Journal of Computer Science
https://dergipark.org.tr/en/pub/bbd
https://doi.org/10.53070/bbd.990934

Journal of Computer Science
ISSN: 2548-1304
Volume: IDAP-2021, Issue: Special, pp: 162-174
@Anatolian Science

Kablosuz Algılayıcı Ağlarda Hibrit Saldırı Tespit Sistemi Geliştirme

Hybrid Intrusion Detection System Development in Wireless Sensor Networks

Hamza ELBAHADIR[1], Ebubekir ERDEM[2]

[1] Bilgisayar Mühendisliği Bölümü, Fırat Üniversitesi, Elazığ, Türkiye

[2] Bilgisayar Mühendisliği Bölümü, Fırat Üniversitesi, Elazığ, Türkiye

(h.elbahadir@yandex.com.tr, aberdem@firat.edu.tr)

Received: Sep.3, 2021 Accepted: Oct.5, 2021 Published: Oct.20, 2021

Özetçe— Günümüzde yaygın kullanım alanlarına sahip olan kablosuz algılayıcı ağlar (KAA), geleneksel ağ mimarisinden farklı olduğundan, özgün güvenlik çözümleri üretilmelidir. Bu çalışmada KAA güvenliği için saldırı tespit sistemi (STS) önerilmiştir. Etkili bir güvenlik için, saldırı tespit sistemlerinde kullanılan anomali ve yanlış kullanım tabanlı algılama metotlarını ihtiva eden hibrit bir model üzerinde çalışılmıştır. Sistemin normal ve saldırı trafiğini sınıflandırabilmesi için veri madenciliği algoritmalarından BayesNet, J48, JRip, PART ve RandomForest algoritmaları kullanılmış ve söz konusu algoritmaların performans değerleri paylaşılmıştır. Bu çalışmada literatürdeki mevcut çalışmalardan farklı olarak, güncel bir veri seti olan CSE-CIC-IDS2018 kullanılmıştır. Veri setindeki veriler ise, KAA performans kriterleri göz önünde bulundurularak ön işleme tabi tutulmuştur. Sonuçlar önerilen sistemin yüksek doğruluk oranına sahip olduğunu göstermiştir.

Anahtar Kelimeler : *Kablosuz algılayıcı ağ, saldırı tespit, veri madenciliği, hibrit*

Abstract— Since wireless sensor networks (WSNs), which have widespread usage areas today, are different from traditional network architecture, security solutions specific to WSNs should be produced. In this study, an intrusion detection system (IDS) is proposed for WSN security. For an effective security, a hybrid model including anomaly and misuse-based detection methods used in intrusion detection system has been studied. Data mining algorithms BayesNet, J48, JRip, PART and RandomForest were used to classify the normal and attack traffic of the system and the performance values of the algorithms are shared. In this study, CSE-CIC-IDS2018, an up-to-date data set, was used, unlike the existing studies in the literature. Considering the WSN performance criteria, the data in the data set was preprocessed. The results showed that the proposed system has high accuracy.

Keywords : *Wireless sensor networks, intrusion detection, anomaly, hybrid*

1. GİRİŞ

Fiziksel ortamlardaki sıcaklık, basınç, titreşim, nem, ses, manyetik kuvvet gibi çeşitli verileri algılamak, sınıflandırmak ve değerlendirmek amacıyla kendi kendine organize olabilen algılayıcı düğümlerden oluşan Kablosuz Algılayıcı Ağlar (KAA), günümüzde çok çeşitli uygulama alanlarına sahiptir. KAA'ların uygulama alanları arasında askeri ve istihbari faaliyetler, sınır hatlarının izlenmesi, sağlık uygulamaları gibi verilerin gizli ve güvenli aktarımının hayati öneme sahip olduğu alanlar da bulunmaktadır. Bununla birlikte, KAA'lar kaynak kısıtları, iletişim ortamı ve altyapısı ve algılayıcıların yerleştirildikleri savunmasız bölgeler gibi faktörlerden dolayı birçok saldırıya açıktırlar. Dahası gerek

geleneksel ağlardan farklı altyapı özelliklerine sahip olması gerekse de fiziksel kaynak kısıtlarından dolayı, bu ağlara özgü güvenlik çözümlerinin geliştirilmesi gerekir.

Katmanlı mimariye sahip KAA'ların her katmanına yönelik çeşitli saldırılar gerçekleştirilebilmektedir. En fazla saldırıya uğrayan katman ise ağ ve yönlendirme katmanı olup; sahte/seçici yönlendirme, sybil, hello seli, sinkhole, wormhole ve DOS yaygın saldırılar arasındadır. Literatürde kablosuz algılayıcı ağlara yönelik gerçekleştirilen saldırılar ve birtakım güvenlik önlemleri önerilmiştir (Padmavathi ve arkadaşları, 2009; Kumar ve arkadaşları, 2014; Chelli, 2015; Deng ve arkadaşları, 2017).

Tomić ve arkadaşları, kablosuz algılayıcı ağlarda yaygın olarak kullanılan iletişim protokollerinin ve standartlarının güvenliği hakkında bir analiz gerçekleştirmişlerdir. Yazarlar, kablosuz algılayıcı ağların karakteristiklerini, güvenlik gereksinimlerini, güvenlik açıklarını ve gerçekleştirilebilecek saldırıları da ele alarak söz konusu saldırıların değerlendirmesini yapmışlardır (Tomić ve arkadaşları, 2017). Ozcelik ve arkadaşları, kümelenmiş KAA'lar için hibrit bir STS önermişlerdir. Yazarlar, saldırıları yalnızca düğüm düzeyinde tespit etmek yerine, tüm ağ bileşenleri arasındaki karşılıklı güven değerlendirmesini kullanarak işbirliğe dayalı merkezi bir tasarım önermişlerdir. Burada her bir algılayıcı düğüm, komşularının işlevsel itibar değerlerini gözlemleyerek hesaplar. Metodolojileriyle ilgili temel sorun, yalnızca enerji tüketim sonuçlarını ifade etmiş olmaları ve tespit edilebilir saldırı türleri ve tespit oranları hakkında herhangi bir tartışma sunmamış olmalarıdır (Ozcelik ve arkadaşları, 2017). Acharya ve arkadaşları, akıllı su damlaları algoritmasına dayalı ve öznitelik seçim yönteminin önerildiği bir STS önermişlerdir (Acharya ve arkadaşları, 2018). Ghugar ve arkadaşları, kablosuz algılayıcı ağın farklı katmanlarına yönelik güven ölçümlerini hesaplayarak, düğümlerin güvenilir veya kötü niyetli olduğunu tespit eden bir sistem önermişlerdir (Ghugar ve arkadaşları, 2019). Çavuşoğlu ve arkadaşları, zararlı trafiğin tespiti için kullanılan veri madenciliği algoritmalarının performans analizini yapmışlardır (Çavuşoğlu ve arkadaşları, 2019).

Bu çalışmada, KAA'lara yönelik gerçekleştirilen saldırılara karşı, etkili bir saldırı tespit sistemi modellenmiştir. Bununla birlikte, saldırı tespiti için literatürde önerilen imza, yanlış kullanım ve anomali tabanlı yöntemler, günümüzde tek başına yeterli güvenliği garanti altına alamamaktadır. Bu nedenle saldırı tespiti için önerilen tüm algılama yöntemlerini ihtiva eden hibrit bir sistem modellenmiştir. STS'de metotların hibrit olarak kullanılması, hesaplama karmaşıklığını ve kaynak tüketimini arttıracağından; hem kaynakların daha verimli kullanılması hem de etkili bir güvenlik için, STS'nin modellenmesinde veri madenciliği yöntemleri kullanılmıştır.

Sistemin ilk savunma hattında anomali tabanlı algılama metodu kullanılmış olup; bu aşamada normalden sapan trafik, saldırı olarak nitelenmiştir. Savunmanın ikinci hattında ise yanlış kullanım tabanlı algılama metodu uygulanmıştır. Bu aşamada ise savunma hattının ilk adımından geçen paketler, normal veya saldırı olarak sınıflandırılmaktadır. Bu modele göre sınıflandırma yapabilmek için güncel bir veri seti olan CSE-CIC-IDS2018 kullanılmıştır. Veri setinde büyük miktarlarda veriler bulunduğundan ve bu miktardaki verilerin KAA'da işlenmesi mümkün olmadığından, veriler ön işleme alınmıştır. Akabinde BayesNet, J48, JRip, PART ve RandomForest algoritmaları ile sınıflandırma işlemi yapılmıştır. Son olarak, algoritmaların performans analizi yapılmış ve sonuçlar sunulmuştur.

Çalışmanın geri kalanı şu şekilde organize edilmiştir: 2. bölümde kablosuz algılayıcı ağların genel karakteristikleri ele alınmış, 3. bölümde kablosuz algılayıcı ağlara yönelik saldırılar incelenmiş, 4. bölümde STS'lerin çalışma mekanizmalarına değinilmiş, 5. bölümde önerilen STS modellenmiş ve 6. bölümde çeşitli kriterlere göre algoritmaların performans analizi yapılmış ve benzetim sonuçları sunulmuştur. Önerilen modelin genel değerlendirmesi ve literatürdeki mevcut çalışmalardan farklılık ve üstünlüklerine ise son bölümde değinilmiştir.

2. KABLOSUZ ALGILAYICI AĞLAR

Kablosuz algılayıcı ağlar geleneksel ağlardan farklı mimarilere sahiptir. KAA'ların karakteristikleri, geleneksel güvenlik çözümlerinin bu ağlarda uygulanabilirliğini geçersiz kılmaktadır. Bu bölümde KAA'ların söz konusu karakteristiklerine değinilmiştir.

2.1. Kısıtlı Kaynak

Algılayıcı düğümlerin kısıtlamaları; düğümün üretim maliyeti, enerji tüketimi ve işlem kapasitesi şeklindedir.

KAA'lar mekânsal olarak çok geniş alana yayılmış yüzlerce hatta binlerce düğümden oluşabileceğinden, düğümlerin üretim maliyeti büyük önem taşımaktadır. Bu sebeple kurulum ve kaynaklar açısından minimal tercihlerde bulunulmalıdır.

Benzer şekilde enerji verimliliği de KAA ömrü açısından birincil öneme sahiptir. Zira sayısız algılayıcı düğümün pilinin şarj edilmesi veya değiştirilmesi son derece zahmetli, hatta imkânsız bir süreçtir. KAA'ların yaşam süresi ise söz konusu bataryaların ömrüne bağlıdır. Bu sebeple mevcut enerjinin verimli olarak kullanılmasına olanak tanıyan algoritmalar kullanılmalıdır.

KAA'ların donanım kısıtları ise işlem ve iletişim üzerinde doğrudan etkili olan önemli bir faktördür. Tablo 1'de bazı ticari sensör düğümlerinin donanım özelliklerine yer verilmiştir. (Altun, 2016).

Tablo 1. KAA katmanlarına yönelik saldırı tipleri

Mote	CPU	Frekans (MHz)	İletişim	Toplam Aktif Güç (mW)
Mika	ATMega103	4	TR1000, ASK	27
MICA2Dot	ATMega128	7.4	CC1000, FSK	44
MİkA2	ATMega128	7.4	CC1000, FSK	89
Tmote Sky	MSP430	8	CC2420, OQPSK	32
Telos B	MSP430	8	CC2420, OQPSK	32
Imote2	PXA271	13–400	CC2420, OQPSK	86.8

Neticede; söz konusu kaynak kısıtları, KAA'lara özgü güvenlik protokollerinin tanımlanmasını gerekli kılmaktadır.

2.2. Büyük Ölçek

Hassas verileri algılayan ve analiz eden KAA uygulamalarının çoğu, geniş bir coğrafi alana kurulmaktadır. Böylesi alanlara bazen on binlerce düğüm kurulabilmektedir. Algılayıcı düğümlerin sayı olarak fazla olmasının bir diğer sebebi de; düğümlerin kısıtlı radyo kapasiteleri, yüksek ölüm oranları gibi faktörlerdir. Böylece bir düğümün beklenmedik bir şekilde kullanılamaz hale gelmesi, sistemin işlerliğini hayati derecede etkilemez. Öte taraftan; ağdaki trafiğin kesintiye uğramaması için düğümlerin bu şekilde kullanımı, düğümlerin baz istasyonuna gönderdikleri veri miktarını artırmakta, dolayısıyla ağın yaşam süresini azaltmaktadır (Martins ve arkadaşları, 2010)

2.3. Güvenlik

Karakteristiklerinden dolayı geleneksel güvenlik protokolleri KAA'larda uygulanamamaktadır. Bununla birlikte algılayıcı düğümlerin fiziksel olarak ele geçirilmesi, akabinde kriptografik anahtarlar veya düğümdeki diğer hassas verilerin çıkarılabilmesi gibi riskler de mevcuttur. Düğüm ayrıca,

saldırganın kontrol ettiği tehlikeye atılmış bir düğüm oluşturmak için değiştirilebilir. Genellikle ek maliyet nedeniyle KAA'larda sensör düğümlerinin kurcalamaya karşı korumalı olmadığı varsayılır. (Dolay, 2009)

2.4. Dinamik Topoloji

Aktif olarak çalışan bir KAA'da bazı düğümlerin beklenmedik şekilde bataryaları tükenebilir ve iletişimin aksamasına sebebiyet verebilir. Hem yaşamı sona eren düğümün faaliyetini icra etmesi için hem de başka görevler için ağa yeni düğümlerin eklenmesi söz konusu olabilir. Dolayısıyla bir algılayıcı ağ, beklenmedik durumlara karşı kendisini yeniden organize edebilmelidir (Akyildiz ve arkadaşları, 2002)

3. KAA'LARA YÖNELİK SALDIRILAR VE GÜVENLİK MEKANİZMALARI

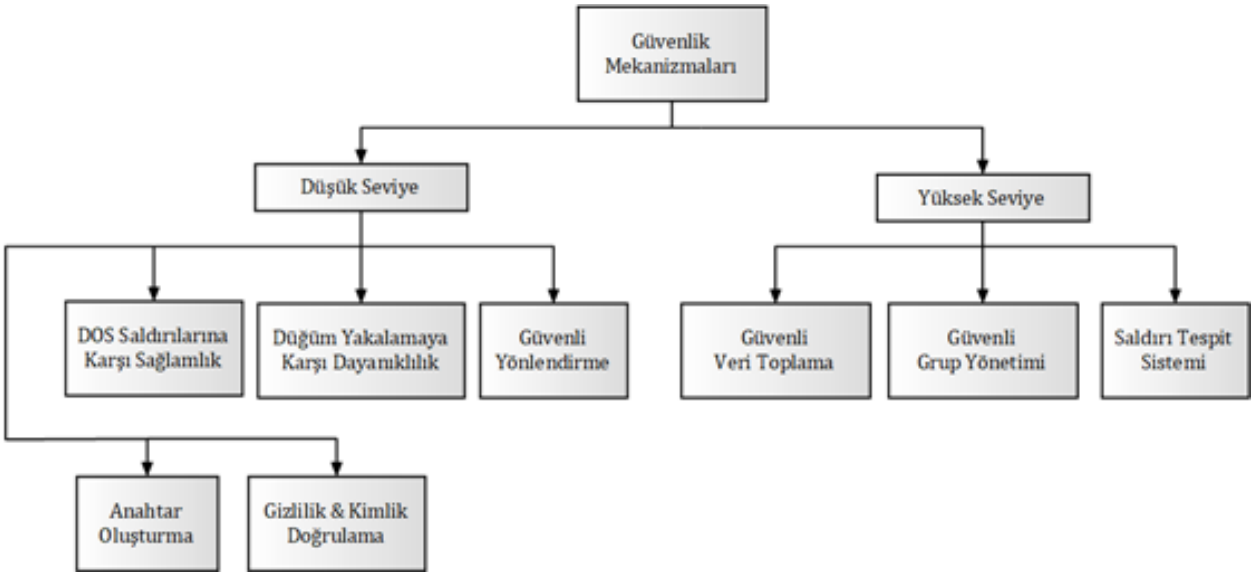
KAA'lar genellikle katmanlara ayrılır ve bu katmanlı mimari, genellikle OSI modelini kullanır. Bir algılayıcı ağın her katmanına yönelik çeşitli saldırılar gerçekleştirilebilmektedir (Amara ve arkadaşları, 2013). Bir algılayıcı ağın katmanlarına yönelik gerçekleştirilebilecek saldırılar Tablo 2'de özetlenmiştir.

Tablo 2. KAA katmanlarına yönelik saldırı tipleri

Ağ Katmanı	Saldırı Tipi
Fiziksel Katman	Sıkışma
	Kurcalama
Bağlantı Katmanı	Çarpışma
	Tükenme
	Adaletsizlik
Ağ ve Yönlendirme Katmanı	Sahte, Değiştirilmiş veya Yeniden Oynatılmış Yönlendirme Bilgileri
	Seçici Yönlendirme
	Sinkhole
	Sybil
	Solucan Delikleri
	Hello Seli Saldırıları
	Onay Sahteciliği
Taşıma Katmanı	Sel Saldırıları
	Desenkronizasyon Saldırıları

Tablo 2'deki saldırı tiplerine karşı KAA'ların güvenliğini sağlamak için birtakım güvenlik mekanizmaları önerilmiştir. Bu mekanizmalar, saldırıları tespit etmek ve önlemek için kullanılır. Kötü niyetli saldırılara karşı koymak için çok çeşitli güvenlik şemaları oluşturulabilir ve bunlar, yüksek seviye ve düşük seviye olarak kategorize edilebilir.

Düşük seviyeli güvenlik mekanizmaları; hizmet reddi saldırılarına karşı sağlamlık, düğüm yakalamaya karşı dayanıklılık, güvenli yönlendirme, anahtar oluşturma, gizlilik ve kimlik doğrulamadır. Yüksek seviyeli güvenlik mekanizmaları ise; güvenli veri toplama, güvenli grup yönetimi ve saldırı tespit sistemi olarak özetlenebilir. Şekil 1'de, güvenlik mekanizmaları gösterilmiştir (Dewal, 2018)



Şekil 1. KAA'larda güvenlik mekanizmaları

4. SALDIRI TESPİT SİSTEMLERİ

KAA'lar için önerilen etkili güvenlik çözümlerinden biri Saldırı Tespit Sistemi'dir (STS). STS, geçici ağlardaki mobil düğümlerin anormal davranışını tespit etmek için geliştirilen güvenlik sistemidir. STS, istemcilere veya sunucuya veya her ikisine birden kurulan ek bir birim olup, bu birime STS ajanı denir. STS ajanı, temelde üç ana bileşenden oluşmaktadır (Alrajeh ve arkadaşları, 2013)

- *İzleme Ünitesi:* Gerek yerel olayları gerekse de komşu düğümlerin faaliyetlerini izleme ve trafiği analiz etme işlemleri bu birim tarafından yürütülür.
- *Analiz ve Tespit Ünitesi:* Ağ trafiğini analiz eden ve paketlerin normal veya saldırı şeklinde sınıflandırılmasını icra eden ana birimdir.
- *Uyarı Ünitesi:* Saldırı tespit edilmesi durumunda, alarm oluşturma yani saldırıya tepki verme görevini gerçekleştiren birimdir.

STS ajanı; ağı, saldırılardan korumada önemli bir görev üstlenir. KAA'lara STS ajanı kurmanın temelde üç farklı yolu bulunmaktadır (Farooqi ve arkadaşları, 2009).

- *Tamamen merkezileştirilmiş:* KAA'larda, algılayıcı düğümler ortamı algılar ve işlenmiş bilgileri havuza veya baz istasyonuna iletir. Tamamen merkezi STS yaklaşımında, STS ajanı havuza veya baz istasyonuna kurulur.
- *Tamamen dağıtılmış:* Bu yaklaşımında, STS ajanı her algılayıcı düğüme kurulur. Yerel olarak komşu düğümlerin anormal davranışını ve telsiz menzilindeki düğümlerden aldığı verileri analiz eder ve anormal faaliyetler için uyarılar oluşturur.
- *Dağıtılmış-merkezileştirilmiş:* Bu yaklaşımda, STS ajanı yalnızca izleme düğümü adı verilen ve diğer düğümlerden daha yetenekli olan küme başı (Cluster-Head, CH) düğümlerine kurulur. Bu düğüm aynı anda iki tür işlevi gerçekleştirir: Normal düğüm faaliyetlerini gerçekleştirme ve saldırı tespitini sağlama. Bu yaklaşımın arkasındaki mantık, tamamen dağıtılmış yaklaşımların karşılaştığı tespit yükünü en aza indirmektir.

Saldırı tespiti için, çeşitli algılama metotları kullanılmaktadır. Bu metotlar; anomali, yanlış kullanım ve imza tabanlı algılama metotlardır (Ozgur ve arkadaşları, 2005).

- *Anomali tabanlı tespit:* Bu yöntemde, önce ağın normal trafiğinin kuralları tanımlanır, akabinde bu kurallara aykırı trafikler saldırı olarak nitelendirilir. Bu yöntemde yüksek algılama oranı bulunmakla birlikte yüksek yanlış alarm oranı da söz konusudur.

- *Yanlış kullanım tabanlı tespit:* Bu yöntem, önceden bilinen saldırılara ait profillerin tanımlanmasına dayanır. Sisteme yönelik saldırılar ise mevcut kalıplar referans alınarak tespit edilir. Bu yöntemde bilinen saldırılar verimli bir şekilde tespit edilebilmesine rağmen, saldırı profili oluşturma gereksinimi ve yeni saldırılara karşı yetersiz olması dezavantajdır.
- *İmza tabanlı tespit:* Bu yöntemde ise trafiğin doğru akışını tanımlayan kurallar ve kısıtlamalar belirlenir ve trafik akışı, söz konusu tanımlara göre izlenir. Bu yöntemde gerekli tanımlamaların manuel olarak ayarlanması ihtiyacı, zaman bakımından dezavantaj olarak değerlendirilebilir.

Bu çalışmada anomali ve yanlış kullanım tabanlı algılama metotları birlikte kullanılarak hibrit bir sistem modellenmiştir.

5. SİSTEMİN GERÇEKLEŞTİRİLMESİ

Saldırı tespit sistemleri, kablosuz algılayıcı ağların güvenliği için kullanılan etkili ve verimli sistemlerdir. Ancak STS'ler için önerilen anomali ve yanlış kullanım tabanlı algılama metotları, günümüz şartlarında, KAA güvenliğini sağlamak için yalnız başlarına kullanımı yeterli değildir. Zira, anomali tabanlı saldırı tespiti, yüksek bir algılama oranına sahip olsa da yüksek yanlış alarm oranı gibi bir dezavantajı vardır. Öte yandan, imza tabanlı tespit yöntemi, saldırıları tespit etmede yüksek doğruluğa ve düşük bir yanlış alarm oranına sahiptir, ancak yeni saldırıların tespiti konusunda yetersizdir. Bu nedenle, söz konusu algılama yöntemlerinin birleştirildiği hibrit bir STS modellenmesi KAA güvenliğini garanti altına alacaktır.

5.1. Anomali Tabanlı Algılama Metodu

Savunma hattının ilk basamağında anomali tabanlı algılama metodu kullanılarak, ağ trafiğinde normalden sapan trafik tespit edilmiştir. Ağın normalden sapan trafiğinin belirlenmesi için çeşitli kurallar belirlenmiştir (Silva ve arkadaşları, 2005).

Aralık Kuralı: Ardışık iki mesaj arasındaki alt ve üst süre limitini belirler. Bu kural ile hizmet reddi saldırıları tespit edilir.

Bütünlük Kuralı: Orijinal mesajın, kaynak ve hedef düğümler arasındaki yol boyunca değişip değişmediğini gösterir. Bu kural ile içeriğin değiştirilmediği garanti altına alınır.

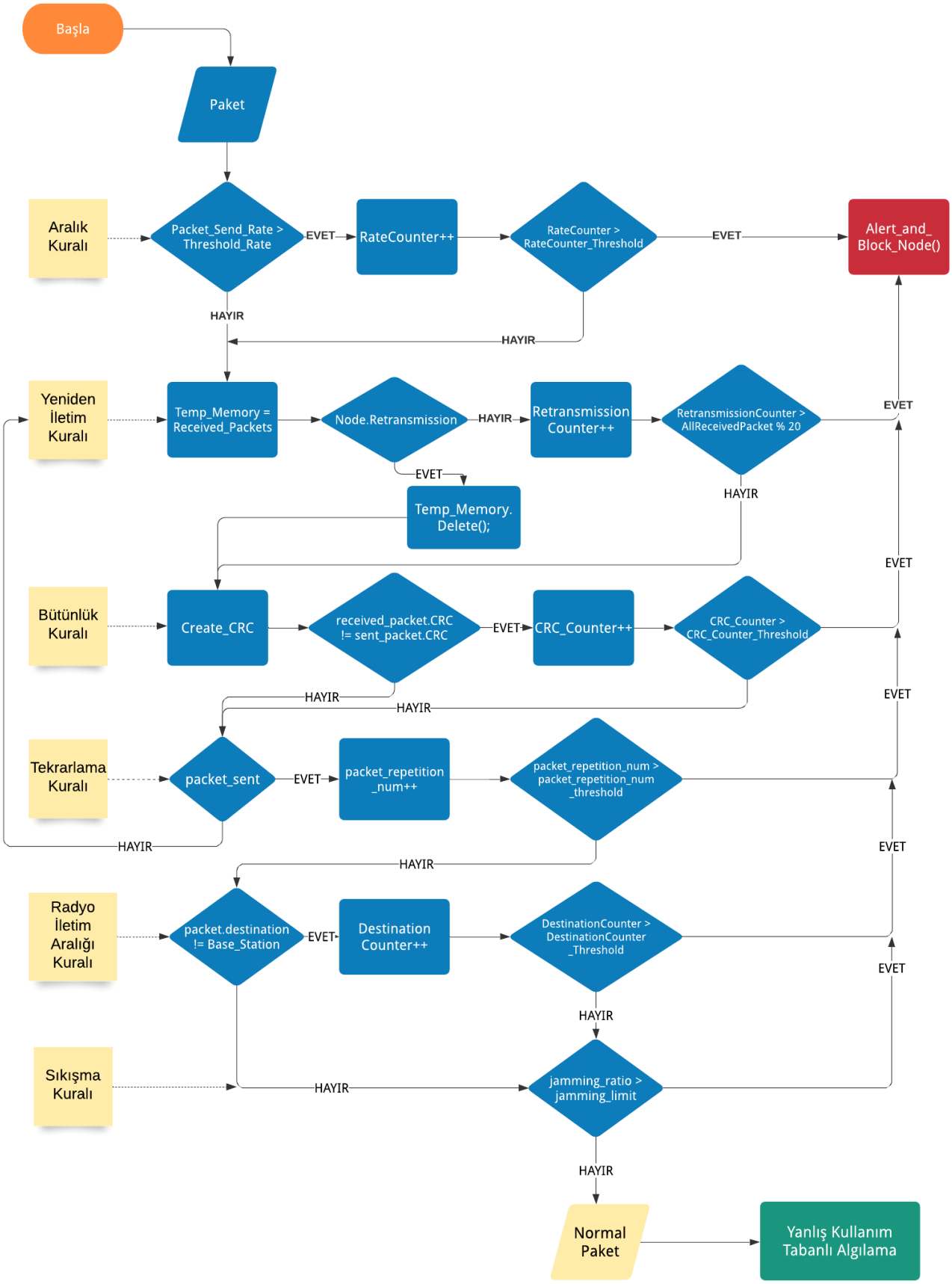
Yeniden İletim Kuralı: İletilmesi gereken bir mesajın, bir düğüm tarafından iletilip iletilmediğini tespit eder. Sinkhole ve seçici yönlendirme saldırılarının tespitinde kullanılır.

Tekrarlama Kuralı: Komşu düğüm tarafından bir mesajın yeniden iletim sayısının, sınırı aşıp aşmadığını belirler. Bu kural ile hizmet reddi saldırıları tespit edilebilir.

Gecikme Kuralı: Belirli bir zaman aşımından önce, bir düğümün komşusu tarafından bir mesajın yeniden iletilip iletilmediğini tespit eder. Bu kural ile hizmet reddi saldırıları tespit edilebilir.

Radyo İletim Aralığı Kuralı: Gelen paketlerin, komşu düğümden gelip gelmediğini tespit eder. Böylece Sybil, solucan deliği ve Hello sel saldırıları tespit edilebilir.

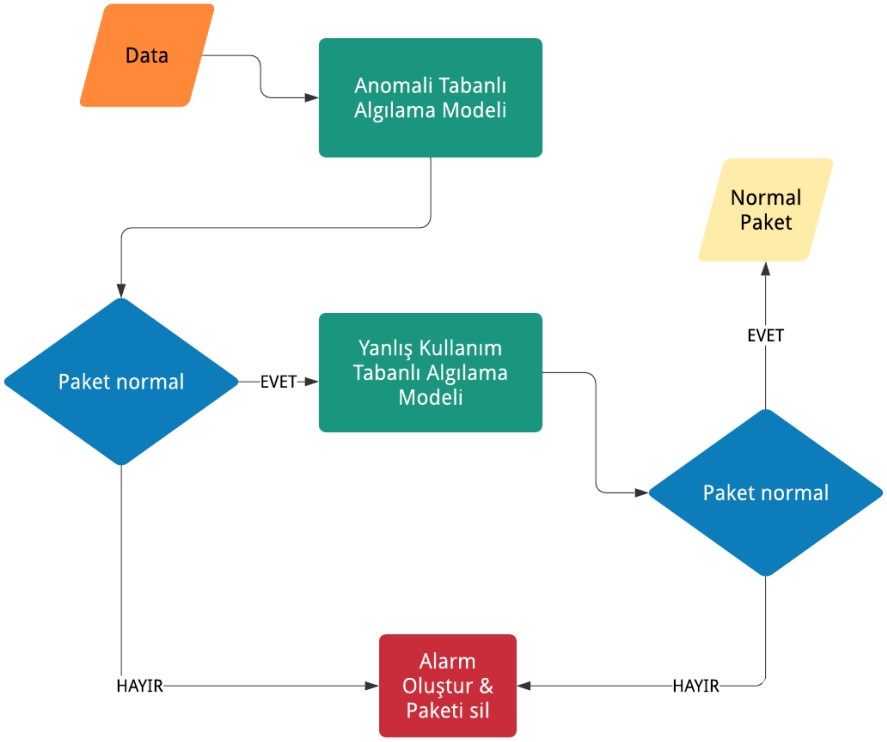
Anomali tabanlı algılama modelinin akış diyagramı Şekil 2'de gösterilmiştir.



Şekil 2. Anomali tabanlı algılama modelinin akış diyagramı

5.2. Yanlış Kullanım Tabanlı Algılama Metodu

Anomali tabanlı algılama modelinin temel avantajı, yüksek algılama oranı iken diğer yandan dezavantajı, yüksek bir yanlış alarm oranı oluşturmasıdır. Söz konusu hatalı sınıflandırma işlemine maruz kalan paketleri yeniden incelemeye tabi tutmak için yanlış kullanım tabanlı algılama modeli devreye girecektir. Bu model, bilinen saldırı çeşitlerinin profillerini kullandığından, söz konusu saldırıların profillerinin oluşturulması gerekir. Gelen paketlerin, oluşturulan saldırı profilleri ile eşleşmesi durumunda, söz konusu paket anormal/saldırı paketi olarak tespit edilecek ve bu paketler silinecektir. Önerilen sistemin çalışma mekanizması Şekil 3'te gösterilmiştir.

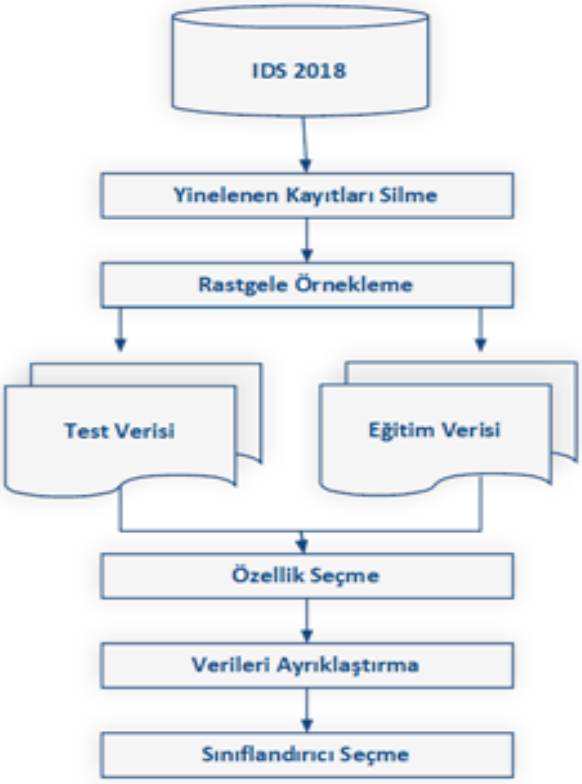


Şekil 3. Önerilen STS'nin çalışma mekanizması

Saldırı profilleri oluştururken, saldırılar için belirleyici nitelikte olan özellikler, ilgili ağın gözlemlenmesi yoluyla elde edilebilse de bu yöntem maliyetlidir. Bu sebeple yapay veriler/veri setleri kullanılmaktadır.

Etkili bir STS tasarımı için veri setlerinin gerçek ağ trafiğinden elde edilmiş ve güncel olması hayati öneme sahiptir. Bu sebeple yapılan çalışmada; İletişim Güvenliği Kuruluşu (CSE) ve Kanada Siber Güvenlik Enstitüsü (CIC) arasında ortak bir proje olarak geliştirilen CSE-CIC-IDS2018 veri seti kullanılmıştır. Söz konusu veri seti; Brute-force, Heartbleed, Botnet, DoS, DDoS, Web saldırıları ve ağa içeriden sızılması şeklinde 7 farklı senaryo içermektedir. Veri kümesi, CICFlowMeter-V3 kullanılarak yakalanan trafikten çıkarılan 80 özellik ile birlikte her makinenin ağ trafiğini ve sistem günlüklerini içerir (Canadian Institute for Cybersecurity, 2021).

IDS 2018 veri seti yaklaşık 16 milyon örnek kayıt içermekte olup, bu kadar fazla örneğin KAA'da uygulanması neredeyse imkânsızdır. Ayrıca STS'de metotların hibrit olarak kullanılması, hesaplama karmaşıklığını ve kaynak tüketimini arttıracağından, söz konusu kısıtlar göz önünde bulundurularak, veriler bir ön işleme tabi tutulacaktır. Ön işleme adımları Şekil 4'te gösterilmiştir.



Şekil 4. Veri setinde uygulanacak ön işlemler

5.2.1. Ön işleme adımları

Ön işleme kapsamında, veri setinde tekrar eden veriler ile etkisiz niteliklerin silinmesi, rasgele örnekleme, veri ayrıklaştırma ve uygun sınıflandırma algoritmasının seçimi gerçekleştirilecektir.

Yinelenen kayıtları silme: Veri setinde toplam 433.261 yinelenen kayıt vardır ki, bu yüksek bir orandır. Yinelenen kayıtlar silindikten sonra geriye 15.799.751 benzersiz kayıt bırakılmıştır. Önişleme yapılmadan önceki normal ağ trafiği %83.07, saldırı trafiği ise %16.93'tür. Yinelenen kayıtlar silindikten sonra ise normal ve saldırı trafiği istatistikleri Tablo 3'te gösterilmiştir.

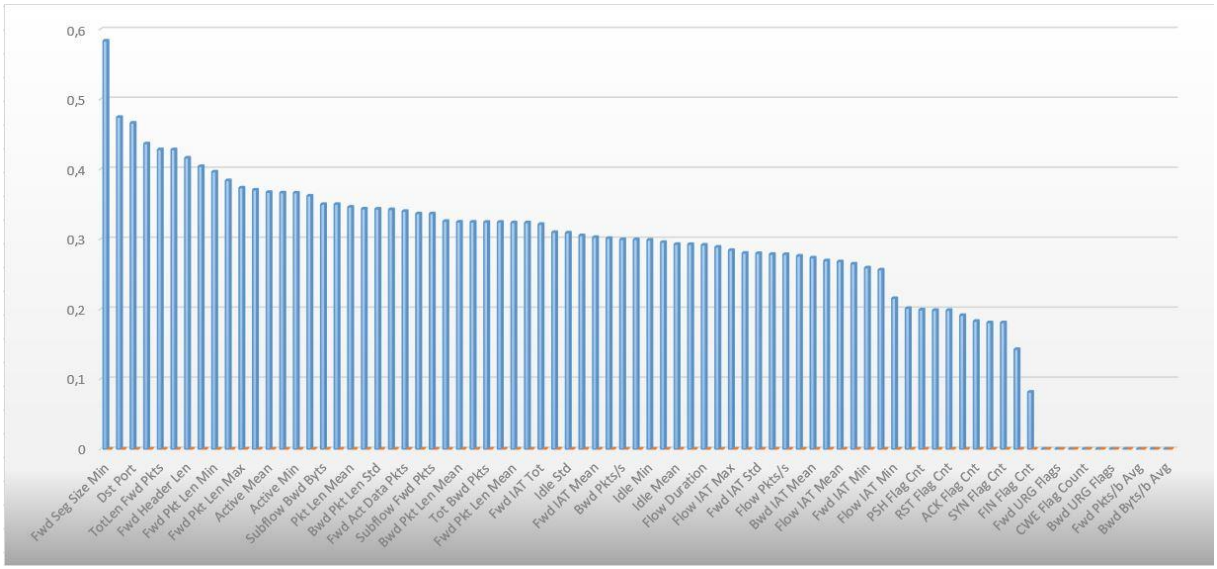
Tablo 3. Yinelenen kayıtlar silindikten sonra ağ trafiği istatistiği

Kategori	Sayı	Yüzde
Normal	13445501	85,099%
FTP-BruteForce	39353	0,249%
SSH-Bruteforce	117323	0,743%
DoS attacks-Slowloris	10286	0,065%
DoS attacks-GoldenEye	41456	0,262%
DoS attacks-SlowHTTPTest	19463	0,123%
DoS attacks-Hulk	434874	2,752%
DDoS attacks-LOIC-HTTP	576176	3,647%
DDOS attack-HOIC	668462	4,231%
DDOS attack-LOIC-UDP	1731	0,011%
Brute Force -Web	612	0,004%
Brute Force -XSS	231	0,001%
SQL Injection	88	0,001%
Infilteration	161898	1,025%
Bot	282311	1,787%

Rastgele örnekleme: Veri setinde bulunan mevcut kayıt sayısı hâlâ KAA'ların performans ve kaynak kısıtları için optimum düzeyde değildir. Bu nedenle tüm kayıtlar arasından rastgele örnekleme yapılarak toplam 20 bin kayıt seçilmiş ve bu kayıtlar normal/saldırı profili oluşturmak için kullanılmıştır. Seçilen kayıtların 10 bini normal trafiği yansıtırken, 10 bini ise saldırı özellikleri taşımaktadır.

Veri setinde bulunan bazı saldırı tiplerinin kayıt sayıları çok düşükken, bazı saldırı tiplerinin ise çok yüksektir. Rastgele örnekleme yapılırken düşük kayıt içeren saldırı tiplerinin tüm kayıtları alınmış, yüksek kayıt sayısına sahip saldırı tipleri ise indirgenmiştir.

Özellik Seçme: Veri setinde bulunan her kayıt toplam 80 özellik içermektedir. Bu özelliklerin bir kısmı ise saldırı profili oluşturmak için gereksiz/ilgisizdir. Normal ve saldırı profilleri oluşturmak için en faydalı ve en önemli özellikleri belirlemek, gereksiz ve ilgisiz özellikleri ise veri setinden çıkarıp, performans ve doğruluk kriterlerini iyileştirmek için GainRatio algoritması kullanılmıştır. Algoritmanın veri setinde uygulanmasının akabinde özelliklerin etki oranı Şekil 5'te gösterilmiştir.



Şekil 5. IDS2018 veri setinde bulunan özelliklerin, saldırı profili oluşturmaya etki oranı

GainRatio algoritmasına göre özelliklerin ortalama etki oranı 0,271299'dur. Veri setindeki niteliklerin bir kısmı ortalama etki oranının altında iken, bir kısmının ise etki değeri yoktur. Neticede etki oranı az veya sıfır olan nitelikler kaldırılmış, saldırı profili oluşturmak için 39 özellik kullanılmıştır.

Özellik seçiminin performans ve doğruluğa etki oranı J48 algoritması ile test edilmiştir. Özellik seçiminden sonra yaklaşık iki buçuk kat daha hızlı ve %2 daha fazla doğruluk oranında model oluşturulmuştur.

Verilerin Ayrıklaştırılması: Numerik verilerin kategorik karşılıklarına dönüştürülmesi işlemi olan ayrıklaştırma işlemi, verideki gürültüyü ve doğrusalsızlığı (non-linearity) azaltarak tahminleme modelinin kesinliğini artırabilir. Bununla birlikte veri boyutunu azaltmak ve dolayısıyla performansı artırmak için kullanışlı bir yöntemdir. Yapılan çalışmada, gözetlenen (supervised) ayrıklaştırma yöntemi kullanılarak veriler ayrıklaştırılmıştır. Verilerin ayrıklaştırma işlemine tabi tutulmasının ardından J48 algoritması ile test edilmiştir. Neticede ayrıklaştırma sonrası algoritmanın yaklaşık 6 kat daha hızlı çalıştığı ve %0.27 daha doğru tespit yaptığı görülmüştür.

5.2.2. Sınıflandırıcı seçimi

Ön işleme adımlarının tamamlanmasından sonra ağ trafiğindeki paketlerin normal veya saldırı olarak sınıflandırılması için BayesNet, J48, JRip, PART ve RandomForest algoritmaları kullanılmıştır.

BayesNet, meydana gelen olayları neden sonuç bağlamında açıklamak için kullanılan bir modelleme algoritmasıdır. Örneğin, hastalık ve semptomları arasındaki ilişkiyi ortaya koymak için kullanılabilir (Wikipedia, 2021).

J48 algoritması, değişkenlerin entropi değerlerine dayanan ve Shannon'un bilgi teorisinden (Wikipedia, 2021) yararlanarak, karar ağaçlarını optimize etmeyi hedefleyen bir algoritmadır. Quinlan tarafından geliştirilen C4.5 algoritmasına dayanır.

JRip algoritması, William W. Cohen tarafından IREP'in optimize edilmiş bir versiyonu olarak geliştirilmiştir. Karar ağacı algoritmalarında bulunan çok yaygın ve etkili bir teknik olan azaltılmış hata budama (REP) ile birliktelik kurallarına dayanır (Wikibooks, 2021). JRip, eğitim verilerindeki örneklerin belirli yargılarını bir sınıf olarak ele alarak ve sınıfın tüm üyelerini kapsayan kurallar kümesini bularak kuralları öğrenen aşağıdan yukarıya bir yöntemdir (Sonawani ve arkadaşları, 2013).

PART; C4.5 ve RIPPER algoritmalarının geliştirilmiş versiyonu olan kısmi karar ağacı algoritmasıdır. PART algoritmasının temel özelliği, uygun kuralları üretmek için C4.5 ve RIPPER gibi global optimizasyon gerçekleştirmesine gerek olmamasıdır (Ali ve arkadaşları, 2006).

Random Forest algoritması, tahmin modelleri oluşturabilmek için birçok karar ağacı üretip, birleştirmeye dayanan bir mekanizmayı baz alır. Algoritma, Breiman tarafından geliştirilmiştir (Liu ve arkadaşları, 2012).

6. BENZETİM SONUÇLARI

Sınıflandırma algoritmalarının veri setinde uygulanması için Weka (Waikato Environment for Knowledge Analysis) uygulaması kullanılmıştır. Weka, Waikato Üniversitesinde geliştirilmiş olup, yaygın olarak kullanılan veri madenciliği algoritmalarını içeren, açık kaynak kodlu, ücretsiz bir araçtır.

Ön işlemden geçirilen veri setindeki kayıtların %66'sı eğitim için, geri kalanı ise test için kullanılmıştır. Algoritmaların performans analizi için; doğruluk, kesinlik, duyarlılık, F-ölçütü, Matthews Correlation Coefficients (MCC), ROC, Precision-recall Curve (PRC) ve işlem süresi kriterleri kullanılmıştır. Doğruluk; doğru sınıflandırılan kayıtların, toplam kayıtlara oranıdır. Kesinlik; doğru sınıflandırılan kayıtların, pozitif tahminlere oranıdır. Duyarlılık ise doğru sınıflandırılan kayıtların, toplama oranıdır. F-Ölçütü duyarlılık ve kesinlik değerlerinin harmonik ortalamasıdır. MCC, dengesiz dağılıma sahip veri setlerinde gerçekçi sonuçlara ulaşılmasını sağlar. ROC eğrisi, duyarlılık ve özgüllük değerlerinin kesişimlerinden elde edilir. PRC eğrisi ise X ve Y eksenlerinde, duyarlılık ve kesinlik değerlerinin kesişimlerinden elde edilir.

Tablo 4. Algoritmaların performans değerleri

Algoritma	Doğruluk (%)	Kesinlik	Duyarlılık	F-Ölçütü	MCC	ROC	PRC	Süre (sn.)
BayesNet	77.5	0,856	0,775	0,796	0,737	0,979	0,918	**0.17**
J48	**93.9706**	**0,935**	**0,940**	0,931	**0,914**	0,980	0,933	0.22
JRip	93.9265	0,934	0,939	0,930	0,913	0,974	0,912	15.12
PART	93.7353	0,931	0,937	0,930	0,911	0,983	0,939	1.18
RandomForest	93.8824	0,932	0,939	**0,932**	0,913	**0,983**	**0,941**	5.36

Algoritmaların performans sonuçları Tablo 4'te gösterilmiştir. Tablodaki sonuçlardan da anlaşılacağı üzere doğruluk, kesinlik, duyarlılık, F-Ölçütü ve MCC kriterleri ile süre performansı bakımından en iyi sonuçlar, J48 algoritması ile elde edilmektedir. BayesNet algoritması süre bakımından en iyi sonucu vermesine rağmen, doğruluk oranı bakımından kabul edilebilir bir seviyede değildir. JRip, PART ve RandomForest algoritmaları ise doğruluk ve diğer kriterler bakımından birbirlerine yakın performanslar göstermektedirler. Bununla birlikte JRip ve RandomForest algoritmaları işlem süresi bakımından PART algoritmasından çok daha ağır çalışmaktadır. Dolayısıyla KAA'larda çalışmaya pek elverişli değildir. PART algoritması ise işlem süresi bakımından makul denebilecek seviyede çalışsa da, J48 algoritmasından yaklaşık 5 kat daha ağır çalıştığı ve daha düşük doğruluk oranına sahip olduğu göz önüne alındığında, isabetli bir tercih sayılmaz. Sonuç olarak, J48 algoritması en iyi seçenek olarak görünmektedir.

7. SONUÇ

Saldırı tespit sistemleri, kablosuz algılayıcı ağlarda güvenliği sağlamak için etkili bir yöntem olmakla birlikte literatür çalışmalarının kahir ekseriyeti güncelliğini yitirmiş veri setleri referans alınarak hazırlanmıştır. Dolayısıyla elde edilen doğruluk ve performans sonuçları, güncel saldırılar karşısında tutarlı sonuçlar vermeyecektir.

Yapılan çalışmada hibrit bir saldırı tespit sistemi modellenmiş ve sınıflandırma algoritmalarının uygulanması aşamasında, güncel ve modern bir veri seti olan IDS2018 tercih edilmiştir. Ayrıca KAA'ların donanım kısıtları, pil kapasiteleri ve diğer karakteristikleri göz önünde bulundurularak, veri seti üzerinde ön işleme adımları uygulanmıştır. Kullanılan algoritmalardan J48 algoritmasının en iyi sonuçlar verdiği tespit edilmiştir. Benzetim sonuçları önerilen STS modelinin yüksek bir doğruluğa ve düşük bir çalışma süresine sahip olduğunu, dolayısıyla KAA güvenliği için kullanılabilecek nitelikte olduğunu göstermektedir.

Kaynaklar

Kumar, V., Jain, A., Barwal, P. N. (2014) "Wireless Sensor Networks: Security Issues, Challenges and Solutions", International Journal of Information & Computation Technology. ISSN 0974-2239 Volume 4, Number 8, pp. 859-868.

Chelli, K. (2015) "Security Issues in Wireless Sensor Networks: Attacks and Countermeasures", Proceedings of the World Congress on Engineering Vol I

Deng, R., Zhuang, P., Liang, H. (2017) "CCPA: Coordinated Cyber-Physical Attacks and Countermeasures in Smart Grid," IEEE Transactions on Smart Grid, vol. 8, no. 5, pp. 2420–2430

Padmavathi, G., Shanmugapriya, D. "A Survey of Attacks, Security Mechanisms and Challenges in Wireless Sensor Networks", International Journal of Computer Science and Information Security (IJCSIS), Vol. 4, No. 1 & 2

Tomić I., McCann, J. A. (2017) "A Survey of Potential Security Issues in Existing Wireless Sensor Network Protocols", in IEEE Internet of Things Journal, vol. 4, no. 6, pp. 1910-1923, doi: 10.1109/JIOT.2017.2749883.

Ozcelik, M., Irmak, E., Ozdemir, S. (2017) "A Hybrid Trust Based Intrusion Detection System for Wireless Sensor Networks", International Symposium on Networks, Computers and Communications, Marrakech, pp. 1-6

Ghugar, U., Pradhan, J., Bhoi, S. K., Sahoo, R. R. (2019) "LB-IDS: Securing Wireless Sensor Network Using Protocol Layer Trust-Based Intrusion Detection System", Journal of Computer Networks and Communications Volume, Article ID 2054298

Çavuşoğlu, Ü., Kaçar, S. (2019) "Anormal Trafik Tespiti için Veri Madenciliği Algoritmalarının Performans Analizi", Academic Platform Journal of Engineering and Science 7-2, 205-216

Acharya, N., Singh, S. (2018) "An IWD-based feature selection method for intrusion detection system", Soft Comput 22, 4407–4416

Altun, B. (2016) "Kablosuz Sensör Ağları ve Uygulama Alanları", Karabük Üniversitesi Mühendislik Fakültesi, 61-62

Martins, D., Guyennet, H. (2010) "Wireless Sensor Network Attacks and Security Mechanisms: A Short Survey," 13th International Conference on Network-Based Information Systems, pp. 313-320,

Dolay, B. (2009) "Kablosuz Sensör Ağları", https://e-bergi.com/y/kablosuz-sensor-aglari, Son Erişim: 3 Eylül 2021

Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., Cayirci, E. (2002) "Wireless Sensor Networks: A Survey". Computer Networks, 38. 393-422. 10.1016/S1389-1286(01)00302-4.

Amara, S., Beghdad, R., Oussalah, M. (2013). "Securing Wireless Sensor Networks: A Survey." EDPACS. 47. 10.1080/07366981.2013.754207.

Dewal P., Narula G.S., Jain V., Baliyan A. (2018) Security Attacks in Wireless Sensor Networks: A Survey. In: Bokhari M., Agrawal N., Saini D. (eds) Cyber Security. Advances in Intelligent Systems and Computing, vol 729. Springer, Singapore. https://doi.org/10.1007/978-981-10-8536-9_6

Alrajeh, N. A., Khan, S., & Shams, B. (2013). Intrusion Detection Systems in Wireless Sensor Networks: A Review. International Journal of Distributed Sensor Networks. https://doi.org/10.1155/2013/167575

Farooqi, A., Khan, F. (2009). "Intrusion Detection Systems for Wireless Sensor Networks: A Survey." International Journal of Ad Hoc and Ubiquitous Computing. 9. 234-241. 10.1504/IJAHUC.2012.045549.

Ozgur, D., Topallar, M., Anarim, E., Ciliz, M.K. (2005). An intelligent intrusion detection system (IDS) for anomaly and misuse detection in computer networks. Expert Syst. Appl.. 29. 713-722. 10.1016/j.eswa.2005.05.002.

Silva, A. P., Martins, M., Rocha, B., Loureiro, A., Wong, H.(2005) "Decentralized intrusion detection in wireless sensor networks." 16-23. 10.1145/1089761.1089765.

Canadian Institute for Cybersecurity. https://www.unb.ca/cic/datasets/ids-2018.html, Son Erişim: 3 Eylül 2021

Information theory, Wikipedia https://en.wikipedia.org/wiki/Information_theory, Son Erişim: 3 Eylül 2021

Bayes Ağı, Wikipedia, https://tr.wikipedia.org/wiki/Bayes_a%C4%9F%C4%B1, Son Erişim: 3 Eylül 2021

Data Mining Algorithms In R/Classification/JRip, https://en.wikibooks.org/wiki/Data_Mining_Algorithms_In_R/Classification/JRip, Son Erişim: 3 Eylül 2021

Sonawani, S., Mukhopadhyay, D. (2013) "A Decision Tree Approach to Classify Web Services using Quality Parameters"

Ali, S., Smith, K. "On learning algorithm selection for classification." Applied Soft Computing. 6. 119-138. 10.1016/j.asoc.2004.12.002

Liu Y., Wang Y., Zhang J. (2012) New Machine Learning Algorithm: Random Forest. In: Liu B., Ma M., Chang J. (eds) Information Computing and Applications. ICICA 2012. Lecture Notes in Computer Science, vol 7473. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-34062-8_32