FEN BİLİMLERİ ENSTİTÜSÜ

ANOMALİ TABANLI SALDIRI TESPİT SİSTEMLERİNDE MAKİNE ÖĞRENİMİ MODELLERİNİN PERFORMANS DEĞERLENDİRMESİ

Mehmet Salih KARAMAN

Danışman
Dr. Öğr. Üyesi Metin TURAN

Eş Danışman
Doç. Dr. Muhammed Ali AYDIN

YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
İSTANBUL - 2020

KABUL VE ONAY SAYFASI

Mehmet Salih KARAMAN tarafından hazırlanan "**Anomali Tabanlı Saldırı Tespit Sistemlerinde Makine Öğrenimi Modellerinin Performans Değerlendirmesi"** adlı tez çalışması 08/09/2020 tarihinde aşağıdaki jüri üyeleri önünde başarı ile savunularak, İstanbul Ticaret Üniversitesi Fen Bilimleri Enstitüsü **Bilgisayar Mühendisliği Anabilim Dalı**'nda **Yüksek Lisans Tezi** olarak kabul edilmiştir.

Danışman	**Dr. Öğr. Üyesi Metin TURAN** İstanbul Ticaret Üniversitesi	.....
Jüri Üyesi	**Doç. Dr. Serhan YARKAN** İstanbul Ticaret Üniversitesi	.....
Jüri Üyesi	**Dr. Öğr. Üyesi Zeynep TURGUT** İstanbul Haliç Üniversitesi	.....

Onay Tarihi : 18.09.2020

İstanbul Ticaret Üniversitesi, Fen Bilimleri Enstitüsünün 18.09.2020 tarih ve 2020/290 numaralı Yönetim Kurulu Kararının 2. maddesi gereğince, ders yüklerini ve tez yükümlülüğünü yerine getirdiği belirlenen "Mehmet Salih KARAMAN" (TC: 22078729478) adlı öğrencinin mezun olmasına oy birliği ile karar verilmiştir.

Prof. Dr. Necip ŞİMŞEK
Enstitü Müdürü

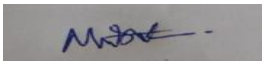
AKADEMİK VE ETİK KURALLARA UYGUNLUK BEYANI

İstanbul Ticaret Üniversitesi, Fen Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmasında,

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- ve bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

18/09/2020



Mehmet Salih KARAMAN

İÇİNDEKİLER

Sayfa

ÖZET

Yüksek Lisans Tezi

ANOMALİ TABANLI SALDIRI TESPİT SİSTEMLERİNDE MAKİNE ÖĞRENİMİ MODELLERİNİN PERFORMANS DEĞERLENDİRMESİ

Mehmet Salih KARAMAN

İstanbul Ticaret Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Metin TURAN

Eş Danışman: Doç. Dr. Muhammed Ali AYDIN
2020, 81 sayfa

Bu çalışmada CSE-CIC-IDS2018 veri seti içerisinde bulunan DDOS, DOS, Bot, Brute-Force ve Infilteration tehditleri toplamda 79 özellik ile ifade edilmiştir. DDOS için 2, DOS için 6, Bot için 4, Brute-Force için 6 ve Infilteration saldırısı için 16 özellik seçilerek tehditlerin tespit edilmesi için oluşturulacak veri setleri küçültülmüştür. Ayrıca herhangi bir tehdidin varlığını bulmak için ise 17 özellik kullanılarak burada da veri setleri küçültülmüştür. Model eğitim ve sınamaları için CSE-CIC-IDS2018 veri seti seçilmiş, Anomali Tabanlı Saldırı Tespit Sistemlerin içinde bu veri setinden DDOS, DOS, Botnet, BruteForce ve Infilteration tehditleri kullanılmıştır. Sınamalar sonucunda, bir saldırı olup olmadığını anlamak üzere Karar Ağaçı modeli diğer yöntemlere göre daha yüksek oranda doğruluk tahmininde (%99,974) bulunmuştur. Ayrıca saldırı tespit türünü belirleme başarısı olarak da, ortalamada Karar Ağaçı modelinin daha başarılı olduğu (%99,81) görülmüştür.

Anahtar Kelimeler: CSE-CIC-IDS2018, makine öğrenmesi yöntemleri, saldırı tespit sistemleri, siber güvenlik.

ABSTRACT

M.Sc. Thesis

PERFORMANCE EVALUATION OF MACHINE LEARNING MODELS IN ANOMALY-BASED INTRUSION DETECTION SYSTEMS

Mehmet Salih KARAMAN

İstanbul Commerce University
Graduate School of Applied and Natural Sciences
Department of Computer Engineering

Supervisor: Assist. Prof. Dr. Metin TURAN

Co-Supervisor: Assoc. Prof. Dr. Muhammed Ali AYDIN
2020, 81 pages

In the this study, DDOS, DOS, Bot, Brute-Force and Infilteration threats in the CSE-CIC-IDS2018 data set were expressed with a total of 79 features. Data sets to be created for detecting threats were minimized by selecting 2 for DDOS, 6 for DOS, 4 for Bot, 6 for Brute-Force and 16 for Infilteration attack. In addition, in order to find the presence of any threat, the data sets have been minimized here by using 17 features. CSE-CIC-IDS2018 data set was selected for model training and testing, DDOS, DOS, Botnet, BruteForce and Infilteration threats were used from this data set in Anomaly Based Attack Detection Systems. As a result of the tests, the Decision Tree model has found a higher rate of accuracy (99.974%) compared to other methods to understand whether there is an attack or not. In addition, it was observed that the Decision Tree model was more successful (99.81%) on average, as a success in determining the type of attack detection.

Keywords: CSE-CIC-IDS2018, cyber security, intrusion detection systems, machine learning methods.

TEŞEKKÜR

Tez çalışmam sırasında gösterdiği her türlü ilgi, destek ve anlayış için değerli danışman hocam sayın Dr. Öğr. Üyesi Metin TURAN 'a, kıymetli bilgi, birikim ve tecrübeleri ile beni yönlendiren ve insani ve ahlaki değerleri ile örnek edindiğim değerli eş danışman hocam sayın Dr. Öğr. Üyesi Muhammed AYDIN 'a, bu günlere gelmemde büyük pay sahibi olan ve beni hiçbir zaman yalnız bırakmayan aileme ve destek ve yardımlarını esirgemeyen başta Tarkan KILIÇ'a ve dostlarıma sonsuz teşekkürler ederim.

Mehmet Salih KARAMAN
İSTANBUL, 2020

ŞEKİLLER

Sayfa

ÇİZELGELER

Sayfa

SİMGELER VE KISALTMALAR

ATSTS	Anomali Tabanlı Saldırı Tespit Sistemleri
DDOS	Distrubuted Denial of Service
DO	Doğruluk Oranı
DOS	Denial of Service
DVM	Destek Vektör Makineleri
ES	Eğitim Süresi
KA	Karar Ağacı
KNN	K-En Yakın Komşu
LR	Logistik Regresyon
MAKS	Maksimum
MEAN	Aritmetik Ortalama
MIN	Minimum
MÖA	Makine Öğrenmesi Algoritması
NB	Naive Bayes
RO	Rastgele Orman
STD	Standart Sapma
STÖS	Saldırı Tespit ve Önleme Sistemleri
STS	Saldırı Tespit Sistemleri
TBA	Temel Bileşen Analizi
TS	Test Süresi
YSA	Yapay Sinir Ağları

1. GİRİŞ

Teknolojinin hızlı bir şekilde ilerlemesiyle, hayatımızın önemli bir bölümünü siber ortamlara yansıtmaktayız. Bununla birlikte devletler, kamu kurumları, sosyal platformlar, özel şirketler; sunmuş oldukları hizmetleri ve ürünleri tüketicilere ve diğer müşterilere ulaştırmak için siber ortamları kullanmaktadır. Bu durumdan faydalanmak ve kazanç sağlamak isteyen kişiler çeşitli yöntemler kullanarak siber ortam kullanıcılarını tehdit etmektedir. Siber saldırıları sayısının ve çeşidinin her geçen gün artmasıyla, siber güvenlik kavramı daha da önem kazanmaktadır. Siber güvenliği tehdit eden unsurların ortaya çıkmasına sebep olan üç boyut bulunmaktadır (Clarke ve Knake, 2015).

- İnternetin tasarımından kaynaklanan zafiyetler (yönetim eksikliği, adresleme sistemi, sistemlerin çoğunun açık ve şifresiz olması, zararlı yazılımları dağıtma kabiliyeti ve internetin merkezi olmayan büyük bir ağ olması)
- Donanımla yazılımlardaki hatalar veya eksiklikler
- Kritik olan sistemlere çevrim içi erişebilme imkanının olması

Siber ortamdaki tehditlerin önüne geçmek için bir takım güvenlik mekanizmalarının geliştirilmesini her geçen gün daha fazla ihtiyaç duyulmaktadır. Bu mekanizmanın parçalarından biri de Saldırı Tespit ve Önleme Sistemleridir (STÖS).

STÖS bir ağdaki kayıtları tutmak, olası tehditleri tanımlamak, saldırıları durdurmak ve ağın güvenliğinden sorumlu yöneticilere durum raporlarını ileterek gerekli önlemlerin alınmasını sağlar. Bu sistemleri bazı kurumlar, bazı durumlarda güvenlik politikalarındaki zayıflıkları tespit etmek içinde kullanabilmektedir. Ayrıca STÖS saldırganların ağla ilgili bilgi toplama faaliyetlerini algılayıp erken aşamada tehlikeyi durdurabilme işlemini gerçekleştirmektedirler.

Saldırı Tespit Sistemleri (STS) genel olarak iki alt başlıkta sınıflandırılırlar (Özgür ve Erdem, 2012):

- Sistemlerin kurulumuna göre
- Ağ STS (Network IDS)
- Ev Sahibi STS (Host IDS)
- Sistemin Tespit Mantığına Göre
- İmza Tabanlı STS
- Anomali Tabanlı STS

Anomali Tabanlı STS (ATSTS), daha önceden belirlenmiş olan normal trafiği ağdaki trafik ile karşılaştırarak normalden sapmaları tespit etmeye çalışır. ATSTS 'nin en büyük avantajı bilinmeyen tehditleri tespit edebilmesidir.

ATSTS' lerin geliştirilmesinde Makine Öğrenme Yöntemleri kullanılmaktadır. Bu çalışmada sınıflandırma algoritmalarından olan Yapay Sinir Ağları (YSA), K-En Yakın Komşu (KNN), Lojistik Regresyon (LR), Destek Vektör Makineleri (DVM), Karar Ağacı (KA), Naive Bayes (NB) ve Rassal Orman (RO)'ın karşılaştırmalı analizleri yapılarak en iyi sonucun bulunması sağlanmıştır.

ATSTS için uygun modellerin oluşturulmasında birçok Makine Öğrenmesi Yöntemi ve birçok veri seti kullanılmıştır. Kullanılan veri setlerindeki bazı eksiklikler de göz önüne alınarak Kanada Siber Güvenlik Enstitüsü tarafından paylaşılan CSE-CIC-IDS2018 veri seti oluşturulmuştur. Bu çalışmada CSE-CIC-IDS2018 (CyberSecurity, 2018) veri seti kullanılmıştır.

Bu çalışmada CSE-CIC-IDS2018 (CyberSecurity, CSE-CIC-IDS2018, 2018) veri seti içerisinde bulunan her bir özellik 79 özellik ile ifade edilmiştir. Modellerin performansının iyileştirilmesi için özellik sayıları BruteForce yöntemi kullanılarak küçültülmüştür. Distrubuted Denial of Service (DDOS) için 2, Denial of Serive (DOS) için 6, Bot için 4, Brute-Force için 6 ve Infilteration saldırısı için 16 özellik kullanılarak veri setleri oluşturulmuştur. Bununla birlikte herhangi bir tehdidin varlığını bulunması için 17 özellik kullanılarak veri seti

oluşturulmuştur. Makine Öğrenmesi Yöntemleri, en iyi sonucu gösterdikleri başarım oranları kıyaslanarak performans değerlendirmeleri yapılmıştır. Modellerin eğitim ve sınamaları için CSE-CIC-IDS2018 veri seti seçilmiştir. Sınamalar sonucunda, herhangi bir saldırı olup olmadığını anlamak üzere KA modeli diğer yöntemlere göre daha yüksek oranda doğruluk tahmininde (%99,974) bulunmuştur. Saldırı tespit türünü belirleme başarısı olarak da ortalamada KA modelinin yine daha başarılı olduğu (%99,81) görülmüştür.

1.1 Tezin Amacı

Bu tez kapsamında güncel veri seti olan CSE-CIC-IDS2018 kullanılarak saldırı türlerine ait en iyi özelliklerin bulunarak veri seti içerisindeki örneklemleri daha az sayıda özellik ile ifade ederek veri setlerinin eğitim ve test zaman maliyetleri düşürülmesine odaklanılmıştır. BruteForce Yöntemi' nin kullanıldığı bu aşamada her bir saldırı türü için performansının yüksek ve özellik sayısının az olması hedeflenmiştir.

Veri setlerinin oluşturulması ile veri setleri üzerinde Makine Öğrenmesi Yöntemleri uygulanmıştır. Her bir saldırı türüne ait veri seti için sınıflandırma algoritmaları kullanılarak performanslarının değerlendirilmesi amaçlanmıştır.

1.2 Tezin Katkısı

CSE-CIC-IDS2018 veri seti güncel ve eski veri setlerinde görülen eksiklikler dikkate alınarak uygun test ortamında hazırlanmıştır. Bu veri seti içerisinde, yaygın saldırı türleri olmasıyla birlikte her saldırı türüne ait yeteri kadar örneklem vardır.

Bu çalışma güncel veri seti içerisindeki her bir tehdit için en iyi özelliklerin bulunmasını sağlayarak oluşturulan veri setlerinin boyutlarının azaltılması sağlanmıştır. Ayrıca oluşturulan veri setlerine uygulanan sınıflandırma algoritmaları ile %99 oranında başarı sağladığı gözlemlenmiştir.

1.3 Tezin Düzeni

Giriş bölümünde tezin genel amacı, katkıları ve kapsamından bahsedilmiştir.

Literatür Özeti bölümünde konuyla ilgili yapılmış çalışmalardan ve tezin sağladığı katkı belirtilmiştir.

Siber güvenlik ve saldırı tespit sistemleri bölümünde genel bilgi güvenliğin öneminden ve bilgi güvenliğini sağlayan araçlardan biri olan saldırı tespit sistemleri anlatılmıştır.

Makine öğrenme yöntemleri bölümünde saldırının tespitinde uygun modelin oluşturulması için çalışmada kullanılan makine öğrenme yöntemleri açıklanmıştır.

Veri seti ve uygulama bölümünde çalışmada kullanılan veri setinden ve kullanılan makine öğrenmesi yöntemlerinden elde edilen sonuçlar verilmiştir.

Sonuç ve öneriler bölümünde yapılan çalışma sonucunda elde edilen sonuçlardan ve gelecekte yapılabilecek çalışmalardan bahsedilmiştir.

2. LİTERATÜR ÖZETİ

ATSTS'lerin geliştirilmesinde birçok makine öğrenmesi yöntemi kullanılmaktadır. Kullanılan Makine Öğrenmesi Yöntemleri veri setinin yapısına ve oluşturulan modele göre farklı sonuçlar vermektedir. Elde edilen sonuçların doğruluk oranları eğitim ve test süreleri performanslarının ölçülmesinde kullanılan önemli parametrelerdir.

Kaynak kullanımını iyileştirmek ve zaman karmaşıklığını azaltmak üzere yapılan çalışmada, 41 öz nitelikli KDD-99 veri seti ve bu verinin 25 öz nitelikli hali, YSA kullanılarak test edilmiştir. Sonuçlar incelendiğinde, azaltılmış veri setinin saldırının olup olmadığımı tahmin yüzdesinin daha yüksek olduğu görülmüştür (Manzoor ve Kumar, 2017). KDD-99 veri seti üzerinde yapılan çalışmada Geri Beslemeli Sinir Ağı algoritması kullanılarak, saldırıları sınıflandıran bir STS tasarlanmıştır. Test sonuçları incelendiğinde, KDD-99 ile yapılmış diğer çalışmalara göre farklı saldırıları tespit etmede daha az sayıda yanlış pozitif ve yanlış negatif verdiği gözlemlenmiştir (Poojitha vd., 2010). KDD-99 veri kümesindeki tüm özellikleri kullanılarak tasarlanan ATSTS'de YSA ve DVM performansı karşılaştırılmıştır. YSA'nın saldırı tespit doğruluğunun DVM' den daha iyi bir performansa sahip olduğu görülmüştür (Cahyo vd., 2016).

NSL-KDD veri seti için TBA (Temel Bileşen Analizi / PCA) kullanılarak 6, 7, 11, 21 öz nitelik belirlenmiş, Naive Bayes, J-48, KNN-1 ve KNN-3 algoritmaları üzerinde 4 farklı veri seti kullanılarak, eğitim ve test işlemleri uygulanmıştır. Tehdit olup olmadığı, %99,6871 doğruluk oranı ile TBA-21 veri setinde, KNN-1 algoritmasıyla elde edilmiştir. Fakat TBA-21 verisinin eğitim ve test süresi, diğer verilerin test ve eğitim sürelerine oranla oldukça yüksek çıktığı görülmektedir (Çavuşoğlu ve Kaçar, 2019).

Popüler bir ağ izleme aracı olan Bro kullanılarak log kayıtları oluşturulmuş, Bot saldırıların tespiti için kullanılmıştır. Burada LR algoritması uygulanarak oluşturulan modelde yapılan test sonuçlarında doğruluk tahmininin %96.7

olduğu görülmüştür (Bapat, et al., 2018). Yapılan başka bir çalışmada (Mok, Sohn, ve Ju, 2010), KDD-cup 1999 (Veri Madenciliği ve Bilgi Keşfi yarışması) veri setinin "normal" ve "anormallik" bağlantılarını içeren 42 değişkeni için 49.427 rasgele gözlem örneğini içeren veri setini üzerinde lojistik regresyon modeli uygulanmıştır. Önerilen model, eğitim veri seti için %98.94'lük bir sınıflandırma doğruluğuna sahipken, doğrulama veri seti için başarısı %98.68'dir.

YSA ve DVM kullanılarak, DARPA veri seti üzerindeki performansının değerlendirildiği çalışmada (Mukkamala vd., 2002), test sonucunda iki yöntemin doğruluk oranları birbirine çok yakın çıkmıştır (%99'un üzerinde). Ama yeniden eğitimin yapılması gereken durumlar göz önüne alındığında, bu iki yöntemin eğitim süreleri: DVM 17.77 sn. YSA 18 dk sürmüştür. Bu durum göz önüne alındığında DVM'nin performansının daha iyi olduğu değerlendirilmiştir. Nitelik değerlerinin model üzerindeki etkilerini değerlendiren diğer bir çalışmada (Heba vd., 2010), NSL-KDD veri setinin öz nitelik sayısı TBA kullanılarak 41'den 23'e indirilmiştir. Öz niteliği azaltılmış veri seti ve azaltılmamış veri seti üzerinde, DVM kullanılarak yapılan deneylerde STS' nin, doğruluk oranlarının iki veri seti içinde birbirine yakın (%95) olduğu gözlemlenmiştir. Fakat azaltılmış veri setinin eğitim ve test süreleri 2 kata yakın daha az süre olduğu gözlemlenmiştir.

Yeni bir etiketli ağ veri kümesi olarak Kyoto 2006+ veri kümesi kullanılarak, NIDS ağ paketini sınıflandırmak için KA (J48) algoritmasını kullanılmıştır. Bu veri setinde eğitim ve test için 134665 ağ örneği kullanılmış olup, oluşturulan kurallar neticesinde modelin % 97,2 doğrulukla çalıştığı gözlemlenmiştir (Sahu ve Mehtre, 2015). KDD CUP1999 veri seti kullanılarak yapılan bir başka önemli çalışmada ise, C4.5 KA sınıflandırma yöntemi kullanılmıştır. Yeni ağ davranışının normal veya anormal olup olmadığını değerlendirmek üzere kullanılan modelin tespit doğruluk oranı % 90 üzerinde olup, saldırı tespitinde etkili bir yöntem olduğu değerlendirilmiştir (Wang, vd., 2009).

KA ve RO yöntemleri kullanılarak yapılan çalışmada, DOS, DDOS ve Port Tarama saldırıları anormal olarak etiketlenmiş ve deney sonuçlarında doğruluk oranı %99,966 ile RO daha başarılı sonuç vermiştir (Özekes ve Karakoç, 2019). Dikkat çeken bir başka çalışmada (Sharafaldin vd., 2018), her tehdidi belirleyen en iyi dört özelliğin bulunması amacı ile RandomForestRegressor kullanılmış KNN, RO, ID3, Adaboost, çok katmanlı algılayıcı, NB ve ikinci dereceden ayırım analizi yöntemleri kullanılarak eğitildiğinde, RO ve ID3'ün % 98 ile en iyi tahminleri yaptıkları gözlemlenmiştir. Test ve eğitim süreleri de dikkate alındığında RO, 74,39 ID3 ise 235,02 sn. olduğu ve en iyi performansın RO yönteminin olduğu görülmüştür.

KDD-99 veri seti kullanılarak, DOS, R2L, U2R, Probing tehditlerinin tespitinde NB ve KA tekniklerinin performansı test edilmiş ve test sonuçlarında NB'in R2L ve Probing tehditleri için doğruluk oranının daha yüksek olduğu gözlemlenmiştir (Amor vd.,2004). Bu modeli kullanan bir başka çalışmada (Mukherjeea ve Sharma, 2012), 41 özellikli NSL-KDD veri setinden öz nitelik sayısı 10, 14, 20, 24 olarak veri setleri oluşturulmuş ve azaltılmış ve azaltılmamış özellikli veri setlerine NB algoritması uygulanarak test edilmiştir. Test sonuçlarına göre, doğruluk oranı 24 öz nitelikli veri setinde en yüksek çıkmıştır. Tüm öz niteliklerin kullanılması ile elde edilen doğruluk oranı %95.11 iken, 24 öz nitelikli veri ile elde edilen doğruluk oranı %97.78'dir.

3. SİBER GÜVENLİK VE SALDIRI TESPİT SİSTEMLERİ

3.1. Bilgi Güvenliği

Bilgi, işlenmiş, derlenmiş, organize edilmiş, yorumlanmış, karar alabilmede kullanılabilmek için bir işlem sürecinden geçirilerek, anlamlı ve değerli hale dönüştürmüş, kararları ve davranışları etkileyen değerdir (Çetin, 2014). Bilginin değerinin arttığı günümüzde şirketlerin gücü ve sahip oldukları değerler yalnızca mal varlıklarıyla değil, sahip oldukları insan gücü, var olan müşteri potansiyeli ve kendi yarattıkları bilgi birikimleriyle ölçülmektedir. 1998 yılında 25 milyon dolar sermaye ile kurulmuş olan ve internet üzerinde bilgi arama motoru olarak kullanılan ve 2013 yılında 110,92 trilyon dolar piyasa değerine ulaşmış olan Google buna en çarpıcı örnektir (Vikipedi, 2014).

Bilgi sistemlerinin teknolojik olarak gelişmesi, kişilerin, toplumun ve kamunun yararına çok fazla olumlu katkılar sağlamış olmasına rağmen suç kavramını ve farklı tehditlerin ortaya çıkarak bir takım olumsuz sonuçlar doğurmasına sebep olmuştur. Çünkü bilgi teknolojileri birçok yeni tehdit türünün ve suç unsurunun ortaya çıktığı bir alan haline gelmiştir. Bilgi teknolojilerini tehdit eden bu unsurlara karşı bilgi güvenliği kavramı karşımıza çıkmaktadır.

Bilgi güvenliği, bir kaynak türü olarak bilginin izinsiz veya yetkisiz bir biçimde erişebilme, kullanabilme, değiştirilebilme, ifşa edebilme, ortadan kaldırılabilme, el değiştirme veya hasar görmesinin önüne geçmek olarak tanımlanır. Bilgi güvenliğinin tam olarak sağlanabilmesi için gizlilik, bütün ve erişilebilirlik unsurlarının sağlanabilmesi gerekmektedir. Bu unsurlardan herhangi bir tanesinin sağlanamaması durumunda zafiyet oluşmaktadır.

Gizlilik: Bilginin bilgisayar sistemlerinde işlenirken, saklama ortamlarında depolanırken ve ağ üzerinde alıcı ve gönderici arasında iletilirken yetkisiz kişiler tarafından ele geçirilmesinin önüne geçmeyi amaçlamaktadır. Burada önemli noktalardan bir tanesi herhangi bir yazılım hatası veya açıklıktan ya da

sosyal mühendislik yöntemleri ile yetkili insanların hatalarını istismar ederek erişebilmesinin önüne geçmek için bilginin tamamen gizlenmesi değil yetkisiz bir şekilde erişebilmesini engellemek ve bu yöndeki herhangi bir girişimde bilgi verilmesini sağlamaktır.

Bütünlük: Bilginin bozulmasını, değiştirilmesini, yeni bilgiler eklenmesini bir kısmını veya tamamının silinmesini engelleyerek bilgiyi olması gerektiği şekilde tutmayı ve korumayı hedeflemektedir. Bütünlük prensibi temel olarak Sistem ve veri bütünlüğünü kapsar. Bu prensip için önemli olan nokta ise bilginin değiştirilmemesini sağlamak olmayıp yetkisiz bir şekilde değiştirilmesini engellemek ve yetkisiz kişiler tarafından değiştirildiği durumda haber verilmesini sağlamaktır.

Erişilebilirlik: Bilginin hedeflenen ve ihtiyaç duyulan süre zarflarında ulaşılabilir ve kullanılabilir olmasını tam ve eksiksiz bir biçimde yapılmasını gerektiren prensiptir. Bu prensip için önemli olan nokta hedef sistem sürekli olarak değil belirlenen süre boyunca erişilebilir olmasının sağlanmasıdır.

Bilgi güvenlik sistemlerde bu üç unsurun dışında sağlanması gereken diğer unsurlar şu şekildedir:

- Bilginin güvenilir olması
- Bilginin inkar edilememesi
- Kimliklendirme
- Kimlik sınaması
- Yetkilendirme
- İzlenebilirlik
- Hesap verilebilirlik

3.2. Siber Saldırı

Siber saldırılar, bilgisayarları, sunucuları, mobil cihazları, elektronik sistemleri ağları ve verileri hedef alarak tehdit oluşturan unsurlardır. Küresel siber tehditler her geçen yıl sayısı artarak veri ihlalleri hızlı bir şekilde büyümektedir.

3.2.1 Temel kavramlar

Siber Uzay: Tüm dünyaya yayılmış durumda bulunan bilişim sistemlerinden ve bunları birbirine bağlayan ağlardan veya bağımsız bilgi sistemlerinden oluşan ortamdır.

Siber Varlık: Bilgisayarlar, telefonlar, sunucular ve ağ cihazlarının yanı sıra internete erişebilen diğer elektronik cihazların tümüne verilen addır.

Siber Olay: Bilişim sistemlerinin veya bu sistemler aracılığıyla işlenen verilerin gizlilik, bütünlük ve erişilebilirliğinin ihlal edilmesidir.

Zayıflık: Bilişim sistemlerindeki yanlışlıkla ya da kasıtlı olarak bırakılmış ve saldırganlar tarafından bu durumun kullanılarak sisteme zarar vermesine sebep olan güvenlik açıklarına verilen genel addır. Diğer bir deyişle sistemdeki zayıflık ve boşluk olarak tanımlanabilir. Güvenlik açıklıklarının altında yatan en önemli sebeplerden bir tanesi geliştirme sürecinde geliştiricinin ön göremediği durumlardan kaynaklanmaktadır. Güvenlik açıklıklarını ve sistemin zayıflığını ölçmek için sızma testleri kullanılır.

Tehdit: Sistemin zayıflıklarını kullanarak sisteme zarar verici davranışlarda bulunma olasılığı olan eylemlerin tamamına denmektedir. Bilişim sistemlerinde bilgiye erişmek ve bilgiyi kendi çıkarları için kullanarak ya da sistemi çalıştırılamaz veya güvenilemez hale getirmek için mevcut imkanlardır (Anderson, 1980).

Siber Tehdit İstihbaratı: Bir kurum veya kuruluşun bilişim sistemleri varlıklarına karşı oluşturulan zararlı yazılımların, tehditlerin; bu tehdit ve zararlı yazılımların kim tarafından, hangi nedenle, kimi hedef aldığına dair gibi bilgilerin toplanarak analiz edilmesi, buna karşı gerekli aksiyonların alınarak elde edilen bilgilerin iş birliği içerisinde bulunduğu kurum ve kuruluşlar ile paylaşılmasını esas alır (Yelken, 2019).

Saldırı: Bir kurum veya kuruluşun bilgi sistemlerine maddi veya manevi menfaat sağlamak amacıyla gerçekleştirdiği eylemlerin tamamına denir. Bilgi sistemlerinin gelişmesiyle birlikte bunlara yönelik saldırı türleri de her geçen gün artmaktadır. Saldırının türü sistemin zafiyetlerine ve saldırganın amacına yönelik olarak farklılık göstermektedir.

Hacker: Bir sistemin engellerini aşarak sisteme sızan ve gizli ve erişilemeyen bilgilerini ele geçirmek isteyen kişilere denir. Hackerler sistem içerisinde izinsiz olarak dolaşabilir, eriştiği bilgileri çalabilir, değiştirebilir ve silebilir.

Hackerler niyetlerine göre üç gruba ayrılmaktadırlar:

- Beyaz şapkalı hacker
- Gri şapkalı hacker
- Siyah şapkalı hacker

Beyaz Şapkalı Hacker: Bu kişiler sistem ile ilgili güvenlik açıklarını bularak bulmuş oldukları açıklıkları kapatma yöntemleri ile birlikte sistem yetkililerine bilgi verirler. Bunlara iyi niyetli hacker denmektedir.

World Wide Web (WWW) sistemini ve HTTP sistemini hayatımıza kazandıran Web'in babası Tim Berners-Lee beyaz şapkalı hackerdir (Kara M. , 2013).

Gri Şapkalı Hacker: Menfaatleri doğrultusunda iyi veya kötü niyetli olabilen hackerlerdir. Kimilerine göre etik, kimilerine göre etik olmayan hackerlerdir.

Siyah Şapkalı hacker: Kötü niyetli olan bu hacker türü oldukça tehlikelidirler. Sistem üzerinde bulmuş oldukları açıkları kullanarak gizli bilgileri ele geçirirler, sistemi çalışamaz hale getirebilirler.

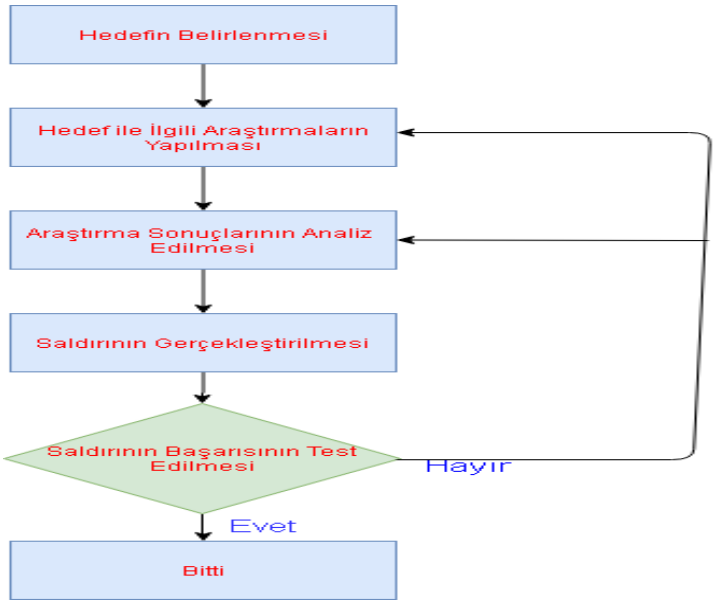
Siyah şapkalı hackerler, ulaşmış oldukları gizli bilgiler ile maddi ve manevi zarar vermesi ile birlikte korunan sistemin açıklıklarını ortaya koyarlar.

İstismar: Bilgisayar sistemine, özellikle de yetkili bir hizmet içindeki istenmeyen, beklenmeyen veya ön görülmemiş bir güvenlik açığından faydalanan saldırıdır. İstismar sonucunda bir bilgisayar sistemini ele geçirme, ayrıcalık yükseltmesine izin verme ve hizmet reddi gibi saldırıların başlatmak için ihlale yol açar.

İhlal: Bir kişinin herhangi bir şekilde bilgi işlem kaynaklarını kötüye kullanması ve bilgilere yetkisiz erişimi sağladığında güvenlik ihlali oluşur. Güvenlik ihlalleri genellikle gizliliğin, bütünlüğün veya erişilebilirliğin prensiplerine uyulmaması durumunu veya bilgi veya bilgi kaynaklarının kaybedilmesi gibi durumları ortaya çıkarır.

Sosyal Mühendislik: Hassas ve önemli bilgilere ve ağ sistemlerine, yasal kullanıcılar üzerinden erişim sağlama amacıyla saldırganlar tarafından kullanılan düşük teknolojili ve insan zafiyetlerine dayalı yöntemlerin tamamıdır (Gündüz ve Daş, 2016). En gelişmiş teknolojilerin başında bile insanların kontrolü olduğunu göz önünde bulundurursak, saldırganlar için o sistemlere sızmak kişinin zafiyetlerini kullanarak daha etkili olabilmektedir. Şekil-3-1'de sosyal mühendislik süreci gösterilmektedir.

Sosyal mühendislik kullanılarak yapılan birçok saldırı mevcuttur. Bu saldırılar için yaygın olarak kullanılan yöntemler ve alınması gereken önlemler çizelge 3-1'de verilmiştir (Canberk ve Sağıroğlu, 2017).



Şekil 3-1 Sosyal mühendislik akış şeması

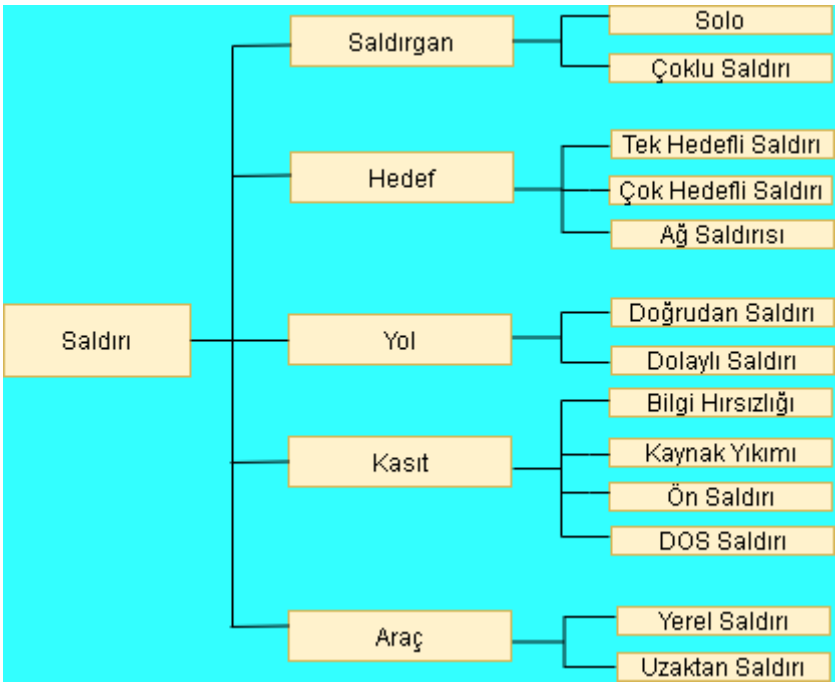
Çizelge 3-1 Sosyal mühendislik taktikleri ve alınması gereken önlemler

Risk Alanı		**Korsan Taktiği**	**Mücadele Stratejisi**
Telefon (Yardım Masası)		Taklit ve inandırma	Çalışanların ve yardım masasının telefonla hiçbir şekilde şifre veya diğer gizli bilgilerin paylaşılmaması konusunda eğitilmesi
Binaya giriş		Yetkisiz fiziksel erişim	Sıkı kimlik kartı güvenliği, çalışanların eğitilmesi ve güvenlik görevlisi çalıştırılması
Ofis		Omuz sörfü (Klavye ile yazı yazarken çevredeki birilerinin sizi izlemesi)	Sizden başka birilerinin ortamda bulunduğu durumlarda şifrelerinizi girmeyiniz
Telefon (Yardım Masası ofis)		Yardım masası aramalarında taklit etme Kimsenin olmadığı açık odalar bulabilmek için koridorlarda dolaşma	Bütün çalışanlara yardım masası desteği alabilmesi için tekil bir PIN numarası atanması Bütün misafirlere iş yerinden birilerinin refakat etmesi

Posta odası		Sahte notların sokulması	Posta odası kilitlemek ve izlemeye tabi tutma
Makine odası-santral		Erişmeye teşebbüs, cihazların kaldırılması ve gizli bilgileri elde edebilmek için bir protokol analizcisi eklenmesi	Santral, sunucu odaları vs. her zaman kilitli tutulmalı ve cihazların güncel envanterlerini tutma
Telefon ve PBX		Telefon görüşme ücreti erişimi çalma	Şehirlerarası, milletlerarası ve cep telefonları aramalarını kontrol et, konuşmaları izle ve aktarmaları reddet
İş yeri atık deposu		Çöplük karıştırma	Bütün çöp kutularını güvenli ve izlenen alanlarda tutulmalı, önemli belgeler kesme makinesiyle yok edilmeli ve manyetik ortamdaki veriler silinmeli
İntranet-İnternet		Şifrelerin çalınması için intranet ve internet üzerinde sahte yazılımların oluşturulması ve konulması	Sistem ve ağ değişikliklerinden sürekli haberdar olma ve şifre kullanımı eğitimi
Ofis		Hassas belgelerin çalınması	Belgelere gizlilik derecesinin verilmesi ve bu belgelerin kilitli yerlerde saklanması
Genel-psikolojik		Taklit ve ikna	Çalışanların sürekli uyanık tutulması ve eğitim programları ile bilinçlendirilmesi

3.3. Siber Saldırı Çeşitleri

Saldırıların Sınıflandırılması Şekil 3-2' de saldırgan sayısına, hedef türüne, kullanılan yola, kasıt ve araçlara göre sınıflandırılmıştır (Canberk ve Sağıroğlu, 2017).



Şekil 3-2 Saldırıların sınıflandırılması

3.4. Saldırı Çeşitleri

3.4.1 DOS ve DDOS saldırıları

İnternete bağlı bir hostun hizmetlerini geçici veya süresiz bir şekilde aksatarak, bir makinenin veya ağ kaynaklarının asıl kullanıcılar tarafından erişilmesinin engellemeye dayanan bir siber atak türüdür. Saldırıyı gerçekleştiren bilgisayar web sitesine istek yollar ve bunu sürekli olarak tekrarlar. Bunun sonucu olarak ana bilgisayarda oluşan aşırı yüklenme ile hizmetler erişilemez duruma gelir.

DDOS saldırısı, kurban web kaynağına çok fazla istek gönderilerek web sitesinin çok sayıda isteği işleme kapasitesini aşmasını ve doğru şekilde çalışmasını engellemeyi hedefler. Bu saldırılar ilk olarak 1998 yılının haziran ayında ortaya çıkmıştır (Lin ve Tsen, 2004).

DOS ve DDOS Saldırılarının bazı belirtileri (Gezgin ve Buluş, 2013):

- Beklenmedik bir biçimde ağın düşük performans göstermesi
- Web sitelerine erişimde yaşanan yavaşlıklar

- Ağ bağlantılarının kesilmesi
- Spam e-postaların sayısında yaşanan artış
- Web sitesinin belli bölümlerine erişimin imkansız hale gelmesi
- Google Analystıc gibi istatistiki bilgiler veren sitelerde, sitenizde aniden artan istatistiki veriler

DOS ve DDOS saldırı türleri:

- Hacim odaklı saldırılar
- Protokol odaklı saldırılar
- Uygulama katmanlı saldırılar
- HTTP Flood
- UDP Flood
- Icmp flood
- Ping of Death
- Syn flood
- Teardrop
- Smurf
- DNS zehirlenmesi

3.4.2 Bot saldırıları

Botnet sözcüğü, "robot" ve "network" sözcüklerinin bir araya gelmesi ile oluşturulmuştur. Botnetler genelde tek bir merkezden yönetilerek botların bir koordinasyon içerisinde belli amaçlar için yönlendirilmesinde kullanılmaktadırlar.

Botnet saldırıları 1999'dan beri bilinmelerine rağmen 2007 yılının sonunda tüm güvenlik endüstrileri, botnetleri güvenlik sistemlerinin en önemli tehdidi olarak kabul etmiştir (Kara ve Şişeci, 2011). Çizelge 3-2'de bazı Botnetlerin tarihsel listesi verilmiştir (Vikipedi, 2020).

Çizelge 3-2 Botnetlerin tarihsel listesi

Oluşturma Tarihi	Sökülme Tarihi	Adı	Botların Tahmini Sayısı	Spam Hacmi (milyar /gün)	Takma Adları
2006	2011	Rustock	150 bin	30	RKRustok,Costrat
2007	2008	Srizbi	450 bin	60	Cbeplay,Exchanger
2008	2010	Waledac	80 bin	1,5	Waled,Waledpak
2009	2010	Bredolab	30.milyon	3,6	Oficla
2009	2012	Grum	560 bin	39,9	Tedroo
2010	2011, 2012	Kelihos	300 bin+	4	Hlux

3.4.3 Kaba kuvvet saldırısı (brute force attack)

Kaba Kuvvet saldırıları hedef kaynak üzerinde kayıtlı şifrelerin bulunması maksadıyla deneme yanılma yöntemleri kullanılarak yapılmaktadır. Kaba kuvvet saldırılarının çalışma mantığı algoritmalar oluşturularak bütün ihtimallerin dene yanılma yolu ile denenmesidir. Tehdit bu saldırıyı kullanarak bilgisayarların kullanıcı bilgilerini, kredi kartı bilgileri, sosyal medya platformlardaki hesap bilgileri, kurumsal bilgiler ve özel bilgileri elde etmeyi amaçlamaktadır.

Kaba kuvvet saldırılarından korunmak için alınması gereken bazı önlemler:

- Karmaşık ve tahmin edilemeyecek şifrelerin kullanılması
- Verileri gönderirken ve saklarken şifreleme algoritmalarının kullanılması
- İki seviyeli kimlik doğrulamanın kullanılması

3.4.4 Sızma saldırısı (penetration attack)

Bir ağın veya sistemin zafiyetlerinden faydalanarak bilgi toplama, bilgi değiştirme veya bilgi silme gibi amaçları gözeterek sızmaya çalışır. Saldırgan

bunun için keşifler yapabilir veya sistem hakkından daha öncesinde bilgi sahibi olduğu için amacına kolay bir şekilde erişebilmektedir.

Bilgisayar sistemlerini sızmalardan koruyabilmek için ağ ve sistemdeki güvenlik açıklarının bulunarak bu açıklıkların kapatılması ve saldırının önüne geçilmesi gerekmektedir. Bu durum için çoğu kurum sistemdeki ve ağdaki açıklıkları bulmak için sızma testlerine tabii tutarlar.

Sızma testleri, sisteme veya ağa yapılan gerçek zamanlı yapılan saldırılara karşı ne kadar savunmasız olduklarını ortaya koymak amacıyla işlenen bir prosedürdür.

Sızma testleri üç farklı yaklaşım ile uygulanabilmektedir:

- Beyaz Kutu Testi: Bu yaklaşım türünde saldırıyı gerçekleştirmeden önce bilgi toplama aşaması sistem yöneticisi tarafından sağlanır ve sadece zafiyet tarama işlemleri ile test yapılır.
- Gri Kutu Testi: Bu yaklaşım türünde saldırıyı gerçekleştirmeden önce sistem ile ilgili olarak çok az bilgiye sahip olarak yapılan sızma testidir.
- Siyah Kutu Testi: Bu yaklaşım türünde sistem hakkında herhangi bir bilgi olmaksızın gerçek bir saldırgan bakış açısı ile yapılan sızma testidir.

3.4.5 SQL injection

SQL verilerin yönetilmesi ve tasarlanması için kullanılan bir veri tabanı yönetim sistemidir. Günümüzde birçok veri tabanı SQL ile yazılmış komutları uygulamak üzere tasarlanmış ve web sitesi kullanıcılarının bilgileri ise kendi SQL veri tabanlarında saklanmaktadır. SQL Injection saldırısı, veri tabanını ele geçirebilecek veya yok edebilecek kodun SQL işlenmesi ile gerçekleştirilir. SQL Injection ile kötü amaçlı kod web sayfalarına SQL ifadelerine yerleştirilebilir. SQL Injection, uygulama yazılımları içindeki güvenlik açıklıklarından faydalanarak uygulanan en yaygın web saldırı yöntemlerinden biridir.

3.4.6 Sıfır gün saldırısı (zero day attack)

Daha öncesinde karşılaşılmayan veya hiç görülmemiş saldırılar sıfır gün saldırısı olarak kabul edilir. Sıfır gün açıkları, saldırı gerçekleştirilene kadar yazılım/donanım geliştiricisi tarafından anlaşılması zor olan zafiyetlerdir. Şekil-3-3'te Sıfır Gün Saldırısı'nın aşamaları verilmiştir.

Şekil 3-3 Sıfır gün saldırısının süreci

3.4.7 Zararlı yazılımlar (malware)

Kötü amaçlı yazılımlar ağa, sisteme ve bilgisayara zarar vermek için tasarlanmış her türlü yazılımlara denir. Kötü amaçlı yazılımlar kullanılarak bilgisayardan hassas bilgiler çalınabilir, bilgisayarın performansı düşürebilir, dosyalar şifrelenebilir veya sahte e-posta hesaplarına sahte e-postalar gönderilebilir.

- Truva Atı: Gerçek bir yazılım gibi görünen bir tür kötü amaçlı yazılımdır. Truva atları etkileşime geçtikten sonra siber tehditler tarafından veri silme, veri engelleme, verileri değiştirme, verileri kopyalama ve ağ

performansının düşürülmesi gibi birçok kötü eylemler gerçekleştirilebilmektedir.

Truva atları, virüsler ve solucanlar gibi kendi kendilerine çoğalamazlar.

- Solucanlar: Yerel sürücü de ya da kendini sürekli olarak kopyalayan bir programdır. İşletim sistemlerinin açıklıklarını kullanarak yayılırlar. Herhangi bir sisteme ya da dosyaya zarar vermez fakat sürekli olarak çoğaldığı için sistemin performansını düşürür.
- Spam: Aynı mesajdan çok sayıda yollanması ile mail adresini, forumu boğmaya çalışır. Spamların çoğu reklam amaçlı olup kullanıcıların istekleri dışında gönderilmektedir.
- Casus Yazılım: Casus yazılımlar, kullanıcı izniyle veya kullanıcı izni olmadan bilgisayara yüklenen ve bilgisayar içerisinde yer alan bilgileri toplayarak uzaktaki bir kullanıcıya gönderen bir program türüdür. Bununla birlikte bilgisayarlara başka zararlı yazılımları da indirip yükleyebilmektedir.
- Adware: Bir program çalışırken bununla birlikte reklam açan yazılımlardır. Adware yazılımları, daha çok reklam amacıyla şirketler tarafından kullanılmaktadır.
- Fidye Yazılımları: Saldırganların cihazı kilitleme veya dosyaları şifreleme yolunu kullanarak kurbandan maddi veya manevi menfaat sağlamaya çalıştığı yazılımlardır. Bu saldırının en yaygın biçimi, kullanıcılara e-posta yolu ile zararlı yazılımı gönderip bunu kurmasını sağlayarak hedefine ulaşmaktır.

3.4.8 Oltalama saldırıları (fishing attack)

Web sitesi ve e-posta kanallarını kullanarak kişisel bilgileri toplamak için kullanılan saldırı türüdür. Amaç, kurbanı ihtiyaç duyulan bir şeye inandırmak ve bu durumdan menfaat sağlamaktır.

Oltalama saldırıları her geçen gün daha yaygın hale gelmektedir. Gartner tarafından yapılan araştırmaya göre, 57 milyon ABD internet kullanıcısı oltalama dolandırıcılığı ile bağlantılı olarak e-posta aldığı ve bunların 2 milyonun hassas bilgi vermek için kandırıldıkları düşünülmektedir (Kirda ve Kruegel, 2006).

3.4.9 Ortadaki adam saldırısı (main in the middle attack)

Ortadaki adam saldırısı, hedef bilgisayar ve o bilgisayarın iletişim kurduğu ağ araçlarının (server, modem, router, switch) arasındaki iletişimi dinleyerek gerçekleştirilen saldırıdır.

Ağ ortamlarında veri paketleri şifrelenmeden iletilmektedir. Bu durumda saldırganlardan tarafından kullanılarak paketlerdeki bilgiler dinlenebilir ve değişiklik yapılabilir. Ortadaki adam saldırısının en çok kullanılan yöntemi ARP zehirlenmesi yöntemidir.

ARP Zehirlenmesi, saldırganların ağ içerisindeki IP ve MAC adresleri eşleştirmelerine müdahale ederek ağ cihazı ile bilgisayarın arasına girmesi olarak tanımlanır. Ağ içerisindeki bilgisayarın haberleşebilmesi için öncelikli olarak ARP protokolünü kullanmaktadır. Saldırgan ağ içerisinde hedefin ve ağ cihazının ARP Tablolarını zehirleyebilirse, yani kendi MAC adresini hedef bilgisayarın tablosuna Ağ Cihazı MAC adresi olarak ağ cihazı ARP Tablosuna ise Hedef Bilgisayarın MAC adresi yazılırsa saldırgan araya girmiş olur ve saldırgan bu trafiği dinleyebilir ve paketlerle ilgili değişiklik yapabilir.

Ortadaki Adam Saldırısından korunma yöntemleri:

- Web sitelerini kullanırken HTTPS protokolü kullanılmaya dikkat edilmelidir.
- Wi-fi şifreleri güvenli şifre kurallarına uygun bir şekilde belirlenmelidir.
- Herkese açık ağlara bağlanmamaya dikkat edilmesi gerekmektedir.

3.4.10 Cryptojacking saldırısı

Kötü amaçlı şifreleme olarak da tanımlanan bu saldırı türü, bilgisayar ve mobil cihazlarda gizlenen kripto para birimleri olarak da tanımlanan çevrimiçi para türlerini elde etmek için ortaya çıkan çevrimiçi bir tehdittir. Web tarayıcıları ele geçirebilen, bilgisayarları, akıllı telefonları ve ağ sunucularını tehlikeye atabilecek bir tehdittir.

3.4.11 Birthday saldırısı

Kaba kuvvet saldırıları sınıfına ait bir tür kriptografik saldırıdır. Olasılık teorisindeki doğum günü probleminin arkasındaki matematiğe dayanır. Bu saldırının başarısı, doğum günü paradoks probleminde açıklandığı gibi, rastgele saldırı girişimleri ile sabit bir permütasyon arasında bulunan çarpışma olasılığının yüksek olmasına bağlıdır.

3.5 Tarihteki Siber Olaylar

Siber ortamın oluşmasıyla beraber, bu ortam üzerinde birbirine üstünlük sağlamak isteyen ve sonuçları çok ağır olan bir çok siber olay yaşanmıştır. Bu bağlamda siber savaş; kara, deniz, hava ve uzaydan sonra 5. Savaş alanı olarak belirlenmiştir (Kara M. , 2013).

Bu bölümde tarihte yaşanmış bazı önemli siber saldırılar ele alınmıştır.

3.5.1 Stuxnet

İran'ın Nükleer Tesisleri'ni hedef alan Stuxnet virüsü sistemlere büyük zarar vermiştir. Bu virüsün amacı, uranyum zenginleştirmede kullanılan ve son derece hızlı dönen makineler olan santrefüjileri kontrol etme amacıyla kullanılan Programlanabilir Mantık Denetleyicisi' nin kontrol devrelerini hedef alarak kontrolü ele geçirmektir. Yazılımı bozulan sistem, hızla dönen

makinelerin çok daha fazla dönmesine sebep olarak makinelerin parçalanmasına sebep olmuştur. Merkez bilgisayarlar tarafından herşeyin normal gösterilmesi virüsün uzun süre fark edilmemesine ve zararın boyutunun daha çok artmasına sebep olmuştur.

3.5.2 Flame

2012 yılında tespit edilen virüs, Stuxnet virüsünden daha karmaşık bir kod ile yazılmıştır. Flame virüsü özellikle Orta Doğu'yu hedef almış olup verilerin sızdırılması amacıyla kullanılmıştır (Kara M. , 2013). Uzun bir süre fark edilememiştir. Etkilemiş olduğu bilgisayarlardan uzunca bir süre veri sızdırmıştır.

Karpersky Laboratuvarlarında yapılan incelemeler neticesinde, Flame virüsünden en çok etkilenen ülkeler sırasıyla; İran, İsrail, Batı Şeria, Sudan, Suriye, Lübnan, Suudi Arabistan ve Mısır'dır (Kara M. , 2013).

3.5.3 Gauss

Orta Doğu'yu etkileyen diğer bir virüs Gauss 2012'de tespit edilmiştir. Çevrimiçi bankacılık hesaplarını izlemek üzere tasarlanmış siber tehdittir. Virüs, bulaştığı bilgisayardaki hassas verileri, tarayıcı parolalarını ve bankacılık bilgilerini çalmak için tasarlanmıştır.

Bu virüsün diğer virüslerden ayıran en önemli özelliği, bankaları hedef alıyor olması olup Lübnan, İsrail ve Filistin'de bir çok bilgisayarı ele geçirmiştir.

3.5.4 Tinba

Virüs 2012 yılında, Danimarka Güvenlik Laboratuvarı olan CSIS tarafından tespit edilmiştir. Bu virüs, verilerin çalınması amacıyla geliştirilmiş olup ayrıca ağları dinleyebilme özelliğine de sahiptir.

Trendmicro güvenlik firması tarafından yayınlanan rapora göre, bu yazılım Türkiye için özelleşmiş ve Türkiye üzerinde aktif olan bir yazılımdır. Bu virüsten Türkiye'de etkilenen 60.000' den fazla kullanıcı tespit edilmiştir (Kara M. , 2013).

3.5.5 Shamoon virüsü

Bu virüs enerji sektörünü hedef alarak bulaştığı sistemlere zarar vermiş ve sistemleri çalıştırılamaz duruma getirmişlerdir. Bu virüs, 2012 yılında Suudi Arabistan'ı hedef alarak 30.000 bilgisayarın hasar görmesine sebep olmuştur.

3.5.6 Wikileaks olayları

ABD kıdemli eri olarak Irak operasyonlarında görev alan Bradley Manning, 1966 ve 2010 tarihleri arasında ABD' nin Dış İşleri Bakanlığınca yapılan ve oldukça gizli olan yazışmaları, ordunun veri tabanından indirmiş olduğu belgeleri, 2010 yılında Wikileaks sitesine sızdırmıştır. ABD yönetimini oldukça zora sokan bu görüntüler, tüm dünyanın kolaylıkla erişebileceği bir alanda paylaşılmış ve bilgiler ifşa edilmiştir.

3.5.7 Sibirya doğalgaz patlaması

1982 senesinde gerçekleştirilen bu saldırı, tarihte siber teknolojinin kullanılmasıyla gerçekleştirilen ilk siber saldırıdır. Saldırı ABD tarafından Rusların kullanmış oldukları yazılıma zararlı yazılımı yüklemesi sonucunda

gerçekleştirilmiştir. Saldırı sonucunda Sibirya doğal gaz boru hattında yaşanan bu patlama nükleer olmayan en büyük patlamadır.

3.5.8 Ay ışığı labirenti

Mart 1998' de gerçekleştirilen saldırı ile Pentagon, ABD Enerji Bakanlığı, NASA ve üniversitelere ait olan bir çok gizli belge bu saldırı ile çalınmıştır. Amerikan İstihbaratı, bu saldırının arkasında Rusya' nın olduğunu belirtmiştir (Kara M. , 2013).

3.5.9 NATO Kosova krizi

NATO direktifi ile Sırp hedeflerin bombalanmasıyla başlayan saldırıya karşılık olarak NATO karargahlarına ve üye ülke askeri haberleşme sistemlerine siber saldırılar yapılmıştır. Bu saldırılarda DDOS ve zararlı yazılımlar kullanılmıştır. NATO'nun remi web sitesi kesintiye uğramış ve NATO'nun e-posta hesapları günlerce kapalı kalmıştır. Bu saldırıların ardında Sırplı, Çinli ve Rus hackerlerin olduğu belirtilmektedir (Kara M. , 2013).

3.5.10 Estonya siber savaş

2007 tarihinde siber saldırıya maruz kalan Estonya, devletin internet siteleri, gazetenin web sayfaları erişilemez duruma getirilmiştir. Ayrıca ülkenin en büyük bankası olan Hansabank etikisiz hale gelmiş, internet hizmet sağlayıcıları çalışamaz duruma getirilmiştir. Uzmanlar tarafından önlemler alınmaya çalışılsa da zombi bilgisayarlar önlemlere karşı adapte olarak önlemleri etkisiz hale getirmiştir. Saldırının arkasında Rusya' nın olduğu düşünülse de Rusya bu iddiayı kabul etmemiştir (Kara M. , 2013).

3.5.11 Gürcistan siber savaşı

2008' de gerçekleştirilen bu saldırıda Rus siber güvenlik güçleri, Gürcistan' ı hedef almıştır. Bu saldırı da Gürcü medyasını ve devletin web sitelerini hedef alarak dış dünya ile olan bağlantılarını kesmeyi hedeflemişlerdir. Rus güçleri Gürcistan' a trafiği destekleyen yönlendiricileri ele geçirmiş ve Gürcistan dışarıya bir e-posta bile gönderememiştir.

3.5.12 Mavi Marmara saldırısı

Gazze' ye yardım götüren gemilere saldıran İsrail' e tepki olarak değişik hackerler çok sayıda banka, şirket ve devlete ait web sitelerine saldırı yapmıştır. İsrail' in özür dilememesine tepki olarak farklı zamanlarda İsrail Web sitelerine saldırılar gerçekleştirilmiştir

3.6 Saldırı Tespit Sistemleri

Bilgisayar ağlarında gerçekleştirilen her türlü yetkisiz erişimin tespiti ve ağ içerisinde meydana gelen anormal davranışların tespit edilmesi, akan trafiğin izlenmesi ve analiz edilmesi ile mümkündür. Bu durumların tespiti için bilgisayar sistemlerinde STS kullanılmaktadır. Saldırı tespit sistemleri bilgisayar sisteminde veya ağda meydana gelen olayları izleme ve bunları güvenlik problemleri açısından analiz etmek için kullanılan sistemlerdir.

İlk olarak STS konsepti Anderson tarafından 1980 yılında önerilmiştir (Anderson, 1980). STS' nin görevi, bilgisayar ve ağ sistemlerinde gerçekleşen tüm olayları gözlemlemek, denetlemek, kontrol etmek, güvenlik sorunlarının ortaya çıktığı durumlarda alarm göndererek ilgili personelleri ve birimleri uyarmak ve olası riskleri ortadan kaldırmak için gerekli tedbirlerin alınmasını sağlamaktır. Bununla birlikte STS' nin çalışma döngüsünde, bilgilerin toplanması, analizlerin yapılması ve buna karşılık bir cevabın oluşturulması olarak 3 aşamada gerçekleştirilir (Bace ve Mell, 2001).

Bilgisayar sistemlerindeki saldırıların tespitinin gerekliliği şu şekilde sıralanabilir (Denning, 1987):

- Mevcut olan sistemlerin birçoğunda saldırıya imkân verebilecek durumlar mevcuttur. Bu durumların bulunması ve düzeltilmesi teknik veya ekonomik nedenlerden dolayı mümkün değildir.
- Zaafları bilinen mevcut sistemler, daha yüksek güvenlik sağlayan alternatifleri ile değiştirilememektedir. Bunun ana sebebi ekonomik sebepler veya mevcut sistemlerdeki bazı özelliklerin daha yüksek güvenlik sağlayan özelliklerinin alternatiflerinde olmamasıdır.
- Tam olarak güvenli bir sistemin geliştirilmesi, imkânsız olmasa bile güçtür. Bu yüzden mevcut sistemlerin korunması daha doğrudur.
- En yüksek güvenlik ile korunan sistemler bile yetkilerini kötüye kullanan kullanıcılar tarafından zarar verebilir durumdadırlar

STS' nin kullanımının faydalarını üç ana başlıkta inceleyebiliriz (Dayıoğlu ve Özgit, 2001):

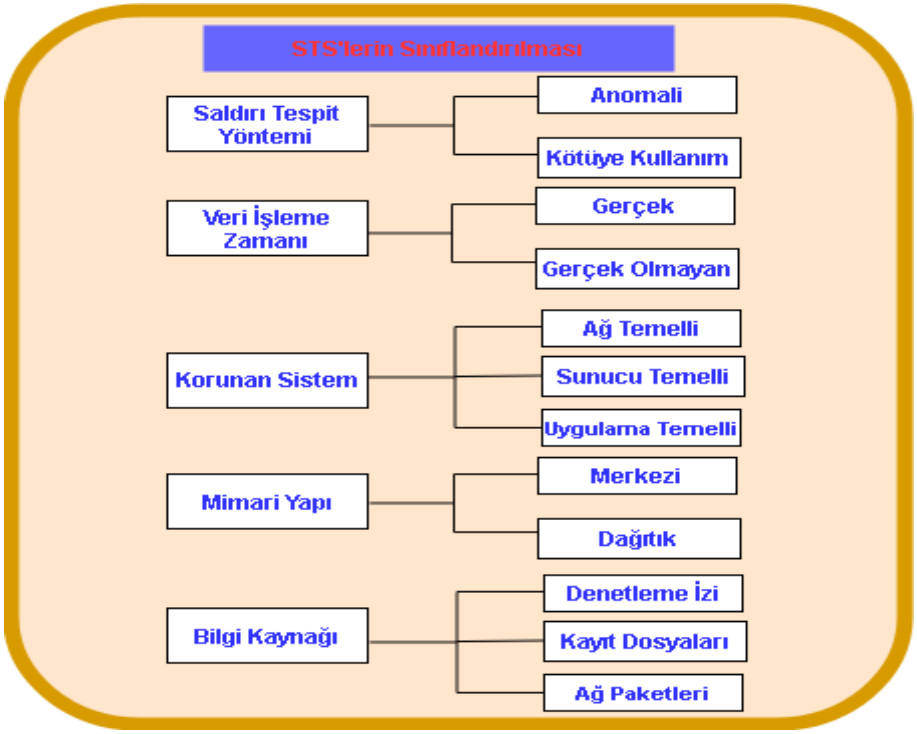
Erken Tespitin Yapılması: STS' ler, yaşanan herhangi bir ihlali sistem yöneticilerinden çok daha önce tespit edebilir. Bu sayede, sistem yöneticisi öncesinden bilgilendirilerek olaya müdahale etmesi sağlanır.

Detaylı Bilginin Toplanması: Sistem yöneticileri geçmişte veya sürmekte olan saldırılarla ilgili detaylı bilgi edinebilir. Bu bilgiler, saldırının kaynağı, büyüklüğü ve hedefler üzerindeki etkilerinin incelenmesi noktasında son derece önemlidir.

Toplanan Bilgilerin Kayıt Niteliği Taşıması: Sistem tarafından toplanan bilgiler, herhangi bir durumda kanıt olarak kullanılabilmektedir.

3.7 STS'lerin Sınıflandırılması

STS'lerin sınıflandırılmasında birçok kriter kullanılmaktadır. STS'ler genel olarak saldırıyı tespit etme yöntemi, verinin işlenme zamanı, korunan sistemin türü, bilgi kaynağı ve mimari yapısı kriterlerine göre sınıflandırılabilir. Şekil 3-4'te belirtilen kriterlere göre sınıflandırılmıştır (Güven, 2007).



Şekil 3-4 STS'lerin sınıflandırılması

3.8 Saldırı Tespit Yöntemine Göre

3.8.1 Kötüye kullanım

Kötüye kullanımı kontrol eden dedektörler ağdaki hareketleri izler, bu olayları daha öncesinden belirlenmiş olan saldırılar ile karşılaştırılır. Eşleşme olursa saldırı tespit edilir. Daha öncesinde tespit edilmiş saldırıları modelleyen şablona imza adı verildiği için bu sistemlere imza temelli tespit adı verilir. Çizelge 2-3' te

Kötüye Kullanım Tabanlı STS'lerin avantajları ve dezavantajları verilmiştir (Aydın, 2005).

Çizelge 3-3 Kötüye kullanım tabanlı STS' nin avantajları ve dezavantajları

Avantajlar	Dezavantajlar
• Kötüye Kullanım dedektörleri, saldırıların tespitinde çok sayıda yanlış alarm üretmezler. • Kötüye Kullanım dedektörleri özel bir saldırı aracı veya tekniğini hızlı bir şekilde tespit eder. • Kötüye Kullanım dedektörleri, sistem yöneticilerine, uzman düzeyinde bilgi sahibi olmasına gerek duymadan sistemlerini izleme olanağı sağlar.	• Dedektörler yalnızca bilinen saldırıları sezer ve yeni saldırılar gözlemlendiğinde bu saldırıların imzalarının eklenerek güncelleme yapılması gereklidir. • Dar kapsamlı olarak tanımlanmış imzalara göre tanımlanan dedektörler, sık rastlanan saldırıların ufak değişikliklerle gerçekleştirilmesi durumunda tespitte zorlanır.

3.8.2 Anormallik

Sistemdeki anormallik dedektörleri, bir sistem üzerindeki normal olmayan davranışların analizini ve tespitini yapar. Bu dedektörler, "normal" olarak tanımlanan davranışlardan farklı olan hareketleri saldırı olarak tanımlar. Bu sistemin çalışma mantığı, kullanıcıların normal çalışma sürelerindeki davranışlarının takip edilerek kullanıcıların normal davranış profillerinin çıkartılmasına dayanır. Bu profillerin belirlenmesi ile anormallik dedektörleri, sistem üzerindeki hareketleri izleyerek saldırıyı tespit eder. Anormallik tabanlı sistemlerde saldırıların doğru tespit edilebilmesi için normal davranış profillerinin doğru bir şekilde belirlenmiş olması gerekir.

Çizelge 4-4'te anormallik durum tespiti için kullanılan STS'lerin avantajları ve dezavantajları verilmiştir (Aydın, 2005).

Çizelge 3-4 ATSTS' nin avantajları ve dezavantajları

Avantajlar	Dezavantajlar
• ATSTS'ler sıra dışı davranışları tespit edebildikleri için saldırı ile ilgili çok fazla bilgi sahibi olmasına gerek yoktur.. • Kötüye Kullanım dedektörleri için tanımlanan imzaları tanımlayan bilgi üretebilirler.	• Kullanıcı ve ağ davranışları hakkında çok fazla bilgi sahibi olmadıkları için çok fazla sayıda hatalı alarm üretebilirler. • Bu yaklaşımda normal davranış profilini karakterize edebilmek için çok büyük eğitim verisine ihtiyaç vardır.

3.8.3 Veri işleme zamanına göre

Veri işleme zamanına göre Saldırı Tespit Sistemleri gerçek zamanlı ve gerçek zamanlı olmayan olarak iki ana başlıkta incelenir.

3.8.3.1 Gerçek zamanlı STS'ler

Bilgi kaynaklarından sürekli olarak bilgi akışı gerçekleşir. Gelen paket sistem tarafından o anda analiz edilir ve bir saldırı tespit edilirse bununla ilgili karar ve işlemler o anda gerçekleştirilir. Ticari amaçlı kullanılan sistemler genellikle gerçek zamanlı olarak tasarlanmıştır. Bir saldırının çok fazla kayba sebep olacağı ticari sistemler için, gerçek zamanlı sistemlerin kullanılması en doğru çözümdür.

3.8.3.2 Gerçek zamanlı olmayan STS'ler

Bilgi kaynağından aralıklar ile bilginin akışı gerçekleşir. Saldırı tespitinin gerçek zamanlı olmasına gerek duyulmayan yerlerde kullanılır. Ayrıca bazı STS'ler hem gerçek zamanlı hem de gerçek zamanlı olmayan biçimde çalışabilirler. Bu sistemlerin kullanım alanlarından bir tanesi de saldırı ve normal durum profillerinin geçmiş kayıtlara bakılarak çıkarıldığı durumlardır.

3.8.4 Korunan sisteme göre

STS'ler korudukları sistemlere göre, ağ, sunucu ve uygulama temelli üç gruba ayrılırlar.

3.8.4.1 Ağ temelli

Ağ Temelli STS'ler mevcut sistem üzerindeki ağ paketlerini yakalar ve bunların analizini yapar. Bu STS'ler çalıştığı ağ üzerindeki bütün haberleşmeleri dinleyebildiği için ağdaki bütün bilgisayarları koruyabilmektedir. Ticari amaçlı kullanılan STS'lerin bir çoğu ağ-temelli STS'dir.

Ağ Temelli STS' lerin avantajları ve dezavantajları Çizelge 3-5' te verilmiştir (Aydın, 2005)

Çizelge 3-5 Ağ tabanlı STS'lerin avantajları – dezavantajları

Avantajları	Dezavantajları
• Geniş bir ağı denetleyebilirler. • Genel olarak pasif aygıtlardır. Bir ağdaki normal işleyişi etkilemezler. • Oldukça gizlidirler ve çoğu saldırgan tarafından tespit edilemezler.	• Çok büyük bir ağdaki tüm paketleri işlemek zordur. Yoğun trafikte gerçekleşen bir saldırıyı tespit edemeyebilir. • Anahtarlar, ağları küçük bölümlere ayırdığı için çoğu anahtarlar evrensel port denetimini sağlayamaz. • Şifrelenmiş bilgiyi analiz edemez. • Yalnızca başlatılan saldırıyı fark edebilirler. • Bazı STS'ler parçalanmış ağ paketlerini içeren ağ tabanlı saldırıları tespit edemez.

3.8.4.2 Sunucu temelli

Bu STS'ler, işletim sisteminin hesap izlerini ve sistemdeki hareketlerin kaydını alarak depolar ve bu veriler üzerinde çalışırlar. Sunucu Temelli STS'lerin avantajları ve dezavantajları Çizelge 3-6' da verilmiştir (Aydın, 2005).

Çizelge 3-6 Sunucu temelli STS'lerin avantajları ve dezavantajları

Avantajları	Dezavantajları
• Ağ temelli STS'lerin yakalamadığı saldırıları tespit edebilirler. • Ağ trafiğinin şifrelendiği bir ortamda çalışabilirler. • Anahtarlamalı ağlar üzerinde iyi çalışırlar • Truva atları ve yazılım eksikliklerinden kaynaklı saldırıları tespit edebilirler.	• Yönetimi oldukça zordur • Üzerinde çalıştığı sunucunun saldırıya maruz kalması durumunda devre dışı kalabilir. • İşletim sistemi izleme modunda bilginin miktarı çok büyük olabilir. • Ağın taranmasından haberdar olamazlar.

3.8.5 Uygulama temelli

Uygulama Temelli STS'ler bilgi kaynağı olarak uygulamaya ait günlük dosyaları kullanırlar. Sunucu Temelli STS'lerin özel bir türüdür. Bu STS' lerin avantajları ve dezavantajları Çizelge 3-7'de verilmiştir (Aydın, 2005).

Çizelge 3-7 Uygulama temelli STS'lerin avantajları ve dezavantajları

Avantajları	Dezavantajları
• Kullanıcının ve uygulamanın birbirlerini etkilemeleri incelenerek bireysel kullanıcıların yetkilerini aşan davranışları tespit edilebilir. • Şifrelenmiş veri işlenebilir.	• Saldırılara daha fazla açıktırlar. • Truva atları ve diğer yazılımsal saldırıları sezemezler.

3.8.6 Mimari yapı

STS' lerin mimari yapısı, fonksiyonel bileşenlerinin birbirlerine göre nasıl yerleştirildiklerini anlatmaktadır. Temel fonksiyonel bileşenler, analizin

yapıldığı sunucu ile çevresi ve izlenen sistem ve problemler için izlenen hedef olarak belirtilebilir (Marttin, 2014).

3.8.7 Merkezi sistem

İzleme, tespit etme ve raporlama işlemlerinin tümünün merkezde kontrol edildiği sistemlerdir. Bu sistemler için fiziksel yakınlık önemli bir etken değildir.

3.8.8 Dağıtık sistem

İzleme, tespit etme ve rapor etme işlemlerinin merkezi bir sistemde olmayıp kurulan ajan tabanlı programlar ile kontrol edilirler. Çok geniş ağ yapısına sahip ağlarda genellikle bu sistem kullanılmaktadır.

3.8.9 Bilgi kaynağı

STS' lerde üç farklı bilgi kaynağı kullanılır.

3.8.9.1 Denetim izi

Denetleme İzi, Verilerin yönetim ve iç kontrol sistemlerine girişlerinden çıkışlarına kadar izlenmelerine imkân veren yoldur. Sistemde gözlemlenen olayların sıralaması değiştiği durumlarda, yeniden yapılandırmayı ve test etmeyi mümkün kılar. STS'ler denetim izlerini kullanıcı profilini çıkartmak için kullanırlar.

3.8.9.2 Ağ paketleri

Koklayıcı programlar sayesinde trafiklerin dinlenmesi ile edilir. Bu veriler ile ağ katmanında gerçekleşen saldırıların tespit edilmesi sağlanır.

3.8.9.3 Kayıt dosyaları

Uygulama katmanında gerçekleşen saldırıları tespit etmek için kullanılan veri kaynağıdır. Bu veriler sınırlı sayıda saldırıyı tespit etmek için kullanılabilir.

3.9 STS Uygulamaları

Bilgi güvenliğinin önemiyle beraber saldırı tespit sistemleri ile ilgili birçok çalışma yapılmıştır. Bu konu ile ilgili olarak araştırmacılar farklı yöntem ve teknikler kullanarak birçok saldırı tespit sistemi uygulaması geliştirmişlerdir.

Günümüzde Snort, Snortsam, Bro, Ripper, Haystack, IDES, MIDAS yaygın olarak kullanılan saldırı tespit sistemleridir. Bu konuda yapılan çalışmalar şu özellikler kullanılarak kategorize edilmiştir:

- STS' nin geliştirilmesinde kullanılan yöntem
- Saldırı ile ilgili olarak veriyi işleme zamanı
- Veri kaynağını koruyan sistem
- Saldırı Tespit Sisteminin Mimari yapısı

Çizelge 3-8'de bazı STS'ler ve karakteristiklerine göre sınıflandırılmıştır.

Çizelge 3-8 Kullanılan bazı STS uygulamaları ve özellikleri

Yıl	STS Uygulaması	STS Geliştirilme Yöntemi	Verinin İşlenme Zamanı	Korunan Sistem	Mimari Yapı
1988	Haystack	Hibrid	Gerçek olmayan	Sunucu	Merkezi
1988	MIDAS	Hibrid	Gerçek	Sunucu	Merkezi
1988	IDES	Anormallik	Gerçek	Sunucu	Merkezi
1989	W&S	Anormallik	Gerçek	Sunucu	Merkezi
1990	Comp-Watch	Anormallik	Gerçek Olmayan	Ağ	Merkezi
1990	NSM	Hibrid	Gerçek	Sunucu	Merkezi
1991	NADIR	İmza	Gerçek olmayan	Sunucu	Merkezi

1991	Computer Misuse Detection System	Hibrid	Gerçek	Sunucu	Dağıtık
1992	DIDS	Hibrid	Gerçek	Hibrid	Dağıtık
1992	ASAX	İmza	Gerçek	Sunucu	Merkezi
1992	Intruder Alert	İmza	Gerçek	Hibrid	Dağıtık
1993	USTAT	İmza	Gerçek	Sunucu	Merkezi
1994	DPEM	İmza	Gerçek	Sunucu	Merkezi
1994	IDIOT	İmza	Gerçek	Sunucu	Merkezi
1995	NIDES	Hibrid	Gerçek	Sunucu	Merkezi
1996	GrIDS	Hibrid	Gerçek Olmayan	Hibrid	Dağıtık
1996	CSM	İmza	Gerçek	Sunucu	Dağıtık
1996	JANUS	İmza	Gerçek	Sunucu	Merkezi
1996	Kane Security Monitor	İmza	Gerçek	Sunucu	Dağıtık
1996	Netranger	İmza	Gerçek	Ağ	Dağıtık
1996	RealSecure	İmza	Gerçek	Ağ	Dağıtık
1997	JiNao	Hibrid	Gerçek	Sunucu	Dağıtık
1997	EMERALD	Hibrid	Gerçek	Hibrid	Dağıtık
1997	Anzen Fligh Jacket	Hibrid	Gerçek	Ağ	Hibrid
1997	Netprowler	İmza	Gerçek	Sunucu	Dağıtık
1997	Securenet Pro	İmza	Gerçek	Ağ	Dağıtık
1997	Sessionwall-3	İmza	Gerçek	Ağ	Dağıtık
1998	Bro	İmza	Gerçek	Ağ	Merkezi
1998	Cenrax Security Suite	Hibrid	Gerçek	Hibrid	Dağıtık
1998	Cross-site for Security	İmza	Gerçek	Ağ	Dağıtık
1998	Smartwatch	İmza	Gerçek	Sunucu	Dağıtık
1998	Stake out	İmza	Gerçek	Ağ	Dağıtık
1998	Tripware	İmza	Gerçek Olmayan	Sunucu	Merkezi
1998	Snort	Hibrid	Gerçek	Sunucu	Merkezi
1999	Cybercop monitor	İmza	Gerçek	Sunucu	Dağıtık
2000	FIRE	Anormallik	Gerçek Olmayan	Ağ	Merkezi

3.9.1 Snort

Snort, 1998 yılında Martin Roesch tarafından geliştirilen ve açık kaynak kodlu bir ağ saldırı tespit ve önleme sistemidir. Snort çok iyi şekilde belgelenmiş ve kolayca indirilebilecek iyi bir kural arşivine sahiptir Protokol ve içerik analizleri yaparak birçok saldırıyı tespit edebilmektedir. Snort'un genel özellikleri şu şekildedir (snort-org):

- Açık kaynak kodlu ağ saldırı tespit ve önleme yapabilir.
- IP ağlarında gerçek zamanlı trafik analizi ve paket kaydı yapabilir.
- Protokol analizi, içerik arama ve eşleştirme yapabilir ve çeşitli saldırı ve probları tespit edebilir.
- Bir paket dinleyicisi olarak kullanılabilir.
- Paket kaydedici olarak kullanılabilir
- Gelişmiş bir ağ saldırı önleme sistemi olarak kullanılabilir.
- Snort üç modda çalıştırılabilmektedir (Ilgaz, 2018):
- Paketleri İzleme modu: Ağdan paketleri okuyarak sürekli olarak ekrana aktarır. Komut satırında "./snort-v" şeklinde çalıştırılabilmektedir.
- Paketleri Günlükleme modu: Paketler diske yazılmaktadır. Komut satırında "snort –dev –l ./log" komutu ile çalışmaktadır.
- Ağ Saldırı Tespit ve Engelleme Sistemi modu: Snort'un en karmaşık ve yapılandırılabilir modudur. Snort bu modda temel olarak trafiği analiz edip kullanıcı tarafından tanımlanabilen bir kural seti ile gördüklerine karşı eylemlerde bulunur.

3.9.2 Haystack

1988'de Amerikan Hava Kuvvetlerinde kullanılan OS/100 işletim sistemini çalıştıran çok kullanıcılı olan Unisys 1100/60 ana bilgisayarları için tasarlanmış olan saldırı tespit sistemidir (Smaha, 1988).

Haystack büyük boyutlu sistem denetim izlerini, kullanıcı davranışları, anormal olaylar ve güvenlik olaylarının özetine indirgemektedir. Bu durum, güvenlikten sorumlu kişinin izinsiz girişleri daha kolay tespit etmesine yardımcı olmak için tasarlanmıştır. Haystack'ın çalışması, güvenlik politikalarından doğan davranış kısıtlamalarına ve kullanıcı grupları ve bireysel kullanıcılar için tipik davranış modellerine dayanmaktadır (Smaha, 1988)

Haystack altı farklı saldırı grubunun tespiti için tasarlanmıştır (Güven, 2007):

- Kötü niyetli kullanım
- Hizmet aksattırma
- Sızma
- Kırma girişimleri
- Kılık değiştiren saldırılar
- Güvenlik kontrol sistemine sızma

3.9.3 IDES

IDES bağımsız olarak kendi donanımında çalışır ve bir ağ üzerinde bir veya daha fazla hedef sistemden alınan denetim verilerini işler. Bir sisteme izinsiz sızmaya çalışan veya ayrıcalıklarını kötüye kullanmaya çalışanlar tarafından kaynaklı güvenlik ihlallerinin gerçek zamanlı tespiti için sistemden bağımsız bir mekanizma sağlamayı amaçlamaktadır (Teresa F. Lunt, 1992).

IDES, şüpheli davranışları tespit etmek için çeşitli yaklaşımlar kullanmaktadır. IDES bireyler için belirlenmiş olan kullanıcı profillerinden uzaklaşmaları gözlemleyerek saldırıları tespit eder. Bunu, geçmiş kullanıcı davranışlarının istatistiksel profillerini tutarak yapar. Ayrıca IDES belirli saldırı türlerini karakterize eden uzman sistemlerin kurallarını da içerir. IDES, gözlemlenen etkinlik, daha önceden belirlenmiş izinsiz giriş senaryolarından herhangi biriyle eşleşirse alarm verir.

IDES'in ilk sürümü tek bir sunucu için tasarlanmış ve daha sonra yapılan çalışmalarda tekrar dizayn edilerek farklı birçok hedef sistem için kullanılabilir hale getirilmiştir. Pek çok sürümü olan yazılım son olarak NIDES (Next-Generation Intrusion Detection Expert Sistem) adını almıştır (Anderson, 1980).

3.9.4 MIDAS

Multics Saldırı tespit ve önleme sistemi, Amerikan Ulusal Bilgisayar Güvenlik Merkezi tarafından yüksek güvenlikli işletim sistemi multics üzerinde çalışan Honeywell DPS 8/70' izlemek için tasarlanmıştır. MIDAS liman yöneticisinin yanıtlama sistemi denetim günlüğünden veri almak üzere tasarlanmıştır. Bu veriler düzenlenmiş olan oturum profilleri oluşturmak için kullanılmıştır ve daha sonra normal davranıştaki kullanıcı profilleri ile karşılaştırılmıştır (Bace, 2000).

IDES ve dönemin diğer sistemleri gibi MIDAS için, istatistiksel anormal durum tespit ve kural tabanlı uzman sistem yaklaşımları birleştiren hibrit bir analiz stratejisi geliştirilmiştir

3.9.5 NADIR

Ağ denetim direktörü ve izinsiz giriş raporlayıcısı, Los Alamos ulusal laboratuvarının bilgi işlem bölümü tarafından Los Alamos'taki tümleşik bilgisayar ağlarındaki kullanıcı davranışlarını izlemek için tasarlanmıştır. Bu ağ Los Alamos'un ana bilgisayar ağıdır ve 9000 kullanıcıya bağlı süper bilgisayarları, yerel ve uzak terminalleri, iş istasyonlarını, ağ hizmetleri makinelerini ve veri iletişim ara yüzlerini sunar.

NADIR, özel ağ servis düğümleri tarafından oluşturulan denetim yollarını işleyerek ağı izler. Sun Unix üzerinde çalışacak şekilde tasarlandı ve zamanın diğer birçok sistemi gibi uzman kural tabanlı analiz ve istatistiksel profillerin bir kombinasyonunu gerçekleştirilir. NADIR SQL'de yazılmıştır ve sybase'in dahili

tetikleyicileri ve diğer özellikleri kullanılarak bir Sybase veritabanı yönetim sisteminde çalışır (Bace R. G., 2000).

3.9.6 NSM

Ağ sistemi monitörü, Davis' teki Kaliforniya Üniversitesi'nde Sun Unix üzerinde çalışacak şekilde geliştirilmiştir. Ağ trafiğini izleme ve bu trafiği birincil veri kaynağı olarak kullanma konusundaki ilk atılımı yaptılar. NSM' nin genel mimarisi, birçok ticari saldırı tespit ürününe hala yansımaktadır.

NSM'in işlevleri genel olarak şöyledir (Bace, 2000).

- Sistemin Ethernet ağ arabirim kartını karışık moda geçirme
- Yakalanan ağ paketleri
- İlgili özelliklerin çıkarılmasına izin vermek için protokolü ayrıştırma
- Hem normal davranıştan kaynaklanan istatistiksel varyanslar hem de önceden belirlenmiş kuralların ihlali için özellikleri arşivlemek ve analiz etmek için matris tabanlı bir yaklaşım kullanma

3.10 Veri Setleri

STS'lerinin geliştirilmesi ve araştırma çalışmalarının yapılması için birçok veri seti üretilmiştir. Bu veri setlerinden en çok kullanılanı DARPA ve KDD99 veri setleridir. Tavallaee ve arkadaşlarının makaleler üzerinde yapmış olduğu tarama çalışmasında, 276 çalışmanın 209 tanesinde KDD99 ve DARPA veri setleri kullanılmıştır. Yapılan araştırmanın sonucu Çizelge 3-9'da gösterilmiştir (Özgür ve Erdem, 2012).

Çizelge 3-9 Veri setlerinin kullanım istatistiği

Veri Setinin Adı	Kullanım Sayısı	Kullanım Yüzdesi (%)
DARPA	67	24
KDD99	77	28

KDD99-DARPA+ Ek Saldırı	41	15
Veri Seti Belirtilmemiş	86	31
Diğer	16	6

3.11 DARPA

DARPA veri setleri DARPA 1998, 1999 ve 2000 veri setleridir. Bu veri setileri DARPA' nın sponsorluğu ile MIT Lincoln Sponsorluğunda STS'ler için bir karşılaştırma ortamı sunmak üzere oluşturulmuştur (Tanrıkulu ve Sazlı, 2007). Siber Sistemler ve Teknoloji Grubu tarafından bu veriler paylaşılmaktadır.

Bu veri kümesi e-posta, tarama, FTP, Telnet, IRC ve SNMP olaylarını içermektedir. Bu veri kümesi, gerçek dünyadaki ağ trafiği tam olarak yansıtmaz Çünkü yanlış pozitiflerin olmaması gibi eksiklikleri barındırır.

3.12 KDD Cup'99

KDD Cup '99 veri seti, Üçüncü Bilgi Keşfi ve Veri Madenciliği Araçlarının Yarışması'nda kullanılmak için sanal ortamdan transfer edilen veriler topluluğudur. Yapılan araştırmalara göre yapılan çalışmalarda ve deneylerde en çok kullanılan veri setlerinden biridir. Bu veri kümesi DARPA veri setinin güncellenmiş sürümüdür. Neptune-DOS, pod-DOS, Smurf-DoS ve arabellek taşması gibi farklı saldırı türleri ile ilgileri verileri tutar.

Bu veri setinde normal durum ve saldırı durumu simüle edilerek veri kümesinde bu veriler birleştirilmiştir. Yani veri seti gerçek olmayıp gerçek bir ağın simüle edilmesi sonucunda elde edilmiştir.

3.13 NSL-KDD

KDD Cup'99 veri kümesi, diğer kayıtların sınıflandırılmasını engelleyen çok sayıda gereksiz kayıt (%78) ve yinelenen kayıtlar (%75) içermektedir (Revathi ve Malathi, 2013). NSL-KDD veri kümesi, KDD Cup '99 veri kümesinin eğitim setindeki gereksiz kayıtlar veya test setinde kopyalar içermeyen daha az sayıdaki seçilmiş özellikten oluşur.

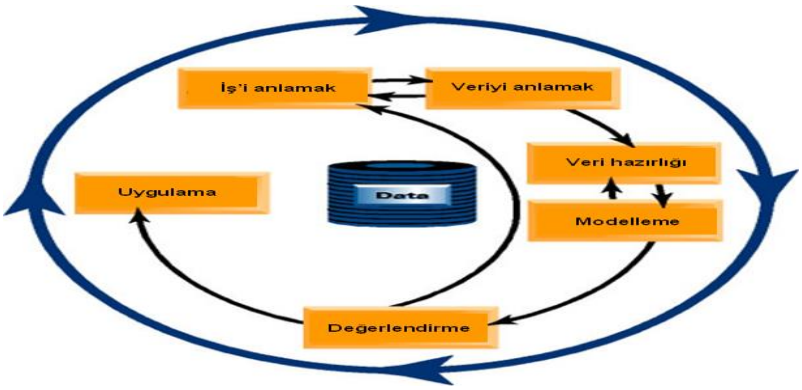
Eğitim veri setinde normal trafiği yansıtan 67.343, tehdit trafiğini yansıtan 59.277 ve toplam da 126.620 örnek bulunur. Test veri setinde ise normal trafiği yansıtan 9.711, tehdit durumunu yansıtan 13.139 ve toplam da 22.850 örnek vardır (Protić, 2018).

3.14 Kyoto 2006+

Kyoto 2006+ veri seti, 2006 ve 2009 yılları arasında gerçek trafik üzerinden elde edilen verilerle oluşturulmuştur. KDD Cup '99 veri kümesinden türetilen 14 istatistik özelliğin yanı sıra STS ağının analiz edilmesi ve değerlendirmenin yapılması için kullanılabilecek 10 ek özellikten oluşur.

4. MAKİNE ÖĞRENMESİ

Makine Öğrenmesi Algoritmaları (MÖA), karmaşık veri kümelerini keşfetme, analiz etme ve bunlarda anlam bulmamızı sağlayanlar yazılımlardır. Her algoritmanın hedefe ulaşmak için izlediği yöntemler farklıdır. MÖA' lar verilerin görselleştirilmesi, görüntü işleme, nesneleri tanıma, sanal gerçeklik, arttırılmış gerçeklik, ses tanıma, robotların hareketleri, pazarlama, tavsiye algoritmaları, sağlık, yer bilimi, sahtekarlık tespiti, veri güvenliği, davranış analizi, arama motorları, endüstri 4.0, Nesnelerin İnterneti gibi konularda yoğun bir şekilde kullanılmaktadır. Kullanım alanlarının geniş olması ile birlikte, bir problemin çözümlenmesi için izlenen yöntem birbirlerine çok benzerdir. Şekil 4-1' de CRISP-DM (Shearer, 2000) yöntemi ile süreç yönetimi anlatılmıştır.



Şekil 4-1 CRİSP-DM iş akışı

4.1 CRISP-DM Modeli İş Akışı Sürecinin Aşamaları

Bu iş modeli için süreç 6 aşamadan oluşmaktadır.

4.1.1. İş sürecinin anlaşılması

Bu süreçte problem tanımı ve problemin çözülebilmesi için gerekli veri kaynaklarının araştırması yapılarak iş sürecinin bitiminde nasıl bir sonuç ile karşılaşılması gerektiği belirlenir. Bu çalışma da problem olarak ağa yapılan

siber saldırıları ele alırsak problemin çözümü için gerekli veri setini elde etmek ve sonuç olarak saldırıları yüksek performans ile tespit edebilen bir sistemin ortaya çıkması beklenmektedir.

4.1.2. Verinin anlaşılması

Problem tanımı ve beklentinin belirlenmesi ile birlikte bu süreçte verilerin elde edilmesi ve verilerin neyi ifade ettiğinin anlaşılması gerekmektedir. Verinin iyi anlaşılması için problemin doğru tespit edilmiş olması önemlidir. Bu çalışma da güncel olduğu ve diğer veri setlerindeki eksiklikler dikkate alınarak üretildiği için CSE-CIC-IDS2018 veri seti üzerinde çalışılmıştır.

4.1.3. Veri ön işleme

Verilerin makine öğrenmesi yöntemleri ile doğru bir şekilde işlem görebilmesi için bazı ön işlemlerden geçmesi gerekebilir. Eksik veriler, metin verileri veya kategorik veriler kullanılacak yönteme için uygun hale getirilmelidir. Bu çalışma da eksik örneklemler veri setlerinden çıkartılmış ve dizgi olan zaman ifadesi numerik değere dönüştürülmüştür.

4.1.4. Uygun modelin seçilmesi

İş sürecinin bu aşamasında veri setine ve probleme uygun makine öğrenmesi algoritması seçilir. Bu çalışma da Yapay Sinir Ağları, Lojistik Regresyon, Destek Vektör Makineleri, K-En Yakın Komşu, Naive Bayes, Rastgele Orman ve Karar Ağaçları sınıflandırma algoritmaları kullanılmıştır.

4.1.5. Modelin değerlendirilmesi

Seçilen modellerin problemin çözümünde kullanılarak performanslarının değerlendirilmesi yapılmaktadır. Bu çalışmada kullanılan makine öğrenmesi

algoritmaları ile oluşturulan modellerin test sonucunda elde edilen tahmin yüzdesi ve test süreleri değerlendirilmiştir.

4.1.6. Ürün elde edilmesi

Elde edilen sonuçlar değerlendirilerek problemin çözülebilmesi için çalışma ortamına uygun halde geliştirme sürecini kapsar.

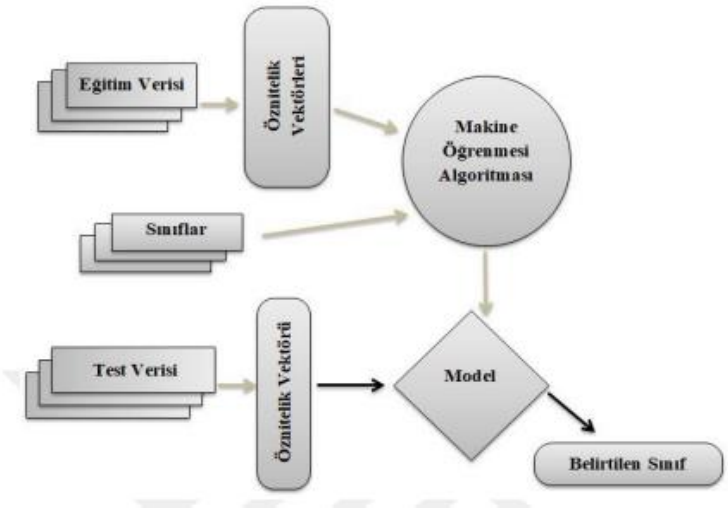
4.1.7. Öğrenme Yöntemine Göre Makine Öğrenmesi Algoritmaları

Makine öğrenmesi teknikleri öğrenme yöntemine göre üç gruba ayrılırlar:

- Denetimli Öğrenme (Supervised Learning)
- Denetimsiz Öğrenme (Unsupervised Learning)
- Pekiştirmeye Dayalı Öğrenme (Reinforcemend Learning)

4.2.1. Denetimli öğrenme

Denetimli öğrenmede öğrenme, modelin etiketlenmiş örneklemlerin kullanılarak eğitilmesi sonucu sağlanmış olur. Model eğitilirken veri setindeki örneklemlere ait giriş çıkışlar verilir. Girişler modelin değişkenlerini, çıkışlar ise modelin kategorisini tanımlamaktadır. Denetimli Öğrenme Süreci modelinin süreci Şekil 4-2'de gösterilmiştir (Afrin ve Nahar, 2015). Naive Bayes, Karar Ağaçları, K-En Yakın Komşu, Rastgele Orman, Karar Ağaçları, Yapay Sinir Ağları, Lojistik Regresyon denetimli öğrenme modeli için yaygın olarak kullanılan Makine Öğrenmesi algoritmalarıdır (Caruana ve Niculescu-Mizil, 2006).



Şekil 4-2 Denetimli öğrenme

Denetimsiz öğrenmede ise örneklemler etkilenmez. Algoritma, verileri düzenleyerek veya bunların yapısını açıklayarak örneklemleri etiketlemektedir. Hiyeraşik ve parçalayıcı kümeleme algoritmaları genellikle denetimsiz öğrenme modelleri oluşturulurken kullanılırlar.

4.1.8. Pekiştirmeli öğrenme

Pekiştirmeli öğrenme modelleri, eylemlerinden sonra seçeneğin doğru mu, nötr mü, yoksa yanlış mı olduğunu belirlemeye yardımcı olan geri bildirim alır. Öğrenme işlemi için kurallara ve ödüllere ihtiyaç duymaktadır. İnsan kılavuzluğu gerektirmeyen birçok küçük kararlar alması gereken otomatikleştirilmiş sistemlerin geliştirilmesinde kullanılabilmektedirler.

4.1.9. Tahmin Yöntemine Göre Makine Öğrenmesi Algoritmaları

4.1.10. Sınıflandırma algoritmaları

Bir veri setindeki örneklemlerin farklı sınıf içerisinde dağıtılarak gelen bir örneğin hangi sınıfa ait olduğunun bulunmasını sağlayan algoritmalardır. Örneklemlerin girişleri verinin özelliğini, etiketleri ise verinin sınıfını

belirtmektedir. Sınıflandırma algoritmaları, verilen eğitim kümesinden dağılımı öğrenirler ve daha sonra test verilerinin giriş değerleri kullanılarak çıkış sınıfı tahmin edilmeye çalışılır.

4.1.11. Anomali algılama algoritmaları

Normal olarak tanımlanmış davranışlara uymayan verilerdeki kalıpları ve örüntüleri bulmak için kullanılan algoritmalardır. Dolandırıcılık tespiti, olağan dışı davranışların tespiti, şüpheli faaliyetlerin tespiti için kullanılmaktadırlar.

4.1.12. Regresyon algoritmaları

Bağımlı değişkenin diğer bağımsız değişkenler arasındaki ilişkiye göre sonuç üreten değerlerdir. Eğitim modelinde bağımsız değişkenler ve bağımlı değişkenler arasındaki ilişki belirlenir. Test modelinde ise bağımsız değişkenler verilerek bağımlı değişkenin bulunması sağlanır.

4.1.13. Kümeleme algoritmaları

Veri seti içerisindeki örneklemler arasındaki benzerlik düzeyini belirleyerek veriyi birden fazla gruba bölen algoritmalardır. Hiyerarşik bölütleme ve K-Means yaygın olarak kullanılan kümeleme algoritmalarıdır.

4.1.14. Sınıflandırma Algoritmaları

4.1.15. Lojistik regresyon

LR analizi adını, bağımlı değişkene uygulanan logit dönüştürmeden almaktadır. En az değişken kullanılarak en iyi uyuma sahip olacak şekilde bağımlı ve bağımsız değişkenler arasındaki ilişkiyi tanımlayabilen ve biyolojik olarak kabul edilebilir bir model kurulur (Bircan, 2004). Bu algoritma ile bağımlı değişken üzerindeki açıklayıcı değişkenlerin etkileri olasılık olarak

hesaplanarak bu faktörlerin olasılıksal olarak belirlenmesi sağlanır (Özdamar, 2004).

4.1.16. K-En yakın komşu

KNN algoritması, örnek veri noktasının bulunduğu sınıfın ve en yakın komşusunun k değerine göre belirlendiği sınıflandırma algoritmasıdır. Veri setine katılacak olan verinin mevcut verilere olan uzaklığına bakılır ve k sayıda yakın komşuluğuna bakılır. Uzaklığın hesaplanması için genelde 3 tip uzaklık fonksiyonu kullanılmaktadır:

- Euclidean uzaklık
- Manhattan uzaklık
- Minkowski uzaklığı

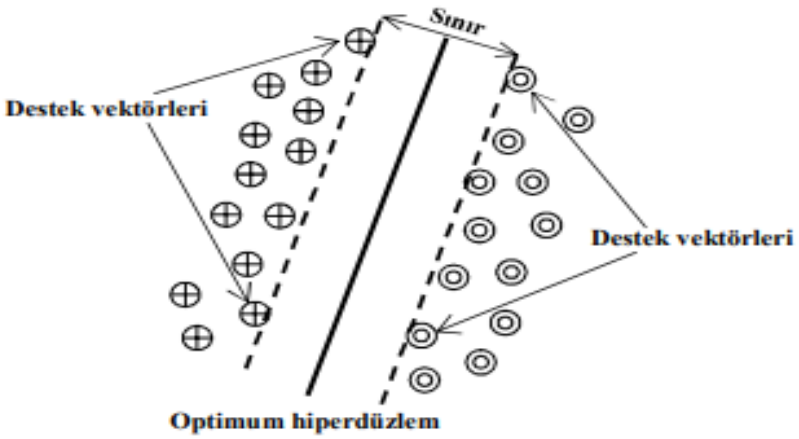
KNN gürültülü veri setlerine karşı dirençli olması ile birlikte, büyük verilerde uzaklık hesabını yaparken çok fazla bellek kullanımı yapmaktadır.

4.1.17. Destek vektör makineleri

DVM örüntü tanıma ve sınıflandırma problemlerinin çözümü için Vapnik tarafından geliştirilmiştir (Cortes ve Vapnik, 1995). DVM' nin temelleri istatistiksel öğrenme teorisine dayanmaktadır.

DVM' de amaç, sınıfları birbirinden ayıracak en iyi düzlemin elde edilmesidir. Yani sınıflara ait destek vektörleri arasındaki mesafenin maksimize edilmesi gerekmektedir.

Sınırı maksimuma çıkararak en uygun ayrımı yapabilen hiperdüzleme optimum hiper düzlem denir ve sınır genişliğini belirleyen noktalar ise destek vektörleridir. Şekil 4-3'te (Kavzoğlu vd., 2012) ikili sınıflandırma problemlerinde optimum hiper düzlemin belirlenmesi gösterilmiştir.



Şekil 4-3 İkili sınıflandırmada optimum hiperdüzlemin belirlenmesi

4.1.18. Çekirdek hilesi

Çekirdek hilesi çoğunlukla, doğrusal olarak ayrılamayan verilerdeki problemleri çözmek için kullanılmaktadır. Çekirdek hilesi ile veri setinin boyutu arttırılarak verinin daha doğru bir şekilde sınıflandırılması sağlanmaktadır. Bu durum "Cover Teoremi" nde şu şekilde açıklanmıştır: karmaşık kalıba sahip bir sınıflandırma problemi yüksek boyutlu bir uzayda doğrusal olmayacak bir şekilde oluşmuş ise düşük boyutlu uzaydakine kıyasla doğrusal olarak ayrılabilir (Haykin, 1994).

4.1.19. Naive bayes

Bayes teoremi, rassal değişken için olasılık dağılımı içinde koşullu olasılıklar ile marjinal olasılıklar arasındaki ilişkiyi gösterir. NB sınıflandırıcıları Basit Bayes Algoritmaları olarak bilinir (Lewis, 1992).

Tembel öğrenme algoritması olmasıyla birlikte dengesiz veri setlerinde oldukça başarılı bir şekilde çalışmaktadır. Algoritmanın çalışma mantığı: herhangi bir eleman için her durumun olasılığını hesaplar ve olasılık değeri en yüksek olana göre hesaplama yaparak sınıflandırma yapar. Küçük veri setlerinde de oldukça

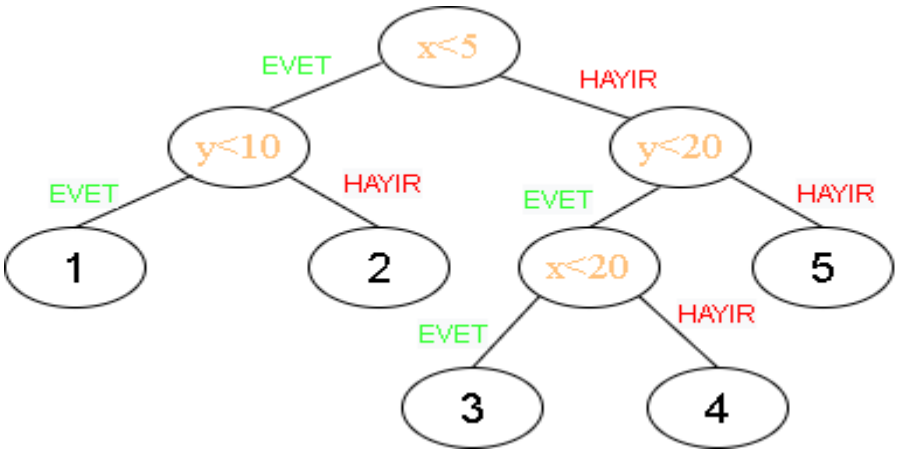
başarılı bir şekilde çalışabilmektedir. Uygulamalarda Naive Bayes' in basit yapısı ve uygulama kolaylığı göz önüne alındığında genellikle iyi performans gösterir (Yousef, et al., 2006).

4.1.20. Karar Ağacı

Sınıflandırma özellik ve hedefe göre karar düğümleri ve yaprak düğümlerinden oluşan ağaç yapısı formunda bir model oluşturan sınıflandırma algoritmasıdır. Bir karar düğümü birden fazla düğüm içerebilmektedir.

İlk düğüme kök düğüm denmektedir. Kök yapısı verinin yapısına ve değerlerine göre dallanarak sınıflandırma yapabilmek için uygun bir yapı oluştururlar. Bir karar ağacı hem kategorik hemde sayısal verilerden oluşabilir.

Sınıflandırma, bilimsel deneyler, tıbbi teşhis, hava tahmini, kredi onayı, müşteri bölümlemesi, pazarlama ve sahtekarlık tespiti gibi çok çeşitli uygulama alanlarına başarıyla uygulanmıştır (Lavanya ve Rani, 2011). Şekil 4-4'te KA' nın çalışma mantığı verilmiştir. İlk düğümden başlayarak verilen cevaplar dikkate alınarak farklı sonuçlar üretebilmektedir.



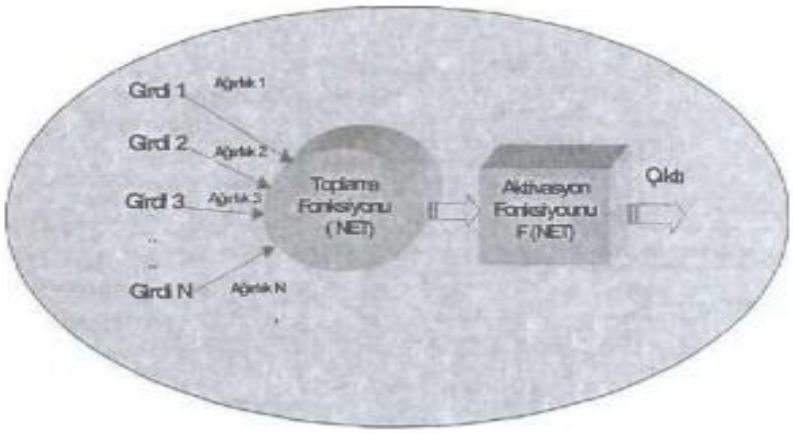
Şekil 4-4 KA modeli

4.1.21. Rassal orman

RO sınıflandırıcısı, her sınıflandırıcının giriş vektöründen bağımsız olarak örneklenen rastgele bir vektör kullanılarak oluşturulduğu ağaç sınıflandırıcılarının bir kombinasyonundan oluşur ve her ağaç bir girdi vektörünü sınıflandırmak için en popüler sınıf için bir birim oyu verir (Breiman, 1999). Ağaç modelleri, özelliklerinin kombinasyonunu kullanarak yeni veriler üzerinde maksimum derinliğe kadar büyür. Bu algoritmanın ağaç sayısı ve parametre sayısı olmak üzere iki önemli parametresi bulunmaktadır.

4.1.22. Yapay sinir ağları

YSA ağırlıklandırılmış biçimde birbirine bağlanan birçok nörondan oluşan matematiksel modellerdir. YSA, temelinde tamamen insan beyni örneklenerek geliştirilmiş bir teknolojidir (Ergezer vd., 2003). Şekil 4-5'te YSA' nın yapısı gösterilmiştir (Öztemel, 2012).



Şekil 4-5 YSA yapısı

Girdiler: Bir yapay sinir hücresine dışardan gelen bağımsız değişkenlerdir.

Ağırlıklar: Yapay sinir hücresine gelen bağımsız değişkenin önemini ve etkisini gösterir.

Toplama Fonksiyonu: Bir hücreye gelen bağımsız değişkenlerden net girdiyi hesaplar. Bunun hesabını yaparken ağırlıkların ve bağımsız değişkenlerin çarpımından gelen değerleri kullanır.

Aktivasyon Fonksiyonu: Toplama fonksiyonunun üretmiş olduğu net girdiyi alarak aktivasyon fonksiyonunun işlevine yapısına bağlı olarak sonuç üretir.

Çıktı: Aktivasyon fonksiyonunun ürettiği değeri temsil eder.

4.1.23. Modelin Oluşturulması ve Değerlendirilmesi

MÖA kullanılarak oluşturulacak birçok model vardır. Modellerde performansın yüksek olması için veriye uygun algoritmaların seçilmesi önemlidir. Modelin oluşturulmasında eğitim için ayrılmış olan veri seti kullanılmaktadır.

Modelin değerlendirilmesinde birçok yöntem kullanılmaktadır. Modelin değerlendirilmesi eğitim veri seti kullanılarak oluşturulan modelin test veri seti ile sınanması ile gerçekleştirilir. Bu çalışmada doğruluğunun ölçülmesi için karmaşıklık matrisi (confusion matrix) ve buna bağlı parametreler kullanılmıştır.

Karmaşıklık Matrisi: Sınıflandırma yapmak için oluşturulan modellerin performanslarının ölçülmesi için kullanılan tablodur. İki veya daha fazla etiketli veriler için kullanılır ve beklenen ve tahmin edilen değerler arasında 4 farklı kombinasyon vardır. Değerlendirme sonucunda verilerde herhangi bir eksiklik yok ise örneklem bu kombinasyonlardan birinin içerisinde olmak zorundadır. Bu kombinasyonlar Çizelge 4-1'de gösterilmiştir.

Çizelge 4-1 Karmaşıklık matrisi

		Gerçek Değer	
		Pozitif(1)	Negatif(0)
Tahmin	Pozitif(1)	DP(Doğru Pozitif)	YP(Yanlış Pozitif)

Edilen Değer	Negatif(0)	YN(Yanlış Negatif)	DN(Doğru Negatif)

DP: Tahmin edilen değer ve gerçek değerin pozitif olduğu durumu kapsar. Bu durumda değerlendirme doğru yapılmıştır.

YP: Tahmin edilen değer pozitif gerçek değerin ise negatif olduğu durumu kapsar. Bu durumda değerlendirme doğru yapılamamıştır.

YN: Tahmin edilen değer negatif gerçek değerin ise pozitif olduğu durumu kapsar. Bu durumda değerlendirme doğru yapılamamıştır.

DN: Tahmin edilen değerin negatif ve gerçek değerinde negatif olduğu durumu kapsar. Bu durumda değerlendirme doğru yapılmıştır.

Yukarıda belirtilen parametreler kullanılarak performansın ölçülmesinde kullanılan: doğruluk, hata oranı, kesinlik, duyarlılık, F-Ölçütü parametreleri hesaplanmaktadır.

Doğruluk (accuracy): Modelin başarımının değerlendirilmesinde kullanılan popüler ve basit parametrelerdendir. Doğruluk değeri, Çizelge-3-1'deki veriler kullanılarak aşağıdaki gibi hesaplanmaktadır.

Doğruluk= (DP+DN)/(DP+YP+YN+DN)

Hata Oranı: Doğru sınıflandırılmış olan örnek sayısı oranın 1'e tamlayanı ya da yanlış sınıflandırılmış örnek sayısının test verisindeki örnek sayısına oranı olarak tanımlanır. Hata Oranı, Çizelge-3-1'deki veriler kullanılarak aşağıdaki gibi hesaplanır.

Hata Oranı=(YP+YN) /(DP+YP+YN+DN)

Kesinlik: Sınıfı 1 olarak tahmin edilen örnek sayısının, sınıfı 1 olarak tahmin edilmiş tüm örnek sayısına oranıdır.

Kesinlik= (DP)/(DP+YP)

Duyarlılık: Doğru sınıflandırılmış pozitif örnek sayısının, toplam pozitif örnek sayısına oranıdır.

Duyarlılık=(DP)/(DP+YN)

F-Ölçütü: Kesinlik ve duyarlılık ölçütleri, tek başlarına anlamlı sonuç çıkartmamız için yeterli değildir. Her iki ölçütü beraber değerlendirmek daha doğru bir yaklaşım olacaktır. Bunun için f-ölçütü(F) tanımlanmıştır. F, kesinlik(K) ve duyarlılık(D) ölçütlerinin harmonik ortalamasına eşittir.

F= (2*D*K)/(D+K)

5. VERİ SETİ VE UYGULAMA

5.1. Veri Seti

STS' lerin daha iyi sonuç verebilmesi için bu alanda birçok çalışma yapılmıştır. Yapılan çalışmaların birçoğu geçerli ve güncel bir veri seti bulma konusunda sorun yaşadıkları için var olan birkaç veri seti üzerinde deneyler gerçekleştirilmiştir. DARPA, KDD'99, DEFCON, CDX, KYOTO, TWENTE, UMASS, ISCX (Sharafaldin vd.,2018) bu veri setlerinden bazılarıdır.

Kanada Siber Güvenlik Enstitüsü tarafından paylaşılan, uygun bir test ortamında hazırlanmış CSE-CIC-IDS2018 (CyberSecurity, 2018) veri seti, daha önce kullanılan veri setlerindeki eksiklikler dikkate alınarak oluşturulmuştur. Bu veri seti; 5 günlük bir test sürecinde, tehdit profilleri ve iyi huylu davranış profillerinin trafikleri göz önüne alınarak üretilmiştir. Veri seti PCAP dosya formatında üretilmiş ve ağ trafiği akış üreteci olan CICFlowMeter uygulaması kullanılarak CSV formatına dönüştürülmüştür. Kanada Siber Güvenlik Enstitüsü tarafından paylaşılan bu veriler, halkın kullanımına açılmış ve siber güvenlik alanında çalışma yapmak isteyen araştırmacılar bu veri setinin PCAP ve CSV formatına erişebilmektedir (CyberSecurity, 2018).

Veri seti Amazon Web Service üzerinden yerel bilgisayara indirilip incelendiğinde, içerisinde yer alan etiketler ve her bir etikete ait örnek sayısı Çizelge 5-1'de verilmiştir. Veri setinin test edilebilmesi için bazı ön işlemlerden geçmesi gerekmektedir. Ön işlemler için "dizgi" veri türünden olan zaman parametresi sayısal değere dönüştürülmüştür. Ayrıca parametresi eksik olan verilerin olduğu tespit edilmiş ve bu örneklemler silinerek (normalleştirme) veri seti uygulama ortamı için uygun hale getirilmiştir.

Veri seti içerisinde yer alan hedef port, protokol ve saldırının yapıldığı zaman ifadesi özellik olarak veri setlerinden çıkartılmışlardır.

Çizelge 5-1 Veri setlerindeki her bir etikete ait örnek sayısı

ETİKET	ÖRNEKLEM SAYISI
Benign (zararsız)	6000000
Bot	290000
BruteForce-Web	612
BruteForce-XSS	231
DDOS	690000
DoS attacks-Slowloris	11000
DoSattacks-GoldenEye	41500
DoSattacks-Hulk	462000
DoSattacks-SlowHTTPTest	140000
FTP-BruteForce	196000
Infilteration	161000
SQLInjection	90
SSH-Bruteforce	188000

5.1.1. Veri seti içerisinde yer alan tehditler

Veri setlerinde yer alan tehditlere ait 79 özellik göz önüne alınarak değerlendirilmiştir. Bazı araçlar kullanılarak belli ortam koşullarında test edilmişlerdir. Bu test sonuçlarında elde edilen değerler veri setlerine yansıtılmışlardır.

5.1.1.1. Brute force saldırıları

Kaba kuvvet saldırısı ve şifrelerin kırılması için Hydra, Medusa, Ncrack, Metasploit modülleri ve Nmap NSE komut dosyaları gibi birçok araç kullanılmaktadır. Ancak çok kapsamlı araçlardan bir tanesi de Patator aracıdır.

Bu saldırı türünde veri seti oluşturmak için saldırgan Kali Linux makine üzerinden SSH ve FTP protokollerini kullanarak kurban makine olan Ubuntu14.0 saldırmıştır. Şifrelerin kırılabilmesi için 90 milyon kelime içeren bir sözlük kullanılmaktadır (CyberSecurity, 2018).

5.1.1.2. Botnet

Birçok saldırı türünde kullanılan araçlardan olan Zeus ve Ares birçok yeteneğe sahiptir. Zeus Microsoft Windows sürümlerinde çalışan bir Truva atı kötü amaçlı yazılım paketidir. İlk olarak Temmuz 2007'de Amerika Birleşik Devletleri Ulaştırma Bakanlığı'ndan bilgi çalmak için kullanılmıştır (Abrams., 2013). Gizli teknikleri kullanarak kendini gizleyebildiğinden Zeus'un güncel anti virüs ve diğer güvenlik yazılımlarıyla bile algılanması zordur. Bununla birlikte tamamlayıcısı olabilmesi için uzaktan erişim, dosya yükleme/indirme ekran görüntüsü ve anahtar kaydı gibi yetenekleri olan Ares kullanılmıştır (CyberSecurity, 2018).

Bu saldırı için oluşturulan veri setleri, iki farklı botnet (Zeus ve Ares) bulunan makinelerden bulaşmaktadır. Bununla birlikte her 400 saniyede bir zombilerin ekran görüntülerini talep edilmektedir (CyberSecurity, 2018).

5.1.1.3. DDOS saldırıları

Bu saldırı türünü oluşturmak için HOIC aracı kullanılmıştır. HOIC, eşzamanlı 256 adet saldırı yürütebilir ve meşru istekler işleninceye kadar sürekli trafik akışı göndererek hedef sistemi çökertebilir. HOIC'in aldatıcı ve varyasyon teknikleri, geleneksel güvenlik araçlarının ve güvenlik duvarlarının DDOS saldırılarını tam olarak saptamasını ve engellemesini zorlaştırmaktadır (imperva, 2020).

Bu tehdidin veri setini oluşturmak için, 4 farklı bilgisayar kullanılmış ve DDOS saldırısı yapmak için ücretsiz HOIC aracı kullanılmıştır (CyberSecurity, CSE-CIC-IDS2018 on AWS , 2018).

5.1.1.4. DOS saldırıları

Tek bir makinenin başka bir makinenin web sunucusunu, ilgisiz hizmetler ve bağlantı noktaları üzerinde minimum bant genişliği ve yan etkilerle almasını

sağlayan bir tür hizmet reddi saldırı aracı olan Slowloris, Robert Hansen tarafından üretilmiştir.

Bu saldırı türünde veri setini oluşturulurken Slowloris Perl tabanlı bir araç kullanılmıştır (CyberSecurity, 2018).

5.1.1.5. Web attacks

Damn Vulnerable Web App (DVWA)'ın temel hedefi, güvenlik profesyonellerinin becerilerini ve araçlarını yasal bir ortamda test etmelerine yardımcı olmaktır.

Bu saldırı türleri için Çok Savunmasız Web Uygulaması (DVWA) kullanılmıştır (CyberSecurity, 2018).

5.1.1.6. Heartbleed saldırıları

Heartleech hataya açık olan sistemleri tarayabilen ve daha sonra bunlardan yararlanmak ve verileri dışarı sızmak için kullanılabilen bir araçtır.Bu veri seti, Heartbleed güvenlik açığından faydalanmış olup ve heartleech kullanılarak oluşturulmuştur (CyberSecurity, 2018).

5.1.1.7. Infiltration saldırıları

Bu saldırı türü için öncelikle mağdura e-posta yoluyla bir belge yollanmıştır. Ardından, Metasploit çerçevesi kullanılarak başarılı bir şekilde sömürülüp kurbanın bilgisayarında bir yazılım (back door) çalıştırılmıştır. Bu şekilde kurbanın ağına IP taraması, tam port tarama ve Nmap kullanarak saldırılabilmesi sağlanmıştır.

5.1.2. CICFlowMeter

CICFlowMeter, 84 ağ trafiği özelliği oluşturabilen bir ağ trafiği akış üretecidir. Java dili ile yazılmış açık kaynak kodlu bir uygulama olup GitHub' dan indirilebilmektedir. Kaynak kod, hesaplamak istenilen özellikleri seçme, yenilerini ekleme ve akış zamanı aşım süresini daha iyi kontrol etme konusunda esneklik tanıdığı için projelere entegre olabilmektedir. Bu uygulama PCAP dosyasını okur ve çıkarılan özelliklerin grafiksel raporunu oluşturur, ayrıca raporun CSV formatlı dosyasını da çıkartır.

Paketler ileri (kaynaktan hedefe) ve geri (hedeften kaynağa) yönleri belirlediği çift yönlü akışlar üretmek için kullanılabilmektedir. Bu nedenle süre, paket sayısı, bayt sayısı, paketlerin uzunluğu vb. ileri ve geri yönlerde ayrı ayrı hesaplanır (CyberSecurity, CSE-CIC-IDS2018, 2018).

5.1.3. Veri setinde yer alan özellikler ve açıklamaları

Veri seti içerisinde yer alan her bir örnekleme ait 79 özellik ve 1 etiket bulunmaktadır. Çizelge 5-2'de veri seti içerisinde yer alan örneklemlere ait özelliklerin açıklamaları verilmiştir.

Çizelge 5-2 Veri setine ait özellikler

Özellik	Açıklama
Dst Port	Hedef Port
Protocol	Protokol
Timestamp	Saldırının yapıldığı tarih ve zaman bilgisi
Flow Duration	Mikro-saniyedeki akış süresi
Tot Fwd Pkts	İleri yönde toplam paket sayısı
Tot Bwd Pkts	Geri yönde toplam paket sayısı
TotLen Fwd Pkts	İleri yönde paketlerin toplam uzunluğu
TotLen Bwd Pkts	Geri yönde paketlerin toplam uzunluğu
Fwd Pkt Len Max	İleri yönde paketlerin maksimum uzunluğu
Fwd Pkt Len Min	İleri yönde paketlerin minimum uzunluğu
Fwd Pkt Len	İleri yönde paketlerin ortalama uzunluğu

Mean	
Fwd Pkt Len Std	İleri yönde paket uzunluklarının standart sapması
Bwd Pkt Len Max	Geri yönde paketlerin maksimum uzunluğu
Bwd Pkt Len Min	Geri yönde paketlerin minimum uzunluğu
Bwd Pkt Len Mean	Geri yönde paketlerin ortalama uzunluğu
Bwd Pkt Len Std	Geri yönde paket uzunluklarının standart sapması
Flow Byts/s	Saniyede akan byte sayısı
Flow Pkts/s	Saniyede akan paket sayısı
Flow IAT Mean	Paketlerin ortalama varış zamanı
Flow IAT Std	Paketlerin varış zamanlarının standart sapması
Flow IAT Max	Paketlerin maximum varış zamanı
Flow IAT Min	Paketlerin Minimum varış zamanı
Fwd IAT Tot	İleri yönde gönderilen iki paket arasındaki toplam zaman
Fwd IAT Mean	İleri yönde gönderilen iki paket arasındaki ortalama zaman
Fwd IAT Std	İleri yönde gönderilen iki paket arasındaki zamanın standart sapması
Fwd IAT Max	İleri yönde gönderilen iki paket arasındaki maksimum zaman
Fwd IAT Min	İleri yönde gönderilen iki paket arasındaki minimum zaman
Bwd IAT Tot	Geri yönde gönderilen iki paket arasındaki toplam zaman
Bwd IAT Mean	Geri yönde gönderilen iki paket arasındaki ortalama zaman
Bwd IAT Std	Geri yönde gönderilen iki paket arasındaki zamanın standart sapması
Bwd IAT Max	Geri yönde gönderilen iki paket arasındaki maksimum zaman
Bwd IAT Min	Geri yönde gönderilen iki paket arasındaki minimum zaman
Fwd PSH Flags	İleri yönde hareket eden paketlerde PSH bayrağının aktif olma sayısı (UDP için 0)
Bwd PSH Flags	Geri yönde hareket eden paketlerde PSH bayrağının aktif olma sayısı (UDP için 0)
Fwd URG Flags	İleri yönde hareket eden paketlerde URG bayrağının aktif olma sayısı (UDP için 0)
Bwd URG Flags	Geri yönde hareket eden paketlerde URG bayrağının aktif olma sayısı (UDP için 0)
Fwd Header Len	İleri yöndeki başlıklar için kullanılan toplam byte
Bwd Header Len	Geri yöndeki başlıklar için kullanılan toplam byte
Fwd Pkts/s	Saniyedeki ileri yön paket sayısı
Bwd Pkts/s	Saniyedeki geri yön paket sayısı
Pkt Len Min	Bir akışın minimum uzunluğu

Pkt Len Max	Bir akışın maksimum uzunluğu
Pkt Len Mean	Bir akışın ortalama uzunluğu
Pkt Len Std	Bir akışın standart sapması
Pkt Len Var	Bir akışın uzunluk varyansı
FIN Flag Cnt	FIN içeren paket sayısı
SYN Flag Cnt	SYN içeren paket sayısı
RST Flag Cnt	RST içeren paket sayısı
PSH Flag Cnt	PUSH içeren paket sayısı
ACK Flag Cnt	ACK içeren paket sayısı
URG Flag Cnt	URG içeren paket sayısı
CWE Flag Count	CWE içeren paket sayısı
ECE Flag Cnt	ECE içeren paket sayısı
Down/Up Ratio	İndirme ve yükleme oranı
Pkt Size Avg	Ortalama paket boyutu
Fwd Seg Size Avg	İleri yönde gözlenen ortalama boyut
Bwd Seg Size Avg	Geri yönde gözlenen ortalama boyut
Fwd Byts/b Avg	İleri yönde byte/kütle oranının ortalama sayısı
Fwd Pkts/b Avg	İleri yönde paket/kütle oranının ortalama sayısı
Fwd Blk Rate Avg	İleri yönde kütle oranının ortalama sayısı
Bwd Byts/b Avg	Geri yönde byte/kütle oranının ortalama sayısı
Bwd Pkts/b Avg	Geri yönde paket/kütle oranının ortalama sayısı
Bwd Blk Rate Avg	Geri yönde kütle oranının ortalama sayısı
Subflow Fwd Pkts	İleri yönde bir alt akıştaki ortalama paket sayısı
Subflow Fwd Byts	İleri yönde bir alt akıştaki ortalama byte sayısı
Subflow Bwd Pkts	Geri yönde bir alt akıştaki ortalama paket sayısı
Subflow Bwd Byts	Geri yönde bir alt akıştaki ortalama byte sayısı
Init Fwd Win Byts	İleri yönde ilk pencere içinde gönderilen bayt sayısı
Init Bwd Win Byts	Geri yönde ilk pencere içinde gönderilen bayt sayısı
Fwd Act Data Pkts	İleri yönde en az 1 bayt TCP veri yüküne sahip paket sayısı
Fwd Seg Size Min	İleri yönde gözlenen minimum segment boyutu
Active Mean	Boşta kalmadan önce bir akışın aktif olduğu ortalama zaman
Active Std	Boşta kalmadan önce bir akışın aktif olduğu zamanın standart sapması
Active Max	Boşta kalmadan önce bir akışın aktif olduğu maksimum zaman

Active Min	Boşta kalmadan önce bir akışın aktif olduğu minimum zaman
Idle Mean	Aktif hale gelmeden önce bir akışın boşta olduğu ortalama zaman
Idle Std	Aktif hale gelmeden önce bir akışın boşta olduğu zamanın standart sapması
Idle Max	Aktif hale gelmeden önce bir akışın boşta olduğu maksimum zaman
Idle Min	Aktif hale gelmeden önce bir akışın boşta olduğu minimum zaman

5.1.4. Veri setindeki örneklemlerin tanımlanması

Veri setlerinin tanımlanması için bazı kavramlar kullanılmaktadır. Bu kavramlar etiketli veriler arasındaki dağılımın gösterilmesi bakımından önemlidir. İstatistikte kullanılan bu hesaplamalar Merkezi Eğilim Ölçüleri olarak tanımlanır. Veri setinde elde edilen sonuçların verinin merkezine olan uzaklığı dikkate alınır. Bu veriler doğrultusunda, veri seti içerisindeki örneklemler veya veri setinin tamamı hakkında kararlar verilir.

Bu kavramlar: mod, aritmetik ortalama, medyan, standart sapma, minimum, maksimum ve çeyreklerdir (%25, %50, %75).

Mod: Bir sayısal veri serisi içerisinde en çok tekrar edilen sayıya mod denir.

Aritmetik Ortalama (MEAN): Bir sayı serisinde yer alan sayıların toplamının, serinin eleman sayısına bölünmesiyle elde edilen değerdir.

Medyan: Bir sayısal veri serisi sıralandığında ortada kalan sayıya denir.

Standart Sapma (STD): Standart sapma verilerin ne kadarının ortalamaya yakın olduğu hesaplanır. Bu durumda eğer standart sapma küçük ise veriler ortalamaya yakın, aksi takdirde veriler ortalamaya uzaktır.

Standart sapmanın hesaplanması için izlenen adımlar:

- Sayıların aritmetik ortalamasının hesaplanması

- Her bir sayı için aritmetik ortalamadan farkı bulunur
- Bulunan her bir farkın karesi alınarak toplanır.
- Elde edilen değer, örneklem sayısının bir eksiğine bölünür.
- Bulunan sayının karekökü alınarak standart sapma hesaplanmış olur.

Minimum(MIN): Seri içerisinde yer alan değerlerden en küçüğünü ifade eder.

Maksimum(MAKS): Seri içerisinde yer alan değerlerden en büyüğünü ifade eder.

Çeyrekler: Seri içerisinde üç tane çeyrek hesaplanır. Bunlar Çeyrek 1 %25, Çeyrek 2 %50, Çeyrek 3 %75 olarak da adlandırılır. Serinin sıralanması sonucunda, ortadaki sayı(medyan) Çeyrek 2'yi verir. Çeyrek 1 ise serinin sıralanması sonucunda ilk yarısının ortancası, Çeyrek 3 ise serinin sıralanması sonucunda ikinci yarısının ortancasıdır.

5.2. Uygulama

Verilerin eğitim ve test için daha uygun bir şekilde kullanılabilmesi için her tehdit türüne ait veriler farklı dosyalara yazılmıştır. Her bir model de kullanılmak üzere oluşturulan veri setleri, BruteForce yöntemi kullanılarak daha az özellikte ifade edilecek şekilde oluşturulmuştur. Denemeler sonucunda uygun sayıda özellikler seçilerek eğitim ve test verisi için 6 farklı dosya hazırlanmıştır. Çizelge 5-3'te her bir veri seti için eğitim ve test veri setlerindeki örneklem sayısı verilmektedir.

Çizelge 5-3 Veri setlerine ait eğitim ve test veri örneklem sayıları

Veri Seti	Eğitim Verisi	Test Verisi
VERI SETI-1	3.019.056	1.006.352
VERI SETI-2	918.494	459.247
VERI SETI-3	820.466	410.233
VERI SETI-4	382.470	191.235
VERI SETI-5	502.392	251.196
VERI SETI-6	212.796	106.398

5.2.1. Herhangi bir tehdidinin tespiti için seçilen özellikler ve sonuçlar

1.Veri seti herhangi bir tehdidin olup olmadığını tespit etmek için oluşturulmuştur. Bu veri setinde, DDOS, DOS attacks Slowris, DOSattacks-GoldenEye, DOSattacks-Hulk, DOSattacks-SlowHTTPTest, Bot, BruteForce-Web, BruteForce-XSS, FTP-BruteForce, BruteForce-SSH ve Infilteration tehdit olarak kullanılmış ve Benign (iyi huylu) normal durum (tehdit yok) olarak kullanılmıştır. Denemeler sonucunda 79 özellikten sadece 17 özellik ile iyi sonuçlar alındığı gözlemlenmiştir. Çizelge 5-4'te veri setini oluştururken kullanılan özellikler ve açıklamaları verilmiştir.

Çizelge 5-4 Veri setlerinin özellikleri ve açıklamaları

ÖZELLİK	AÇIKLAMA
Flow Duration	Mikro-saniyedeki akış süresi
Flow Byts/s	Saniyede akan byte sayısı
Flow Pkts/s	Saniyede akan paket sayısı
Flow IAT Std	Paketlerin varış zamanlarının standart sapması
Flow IAT Max	Paketlerin maksimum varış Zamanı
Fwd IAT Tot	İleri yönde gönderilen iki paket arasındaki toplam zaman
Bwd Pkts/s	Saniyedeki geri yön paket sayısı
Pkt Len Std	Bir akışın standart sapması
Pkt Len Var	Bir akışın uzunluk varyansı
RST Flag Cnt	RST içeren paket sayısı
ECE Flag Cnt	ECE içeren paket sayısı
Down/Up Ratio	İndirme ve yükleme oranı
Bwd Blk Rate Avg	Geri yönde kütle oranının ortalama sayısı
Init Fwd Win Byts	İleri yönde ilk pencere içinde gönderilen bayt sayısı
Fwd Act Data Pkts	İleri yönde en az 1 bayt TCP veri yüküne sahip paket sayısı
Fwd Seg Size Min	Geri yönde kütle oranının ortalama sayısı
Active Max	Boşta kalmadan önce bir akışın aktif olduğu maksimum zaman

Veri seti içerisinde yer alan özellikler göz önünde bulundurularak tehdide ve normal duruma ait verilerin dağılımı gösterilmeye çalışılmıştır. Bunun için tehdidi tanımlayan veriler Çizelge 5-5'te normal durumu tanımlayan veriler ise Çizelge 5-6'da gösterilmiştir.

Çizelge 5-5 1.veri setindeki tehdide ait özelliklerin özeti

Özellik	MEAN	STD	MIN	25%	50%	75%	MAX
Bwd Pkts/s	839942	8098043	1	483	4388	16171	120000000
Flow Duration	26606	234532	0	0	0	0	230000000
RST Flag Cnt	181602	438226	0	143	690	4292	2000000
Down/Up Ratio	198662	2799586	0	0	0	395	82170520
Flow IAT Std	553863	5616572	1	483	4286	15440	119464300
Fwd IAT Tot	795228	8070666	0	474	2089	15015	120000000
Active Max	90422	219251	0	0	0	432	1000000
Init Fwd Win Byts	53	108	0	0	0	0	1286
Pkt Len Var	14610	34275	0	0	0	0	1652619
Flow Byts/s	0	0	0	0	0	0	1
ECE Flag Cnt	0	0	0	0	0	0	1
Bwd Blk Rate Avg	0	1	0	0	0	1	57
Fwd Seg Size Min	0	0	0	0	0	0	0
Fwd Act Data Pkts	20983	19080	-1	241	26883	32738	65535
Flow IAT Max	102	3540	0	0	0	0	309628
Flow Pkts/s	27	8	0	20	32	32	40
Pkt Len Std	21147	395344	0	0	0	0	13823000

Çizelge 5-6 1.veri setindeki tehdit olmayan duruma ait özelliklerin özeti

Özellik	MEAN	STD	MİN	25%	50%	75%	MAX
Bwd Pkts/s	1E+07	3,2E+07	1	0	2E+06	5E+06	120000000
Flow Duration	206027	3247232	0	39	334	6233	788000000
RST Flag Cnt	21093	179054	0	2	7	2053	4000000
Down/Up Ratio	1E+06	3324442	0	0	79327	2E+06	84500000
Flow IAT Std	9E+06	2,2E+07	1	0	1E+06	4E+06	120000000
Fwd IAT Tot	1E+07	3,2E+07	0	86	2E+06	4E+06	120000000
Active Max	2529	36978	0	0	1	51	2000000
Init Fwd Win	165	167	0	9	80	317	1288

Byts							
Pkt Len Var	54907	79571	0	85	6352	1E+05	1657914
Flow Byts/s	0	0	0	0	0	0	1
ECE Flag Cnt	0	0	0	0	0	0	1
Bwd Blk Rate Avg	0	1	0	0	0	1	148
Fwd Seg Size Min	0	0	0	0	0	0	0
Fwd Act Data Pkts	6527	14616	-1	123	254	8192	65535
Flow IAT Max	2	10	0	0	1	3	9262
Flow Pkts/s	20	8	0	20	20	20	48
Pkt Len Std	153234	2119633	0	0	0	0	111000000

1. veri setine MÖA uygulanarak oluşturulan modeller üzerinde elde edilen sonuçlar çizelge 5-7'de verilmiştir. Çizelge de yer alan 1. sütunda uygulanan MÖA, 2. sütunda test sonuçlarında elde edilen Doğruluk Oranı (DO), 3. sütunda modelin eğitilmesi için geçen süre yani Eğitim Süresi (ES), 4. sütunda ise Test Süresi (TS) verilmiştir.

Çizelge 5-7 1. veri setine ait deney sonuçları

MÖA	**DO (%)**	**ES (sec.)**	**TS (sec.)**
YSA	99,11	457	**34**
LR	93,794	119	1
NB	88,065	10	2
KNN	99,814	10.209	1.735
KA	99,974	33	1
RO	99,969	84	2
DVM	94,980	21.351	5.182

5.2.2. DDOS tehdidinin tespiti için seçilen özellikler

2.veri seti tehdidin DDOS olup olmadığını kontrol etmek için oluşturulmuştur. Bu veri seti, zararsız durum için Benign veri dosyasından ve anormal durum için ise DDOS veri dosyasından alınarak oluşturulmuştur. Denemeler sonucunda 79 özellikten sadece 2 özellik ile iyi sonuçlar alındığı gözlemlenmiştir. DDOS saldırı türü için belirlenen özellikler ve açıklamaları Çizelge 5-8'de verilmiştir

Çizelge 5-8 DDOS saldırı türü için belirlenen özellikler ve açıklamalar

ÖZELLİK	AÇIKLAMA
Init Fwd Win Byts	İleri yönde ilk pencere içinde gönderilen bayt sayısı
Pkt Len Var	Bir akışın uzunluk varyansı

Tehdidi (DDOS) tanımlayan veriler çizelge 5-9'da normal durumu tanımlayan veriler ise Çizelge 5-10'da gösterilmiştir.

Çizelge 5-9 2.veri setindeki tehdide ait özelliklerin özeti

ÖZELLİK	MEAN	STD	MİN	25%	50%	75%	MAX
Init Fwd Win Byts	26331	47106	0	0	0	0	113743
Pkt Len Var	40465	14110	-1	32738	32738	32738	65535

Çizelge 5-10 2.veri setindeki tehdit olmayan duruma ait özelliklerin özeti

ÖZELLİK	MEAN	STD	MİN	25%	50%	75%	MAX
Init Fwd Win Byts	47273	85908	0	0	2028	92544	710533
Pkt Len Var	6922	13560	-1	-1	984	8192	65535

Bu veri setine MÖA uygulanması sonucu elde edilen sonuçlar çizelge 5-11'de verilmiştir.

Çizelge 5-11 2. veri setine ait deney sonuçları

MÖA	DO (%)	ES (sec.)	TS (sec.)
YSA	99,31	98	11
LR	97,111	4	1
NB	97,110	2	1
KNN	99,879	1.540	844
KA	99,879	2	1
RO	99,879	6	1
DVM	97,144	2.468	1.855

5.2.3. DOS tehdidinin tespiti için seçilen özellikler

3.veri seti tehdidin DOS olup olmadığını kontrol etmek için oluşturulmuştur. Bu veri seti, zararsız durum için Benign veri dosyasından ve anormal durum için DoS attacks-Slowloris, DoSattacks-GoldenEye, DoSattacks-Hulk, DoSattacks-SlowHTTPTest veri dosyalarından alınarak oluşturulmuştur. Denemeler sonucunda 79 özellikten sadece 6 özellik ile en iyi sonuçlar alındığı gözlemlenmiştir. DOS saldırı türü için belirlenen özellikler ve açıklamaları çizelge 5-12'de verilmiştir.

Çizelge 5-12 DOS saldırı türü için belirlenen özellikler ve açıklamalar

ÖZELLİK	AÇIKLAMA
Flow Pkts/s	Saniyede akan paket sayısı
Pkt Len Var	Bir akışın uzunluk varyansı
Init Fwd Win Byts	İleri yönde ilk pencere içinde gönderilen bayt sayısı
Flow IAT Max	Paketlerin maksimum varış zamanı
Flow Duration	Mikro-saniyedeki akış süresi
Flow Byts/s	Saniyede akan byte sayısı

Tehdidi (DOS) tanımlayan veriler çizelge 5-13'te normal durumu tanımlayan veriler ise çizelge 5-14'te gösterilmiştir.

Çizelge 5-13 3.veri setindeki tehdide ait özelliklerin özeti

Özellik	MEAN	STD	MİN	25%	50%	75%	MAX
Flow Pkts/s	2134685	12523190	1	374	6697	40344	1,19E+08
Pkt Len Var	6409	360136	0	0	0	0	2,3E+08
Init Fwd Win Byts	159031	336119	0	58	310	5348	2000000
Flow IAT Max	1649586	9708580	1	374	6647	38667	1,19E+08
Flow Duration	8178	27750	0	0	0	0	160833
Flow Byts/s	7889	12066	0	225	225	26883	26883

Çizelge 5-14 3.veri setindeki tehdit olmayan duruma ait özelliklerin özeti

Özellik	MEAN	STD	MİN	25%	50%	75%	MAX
Flow Pkts/s	15305510	34210840	1	611	139804	3917790	1,2E+08
Pkt Len Var	280001	3765046	0	3	1147	61765	4,8E+08
Init Fwd Win Byts	28135	208104	0	5	34	5249	3000000
Flow IAT Max	7720172	19998800	1	402	98200	985731	1,2E+08
Flow Duration	46596	86882	0	0	1825	89047	710533
Flow Byts/s	6759	13381	-1	-1	596	8192	65535

Bu veri setine MÖA uygulanması sonucu elde edilen sonuçlar çizelge 5-15'te verilmiştir.

Çizelge 5-15 3. veri setine ait deney sonuçları

MÖA	DO (%)	ES (sec.)	TS(sec.)
YSA	92,26	552	66
LR	67,022	7	1
NB	63,209	2	1
KNN	99,897	983	96
KA	99,988	5	1
RO	99,987	16	1
DVM	65,440	5.351	2.051

5.2.4. Bot tehdidinin tespiti için seçilen özellikler

4.veri seti tehdidin Bot olup olmadığını kontrol etmek için oluşturulmuştur. Bu veri seti, zararsız durum için Benign veri dosyasından ve anormal durum için Bot veri dosyasından alınarak oluşturulmuştur. Denemeler sonucunda 79 özellikten sadece 4 özellik ile en iyi sonuçlar alındığı gözlemlenmiştir. Bot saldırı türü için belirlenen özellikler ve açıklamaları çizelge 5-16'da verilmiştir.

Çizelge 5-16Bot saldırı türü için belirlenen özellikler ve açıklamalar

ÖZELLİK	AÇIKLAMA
Flow Byts/s	Saniyede akan byte sayısı
Flow Pkts/s	Saniyede akan paket sayısı
Pkt Len Var	Bir akışın uzunluk varyansı
Init Fwd Win Byts	İleri yönde ilk pencere içinde gönderilen bayt sayısı

Tehdidi (Bot) tanımlayan veriler çizelge 5-17'de normal durumu tanımlayan veriler ise çizelge 5-18'de gösterilmiştir.

Çizelge 5-17 4.veri setindeki tehdide ait özelliklerin özeti

Özellik	MEAN	STD	MIN	25%	50%	75%	MAX
Flow Byts/s	20369	20976	0	0	0	41674	474576
Flow Pkts/s	2237	1644	0	642	1007	3968	12712
Pkt Len Var	6627	6749	0	0	0	13319	156734
Init Fwd Win Byts	5024	3190	-1	2052	2053	8192	8192

Çizelge 5-18 4.veri setindeki tehdit olmayan duruma ait özelliklerin özeti

Özellik	MEAN	STD	MIN	25%	50%	75%	MAX
Flow Byts/s	264542	3523263	0	14	1195	33407	357000000
Flow Pkts/s	27106	205637	0	5	15	3559	3000000
Pkt Len Var	48150	81756	0	77	2465	96989	710533
Init Fwd Win Byts	7380	14150	-1	-1	8192	8192	65535

Bu veri setine MÖA uygulanması sonucu elde edilen sonuçlar ise çizelge 5-19'da verilmiştir.

Çizelge 5-19 4. veri setine ait deney sonuçları

MÖA	DO (%)	ES (sec.)	TS(sec.)
YSA	93,23	107	14
LR	75,007	2	1
NB	83,826	1	1
KNN	99,914	127	12
KA	99,908	2	1

RO	99,914	5	2
DVM	73,738	2.054	1.955

5.2.5. Brute-Force tehdidinin tespiti için seçilen özellikler

5.veri seti tehdidin BruteForce olup olmadığını kontrol etmek için oluşturulmuştur. Bu veri seti, zararsız durum için Benign veri dosyasından ve anormal durum için BruteForce-Web, BruteForce-XSS, FTP-BruteForce, SSH Bruteforce veri dosyalarından alınarak oluşturulmuştur. Denemeler sonucunda 79 özellikten sadece 6 özellik ile en iyi sonuçların elde edildiği görülmüştür. BruteForce saldırı türü için belirlenen özellikler ve açıklamaları çizelge 5-20'de verilmiştir.

Çizelge 5-20 BruteForce saldırı türü için belirlenen özellikler ve açıklamalar

ÖZELLİK	AÇIKLAMA
Flow Duration	Mikro-saniyedeki akış süresi
Bwd IAT Max	Geri yönde gönderilen iki paket arasındaki maksimum zaman
Flow Byts/s	Saniyede akan byte sayısı
Flow Pkts/s	Saniyede akan paket sayısı
Init Fwd Win Byts	İleri yönde ilk pencere içinde gönderilen bayt sayısı
Pkt Len Var	Bir akışın uzunluk varyansı

Tehdidi (Brute-Force) tanımlayan veriler çizelge 5-21'de normal durumu tanımlayan veriler ise çizelge 5-22'de gösterilmiştir.

Çizelge 5-21 5.veri setindeki tehdide ait özelliklerin özeti

Özellik	MEAN	STD	MIN	25%	50%	75%	MAX
Flow Duration	164078	2335536	1	2	6	31	112641300
Bwd IAT Max	3122	6107	0	0	0	0	1037954
Flow Byts/	680165	696463	0	64516	333333	1000000	2000000

s							
Flow Pkts/s	36427	97388	0	0	0	0	5005224
Init Fwd Win Byts	10505	21232	0	0	0	0	1652619
Pkt Len Var	20292	11488	-1	26883	26883	26883	26883

Çizelge 5-22 5.veri setindeki tehdit olmayan duruma ait özelliklerin özeti

Özellik	**MEAN**	**STD**	**MIN**	**25%**	**50%**	**75%**	**MAX**
Flow Duration	15981720	34855380	1	765	279228	3991629	120000000
Bwd IAT Max	267073	3686430	0	7	1164	41040	394000000
Flow Byts/s	26860	203982	0	5	18	4149	3000000
Flow Pkts/s	2645240	9688727	0	0	0	953177	120000000
Init Fwd Win Byts	47408	83093	0	0	2241	93075	710533
Pkt Len Var	7469	14415	-1	-1	8192	8192	65535

Bu veri setine MÖA uygulanması sonucu elde edilen sonuçlar ise çizelge 5-23'te verilmiştir.

Çizelge 5-23 5. veri setine ait deney sonuçları

MÖA	**DO (%)**	**ES (sec.)**	**TS (sec.)**
YSA	99,26	123	23
LR	97,862	4	1
NB	84,966	1	1
KNN	99,939	285	192
KA	99,991	3	1
RO	99,790	7	1
DVM	97,801	7.156	2.387

5.2.6. Infilteration tehdidinin tespiti için seçilen özellikler

6.veri seti tehdidin Infilteration olup olmadığını kontrol etmek için oluşturulmuştur. Bu veri seti, zararsız durum için Benign veri dosyasından ve anormal durum için Infilteration veri dosyalarından alınarak oluşturulmuştur. Denemeler sonucunda 79 özellikten sadece 16 özellik ile en iyi sonuçların elde edildiği görülmüştür. Infilteration saldırı türü için belirlenen özellikler ve açıklamaları çizelge 5-24'te verilmiştir.

Çizelge 5-24 Infilteration saldırı türü için belirlenen özellikler ve açıklamalar

ÖZELLİK	AÇIKLAMA
Fwd Pkt Len Mean	İleri yönde paketlerin ortalama uzunluğu
Flow Byts/s	Saniyede akan byte sayısı
Fwd Pkt Len Std	İleri yönde paket uzunluklarının standart sapması
Bwd Pkt Len Max	Geri yönde paketlerin maksimum uzunluğu
Fwd Pkt Len Min	İleri yönde paketlerin minimum uzunluğu
Bwd Pkt Len Min	Geri yönde paketlerin minimum uzunluğu
Flow Pkts/s	Saniyede akan paket sayısı
Fwd Pkts/s	Saniyedeki ileri yön paket sayısı
Bwd Pkts/s	Saniyedeki geri yön paket sayısı
Pkt Len Min	Bir akışın minimum uzunluğu
FIN Flag Cnt	FIN içeren paket sayısı
ACK Flag Cnt	ACK içeren paket sayısı
URG Flag Cnt	URG içeren paket sayısı
Down/Up Ratio	İndirme ve yükleme oranı
Init Bwd Win Byts	Geri yönde ilk pencere içinde gönderilen bayt sayısı
CWE Flag Count	CWE içeren paket sayısı

Tehdidi (Infilteration) tanımlayan veriler çizelge 5-25'te normal durumu tanımlayan veriler ise çizelge 5-26'da gösterilmiştir.

Çizelge 5-25 6.veri setindeki tehdide ait özelliklerin özeti

Özellik	MEAN	STD	MIN	25%	50%	75%	MAX
Fwd Pkt Len Mean	14	27	0	0	0	34	1416
Flow Byts/s	39	53	0	0	32	47	1456
Fwd Pkt Len Std	43	86	0	0	0	27	1329
Bwd Pkt Len Max	296	507	0	0	59	157	1880
Fwd Pkt Len Min	33	54	0	0	0	61	549
Bwd Pkt Len Min	271152	3328746	0	0	784	80501	274000000
Flow Pkts/s	106931	348617	0	9	1708	8197	3000000
Fwd Pkts/s	65697	250183	0	6	877	4219	3000000
Bwd Pkts/s	41235	143698	0	0	59	3425	2000000
Pkt Len Min	14	25	0	0	0	34	788
FIN Flag Cnt	0	0	0	0	0	0	1
ACK Flag Cnt	0	0	0	0	0	0	1
URG Flag Cnt	0	0	0	0	0	0	1
Down/Up Ratio	0	0	0	0	0	0	1
Init Bwd Win Byts	1	1	0	0	1	1	11
CWE Flag Count	6154	17849	-1	-1	0	172	65535

Çizelge 5-26 6.veri setindeki tehdit olmayan duruma ait özelliklerin özeti

Özellik	MEAN	STD	MIN	25%	50%	75%	MAX
Fwd Pkt Len Mean	10	25	0	0	0	0	1460
Flow Byts/s	65	63	0	12	44	128	1460
Fwd Pkt Len Std	90	105	0	0	23	213	1014
Bwd Pkt Len Max	527	567	0	0	137	1173	1460
Fwd Pkt Len Min	23	49	0	0	0	0	1232
Bwd Pkt Len Min	230943	3263671	0	16	1142	7902	282000000
Flow Pkts/s	24348	191606	0	5	10	2105	3000000
Fwd Pkts/s	21453	182248	0	3	6	1078	3000000
Bwd Pkts/s	2895	39641	0	0	3	54	2000000
Pkt Len Min	10	22	0	0	0	0	1232

FIN Flag Cnt	0	0	0	0	0	0	1
ACK Flag Cnt	0	0	0	0	0	1	1
URG Flag Cnt	0	0	0	0	0	0	1
Down/Up Ratio	0	0	0	0	0	1	1
Init Bwd Win Byts	0	0	0	0	0	0	1
CWE Flag Count	0	0	0	0	0	0	0

Bu veri setine MÖA uygulanması sonucu elde edilen sonuçlar ise çizelge 5-27'te verilmiştir.

Çizelge 5-27 6. veri setine ait deney sonuçları

MÖA	**DO (%)**	**ES (sec.)**	**TS (sec.)**
YSA	99,02	142	12
LR	91,340	4	1
NB	90,876	1	1
KNN	98,245	239	239
KA	99,094	2	1
RO	99,119	4	1
DVM	87,360	13.156	3.156

6. SONUÇ ve ÖNERİLER

Kanada Siber Güvenlik Enstitüsü tarafından paylaşılan veri seti, daha önceki çalışmalarda kullanılan veri setlerindeki eksiklikler göz önünde bulundurularak oluşturulmuştur. CSE-CIC-IDS2018 veri seti kullanılarak 6 adet veri seti oluşturulmuştur. Bu veri setleri (V1,V2,V3,V4,V5,V6) oluşturulurken zaman maliyetinin azaltılması ve doğruluk oranının daha yüksek olabilmesi için boyut küçültme için Brute Force yöntemi kullanılmıştır.

Bu çalışmada oluşturulan her bir veri setine 7 farklı MÖA uygulanmıştır. Yapılan deney sonuçlarında elde edilen doğruluk oranları ve doğruluk oranlarının ortalamaları Çizelge 6-1'de, eğitim süreleri ve eğitim sürelerinin ortalamaları Çizelge 6-2'de, test süreleri ve test sürelerinin ortalamaları Çizelge 6-3'te verilmiştir. Çizelgelar incelendiğinde, yapılan sınamalar sonucunda eğitim zamanı, test zamanı ve doğruluk oranları göz önüne alındığında herhangi bir tehdidin tespitinde en iyi sonucu KA (99,97) modelinin verdiği gözlemlenmiştir. Bununla birlikte saldırı tespit türünü belirleme başarısı olarak da, ortalamada Karar Ağaçı modelinin daha başarılı olduğu (%99,81) görülmüştür.

Çizelge 6-1 Test sonuçlarında hesaplanan doğruluk oranları

MÖA	**V1**	**V2**	**V3**	**V4**	**V5**	**V6**	**Ortalama DO (%)**
YSA	99,11	99,31	92,26	93,23	99,26	99,02	97,08
LR	93,79	97,11	67,02	75,01	97,86	91,34	87,02
NB	88,07	97,11	63,21	83,83	84,97	90,88	84,68
KNN	99,81	99,88	99,90	99,91	99,94	98,25	99,61
KA	99,97	99,88	99,99	99,91	99,99	99,09	99,81
R0	99,97	99,88	99,99	99,91	99,79	99,12	99,78
DVM	94,98	97,14	65,44	73,74	97,80	87,36	86,08

Çizelge 6-2 Test sonuçlarında hesaplanan eğitim süreleri

MÖA	**V1**	**V2**	**V3**	**V4**	**V5**	**V6**	**Ortalama ES (sec.)**
YSA	457	98	552	107	123	142	247
LR	119	4	7	2	4	4	23
NB	10	2	2	1	1	1	3

KNN	10.209	1.540	983	127	285	239	2231
KA	33	2	5	2	3	2	8
RO	84	6	16	5	7	4	20
DVM	21.351	2.468	5.351	2.054	7.156	13.156	8589

Çizelge 6-3 Test sonuçlarında hesaplanan test süreleri

MÖA	**V1**	**V2**	**V3**	**V4**	**V5**	**V6**	**Ortalama TS (sec.)**
YSA	34	11	66	14	23	12	27
LR	1	1	1	1	1	1	1
NB	2	1	1	1	1	1	1
KNN	1.735	844	96	12	192	239	520
KA	1	1	1	1	1	1	1
R0	2	1	1	2	1	1	1
DVM	5.182	1.855	2.051	1.955	2.387	3.156	2.764

Makine Öğrenmesi kullanılarak oluşturulan modellerde en önemli parçalardan bir tanesi veridir. Verilerin doğru, anlamlı ve çeşitli olması yapılan çalışmadaki başarıyı çok fazla etkilemektedir. Bu çalışmada kullanılan veri setleri ile oluşturulan modellerin ilerleyen çalışmalarda farklı veri setleri ile test edilerek modellerin daha fazla geçerli kılınması sağlanacaktır. Bununla birlikte bu veri seti içerisinde yer alan diğer tehdit türleri içinde modeller oluşturulacak ve Derin Öğrenme yöntemleri (örneğin CNN algoritması), kullanılarak daha farklı sonuçların elde edilmesi sağlanacaktır.

KAYNAKLAR

Abrams, L. 2013. Cryptolocker Ransomware İnformation Guide and FAQ.

Afrin, F., Nahar, I. 2015. Incremental Learning Based İntelligent Job Search System. BRAC University.

Amor, N. B., Benferhat, S., Elouedi, Z. 2004. Naive Bayes vs Decision Trees in Intrusion Detection Systems. Proceedings of The 2004 ACM Symposium on Applied Computing (s. 420-424). Nicosia, SAC.

Anderson, J. P. 1980. Computer Security Threat Monitoring and Surveillance. Washington, PA.

Aydın, M. A. 2005. Bilgisayar Ağlarında Saldırı Tespiti İçin İstatistiksel Yöntem Kullanılması. İstanbul, İstanbul Teknik Üniversitesi.

Bace, R. G. 2000. Intrusion Detection. İndianapolis, Macmillan Technical Publishing.

Bace, R., Mell, P. 2001. NIST Special Publication on Intrusion Detection Systems. Greensboro, Booz Allen, Hamilton.

Bapat, R., Mandya, A., Liu, X., Abraham, B., Brown, D. E., Kang, H., Veeraraghavan, M. 2018. Identifying Malicious Botnet Traffic Using Logistic Regression. In 2018 Systems and Information Engineering Design Symposium (s. 266-271). IEEE.

Bircan, H. 2004. Lojistik Regresyon Analizi: Tıp Verileri Üzerine Bir Uygulama . Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 2, 185-208.

Breiman, L. 1999. Random Forests-Random Features. Berkeley, Dept, of Statistics, Univ. of California.

Cahyo, A. N., Hidayat, R., Adhipta, D. 2016. Performance Comparison of Intrusion Detection System Based Anomaly Detection Using Artificial Neural Network and Support Vector Machine. AIP Conference Proceedings 1755 (s. 070011). AIP.

Canberk, G., Sağıroğlu, Ş. 2017. Bilgisayar Sistemlerine Yapılan Saldırılar ve Türleri. Erciyes Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 1-12.

Caruana, R., Niculescu-Mizil, A. 2006. An Empirical Comparison of Supervised Learning Algorithms. in Proceedings of the 23rd International Conference on Machine Learning, (s. 161-168).

Clarke, R. A., Knake, R. K. 2015. Cyber War: the Next Threat to National Security and What to Do About It? Strategic Analysis, 39(4), 458-460.

Cortes, C., Vapnik, V. 1995. Support-Vector Networks. Kluwer Academic Publishers, 20(3) 273-297.

CyberSecurity, Erişim Tarihi: 03.03.2020, https://www.unb.ca/cic/datasets/ids-2018.html

Çavuşoğlu, Ü., Kaçar, S. 2019. Anormal Trafik Tespiti için Veri Madenciliği Algoritmalarının Performans Analizi . dergipark , 205-216.

Çetin, H. 2014. Kişisel Veri Güvenliği ve Kullanıcıların Farkındalık Düzeylerinin İncelenmesi . Akdeniz İ.İ.B.F. Dergisi, 86-105.

Dayıoğlu, B., Özgit, A. 2001. İnternet'de Saldırı Tespiti Teknolojileri. İletişim Teknolojileri 1. Ulusal Sempozyumu ve Fuarı. Ankara.

Ergezer, H., Dikmen, M., Özdemir, E. 2003. Yapay Sinir Ağları ve Tanıma Sistemleri. PiVOLKA, 2(6), 14-17.

Gezgin, M. D., Buluş, E. 2013. Kablosuz Ağlar için Bir DoS Saldırısı Tasarımı. Bilişim Teknolojileri Dergisi, 6(3), 17-23.

Gündüz, M. Z., Daş, R. 2016. Sosyal Mühendislik: Yaygın Ataklar ve Güvenlik. PD Ş. Sağıroğlu, PDM Alkan, PDE Akyıldız ve DDS Akleylek, 9.

Güven, E. N. 2007. Zeki Saldırı Tespit Sistemlerinin İncelenmesi, Tasarımı ve Gerçekleştirilmesi. Ankara, Gazi Üniversitesi Fen Bilimleri Enstitüsü.

Haykin, S. 1994. Neural Networks, A Comprehensive Foundation. Upper Saddle River, Prentice Hall .

Heba, F. E., Darwish, A., Hassanien, A. E., Abraham, A. 2010. Principle Components Analysis and Support Vector Machine Based Intrusion Detection System. In 2010 10th International Conference on Intelligent Systems Design and Applications (s. 363-367). IEEE.

Ilgaz, B. 2018. Küçük ve Orta Büyüklükteki İşletmeler İçin Veri. Konya, KTO Karatay Üniversitesi Fen Bilimleri Enstitüsü.

Imperva, Erişim Tarihi: 03.03.2020, https://www.imperva.com/learn/application-security/high-orbit-ion-cannon

Kara, M. 2013. Siber Saldırılar - Siber Savaşlar ve Etkiler, İstanbul Bilgi Üniversitesi, Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi, İstanbul.

Kara, M., Şişeci, N. E. 2011. Botnetlerle Mücadelede Dünyadaki ve Türkiye'deki Durum. Akademik Bilişim'11 - XIII. Akademik Bilişim Konferansı Bildirileri (s. 495-501). Malatya, İnönü Üniversitesi.

Kavzoğlu, T., Çölkesen, İ., Şahin, E. K. 2012. Heyelan Duyarlılık Haritasının Üretilmesinde Kullanılan Faktörlerin Etkilerinin Araştırılması: Düzköy Örneği . IV. Uzaktan Algılama ve Coğrafi Bilgi Sistemleri Sempozyumu (UZAL-CBS 2012), Zonguldak.

Kirda, E., Kruegel, C. 2006. Protecting Users Against Phishing Attacks. The Computer Journal, 554-561.

Lavanya, D., Rani, K. U. 2011. Performance Evaluation of Decision Tree Classifiers on Medical Datasets. International Journal of Computer Applications, 1-4.

Lewis, D. D. 1992. Representation and Learning in Information Retrieval. Massachusetts: University of Massachusetts.

Lin, S.-C., Tsen, S.-S. 2004. Constructing Detection Knowledge for Ddos İntrusion Tolerance. Expert Systems with Applications, 379-390.

Manzoor, I., Kumar, N. 2017. A Feature Reduced Intrusion Detection System Using ANN Classifier. ELSEVİER.

Marttin, V. 2014. Saldırı Tespit Sistemlerinde Yapay Sinir Ağlarının Kullanımı ve Başarımlarının İncelenmesi. Bilecik, Bilecik Şeyh Edebali Üniversitesi Fen Bilimleri Enstitüsü.

Mok, M. S., Sohn, S. Y., Ju, Y. H. 2010. Random Effects Logistic Regression Model for Anomaly Detection. Expert Systems with Applications, 7162-7166.

Mukherjeea, D. S., Sharma, N. 2012. Intrusion Detection Using Naive Bayes Classifier with Feature. Procedia Technology, 120-128.

Mukkamala, S., Janoski, G., Sung, A. 2002. Intrusion Detection Using Neural Networks and Support Vector Machines. In Proceedings of the 2002 International Joint Conference on Neural Networks (2, 1702-1707)). IEEE.

Özdamar, K. 2004. Paket Programlar ile İstatistiksel Veri Analizi (Çok Değişkenli Analizler). Eskişehir, Kaan Kitabevi.

Özekes, S., Karakoç, E. N. 2019. Makine Öğrenmesi Yöntemleriyle Anormal Ağ Trafiğinin Tespit Edilmesi. dergipark, 566-576.

Özgür, A., Erdem, H. 2012. Saldırı Tespit Sistemlerinde Kullanılan Kolay Erişilen Makine Öğrenme Algoritmalarının Karşılaştırılması. Bilişim Teknolojileri Dergisi,, 5.2, 41-48.

Öztemel, E. 2012. Yapay Sinir Ağları. Papatya.

Poojitha, G., Kumar, K. N., Reddy, P. J. 2010. Intrusion Detection using Artificial Neural Network. Second International conference on Computing,Communication and Networking Technologies (s. 1-7). Karur: IEEE.

Protić, D. D. 2018. Rewiew oF KDD Cup '99, NSL-KDD AND KYOTO 2006+ Datasets. Vojnotehnički glasnik, 580-596.

Sahu, S., Mehtre, B. M. 2015. Network Intrusion Detection System Using J48 Decision Tree. In 2015 International Conference on Advances in Computing, Communications and Informatics (s. 2023-2026). IEEE.

Sharafaldin, I., Arash, H. L., Ali, A. 2018. Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. 4th International Conference on Information Systems Security and Privacy (ICISSP) (s. 108-116). Portekiz, ICISSP.

Shearer, C. 2000. The CRISP-DM Model: The New Blueprint for Data Mining. Journal of Data Warehousing, 5.4, 13-22.

Smaha, S. E. 1988. Haystack: An a Intrusion Detection System. Fourth Aerospace Computer Security Applications Conference., (s. 37-44). Austin.

Snort-org, Erişim Tarihi: 20.05.2020. https://www.snort.org/production/document_files/files. 5 20, 2020

Tanrıkulu, H., Sazlı, M. 2007. Saldırı Tespit Sistemlerinde Yapay Sinir Ağlarının Kullanılması. Pitonca.

Teresa F. Lunt, A. T. 1992. A Real-Tıme Intrusıon-Detectıon Expert System (IDES). Washington: SRI.

Vikipedi, Erişim Tarihi: 03.03.2014, https://tr.wikipedia.org/wiki/Google

Vikipedi, Erişim Tarihi: 02.05.2020, https://tr.wikipedia.org/wiki/Botnet

Wang, J., Yang, Q., Ren, D. 2009. An Intrusion Detection Algorithm Based on Decision Tree Technology. In 2009 Asia-Pacific Conference on Information Processing (s. Vol. 2, pp. 333-335). IEEE.

Yelken, A. B. 2019, 6 25. siber-tehdit-istihbarati. 5 5, 2020 tarihinde www.cybermagonline.com: https://www.cybermagonline.com/siber-tehdit-istihbarati adresinden alındı

Yousef, M., Nebozhyn, M., Shatkay, H., Kanterakis, S., Showe, L. C., Showe, M. K. 2006. Combining Multi-Species Genomic Data for Microrna İdentification Using A Naive Bayes Classifier. Bioinformatics , 22(11), 1325-1334.

ÖZGEÇMİŞ

Adı Soyadı : Mehmet Salih KARAMAN

Doğum Yeri ve Yılı : Diyarbakır, 01/09/1991

Medeni Hali : Bekar

Yabancı Dili : İngilizce

E-posta : salih.krmn21@outlook.com

Eğitim Durumu

Lise : Özel Amid Lisesi, 2009

Lisans : Dicle Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği

Yüksek Lisans : İstanbul Ticaret Üniversitesi,
Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği

Mesleki Deneyim

Bilgi Otomasyon	2017-2017
Araştırma Merkezi Komutanlığı	2017-…(devam ediyor)

Yayınları

Karaman, S., Turan, M., Aydın, M. A. 2020. Yapay Sinir Ağı Kullanılarak Anomali Tabanlı Saldırı Tespit Modeli Uygulaması, Avrupa Bilim ve Teknoloji Dergisi, Özel Sayı, 17-25