

SES KODLAYICILARININ KODLAMA ÖZELLİKLERİNE GÖRE SINIFLANDIRILMASI

CLASSIFICATION OF AUDIO CODECS BASED ON CODING CHARACTERISTICS

Samet Hiçsönmez, Hüseyin T. Sencar

Bilgisayar Mühendisliği Bölümü
TOBB Ekonomi ve Teknoloji Üniversitesi
{shicsönmez,htsencar}@etu.edu.tr

Ismail Avcıbaş

Elektrik Elektronik Mühendisliği Bölümü
Turgut Özal Üniversitesi
iavcibas@turgutozal.edu.tr

ÖZETÇE

Bu bildiriye verilen bir sesin kodlanmasında kullanılan kodeği, kod çözümü yapmadan ayırt edebilen hızlı, basit ve yeni bir yöntem sunulmaktadır. Yöntem, kodlanmış seslere ilişkin kaotik ve rastgelelik özelliklerini kullanarak farklı kodekler için modeller oluşturmaktadır. Metodun en önemli özellikleri kodlanmış sesin rastgele bir yerinden alınan sadece birkaç kilobayt boyutundaki veri ile işlem yapması ve bir kodeğin kodlama yapısı ve özellikleri hakkında bilgi sahibi olmamasıdır. Bu iki özellik yöntemin basit ve hızlı olmasını sağlamaktadır. Yöntemin başarımını ölçmek amacıyla tek kodlanmış ve birden fazla kodlanmış ses örnekleri üzerinde deneyler yapılmıştır.

ABSTRACT

We propose a fast, simple and new method for identification of audio codecs that does not require decoding of coded audio data. The method uses chaotic and randomness features of coded audio to build models associated with different codecs. The most important features of the method are operating on just a few kilobytes of data sampled randomly from audio file and assuming no knowledge on encoding structure of a codec. These two features make it fast and simple. Experiments are performed on both singly coded and transcoded audio samples to measure accuracy of the technique.

1. GİRİŞ

Dijital ses, birçok farklı amaca yönelik olarak tasarlanmış kodekler ile kodlanmaktadır. Bu amaçlar temel olarak yüksek kaliteli sesin ve müziğin sıkıştırılması, Internet ortamında ses yayını yapılması ya da telefon, GSM ya da VoIP hatları üzerinden ses iletişiminin gerçekleştirilmesi olarak sıralanabilir. Şu anda 100'den fazla ses kodeğinin mevcut olduğunu göz önünde bulundurduğumuzda, bir ses kodlanırken kullanılan kodeğin herhangi bir metadataya bağlı kalınmadan belirlenmesi bazı mevcut problemlerin üstesinden gelmek için yeni çözümler sunabilir.

Bu problemlere örnek olarak, ağ trafiği izleme araçlarının ağ performansını takip etmek için ses içeren ağ akışlarını hızlı bir şekilde sınıflandırmak zorunda olmaları [1], telefon sisteminde akan seslerin kodlama geçmişlerinin belirlenmesi [2], ve ses dosyalarının kalitesinin ve doğruluğunun veya

hatalı kalitede olanların (düşük kalitede olan bir ses dosyasının tekrar yüksek kalitede kodlanması) ayırt edilmesi [3] verilebilir.

Bu çalışmada, belirtilen uygulama alanlarında kullanılacak yeni bir kodek sınıflandırma yöntemi sunulmaktadır. Kodek sınıflandırma ile ifade etmeye çalıştığımız, belirli sayıda kodekten biri ile kodlanmış bir ses verisinin hangi kodek ile kodlandığının bulunmasıdır. Çok sayıda kodek bulunmasına rağmen, temelde kodekler ses kalitesi, veri büyüklüğü (data rate) ve kodlayıcı/kod çözücünün karmaşıklığı gibi tasarım kararlarında farklılık gösterirler. Bizim çalışmamızın temelini de bu tasarım kararlarının ses verisi üzerinde bıraktığı etkilerin karakterize edilmesi oluşturmaktadır. Yöntem bu etkileri çıkartırken, verilen sesteki küçük boyutlu bir bayt dizisi (birkaç KB) olarak, bu bayt dizisinden elde ettiği rastgelelik ve kaotik nitelikleri çıkartarak bu işlemi gerçekleştirmektedir.

Kodlanmış veriden bilgi edinmek için istatistiksel analiz yöntemleri kullanan az sayıda çalışma vardır. Bunlara örnek olarak bir ses steganaliz uygulaması verilebilir [4]. Bizim çalışmamıza temel oluşturan çalışmada [1] ise, ağ akışının istatistiksel özelliklerine göre karakterize edilmesi amaçlanmıştır. Fakat bu çalışmanın amacı bir ses verisinde kullanılan kodeğin belirlenmesinden ziyade, ağdaki paketlerin metin, resim, ses, görüntü verisi ya da şifrelenmiş veri olup olmadığının belirlenmesidir.

Bizim çalışmamız belirtilen çalışmalardan esinlenmekle beraber, bunlar ile şu iki açıdan farklılık göstermektedir. Öncelikle tamamen farklı tip veri türleri birbirlerinden çok farklı istatistiksel özellikler gösterirken, sadece ses kodekları ile kodlanmış ses verileri birbirlerine benzer özellikler göstermektedir. İkincisi biz herhangi bir kodek ile ilgili dosya yapısı hakkında bilgi sahibi olunmadığını varsayıyoruz. Testlerimizde, yüksek kaliteli müzik sıkıştırma kodekları ile PSTN, GSM ve VoIP sistemlerinde kullanılan en popüler konuşma kodeklerini kullandık.

Bildirinin bundan sonraki bölümünde yöntemin işleyişi ve nitelikler açıklanmıştır. Bölüm 3'te ise kullanılan kodeklerin temel özellikleri ve birbirlerinden farklıları hakkında bilgi verilmiştir. En son bölümde ise yapılan testler ve sonuçları sunulmuştur.

2. METODOLOJİ VE NİTELİKLER

Bu bildiriye sunulan yöntemin amacı bir ses kodlanırken kullanılan kodeğin belirlenmesidir. Bu yapılırken herhangi

bir metadata ya da kodlama yapısı göz önüne alınmamaktadır. Teknik temelde, kalite, sıkıştırma oranı ve karmaşıklık gibi tasarım seçeneklerini içeren kodlama sürecini bir sınıflandırıcı vasıtasıyla karakterize etmektedir.

Sistem çevrimdışı ve çevrimiçi olmak üzere iki kısımda incelenebilir. Çevrimdışı kısım kodlama ve örnekleme, nitelik çıkarma ve makine öğrenmesi şeklinde üç aşamaya bölünebilir. İlk aşamada kodlanmamış ses verileri 3. Bölüm'de belirtilen kodekler ile kodlanmış ve bu kodlanmış verilerden farklı boyutlarda bayt dizileri alınmıştır. Bayt dizileri, kodlanmış ses dosyalarının rastgele bir yerinden başlanarak ardışık olacak şekilde alınmıştır. İkinci aşamada ise alınan bu bayt dizileri rastgelelik ve kaotik niteliklerinin hesaplanmasında kullanılmıştır. Son aşamada ise hesaplanmış nitelik vektörü ile makine öğrenmesi algoritmaları eğitilmiştir. Çevrimiçi kısımda ise sistem verilen kodlanmış farklı ses dosyaları ile test edilmiş ve ortalama başarımları ölçülmüştür.

Görüldüğü üzere sistemin en kritik adımı niteliklerin belirlenme aşamasıdır. Kullandığımız nitelikleri temel olarak iki bölümde inceleyebiliriz.

2.1. Rastgelelik Nitelikleri

Rastgelelik nitelikleri çıkartılırken ses örneklerinden alınan bayt dizileri üzerinde istatistiksel hesaplamalar yapılmıştır. Bizi bu istatistiksel testleri yapmaya iten en önemli etken, bunların NIST tarafından yapılan kriptografik uygulamaların rastgelelik özelliklerini değerlendiren testlerden olmalarıdır [5]. Rastgelelik nitelikleri zaman ve frekans alanı nitelikleri olmak üzere ikiye ayrılabilir. Zaman alanı nitelikleri alınan bayt dizileri üzerinden hesaplanmıştır. Ancak frekans alanı niteliklerini çıkartmak amacıyla alınan bayt dizilerine FFT (Fast Fourier Transform) uygulanmıştır.

Zaman alanında çıkartılan nitelikler de basit ve yüksek dereceli istatistikler olmak üzere iki farklı gruba ayrılabilir. Basit istatistikler ortalama değer, varyans, öz-iliinti ve entropidir. Bu istatistikler basit gibi gözükmesine rağmen kodeklerin tanımlanmasında etkileri büyüktür. Örneğin WMA kodeği ile kodlanmış bir ses dosyasının bayt dizisine baktığımızda uzun sıfır serileri görmekteyiz. Bu sebepten dolayı WMA ile kodlanmış seslerin ortalama değerleri düşük çıkmaktadır. Buna karşın G711 kodeklerinin ortalama değeri diğerlerine kıyasla çok yüksek çıkmaktadır. Öz-iliinti verideki tekrarları ortaya çıkarmaya yarayan bir istatistiktir. Farklı kodlama algoritması kullanan kodekler ses verisi üzerinde farklı etkiler bırakır ve öz-iliinti bu farklılıkları ortaya çıkarabilir. Bu amaçla ilk 21 katsayı nitelik vektörüne eklenmiştir. Entropi ise verideki rastgeleliği ortaya çıkarır. Bu istatistik, yüksek entropiye sahip olan kodlanmamış veri (WAV) ile kodlanmış verinin ayrılmasında kullanılır. Bu basit istatistikler grubundaki bir başka istatistik varyanstır.

Zaman alanındaki yüksek dereceli istatistikler ise yamukluk, savrukuk ve bicoherencetir. Yamukluk ve savrukuk verinin dağılımı ile ilgili istatistiklerdir. Verideki değerlerin çoğunluğu ortalama değerden fazla ise yamukluk negatif değerler, tam tersi durumda ise pozitif değerler almaktadır.

Frekans alanındaki istatistikler, alınan bayt dizilerinin FFT dönüşümlerinin alınıp, ortaya çıkan güç spektrumunun 4 eşit banda bölünüp her bant için ayrı ayrı hesaplanması ile çıkarılmıştır. Bu alandaki nitelikler ortalama değer, varyans

ve yamukluktur. Toplam 12 tane frekans alanı niteliği vektöre eklenmiştir.

2.2. Kaotik Nitelikler

Konuşma sinyallerinde doğrusal modeller ile belirlenememiş kaotik olgular olduğuna dair teorik ve deneysel kanıtlar mevcuttur [6]. Konuşma sinyallerinin kaotik bir sistem tarafından üretildiği göz önüne alındığında, farklı kodlama teknikleri bu konuşma sinyallerinin kaotik yapısını değiştirecektir ve Lyapunov Exponents (LE) ve False Neighbour Fraction (FNF) gibi kaotik nitelikleri farklı olacaktır. Bu sebeple 11 tane LE ve 15 tane FNF olmak üzere toplam 26 kaotik nitelik, vektöre eklenmiştir.

LE ve FNF'nin temeli, faz uzayında konuşma sinyal vektörlerinin komşuluğuna dayanmaktadır. Sıkıştırma işlemi uzay alanında komşuluk mesafelerini değiştirmektedir. Sıkıştırma algoritmaları sinyaldeki tekrarları açığa çıkarırlar ve performansları çıktıda bıraktıkları ilinti miktarı ile ölçülür. Uzay alanından elde edilen kaotik tip nitelikler ise bu ilintileri ölçmektedir.

3. SES KODEKLERİ

Ses kodlama için kullanılan birçok kodek mevcuttur ve bu kodekler temelde yüksek kaliteli sesin sıkıştırılması, internette canlı yayın ve PSTN, GSM ya da VoIP hatları üzerinden sesli konuşma amacıyla kullanılmaktadırlar. Fazla sayıda kodek olmasına rağmen, kodekler birbirlerinden ses kalitesi ve sıkıştırma oranı arasında yaptıkları dizayn kararı ile ayrılırlar [7]. Bu tasarım kararına kodlayıcı ve kod çözücünün karmaşıklığı da dahil edilebilir.

Kodekleri amaçlarına göre 2 gruba ayırabiliriz, yüksek kaliteli ses sıkıştırma amacıyla kullanılan kodekler ve konuşma kodlama amacıyla kullanılan kodekler. Yüksek kaliteli ses sıkıştırma için kullanılan kodekler müzik dosyalarının kodlanmasında kullanılırlar ve kullanım amacına uygun şekilde yüksek kaliteli çıktı verecek şekilde kodlama yaparlar. Bu kodeklerin sıkıştırma oranı konuşma (speech) kodeklerine göre daha düşüktür. Buna karşın konuşma kodekleri farklı iletişim kanallarında, haberleşme amacıyla kullanılmaktadırlar. İletişim ortamlarının temel sorunu olan bant genişliği faktörü bu kodeklerin, yüksek sıkıştırma yapan bir kodlama algoritması kullanmasını gerektirmektedir. Bu yüksek sıkıştırma sağlanırken konuşmanın belirli bir kalitenin (toll quality) altına düşmemesi amaçlanmıştır [8].

Biz bu sistemde testleri yaparken PSTN, GSM ve VoIP hatlarında kullanılan konuşma kodekleri ile yüksek kaliteli ses sıkıştırmasında kullanılan toplam 16 adet kodek kullandık. Günümüzde PSTN'lerde en yaygın kullanılan kodek G711'dir ve bu kodeğin a-law ve u-law olmak üzere 2 farklı türü vardır [9]. Bu gruba kattığımız diğer kodek ise PCM'dir. Bu 3 kodek testlerimizde PSTN kodekleri olarak kullanılmıştır. GSM hatlarında kullanılan en yaygın kodekler olan AMR, AWB, GSM 6.10 ve GSM 6.10(WAV) [9] testlerimizde GSM grubu kodekler olarak kullanılmıştır. VoIP uygulamalarında birçok farklı kodek kullanılmaktadır. Biz VoIP grubu kodekler olarak G729, G726, iLBC ve Speex [9] kodeklerini kullandık. Yüksek kaliteli ses kodekleri grubuna [10] ise günümüzün en popüler müzik kodeği olan MP3 başta olmak üzere AAC, OGG, WMA ve FLAC [11] dahil edilmiştir.

4. TESTLER ve SONUÇLAR

Testlerimizde 2 farklı veri seti kullandık. Birincisi pop, rock, folk, klasik ve enstrümantal gibi farklı müzik türlerinden oluşan 50 farklı müzik CD'sinden toplanmış 1000 adet müzik örneğinden oluşmaktadır. Bu örnekler yaklaşık 5 sn. uzunluğunda ve 850 KB dosya boyutundadır. Bu veri setini oluştururken her şarkıdan 2 örnek alınmıştır ve bu 2 örnek sınıflandırıcı oluşturulurken her zaman birlikte test ya da eğitime (train) grubunda yer almıştır. Aynı şarkıdan alınan 2 örnek benzer karakteristik gösterebilir dolayısıyla sınıflandırıcı oluşturulurken bu 2 örneğin farklı gruba gitmesi hatalı yüksek sınıflandırma başarılarına yol açabilir.

Testlerde kullanılan diğer veri kümesi ise VoxForge [12] konuşma veri tabanından alınan uzunlukları 1-13 sn. arasında, boyutları ise 41-438 KB arasında değişen 2000 tane konuşma örneğidir. Bu ses örnekleri 256 kbps bit hızına sahiptir.

Tüm testlerde veri kümesinin yarısı eğitim işlemi için, kalan yarısı ise test işlemi için kullanıldı. Bildirinin bundan sonraki kısmında ilk veri kümesi müzik, ikinci veri kümesi ise konuşma kümesi diye adlandırılacaktır.

Testler yapılırken Bölüm 2'de bahsettiğimiz adımlar sırası ile uygulanmıştır. İlk adımda veri seti örnekleri, 3.Bölüm'de belirttiğimiz kodekler ile kodlandı. Kodlama tamamlandıktan sonra kodlanmış olan ses dosyalarının rastgele bir yerinden başlanarak 2 KB ya da 4 KB boyutlarında bayt dizileri alındı. Daha sonra alınan bu bayt dizileri üzerinden kaotik ve rastgelelik nitelikleri çıkartıldı. Toplam 65 adet nitelikten 39 tanesi rastgelelik niteliklerine, kalan 26 tanenin 15 tanesi FNF, 11 tanesi ise LE grubuna dâhildir. Daha sonra ise sınıflandırıcı olarak, standart bir makine öğrenme tekniği olan Support Vector Machine (svm), radial basis function kernel ile kullanıldı. Svm için kullanılan paket ise Libsvm'dir [13]. Testlerin başarımları doğru sınıflandırılan örnek sayısının toplam örnek sayısına bölünmesi ile hesaplanmıştır.

Yapılan testleri 2 grupta inceleyebiliriz. İlk grup testlerde kodlanmamış örnekler belirtilen kodekler ile bir defa kodlandı ve tüm kodlanmış örnekler birbirlerinden ayrılmaya çalışıldı. İkinci grup testlerde ise kodlanmamış sesler iki farklı kodek ile kodlanarak kodlama işleminde kullanılan ilk kodeğin tespit edilmesi ve iki kere kodlanmış sesler ile tek kodlanmış seslerin ayırt edilmesi amaçlandı.

4.1. Tekli Kodlama Testleri

Bu testlerin amacı farklı kodekleri birbirlerinden ayırt etmek olduğu için kodlanmamış ses örnekleri 3.Bölüm'de belirtilen kodekler ile kodlandı. Bu 16 kodek ile kodlanmış ses örneklerine ek olarak kodlanmamış ses örnekleri de sınıflandırma işlemine dahil edildi ve sonuç olarak 17-sınıf sınıflandırma işlemi yapıldı.

Müzik veri seti ile kodlanan ses örneklerinden 1 KB ve 4 KB olmak üzere farklı uzunluklarda bayt dizileri alındı. Bu iki farklı boyut için iki farklı sınıflandırma işlemi yapıldı. 1 KB boyutundaki bayt dizisi için sınıflandırma başarımları %85.11 iken 4 KB için başarımlar %95.88 olarak ölçüldü ve Tablo 1'in sağ tarafındaki değerler bu sınıflandırmaya ait doğruluk matrisini oluşturmaktadır.

Konuşma veri seti ile yapılan testlerde ise örneklerden 2 KB uzunluğunda bayt dizisi alındı. Bu teste ait başarımlar

%97.34 olarak hesaplandı ve Tablo 1'in sol tarafındaki değerler bu teste ait doğruluk matrisini göstermektedir. Bu testin sonuçları incelendiğinde benzer kodlama tekniği kullanan kodeklerin birbirleri ile karıştırıldığı görülmektedir. Kodlama teknikleri (ACELP) aynı olan AMR, AWB ve G729 kodeklerinin sadece birbirleri ile karıştırıldığı görülmektedir. Benzer şekilde a-law ile u-law da neredeyse aynı kodlama tekniğini kullanırlar ve bu iki kodek te birbirleri ile karışmıştır. Bir diğer örnek ise GSM ile GSM(WAV) kodeklerinin sadece birbirleri ile karıştırılmasıdır. Bu iki kodek tamamen aynı kodlama tekniği kullanırlar fakat bit hızları farklıdır. Müzik veri seti testlerinde ise bu örneklere ek olarak MP3, OGG ve WMA kodeklerinin karıştırılmasının sebebi olarak üçünün de MDCT tabanlı kodekler olması verilebilir.

Bu grupta düşünülen bir diğer senaryo ise bir kodeğin sadece kendi grup kodekleri ile test edilmesidir. Bu amaca yönelik olarak 4 farklı sınıflandırıcı kurulmuştur. Müzik veri seti ile kodlanmış örneklerden alınan 1 KB'lık bayt dizisi için sınıflandırma başarımları PSTN, GSM, VoIP ve yüksek kaliteli ses kodekleri için sırası ile %98.53, %96.20, %80.05, %83.33 olarak ölçüldü. Aynı veri seti ile kodlanmış örneklerden alınan 4 KB bayt dizisi için sonuçlar ise %98.80, %99.05, %99.95 ve %95.33 olarak hesaplanmıştır. Konuşma veri seti ile kodlanmış örneklerden alınan 2 KB'lık bayt dizisi için ise başarımlar %97.80, %95.65, %99.02 ve %100'dür. Sonuçlar incelendiğinde konuşma veri setine ait testlerde görece olarak düşük başarımlar elde edilen grup GSM grubudur. Bunun sebebi ise bu gruptaki 4 kodeğin 2'şerli olarak (AMR ile AWB, GSM ile GSM(WAV)) tamamen aynı kodlama tekniğini kullanmalarıdır.

Bu gruptaki tüm test sonuçlarından da görüldüğü üzere önerilen yöntem benzer hatta aynı tekniği kullanan kodekleri daha yüksek başarımla ayırt edebilmektedir.

4.2. Çoklu Kodlama Testleri

Günümüzde farklı iletişim ortamları birbirleri ile haberleşmektedir ve bu haberleşme gerçekleşirken iki ortamın sınırında ses diğer ortama ait uygun bir kodek ile tekrar kodlanmaktadır. Örneğin bir VoIP uygulamasından GSM hattına ait bir telefon arandığında ses öncelikle uygun VoIP kodeği ile kodlanır fakat daha sonra bu ses GSM hattına geçiş yaparken uygun bir GSM kodeği ile tekrar kodlanır. Bu tip durumlarda sesin kaynağının tespit edilebilmesi için ilk kodeğin bilinmesi gerekmektedir. Bunun yanında aynı örnekten yola çıkarak GSM hattına gelen sesin birden fazla kodlanmış olup olmadığı kaynak tespiti için gereklidir.

Çoklu kodlamanın olup olmadığı ve varsa ilk kodlama yapılan kodeğin tespiti için 3 farklı senaryo üzerinde testler yapıldı. Bu senaryolar GSM'den PSTN hattına geçiş, GSM hattından VoIP hattına geçiş ve VoIP ağından PSTN'e geçiştir. Bu testler yapılırken sınıflar şu şekilde oluşturuldu. Örneğin GSM ağından PSTN hattına geçiş senaryosu göz önüne alınırsa, kodlanmamış ses örnekleri önce 4 GSM kodeği ile kodlandı sonra bu seslerin kod çözümü yapıldı daha sonra a-law kodek ile kodlandı ve bu 4 sınıfın yanına tekli a-law kodlanmış ses örnekleri de eklendi. Aynı işlem u-law ve PCM kodekleri için de tekrarlandı. Sonuç olarak bu senaryo için PSTN grubundaki kodek sayısı kadar yani 3 adet sınıflandırma işlemi yapıldı. Diğer senaryolar için de aynı prosedür takip edildi. Her iki veri seti ile yapılan testlerde 4

KB uzunluğunda bayt dizisi üzerinden nitelik çıkartma işlemi yapıldı.

İlk senaryo olan GSM'den PSTN hattına geçişte 3 farklı sınıflandırma yapıldı. GSM – a-law testine ait sınıflandırma sonuçları %92.28 müzik veri seti için ve %89.78 konuşma veri seti için olarak hesaplanmıştır. GSM – u-law için ise başarımları aynı veri seti için %85.20 ve %90.66 olarak ölçüldü. GSM – PCM testine ait sonuçlar ise %84.48 ve %90.20'dir.

GSM ağından VoIP hattına geçiş senaryosu için ise 4 farklı test yapıldı. GSM – G729 testine ait başarımları müzik veri seti için %62, konuşma veri seti için ise %45.46 çıkmıştır. GSM – G726 testine ait başarımları ise sırası ile %87.44 ve %81.30 olarak hesaplanmıştır. GSM – iLBC testinin sonuçları %95 ve %87.54'tür. GSM – Speex testine ait sonuçlar %91.20 ve %79.22 olarak ölçülmüştür. Buradaki en dikkat çekici test sonucu GSM – G729 testidir. Bu testin sonucu beklenildiği gibi son derece düşük çıkmıştır. Birden fazla kodlama yapıldığı durumlarda kodlama yapılan son kodek önceki kodekten daha düşük bit hızında kodlama yapıyorsa önceki kodeğin ses üzerinde bıraktığı izler takip edilememektedir. G729 kodeği bizim kullandığımız GSM kodeklerinden daha düşük bit hızında kodlama yaptığı için bu testin sonucu bu şekilde düşük çıkmıştır.

Son senaryo olan VoIP hattından PSTN'e geçişte ise 3 farklı test yapılmıştır. İlk test olan VoIP – a-law'ın sonuçları sırası ile müzik ve konuşma veri setleri için %97.15 ve %88.52'dir. VoIP – u-law testinin başarımları ise %99,2 ve %88.67 olarak hesaplanmıştır. Son test olan VoIP – PCM testine ait başarımları ise %96,8 ve %98.52'dir.

Tüm bu testlerin sonuçları incelendiğinde tekli kodlanmış sesler ile çoklu kodlanmış sesler yüksek başarımla ayırt edilebilmekte ve çoklu kodlanmış seslerin kodlandığı ilk kodek tespit edilebilmektedir.

5. KAYNAKÇA

- [1] Shanmugasundaram, K., Kharrazi, M., Memon, N., "Nabs: a system for detecting resource abuses via characterization of flow content type", *Computer Security Applications Conference, 2004. 20th Annual*, vol., no., pp. 316- 325, 6-10 Dec. 2004.

- [2] V. Balasubramanian, A. Poonawalla, M. Ahamad, M. Hunter and P. Traynor, "PinDr0p: Using Single-Ended Audio Features to Determine Call Provenance", 2010.
- [3] R. Yang, Y. Q. Shi and J. Huang "Defeating fake-quality MP3", *Proc. ACM Multimedia and Security'09*, 2009.
- [4] Rainer B., AndreasW., "Statistical characterisation of MP3 encoders for steganalysis", *Proc. of the 2004 workshop on Multimedia and security*, pp. 25-34, 2004.
- [5] A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications, NIST Special publication 802-22,2001 <http://csrc.nist.gov/groups/ST/toolkit/rng/documents/SP800-22b.pdf>
- [6] O. H. Kocal, E. Yuruklu, I. Avcibas, "Speech steganalysis using chaotic-type features", *IEEE Transactions on Information Forensics and Security* (2008) 651-661.
- [7] M. Bosi, R. E. Goldberg, "Introduction to Digital Audio Coding and Standards", Kluwer Academic Publishers, Boston, 2003.
- [8] Gibson, J.D.; , "Speech coding methods, standards, and applications," *Circuits and Systems Magazine, IEEE* , vol.5, no.4, pp. 30- 49, Fourth Quarter 2005
- [9] S. Karapantazis and F.N. Pavlidou, "VoIP: a comprehensive survey on a promising technology", *Computer Networks* 53 (12) (2009), pp. 2050–2090.
- [10] D. Avelar, B. Morrissette, D. Forbes, G. Albert, "Audio Codecs: Evaluation and Comparison of Popular Format", 2008.
- [11] FLAC. FLAC: Free Lossless Audio Codec. http://flac.sourceforge.net/documentation_format_overview.html, 2008.
- [12] <http://voxforge.org/>
- [13] C.-C. Chang and C.-J. Lin, LIBSVM: A library for support vector machines, <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

Tablo1: İki Farklı Veri Kümesi (Konuşma/Müzik) için 17 Sınıf Kodek Sınıflandırma İşlemine Ait Doğruluk Matrisleri

Gerçek	Tahmin Edilen																
	PSTN			VoIP				GSM				Yüksek Kaliteli Ses					
	A-law	u-law	PCM	AMR	AWB	GSM	GSM (WAV)	G.729	G.726	iLBC	Speex	AAC	MP3	OGG	FLAC	WAV	WMA
PSTN	A-law	100/93,6	0/4,6	0/1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0,6	0/0
	u-law	6,7/3	93/96	0,3/1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0
	PCM	0/0	0/0,4	96/99,6	0/0	0/0	0/0	0/0	0/0	3,7/0	0/0	0,1/0	0,2/0	0/0	0/0	0/0	0/0
VoIP	AMR	0/0	0/0	0/0	96,7/99,4	2/0	0/0	0/0	1,3/0	0/0	0/0	0/0	0/0	0/0	0/0,6	0/0	0/0
	AWB	0/0	0/0	0/0	1,1,6	99/96,4	0/0	0/0	0/0,2	0/0	0/0	0/0,4	0/0	0/0	0/1,4	0/0	0/0
	GSM	0/0	0/0	0/0	0/0	0/0	97,5/92	2,5/8	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0
	GSM (WAV)	0/0	0/0	0/0	0/0,2	0/0	14,2/1	85,8/96,6	0/0,8	0/0	0/0	0/0,4	0/0	0/0	0/0,1	0/0	0/0
GSM	G.729	0/0	0/0	0/0	2,3/0	0,4/0	0/0	0/0,8	96,2/96,8	0/0	0/0	0/0	1,1/0	0/0	0/0,2,4	0/0	0/0
	G.726	0/0,2	0/0	0/0,2	0/0	0/0	0/0	0/0	96/99,6	4/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0
	iLBC	0/0	0/0	0/0	0/0	0/0	0/0,2	0/0	0/0	99,9/98,6	0/0	0/1	0,1/0	0/0,2	0/0	0/0	0/0
	Speex	0/0	0/0	0/0	0/0	0/0	0/0,2	0/0,2	0/0	0/0	100/99	0/0,2	0/0	0/0,2	0/0,2	0/0	0/0
Yüksek Kaliteli Ses	AAC	0/0	0/0	0/0	0,3/0	0/0	0/0	0/0	0/0	0/0	0/0,2	99,7/97,4	0/1,4	0/0,4	0/0,6	0/0	0/0
	MP3	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0,2	0/2	100/88,4	0/6,4	0/3	0/0	0/0
	OGG	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0,8	0/0,2	100/91	0/0	0/0	0/0,2	
	FLAC	0/0	0/0	0/0	0/0	0/0	0/0,2	0/0,6	0/0	0/0	0/0	0/0	0/2	100/97	0/0	0/0,2	
	WAV	0,5/0	2,1/0	0,1,1,6	0/0	0/0	0/0	0/0	0/0	2,3/0,6	0/0	0/0	0/0	0/0	0/0	95/97,8	0/0
WMA	0/0	0/0	0/0,8	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/6,8	0/1,6	0/0	0/0	100/90,8