KAMPÜS BİLGİSAYAR AĞLARINDA MEYDANA GELEN SORUNLARIN OTOMATİK TESPİT EDİLMESİ

İlhan Fırat KILINÇER

Yüksek Lisans Tezi
Elektrik - Elektronik Mühendisliği Anabilim Dalı
Danışman: Dr. Öğr. Üyesi Ayhan AKBAL

HAZİRAN-2018

T.C
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

KAMPÜS BİLGİSAYAR AĞLARINDA MEYDANA GELEN SORUNLARIN OTOMATİK TESPİT EDİLMESİ

YÜKSEK LİSANS TEZİ

İlhan Fırat KILINÇER

(161113114)

Anabilim Dalı: Elektrik - Elektronik Mühendisliği

Programı: Telekomünikasyon

Danışman: Dr. Öğr. Üyesi Ayhan AKBAL

Tezin Enstitüye Verildiği Tarih: 25.05.2018

HAZİRAN-2018

T.C
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

KAMPÜS BİLGİSAYAR AĞLARINDA MEYDANA GELEN SORUNLARIN OTOMATİK TESPİT EDİLMESİ

YÜKSEK LİSANS TEZİ

İlhan Fırat KILINÇER

(161113114)

Tezin Enstitüye Verildiği Tarih : 25.05.2018
Tezin Savunulduğu Tarih : 20.06.2018

Tez Danışmanı : Dr. Öğr. Üyesi Ayhan AKBAL (F.Ü)

Diğer Jüri Üyeleri : Doç.Dr. Muhammed Fatih TALU (İ.Ü)

Dr.Öğr. Üyesi Turgay KAYA (F.Ü)

HAZİRAN-2018

ÖNSÖZ

Bu çalışmada, internet erişimi sağlayan ağ cihazlarında ve bu cihazların hatlarında oluşabilecek temel arızaların erkenden tespit edilebilmesi için bir yöntem geliştirilmiştir. Geliştirilen yöntem üniversite kampüsü ortamında gerçekleştirilmiştir. Bu kapsamda geliştirilen yazılım ile ağ cihazlarından belirli veriler gerçek zamanlı olarak alınmıştır. Alınan ham veriler üzerinde ön işlem yapılarak analiz için uygun hale getirilmektedir. Bu amaçla makine öğrenmesi yaklaşımlarından faydalanılmıştır. Uzman görüşü ile etiketlenen veriler eğitim için kullanılmış, daha sonra alınan veriler Karar Ağaçları, Destek Vektör Makinesi ve Naive Bayes makine öğrenmesi metodları ile sınıflandırılmıştır. Bu tez çalışmasının da bu amaç doğrultusunda faydalanılabilir bir kaynak olmasını umuyorum.

Bu çalışmada değerli vaktini bana ayırarak çalışmamın bitirilmesinde her türlü desteğini esirgemeyen sayın danışman hocam Dr.Öğr.Üyesi Ayhan AKBAL' a teşekkürlerimi sunarım.

Tez süresince her türlü desteğini esirgemeyen, başta Orhan YAMAN ve Fatih ERTAM olmak üzere, Fırat Üniversitesi Bilgi İşlem Dairesi Başkanlığı'nda görev yapan mesai arkadaşlarıma teşekkür ediyorum.

Ayrıca hayatımın her anında ilgi, anlayış ve her türlü desteğini esirgemeyen aileme çok teşekkür ediyorum.

İÇİNDEKİLER LİSTESİ

ÖZET

Son yıllarda, gelişen teknolojiler ile birlikte internet kullanımı da yaygın hale gelmiştir. Sosyal medya başta olmak üzere birçok ihtiyaç için internet kullanılmaktadır. Günümüzde ev ve ofislerde kullanılan elektronik cihazların uzaktan kullanımı için de internet gerekmektedir. Bu tez çalışmasında, internet erişimi sağlayan ağ cihazlarında ve bu cihazların hatlarında oluşabilecek temel arızaların erkenden tespit edilebilmesi için bir yöntem geliştirilmiştir. Geliştirilen yöntem üniversite kampüsü ortamında gerçekleştirilmiştir. Bu kapsamda geliştirilen yazılım ile ağ cihazlarından belirli veriler gerçek zamanlı olarak alınmıştır. Alınan ham veriler üzerinde ön işlem yapılarak analiz için uygun hale getirilmektedir. Bu amaçla makine öğrenmesi yaklaşımlarından faydalanılmıştır. Uzman görüşü ile etiketlenen veriler eğitim için kullanılmış, daha sonra alınan veriler Karar Ağaçları, Destek Vektör Makinesi ve Naive Bayes makine öğrenmesi metodları ile sınıflandırılmıştır. İnternet hattında oluşabilecek arızaların giderilmesi uzun süre ve iş yükü gerektirmektedir. Geliştirilen bu yöntem sayesinde, üniversite kampüs ortamında oluşabilecek arızaların erken tespit edilmesi sağlanarak bakım çalışmaları yapılmaktadır.

Anahtar Kelimeler: Ağ yönetimi, sınıflandırma, karar ağaçları, destek vektör makinesi, bayes metodu, ağ durum izleme, hata tespiti.

SUMMARY

AUTOMATIC NETWORK PROBLEM DETECTION OCCURRING IN THE CAMPUS COMPUTER NETWORK

In recent years, the use of the internet has become widespread with developing technologies. Internet is used for many needs, especially social media. Today, internet is also needed for remote use of electronic devices used in homes and offices. Continuous access to the internet has become very important for people's quality of life. In this thesis, a method has been developed for early detection of basic faults that may occur in network devices providing internet access and in the lines of these devices. The developed method was carried out in the university campus environment. With this software developed, certain data are taken in real time from network devices. This raw data is pre-processed and made suitable for analysis. Machine learning approaches have been used for this purpose. The data labeled with the expert opinion were used for training, then the data received were classified by Decision Trees, Support Vector Machine and Naive Bayes machine learning methods. Eliminating the malfunctions that may occur in the Internet line requires long time and workload. Thanks to this developed method, maintenance studies are performed by providing early detection of failures that may occur in the university campus environment.

Key Words: Network management, classification, decision trees, support vector machine, bayes method, network status monitoring, error detection.

ŞEKİLLER LİSTESİ

TABLOLAR LİSTESİ

SEMBOLLER LİSTESİ

x_i : Öğrenme verisi

y_i : Öğrenme verisi

w : Normal vektörü

b : Ofset değeri

n : Birbirinden bağımsız üretilen veri sayısı

p_i : Her veri için üretim olasılığı

v_i : Veriler

X : Örnekler kümesi

C : Veri kümesindeki sınıflar

KISALTMALAR LİSTESİ

SLA : Service Level Agreement
LAN : Local Area Network
WAN : Wide Area Network
MAN : Metropolitan Area Network
ATM : Asynchronous Transfer Mode
FDDI : Fiber Distributed Data İnterface
TCP/IP : Transmission Control Protocol/ İnternet Protokol
UDP : User Datagram Protokol
MAC : Media Access Control
UTP : Unshielded Twisted-Pair
STP : Shielded Twisted-Pair
CAT : Category
SNMP : Simple Network Management Program
VLAN : Virtual Local Area Network
QoS : Quality of Service
Rx : Received
Tx : Transmitted
CRC : Cylic Redundancy Check
NV : Naive Bayes
Tx : Transmitted
CRC : Cylic Redundancy Check
IoT : İnternet of Things
BT : Bilgi Teknolojileri
ABSAMN : Agent Based Activity on Network
RCV : Route Convergence Visualizer
DVM : Destek Vektör Makineleri
TP : True Pozitive
FP : False Positive
TN : True Negatif
FN : False Negatif
BER : Bit Error Rate

PC : Personal Computer

NB : Naive Bayes

SSH : Secure Shell

ROC : Receiver Operating Characteristic

SVM : Support Vector Machine

1. GİRİŞ

Teknolojinin gelişmesi ile birlikte internet, günlük hayatın vazgeçilmez unsurlarından biri haline gelmiştir. En genel haliyle bilgisayar ağları; satış, pazarlama, kişisel verilerin saklanması gibi birçok alanda kullanılmaktadır. Buna ek olarak bu bilgilerin farklı coğrafi bölgede bulunan kişi veya kurumlarla paylaşılabilmesi bilgisayar ağları ile mümkün olabilmektedir. İnternet, günümüzde birçok alanda kullanılan en büyük bilgisayar ağıdır. Dolayısıyla internet, dünyanın her lokasyonundaki ağları belirli protokoller vasıtasıyla haberleştiren bir bağlantıdır.

Kurumsal ağlar gibi büyük ağlarda, internet erişiminin kesintisiz bir şekilde verilmesi büyük bir önem taşımaktadır. Bu bağlamda günümüzde özellikle bilişim sektöründe five 9 olarak da bilinen %99,999 çalışma süresi kavramı sıklıkla kullanılmaktadır. Bu kavram bilgisayar ağının çalışma süresini ve güvenirliği gösterir. Five 9 kavramına göre bir sistemin verimli ve güvenli olması için bu sistem % 99,999'luk bir kullanılabilirliğe sahip olmalıdır. Kurumlar %99,999'luk kullanılabilirlik hedefine ulaşabilmek için hizmet aldıkları firma ile Hizmet Düzeyi Sözleşmesi (SLA) imzalamaktadırlar. Bu anlaşma, kuruma hizmet sağlayan firmanın hizmet kalite standardını belirtmektedir. Bu anlaşma ile sağlanan hizmetlerin niteliği, miktarı, hizmetin teslim süresi, teknik destek, yedek almak ve sorun çözme süreleri gibi kriterler yer almaktadır. Böylece olası bir hizmet kesintisi durumunda çok hızlı müdahale edilebilmesini garantilemektedirler.

Özelikle bankacılık, borsa, telekomünikasyon gibi kritik öneme sahip sektörde meydana gelebilecek İnternet kesintileri iş akışını sekteye uğratmakta ve zaman zaman büyük maddi kayıplara neden olmaktadır. Kurumlarda internet hizmetinin kesintisiz verilebilmesi için verimli bir ağ yapısının kurulması gerekmektedir. Verimli bir ağ yapısının en önemli unsurlarından biri kurulan ağ cihazlarının izlenebilmesidir. Çünkü önceden fark edilmeyen bir sorundan dolayı ağ cihazında meydana gelebilecek bir sorun büyük hizmet kesintilerine neden olmaktadır. Ayrıca kullanılan ağ cihazları pahalı olup arıza sırasında yenisinin temin edilmesi oldukça zaman almaktadır. Bu nedenle kurumsal ağlarda ağ cihazlarının daha uzun ve performanslı çalışabilmesi için sıcaklığı ve işlemci bilgisi gibi birçok parametresi sürekli izlenmelidir.

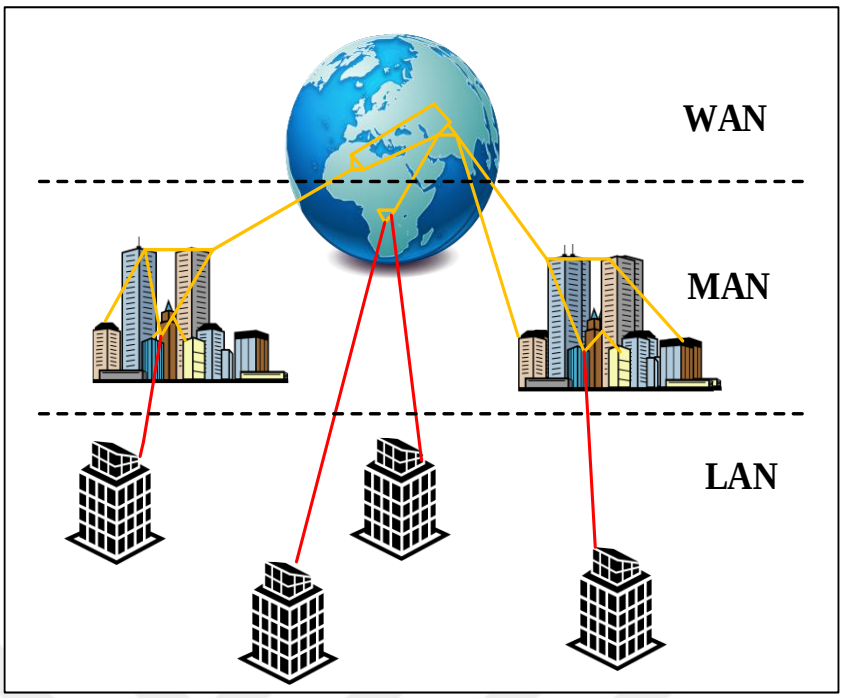
1.1. Ağ Teknolojileri

Bilgisayar, yazıcı, modem, kamera gibi birçok cihazın belirli protokoller vasıtasıyla haberleşmesini sağlayan yapıya network (ağ) denir. Bu haberleşme bir kampüs ya da odada olabileceği gibi kıtalar ve ülkeler arasında da olmaktadır. Bu şekilde aynı ortam ya da farklı ortamdaki birden fazla cihazın birbiriyle iletişime geçebilmesi işgücü ve maliyet açısından çok büyük avantajlar sağlamaktadır. Ağ teknolojisi ile birlikte bilgi paylaşımı, merkezi yönetim gibi konular oldukça kolay hale gelmiştir.

1.1.1. Ağ Yapıları

Ağlar büyüklüklerine göre LAN (Local Area Network), WAN (Wide Area Network), MAN (Metropolitan Area Network) olarak ayrılırılar. Bu yapıları tanımlayacak olursak;

- ***LAN:*** Aynı ortam ya da bölgede bulunan cihazların aynı haberleşme ortamı üzerinden birbirine bağlanması sonucu oluşan yapılardır.
- ***WAN:*** Farklı yada aynı bölgelerde bulunan iki yada daha fazla LAN'ın birbirlerine çeşitli protokoller aracılığıyla bağlanmasıdır.
- ***MAN:*** Daha çok deniz aşırı ülkeleri birbirine fiber optik kablolar ile birbirlerine bağlanmasıdır.



Şekil 1.1. LAN, WAN, MAN yapısı

LAN, WAN, MAN yapıları Şekil 1.1'de özet şekil halinde verilmiştir. LAN, WAN ve MAN yapılarının haberleşmesi için kullanılan bazı ağ teknolojilerini aşağıdaki gibi sıralayabiliriz.

a) Ethernet
b) Token ring
c) ATM (Asynchronous Transfer Mode)
d) FDDI (Fiber Distributed Data İnterface)
e) Frame Relay

Bu seminer kapsamında Ethernet teknolojisi üzerinde çalışılacak olup Ethernet teknolojisi ile ilgili tanımlar önemsenecektir.

1.1.2. OSI Referans Modeli

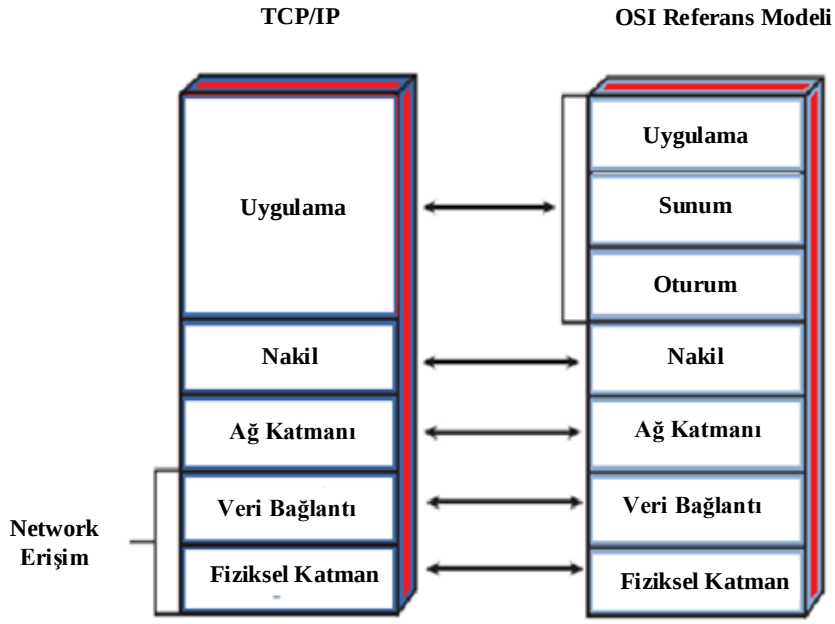
Ağlarda bulunan değişik işletim sistemlerine sahip cihazların birbirleriyle haberleşmesine olanak sağlayan modeldir. OSI modeli katmanlı bir yapıya sahiptir ve her bir katmana belirli görevler tahsis edilmiştir. Model tasarlanırken her katman kendisinden bir üst seviyede bulunan katmana hizmet edecek şekilde tasarlanmıştır. OSI referans modeli katmanları ve katmanların işlevleri Tablo 1.1'de verilmiştir.

Tablo 1.1. OSI referans modeli

Katman	İşlev	Örnek
Uygulama Katmanı	Kullanıcıya en yakın katmandır.	Telnet, HTTP
Sunum Katmanı	Gönderilecek verinin, veriyi alacak bilgisayar tarafından da anlayabileceği ortak bir yapıya dönüştürür.	ASCII/EBCDIC, JPEG/MP3
Oturum Katmanı	Sunum katmanının çalışması için gerekli servisleri sağlar. Port katmanı olarak da bilinir. Uygulamaları yönetir.	RPC, SQL
İletim Katmanı	Güvenli ve verimli iletişimi sağlar. Hata denetim mekanizmalarını barındırır.	TCP, UDP
Ağ Katmanı	Veri paketinin ağ içinde yada dışında gideceği yeri belirler. Veri için en iyi yol seçimi yapar.	BGP, RIP, OSPF
Data Link Katmanı	Fiziksel adreslemenin yapıldığı katmandır. Son kullanıcıların MAC adreslerini kullanarak iletişimi sağlar.	Ethernet, FDDI, Frame Relay
Fiziksel Katman	Bağlı uçlara arasında elektriksel tanımlamalar yapar ve veri hareketinin başlatılmasını, sürdürülmesini ve bitmesini sağlar.	EIA/TIA

1.1.3. İletim Protokolleri

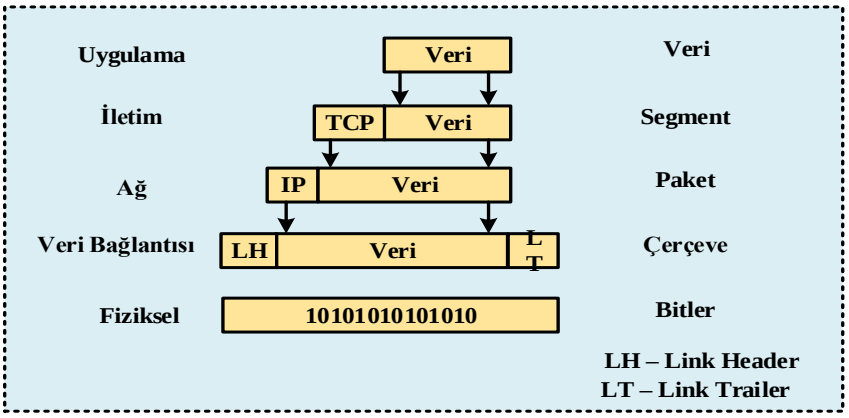
İletim protokolleri başlığı altında bilinmesi gereken en temel şey şüphesiz TCP/IP (Transmission Control Protocol/İnternet Protocol) protokolüdür. TCP/IP, OSI (Open Systems Interconnections) referans modeli içerisinde katmanlar arasındaki iletişimi sağlayan protokoldür. TCP/IP protokolü de OSI referans modeli gibi katmanlı bir yapıya sahiptir. Her iki modelin katmanları benzerlik gösterir. OSI referans modelindeki Oturum, Sunum ve Uygulama katmanları TCP/IP protokolünde Uygulama Katmanı olarak birleştirilmiştir. TCP/IP modeli ile OSI referans modeli katmanlarının karşılaştırması Şekil 1.2'de verilmiştir.



Şekil 1.2. TCP/IP ve OSI modelleri karşılaştırması

TCP/IP protokolü kendi içerisinde haberleşmeyi sağlamak için TCP ve UDP (User Datagram Protocol) protokollerinden yararlanır. TCP iletişim halinde olan iki cihaz arasında güvenli bir veri akışı ve hata denetimi yaparken, UDP protokolü güvenilir veri akışı ve akış kontrolü sağlamaz. Dolayısıyla veri güvenliğinin önemsenmediği, veri paylaşım hızının daha önemli olduğu uygulamalarda UDP protokolü kullanılırken veri güvenliğinin önemli olduğu uygulamalarda TCP protokolü kullanılır.

Günümüzde bilgisayar ağlarında en çok kullanılan iletişim teknolojisi ethernet teknolojisidir. Ethernet teknolojisinde haberleşme TCP/IP protokolü ile sağlanır. Bu teknolojiyi kullanan cihazlar üzerinde bulundurdukları ethernet kartı aracılığıyla iletişim kurarlar. Her Ethernet kartının kendine ait dünyada eşi bulunmayan bir numaraya sahiptir. Bu numaraya MAC (Media Acces Control) ya da fiziksel adres denir ve 48 bitten oluşmaktadır. Ethernet IEEE 802.3 standardı ile tanımlanmıştır. Ethernet teknolojisinde veriler encapsulation işleminden geçirdikten sonra frame'ler halinde istemciye gönderilir. Bir ethernet teknolojisindeki veri kapsüllemesi genel olarak Şekil 1.3'teki gibi verilmiştir [6].



Şekil 1.3. Veri kapsüllemesi [6]

1.1.4. Kablo Standartları

Ağların yapısına göre kullanılabilecek değişik kablo standartları bulunmaktadır. Bu tez kapsamında bakır kablo çeşitlerinden olan UTP (Unshielded Twisted-Pair) kablolar, STP (Shielded Twisted-Pair) kablolar ve fiber optik kablo standartları hakkında bilgi verilecektir.

Günümüz LAN ortamlarında en çok kullanılan bakır kablo çeşidi UTP ve STP kablolardır. UTP kablolar izole edilmediği için STP kablolara oranla çevre etkenlerden daha çok etkilenirler. Günümüzde en çok kullanılan UTP ve STP bakır kablo standartları Tablo 1.2'de verilmiştir [7].

Tablo 1.2. Bakır kablo standartları [7]

Kategori	Maximum Veri Hızı	Kablo Tipi	Kullanıldığı Uygulamalar
CAT 1	1 Mbps	UTP, STP	Eski telefon hatları
CAT 2	4 Mbps	UTP, STP	Token Ring ağlar
CAT 3	10 Mbps	UTP, STP	Token Ring ağlar
CAT 4	16 Mbps	UTP, STP	Token Ring ağlar
CAT 5	100 Mbps	UTP, STP	Ethernet, Fastethernet
CAT 5E	1 Gbps	UTP, STP	Ethernet, Fastethernet, Gigabitethernet
CAT 6	1 Gbps	UTP, STP	Gigabitethernet, 10 GigabitEthernet
CAT 6A	10 Gbps	STP	Gigabitethernet,10 GigabitEthernet
CAT 7	10 Gbps	STP	Gigabitethernet,10 GigabitEthernet
CAT 7A	40 Gbps		40 Gigabitethernet

Bakır kabloların yanı sıra daha hızlı veri iletimi sağlayan ve verileri ışık yoluyla yollayan fiber optik kablolar günümüz ağlarında sıkça kullanılmaktadır. Günümüzde en çok kullanılan fiber optik kablo standartları Tablo 1.3'te belirtilmiştir [7].

Tablo 1.3. EN 50173 sınıflandırmasına Göre fiber kablo limit değerleri [7]

EN 50173 Sınıflandırması	**OS1**	**OS2**	**OM1**	**OM2**	**OM3/OM4**
	SM -9/125 µm	SM -9/125 Mm	MM - 62.5/125µm	MM -50/125 µm	MM -50/125 µm
1000 BASE-SX					
-850nm			275 m	550 m	550 m
1000 BASE-LX					
1300nm			550 m	550 m	550 m
-1310nm	2,000 m	5,000 m			
10 GBASE-SR/SW					
-850nm			32 m	82 m	300/400 m
10 GBASE-LX4					
-1300nm			300 m	300 m	300 m
-1310nm	2,000 m	10,000 m			
10 GBASE-LR/LW					
-1310nm	2,000 m	10,000 m			
10 GBASE-ER/EW					
-1550nm	2,000 m	22,250 m			
40 GBASE-LR4					
-1310nm	4,700 m	10,000 m			
40 GBASE-SR4					
-850nm					100/150 m
100 GBASE-SR10					
-850nm					100/150 m
100 GBASE-LR4					
-1310nm	10,000 m	10,000 m			
100 GBASE-ER4					
-1550nm	16,000 m	40,000 m			

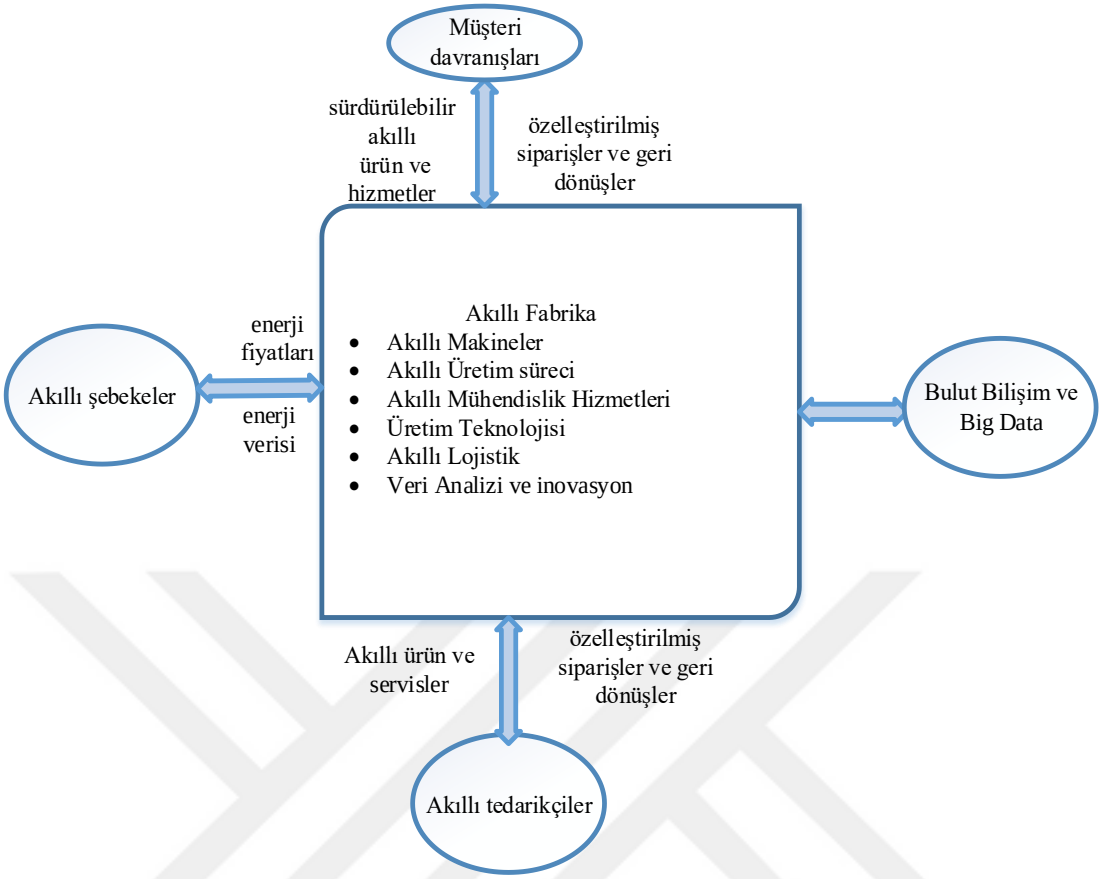
1.1.5. Ağ Cihazları

Genel olarak bir ağı genişletmek, yönetimini kolaylaştırmak, güvenliğini sağlamak gibi amaçlarla ağ içinde kullanılan bir takım cihaz vardır. Bu bölümde bu cihazların kısaca tanımları yapılacaktır.

- ***Hub:*** OSİ refrans modelinde 1. katmanda çalışan en basit ağ cihazıdır. Kendisine bağlanan tüm cihazlara paylaşılan bir yol sunar
- ***Switch (Anahtarlama Cihazı):*** OSI referans modelinin 2. ya da 3. Katmanlarında çalışabilir. Kendisine bağlı cihazların MAC adreslerini öğrenir ve öğrendiği bu bilgiye göre anahtarlama yaparak cihazların iletişime geçmelerine olanak sağlar.
- ***Router (Yönlendirici):*** OSİ referans modelinin 3. Katmanında çalışır. Bir ağdan daha uzak bir ağa erişebilmek için gidilebilecek en kısa yol seçimini yapar ve paketleri bu yol seçimine göre yönlendirir. En kısa yol seçimi esnasında routing protokolleri kullanılır.
- ***Firewall (Güvenlik Duvarı):*** internet ve kurumların ağları arasına kurulan ve her iki taraftan da gelecek zararlı yazılımları (virüs, botnet vb..) önleyen ağ cihazlarıdır. Güvenlik duvarının efektif kullanılması için yerel ağ ve internetten gelen tüm trafiği üzerinden geçirmelidir. Daha sonra ağ güvenliği için kurum tarafından belirlenmiş ölçütler ile erişim kuralları tanımlanır.

1.2. Literatür Özeti

Endüstri 4.0 devrimi ile birlikte Nesnelerin interneti kavramı günlük hayatta karşımıza çıkmıştır. IoT kavramı fabrikaların, kurumların, hükümetlerin sistemlerini güncellemeye teşvik etmiştir. Bu ile birlikte şirketler akıllı fabrikalar oluşturma vizyonu geliştirmektedirler [1]. F.Shrouf ve arkadaşları akıllı fabrikalarda IoT tabanlı enerji yönetimini desteklemek için bir yaklaşım önererek IoT'nin enerji verimliliğini nasıl artıracağını göstermişlerdir. Şekil 1.4'te F.Shrouf ve arkadaşlarının IoT tabanlı akıllı fabrikalar için geliştirdikleri referans mimari verilmiştir [1].



Şekil 1.4. F.Shrouf ve Arkadaşlarının Iot Tabanlı Referans Mimari [1]

F. Wortmann yaptığı çalışmada IoT çözümlerinin tipik olarak fiziksel şeyleri donanım ve yazılım biçiminde IoT ile birleştirilebildiğini göstermiştir. Sonuç olarak, bir şeyin birincil temel işi yada işlevleri, global düzeyde erişilebilen ek IoT temelli dijital hizmetler ile geliştirilebilir [2]. F. Wortmann Şekil 1.5'te, böyle bir değer yaratma mantığını göstermektedir. Örneğin, Şekil 1.5'e göre bir ampulün esas görevi, bir yerde ışık sağlamaktır. Bu aşamadan sonra ampul IoT teknolojisi ile geliştirilir ve ampüle insan varlığını algılayıp saldırı durumunda alarm üretmesini sağlayabilecek ek özellikler katılarak ampül düşük maliyetli bir güvenlik cihazına dönüştürülebilir. Benzer şekilde, bir deponun asıl işi temelli işlevi, depolama kapasitesi sağlamaktır. Ancak, depo IoT teknolojisi ile zenginleştirildiğinde ayrıca kendi ağırlığını ölçebilir ve izleyebilir, böylece düşük stok seviyelerini tespit edebilir ve otomatik yenileme hizmeti sunabilir. Bir traktörün iş temelli işlevi, diğer tarım ekipmanlarını çekmek için olabilirken, traktör ile IoT arasındaki bağlantı, BT temelli tahmini bakım ve optimizasyon hizmetlerini kolaylaştırabilir [2,3].

NESNE	+	Yazılım & Donanım	=	Fiziksel veya Lokal Fonksiyonlar	+	Bilişim Tabanlı Servisler
Ampül		IoT	⇨	Işık		Güvenlik
Çöp		IoT	⇨	Yük Kapasitesi		Otomatik ikmal
Traktör		IoT	⇨	Çekme Aracı		Tahmini bakım, optimizasyon

Şekil 1.5. IoT Ürün ve Hizmet Mantığı [2,3]

IoT devrimi, hizmet, üretim, lojistik gibi birçok ortamdan verilerin toplanmasını ve iyi bir şekilde analiz edilebilmesini mümkün kılmıştır. Endüstri 4.0 genel olarak üç yapıdan oluşmaktadır [4]. Bunlar;

1) Nesnelerin interneti

2) Hizmetlerin interneti

3) Siber fiziksel sistemler

Endüstri 4.0'ın en büyük avantajlarından birisi; sistemlerin izlenebilmesine ve bu sayede arıza teşhisinin kolaylıkla yapılabilmesine olanak sağlamasıdır. Bunun sonucunda yüksek verimli bir alt yapı ile hizmet verilebilmektedir. Bu tez çalışmasında Endüstri 4.0 yapılarından olan hizmetlerin interneti yapısı ile çalışılacak olup sistem ve ağ verimliliğinin en yüksek dereceye getirilmesi amaçlanmaktadır.

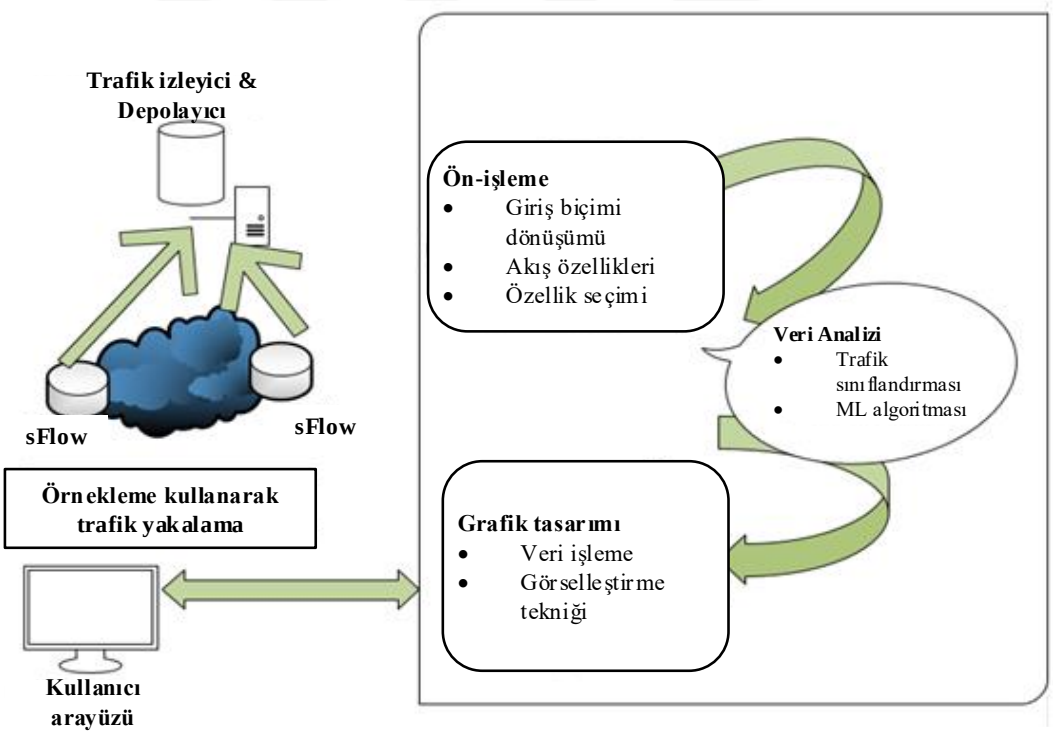
Günümüzde hemen her kurumda bilgisayar ağları çok gelişmiş ve karmaşık bir hale gelmiştir. Aynı network içinde birden fazla uygulamanın kullanılması ağ sistemlerini yormakta ve bu sistemlerin yönetimini zorlaştırmaktadır. Ağ yöneticilerinin en önemli görevlerinden biri, ağ üzerinde oluşan sorunların tespit edilip sorunun kısa zamanda çözülmesini sağlamaktır [8].

Ağ yöneticileri sorunlar oluşmadan sorunları tespit edebilmek için ağ izleme teknolojilerinden sıkça yararlanırlar. Shubhangi U. Pawar ve arkadaşları ağ içerisinde çevirim dışı çalışan bir ajan ile ağ içerisindeki cihazların çalışan işlemlerini, servisleri, cihazların CPU'ları gibi birçok değer elde etmiş ve bunlardan faydalanarak şüpheli davranışları tespit etmeye çalışmışlardır. Bu sistemi yaparken JADE multi agent framework kullanılmışlardır [9].

Umar Manzoor ve SamiaNefti ajan bazlı bir ağ izleme sistemi geliştirmiş ve bu sisteme Agent Based Activity on Network (ABSAMN) ismini vermişlerdir. Bu ağ izleme sistemi ile birlikte Umar Manzoor ve SamiaNefti sistem konfigürasyonlarının, sistem yöneticilerinin aktivitelerini ve internet kullanımını monitör edebilmişlerdir [9,10].

Soya Park KAIST ve arkadaşları Route Convergence Visualizer (RCV) isimli bir sistem geliştirmiş ve bu sistemle networkte routing komşuluklarının oluşmamasının nedenini bulmayı amaçlamışlardır [11].

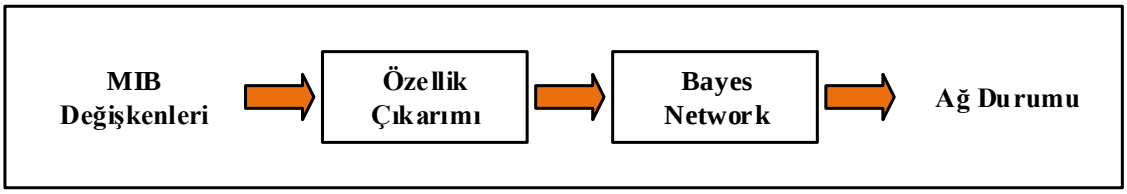
Meryem Elbaham ve arkadaşları trafiği tanımlamak, sınıflandırmak ve karakterize etmek için bir yöntem önermişleridir. Bu amaçla bir trafik görüntüleme mekanizması geliştirmişlerdir. Geliştirilen bu yöntem ile networkteki geçici durumlar, akış bazlı analizler, uygulama sınıflandırmaları ve port aktiviteleri görüntülenebilmiştir. Bu yönteme ait mimari Şekil 1.6'da verilmiştir [12].



Şekil 1.6. Trafik Görüntüleme Mimarisi [11].

Choi ve arkadaşları node ve link hatalarını SNMP protokolü ile alıp bu hataların network performansı üzerindeki etkilerini incelemişler [13]. Hood ve arkadaşları SNMP ile ağ cihazlarının loğlarını toplamış ve topladıkları bu verileri Bayes Networks algoritmasından

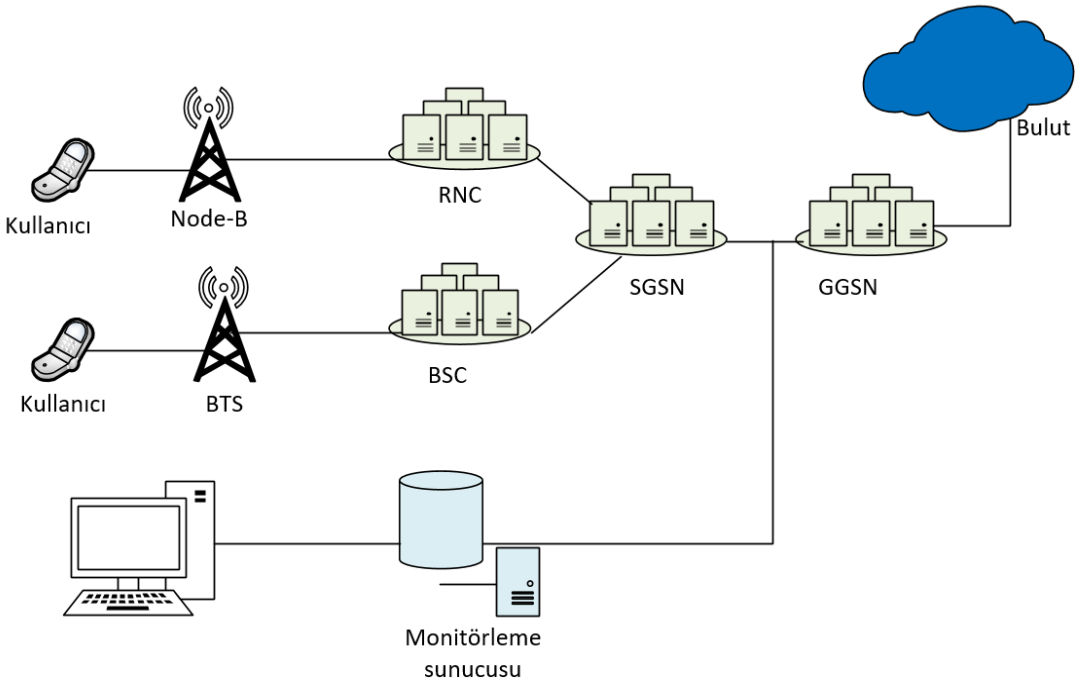
geçirmişlerdir. Böylelikle network arızası olmadan sorunu daha önceden çözebilme imkânı bulabilmişlerdir [14]. Hood ve arkadaşlarının kullandığı algoritma Şekil 1.7'de verilmiştir.



Şekil 1.7. Hood Ve Arkadaşlarının Yaptığı Çalışmanın Akış Diyagramı [14]

Wang ve arkadaşları network yönetimini hata yönetimi, performans yönetimi, trafik yönetimi, konfigürasyon yönetimi, güvenlik yönetimi ve yönlendirme yönetimi olarak 6 parçaya ayırmış ve buna göre bir çözüm geliştirmişlerdir [15].

Barış Kurt ve arkadaşları yüksek hızlı telefon şebekelerinde ağ izleme yaparak, uygulama kullanım analizleri, güvenlik analizleri, altyapı planlaması gibi analizler gerçekleştirmişlerdir. Sistemin fiziksel implemantasyonu Şekil 1.8'de gösterilmiştir [16].



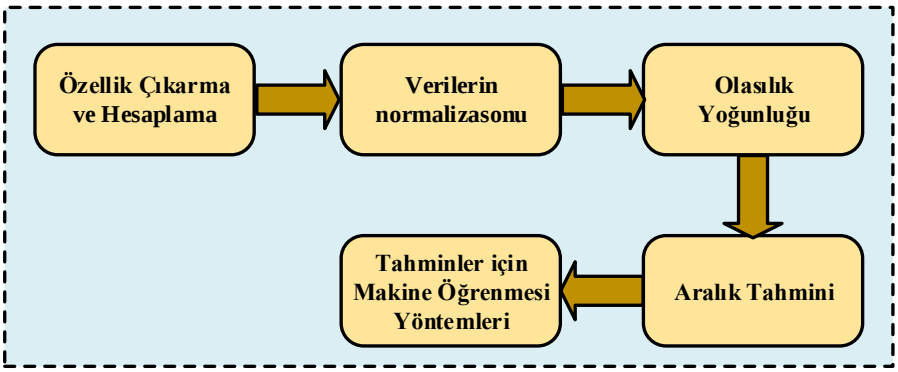
Şekil 1.8. Monitörleme Sisteminin Fiziksel Konumlandırması [16]

Y Zhang ve Q Yuan yaptıkları çalışmada, gürültülerden ya da iletim hatalarından dolayı iletilen veri ile alıcı tarafta alınan verinin aynı olmadığı durumları en aza indirgemek

amacıyla veri iletimi esnasında CRC hata denetim metodunu kullanmışlardır. CRC, bir çeşit yüksek performans ve kolay gerçekleştirilen kontrol kodudur ve yüksek güvenilirlikte iletim hatalarını tanıyabilir. Yaptıkları simülasyon sonuçlarına göre, CRC bazlı çoklu bit hata düzeltme yönteminin, bit hata oranını ve çerçeve hata oranını etkili bir şekilde artırabildiğini göstermişlerdir [17].

Literatürde ağların performanslarını ölçmek için, ağlara yapılan saldırıların tespiti için, ağ cihazların durumlarının kontrolü için makine öğrenmesi yöntemleri kullanılarak birçok çalışma yapılmıştır.

Neha Bansal ve Rishabh Kaushal yaptıkları çalışmada makine öğrenme algoritmalarını kullanarak ağdaki İnternet erişim davranışını analiz edilmiş ve bazı tahminler yapılarak, ağ yöneticilerinin kullanabileceği bir yöntem sağlanmıştır. Çalışmada squid proxy sunucusu kullanılmış ve bu sunucuda şebeke ve kullanıcı düzeyinde ağ trafiğinin çeşitli özellikleri hakkında analizler yapılmıştır. Çeşitli özellikler için en olası değer aralığı tahmin edilmiş ve normal ağ erişim özellik değerlerinden sapan IP adreslerini belirlenmiştir. Bundan sonra, elde edilen veri kümesinde dört farklı makine öğrenme algoritması uygulanmış ve bu algoritmalar TP, FP, TN ve FN gibi çeşitli sınıflandırma matrislerinde karşılaştırılmıştır. Sonuçta, Karar Ağacı algoritmasında % 95'e yakın bir genel doğruluk sağladığını, buna karşılık Naive Bayes ve DVM'in genel bir doğruluk oranının % 85 civarında olduğu gözlemlenmiştir [20]. Bu çalışmada kullanılan akış diyagramı Şekil 1.9'da verilmiştir [20].



Şekil 1.9. Neha Bansal Ve Rishabh Kaushal' In Kullandığı Yönteme Ait Diyagram [20]

Oleksii Sheremet, Oleksandr Sadovoy çalışmalarında telekomünikasyon ağlarının izlenmesi için destek vektör makinesi metodu kullanılmıştır. Yapılan çalışmalar sonucunda DVM algoritması sonucunda maksimum tahmini hata oranı 0,37 olarak ölçülmüştür.

Belirlenen hata değeri olan 0,1 değerine çok yakın bir hata tespiti yapan DVM metodunun, yapılan çalışma için çok iyi bir metod olduğu kanısına varılmıştır [21].

Raguram Sasisekharan ve arkadaşları yaptıkları çalışmada makine öğrenmesi yöntemlerini, AT & T'nin dijital iletişim ağındaki kronik iletim hatalarının tespit edilmesi için AT&T ağında uygulamışlardır. Büyük hacimli verilere bir pencere tekniği uygulanmış ve bu veriler analiz edilmiş ve karar kuralları uygulanmıştır. Kronik problemlerinin öngören bir dizi koşul tespit edilmiştir. Yeni yaklaşımı kullanarak düzenli aralıklarla ağın sürekli izlenmesi sayesinde, belirli kronik sorunların ilerledikçe birkaç yerel ağ eğilimini tespit edebilmiştir. Rule induction (Kural indüksiyonu) metodu bu uygulama için tercih edilen öğrenme yöntemi olduğu sonucuna varılmıştır. Ayrıca rule induction metodu dışındaki diğer makine öğrenmesi metodlarının hata oranları Tablo 1.4'te verilmiştir [22].

Tablo 1.4. Literatürdeki Çalışmanın Hata Oranları [22].

Metod	Hata oranı (%)
Prior	6.4
Linear Discriminant	6.4
Nearest Neighbor	5.6
Neural Networks	4.7
Decision Tree	4.5
Decision Rules	4.4

2. AĞ ÜZERİNDE OLUŞABİLECEK HATALAR

Ağ üzerinde meydana gelen hatalar, ağın çalışma performansını etkiler. Dolayısıyla ağ üzerindeki hataların karakteristiklerinin bilinmesi ağdaki sorunların kaynaklarının bulunmasında önemli derecede bir etkiye sahiptir. Bir ethernet ağında oluşan sorunların genelde iki problemden dolayı meydana gelir. Bunlar;

- Fiziksel Katman problemleri: Ağ yöneticilerinin en çok karşılaştığı problem tiplerindendir. Bu problemler genel olarak yanlış sonlandırma, kablo ezilmesi, kablo bant genişliğinin yetersiz olması gibi durumlardan dolayı oluşmaktadırlar.
- Layer 2 ve Layer 3 Problemleri: Genel olarak switchlerde veya routerlarda yapılan yanlış VLAN (Virtual Local Area Network), QoS (Quality of Service) konfigürasyonları, hasarlı ağ cihazı, yetersiz bant genişliği nedeniyle meydana gelen problemlerdir.

Bu tez kapsamında fiziksel katman problemlerinin otomatik olarak çözülmesi amaçlanmaktadır. Fiziksel katman problemlerinin belirlenebilmesi için ağ yöneticilerinin sıklıkla kullandığı, ağ cihazı üzerinden alınabilecek bir takım ağ hataları bulunmaktadır. Bölüm 2.1' de verilen ağ hatalarının ağ yöneticileri tarafından bilinmesi ve doğru analizi ağ verimliliği açısından oldukça büyük öneme sahiptir. Bu nedenle ağ yöneticilerinin hizmet kesintisi olmadan bu hataları bilmesi gerekmektedir.

2.1. Fiziksel Katman problemleri

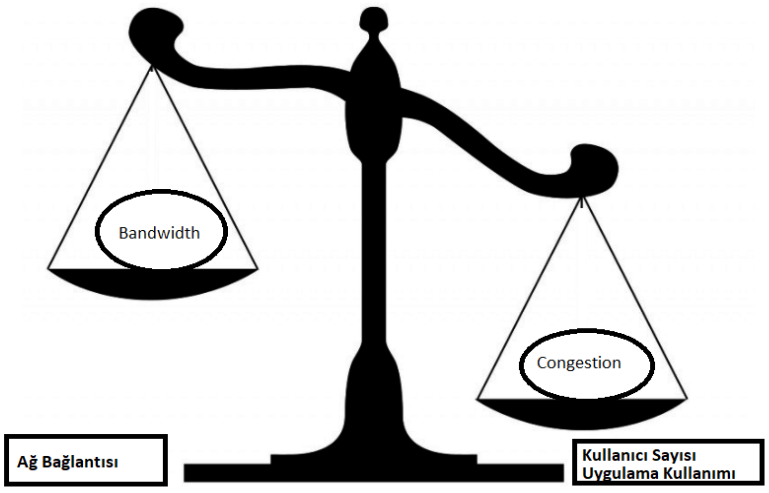
OSI referans modelinin ilk katmanı olan fiziksel katman bilgisayar ağlarında haberleşme için kullanılacak altyapıyı tanımlar. Gönderilecek verinin gönderilme biçimi (ışık, elektriksel, kablosuz vb.) fiziksel katmanda belirlenir. Bu katmanda veriler bitler halinde aktarılır ve bitler 1 ve 0 şeklinde kodlanır. Fiziksel katmanda meydana gelebilecek bir sorun bütün veri haberleşmesini etkilediğinden, ağ yöneticileri bu katman problemlerini en aza indirgemek zorundadırlar. Bu katmanda meydana gelen sorunlar, sabit sorunlar olmamak ile birlikte çok farklı şekillerde ortaya çıkabilmektedirler. Bu tez çalışmasında ağ yöneticilerinin en fazla karşılaştıkları fiziksel katman problemleri ele alınmıştır.

2.1.1. Congestion

Anahtarlama cihazının özellikle uplink noktalarında aşırı yük nedeniyle görülen paket kayıplarıdır [18]. Kaynaklardaki sınırlamalardan dolayı veri paketleri vericiden alıcıya doğru iletilirken herhangi bir noktada tıkanıklığa maruz kalabilirler. Bu tıkanıklıklar, paket kayıplarına paket iletiminde gecikmelere neden olabilirler. Bu nedenle ağ kaynaklarının verimli kullanımı açısından tıkanıklıkların kontrol altına alınması önem arz etmektedir.

Senthilkumaran ve diğ. [19] yaptıkları çalışmada ad hoc ağlardaki dinamik tıkanıklık saptama ve kontrol yönlendirme (DCDR) için, düğüm düzeyinde ortalama kuyruk uzunluğunun tahminlerine dayanan bir yöntem önermişlerdir. Ortalama sıra uzunluğunu kullanarak, bir düğüm mevcut tıkanıklık seviyesini tespit eder ve komşularına bir uyarı mesajı gönderir. Komşu düğümler, daha sonra, hedefin tıkanık olmayan alternatif bir yolunu bulmaya çalışmışlardır. Geçici ağlarda tıkanıklığı kontrol etmeyi destekleyen bu dinamik tıkanıklık tahmini mekanizması, MANET içinde güvenilir iletişim sağlamıştır [19].

MM Monowar ve arkadaşları çoklu yol tıkanıklığını kontrol etmek için bir plan önermişlerdir. Sıkışıklığı tespit etmek için paket hizmet oranını ve bu metriğe dayanan hop-by-hop çoklu yol tıkanıklığı kontrolü kullanılmıştır [20].

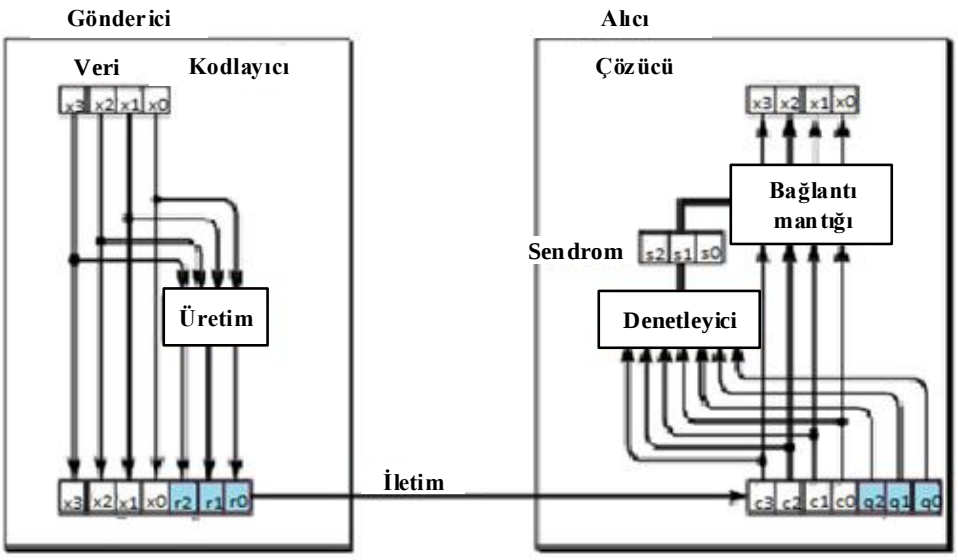


Şekil 2.1. Congestion ve bantgenişliği arasındaki ilişki [38]

Şekil 2.1'de görüleceği üzere denge, kullanıcıların ve uygulamaların ihtiyaçlarını karşılamak için yeterli bant genişliğine sahip olmasına bağlıdır [38].

2.1.2. Rx_CRC

Port tarafından alınan, ancak orijinal frame'e oranla kötü FCS (frame check sequence) değeri içeren doğru uzunlukta çerçeve sayılarının toplamıdır [21]. CRC, dijital iletişimde, verilerin paketler halinde gönderilmesi esnasında meydana gelen hataları tespit eden bir teknolojidir. Token Ring, Ethernet, Asenkron Transfer Modu (ATM) ve Senkron Optik Ağ (SONET) gibi verileri paketler halinde ileten teknolojiler hata algılama ve düzeltme tekniklerini kullanırlar. Bir veri paketi bir kaynaktan hedefe doğru gönderildiğinde, alıcı, iletimin bütünlüğünü doğrulamak için CRC'yi kullanır [22]. CRC, iletilen veri bitlerinde çerçevenin sonuna eklenir. Alıcı aldığı çerçeve üzerinde CRC'yi hesaplar ve hesaplanan CRC'yi, vericiden gönderilen çerçevenin sonuna eklenen CRC ile karşılaştırır. Eşleşme olursa frame alıcı tarafta alınır, Eşleşme gerçekleşmezse frame bozulmuştur [23]. Ethernet ağlarında veri iletimi esnasında, veriye eklenen CRC bilgisi Şekil 2.2'de şematik olarak verilmiştir [24].



Şekil 2.2. Veri iletimi esnasında CRC bilgisi

2.1.3. Tx_Error

Ağ hatalarından dolayı bağlantı noktası tarafından tamamen iletilmeyen toplam çerçeve sayısıdır [26]. P.Chatzimisios ve diğ. [25] yaptıkları çalışmada IEEE 802.11a protokolünün

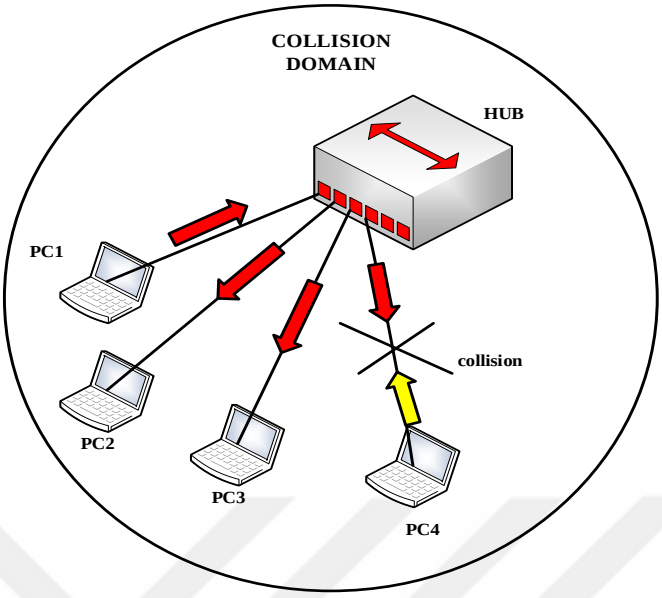
performansı için, geliştirilmiş bir analitik model önermişlerdir. Çalışmada OPNET simülasyon programı kullanılmıştır. Önerilen model, temel erişim için çıktı verimliliği, ortalama paket gecikmesi, paket düşüş olasılığı ve ortalama süreyi hesaplamıştır. Çalışmada protokol performansının bit hata oranına (BER) bağlı olduğu analitik sonuçlar ile elde edilmiştir. Çalışma sonucuna göre BER arttıkça paket gecikmesi ve paket düşüş olasılığının arttığı gözlemlenmiştir. Ayrıca veri hızının da paket gecikmesi ve paket kayıpları üzerinde önemli bir rol üstlendiği gözlemlenmiştir [25].

2.1.4. Rx_Fragmentation

İnternet Protokolü ağlar arasında haberleşmeyi sağlayan bir teknolojidir. Haberleşme esnasında önem arz eden parametrelerden biri de Maximum Transmission Unit (MTU) olarak bilinen gönderilebilecek en büyük veri boyutudur. Bu parametre ağ çeşidine göre değişen bir parametredir. Örneğin Ethernet ağları için MTU boyutu 1500 Byte iken FFDI teknolojisi için bu değer 4352 Byte' tır. Ağ standartlarının sağladığı MTU değerinden daha büyük boyutta veri paketi geldiğinde, veri paketlerinin ağlar arasında dolaşabilmesi için iletim esnasında paketler yeniden boyutlandırılır. Bu işleme Fragmentation denir. Bu özellik farklı ağ türleri arasında iletişim yapmaya olanak sağlamıştır. Fragmente edilmiş veri paketinin alıcı tarafta birleştirilmesi için her veri parçasına IP header (IP başlığı) eklenir [27]. Rx_Fragmantation, ağ hatalarından dolayı anahtarlama cihazının portunda görülen normal boyutlarda olmayan parçalanmamış frame sayısıdır [26]. Anahtarlama cihazları üzerinde Rx_Fragmentation değerinin artması ağ yöneticilerine çok fazla büyük boyutta veri paketi geldiğini gösterir.

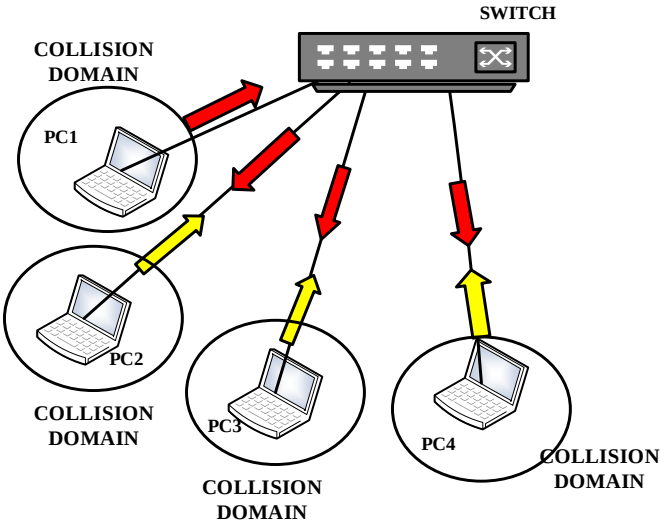
2.1.5. Collision

Anahtarlama cihazı portunda görülen toplam çarpışma sayısıdır [26]. Bilindiği üzere switch, gelen trafiği sadece gitmesi gereken yere gönderken, hub gelen trafiği tüm portlarına iletir. Şekil 2.3'te 4 portluk bir hub'a 4 adet bilgisayar bağlanmıştır. Bu durumda bilgisayarlardan biri başka bir bilgisayara veri gönderdiğinde hub bu veriyi tüm portlarından yayar ve hub' a bağlı tüm bilgisayarlar bu veriyi dinlemek zorundadır. Bu esnada başka bir bilgisayar veri gönderdiğinde hatta collision (çarpışma) meydana gelir ve veri kaybı oluşur.



Şekil 2.3. Hub ve Collision Domain İlişkisi

Ağda oluşabilecek çarpışma sayısını en aza indirgemek için collision domain sayısını artırmak fakat collision domaindeki bilgisayar sayısını azaltmak gerekmektedir. Bu durumun sağlanabilmesi için anahtarlama cihazları kullanılmaktadır.



Şekil 2.4. Anahtarlama Cihazı ve Collision Domain İlişkisi

Şekil 2.4'te anahtarlama cihazına bağlanan 4 adet bilgisayar bağlanmıştır. Bu durumda herhangi bir bilgisayar başka bir bilgisayara veri gönderdiğinde anahtarlama cihazı bu veriyi

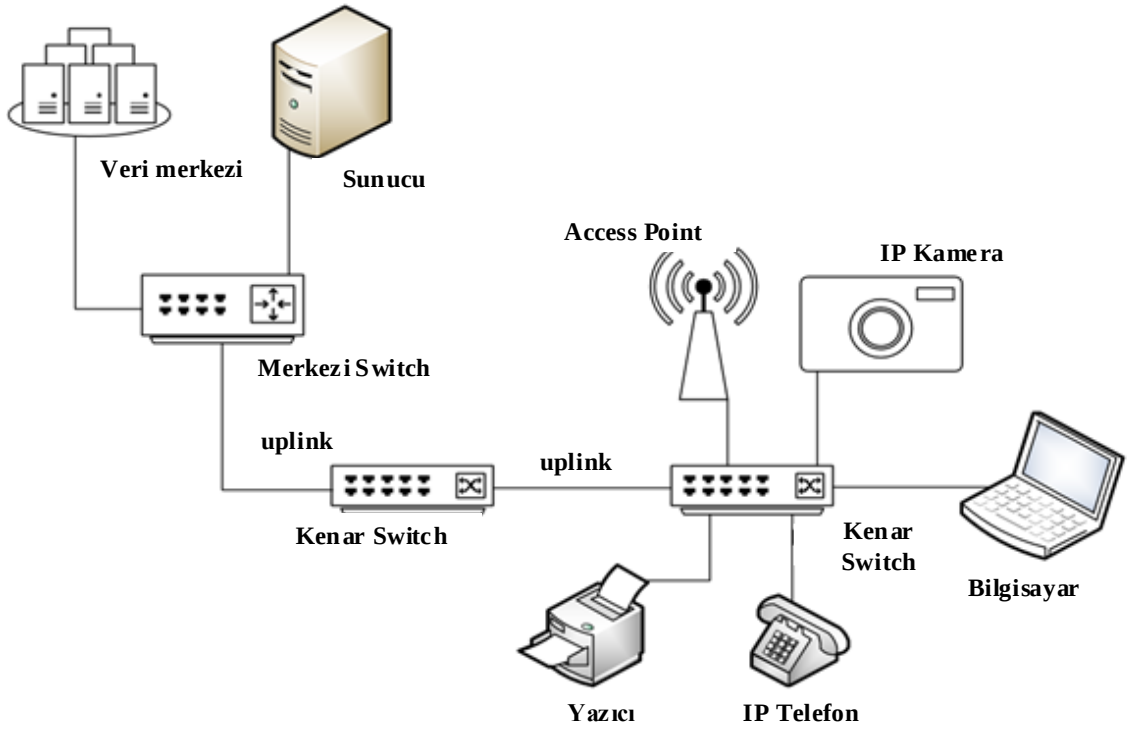
sadece ilgili porta bağlı bilgisayara gönderirdi. Bu da tüm bilgisayarların aynı anda veri alışverişi yapabilmelerine olanak sağlamıştır. Bu durumda anahtarlama cihazının her bir portu farklı birer Collision Domain oldu. Buna göre Şekil 2.4'te de görüleceği gibi 4 adet collision domain meydana gelmiştir. Bu da anahtarlama cihazlarının, collision domain sayısını artırdığını fakat collision domain boyutunu azalttığını göstermektedir.

2.1.6. Rx_Over

Port tarafından alınan paket boyutu 1522 byte'dan fazla olan frame sayısıdır [26].

3. ÖNERİLEN YÖNTEM

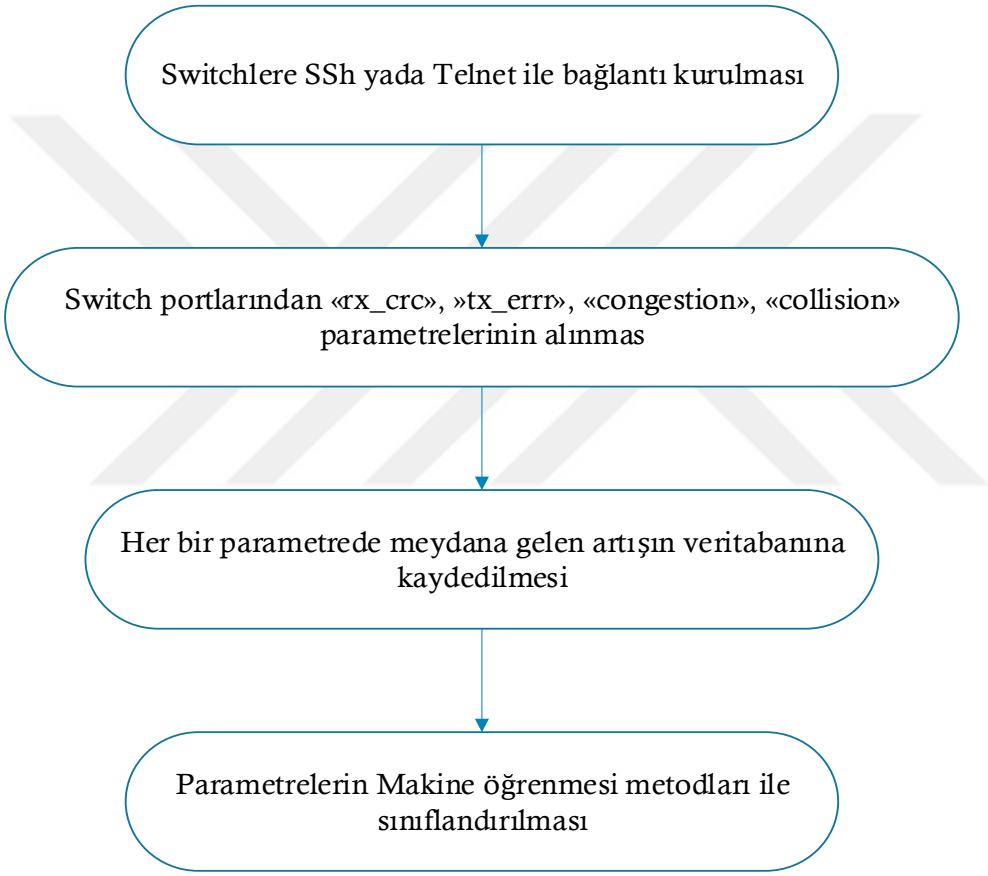
Bu çalışmada Fırat üniversitesi kampüs networkünde çalışacak bir uygulama geliştirilmiştir. Geliştirilen bu uygulama ile Fırat Üniversitesi kampüsünde bulunan switchlerin portlarındaki hatalar gözlemlenerek bir veri kümesi oluşturulmuştur. Bu çalışmada önerilen yöntemin uygulandığı kampüs network yapısı Şekil 3.1'de verilmiştir.



Şekil 3.1. Önerilen yöntemin uygulanacağı kampüs ağının yapısı

Şekil 3.1'de görüldüğü üzere kampüs ağı içinde access point, kamera, bilgisayar, yazıcı gibi birçok farklı cihaz bulunmakta ve birbirleri ile etkileşim halindedirler. Bu cihazların haberleşmesini etkileyen en önemli unsurlardan biri cihazlar arasındaki kabloların sağlamlığıdır. Bu çalışmada kablolarda oluşan arızaların herhangi bir hizmet kesintisine neden olmadan önce tespit edilebilmesi için ağ üzerindeki switchlerden çeşitli bilgileri almak amacıyla bir yazılım geliştirilecektir. Geliştirilen yazılım ile switchlere ssh ile bağlantı kurulacak ve switchlerdeki aktif portlara ait bazı hata parametreleri alınarak, bu hata parametrelerinden bir veri kümesi oluşturulacaktır. Oluşturulacak veri kümesinde belirtilen

“Rx_CRC”, “Tx_Errors”, “Collision” ve “Congestion” parametreleri incelenecektir. Bu parametreler cihazların ağ ile olan bağlantılarındaki hataları bulma açısından oldukça önemli parametrelerdir. Geliştirilen yazılımda switch portlarındaki bu parametreler alınarak bir veri dizisi elde edilecektir. Bu parametre sonuçlarına göre mevcut kablo yapısının sağlamlığı ve hattın yoğunluğu hakkında bir takım bulgular elde edilecektir. Eğer sorunlu bir durum var ise sorunun kaynağı tespit edilmeye çalışılacaktır. Bu çalışmada önerilen yöntemin akış şeması Şekil 3.2’de verilmiştir.



Şekil 3.2. Önerilen yöntemin akış şeması

Bu çalışmada, switchlerin portlarından okunan “Rx_CRC”, “Tx_Errors” ve “Collision” parametrelerinde oluşan artışlar cihazın ağ bağlantısında fiziksel arızanın olduğunu göstermektedir. “Congestion” parametresi ise hattaki düşen paket sayısını ifade etmekte ve oluşan artış miktarı hattın yoğun olarak kullanıldığını göstermektedir.

Bu tez kapsamında Fırat Üniversitesi kampüs ağında buluna switchlerden yararlanılmıştır. Üniversite kampüs ağında buluna switchlerin genel özellikleri aşağıda verilmiştir.

- 48 port gigabit Ethernet
- 4 port 1/10 gigabit Ethernet (SFP (small form pluggable))
- 850 W PoE
- 260 Gigabit Bandwidth
- 68000 MAC address
- 4094 VLAN desteği
- 64-bit MIPS Processor, 1 GHz clock, dual core
- 1GB ECC DDR3 DRAM
- 4GB eMMC Flash
- 4MB packet buffer

3.1. Makine Öğrenmesi

Bilimsel çalışmalarda sıkça kullanılan makine öğrenmesi algoritmaları verilerden kurallar ve davranışlar çıkartmamızı sağlar. Daha sonra verileri sınıflandırır ve bu yaptığı bu sınıflandırma ile veriler hakkında ileriye dönük tahminler yapılabilmesini sağlar.

İlk etapta elimizde ham veri bulunur. Bu ham veri içerisinde tutarsızlıklar ve farklılıklar içerdiği için bu verinin analiz edilmesi çok zordur. Ham veriden, verinin tamamını temsil edecek anlamlı bir veri kümesi elde etmek için bir takım ayıklamalar yapılır. Bu aşamadan sonra makine öğrenmesi algoritmalarının uygulanabileceği veriler elde edilmiş olur. Literatürde sıkça kullanılan bazı makine öğrenmesi algoritmaları şunlardır [28];

- Destek Vektör Makineleri
- Karar Ağaçları
- Naive Bayes Sınıflandırma Algoritması

3.1.1. Destek Vektör Makineleri

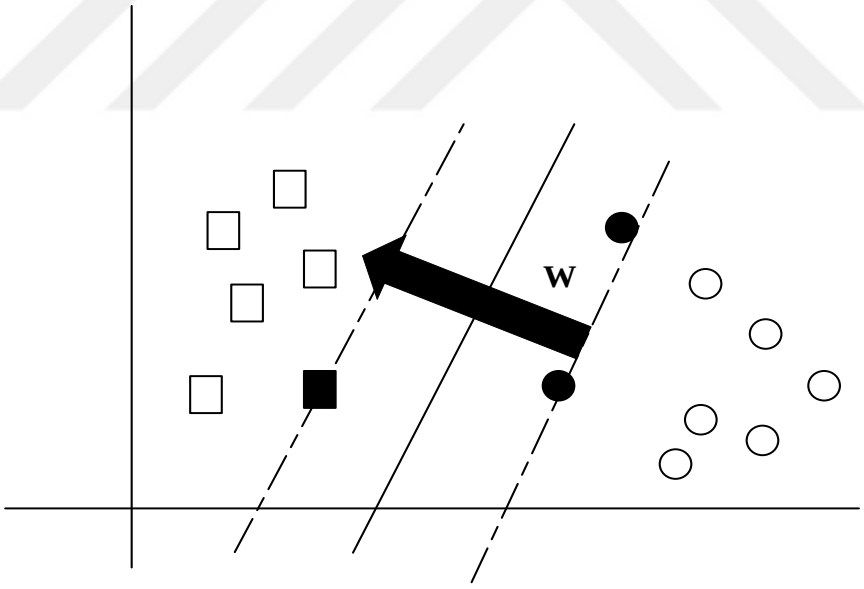
Temeli istatiksel yöntemlere dayanan güçlü bir sınıflandırıcıdır. Destek vektör makineleri DVM adı ile etiketlenmiş bir veri setine ihtiyaç duyar. DVM eğitim algoritması, yeni gelen her örneğin sınıflandırılması için bir model yapı oluşturur. DVM'de oluşturulan model, uzayda noktalar gibi örneklerin temsil etmektedir. Sınıflandırılan örnekler, geniş ve net bir

düzlem ile birbirinden ayrılır. Yeni örnekler aynı uzaya dâhil edilir ve hangi sınıflara ait oldukları hesaplanır [28].

DVM' de öğrenme verileri $x_i \in R^d,\ i = 1 \ldots . n\ , y_i \in \{-1, +1\}$ olmak üzere $\{x_i, y_i\}$ile etiketlenirse Şekil 3.3'de verilen iki boyutlu uzayda iki sınıfa ait veri kümesi örneği oluşur. Şekil 3.3'e göre kalın çizgi ayırıcı hiper düzlemi, w: normal vektörünü göstermiştir. Ayrıca taranmış veriler, destek vektörlerini göstermektedir. Hiper düzlemin pozitif ve negatif örnekleri ayırdığını kabul edersek, iki sınıflı doğrusal sınıflandırıcı probleminde hiper düzlemin normal vektörü w ve ofset değeri b ile tanımlanır. Buna göre karar sınırı $w^T x + b = 0$ doğrusudur. Bu durumda denklem 3.1 ve denklem 3.2'deki koşullar sağlanmalıdır [29].

$$w^T x_i + b \geq 1 \tag{3.1}$$

$$w^T x_i + b \leq -1 \tag{3.2}$$



Şekil 3.3. İki Boyutlu DVM [15]

Destek vektör makineleri başlangıçta birbirinden doğrusal bir şekilde ayrılmış iki sınıflı problemlerde kullanılmıştır. Daha sonra iki sınıflı problemleri temel alınarak çoklu sınıf problemlerinin çözümü için de kullanılmaya başlanmıştır. Bu amaç ile sınıf etiketleri genelleştirilmiş $y_i = sgn(w^T x_i + b)$ denklemi elde edilmiştir. Hiper düzlemin öğreneme verilerine olan en kısa uzaklık doğrusal sınıflandırıcının sınırı olarak kabul edilir. Destek

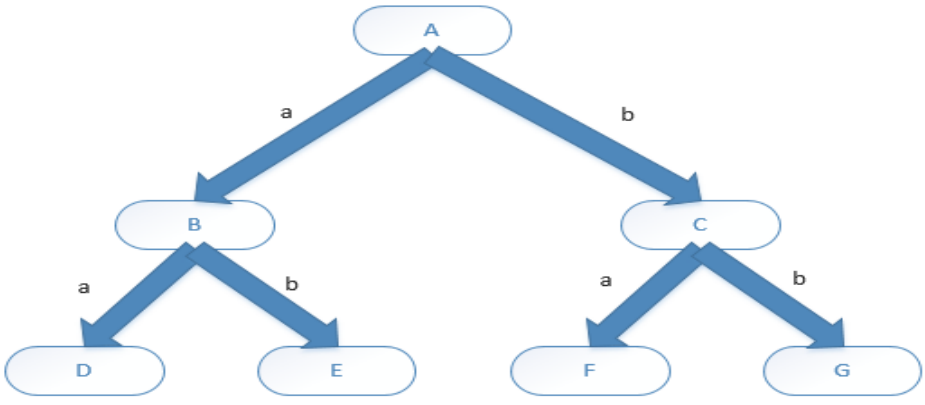
vektörleri ayırt etme yüzeyine en yakın olan vektörlerdir ve karar sınırı bu vektörlere göre belirlenir. DVM algoritması sınıflandırma yaparken ayırıcı hiper düzlemi kullanır ve böylece öğrenme hatasını en aza indirger. Sınıflar doğrusal olarak ayırt edilemiyorsa çekirdek fonksiyonları devreye girer ve bu şekilde sınıflandırma yapılır [29,30].

Literatürde ağ görüntüleme konusunda Destek vektörlerden yararlanılmaktadır. Oleksii Sheremet, Oleksandr Sadovoy çalışmalarında telekomünikasyon ağlarının izlenmesi için destek vektör makinesi metodu kullanılmıştır. Yapılan çalışmalar sonucunda DVM algoritması sonucunda maksimum tahmini hata oranı 0,37 olarak ölçülmüştür. Belirlenen hata değeri olan 0,1 değerine çok yakın bir hata tespiti yapan DVM metodunun, yapılan çalışma için çok iyi bir metod olduğu kanısına varılmıştır [31].

3.1.2. Karar Ağaçları

Karar ağaçları, eğitim ve testinin hızlı ve etkin olması sebebiyle sınıflandırmada sıklıkla kullanılan makine öğrenmesi yöntemlerden biridir [28]. Karar ağaçları ile sınıflandırma iki adımda gerçekleştirilir. İlk adımda ağaç oluşturulur. İkinci adımda bu ağaç yapısından sınıflandırma kuralları elde edilir.

Karar ağaçları temelde kök, dallar ve yapraklardan oluşmaktadır. Karar verme süreci kökten yaprağa doğrudur. Karar verme sürecinde dallar takip edilir. Şekil 3.4'te basit bir karar ağacı yapısı verilmiştir [32].



Şekil 3.4. Basit bir karar ağacı modeli

Karar ağaçları oluşturulurken kullanılan algoritmanın türü oldukça önemli bir parametredir. Kullanılan algoritma ağaç yapısını değiştirir. Ağaç yapısını değişmesi sınıflandırma sonucunu etkilemektedir [28].

Karar ağaçlarına dayalı olarak geliştirilen birçok yöntem bulunmaktadır. Bu yöntemler birbirlerinden kök, düğüm ve dallanma özelliklerine göre farklı sınıflara ayrılırlar. Yaygın olarak kullanılan yöntemler ID3, C4.5 ve C5dir. C4.5 algoritması, ID3 algoritmasının genişletilmiş hali olup, ID3 algoritmasını kapsamaktadır. Bu çalışmada C4.5 algoritması kullanılmıştır. C4.5 algoritması entropi tabanlı olup denklem 3.3'te verilmiştir [32].

$$H(S) = -\sum_{i=1}^{n} p_i \log_2(p_i) \tag{3.3}$$

n birbirinden bağımsız üretilen veri sayısı ($\{a_1, a_2, \dots, a_n\}$) olmak üzere, her veri için üretim olasılığı , p_i'dir. ID3, sınıflandırma işlemini yapan matematiksel bir algoritmadır. Dallanmaları belirleyen bir nitelik seçim ölçüsüdür. T, hedef niteliği ve X, hedef niteliği olmayan bir nitelik ise; T, X'e bağlı olarak $T_1, T_2, \dots, T_n$ alt kümelerine ayrılırsa, T'nin bir elemanının sınıfını belirlemek için gerekli bilgi, T_i'nin herhangi bir elemanının sınıfının belirlenmesinde gerekli olan bilginin ağırlıklı ortalamasıdır. O halde, T'nin sınıfını belirlemek için bilgi miktarı denklem 3.4'te verilmiştir [32].

$$H(X,T) = \sum_{i=1}^{n} \frac{\left|T_i\right|}{T} H(T_i) \tag{3.4}$$

En yüksek bilgi kazancına sahip parçanın bulunması için bilgi kazancı değeri hesaplanır. Bilgi kazancını hesaplamak için, orijinal bilgi ile parçalanma sonrasında elde edilen bilginin farkı alınarak bulunur. Kazanç denklemi denklem 3.5'te verilmiştir [32].

$$Gain(X,T) = H(T) - H(X,T) \tag{3.5}$$

Bilgi kazancı ölçütünün işe yaramadığı durumlarda bölünme bilgisi kavramı hesaplanmaktadır. Bu kavram bilgi kazancının bir çeşit normalizasyon işlemidir. Denklem 3.6'da bu işlem verilmiştir [32].

$$H(P_X, T) = -\sum_{i=1}^{k} \frac{|T_i|}{T} * \log_2\left(\frac{|T_i|}{|T|}\right) \quad (3.6)$$

Bölünme bilgisi kavramı, bilgi kazancından farklı olup T veri kümesinin, X nitelik işleminin k parçaya bölünmesiyle elde edilir. Buna göre kazanç oranı denklem 3.7'de verilmiştir [32].

$$GainRate(X, T) = \frac{InformationGain(X, T)}{DivisionInformation(X, T)} \quad (3.7)$$

Maksimum kazanç oranına sahip nitelik bölme niteliği olarak kullanılır. Bu işlem için öncelikle verilen sayısal eğitim verileri için eşik değeri bulunur. Bunun için tüm veriler sıralanır ve $\{v_i, v_2, \dots, v_n\}$ haline getirilir. Bu sıralama iki eşit parçaya ayrılır ve $\{v_1, \dots, v_i\}$ ve $\{v_{i+1}, \dots, v_n\}$ şeklini alır. Bu işlem sonucunda eşik değeri denklem 3.8'de verilmiştir [32].

$$t_i = \frac{v_i + v_{i+1}}{2} \quad (3.8)$$

Eşik değer bulunduktan sonra, verilen sayısal eğitim verileri nominal hale getirilir:

- $v_i < t_i$ şartına uyan veriler, "küçük"
- $v_i \geq t_i$ şartına uyan veriler ise, "büyük eşit"

şeklinde yeni değerlere sahip olur [32].

3.1.3. Naive Bayes Sınıflandırıcısı

Bir olasılıksal sınıflandırma olan Naive Bayes, değişkenler arasında olan ilişkiyi temel alır. Daha sonra verinin öğrenilmesi amacıyla, analiz ve tahmine dayalı sınıflandırma yapar. Modelin öğrenilebilmesi için eğitim verileri kullanılır. NB, kullanılan eğitim verileri ile her bir çıktının kaç kez meydana geleceği hesaplamaya çalışır. Hesaplamalar yapılırken olayın oluşma sıklığı her bir bağımsız değişkenin bağımlı değişkenlere oranının birleşimi ile meydana gelir. Veri kümesindeki tahminler için hesaplanan sıklık değeri kullanılır [33,34]. Naive Bayes sınıflandırıcısı yöntemiyle bir hedef niteliğin sınıf değerlerine ait olma olasılıkları bulunabilir [35]. NB sınıflandırıcısının çalışma prensibi şu şekildedir: $X =$

$(x_1, x_2, \dots, x_n)$ örnekler kümesini, $C = (c_1, c_2, \dots, c_n)$ veri kümesindeki sınıflar olmak üzere $P(c_i|X)$ değerinin maksimize edilmesi amaçlanmaktadır. Her c_i sınıfı için $P(c_i|X)$ değerinin maksimize edilmesi maksimum sonlu hipotez olarak adlandırılır. Bayes teoreminden denklem 3.9 elde edilir.

$$P(c_i \mid X) = \frac{P(c_i)P(X \mid c_i)}{P(X)} \tag{3.9}$$

Denklemde $P(X)$ tüm sınıflar için eşittir. Dolayısıyla sabit olarak alınabilir. Bu denklemde pay maksimum yapılmaya çalışılmaktadır. $P(X|c_i)$ değerinin hesaplanması çok fazla örneğin bulunduğu veri kümesinde zor olmaktadır. Bu nedenle sınıf şartlı bağımsızlık varsayımı meydana gelecektir. Bu varsayım, örneklerin arasında birbirine bağımlı ilişki bulunmayacağını belirtir. Dolayısıyla $P(X|c_i)$ değeri denklem 3.10'daki haliyle yazılabilir.

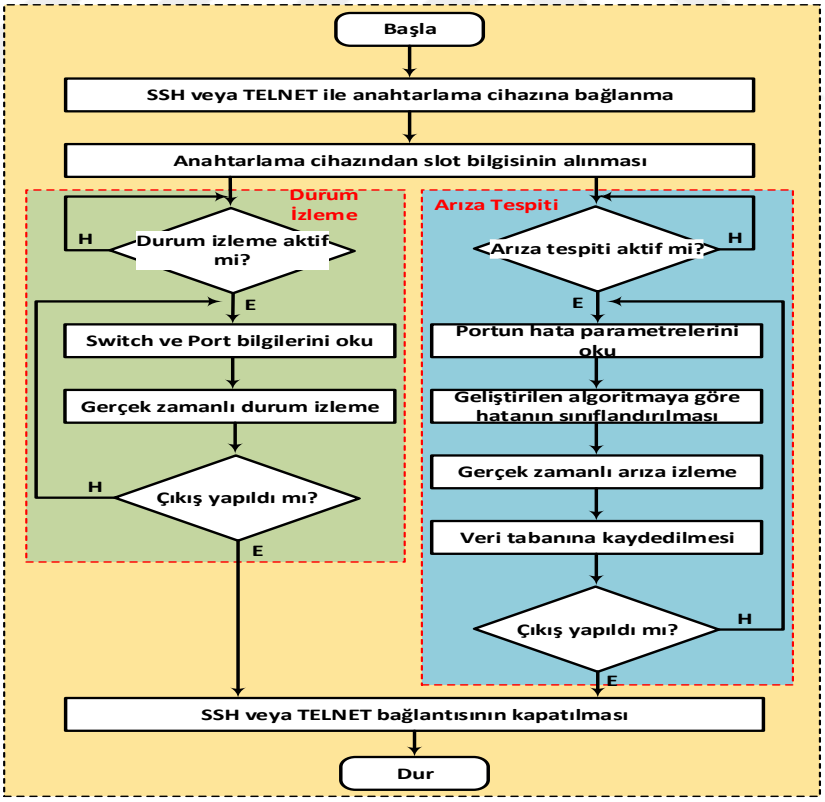
$$P(X \mid c_i) = \prod_{t=1}^{n} P(x_t \mid c_t) = P(x_1 \mid c_i)P(x_2 \mid c_i)\dots P(x_n \mid c_i) \tag{3.10}$$

4. UYGULAMA SONUÇLARI

Günümüzde ağ görüntüleme teknolojisi üzerine çalışan birçok yazılım bulunmaktadır. Bu yazılımlar genelde ağ cihazlarına ait güç, fan, sıcaklık gibi parametreleri izlerken, tez kapsamında geliştirilen yazılıma başka özellikler de eklenmiştir. Tez kapsamında geliştirilen yazılımın kabiliyetleri üç başlık altında incelenmiştir. Bunlar;

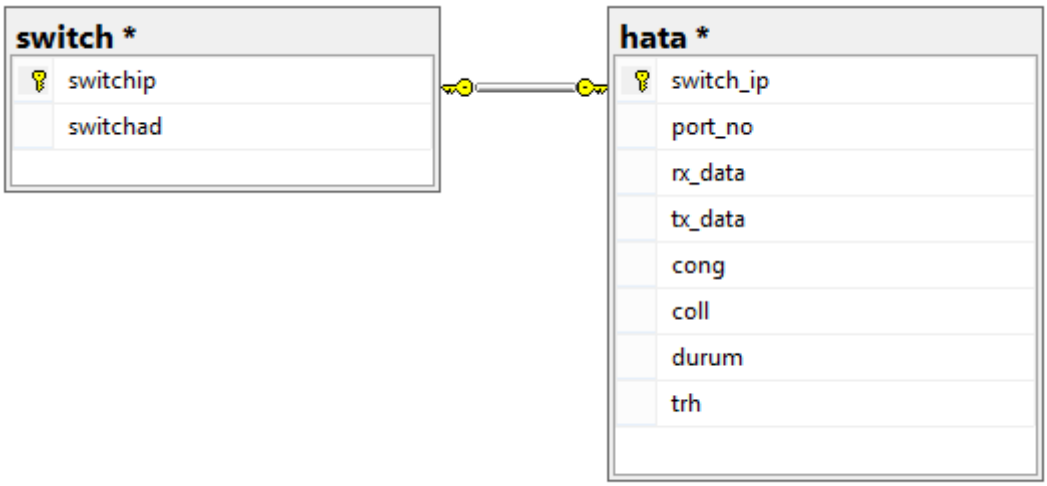
- Durum İzleme
- Arıza Parametrelerinin Eldesi
- Arıza Parametrelerinin Makine Öğrenmesi Metodları ile Sınıflandırılması

Bu tez çalışmasında anahtarlama cihazlarına bağlantı yapmak için geliştirilen yazılım Visual Studio 2013 ortamında C# programlama dilinde hazırlanmıştır. Geliştirilen yazılımın akış şeması Şekil 4.1'de verilmiştir.



Şekil 4.1. Geliştirilen yazılımın algoritması

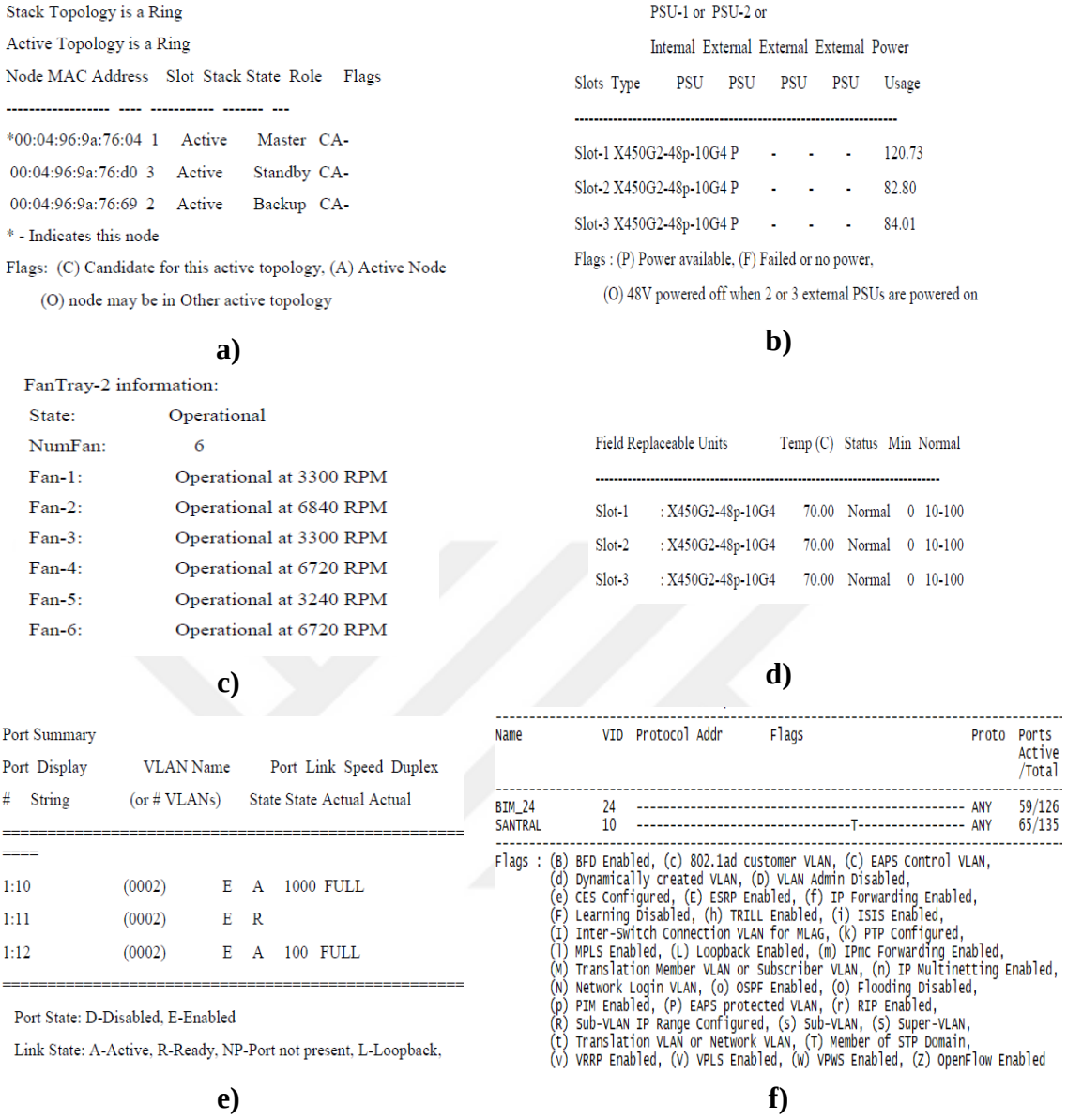
Anahtarlama cihazlarından alınan veriler SQL Server Management Studio versiyon 17 programı kullanılarak veri tabanına kaydedilmiştir. Bu tez çalışmasında “ariza_tespiti” adında veri tabanı oluşturulmuş, “switch” ve “hata” adında tablo oluşturulmuştur. Oluşturulan tablo Şekil 4.2’de verilmiştir.



Şekil 4.2. Veri tabanı diyagramı

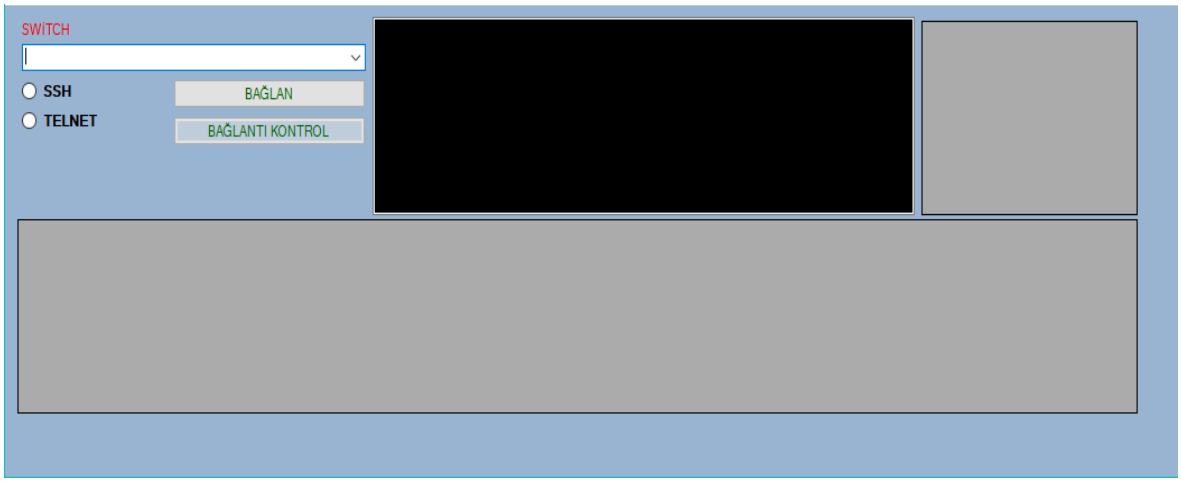
4.1. Durum İzleme

Ağ cihazlarının sağlıklı çalışabilmesi için cihazlara ait sıcaklık, fan, güç gibi değerler optimum seviyede çalışmalıdır. Fakat cihazların bulunduğu ortamlardan dolayı ya da cihazlarda meydan gelebilecek küçük arızalardan dolayı bu parametreler istenmeyen bir şekilde bozulabilir. Bunun sonucunda cihazlar verimsiz çalışır. Bu noktada ağ cihazlarının durumlarını görüntülemek amacıyla birçok yazılım geliştirilmiştir. Bu tez kapsamında geliştirilen yazılımın görevlerinden biri de ağ cihazlarının performansını etkileyen fan, sıcaklık, güç, yığın sayısı gibi parametreleri de görüntülemektir. Şekil 4.3’te ağ cihazının yığın bilgisi, sıcaklık bilgisi, fan bilgisi, güç bilgisi ve port bilgisi alınmıştır.



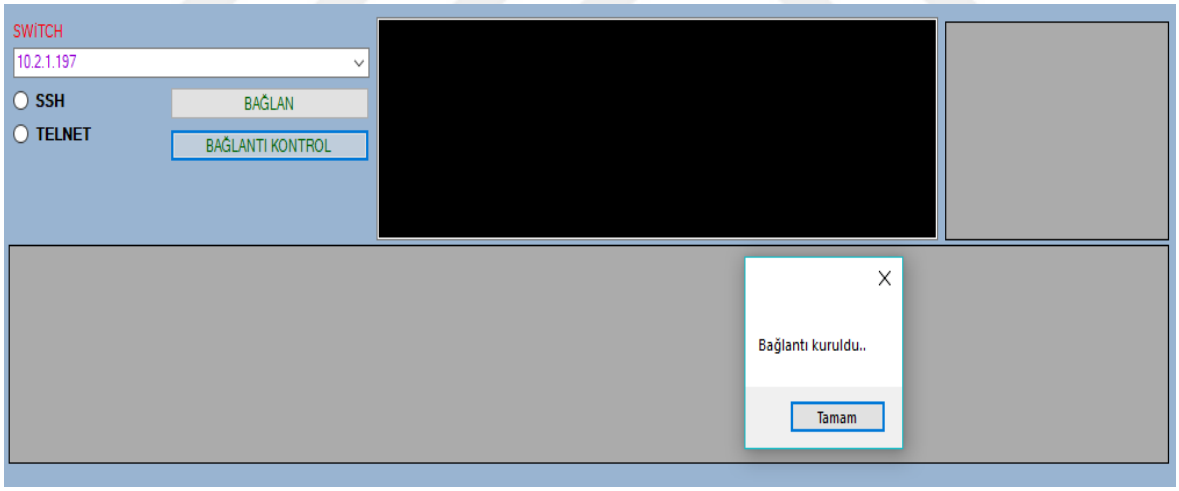
Şekil 4.3. a) Yığın bilgisi b) Güç bilgisi c) Fan bilgisi d) Sıcaklık bilgisi e) Port bilgisi f) Vlan bilgisi

Tez kapsamında geliştirilen yazılım ile ağ cihazlarına bağlanılmakta ve Ağ cihazlarından Şekil 4.3'te verilen bilgiler elde edilmektedir. Şekil 4.4'te geliştirilen yazılımın ara yüzü verilmiştir.



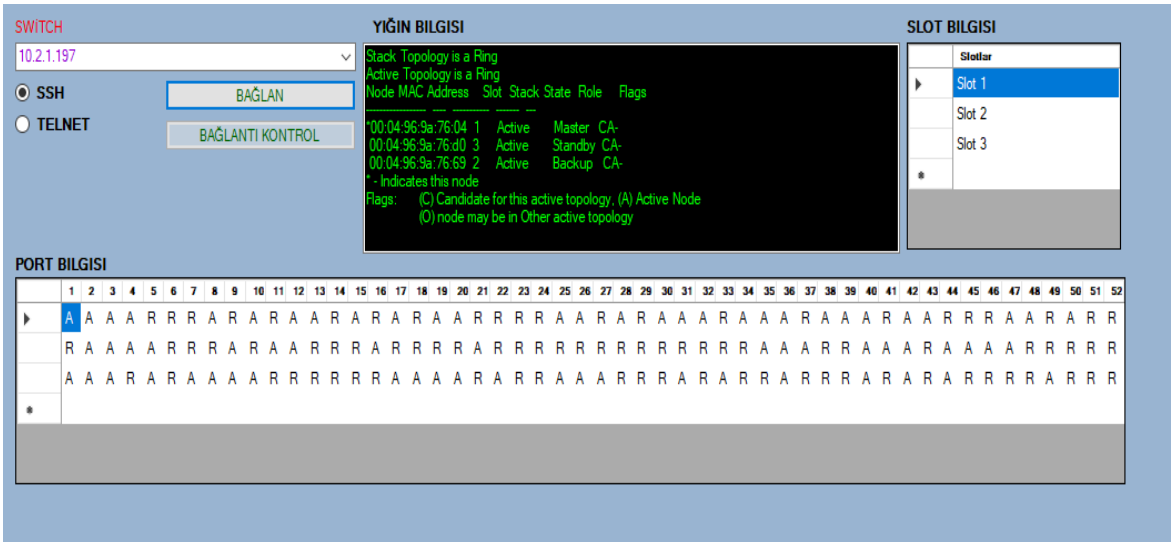
Şekil 4.4. Geliştirilen yazılıma ait ara yüz

Ara yüzde Switch sekmesi altında kuruma ait izlenecek ağ cihazları listelenmektedir. Daha sonra Bağlantı kontrol sekmesi ile seçilen ağ cihazı ile bağlantı olup olmadığı kontrol edilmektedir. Bu kontrol yapılırken yazılım cihaza ping atmaktadır. Şekil 4.5'te Bağlantı Kontrol butonuna tıklandığında meydana gelen ekran verilmiştir.



Şekil 4.5. Bağlantı kontrol butonu çıktısı

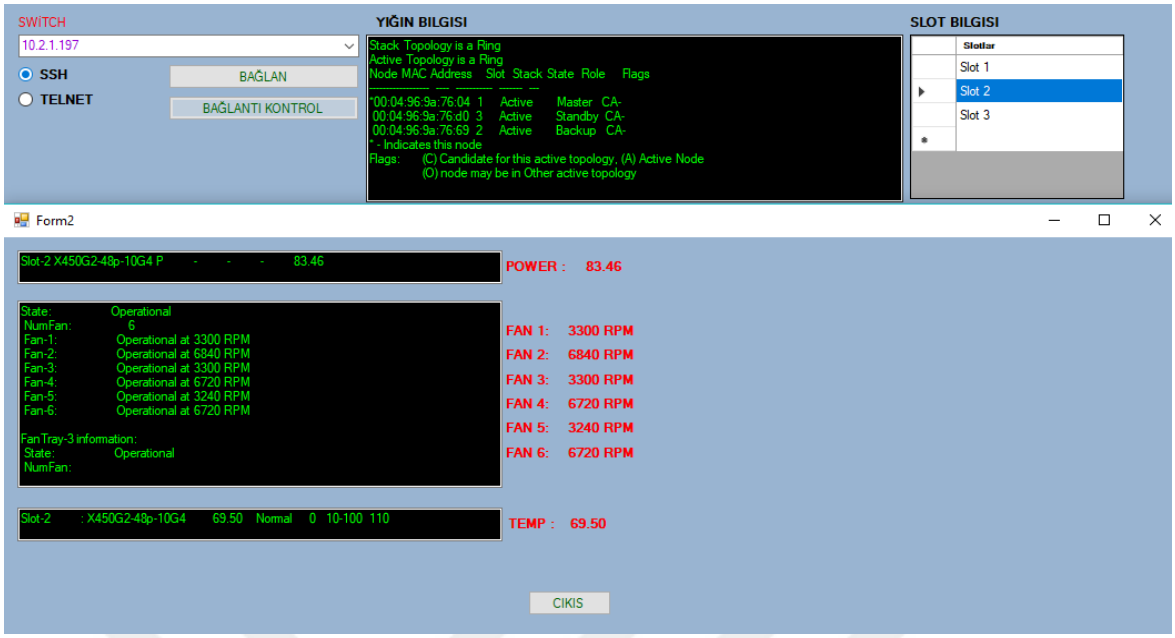
Ağ cihazlarına bağlanabilmek amacıyla SSH ve Telnet olmak üzere iki opsiyon ara yüzde sunulmuştur. Fırat Üniversitesi kampüs ağında SSH etkin olduğundan switche SSH bağlantısı yapılmış ve bağlantı sonucu karşımıza çıkan ekran Şekil 4.6'da verilmiştir.



Şekil 4.6. SSH bağlantısı sonucu oluşan ekran

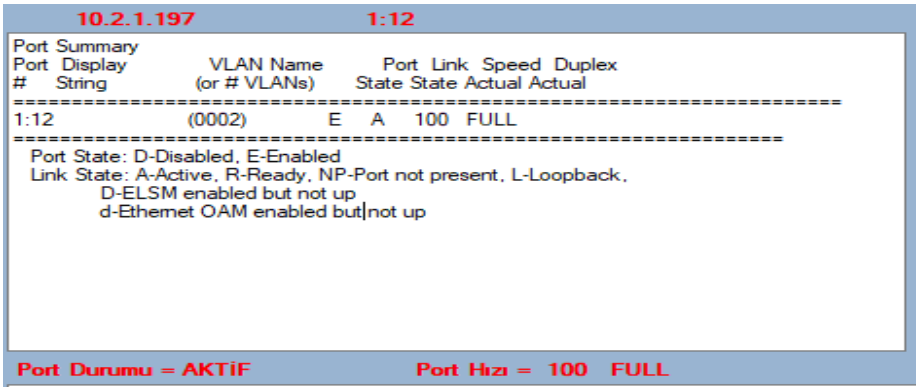
Şekil 4.6'da bağlanılan switche ait yığın bilgisi, slot bilgisi ve switche ait port bilgisi verilmiştir. Yığın Bilgisi yığını oluşturan switchlerin aktif halde çalışıp çalışmadığını, Slot bilgisi yığında kaç adet switch olduğunu, port bilgisi bağlantı kurulan switchteki portların durumunu göstermektedir. Aktif çalışan portlar "A" harfi ile pasif olan portlar "Ready" kelimesinin başındaki "R" harfi ile ifade edilmektedir.

Bağlantı yapıldıktan sonra Şekil 4.6'da bulunan slot bilgisi tablosundan herhangi bir slota tıklandığında, yığında bulunan ilgili switch ile ilgili fan, sıcaklık ve switchin güç durumu görüntülenebilmektedir. Şekil 4.7'de yığını oluşturan herhangi bir switche ait bilgiler verilmiştir.



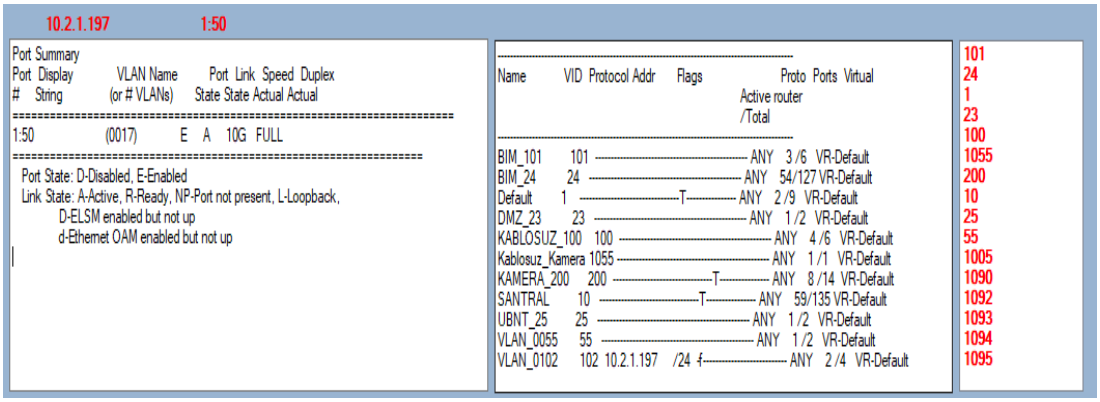
Şekil 4.7. Switche ait güç,fan ve sıcaklık bilgisi

Şekil 4.7'de de görüleceği üzere switche ait güç kullanımı, fanların durumu ve çalışma hızları, switchin çalışma sıcaklığı elde edilmiştir. Şekil 4.6'da Port Bilgisi tablosunda anahtarlama cihazına ait portlar listelenmektedir. Anahtarlama cihazından bir kullanıcı portu seçilip tıklandığında, portun hızı, portun aktif olup olmadığı bilgisi alınabilmektedir. Şekil 4.8'de bir kullanıcı portuna ait bilgiler verilmiştir.



Şekil 4.8. Switch port durum bilgisi

Ayrıca anahtarlama cihazının portundaki vlan bilgisi Şekil 4.9'daki gibi alınabilmektedir. Şekil 4.9'da anahtarlama cihazının uplink portu görüntülenmiştir. Böylece uplink portunun aktif vlan bilgisi otomotik olarak görüntülenebilmektedir.



Şekil 4.9. Anahtarlama cihazı portundaki vlan bilgisi

4.2. Arıza Parametrelerinin Elde Edilmesi

Bu tez çalışmasında Visual C# dilinde bir uygulama geliştirilmiştir. Geliştirilen uygulamanın kampüs içindeki tüm aktif cihazlara erişimine izin verilmiştir. Geliştirilen uygulama ağın herhangi bir noktasında konumlandırılabilir, fakat veri güvenliği bakımından uygulamanın ana sistem odasında çalıştırılması önerilmektedir. Geliştirilen uygulamada sistem odalarında bulunan switchlerin portları izlenerek veri kümesi oluşturulmuştur. Geliştirilen uygulama ile "Rx_CRC", "Tx_Errors", "Collision" ve "Congession" gibi hata parametreleri alınmıştır. Şekil 4.10'da anahtarlama cihazı ara yüzünde bu komutlar çalıştırılmış ve ekran çıktıları verilmiştir.

Port Rx Error Monitor

Port	Link State	Rx Crc	Rx Over	Rx Under	Rx Frag	Rx Jabber	Rx Align	Rx Lost
1:1	A	50	0	0	0	0	0	0
1:2	A	209	0	0	0	0	0	0
1:3	A	0	0	0	0	0	0	0
1:4	A	0	0	0	0	0	0	0
1:5	R	0	0	0	0	0	0	0
1:6	R	0	0	0	0	0	0	0
1:7	R	0	0	0	0	0	0	0
1:8	A	0	0	0	0	0	0	0
1:9	R	0	0	0	0	0	0	0
1:10	R	0	0	0	0	0	0	0
1:11	R	0	0	0	0	0	0	0
1:12	A	0	0	0	0	0	0	0
1:13	A	0	0	0	0	0	0	0
1:14	R	0	0	0	0	0	0	0
1:15	A	0	0	0	0	0	0	0
1:16	R	0	0	0	0	0	0	0
1:17	A	0	0	0	0	0	0	0
1:18	R	0	0	0	0	0	0	0
1:19	A	0	0	0	0	0	0	0
1:20	A	0	0	0	0	0	0	0

Link State: A-Active, R-Ready, NP-Port Not Present L-Loopback

a)

Port Tx Error Monitor

Port	Link State	Tx Coll	Tx Late coll	Tx Deferred	Tx Errors	Tx Lost	Tx Parity
1:1	A	0	0	0	0	0	0
1:2	A	0	0	0	0	0	0
1:3	A	0	0	0	0	0	0
1:4	A	0	0	0	0	0	0
1:5	R	0	0	0	0	0	0
1:6	R	0	0	0	0	0	0
1:7	R	0	0	0	0	0	0
1:8	A	0	0	0	0	0	0
1:9	R	0	0	0	0	0	0
1:10	R	0	0	0	0	0	0
1:11	R	0	0	0	0	0	0
1:12	A	0	0	0	0	0	0
1:13	A	0	0	0	0	0	0
1:14	R	0	0	0	0	0	0
1:15	A	0	0	0	0	0	0
1:16	R	0	0	0	0	0	0
1:17	A	0	0	0	0	0	0
1:18	R	0	0	0	0	0	0
1:19	A	0	0	0	0	0	0
1:20	A	0	0	0	0	0	0

> indicates Port Display Name truncated past 8 characters
Link State: A-Active, R-Ready, NP-Port Not Present L-Loopback

b)

Port Congestion Monitor

Port	Link State	Packet Drop
1:1	A	0
1:2	A	0
1:3	A	0
1:4	A	0
1:5	R	0
1:6	R	0
1:7	R	0
1:8	A	0
1:9	R	0
1:10	R	0
1:11	R	0
1:12	A	47911
1:13	A	0
1:14	R	0
1:15	A	24377
1:16	R	0
1:17	A	54026
1:18	R	0
1:19	A	14715
1:20	A	27572

Link State: A-Active, R-Ready, NP-Port Not Present L-Loopback

c)

Port Collision Monitor

Port	Link State	Collision Histogram 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1:1	A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:2	A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:3	A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:4	A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:5	R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:6	R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:7	R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:8	A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:9	R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:10	R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:11	R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:12	A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:13	A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:14	R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:15	A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:16	R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:17	A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:18	R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:19	A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1:20	A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

> indicates Port Display Name truncated past 8 characters
Link State: A-Active R-Ready NP- Port not present L-Loopback

d)

Şekil 4.10. Parametre çıktısı a) Rx Error, b) Tx Error, c) Congestion, d) Collision

Bu parametreler ağlarda sıkça karşılaşılan, fiziksel bağlantı sorunu ve bant genişliği yetersizliği gibi problemlerin bulunması açısından önemli parametrelerdir. Bu tez çalışmasında geliştirilen uygulama switch portlarındaki bu parametreleri alarak veri tabanına kaydetmektedir. Bu parametrelerdeki artış miktarları göz önüne alınarak arızanın kaynak nedeni hakkında bilgi sahibi olunacaktır.

Şekil 4.11'de tez kapsamında geliştirilen yazılıma ait sözde kod parçacığı verilmiştir. Geliştirilen algoritmaya göre switchlerden alınan parametrelerden dört sınıf oluşturulmaktadır. Bu kod parçacığında daha önce Bölüm 2.1'de belirtilen hata parametrelerinden "Rx_crc", "Tx_error", "Collision" ve "Congestion" değerleri kullanılmıştır. Bu sınıflandırma sonucu alınan veriler makine öğrenmesi yöntemleri ile işlenecektir.

```
EĞER((Rx Error <20) && (Tx Error <20) && (Collision <10) && (Congestion <40))
  DURUM=SAĞLAM
EĞER((Rx Error ≥20) || (Tx Error ≥20) || (Collision ≥10) && (Congestion <40))
  DURUM=FİZİKSEL HATA
EĞER((Rx Error <20) && (Tx Error <20) && (Collision <10) && (Congestion ≥40))
  DURUM=YOĞUNLUK ARIZASI
EĞER((Rx Error ≥20) || (Tx Error ≥20) || (Collision ≥10) && (Congestion ≥40))
  DURUM=FİZIKSEL VE YOĞUNLUK ARIZASI
```

Şekil 4.11. Sözde kod parçacığı

Şekil 4.11'de verilen sözde kod parçacığı dikkate alınarak bir veri kümesi elde edilmiştir. Örnek veri kümesi ve yapılan sınıflandırma soncu Tablo 4.1'de verilmiştir.

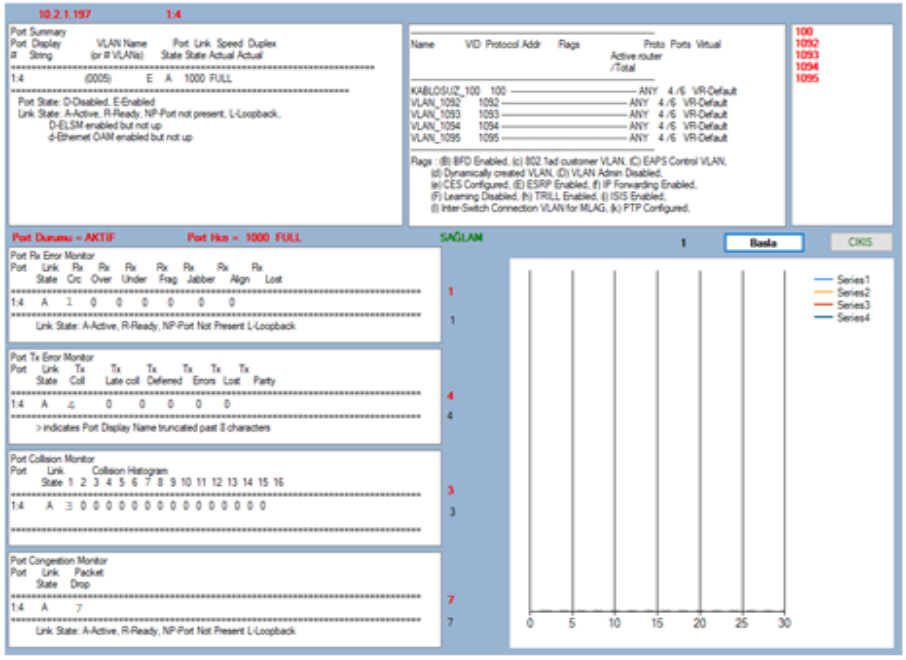
Tablo 4.1. Örnek veri kümesi ve elde edilen sınıflandırma sonucu

Rx CRC	Txerrors	Congestion	Collision	Sınıflandırma Sonucu
35	0	41	9	FY
16	3	15	9	S
27	94	45	5	FY
1	32	15	5	F
98	18	39	2	FY
90	89	39	7	FY
6	64	42	9	FY
4	18	46	5	Y
2	42	35	6	F
5	7	12	7	S
23	12	24	14	F
19	27	45	3	FY
36	21	47	0	FY
2	0	7	0	S

Tablo 4.1'deki sınıflar uzman görüşü ve yapılan testler sonucu belirlenmiştir. Yapılan testlerde her 5 sn. de bir örnek veriler alınmıştır. Yapılan gözlemler sonucu "Rx Error" ve "Tx Error" parametrelerindeki artış 20 den fazla ise, bu hata parametreleri gerçek bir fiziksel arıza yaratmaktayken, "Collision" parametresindeki artış ise "10" dan fazla ise gerçek bir fiziksel arızaya neden olmaktadır. Ayrıca "Congestion" parametresindeki artış 40' tan fazla ise hattın yoğun olduğu anlaşılmaktadır. Tablo 4.1'deki harflerin gösterdiği sınıflar aşağıda verilmiştir.

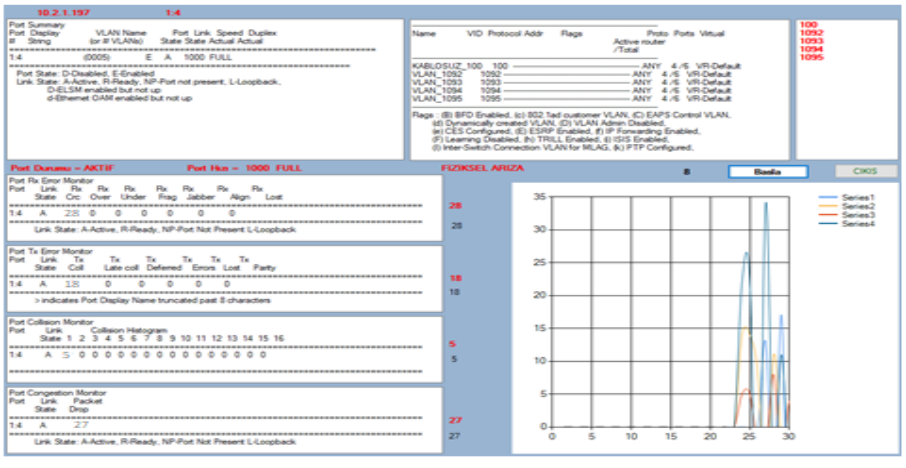
- F- Fiziksel Hata
- S- Sağlam
- Y- Yoğunluk hatası
- FY- Hem fiziksel hem yoğunluk hatası

Tez çalışmasında geliştirilen uygulama ile istenen anahtarlama cihazı portundaki arıza parametrelerini görüntülenebilmektedir. Geliştirilen uygulama kullanılarak alınan bu arıza parametreleri Şekil 4.11'de verilen sözde kod parçacığına göre yorumlanmıştır.



Şekil 4.12. Fiziksel hatta sorun olmaması durumu

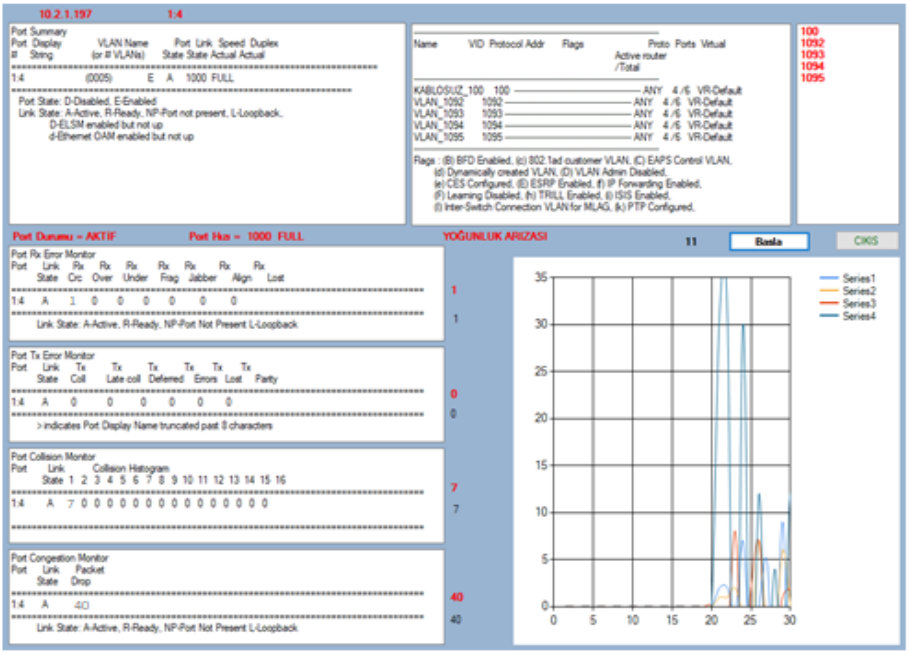
Şekil 4.12'de herhangi bir anahtarlama cihazı portuna ait hata parametreleri gösterilmektedir. Daha önce Şekil 4.11'da verilen sözde kod parçacığına göre "$Rx\ Error\ < 20\ \&\&\ Tx\ Error\ < 20\ \&\&\ Collision\ < 10\ \&\&\ Congestion\ < 40$" durumu Şekil 4.12'de sağlanmıştır. Dolayısıyla bu eşik değerler göz önüne alındığında fiziksel hattın sağlam olduğu sonucu görülmektedir.



Şekil 4.13. Fiziksel arıza olması durumu

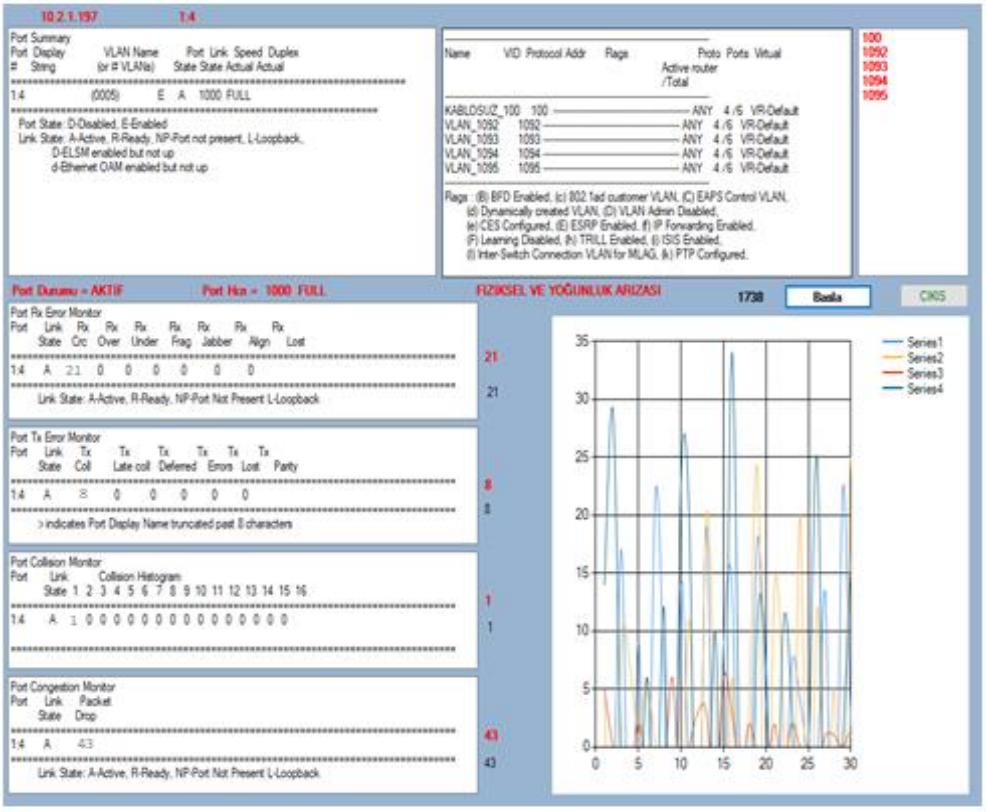
Şekil 4.13'te Fiziksel arıza oluşması durumu gösterilmiştir. Bu durum daha önce Şekil 4.11'de verilen sözde kod parçacığına göre "$Rx\ Error\ \geq 20\ ||Tx\ Error\ \geq$

$20 \,||\, Collision \geq 10 \,\&\&\, Congestion < 40$" durumu için sağlanmıştır. Dolayısıyla bu eşik değerler göz önüne alındığında hatta fiziksel bir arıza olduğu sonucu görülmektedir.



Şekil 4.14. Yoğunluk arızası oluşma durumu

Daha çok kullanıcı yoğunluğundan ya da kullanılan kablonun bant genişliği yetersizliğinden dolayı meydan gelen yoğunluk arızası Şekil 4.14'te verilmiştir. Bu arıza durumu "$Rx\ Error < 20 \,\&\&\, Tx\ Error < 20 \,\&\&\, Collision < 10 \,\&\&\, Congestion \geq 40$" şartları için sağlanmıştır. Dolayısıyla bu eşik değerler göz önüne alındığında hatta yoğunluk arızası olduğu sonucuna varılmaktadır.



Şekil 4.15. Hem fiziksel hem yoğunluk arızası oluşma durumu

Hem bant genişliği yetersizliği, kullanılan kablonun fiziksel bir arızaya uğramış olması ve kullanıcı yoğunluğundan dolayı oluşabilecek sorunlar "Fiziksel ve Yoğunluk" arızası olarak yorumlanmıştır. Bu durum daha önce Şekil 4.9'da verilen sözde kod parçacığına göre "$(Rx\ Error\ \geq 20\ ||Tx\ Error\ \geq 20\ ||\ Collision\ \geq 10\ )\&\&\ Congestion\ \geq 40$" durumu için sağlanmış ve Şekil 4.15'te verilmiştir.

4.3. Arıza Parametrelerinin Makine Öğrenmesi Metodları ile Sınıflandırılması

Tez kapsamında geliştirilen uygulama ile aktif ağ cihazlarından arıza parametreleri sağlıklı olarak alınabilmektedir. Şekil 4.11'de verilen sözde kod parçacığına göre 4 adet arıza parametresine bakılarak 4 adet sınıf oluşturulmuştur. Sınıflandırma için Destek Vektör Makineleri, Karar Ağaçları ve Naive Bayes makine öğrenmesi yöntemleri kullanılmıştır. Daha sonra elde edilen sonuçlar karşılaştırılmıştır.

4.3.1. Naive Bayes Metodu ile Sınıflandırma

Bu tez çalışmasında elde edilen 3142 tane veri için sınıflandırma yapılmıştır. Sınıflandırma için Weka programı kullanılmıştır. Sınıflandırıcı verilen bu 3142 adet verinin 2769 adetini doğru sınıflandırmış ve ortalama doğruluk oranını %88.1286 olarak bulunmuştur. Kullanılan performans metriklerine göre elde edilen sonuçlar Tablo 4.2'de verilmiştir.

Tablo 4.2. Naive Bayes Sınıflandırıcısı Sonucu

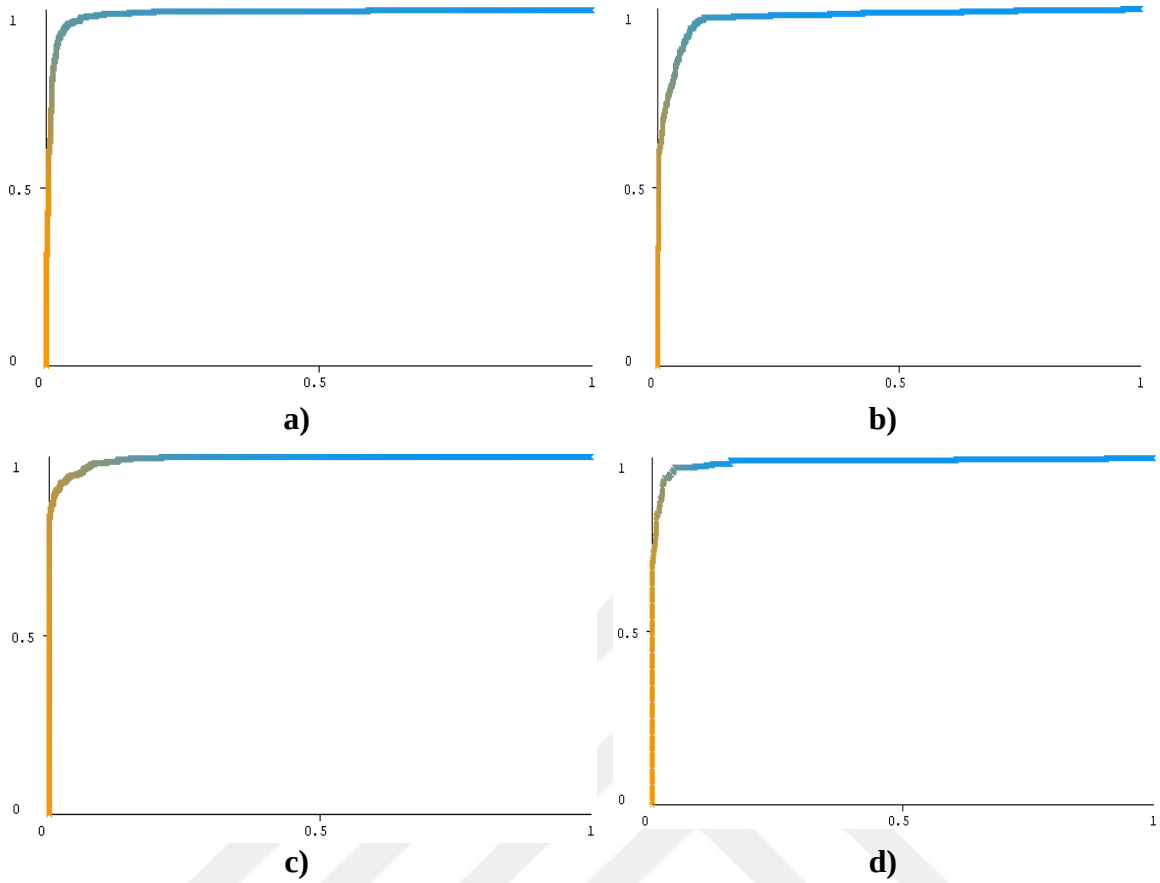
	Doğru Pozitif Oranı	**Yanlış Pozitif Oranı**	**Kesinlik**	**Duyarlılık**	**F-Ölçüsü**	**ROC Alanı**	**Sınıf**
	0,84	0,015	0,983	0,84	0,906	0,987	F
	0,802	0,029	0,833	0,802	0,818	0,973	FY
	0,991	0,082	0,805	0,991	0,888	0,991	S
	0,957	0,027	0,737	0,957	0,832	0,986	Y
Ağırlıklı Ortalama	0,881	0,035	0,896	0,881	0,882	0,986	

Naive Bayes sınıflandırıcısının kullandığı Hata matrisi Tablo 4.3'te verilmiştir.

Tablo 4.3. Hata Matrisi

a	**b**	**c**	**d**	**Sınıflar**
1363	70	189	1	a=F
16	390	2	78	b=FY
6	1	795	0	c=S
1	7	2	221	d=Y

Her bir sınıf için elde edilen ROC eğrileri Şekil 4.16'da verilmiştir.

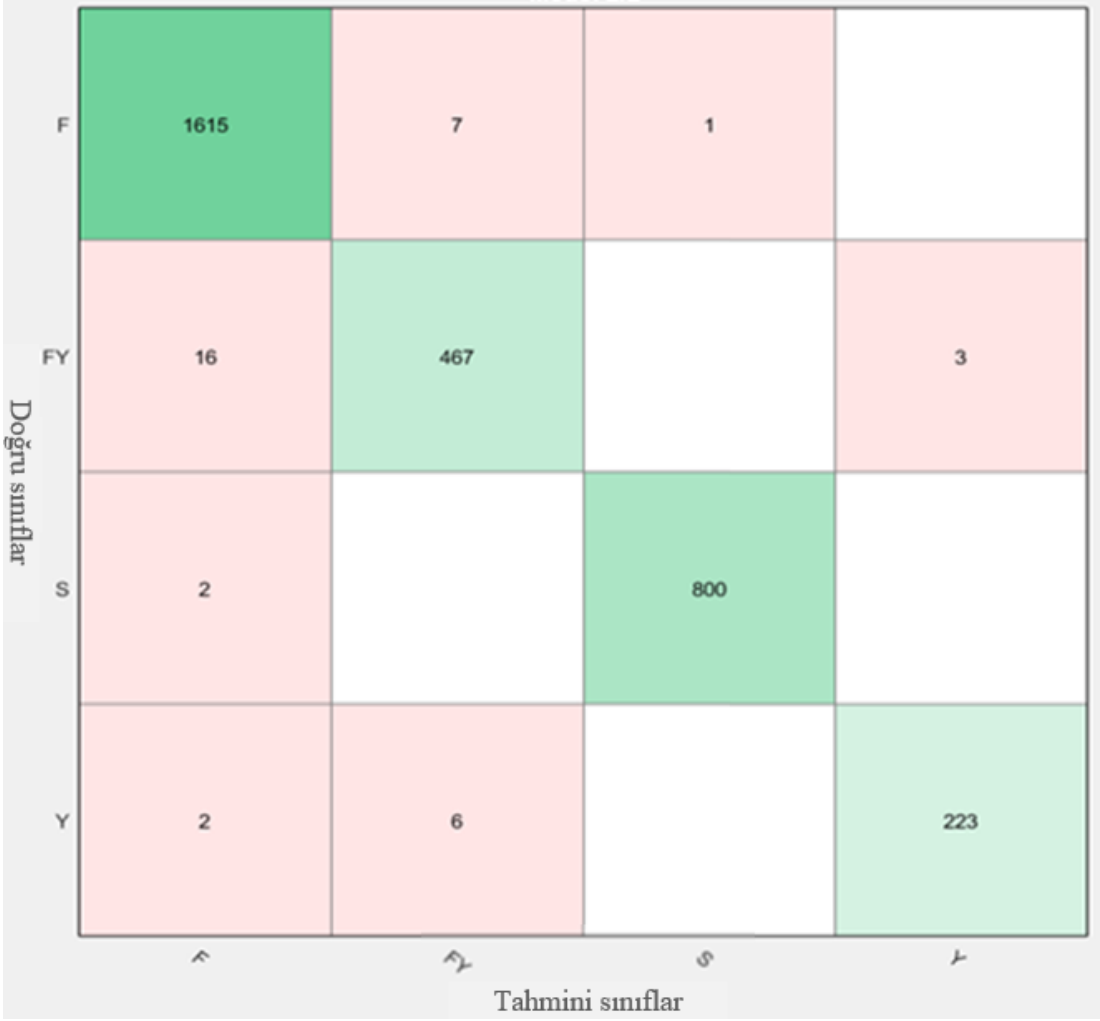


Şekil 4.16. Naive Bayes sınıflandırıcı sonucunda elde edilen ROC eğrileri a) Fiziksel hata b) Fiziksel ve Yoğunluk hatası c) Sağlam d) Yoğunluk arızası

4.3.2. Karar Ağaçları ile Sınıflandırma

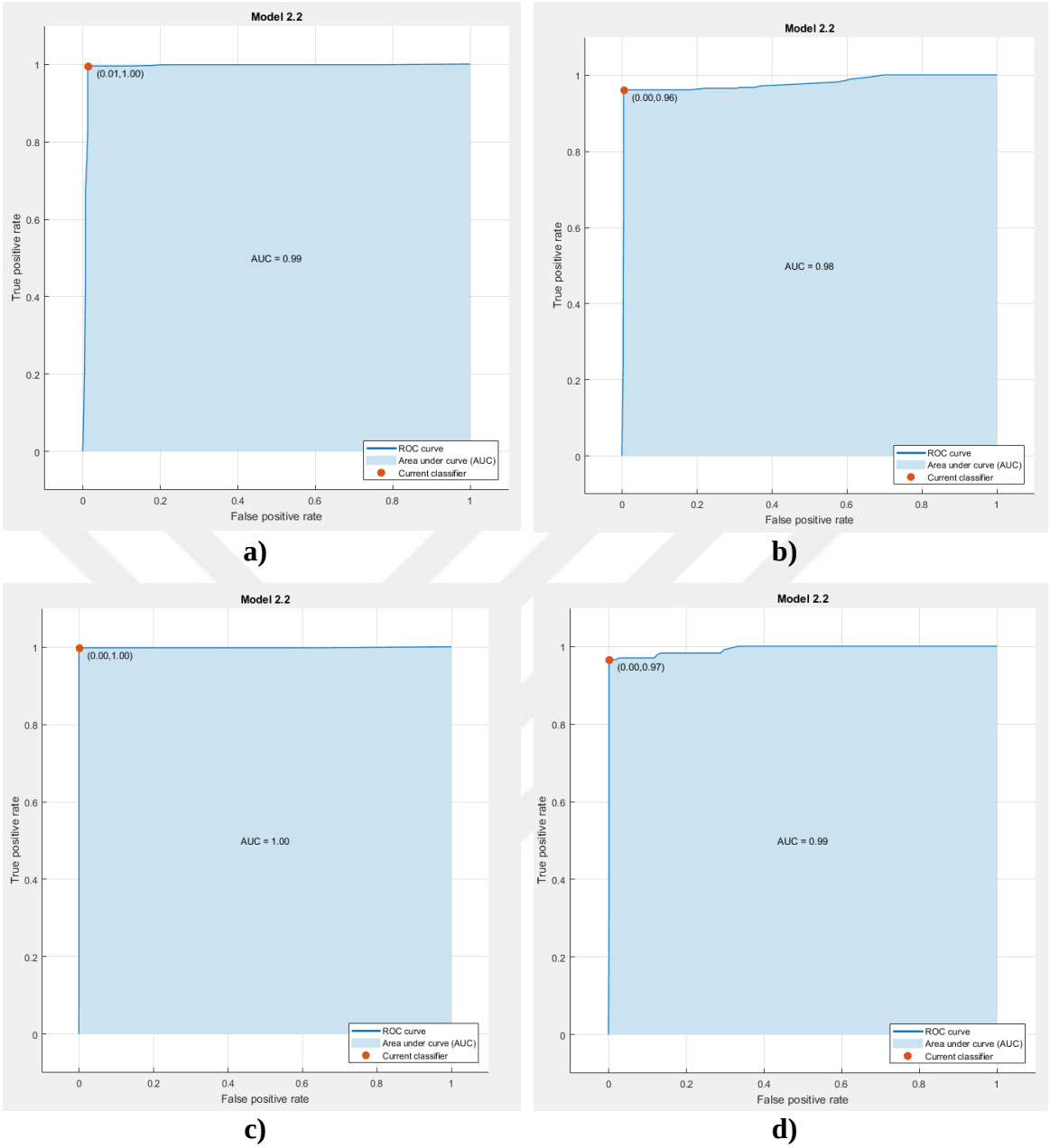
Karar ağaçları sıklıkla kullanılan makine öğrenmesi metodlarından birisidir. Eğitim ve testinin hızlı ve etkin olması karar ağaçlarının sıklıkla kullanılma sebeplerinden biridir. Sınıflandırma yapılırken ilkin ağaç oluşturulur daha sonra oluşturulan bu ağaç yapısından sınıflandırma kuralları elde edilir. Karar ağaçları ile sınıflandırma iki adımda gerçekleştirilir. İlk adımda ağaç oluşturulur.

Tez kapsamında geliştirilen uygulama ile alınan veriler karar ağacı algoritması ile sınıflandırılmış ve bunun sonucunda %98.8'lik doğruluk oranı elde edilmiştir. Şekil 4.17' de karar ağacı sınıflandırması sonucu oluşan hata matrisi verilmiştir.



Şekil 4.17. Karar ağacı hata matrisi

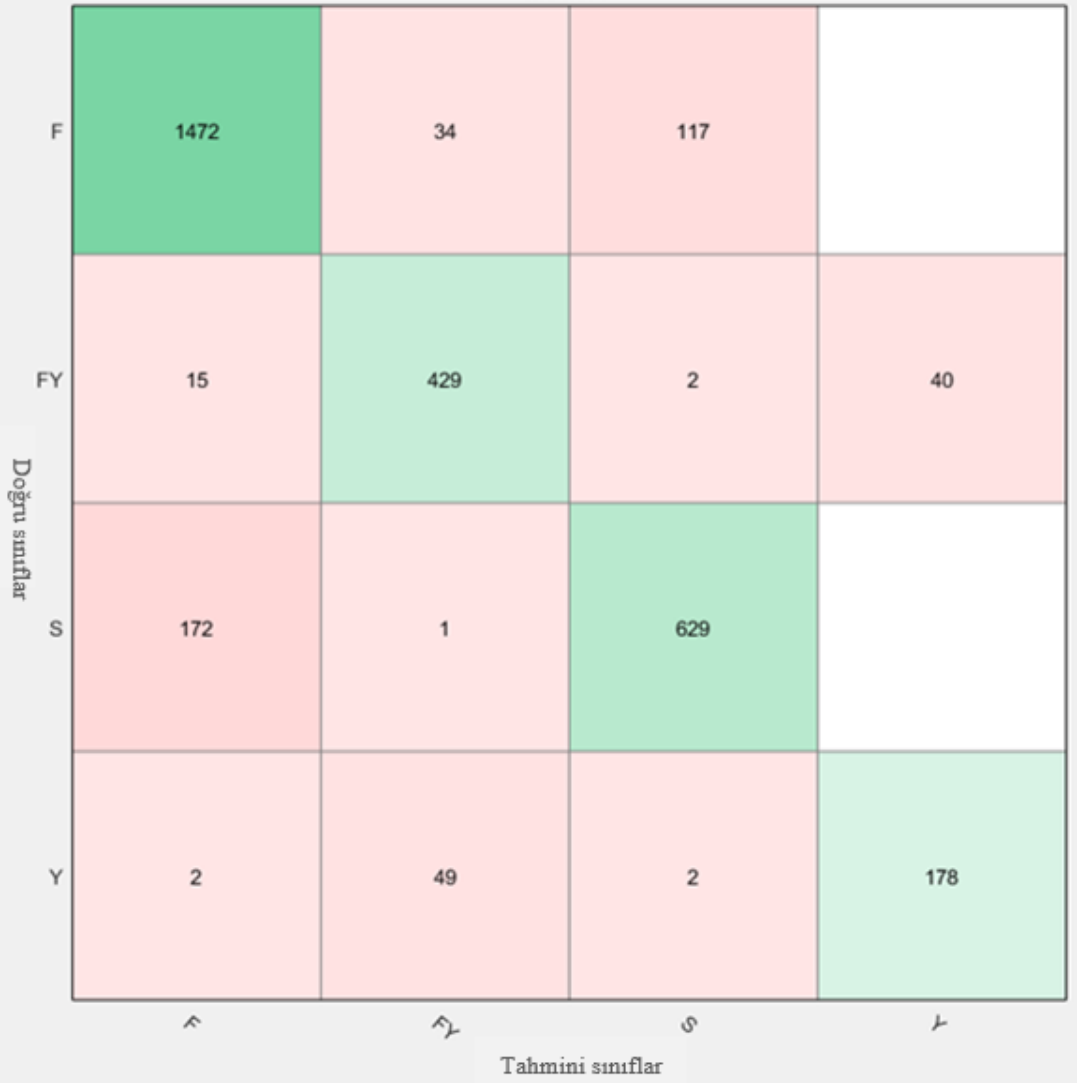
Şekil 4.18'de Karar ağacı sonucunda elde edilen ROC eğrileri her bir sınıf için ayrı ayrı verilmiştir.



Şekil 4.18. Karar ağacı sınıflandırıcısı sonucunda oluşan ROC eğrileri a) Fiziksel hata b) Fiziksel ve Yoğunluk hatası c) Sağlam d) Yoğunluk arızası

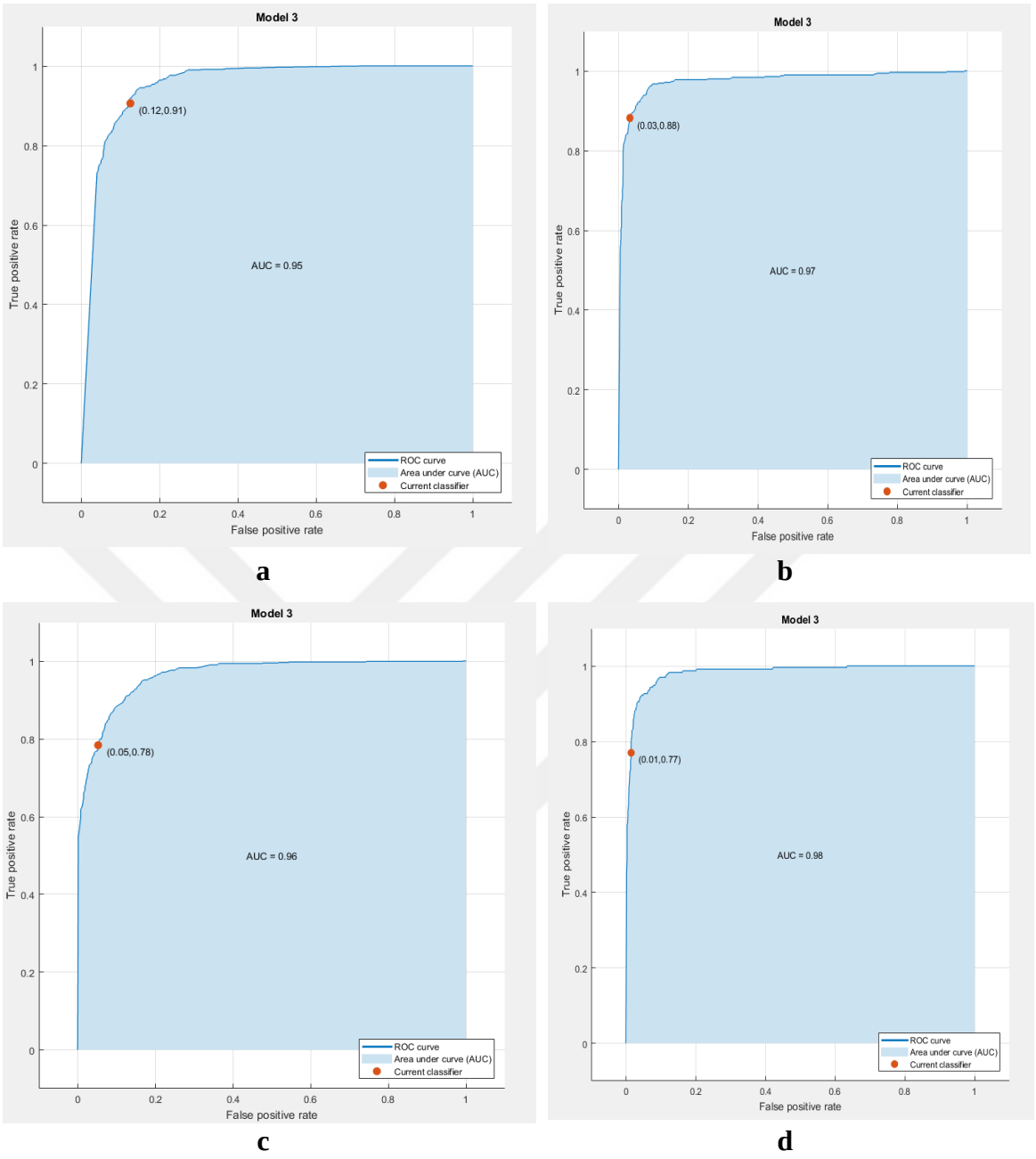
4.3.3. Destek Vektör Makineleri ile Sınıflandırma

Bu tez çalışmasında elde edilen 3142 tane veri için matlab ortamında sınıflandırma yapılmıştır. Lineer Destek Vektör sınıflandırıcısı (Lineer SVM) verilen bu 3142 adet veriyi ortalama % 86,2 doğruluk oranı ile sınıflandırmıştır. Sınıflandırma sonucu elde edilen hata matrisi Şekil 4.19'da verilmiştir.



Şekil 4.19. Lineer DVM hata matrisi

Şekil 4.20'de Lineer Destek Vektör Sınıflandırıcısı sonucunda elde edilen ROC eğrileri her bir sınıf için ayrı ayrı verilmiştir.



Şekil 4.20. Lineer DVM sınıflandırıcısı sonucunda oluşan ROC eğrileri a) Fiziksel hata b) Fiziksel ve Yoğunluk hatası c) Sağlam d) Yoğunluk arızası

5. SONUÇLAR

Ağ yöneticileri için yönettikleri ağ üzerinde meydana gelebilecek iletişim ve fiziksel hataların tespit edilebilmesi çok önemlidir. Hata tespitinin hızlı yapılması çözüme de hızlı bir şekilde gidilebilmesini sağlamaktadır. Bu çalışma için bir ağ kampüs alanında bulunan kenar anahtarlama cihazları üzerinden verilerin çekilebilmesi amacıyla bir yazılım geliştirilmiş ve bu yazılım ile hatanın sınıflandırılabilmesi sağlanmıştır. Sınıflandırmanın akıllı bir sistem haline getirilebilmesi için makine öğrenmesi yöntemlerinden Naive Bayes, Karar Ağaçları ve Destek Vektör Makinesi sınıflandırıcıları ile sınıflandırılmıştır. Sınıflandırma sonuçları ele alındığında en iyi sınıflandırmayı %98,8 doğruluk oranı ile Karar Ağacı sınıflandırıcısı bulmuştur. Bunu %88,12'lik doğruluk oranı ile Naive Bayes sınıflandırıcısı takip etmiştir. Destek Vektör Makinesi sınıflandırıcısı %86,2'lik doğruluk oranı bulmuştur. Elde edilen sonuçlara göre bu çalışmada en iyi performansa Karar Ağacı sınıflandırıcısı ile ulaşılmaktadır. İlerleyen çalışmalarda tespit edilen hatanın doğrudan port üzerinde giderilebilmesi için ağ yöneticisine uyarıda bulunacak ve otomatik olarak port üzerinde işlem yapabilecek yazılımın geliştirilmesi hedeflenmektedir. Ayrıca çok daha fazla verinin alınarak büyük veriler üzerinde farklı makine öğrenmesi yaklaşımları ile tespit işleminin hızlandırılmasına çalışılacaktır.

KAYNAKLAR

[1] **Shrouf, F., Ordieres, J., Miragliotta, G.** "Smart factories in Industry 4.0: A review of the concept and of energy management approached in production based on the Internet of Things paradigm." IEEE International Conference on In Industrial Engineering and Engineering Management (IEEM), 697-701, 2014.

[2] **Wortmann, F., Flüchter, K.**, Internet of things. Business & Information Systems Engineering, *57*(3), 221-224, 2015.

[3] **Fleisch E, Weinberger M, Wortmann F**., Business models for the internet of things. Bosch lab white paper. http://www.iot-lab ch/wp-content/uploads/2014/09/EN_Bosch-Lab-White-PaperGM-im-IOT 1_1.pdf. Accessed, 29 Sep 2014

[4] http://www.endustri40.com/endustri-tarihine-kisa-bir-yolculuk/

[5] **Mahany, R. L., Bunte, A. G., Luse, R. E., West, G. J., Gollnick, C. D.**, U.S. Patent No. 6,359,872. Washington, DC: U.S. Patent and Trademark Office, 2002

[6] **Hoong, Y. C..,** CCNA Complete Guide. CreateSpace, 2008.

[7] http://www.emo.org.tr/ekler/12d73cae3a71238_ek.pdf

[8] **Engel, F., Jones, K. S., Robertson, K., Thompson, D. M., White, G..,** U.S. Patent No. 6,115,393. Washington, DC: U.S. Patent and Trademar Office, 2000.

[9] **Pawar, S. U., Sandeep A. T.,** "Offline network monitoring system using multi-agent framework." International Conference on Computing, Analytics and Security Trends (CAST), IEEE, 2016.

[10] **Manzoor, U., Samia N.,** "An agent based system for activity monitoring on network– ABSAMN." Expert Systems with Applications 36.8 10987-10994, 2009.

[11] **Park, S., Ketan T., Chris M.,** "RCV: Network Monitoring and Diagnostic System with Interactive User Interface." International Conference on Collaboration Technologies and Systems (CTS), 2016.

[12] **Elbaham, M., Kim Khoa N., Mohammed C.,** "A traffic visualization framework for monitoring large-scale inter-datacenter network." 12th International Conference on Network and Service Management (CNSM), 2016.

[13] **Choi, B. Y., Song, S., Koffler, G., Medhi, D.** "Outage analysis of a university campus network," Proc. - Int. Conf. Comput. Commun. Networks, ICCCN, no. September, pp. 675–680, 2007.

[14] **Hood C. S., Ji, C.**, "Proactive network-fault detection," IEEE Trans. Reliab., vol. 46, no. 3, 333–341, 1997.

[15] **Wang, X., Wang, L., Yu, B., Dong, G.,** "Studies on network management system framework of campus network," in CAR 2010 - 2010 2nd International Asia Conference on Informatics in Control, Automation and Robotics, vol.2, 285–289, 2010.

[16] **Kurt, B.,** "A Network Monitoring System for High Speed Network Traffic." 13th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON), 2016.

[17] **Zhang, Y., Yuan, Q.**., A multiple bits error correction method based on cyclic redundancy check codes. 9th International Conference on In Signal Processing, 1808-1810, 2008.

[18] **Wu, H., Feng, Z., Guo, C., Zhang, Y.,** "ICTCP: Incast congestion control for TCP in data-center networks," IEEE/ACM Trans. Netw., vol. 21, no. 2, 345–358, 2013.

[19] **Senthilkumaran, T., Sankaranarayanan, V..,** Dynamic congestion detection and control routing in ad hoc networks. Journal of King Saud University-Computer and Information Sciences, 25(1), 25-34, 2013.

[20] **Monowar, M. M., Rahman, M. O., Hong, C. S..,** Multipath congestion control for heterogeneous traffic in wireless sensor network. 10th International Conference on In Advanced Communication Technology, ICACT 2008. Vol. 3, 1711-17152008.

[21] **P. Koopman T. Chakravarty,** "Cyclic Redundancy Code (CRC) Polynomial Selection For Embedded Networks," Int. Conf. Dependable Syst. Networks, 1–11, 2004.

[22] **Calvignac, J. L., Jeffries, C. D., Verplanken, F. J..,** U.S. Patent No. 6,681,364. Washington, DC: U.S. Patent and Trademark Office, 2004.

[23] **Ng, Y. K.**, U.S. Patent No. 6,609,225. Washington, DC: U.S. Patent and Trademark Office, 2003.

[24] https://buraktahtacioglu.blogspot.com.tr/2016/04/bilgisayar-aglar-vi-computer networks-vi.html

[25] **Chatzimisios, P., Boucouvalas, A. C., Vitsas, V..,** Performance analysis of IEEE 802.11 DCF in presence of transmission errors. IEEE International Conference on In Communications, Vol. 7, 3854-3858, 2004.

[26] **Extreme Networks,** "ExtremeXOS ® Command Reference Guide," vol. 95051, no. April, 257–3000, 2012.

[27] **Dirican C. O.,** Teori ve Uygulamalar ile TCP/IP ve Ağ Güvenliği, Açık akademi Yayınları, 2005

[28] **Çetin, K. A. Y. A., YILDIZ,** "Makine öğrenmesi teknikleriyle saldırı tespiti: Karşılaştırmalı analiz. "Marmara Fen Bilimleri Dergisi, 26(3), 89-104, 2014

[29] **Gürsoy, M. İ., Subaşı, A..,** DVM ile EEG işaretlerinin sınıflandırılmasında TBA BBA ve DAA'nin performansının karşılaştırılması. In Signal Processing, Communication and Applications Conference, 1-4, 2008.

[30] **Vapnik, V.,** "Statistical Learning Theory", Wiley, New York, 1998.

[31] **Sheremet, O., Sadovoy, O.** "Using the Support Vector Regression method for telecommunication networks monitoring. In Problems of Infocommunications Science and Technology (PIC S&T)," Third International Scientific-Practical Conference, 8-10, 2016

[32] **Murat, K.,** "Yeniden düzenleşim için entropi tabanlı arama algoritmalarının geliştirilmesi", Fırat Üniversitesi / Fen Bilimleri Enstitüsü / Bilgisayar Mühendisliği Anabilim Dalı, Yüksek lisans tezi, 2014

[33] **Zhang, J., Chen, C., Xiang, Y., Zhou, W., Xiang, Y.,** "Internet traffic classification by aggregating correlated naive bayes predictions," IEEE Trans. Inf. Forensics Secur., vol. 8, no. 1, 5–15, 2013.

[34] **Wang, S., Jiang, L., Li, C.,** "Adapting naive Bayes tree for text classification," Knowl. Inf. Syst., vol. 44, no. 1, 77–89, 2015.

[35] **Lewis, D. D., Nedellec, C., Rouveirol, C.,** "Naive (Bayes) at Forty: The Independence Assumption in Information Retrieval," Mach. Learn. ECML-98, 4 -15, 1998.

[36] **Davis, J., Goadrich, M.,** "The Relationship Between Precision-Recall and ROC Curves," Proc. 23rd Int. Conf. Mach. Learn, 233–240, 2006.

[37] **Kilinçer, İ. F., Ertam, F., Yaman, O., Akbal, A..,** Automatic fault detection with Bayes method in university campus network. International In Artificial Intelligence and Data Processing Symposium (IDAP), 1-4, 2017.

[38] Ciscopress.com copyright 2003 CCNA 3 Chapter 5

http://www.myshared.ru/slide/895325/

ÖZGEÇMİŞ

İlhan Fırat KILINÇER 1986 yılında Elazığ' da doğdu. İlk, orta ve lise öğrenimini Elazığ'da tamamladı. 2006 yılında Hulusi Sayın Lisesini bitirdikten sonra 2012 yılında Kocaeli Üniversitesi, Mühendislik Fakültesi, Elektronik ve Haberleşme Mühendisliği Bölümünden mezun oldu. 2012-2014 yılları arasında Türk Telekom Genel Müdürlüğünde Network ve Wireless Mühendisi, 2014-2015 yıllarında Netaş Telekomünikasyon A.Ş' de Network ve Güvenlik Uzmanı, 2015-2016 yılları arasında Türk standartları Enstitüsü' nde Network ve Sistem Uzmanı olarak çalıştı. Şubat 2016' dan beri Fırat Üniversitesi Bilgi İşlem Dairesi Başkanlığında Uzman olarak çalışma hayatına devam etmektedir. Aynı yıl içerisinde Fırat Üniversitesi, Fen Bilimleri Enstitüsü, Elektrik - Elektronik Mühendisliği Ana Bilim dalında yüksek lisans eğitimine başladı.