



VoIP Şebekesindeki Ses Kalitesinin Çok Katmanlı Yapay Sinir Ağları ile Modellenmesi

Selçuk Mete^{1*}

¹ Türk Telekom A.Ş., Network Yönetim Müdürlüğü, Kayseri, Türkiye (ORCID: 0000-0001-6842-1088)

(İlk Geliş Tarihi 11 Mayıs 2020 ve Kabul Tarihi 10 Temmuz 2020)

(DOI: 10.31590/ejosat.745810)

ATIF/REFERENCE: Mete, S. (2020). VoIP Şebekesindeki Ses Kalitesinin Çok Katmanlı Yapay Sinir Ağları ile Modellenmesi. *Avrupa Bilim ve Teknoloji Dergisi*, (19), 679-690.

Öz

Son yıllarda internet ağının, çok yaygınlaşması, esnek olması, hızlı ve yüksek kapasite erişimi sunabilmesine bağlı olarak farklı tipteki birçok servis internet protokolü (IP) üzerinden iletmeye başlanmıştır. Bu servis tiplerinden birisi de ses trafiği olup bu durum klasik haberleşme sistemlerini doğrudan etkilemektedir. Ses iletimi, genel anahtarlamalı telefon ağları (Public Switched Telephone Network, PSTN) yerine IP üzerinden yapılmaya başlanmıştır. Bu yeni haberleşme teknolojisi, IP üzerinden ses iletimi (Voice over IP, VoIP) olarak adlandırılmıştır. VoIP teknolojisinde, ses iletiminin kullanıcı tarafından algılanma kalitesi ise ortalama görüş puanı (Mean Opinion Score, MOS) ile belirlenir. MOS değerinin üretilmesi birçok parametreye bağlıdır. ITU-T (International Telecommunications Union - Telecommunication Standardization Sector) tarafından MOS değerinin belirlenmesi için PSQM (Perceptual Speech Quality Measure), PESQ (Perceptual Evaluation of Speech Quality), POLQA (Perceptual Objective Listening Quality Analysis) gibi birçok model standartlaştırılmıştır. Fakat bu model yapıları, birçok avantaja sahip olmasının yanında referans işaret gereksiniminden dolayı uygulamalarda pratik zorluklar oluşturmaktadır. Buyüzden bu çalışmada, ses iletimindeki MOS kalite değerlerini tahmin etmek amacıyla referans sinyal gerektirmeyen çok katmanlı yapay sinir ağları (YSA) tabanlı bir model yapısı tasarlanmıştır. Bu modelde, giriş olarak IP trafiğine ait servis kalite (Quality of Service, QoS) parametreleri ve çıkış olarak ise elde edilen MOS değerleri kullanılmıştır. Giriş olarak kullanılan bu QoS parametreleri ise paket kaybı (Packet Loss) ve gecikme (Delay) değerleridir. Böylece 2 giriş ve 1 çıkışa sahip bir model oluşturulmuştur. Ayrıca model, farklı QoS parametresi kullanılarak MOS tahmini yapabildiği için esnek bir yapıya da sahiptir. Literatürdeki çalışmalardan farklı olarak bu çalışmada, POLQA yöntemi ile ölçülen MOS değerlerinin YSA model ile tahmin edilmesine çalışılmıştır. YSA modelin eğitim ve test sürecinde LM (Levenberg Marquardt), BR (Bayesian Regulation) ve RPROP (Resilient Backpropagation) algoritması kullanılmıştır. Simülasyon sonucunda elde edilen sonuçlar tablolar ve şekiller vasıtasıyla sunulmuştur. Sonuçlara göre geliştirilen bu metodun ITU-T tarafından önerilen modeller ile karşılaştırılabilir seviyede olduğu gösterilmiştir.

Anahtar Kelimeler: VoIP, Ortalama Görüş Puanı, POLQA Model, Yapay Sinir Ağları.

Modeling of Sound Quality in VoIP Network with Multi-Layer Artificial Neural Networks

Abstract

In recent years, different types of services have been started to be transmitted over the internet protocol (IP), due to the widespread use of the internet network, its flexibility and the ability to offer fast and high capacity access. One of these service types is voice traffic and this directly affects classical communication systems. Voice transmission is started over IP instead of public switched telephone network (PSTN). This new communication technology is called voice over IP (VoIP). In VoIP technology, the quality of the perception of the voice transmission by the user is determined by the mean opinion score (MOS). Generation of the MOS value depends on many parameters. Many models such as PSQM (Perceptual Speech Quality Measure), PESQ (Perceptual Evaluation of Speech Quality),

* Sorumlu Yazar: Türk Telekom A.Ş., Network Yönetim Müdürlüğü, Kayseri, Türkiye, ORCID: 0000-0001-6842-1088, selcuk.metes@gmail.com

POLQA (Perceptual Objective Listening Quality Analysis) have been standardized by ITU-T (International Telecommunications Union -Telecommunication Standardization Sector) to determine the MOS value. However, besides having many advantages these model structures also create practical difficulties in applications due to the need of a reference signal. Therefore, in this study, a multi-layered artificial neural networks (ANN) based model structure that does not need a reference signal is designed to estimate MOS quality values in voice transmission. In this model, quality of service (QoS) parameters of IP traffic are used as input and obtained MOS values are used as output. These QoS parameters used as input are packet loss and delay values. Thus, a model with 2 inputs and 1 output was created. In addition, the model has a flexible structure as it can estimate MOS using different QoS parameters. Different from the studies in the literature, in this study, it was tried to estimate MOS values measured by the POLQA method with ANN model. LM (Levenberg Marquardt), BR (Bayesian Regulation) and RPROP (Resilient Backpropagation) algorithm was used in the training and testing process of the ANN model. The results obtained from the simulations are presented by tables and figures. The method, developed according to the results, has been shown to be comparable to the models recommended by ITU-T.

Keywords: VoIP, Mean Opinion Score, POLQA Model, Artificial Neural Networks.

1. Giriş

Günümüzde IP (internet protokol) ağlarının altyapıları geliştikçe üzerinden verilen servis çeşitliliği de artmaktadır. WWW, e-posta, dosya aktarım protokolü (File Transfer Protocol, FTP) gibi iyi bilinen servislerin yanında IP üzerinden ses iletimi (Voice Over IP, VoIP), video konferans, IP-TV gibi yeni nesil ses ve video servisleri de IP ağları üzerinden iletilebilmektedir. Yeni nesil bu haberleşme servislerinden en popülerleri kuşkusuz internet üzerinden ses iletimi olarak adlandırılan VoIP 'dir. Çünkü sabit telefonların kullandıkları ağ olarak da bilinen genel anahtarlamalı telefon ağı (Public Switched Telephone Network, PSTN) sisteminde arayan ve aranan numara arasında uçtan uca bir devre kurulmaktadır. Bu sistemlerde telefon trafik artışına bağlı olarak yeni ilave devreler tahsis edilmekte ve bu durum toplam maliyeti arttırmaktadır. VoIP teknolojisi sayesinde IP santral sistemleri (IP-PBX) kullanılarak kurumsal ve büyük ölçekli firmaların bütün şubeleri ile internet üzerinden ücretsiz görüşmesi sağlanabilmektedir. Ayrıca kullanıcılar arasında gerçekleştirilen aramalarda kanal sınırı bulunmamaktadır. Yani eş zamanlı çağrı sayısını belirleyen tek unsur internet bant genişliğidir. Ayrıca akıllı telefonlara yüklenen uygulamalar ile hızlı bir şekilde VoIP şebekesine erişilebilmektedir. Bu sayede VoIP teknolojisi, kullanıcılara çok esnek bir yapı sunabilmektedir. Bütün bu özellikleri dikkate alındığında yeni nesil Telekom Operatörleri kendi telefon sistemlerini IP tabanlı yapılara dönüştürme eğilimindedirler. PSTN şebekesi üzerinden direkt olarak VoIP şebekesine erişim sağlanamaz. Telekom Operatörleri ağ geçidi (Gateway) cihazları kullanarak PSTN ile VoIP şebeke arasındaki iletişimi sağladıkları tümleşik bir topolojiye sahiptirler [1].

VoIP üzerinden görüşme yapılırken en önemli unsurlardan biri ses iletiminin kalitesidir. VoIP, internet tabanlı olduğundan ve ses iletimi paketler şeklinde sağlandığından internet bağlantı kapasitesi ve trafik yoğunluğuna bağlı olarak bazı paketlerin iletilememesi veya geç iletilmesi olağandır. Bunlar da ses kalitesini bozan etkenlerdendir. Bu yüzden ses kalitesinin ölçümü Telekom Operatörleri için oldukça önemlidir. Servis kalitesi (Quality of Service, QoS) ve deneyim kalitesi (Quality of Experience, QoE) ile VoIP sistemlerin kaliteleri belirlenir.

QoS ile ses paketlerine ait paket kayıp (Packet Loss) oranları ve gecikme (Delay) süreleri tespit edilerek ağın performansı ortaya çıkartılabilir [1]. Delay, bir ses paketinin kaynaktan hedefe teslim edilirken meydana gelen ve tüm iletişim ağındaki genel gecikmeyi ifade etmektedir. VoIP 'te gerçekleşen delay süresi, paket kodlama, paket oluşturma, fiziksel ağ ve yönlendirme cihazları başta olmak üzere kod çözme ve paketin tekrar iletilmesi gibi sebeplerden kaynaklanabilmektedir [2-4]. Packet loss, birçok faktörden oluşabilmektedir. Ancak en temel sebebi, ağ trafığında yaşanan sıkışıklığa bağlı olarak ses paketlerinin iletilememesidir. VoIP servislerinin çoğu iletim katmanı protokollerinden olan UDP protokolünü tercih ettiği için kaybolan (Drop) ses paketleri tekrar gönderilmez. Eğer bir ses paketi zamanında hedefe ulaşamazsa söz konusu paket VoIP uygulaması tarafından atılır [2,3,5]. Bu durumda alıcı tarafında paket kaybı olarak algılanmaktadır.

QoE ise Qos parametrelerine ilave olarak servisin tamamı hakkında kullanıcının kişisel görüşünü de dahil ederek bir kalite derecesi oluşturmaktadır [1]. QoE, bir konuşmacıya ait sesin dinleyici tarafından algılanma netliğini ifade etmektedir. Genel olarak kalite, beklenen ve algılanan arasındaki fark olarak formüle edilebilir [6,7,8]. Konuşma kalitesinin ölçümü, sesli telefon hizmetlerinde müşteri deneyimini ortalama görüş puanı (Mean Opinion Score, MOS) ile ifade etmek anlamına gelmektedir [6]. QoE'nin ölçümü için ITU-T tarafından çeşitli ölçüm yöntemleri sunulmuştur. Bu yöntemlerden bazıları POLQA, PESQ, PSQM, PEAQ 'dir.

PSQM (Perceptual Speech Quality Measure), ITU-T tarafından 1996 yılında P.861 standartı olarak kabul etmiştir. PSQM, hesaplama ve modelleme algoritmasıdır. 300-3400 Hz aralığındaki ses kalitesini değerlendirip ölçer. PAQM'in (Perceptual Audio Quality Measure) daha genelleştirilmiş ve telefon konuşma işaretleri için optimize edilmiş halidir. PSQM algoritması, konuşma işaretinin farklı kodek ve iletim ortamlarında yaşadığı bozulmaları ölçer. Bozulma, zaman veya frekans bölgesinde ölçülmediği için sinyal-gürültü oranından (Signal to Noise Ratio - SNR) farklılık gösterir. Bozulma fiziko-akustik bölgede ölçülür. Böylece gerçek telefon kullanıcılarına benzer bir ses alımı sağlanır [1,9].

PEAQ (Perceptual Evaluation of Audio Quality), ITU-R (International Telecommunication Union's Radiocommunication Sector) tarafından 1998 yılında ITU-R BS.1387 standartı olarak kabul etmiştir. PEAQ, alınan sesin nesnel olarak kullanıcı deneyimini ölçmek amacıyla kullanılan bir standarttır. Bu standartta, insan kulağının özellikleri temel alınarak oluşturulan parametreler birleştirilerek bir çıkış değeri hesaplanır. PEAQ ölçümünde, test ve referans ses örnekleri hizalanarak her örneğin karşılaştırılmasıyla QoE kalitesi MOS cinsinden hesaplanır [1,10].

ITU-T Şubat 2001'de [7], PESQ (Perceptual Evaluation of Speech Quality) kalite ölçüm modelini P.862 standartı olarak kabul etmiştir. ITU-T P.861 kodlu PSQM'in geliştirilmesiyle oluşturulmuştur. PESQ, iletim kanalına gönderilen orijinal referans sinyali ile

iletim ortamı çıkışında elde edilen bozulmuş sinyali karşılaştırabilen bir yöntemdir. Bu yöntemde sonuç hesaplamaları için insan yorumuna ihtiyaç duymayan otomatik bir test mekanizması bulunmaktadır [6]. PESQ'in prensibi, iletim ortamı çıkışındaki işaret ve referans işaretin öncelikle zaman ve frekansları hizalanır. Daha sonra filtreden geçirilen iki sinyal de eş zamanlanır ki arada VoIP nedeniyle oluşan gecikmelerin, zaman kayıplarının önüne geçilebilsin. Sesteki bozulmaları açıklamak için iki işarete frekans bölgesine geçirilerek birbirlerinden farkları alınır. Böylece iki işaretin ses farklılıkları ortaya çıkar. Zamanla gürültünün mü eklendiği yoksa işarete kayıp mı olduğuna bağlı olarak ses farkları ağırlıklandırılıp toplanır. Bundan sonra da MOS değeri hesaplanarak sesin kalitesi kestirilir. PESQ'nun en önemli avantajı ağ üzerinde veya bozulma türleri hakkında herhangi bir varsayım yapılmasına gerek kalmamasıdır [1,11]. Bu nedenle, IP ağında ses kalitesini test etmek amacıyla PESQ tercih edilmektedir [6].

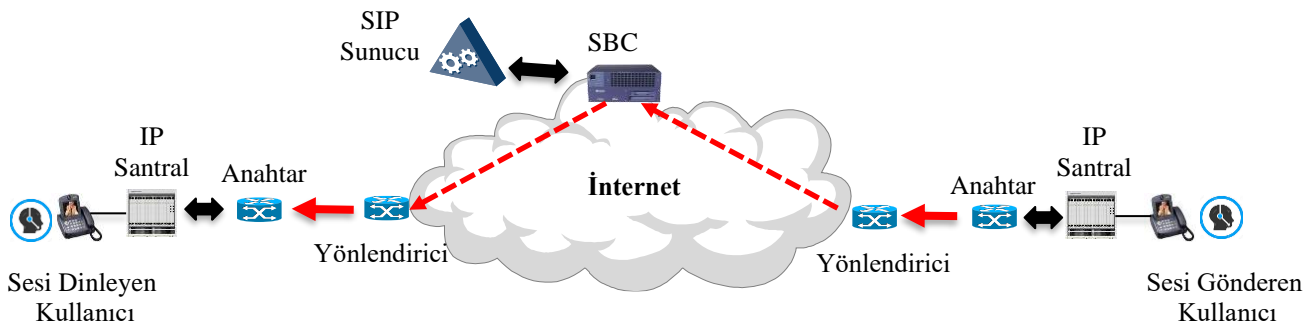
2010 yılında yeni bir konuşma kalitesi ölçüm yöntemi olan POLQA (Perceptual Objective Listening Quality Analysis) modeli geliştirilmiştir [8,12-15]. Bu yöntem aslında PESQ 'nin geliştirilmiş halidir. POLQA, sabit, mobil ve IP tabanlı ağlar için geliştirilen yeni nesil ses kalitesi ölçüm teknolojisidir. POLQA, 2011 yılında ITU-T tarafından HD ses, 3G ve 4G / LTE ağlarının ses kalite analizi için P.863 olarak standartlaştırılmıştır [6]. POLQA 'nın yüksek arka plan gürültülerini ölçebilmesi, süper geniş bant (50Hz - 14kHz) aralığında ses analizi yapabilmesi ve akustik arayüzlerin kaydedilebilmesi gibi avantajları vardır. Bu özellikleri ile PESQ 'ya göre üstünlük sağlamaktadır. Ayrıca literatürdeki çalışmalar, POLQA 'nın PESQ 'ya göre dar bant aralığı için %27 oranında, geniş bant aralığında ise %56 oranında daha başarılı sonuçlar elde ettiğini göstermektedir [16,17]. Günümüzde birçok Telekom Operatörü kendi şebekelerindeki ses iletim kalitesini ölçmek için POLQA standartını tercih etmektedir.

Yukarıda belirtilen modellerin birçok avantajı olmasına rağmen referans işaret gereksiniminden dolayı uygulamalarda pratik zorluklar oluşmaktadır [1]. Buyüzden bu çalışmada, internet üzerinden ses iletimine ait MOS kalite değerlerini tahmin etmek amacıyla referans sinyal gerektirmeyen bir model yapısı tasarlanmıştır. Bu model sadece QoS parametrelerini dikkate alan YSA (Yapay Sinir Ağları) yapısından oluşmaktadır. Önerilen bu modelin eğitim süresi zaman alıcı olsa da en büyük avantajı referans ve iletim ortamında bozulan ses sinyallerine ihtiyaç duymamasıdır. Ayrıca model farklı ağ parametrelerini kullanılarak MOS tahmini yapabildiği için de esnek bir yapıya sahiptir. Literatür taraması yapıldığında, VoIP servis tiplerine ait MOS kalite değerlerini tahmin etmek amacıyla YSA tabanlı yapıların kullanıldığı gözlenmiştir [18-22]. Fakat literatürdeki çalışmalarda, PESQ yöntemi ile ölçülen MOS değerlerinin YSA modelleri ile tahmin edildiği fakat POLQA yöntemi ile elde edilen MOS değerleri için bu tip bir çalışmanın olmadığı tespit edilmiştir. Bu yüzden bu makalede, POLQA yöntemi ile ölçülen MOS kalite değerlerinin YSA model ile tahmin edilmesine çalışılmıştır. Makale bu yönüyle de literatüre katkı sağlamaktadır.

2. VoIP Test Altyapısı ve MOS Ölçüm Yöntemi

2.1. VoIP Test Altyapısı

VoIP teknolojisi ile fiziksel bir işaret olan ses uygun dönüştürücülerle elektriksel işarete dönüştürülerek, IP ağı üzerinden karşı tarafa iletilir. Bu iletim gerçekleştirilirken, ses işaretleri öncelikle kodlenerek bitlere dönüştürülüp karşı tarafın anlayacağı şekilde paketlenir daha sonra internet üzerinden iletilir. Türk Telekom şebekesinde IP ağları üzerinden verilen ses hizmetlerinde SIP (Session Initiation Protocol) protokolü kullanılmaktadır. SIP basit ve genişleyebilir bir protokoldür. RFC 2543 ve 3261 tarafından tanımlanır. Bir veya daha fazla katılımcı arasında multimedia oturumları oluşturur ve sonlandırır. Şekil 1 'de Türk Telekom şebekesinde VoIP hizmetini simüle etmek için oluşturulan test topolojisi bulunmaktadır. Müşterinin lokal ağlarında IP temelli santraller bulunmakta olup bu santraller bir yazılım veya donanımsal bir cihaz olabilir. Gönderici tarafından iletilen ses paketi internet omurgası üzerinden ilk olarak SIP sunucuya ulaşır. Burada gerekli yetkilendirme ve yönlendirme işlemleri yapıldıktan sonra alıcı tarafındaki IP santrale iletilir ve böylece IP ağı üzerinden ses görüşmesi sağlanmış olur. SIP sunucu ile internet omurgası arasında SBC (Session Border Controller) cihazı bulunmaktadır. SBC, internet servis sağlayıcılarının kendi VoIP ağlarına dış dünyadan erişim vermek, güvenlik sağlamak, proxy gibi güvenlik özellikleri için kullandıkları cihazlara verilen isimdir.



Şekil 1. VoIP Test Topolojisi

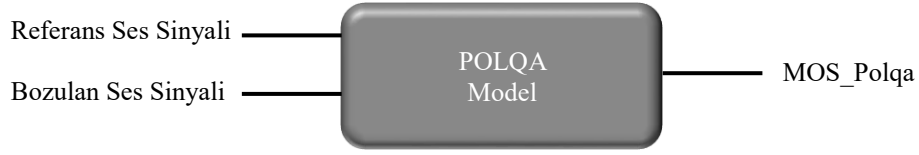
Uzun yıllar süren araştırmalar sonucunda insanların sesleri algılayış biçimleri değerlendirilmiş ve bir notlandırma sistemi oluşturulmuştur. MOS derecelendirmesi olarak adlandırılan bu sistem ITU'nun P.800 tavsiyesini oluşturmaktadır. POLQA model için MOS dereceleri 4,5 ile 1,0 arasında değişirken (Tablo 1' de gösterilmiştir), müşterilerin bu hizmetten tatmin olma yüzdeleri de MOS değeri ile orantılı olarak değişmektedir [16]. Şekil 1 'deki test ortamında, kaynaktan hedefe gönderilen farklı paket kayıp oranına ve delay süresine sahip ses paketleri için MOS kalite değerleri POLQA model kullanılarak ölçülmüştür. Bu test 183 farklı ses görüşmesi için tekrarlanmış ve elde edilen MOS değerleri kaydedilmiştir.

Tablo 1. MOS Değerlerinin Sınıflandırılması.

MOS derecesi	Kalite	Sesteki Bozulma
4,5	Çok iyi	Algılanamaz
4,0	İyi	Algılanabilir fakat rahatsız edici değil
3,0	Orta	Kısmen rahatsız edici
2,0	Yetersiz	Rahatsız edici
1,0	Kötü	Çok rahatsız edici

2.2. MOS Ölçüm Yöntemi

Şekil 2’ de POLQA model yapısı gösterilmiştir. Bu model yapısı temel olarak referans ses sinyali ile bozulan ses sinyali karşılaştırılarak MOS değerini üretmektedir. Referans ses sinyali, kaynak tarafından gönderilen orijinal ses sinyali olarak tanımlanırken, bozulan ses sinyali ise hedef tarafında duyulan ses sinyali olarak tanımlanmaktadır [13]. Gönderilen referans ses sinyali, IP ağı üzerinde ilerlerken gürültü, paket kaybı ve delay gibi faktörlerden olumsuz etkilendiği için alıcı tarafında bozulmuş bir ses sinyali olarak algılanabilmektedir. Burada MOS_Polqa ile ölçülen MOS değeri ifade edilmektedir.

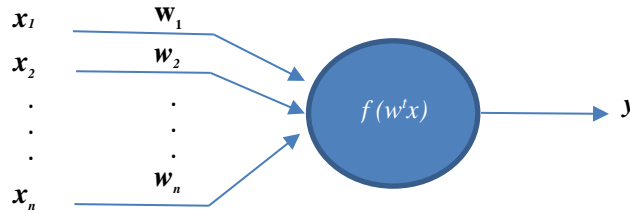


Şekil 2. POLQA Model Yapısı

3. Çok Katmanlı Yapay Sinir Ağları

YSA, biyolojik sinir ağlarının işlevlerinden esinlenilerek oluşturulmuş matematiksel modellerdir. YSA, insan beynini taklit ederek öğrenme yolu ile yeni bilgiler türetebilme, oluşturabilme ve keşfetme yeteneklerine sahip bilgi işleme sistemleridir. Bu model yapısı, sistemlere veya problemlere ait giriş-çıkış verilerini yorumlayarak öğrenir ve gelecekte hangi durumlar için nasıl sonuçlar üretilebileceğini tahmin edebilir [23,24].

Yapay sinir hücreleri (Nöronlar), YSA’larının temel bilgi işleme birimidir. Şekil 3 ‘te nöron yapısı gösterilmiştir. Nöronlar, bağlantı hatları üzerinden işaret göndererek birbirlerini etkilerler. Bağlantı hatlarının her birinin kendine özgü bir ağırlığı vardır. Bilgileri depolayan bağlantı ağırlıkları nöronları paralel ve ardışık bir biçimde birleştirirler. Uyarılamalı hesaplanabilen bu ağırlıkların aldıkları değerler; bilginin en doğru biçimde iletilmesini sağlar. Genellikle, bağlantı ağırlıkları bir öğrenme işlemi ile belirlenir. YSA ‘nın hafızaları olarak nitelendirilen bu ağırlıklar, sistemde modele ait bilginin taşındığı noktalardır. YSA ‘nın eğitilmesi, sisteme verilen örnek girdiler için beklenen çıktı değerlerini üretmesini sağlayacak ağırlık değerlerinin bulunması işlemidir [25].



Şekil 3. Nöron Yapısı

Temel bir nöron çıkış sinyali eşitlik (1) ‘deki gibi ifade edilmektedir [26].

$$y = f(w^t x) \quad (1)$$

Burada w ağırlık vektörü, x ise giriş vektörü olup eşitlik (2) ve (3) ‘te verilmiştir.

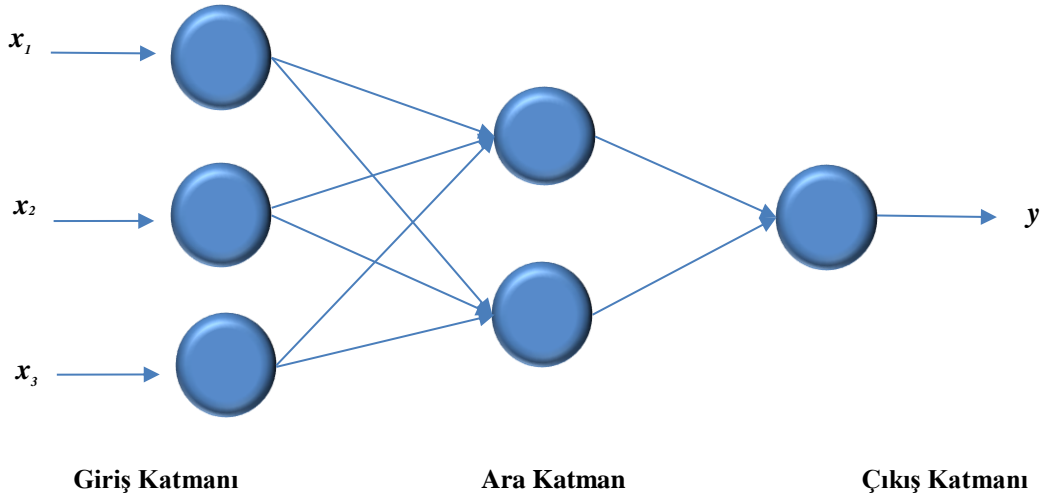
$$w = [w_1 w_2 \dots w_n]^t \quad (2)$$

$$x = [x_1 x_2 \dots x_n]^t \quad (3)$$

Genellikle eşitlik (4) ‘teki gibi bir net değişkeni tanımlanır ve aktivasyon fonksiyonu $f(net)$ biçiminde kullanılır [26].

$$net = w^t x \quad (4)$$

YSA yapısı temelde giriş, ara ve çıkış olmak üzere üç katmandan oluşmaktadır. Giriş katmanı, uygulanan probleme ait eğitim örneklerinden oluşan veri kümesinden gelen bilgileri ara katmana iletmektedir. Ara katman, giriş katmanından gelen bilgiyi işleyip çıkış katmanına göndermektedir. Ara katman sayısı birden fazla olabilmektedir. Çıkış katmanı ise ara katmandan gelen bilgiyi işleyerek, ilgili girişe karşı üretilecek çıkışı hesaplamakta ve çıkış veri kümesinin oluşumunu sağlamaktadır [23,24,27,28]. Şekil 4’te çok katmanlı bir YSA yapısı verilmiştir.



Şekil 4. Çok Katmanlı YSA Yapısı

YSA hücre modellerinde kullanılan çeşitli tipte aktivasyon fonksiyonları bulunmaktadır. En uygun tanımlamanın yapılabilmesi için aktivasyon fonksiyonunun seçilmesi önemlidir. Lineer, basamak, eşik, hiperbolik tanjant gibi türleri olan aktivasyon fonksiyonunun en yaygın kullanılan tipi sigmoid fonksiyonlardır. YSA'nın eğitiminde, tanjant sigmoid (Tansig), logaritmik sigmoid (Logsig) ve lineer (Purelin) aktivasyon fonksiyonları ara katman ve çıkış katmanları için literatürde sıklıkla tercih edilmektedir [23,29].

4. Simülasyon Çalışmaları

Bu çalışmada, internet üzerinden ses iletimi yapılırken ölçülen MOS kalite değerlerini tahmin etmek amacıyla referans sinyal gerektirmeyen YSA model yapısı tasarlanmıştır. YSA modelin tercih edilme sebebi ise öğrenme yolu ile yeni bilgiler türetebilme, oluşturabilme ve keşfetme yeteneklerine sahip olmasıdır. Önerilen model yapısı Şekil 5 'te gösterilmiştir. Model yapısı 2 giriş ve 1 çıkıştan oluşmaktadır. Şekil 5 'te sunulan yapı, delay ve paket kaybı gibi ağ parametrelerini dikkate alarak MOS çıkış değerini tahmin edebilen çok katmanlı YSA 'dır. Burada MOS_Ysa ile tahmin edilen MOS değeri ifade edilmektedir.



Şekil 5. MOS Tahmini İçin Tasarlanan Çok Katmanlı YSA Model

Genellikle YSA mimarisinin ileri beslemeli ve tekrarlamalı olmak üzere iki tip vardır. Bu çalışmada ileri beslemeli tipi tercih edilmiştir. Şekil 1' deki topoloji üzerinde 183 adet test yapılmış ve elde edilmiş veriler kaydedilmiştir. Bu veriler içerisinde rastgele seçilen 157 adet veri YSA 'nın eğitiminde kullanılırken 26 adet veri ise YSA 'nın testi için kullanılmış ve MOS değerleri tahmin edilmeye çalışılmıştır. Oluşturulan model için dikkate alınan veri aralıkları Tablo 2 'de verilmiştir. Bu sınır değerleri çerçevesinde gerçek değerler [-1, +1] aralığına normalize edilerek modelin gerçek MOS değerlerine daha yakın sonuçlar elde etmesi amaçlanmıştır. YSA 'nın simülasyonu Matlab platformunda gerçekleştirilmiştir. Simülasyon çalışmalarında modelin eğitimi için 1000 adımlık bir benzetim çalışması yapılmıştır. Eğitimlerde farklı adım büyüklükleri de testler yapılmış ancak optimum sonuçlar 1000 adımlık testlerde elde edilebildiği için bu şekilde seçilmiştir.

Tablo 2. Model İçin Dikkate Alınan Veri Aralıkları

Delay (ms)	Paket Kayıp Oranı (%)	MOS Değeri
1 - 57	0 - 0,42	4,40 - 4,50

Bu çalışmada, farklı paket kayıp oranı ve delay değerlerine sahip ses paketlerinin MOS değerleri için çeşitli simülasyonlar gerçekleştirilmiştir. POLQA yöntemi ile ölçülen MOS değerlerinin IP ağındaki delay ve paket kaybı verilerine bağlı olarak değiştiği gözlemlenmiştir. YSA modeli içerisindeki ara katmanda, farklı sayılarda nöron kullanılarak simülasyonlar gerçekleştirilmiştir. YSA modelin parametre eğitimleri için üç farklı algoritma kullanılmıştır. Bu algoritmalar LM (Levenberg Marquardt), BR (Bayesian Regulation) ve RPROP (Resilient Backpropagation) 'dır. LM, BR ve RPROP algoritmaları kullanılarak tasarlanan modelin eğitim ve test verileri için sergilediği en iyi minimum, ortalama ve maksimum yüzde hataları Tablo 3 'de gösterilmiştir. Tablo 3' ün oluşturulması amacıyla her algoritma için ara katman tipi 2x2, 2x3, 2x4, 2x5, 3x2, 3x3, 3x4, 3x5, 4x2, 4x3, 4x4, 4x5, 5x2, 5x3, 5x4, 5x5 olmak üzere 16 farklı şekilde tasarlanarak testleri yapılmıştır. Bütün algoritmalara ait testlerdeki en iyi sonuçlar seçilerek karşılaştırma yapılabilmesi amacıyla Tablo 3 'te sunulmuştur.

Tablo 3. MOS Problemi için Tasarlanan Çok Katmanlı YSA Modelin Performansı.

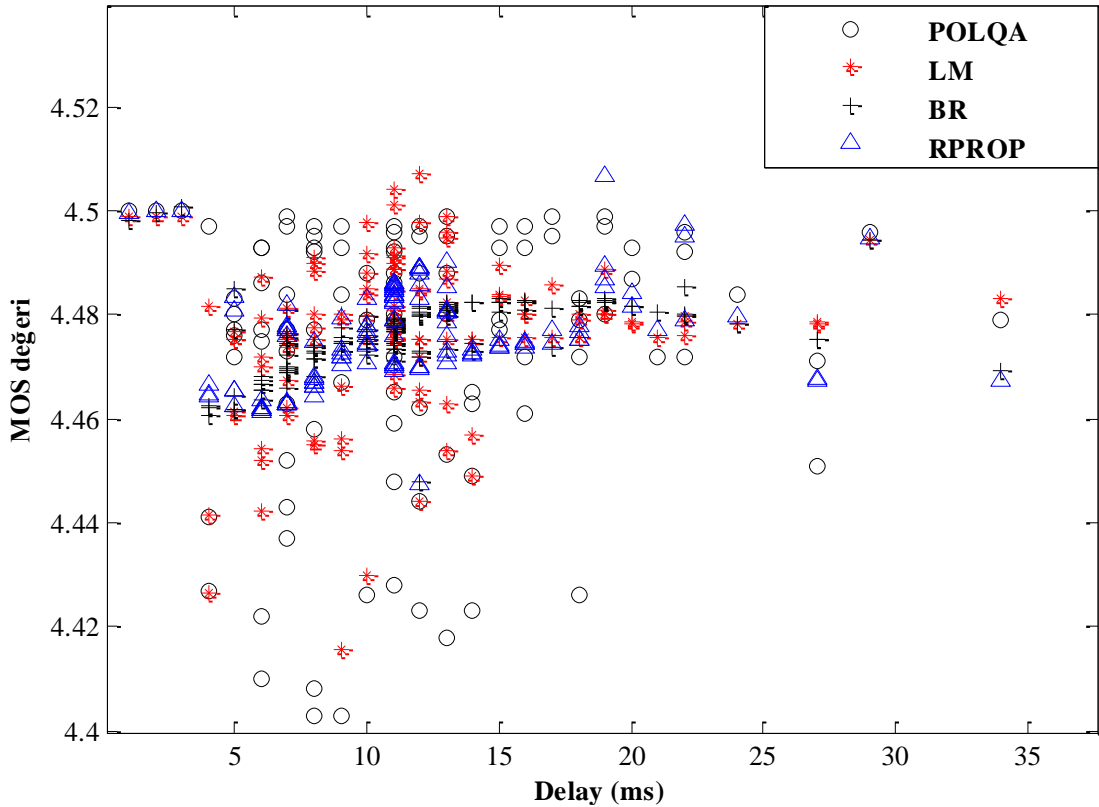
Algoritma Tipi	Ara Katman Tipi	Simülasyon Tipi	Minimum % Hata	Ortalama % Hata	Maksimum % Hata	MSE
LM	(3x4)	Eğitim	1,6712 e-04	2,4039 e-01	1,8227	2,7581 e-04
		Test	6,6030 e-03	5,6452 e-01	1,8227	7,4331 e-04
BR	(4x4)	Eğitim	1,1705 e-03	3,0221 e-01	1,5893	3,9588 e-04
		Test	9,0971 e-02	8,2331 e-01	1,5893	4,5291 e-04
RPROP	(4x5)	Eğitim	3,8259 e-04	2,7987 e-01	1,5290	3,6987 e-04
		Test	2,7954 e-03	7,6368 e-01	1,5290	4,6254 e-04

YSA kullanılarak elde edilen modelleme hataları, eşitlik (5)'te verilen ortalama karesel hata (Mean Square Error, MSE) yöntemi ile hesaplanmıştır. Bu performans kriteri, YSA modelinin başarısını ortaya koymak için yardımcı olmuştur.

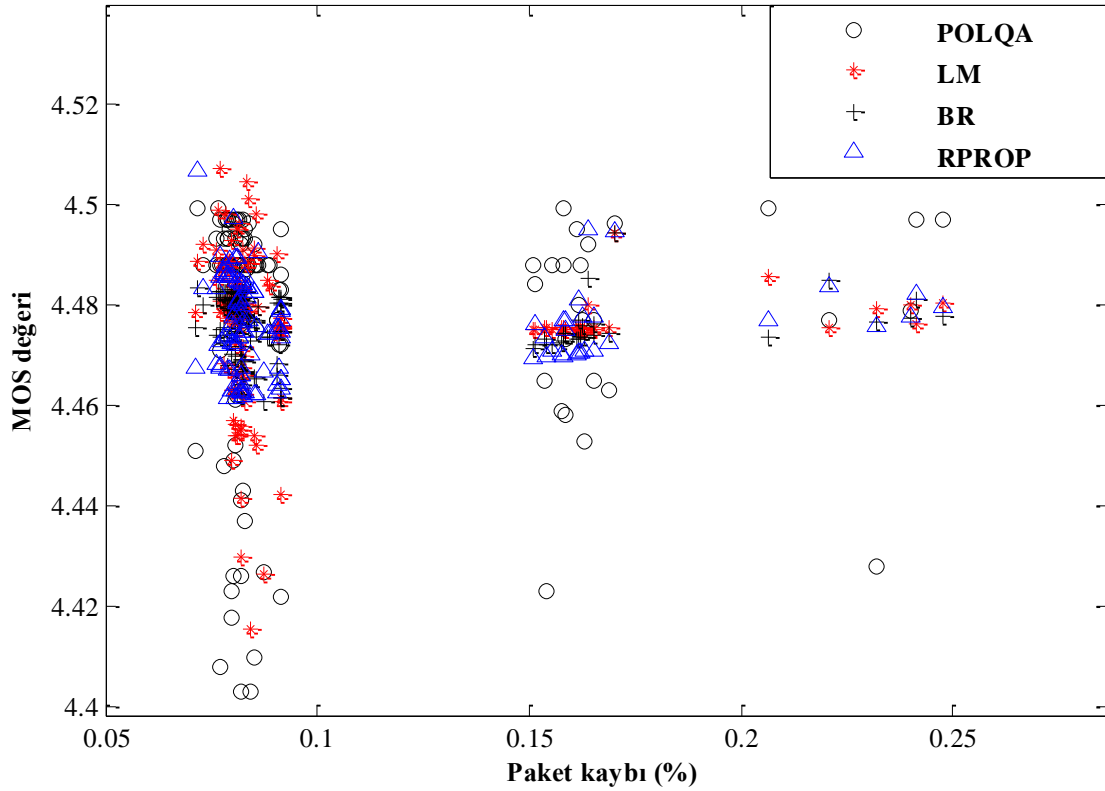
$$MSE = \frac{1}{N} \sum_{n=1}^N (d(n) - y(n))^2 \quad (5)$$

Bu ifadeye verilen $d(n)$, istenilen çıkış değeri, $y(n)$, gerçek çıkış değeri ve N ise örnek sayısıdır. Tablo 3 'de verilen çıktılara göre elde edilen sonuçlar kabul edilebilir değerler arasındadır. Çalışma da kullanılan algoritmalar karşılaştırıldığında ise birbirine yakın sonuçlar elde etmelerine rağmen LM algoritmasının daha başarılı sonuçlar ürettiği gözlenmiştir. Tabloya göre en iyi sonuç ise modelin (3x4) 'lük kural yapısında iken LM algoritması kullanılarak elde edilmiştir. Bu modelde (3x4) olarak gösterilen değerler gizli katmanda yer alan işlemci eleman sayılarını belirtmektedir.

Eğitim sürecindeki POLQA ve tahmin edilen MOS çıkış değerlerinin, delay giriş verileri ile olan ilişkisi Şekil 6 'da verilmiştir. Şekil 7' de ise POLQA ve tahmin edilen MOS çıkış değerlerinin, paket kaybı giriş verileri ile olan ilişkisi verilmiştir. Heriki şekilde de YSA modelin 3 farklı algoritma kullanılarak eğitildiği süreçteki MOS tahminleri gösterilmiştir. Şekillere göre LM algoritması kullanıldığında model çıkışının POLQA çıkışına çok yaklaştığı tespit edilmiştir. Bu şekillerde \circ sembolü POLQA modele ait MOS değerini, $*$ sembolü LM algoritması kullanılan YSA modelin MOS tahminini, $+$ sembolü BR algoritması kullanılan YSA modelin MOS tahminini, \triangle sembolü ise RPROP algoritması kullanılan YSA modelin MOS tahminini temsil etmektedir.

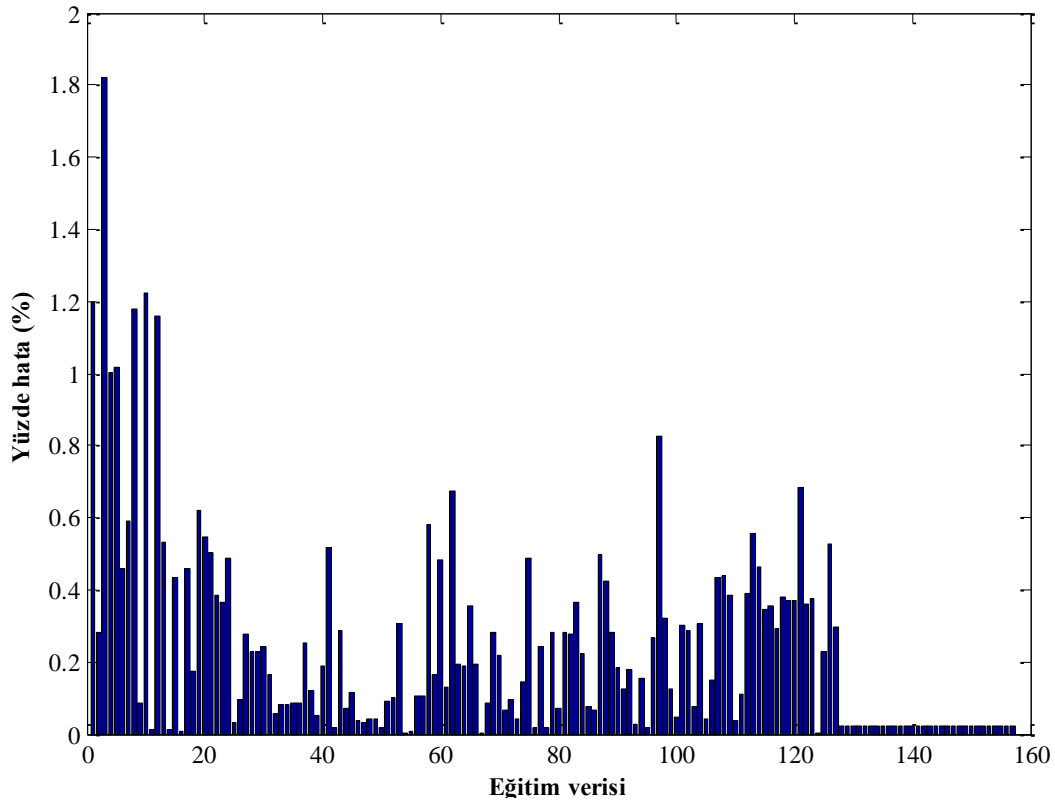


Şekil 6. Model Eğitiminde Delay & MOS İlişkisi

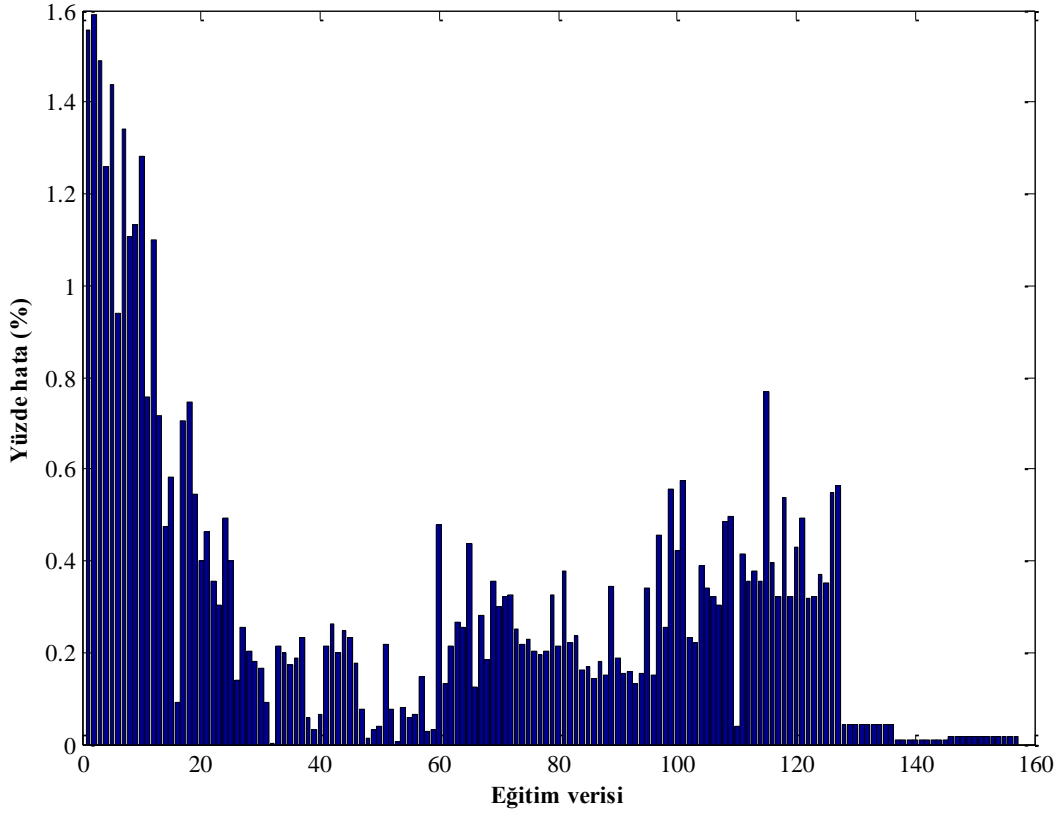


Şekil 7. Model Eğitiminde Paket Kaybı & MOS İlişkisi

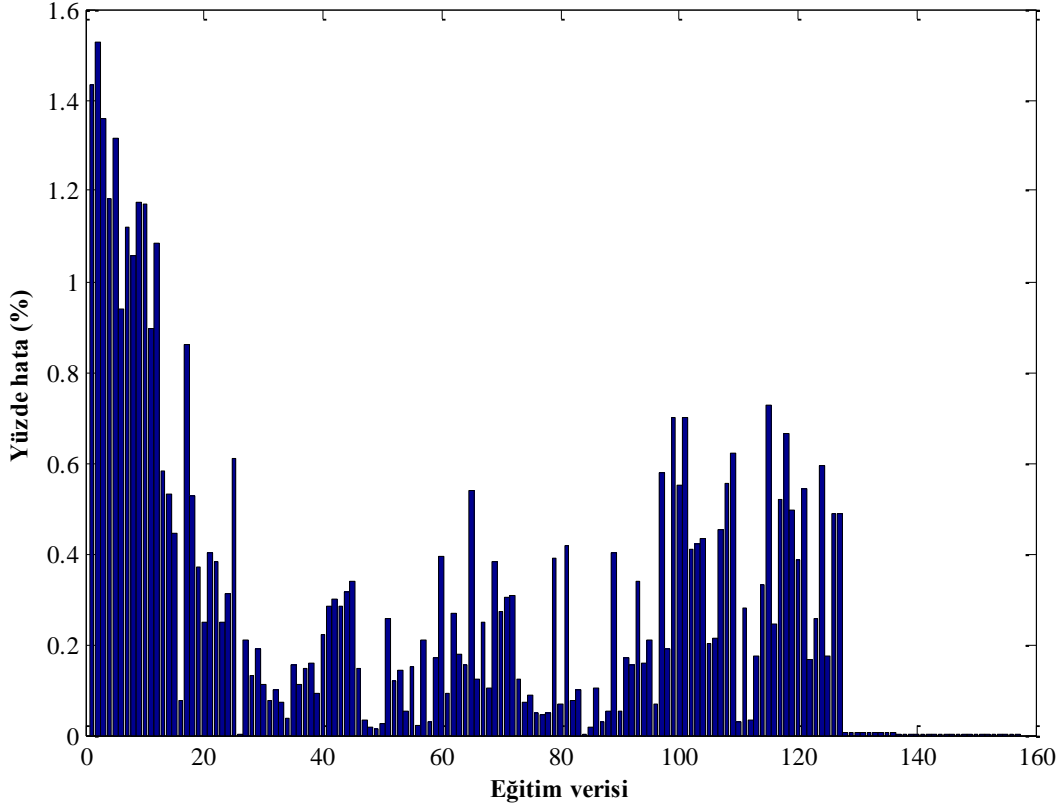
Bu eğitim sürecinde en iyi sonucu veren (3x4) kuralına sahip YSA modelin LM algoritması kullanıldığı durumda elde ettiği hata yüzdeleri ise Şekil 8 'de gösterilmiştir. Hata yüzdesinin ilk başta %1,2 olduğu ve eğitim süreci ilerledikçe hatanın % 0,2 'den daha aşağıya indiđi gözlenmiştir. BR (4x4 kural) ve RPROP (4x5 kural) algoritması kullanıldığı durumda elde edilen hata yüzdeleri ise Şekil 9 ve 10 'da verilmiştir. LM' ye ait hata yüzde grafiğinin, BR ve RPROP hata yüzde grafiklerine göre daha hızlı ideal sonuca yakınsamaya çalıştığı görülmüştür.



Şekil 8. LM Algoritması ile Model Eğitimindeki Hata Yüzdeleri

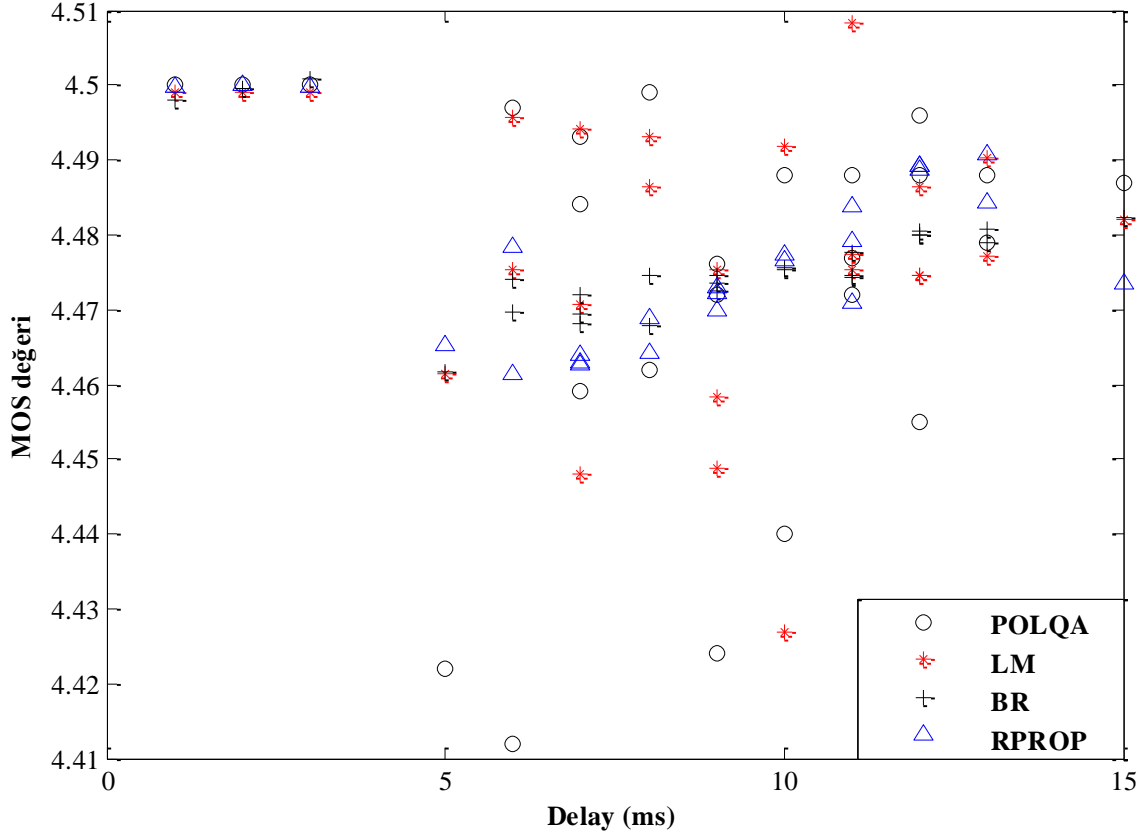


Şekil 9. BR Algoritması ile Model Eğitimindeki Hata Yüzdeleri

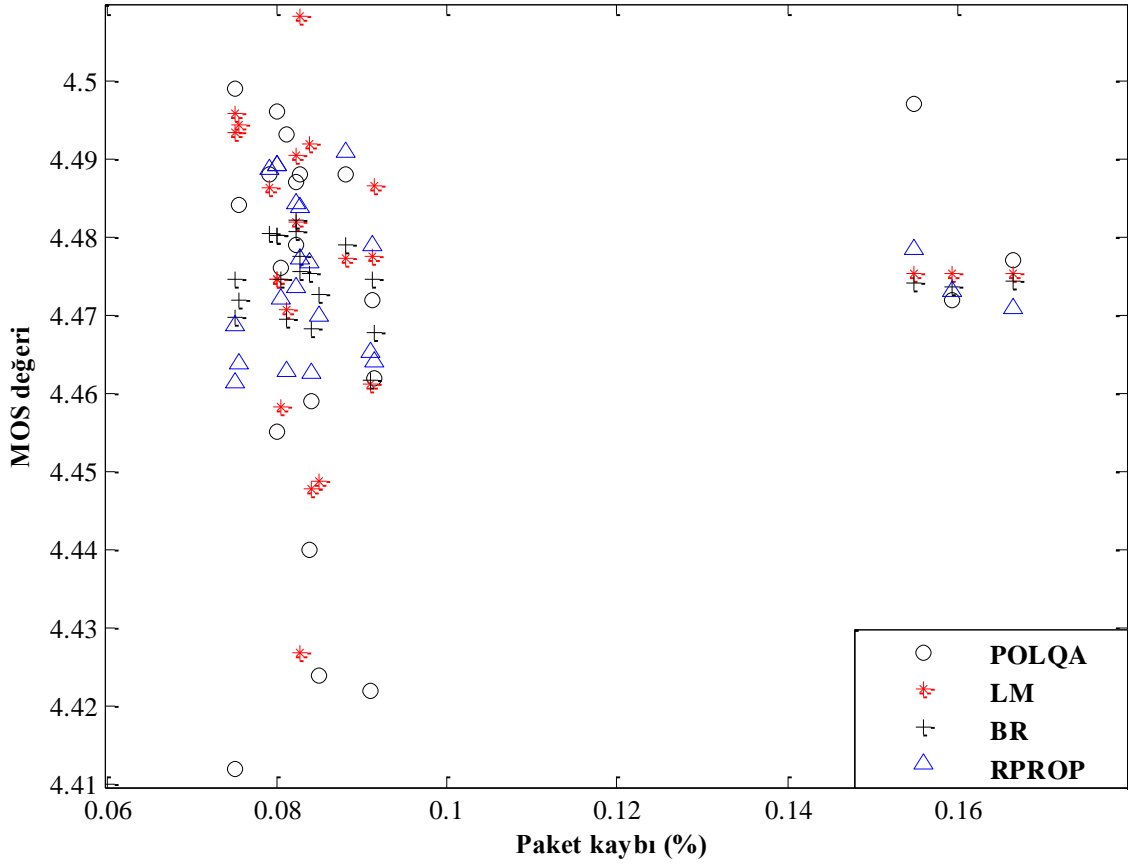


Şekil 10. RPROP Algoritması ile Model Eğitimindeki Hata Yüzdeleri

Test sürecindeki POLQA ve tahmin edilen MOS değerlerinin, delay giriş verileri ile olan ilişkisi Şekil 11 'de verilmiştir. Şekil 12' de ise POLQA ve tahmin edilen MOS değerlerinin, paket kaybı giriş verileri ile olan ilişkisi verilmiştir. Heriki şekilde de YSA modelin 3 farklı algoritma kullanılarak test edildiği süreçteki MOS tahminleri kullanılmıştır. Şekillere göre LM algoritması kullanıldığında model çıkışının POLQA çıkışına çok yaklaştığı tespit edilmiştir.



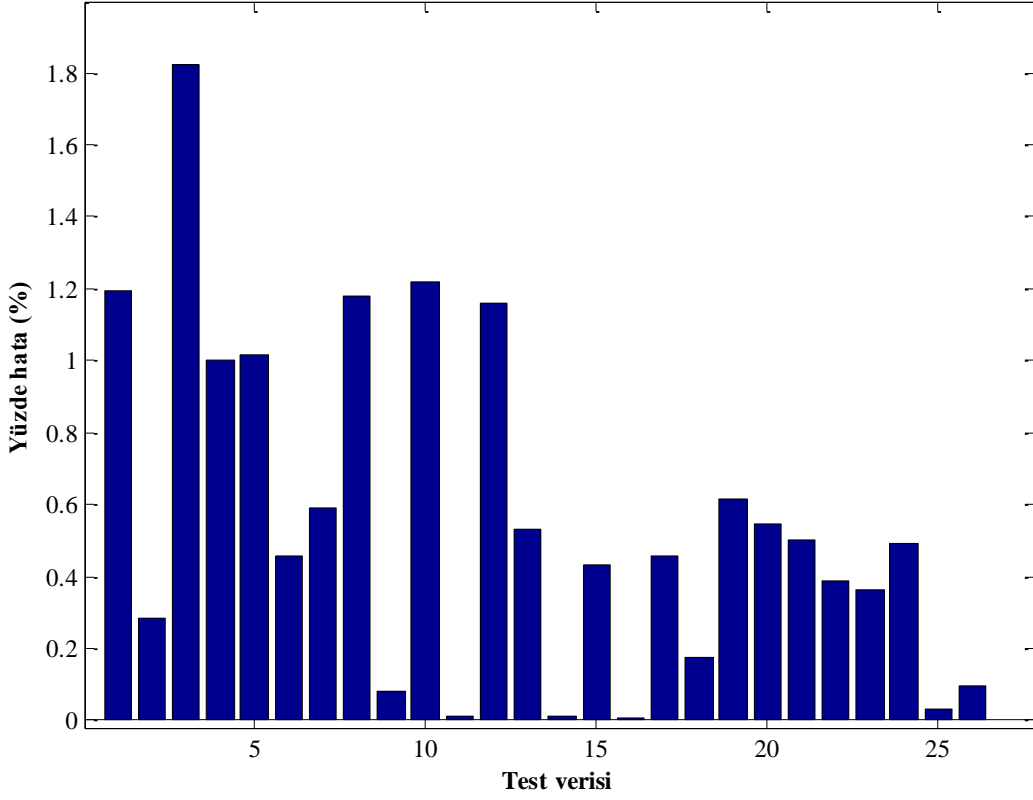
Şekil 11. Model Testinde Delay & MOS İlişkisi



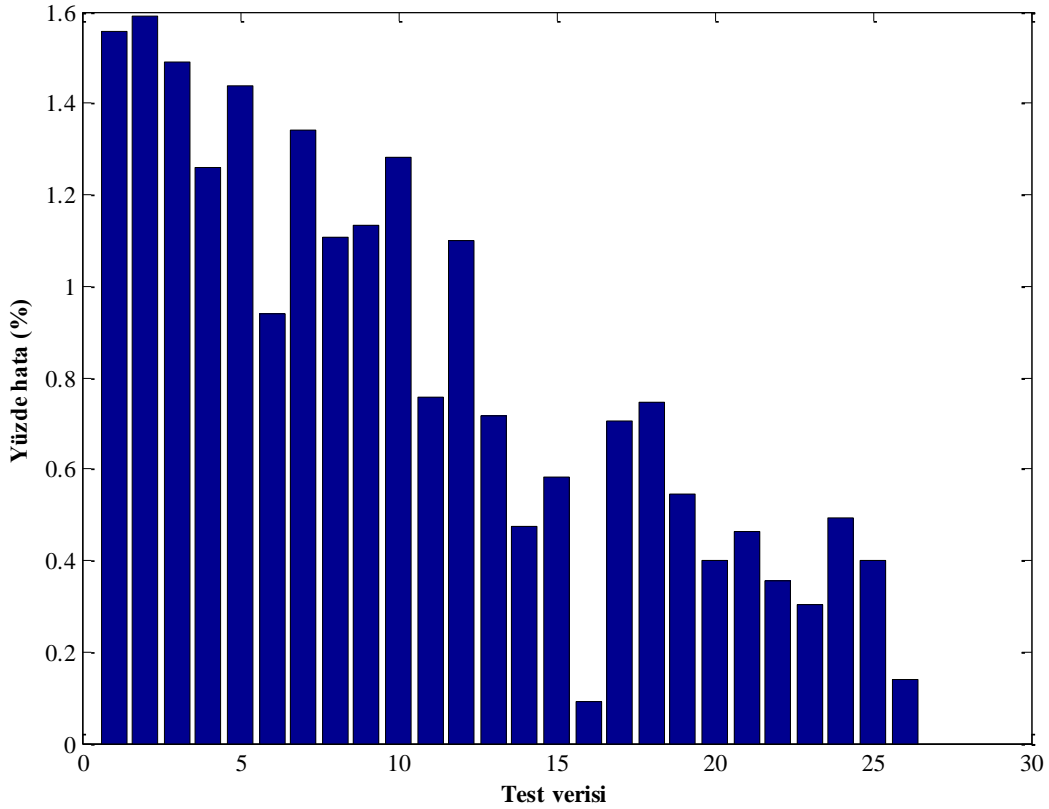
Şekil 12. Model Testinde Paket Kaybı & MOS İlişkisi

Test sürecinde en iyi sonucu veren (3x4) kuralına sahip YSA modelin LM algoritması kullanıldığı durumda elde ettiği hata yüzdeleri ise Şekil 13 'te gösterilmiştir. Hata yüzdesinin ilk başta %1,2 iken üçüncü test verisinde 1,8'e kadar çıktığı ve daha sonraki test verilerinde ise hatanın sürekli azaldığı gözlenmiştir. BR (4x4 kural) ve RPROP (4x5 kural) algoritması kullanıldığı durumda elde edilen

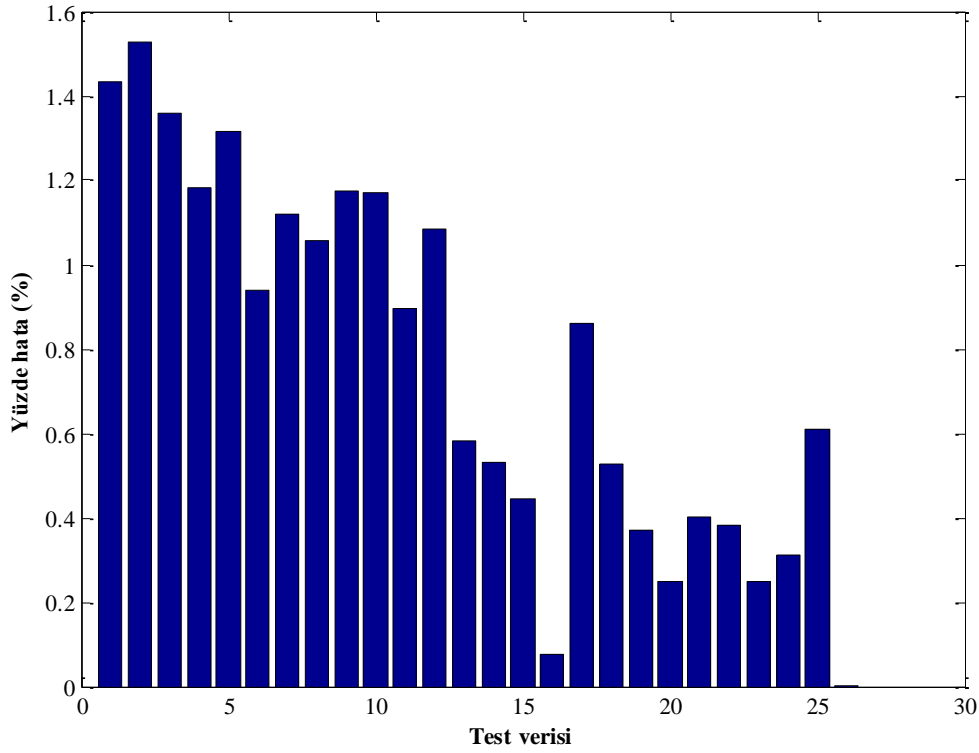
hata yüzdeleri ise Şekil 14 ve 15 'te verilmiştir. LM' ye ait hata yüzde grafiğinin, BR ve RPROP hata yüzde grafiklerine göre daha hızlı ideal sonuca yakınsamaya çalıştığı görülmüştür.



Şekil 13. LM Algoritması ile Model Testindeki Hata Yüzdeleri



Şekil 14. BR Algoritması ile Model Testindeki Hata Yüzdeleri



Şekil 15. RPROP Algoritması ile Model Testindeki Hata Yüzdeleri

Çalışmadaki bütün sonuçlar genel olarak değerlendirildiğinde, tasarlanan YSA tabanlı modelin uygun şekilde eğitildiği zaman POLQA yöntemi ile ölçülen MOS değerlerini yüksek doğrulukta tahmin edebildiği gösterilmiştir. Elde edilen değerler bu sonucu destekler niteliktedir. Modellere ait ortalama yüzde hatalar da dikkate alındığında LM algoritması kullanılan YSA modelin biraz daha öne çıkan bir yaklaşım olduğunu söylemek yanlış olmayacaktır. Ayrıca, MOS tahmin probleminin çözümünde, önerilen YSA tabanlı modelin, mevcut yöntemlere alternatif olacağı da söylenebilir.

5. Sonuçlar

VoIP, internet platformunda en çok talep gören uygulamalardan biridir. Bu çalışmada, gerçek zamanlı IP ağları üzerindeki ses kalitesini (MOS) analiz etmek amacıyla çok katmanlı YSA model tasarlanmıştır. Burada ki esas odak noktası ITU-T tarafından önerilen POLQA yöntemindeki gibi herhangi referans ve bozulmuş ses sinyaline ihtiyaç duymadan alıcı tarafında sesin MOS kalite değerini hesaplayabilmektir. Testlerde farklı paket kayıp oranı ve delay sürelerine sahip ses paketleri için POLQA modeli kullanılarak MOS değerleri ayrı ayrı ölçülmüştür. Simülasyon sonuçları, YSA ile tasarlanan model yapısının ITU-T P.863 standardı olan POLQA model ile elde edilen MOS değerlerini yüksek doğrulukta tahmin edebildiğini göstermiştir. Önerilen model, YSA 'nın öğrenme yeteneği sayesinde MOS kalite değerini, herhangi insan görüşüne ihtiyaç duymadan ağdaki değişimlere kolaylıkla adapte olarak belirleyebilmektedir.

Bu çalışma ile MOS değerinin, IP ağındaki QoS parametreleri dikkate alınarak hesaplanabileceği ortaya konulmuştur. Çalışmada farklı kural yapılarında YSA modelleri tasarlanmıştır. Bu modellerin eğitiminde ise LM, BR ve RPROP algoritmasından yararlanılmış ve algoritmaların performansları karşılaştırılmıştır. En iyi sonuçlar, LM için (3x4) kural yapısında (MSE: 2,7581 e-04), BR için (4x4) kural yapısında (MSE: 3,9588 e-04), RPROP için (4x5) kural yapısında (MSE: 3,6987 e-04) elde edilmiştir. Model yapılarına ait en iyi sonuçların birbirine yakın olduğu gözlenmiştir. Modeller, minimum (% hata), maximum (% hata), ortalama (% hata) ve MSE kriterleri gözönüne alınarak değerlendirildiğinde ise LM 'nin öne çıktığı gözlenmiştir. Tasarlanan YSA modellerinde, 4 ara katmana sahip olanların daha başarılı olduğu tespit edilmiştir. Modellerin başarısını arttırmak için farklı iterasyon sayılarının, kural yapılarının ve algoritmaların kullanılabileceği söylenebilir. Bu çalışma ile YSA temelli modellerin, ITU-T G.107 standardı olan E-model yapısına bir alternatif olarak kullanılabileceği gösterilmiştir. Gelecekte daha başarılı MOS tahmini yapabilmek için farklı ağ parametreleri de dahil edilebilir. Bu çalışmada sadece SIP temelli yaklaşım kullanılmış olup gelecekte farklı VoIP protokolleri kullanılarak elde edilen sonuçlar da karşılaştırılabilir.

Kaynakça

- 1- Çalık, O., Ircioğlu, U., Karabulut Kurt, G., Pusane, A.E., Demiroğlu, A.S., & Kayık, G. (2016). Impact of retransmissions on the quality of experience in VoIP systems. *Elektrik-Elektronik-Bilgisayar Mühendisliği Sempozyumu ve Fuarı (ELECO)*, Bursa, 617-621.
- 2- Pala, Z. (2017). Kampüs ağlarında arkaplan trafiğin IP-tabanlı telefon sistemlerin ses kalitesi üzerindeki etkisi. *Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 7(2), 55-63.

- 3- Al-Wahshat, H., Al-Maitah, M., & Al-Smadi, T. (2017). Voice quality for internet protocol based on neural network model. *Journal of Signal and Information Processing*, 8, 195-202. (<https://doi.org/10.4236/jsip.2017.84013>)
- 4- Agrisani, L., Capriglinoe, D., Ferrigno, L., & Miele, G. (2016). Measurement of the IP packet delay variation for a reliable estimation of the mean opinion score in VoIP services. *IEEE International Instrumentation and Measurement Technology Conference Proceedings*, Taipei, Taiwan, 1-6. (<https://doi.org/10.1109/I2MTC.2016.7520492>)
- 5- Hartpence, B. (2013). Packet Guide to Voice over IP. *O'Reilly press*.
- 6- Kadioğlu, R., Dalveren, Y., & Kara, A. (2015). Quality of service assessment: a case study on performance benchmarking of cellular network operators in Turkey. *Turk J Elec Eng & Comp Sci*, 23, 548-559. (<https://doi.org/10.3906/elk-1302-191>)
- 7- Nipp, O., Kuhn, M., Wittneben, A., & Schweinhuber, T. (2007). Speech quality evaluation and benchmarking in cellular mobile networks. *IEEE 2007 Mobile and Wireless Communications Summit*, Budapest, Hungary, 1-5. (<https://doi.org/10.1109/ISTMWC.2007.4299219>)
- 8- Mossavat, I. (2012). A hierarchical Bayesian approach to modeling heterogeneity in speech quality assessment. *IEEE T Audio Speech*, 20, 136-146. (<https://doi.org/10.1109/TASL.2011.2158421>)
- 9- ITU-T (1996). Recommendation P. 861, Objective quality measurement of telephone band (300 - 3400 Hz) speech codecs. (<https://www.itu.int/rec/T-REC-P.861-199608-S/en>), (Erişim Tarihi: Mayıs 2020).
- 10- Kuipers, F., Kooij, R., De Vleeschauwer, D., & Brunnström, K. (2010). Techniques for measuring quality of experience. *International Conference on Wired/Wireless Internet Communications (WWIC)*, Lulea, Sweden, 216-227.
- 11- Opticom GmbH. (2005). PESQ - Perceptual evaluation of speech quality. (<http://www.opticom.de/technology/pesq.php>), (Erişim Tarihi: Mayıs 2020).
- 12- Jelassi, S., Rubino, G., Melvin, H., Youssef, H., & Pujolle, G. (2012). Quality of experience of VoIP service: a survey of assessment approaches and open issues. *IEE Communications Surveys & Tutorials*, 14, 491-513. (<https://doi.org/10.1109/SURV.2011.120811.00063>)
- 13- Pocta, P., Cinar, Y., & Melvin, H. (2016). Black-box analysis of the extent of time-scale modification introduced by WebRTC adaptive jitter buffer and its impact on listening speech quality. *Journal Communications*, 18(1), 17-22.
- 14- Hines, A., Skoglund, J., Kokaram, A., & Harte, N. (2013). Robustness of speech quality metrics to background noise and network degradations: comparing VISQOL, PESQ and POLQA. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Canada, 3697-3701. (<https://doi.org/10.1109/ICASSP.2013.6638348>)
- 15- Gaoxiong, Y., & Wei, Z. (2012). The Perceptual objective listening quality assessment algorithm in telecommunication: introduction of ITU-T new metrics POLQA. *1st IEEE International Conference on Communications in China (ICCC)*, Beijing, China, 351-355. (<https://doi.org/10.1109/ICCCChina.2012.6356906>)
- 16- POMY, J. (2011). POLQA- The next-generation mobile voice quality testing standard. *ZNIIS / ITU Workshop*. (https://www.itu.int/ITU-D/tech/events/2011/Moscow_ZNIIS_April11/Presentations/09-Pomy-POLQA.pdf), (Erişim Tarihi: Mayıs 2020).
- 17- Gerlach, O. (2012). Next-generation (3G/4G) voice quality testing with POLQA. (https://scdn.rohde-schwarz.com/ur/pws/dl_downloads/dl_application/application_notes/1ma202/1MA202_1e_3G4G_voice_quality_testing_POLQA.pdf), (Erişim Tarihi: Mayıs 2020).
- 18- Mishra, K.C., & Das, P.C. (2015). Measuring quality of service of VoIP based on artificial neural network approach. *International Journal of Advanced Research in Computer Science and Software Engineering*, 5(3), 657-661.
- 19- Ren, J., Mao, D., & Wang, Z.W. (2009). A neural network based model for VoIP speech quality prediction. *Proceedings of the 2nd International Conference on Interaction Sciences: Information Technology, Culture and Human*, Seoul, Korea, 1244-1248. (<https://doi.org/10.1145/1655925.1656152>)
- 20- AL-Akhras, M., ALMomani, I., & Sleit, A. (2011). An improved E-model using artificial neural network VoIP quality predictor. *Neural Network World*, 21(1), 3-26. (<https://doi.org/10.14311/NNW.2011.21.001>)
- 21- Sun, L., & Ifeachor, E.C. (2002). Perceived speech quality prediction for voice over IP-based networks. *IEEE International Conference on Communications (ICC)*, New York, USA, 1-5. (<https://doi.org/10.1109/ICC.2002.997307>)
- 22- Larijani, H., & Radhakrishnan, K. (2010). Voice quality in VoIP networks based on random neural networks. *Ninth International Conference on Networks*, Menuires, France, 89-92. (<https://doi.org/10.1109/ICN.2010.23>)
- 23- Konar, M. (2019). GAO algoritma tabanlı YSA modeliyle İHA motorunun performansının ve uçuş süresinin maksimizasyonu. *European Journal of Science and Technology*, 15, 360-367. (<https://doi.org/10.31590/ejosat.529093>)
- 24- Haykin, S. (1999). Neural networks-a comprehensive foundation, *Prentice Hall*, 2nd ed.
- 25- Öztürk, C. (2011). Yapay sinir ağlarının yapay arı kolonisi algoritması ile eğitilmesi. *Erciyes Üniversitesi, Fen Bilimleri Enstitüsü*, (Doktora Tezi).
- 26- Kılıç, E., Özbalcı, U., & Özçalık, H.R. (2012). Lineer olmayan dinamik sistemlerin yapay sinir ağları ile modellenmesinde MLP ve RBF yapılarının karşılaştırılması. *Elektrik - Elektronik ve Bilgisayar Mühendisliği Sempozyumu (ELECO)*, Bursa, 694-698.
- 27- Bağış, A., & Konar, M. (2010). Uçuş kontrol sistemi yakıt parametresinin yapay sinir ağları kullanılarak belirlenmesi. *Akıllı Sistemlerde Yenilikler ve Uygulamaları Sempozyumu (ASYU'2010)*, Kayseri, Türkiye, 104-108.
- 28- Konar, M., & Bağış, A. (2016). Simultaneous computation of the speed and fuel parameters of flight control system by using Anfis and artificial neural networks. *24th Signal Processing and Communication Application Conference (SIU 2016)*, 1389-1392. (<https://doi.org/10.1109/SIU.2016.7496008>)
- 29- Oktay, T., Arik, S., Turkmen, I., Uzun, M., & Celik, H. (2018). Neural network based redesign of morphing UAV for simultaneous improvement of roll stability and maximum lift/drag ratio. *Aircraft Engineering and Aerospace Technology*, 90(8), 1203-1212. (<https://doi.org/10.1108/AEAT-06-2017-0157>).