, May). Characterizing and classifying IoT traffic in smart cities and campuses. In *2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)* (pp. 559-564). IEEE.

[16] Apthorpe, N., Reisman, D., Sundaresan, S., Narayanan, A., & Feamster, N. (2017). Spying on the smart home: Privacy attacks and defenses on encrypted iot traffic. *arXiv preprint arXiv:1708.05044*.

[17] Engelberg, A., & Wool, A. (2021). Classification of Encrypted IoT Traffic Despite Padding and Shaping. *arXiv preprint arXiv:2110.11188*.

[18] Trimananda, R., Varmarken, J., Markopoulou, A., & Demsky, B. (2020, February). Packet-level signatures for smart home devices. In *Network and Distributed Systems Security (NDSS) Symposium* (Vol. 2020).

[19] Ghasemi, M., Saadaat, M., & Ghollasi, O. (2019). Threats of social engineering attacks against security of Internet of Things (IoT). In *Fundamental research in electrical engineering* (pp. 957-968). Springer, Singapore.

[20] Yekkehkhany, A., Feng, H., & Lavaei, J. (2021, December). Adversarial Attacks on Computation of the Modified Policy Iteration Method. In *2021 60th IEEE Conference on Decision and Control (CDC)* (pp. 49-56). IEEE.

[21] Sagduyu, Y. E., Shi, Y., & Erpek, T. (2019, June). IoT network security from the perspective of adversarial deep learning. In *2019 16th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)* (pp. 1-9). IEEE.

[22] Luo, Z., Zhao, S., Lu, Z., Sagduyu, Y. E., & Xu, J. (2020, July). Adversarial machine learning based partial-model attack in IoT. In *Proceedings of the 2nd ACM Workshop on Wireless Security and Machine Learning* (pp. 13-18).

[23] Winter, P., Pulls, T., & Fuss, J. (2013, November). ScrambleSuit: A polymorphic network protocol to circumvent censorship. In *Proceedings of the 12th ACM workshop on Workshop on privacy in the electronic society* (pp. 213-224).