Türk Kütüphaneciliği 30, 4 (2016), 678-701

Üstverinin Tam-Metin Bilgi Erişim Performansı Üzerindeki Etkisi: Küçük Ölçekli Türkçe Külliyat Üzerinde Deneysel Bir Araştırma*

Impact of Metadata on Full-text Information Retrieval Performance: An Experimental Research on a Small Scale Turkish Corpus

Çağdaş Çapkın**

Öz

Bilgi kurumları üstveri, tam-metin veya hem üstveri hem de tam-metin (melez) içerikleri depolamak, dizinlemek ve eriştirmek için metin tabanlı bilgi erişim sistemleri kullanmaktadır. Araştırmanın amacı, bu içeriklerin bilgi erişim performansı üzerindeki etkisini değerlendirmektir. Bu amaçla, küçük ölçekli bir Türkçe külliyat için varsayılan Lucene bilgi erişim modelini kullanan üstveri (ÜBES), tam-metin (TBES) ve melez (MBES) içerik bilgi erişim sistemleri geliştirilmiştir. Bu üç sistemin performansını değerlendirmek için "duyarlılık - anma" ve "normalize sıralama" testleri yapılmıştır. Deneysel bulgular, ÜBES ve TBES arasında ortalama duyarlılık performansında anlamlı bir fark olmadığını göstermiştir. Diğer taraftan, MBES'in ortalama duyarlılık performansı ÜBES ve TBES'ten anlamlı olarak yüksektir. Bilgi erişim performansı kullanıcı-merkezli olarak değerlendirildiğinde, ÜBES ve MBES'in normalize sıralama performansları TBES'e göre anlamlı olarak yüksektir. Ayrıca, üç bilgi erişim sisteminin eriştiği ilgili doküman ortalamaları arasında anlamlı bir farka ulaşılamamıştır. Bilgi erişim sistemlerinde üstveri ve tam-metin gibi faklı türlerdeki içeriklerin işlenmesinde terim yönetimi bakımından bazı avantajlar ve dezavantajlar bulunmaktadır. Melez içerik işleme (MBES), avantajları bir araya getirmiş ve bilgi erişim performansını artırmıştır.

Anahtar Sözcükler: *Bilgi erişim; dizinleme; otomatik dizinleme; üstveri; performans değerlendirme; Türk Kütüphaneciliği.*

Abstract

Information institutions use text-based information retrieval systems to store, index and retrieve metadata, full-text, or both metadata and full-text (hybrid) contents. The aim of this research was to evaluate impact of these contents on information retrieval performance. For this purpose, metadata (MIR), full-text (FIR) and hybrid (HIR) content information retrieval systems were developed with default Lucene information retrieval model for a small scale Turkish corpus. In order to evaluate performance of this three systems, "precision - recall" and "normalized recall" tests were conducted. Experimental findings showed that there were no significant differences between MIR and FIR in mean average precision (MAP) performance. On the other hand, MAP performance of HIR was significantly higher in comparison to MIR and FIR. When information retrieval performance was evaluated as user-centered, the "normalized recall" performances of

* Bu araştırma, yazarın Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü'ne sunduğu yayımlanmamış yüksek lisans tezine dayanmaktadır.
This study is based on the author's unpublished master's thesis that is introduced to Hacettepe University Graduate School of Social Sciences.

** Arş. Gör., İstanbul Üniversitesi Bilgi ve Belge Yönetimi Bölümü. e-posta: cagdascapkin@gmail.com
Research Ass. İstanbul University Department of Information and Records Management

Geliş Tarihi - ***Received*:** 24.05.2016
Kabul Tarihi - ***Accepted*:** 09.12.2016

MIR and HIR were significantly higher than FIR. Additionally, there were no significant differences between the systems in retrieved relevant document means. Processing different types of contents such as metadata and full-text had some advantages and disadvantages for information retrieval systems in terms of term management. The advantages brought together in hybrid content processing (HIR) and information retrieval performance improved.

Keywords: *Information retrieval; indexing; automatic indexing; metadata; performance evaluation; Turkish Librarianship.*

Giriş

Bilginin sağlanması, düzenlenmesi ve eriştirilmesi bilgi kurumlarının temel sorumluluklarındandır. Bu sorumluluklardan bilginin düzenlenmesi ve eriştirilmesiyle ilgili dönemin koşullarına göre ihtiyaçların karşılandığı çeşitli dizinleme sistemleri 19. yüzyıla kadar büyük ölçüde kütüphane bilimi tarafından merkeze insan emeği alınarak geliştirilmiştir. Öte yandan, II. Dünya Savaşını izleyen soğuk savaş döneminde bilimsel araştırmalar hızlanmış ve bilimsel yayın üretim artışı "patlama" düzeyine ulaşmıştır (Garfield, 1979, s. 6). Yayın sayısındaki yüksek artış mevcut dizinleme sistemleriyle bilginin organizasyonunu ve erişimini zorlaştırmıştır. Büyük ölçekli bilginin organizasyonu ve erişimi için de bilgi erişim sistemlerine ihtiyaç doğmuştur (Sanderson ve Croft, 2012, s. 1444). Bu dönemde, kütüphane bilimiyle birlikte matematik, istatistik ve bilgisayar bilim gibi çeşitli disiplinler bilginin organizasyonu ve erişimi sorununa insan emeğiyle birlikte makinalar aracılığıyla da çözüm üretmek üzere önemli katkılarda bulunmuştur.

Bilgi erişim sistemlerinin gelişim sürecindeki en önemli aşama kuşkusuz bilgisayarın ortaya çıkmasıdır. Bu konuda, kendisi de bilgisayar üzerine çalışmalar yürütmüş Vannevar Bush'un öngörüleri dikkate değerdir. Bush bir makalesinde (1945) matematik ve mantık problemlerini çözebilecek "Düşünen Makina"; kitap ve doküman gibi tüm iletişim araçlarının sıkıştırılıp depolanabileceği, hız ve esneklikle geri çağırılabileceği ve dokümandan dokümana zıplanabileceği "MEMEX" adını verdiği kişisel bir makine öngörüsünde bulunmuştur. Bush'un öngörülerinin günümüzde kullandığımız bilgisayar, hiper-metin ağ ve bilgi erişim sistemlerinin geliştirilmesinde hedef veya ilham kaynağı olduğu söylenebilir.

İkinci dünya savaşından sonraki dönemde bir başka önemli dönemeç de sınıflama, denetimli ve doğal dil dizinleme gibi mevcut sistemlerin Cranfield deneyleri ile zaman, maliyet ve erişim performansları bakımından sorgulanmaya başlanmasıdır (Cleverdon, 1960; Cleverdon ve Keen, 1966; Cleverdon, 1970; Cleverdon, 1977). Cranfield deneylerinin beklenmedik bulgusu ise, inanılanın aksine, bilgi erişim performansında, tek terim doğal dil dizinlemenin eğitimli kütüphanecilerin sofistike denetimli dillerden elle terim atayarak oluşturduğu konu dizinlemelerinden iyi sonuç vermesidir (Cleverdon ve Keen, 1966, ss. 252-253).

İnsan ve makine merkezli yaklaşımların ortak hedefi bilgiyi organize etmek ve eriştirmek olsa da her iki yaklaşımın kullanım alanları ve kullandığı teknikler farklı olabilmektedir. Bununla birlikte, her iki yaklaşımın da birbirine karşı sağladığı çeşitli üstünlükler ileri sürülebilmektedir. Makinaya dayalı organizasyonun maliyet, kapsam ve zaman bakımından insana dayalı bilgi organizasyonundan üstün olduğu üzerinde anlaşılan bir konudur (Anderson ve Pérez-Carballo, 2001; Shields, 2005). Öte yandan, insana dayalı organizasyonun diğer bakımlardan makinaya üstünlük sağlaması veya tam tersi karşılaştırmalar tartışmaya açıktır. Bu tip karşılaştırmalarda yer alan ilgili pek çok değişken izole ve kontrollü olmayan ortamlarda değerlendirilmekte ve ulaşılan sonuçlar inançlara dayandırılmaktadır (Anderson ve Pérez-Carballo, 2001).

İnsana ve makinaya dayalı organizasyonda üstünlük tartışmalarının paralelinde bir tartışma da üstveri ve tam-metin bilgi erişim sistemleri üzerinde yapılmaktadır. MEDLINE külliyatında üstveri ve tam-metin içerik üzerinden bilgi erişim performansı değerlendiren bir araştırmada (McKinin, Sievert, Johnson ve Mitchell, 1991), üstveri dizinlemede duyarlılığın

yüksek, tam-metin dizinlemede ise anmanın yüksek, duyarlılığın düşük olduğu sonucuna ulaşılmıştır. Bu çalışmaya benzer bir çalışma sadece gen isimlerinin sorgulanmasıyla yapılmış ve aynı sonuca ulaşılmıştır (Hemminger, Saelim, Sullivan ve Vision, 2007). Aynı sonuçlara ulaşılan bu iki araştırmada; doğal dil işleme süreçleri, ilgililik algoritması veya bilgi erişim modeli gibi deney düzeneği için son derece önemli olan ve bilgi erişim performansını doğrudan etkileyen unsurların McKinin, Sievert, Johnson ve Mitchell'in çalışmasında (1991) hiç yer almaması, Hemminger, Saelim, Sullivan ve Vision'ın çalışmasında (2007) ise ayrıntılandırılmamış olması dikkat çekmektedir. Diğer taraftan, deney düzenekleri ayrıntılarıyla işlenmiş ve kontrollü ortamda yapılmış benzer bir araştırmanın sonuçları farklı çıkmıştır. Söz konusu çalışmada (Lin, 2009), Okapi BM25 ve koordinasyon faktöründe değişiklik yapılan Lucene bilgi erişim modelleri ile geliştirilen bilgi erişim sistemlerinde; dokümanların (1) başlık ve özleri, (2) tam-metinleri ve (3) tam-metinlerin paragraflara bölünerek dizinlenmesi üzerinden bilgi erişim performansları karşılaştırılmıştır. Ortalama duyarlılık testi sonucunda, başlık ve öz dizinlemede Okapi BM25 algoritmasının Lucene'den etkili olduğu, Lucene algoritmasının tam-metin bilgi erişim performansının başlık ve öz performansından yüksek olduğu, paragraflara ayrılmış düzeyde dizinlenen tam-metin dokümanların bilgi erişim performansının da "bir bütün olarak tam-metin", "başlık ve öz" dizinleme performansından daha iyi sonuç verdiği ortaya konulmuştur.

Çıktıları ile bilgi kurumlarının bilgi erişim sistemi geliştirme veya seçim süreçlerine katkı sağlamayı hedefleyen bu araştırmanın amacı; üstveri, tam-metin ve hem üstveri hem de tam-metin (melez) içeriklerin bilgi erişim performansı üzerindeki etkisini değerlendirmektir. Bu amaçla, büyük ölçekli Türkçe bir külliyat (corpus) olmaması nedeniyle oluşturulan küçük ölçekli bir külliyat üzerindeki tam-metin doküman ve dokümanların Dublin Core (DC) standardına göre üretilmiş üstverisinden yararlanılarak varsayılan Lucene bilgi erişim modeline göre farklı içerikleri alanlara dayalı dizinleyen üç bilgi erişim sistemi tasarlanmış ve bilgi erişim performansları test edilmiştir. Bilgi erişim sistemlerinden Üstveri Bilgi Erişim Sistemi (ÜBES) DC elementlerinin bir kısmını, Tam-metin Bilgi Erişim Sistemi (TBES) sadece dokümanların tam-metnini, Melez Bilgi Erişim Sistemi (MBES) ise DC elementleriyle birlikte dokümanların tam-metnini alanlara dayalı dizinlemektedir. Bilgi erişim sistemlerinin performanslarının değerlendirilmesi için duyarlılık-anma ve normalize sıralama ölçümleri yapılmıştır. Çalışmanın hipotezleri ise aşağıdaki biçimde yapılandırılmıştır:

- MBES'in ortalama duyarlılık performansı ÜBES'ten ve TBES'ten yüksektir.
- ÜBES ve TBES'in ortalama duyarlılık performansları arasında anlamlı bir fark yoktur.
- TBES'in normalize sıralama performansı ÜBES'ten ve MBES'ten düşüktür.
- Üç bilgi erişim sisteminin seçilen sorulara karşı eriştikleri ortalama doküman sayısı birbirinden farklıdır.

Arka Plan

Bilgi Erişim Sistemleri

Bilgi erişim, kısaca "depolanmış bilgiler içerisinden ilgili (relevant) bilgilerin bulunması" biçiminde tanımlanmaktadır (Dominich, 2008, s. 2). Yao (2004, s. 314), bilgi erişimde herhangi bir şablona bağlı kalmadan, yarı yapılandırılmış veya yapılandırılmamış verilerin organizasyonunun sağlandığını ve veri erişime göre belirsizliğin daha büyük rol oynadığını vurgulamaktadır. Ayrıca, "bilgi erişim" ile "veri erişim" arasındaki farkları Tablo 1'deki biçimde toplayabilmek mümkündür.

Tablo 1
Bilgi erişim ve veri erişim arasındaki farklar (Van Rijsbergen, 1979)

Kriter	Veri erişim	Bilgi erişim
Sorgu eşleşmesi	Tam eşleşme	Kısmi eşleşme, en iyi eşleşme
Sonuç çıkarma	Tümdengelim	Tümevarım
Model	Belirleyici	Olasılık
Sınıflama	Tekil sınıflama	Çoğul sınıflama
Sorgulama dili	Yapay	Doğal
Sorgu şartnamesi	Önceden belirlenmiş	Önceden belirlenmemiş
İstenen öğeler	Eşleştirilebilenler	İlgililer
Hata yanıtı	Hassas	Hassas değil

Salton (1986, s. 648), "otomatik metin erişim sistemi" olarak bilgi erişim sistemini "kullanıcı sorgularını cevaplamada doğal dil dokümanlarını aramak amacıyla geliştirilmiş sistem" olarak tanımlamaktadır. Bir bilgi erişim sisteminin temel işlevi ise kullanıcıların bilgi ihtiyaçlarını karşılamak amacıyla derlemdeki "ilgili" dokümanların tümüne erişmek, ilgisizleri de ayıklamaktır (Tonta, Bitirim ve Sever, 2002, s. 9).

Bilgi erişim sistemlerinde kaynaklardan bilginin edinilmesiyle başlayan dizinleme süreci, metin işlemenin ardından dizinin oluşturulmasıyla son bulmaktadır. Metin işlemede, (1) harf olmayan karakterler boşluklarla yer değiştirilir; (2) tek harfli sözcükler silinir; (3) bütün karakterler küçük harfli yapılır; (4) erişim değeri taşımayan terimlerden oluşturulmuş dur listesinde (stop list) geçen terimler silinir; (5) terimler gövdelenir (stemming); ve (6) tek karakterli gövdeler atılır (Sever ve Tonta, 2006, s. 1). Türkçe metin işleme sürecinde terimlerin gövdelenmesi, küçük ölçekli (Sezer, 1999, s. 65; Eroğlu, 2000, ss. 88-89) ve büyük ölçekli (Can ve diğerleri, 2008) külliyatlarda bilgi erişim performansını artırmaktadır. Diğer taraftan, büyük ölçekli Türkçe külliyat üzerinde dur listelerinin kullanılması bilgi erişim performansını etkilememektedir (Can ve diğerleri, 2008). Güncel bilgi erişim sistemlerinde dizinin oluşturulmasında ise "artımlı dizinleme" (incremental indexing) tekniği kullanılmaktadır. INQUERY bilgi erişim sistemi ile gündeme gelen artımlı dizinleme tekniği öncesinde, bilgi erişim sistemlerinin oluşturduğu devrik dizine yeni bir dokümanın eklenmesi, silinmesi veya güncellenmesi durumunda tüm dizinin yeniden yaratılması gerekmekteydi (Brown, Callan ve Croft, 1994). Artımlı dizinleme tekniği ile bilhassa büyük ölçekli külliyatlar için geliştirilen bilgi erişim sistemlerinde zaman ve maliyet tasarrufu sağlanmıştır.

Bir bilgi erişim sistemi, belirsizliğin ele alındığı "dokümanın temsili", "bilgi ihtiyaçlarının temsili" ve "eşleşme fonksiyonu" olmak üzere üç temel bileşeni bünyesinde barındırmakla birlikte, bazı bilgi erişim sistemleri dördüncü bileşen olarak "ilgililik geri bildirimi"ni de bünyesinde barındırabilmektedir (Turtle ve Croft, 1997, ss. 189-190):

1. Dokümanın temsili: Doküman içeriklerini temsil etmek amacıyla saptanan terimler herkes tarafından kabul görmese de dokümanı temsil etmektedir. Bu alanda otomatik tekniklerin kullanımı belirsizliği daha da artırmaktadır ve hangi terimlerin doküman içeriği hangi derecede temsil edeceği karmaşıktır.

2. Bilgi ihtiyaçlarının temsili: Kullanıcıların bilgi ihtiyaçlarını ifade etme sürecinde de aynı temsil sorunuyla karşılaşılmaktadır. Bilgi ihtiyacı açık bir biçimde temsil edilememektedir. Bu durum, kullanıcının sistemde hangi tür dokümanların yer aldığını görmesiyle, arama stratejisini değiştirmesine neden olmaktadır.

3. Eşleşme fonksiyonu: Eşleşme fonksiyonunda belirsizlik, bilgi ihtiyaçlarının ve dokümanın temsilinden miras alınmıştır. Bilgi ihtiyacının ve dokümanın temsilinde kesinlik olsa bile belirsizlik devam etmektedir, çünkü aynı terim farklı biçimde temsil edilebilmekte ve tek bir temsilde yer alan terimler birbirlerinden bağımsız da değildir.

4. İlgililik geri bildirimi: Kullanıcıdan erişim kümesindeki ilgili dokümanları seçmesi istenebilir. Seçilen ilgili dokümanlar, kullanıcının bilgi ihtiyacı temsilini düzenlenmesinde veya eşleşme fonksiyonunun sonraki erişimi geliştirmesinde kullanılabilir.

Yukarıda yer alan temel belirsizliklere yönelik getirilen çözümler de "bilgi erişim modeli" olarak adlandırılmaktadır. Bilgi erişim literatüründe, Boole ve türevleri (Akıllı, Genişletilmiş Boole), Vektör Uzayı ve türevleri (Gizli Anlamsal Analiz, Genişletilmiş Boole), Olasılık, Dil, Bayesian ağ ve PageRank gibi çeşitli temel bilgi erişim modelleri bulunmaktadır (Baeza-Yates ve Ribeiro-Neto, 1999, ss. 25-27; Deerwester ve diğerleri, 1990; Marcus, 1991; Maron ve Kuhns, 1960; Page, Brin, Motwani ve Winograd, 1998; Ponte ve Croft, 1998; Robertson ve Jones, 1976; Salton, Fox ve Wu, 1982; Salton, Wong ve Yang, 1975; Turtle ve Croft, 1989; Turtle ve Croft, 1991). Çalışma kapsamında kullanılan bilgi erişim modellerinden Boole ve Vektör Uzayı modellerinin kısaca işlenmesinde yarar vardır.

Bilgi Erişim Modelleri

Boole Bilgi Erişim Modeli

Boole bilgi erişim modeli (BBEM), küme (set) teorisi ve Boole cebrine dayalı olarak geliştirilmiş basit bir bilgi erişim modelidir. BBEM, ilk klasik bilgi erişim modeli olmakla beraber, geniş çevrelerce en çok benimsenmiş modeldir (Dominich, 2001, s. 97). BBEM neredeyse tüm veri tabanı yönetim sistemi üreticilerince desteklenip, geliştirilmiştir. Bu durum, modele ulaşımı kolaylaştırdığı gibi kullanımını da yaygınlaştırmıştır.

BBEM'de, olası dizin terimleri Boole işleçleriyle (VE, VEYA, DEĞİL) birbirlerine bağlanarak sorgu oluşturulur (bir başka ifadeyle, bilgi ihtiyacı formüle edilir). Belirlenen koşullar çerçevesinde sorgudaki terim(ler)in, dizindeki terim(ler)le çakışması durumunda ilgililik kararı verilir ve *N* kümeye ilgili olduğu varsayılan yeni bir erişim kümesi tanımlanır. Modele göre, dizin terimleri dokümanda ya geçmektedir ya da geçmemektedir. Bu nedenle, erişim kümesinde ilgili kabul edilen tüm terimlerin ağırlığı 1, erişim kümesine giremeyen terimlerin ağırlıkları ise 0'dır (Baeza-Yates ve Ribeiro-Neto, 1999, ss. 26-27).

BBEM, kullanıcı gruplarının özellikleri göz önünde bulundurulduğunda göreli avantajlar sağlamaktadır. Göker ve Davies (2008, s. 3), modelin uzman kullanıcılarda sistem üzerinde kontrol hissi uyandırdığını, gönderilen sorguya karşılık dokümanın neden geldiğinin ve sonuç kümesinin küçük veya büyük gelmesi durumunda hangi işleçlerle istenilen boyutta sonuç kümesi elde edilebileceğinin kolay anlaşılabilir olduğunu vurgulamaktadır. Ayrıca, modelin kolay uygulanabilir olması ve hesaplama verimliliği de modelin avantajları arasında sayılabilmektedir (Spoerri, 1995, s. 31).

Modelin avantajlarının yanı sıra, dikkate alınması gereken bazı temel dezavantajları da bulunmaktadır. Salton (1984) ve Cooper (1988) genel olarak modelin üç önemli dezavantajı üzerinde durmuştur. Bunlar; Boole formülasyonunun zorluğu, boş çıktı veya fazla yüklü çıktı alınması ve ağırlıklandırma eksikliğidir. Modelin temel dezavantajları aşağıdaki biçimde açıklanabilir:

- Kullanıcıların doğal dilde kullandıkları "VE" ve "VEYA" sözcükleri bilgi erişim sistemlerinde farklı anlamlara gelmektedir. Bilgisayar ve Boole cebri hakkında bilgi sahibi olmayan kullanıcılar VE-VEYA işleçlerinin mantığını kavramada ve sorgu formülize etmede zorluk çekmektedir. Özellikle tecrübesiz kullanıcılar karmaşık sorgularda parantez kullanımı konusunda hata yapabilmekte ve sistemi kullanabilmek için çok çaba sarf etmektedir.
- Boole arama taleplerinde VE işleçlerinin fazla kullanılması durumunda boş çıktı ile karşılaşılabilmektedir. VEYA işleçleriyle oluşturulmuş arama taleplerinde ise çok fazla sonuçla karşılaşılabilmektedir.

- BBEM ile erişilen dokümanlarda ilgililik sıralaması yapma konusunda herhangi bir yaklaşım bulunmamaktadır. Sıralama, yapılandırılmış verilerin karakteristiklerine uygun olarak, "artan-azalan", "büyüktür-küçüktür" veya "arasında" gibi çeşitli kıstaslara göre yapılabilmektedir.

BBEM'in yukarıda sıralanan temel eksikliklerini giderebilmek üzere çeşitli çalışmalar yapılmıştır. Boole sorgusu oluşturma ve boş çıktı veya fazla yüklü çıktı sorunlarının üstesinden gelebilmek üzere, sorgu genişletmeye veya daraltmaya odaklı "akıllı Boolean" (smart Boolean) (Marcus, 1991) geliştirilmiştir. Ağırlıklandırma ve ilgililik sıralaması sorunlarının üstesinden gelmek üzere Vektör Uzayı modelinden de faydalanılarak Genişletilmiş (Extended) Boole Modeli (Salton, Fox ve Wu, 1982) geliştirilmiştir.

Genel bir değerlendirme yapıldığında, BBEM'in yaygın kullanım alanının "bilgi erişim"den ziyade "veri erişim" olduğu dikkat çekmektedir. Bunun sebebi ağırlıklandırmanın ikili (binary) olmasına dayanmaktadır. Ağırlıklandırmanın ikili yapılması "bilgi erişim" modeli olarak tatmin edici olmasa da "veri erişim" için idealdir.

Vektör Uzayı (Vector Space) Bilgi Erişim Modeli

Vektör uzayı bilgi erişim modeli (VUBEM), temel olarak, BBEM'in ikili ağırlıklandırma kaynaklı sıralama yeteneğinin olmayışının üstesinden gelebilmek amacıyla istatistiksel yaklaşımla geliştirilmiştir. Modelin istatistiksel dayanağı ise Luhn tarafından ortaya konulmuştur. Luhn (1957), terimlerin dokümanlardaki geçiş sıklıklarının dokümanı temsil etmede veya doküman için önem belirlemede kullanılmasını önermiştir. Ayrıca, kullanıcıların bilgi ihtiyaçlarını ifade etmek için doküman hazırlayabileceğini, hazırlanan doküman ile külliyattaki dokümanların benzerlik derecelerinin ilgililiğe dayalı sıralamalı sorgu sonucunu verebileceğini ileri sürmüştür.

Salton, Wong ve Yang ise, (1975) Luhn'un istatistiksel yaklaşımını geliştirip güçlü bir model olan VUBEM'i ortaya koymuştur. Bu doğrultuda, VUBEM'i temel olarak aşağıdaki gibi özetleyebilmek mümkündür:

- İkili ağırlıklandırma ideal bir bilgi erişim modeli için oldukça kısıtlıdır. Bu kısıtlamayı ortadan kaldırmak üzere dizin terimlerine, sorgulara ve dokümanlara ikili olmayan ağırlıkların atanması gerekmektedir.
- VUBEM'de dokümanlar ve sorgular *t* boyutlu vektörler olarak gösterilir.
- Ağırlıklandırma, dokümanlar ile sorgular arasındaki benzerlik derecesinin hesaplanmasında kullanılmaktadır.
- Sonuç olarak, kullanıcılara benzerliği azalan bir sıralama ile sonuç kümesi döndürülebilmektedir.

VUBEM'de hem doküman terimleri hem de sorgu terimleri ağırlıklandırılmaktadır. Dokümandaki, sorgudaki ve külliyattaki terimlerin önemini terim ağırlıkları belirlemektedir. VUBEM'de terimlerin ağırlıklandırılmasının ardından sorgu ve doküman vektörleri arasındaki benzerlik hesaplanmaktadır. Benzerliğin hesaplanmasında, iki vektör arasındaki derecenin kosinüs bağıntısı kullanılmaktadır. Çok boyutlu uzayda, vektörler dik ise açının kosinüsü 0'dır, eğer açı 0 ise kosinüsü 1'dir. Bu durumda, 90^0 ile 0^0 arasındaki benzerlik 0 ile 1 arasındaki değerlere tekabül etmektedir. Kosinüs bağıntısı ise Formül 1'deki gibidir.

$$sim\left(d_j, q\right) = \frac{\vec{d_j} \cdot \vec{q}}{|\vec{d_j}| \times |\vec{q}|} = \frac{\sum_{i=1}^{t} w_{i,j} \times w_{i,q}}{\sqrt{\sum_{k}^{t} w_{i,j}^2} \times \sqrt{\sum_{j}^{t} w_{i,q}^2}}$$

Formül 1. Kosinüs bağıntısı (Baeza-Yates ve Ribeiro-Neto, 1999, s. 27; Göker ve Davies, 2008, s. 6)

Kosinüs bağıntısının anlaşılabilmesi için bağıntıda yer alan "ağırlıklandırma" ve "normalizasyon" süreçlerine değinmekte yarar vardır.

Doküman vektöründeki bir terimin ağırlığı farklı yaklaşımlarla belirlenebilmektedir. Terim ağırlıklarının belirlenmesinde en çok bilinen ve yaygın olarak kullanılan yaklaşım *tf x idf* biçiminde ağırlıklandırmadır (bkz. Formül 2).

$$w_{i,j} = tf_{i,j} \times idf_j = tf_{i,j} \times \log N/df_j$$

Formül 2. tf x idf biçiminde ağırlıklandırma (Salton ve Buckley, 1988)

tf x idf biçiminde ağırlıklandırmada, terim ağırlıklarının ($w_{i,j}$) belirlenmesinde iki faktör vardır. Bu faktörlerden ilki terim sıklığıdır (term frequency). Terim sıklığı ($tf_{i,j}$), *j* teriminin *i* dokümanındaki geçiş sıklığını ifade etmektedir. Dokümanda beş defa geçen bir terim ile yüz defa geçen bir terimin söz konusu doküman için farklı önem taşıması gerekmektedir. Bu nedenle, terimlerin ağırlıklandırılmasında sadece dokümanda geçen terimlerin sıklıkları kullanılabilmektedir. Öte yandan, sağlıklı bir ağırlıklandırma yapabilmek için terimlerin dokümandaki geçiş sıklıkları tek başına yetersiz kalabilmektedir. Bir dokümanda sıklıkla geçen terimlerin külliyatta da sıklıkla geçmesi durumunda, bahsi geçen terimle oluşturulmuş sorgunun neredeyse tüm külliyatla ilgili olması gibi bir sonuç ortaya çıkabilmektedir. Bu istenmeyen sonucu ortadan kaldırmak üzere ikinci bir faktör olan devrik doküman sıklığı (inverse document frequency, idf) devreye girmektedir (Spärck-Jones, 1972). *tf x idf* biçiminde ağırlıklandırmada, *idf* ($log\ N/df_j$) faktörü logaritmik bir fonksiyondur ve terimin doküman sıklığının artması durumunda azalma özelliği göstermektedir. Sonuç olarak, *tf x idf* biçiminde ağırlıklandırma sayesinde külliyattaki az sayıda dokümanda geçen terimlere yüksek ağırlıklar atanabilmektedir.

VUBEM'de külliyattaki dokümanların uzunlukları göz önünde bulundurulduğunda, sadece *tf x idf* biçiminde ağırlıklandırma yetersiz kalabilmektedir. Bu sorunun üstesinden gelebilmek üzere doküman uzunlukları normalize edilmektedir. Normalizasyon yapmanın temelinde yatan gerekçeler ise aşağıdaki biçimde ayrıntılandırılabilir (Singhal, Salton, Mitra ve Buckley, 1995; Singhal, Buckley ve Mitra, 1996):

- **Yüksek Terim Sıklıkları:** Uzun dokümanlar aynı terimleri tekrarlı olarak kullanmaktadır. Sonuç olarak, uzun dokümanlar için terim sıklığı faktörleri kısa dokümanlara göre geniş olabilmekte ve bu durum uzun doküman terimlerinde sorgu-doküman benzerliğinin artmasına neden olabilmektedir.
- **Fazla Terim:** Uzun dokümanlar pek çok farklı/ayrık terimi/kavramı bünyelerinde barındırmaktadır. Bir başka ifadeyle, uzun dokümanlar fazla sayıda konuyla ilgilidir. Bu durum, kısa dokümanlarda işlenen az sayıda konuya ait terimlerle fazla konuyu işleyen dokümanlardaki terimlerin bir tutulmasına neden olmaktadır. Sonuç olarak, uzun dokümanların bulunduğu bir külliyatta yapılan arama sonuçlarında farklı konularla da ilgili olan dokümanlara erişilmektedir.

Bilgi erişim sistemlerinde *Kosinüs Normalizasyonu* (Salton, Wong ve Yang, 1975), *Maksimum tf Normalizasyonu* (Salton ve Buckley, 1988; Turtle ve Croft, 1989) ve *Byte Uzunluk Normalizasyonu* (Robertson, Walker, Jones, Hancock-Beaulieu ve Gatford, 1995) gibi çeşitli doküman uzunluk normalizasyonları kullanılmaktadır. Bunlar içerisinde en yaygın kullanıma sahip olan *Kosinüs Normalizasyonu*'dur. Öte yandan, Singhal, Buckley ve Mitra'nın (1996) yaptığı çalışmanın bulguları, *Kosinüs Normalizasyonu*'nun kısa dokümanlara iltimas göstermeye meyilli olduğunu göstermiştir. Aynı çalışmada, bu sorunun üstesinden gelmek üzere *Eksen Doküman Uzunluğu Normalizasyonu* (Pivot Document Length Normalization) geliştirilip, test edilmiş ve klasik *Kosinüs Normalizasyonu*'na göre %18,3 gelişme elde edilmiştir.

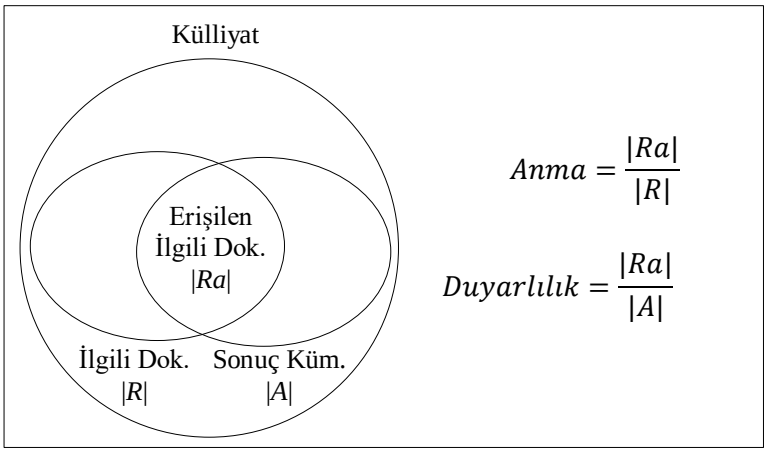
Doküman uzunluklarının bilgi erişim performansını olumsuz yönde etkilemesi, VUBEM'in önde gelen sınırlılıklarından sayılmaktadır. Bu sorunu çözüme kavuşturmak

amacıyla yapılan çalışmalarda önemli ilerlemeler sağlanmış olsa da sorunun tam olarak çözüldüğü söylenemez. VUBEM'in önemli sorunlarından bir başkası da kullanıcıların bilgi ihtiyaçlarını sisteme doğru aktararak bilgi ihtiyacına cevap verebilecek nitelikteki sonuçların istenilen düzeyde alınamamasıdır. Bu amaçla, sorgu formülasyonunda sorgu genişletme ve ilgililik geribildirimi çözümleri üretilmiştir. Küçük ölçekli külliyatlarda yapılan deneylerde bu yaklaşımların duyarlılığı artırdığı tespit edilmiştir (BaezaYates ve Ribeiro-Neto, 1999, s. 30, 118).

Performans Değerlendirme

Bilgi erişim sistemlerinin etkililiklerini veya performansını değerlendirmede veri toplamak veya sorgu oluşturmak amacıyla çeşitli sorular seçilmektedir. Soruların seçiminde kullanıcıların bilgi ihtiyaçlarının ve bilgi erişim sisteminin hangi özelliklerinin değerlendirilmek istendiği önemlidir (Saracevic, 1995, s. 140). Sorgulardan gelen sonuçlara dayalı olarak bilgi erişim sistemlerinin performanslarını değerlendirmede ise farklı yaklaşımlar söz konusudur. Croft, Metzler ve Strohman (2015, ss. 333-334), tüm bilgi erişim sistemlerinin performanslarını doğru ölçebilecek tek bir ölçüm veya yaklaşımın bulunmadığına dikkat çekmekte, değerlendirilecek bilgi erişim sisteminin performansını ortaya koyabilecek bir veya birkaç ölçüm yaklaşımının kombinasyonunu kullanılmayı önermektedir.

Bilgi erişim sistemlerinde performans değerlendirme, çoğunlukla "ilgililiğin" ön plana çıkarıldığı duyarlılık (precision) ve anma (recall) kriterlerine (Kent, Berry, Luehrs ve Perry, 1955) dayanmaktadır (bkz. Şekil 1). Duyarlılık ve anma ölçümünde, sorgulardan dönen sonuç kümelerindeki dokümanlar "ilgili" veya "ilgisiz" olarak ikili (binary) değerlendirilir.



Şekil 1. Duyarlılık ve anma (Baeza-Yates ve Riberio-Neto, 1999, s. 75)

Duyarlılık ve anma ölçümünde 0'dan 1'e kadar (0, 0.1, 0.2, ..., 1) toplam 11 anma basamağı bulunmaktadır. Performans değerlendirmede birden fazla soru sorulduğu için 11 anma basamağında ortalama (average) duyarlılık Formül 3'deki biçimde hesaplanmaktadır.

$$P(r) = \sum_{i=1}^{N_q} \frac{P_i(r)}{N_q}$$

$P(r)$: i. anma basamağındaki ortalama duyarlılığı,
$P_i(r)$: i. sorgu için r anma basamağındaki duyarlığı,
N_q : Kullanılan sorgu sayısını ifade etmektedir.

Formül 3. Ortalama duyarlılık (Baeza-Yates ve Riberio-Neto, 1999, s. 77)

Ayrıca, her sorgu için anma 11 standart anma basamağından ayrık olabilmektedir. Bu durumda, söz konusu anma basamağındaki duyarlılık değerine; kendisi ile kendinden sonra gelen anma basamağındaki maksimum değer atanarak interpolasyon (interpolation) uygulanmaktadır (Baeza-Yates ve Riberio-Neto, 1999, s. 77; Croft, Metzler ve Strohman, 2015, ss. 315-316). İnterpolasyon sonucunda 11 adet anma basamağı elde edildiği gibi, performans

değerlendirmesini görselleştirmede kullanılan duyarlılık-anma grafiği de daha anlamlı hale gelmektedir (Manning, Raghavan ve Schütze, 2008).

İkili ilgililiğe göre performans ölçümünde kullanılan bir başka kriter de kullanıcı merkezli değerlendirme tekniğidir. İngilizce literatürde "anma normalizasyonu" (recall normalization) olarak geçen bu değerlendirme tekniği (Bollmann, 1983; Yao, 1995), Türkçe literatürde "normalize sıralama" (R_{norm}) olarak geçmektedir (Tonta, Bitirim ve Sever, 2002, s. 26). Bu teknikle, ilgisiz dokümanların ilgili dokümanların önüne geçmesi durumunda, performans anma ve duyarlılığa nazaran daha olumsuz etkilenmektedir. R_{norm}, sonuç kümesindeki dokümanların tümü ilgili olduğunda 1, tümü ilgisiz olduğunda 0 değerini almaktadır. 0 ve 1 arasındaki değerleri hesaplamak için Formül 4'ten yararlanılmaktadır.

$$R_{norm} = \frac{1}{2}\left(1 + \frac{C^+ - C^-}{C_{max}}\right)$$

R_{norm} : Normalize sıralama,
C^+ : İlgili dokümanların ilgisiz dokümanların önünde yer aldığı doküman çiftleri sayısı,
C^- : İlgisiz dokümanların ilgili dokümanların önünde yer aldığı doküman çiftleri sayısı,
C_{max} : Mümkün olan en fazla C^+ sayısıdır.

Formül 4. Normalize sıralama (Bollmann, 1983; Yao, 1995)

Ortalama duyarlılık, anma ve R_{norm} kriterleri tüm bilgi erişim sistemlerinin performansını ölçmede kullanılamasa da ilgili dokümanların tamamının tespit edilebildiği küçük ölçekli külliyatlar üzerinde ayrıntılı değerlendirmeler yapabilmek amacıyla kullanılabilmektedir. Arama motorları veya kütüphanelerde kullanılan keşif araçları gibi tüm ilgili dokümanların tespit edilmesinin mümkün olmadığı, büyük ölçekli külliyatlarda "k dokümanda duyarlılık" (P@k) (Buckley ve Voorhees, 2004) veya "birikimli kazanç" (cumulated gain) (Järvelin ve Kekäläinen, 2002; Wang ve diğerleri, 2013) değerlendirme ölçütlerinin kullanılması daha uygun olabilir.

Yöntem, Tasarım ve Sınırlılıklar

Araştırmada "deneysel yöntem" ve "betimleme yöntemi", veri toplamak için "sorgulama" veya "soru sorma" tekniği kullanılmıştır. Külliyat olarak *Türk Kütüphaneciliği (TK)*[1] dergisindeki 2215 tam-metin dokümandan ve dokümanların DC standardına göre oluşturulmuş üstverisinden yararlanılmıştır. Araştırmada kullanılan külliyatın küçük ölçekli olması ve sor(g)ulara karşılık gelen tüm "ilgili" dokümanların tespit edilebilmesi nedeniyle, sonuç kümelerinde ortalama duyarlılık, anma ve R_{norm} performans değerlendirme ölçütleri kullanılmıştır. Sor(g)ular için duyarlılığın 11 standart anma basamağından ayrık olduğu durumlarda interpolasyon uygulanmıştır.

Araştırmada deney düzeneği için aynı bilgi erişim modelini kullanan, farklı içerikleri dizinleyen üç bilgi erişim sistemi tasarlanmıştır. Bilgi erişim sistemlerinin tasarımında; bilgi kurumları ve paydaşlarınca yaygın kullanılması sebebiyle (örneğin; Koha, SolrMarc, Summon, Balcklight, VuFind, DSpace, Europeana, DOAJ vb.) BBEM ve VUBEM'e dayalı varsayılan Apache Lucene algoritması (Similarity, 2010), Türkçe gövdelemede sözlük kullanan birkaç seçenek olsa da (örneğin, Gövdebul (Duran, 1997) ve Zemberek (Akın ve Akın, 2007)) Apache Lucene dışına çıkmadan sadece morfolojik analize dayalı (Eryiğit ve Adalı, 2004) Snowball algoritması (Çilden, 2006) ve pdf dosyalarındaki metinleri çıkarmak için Apache Tika kullanılmıştır.

[1] *Türk Kütüphaneciliği* dergisinin http://tk.kutuphaneci.org.tr/ adresinden yayınladığı tamamı metin tabanlı olan dokümanlar kullanılmıştır. Günümüzde http://www.tk.org.tr/ adresinden yayınlanan dokümanların büyük çoğunluğu imaj tabanlıdır.

Deney düzeneği için tasarlanan bilgi erişim sistemlerinden ÜBES, DC elementlerinden "title/başlık", "author/yazar", "description/öz", "subject/konu" ve "type/tür" elementleri, TBES dokümanın sadece tam-metnini barındıran "fulltext/tam-metin", MBES ise DC elementlerinden "title/başlık", "author/yazar", "description/öz", "subject/konu" ve "type/tür" elementleri ile birlikte dokümanın tam-metnini barındıran "fulltext/tam-metin" ile oluşturulup, Lucene alanlarına (field) eşleştirilerek dizinlenmiştir. Sorgular da bu alanlar üzerinden yapılandırılmıştır.

Bilgi erişim sistemlerinin birbirine göre avantajını ortaya koyabilmek, gövdeleme algoritmasının etkinliğinin ve kullanıcıların yapabileceği muhtemel sorguların değerlendirmesini sağlayabilmek amacıyla aşağıda yer alan dokuz soru seçilmiş ve Ek 1'deki biçimde formülize edilmiştir:

1) İrfan Çakın'ın yazdığı tüm dokümanlar
2) İrfan Çakın'a bilimsel/hakemli dokümanlarda yapılan atıflar
3) "Bilgi arama davranışı"
4) AACR, AACR1 veya AACR2
5) OPAC veya "çevrimiçi katalog"
6) Engelliler veya özürlüler
7) Engelli veya özürlü
8) "Kullanıcılara eğitimler", "okuyuculara eğitimler" veya oryantasyonlar
9) "Kullanıcı eğitimi", "okuyucu eğitimi" veya oryantasyon.

Bulgular

Çalışma kapsamında kullanılan *TK* külliyatına ilişkin tanımlayıcı veri Tablo 2 ve Tablo 3'de yer almaktadır. Devrik dizini oluşturan alanlardaki boşluksuz karakter uzunluklarının sunulduğu Tablo 2'nin "min." sütununda yer alan "öz", "konu" ve "tam-metin" alanlarının 0 olduğu görülmektedir. Bunun nedenleri; 1886 dokümanın (dokümanların yaklaşık %85'i) öz öğesine sahip olmaması, 22 dokümanın İngilizce olması nedeniyle Türkçe konu başlığı verilmemesi ve 5 dokümanın imajdan oluşması, tam-metin barındırmamasıdır. Külliyatta en sık geçen 25 terimin sunulduğu Tablo 3 incelendiğinde "ve", "olarak", "bir", "veya" ve "bu" gibi bilgi erişim için anlamlı olamayan, "dur listesi" oluşturulabilecek terimlerin fazla olduğu görülmektedir. Ayrıca, külliyatta az sayıda tam-metin İngilizce doküman olmasına ve makale başlıklarında, özlerde ve anahtar kelimelerde İngilizce kullanılmasına karşın, "the" (18. sırada) ve "of" (25. sırada) gibi bilgi erişim için anlamlı olmayan İngilizce terimler de listeye girmiştir. Bilgi erişim açısından anlamlı olabilecek ve külliyatta en sık geçen terimler ise sırasıyla, "kütüphane", "bilgi", "kitap", "çalışma", "hizmet", "halk", "kütüphanecilik", "türk" ve "eser"dir.

Tablo 2
Alanlardaki boşluksuz karakter uzunlukları

Alan	Min.	Max.	Ortalama	Ortanca	Toplam
Başlık	5	269	46,2	41	102.342
Yazar	3	108	14,2	14	31.542
Öz	0	2.356	109,3	0	242.261
Konu (Anah. Kel.)	0	260	31,8	28	70.549
Tam-metin	0	294.930	18.793,3	13.050	41.605.165

Tablo 3
Külliyatta en sık geçen 25 terim

Sıra	Terim	Sıklık	Sıra	Terim	Sıklık	Sıra	Terim	Sıklık
1	ve	156.321	10	bilgi	29.798	19	kütüphanecilik	14.829
2	olarak	97.673	11	da	28.590	20	türk	13.965
3	bir	92.524	12	kitap	27.519	21	eser	13.901
4	bu	72.819	13	çalışma	21.728	22	daha	13.815
5	kütüphane	61.213	14	çok	15.891	23	genel	13.401
6	için	41.196	15	hizmet	15.506	24	üzere	12.799
7	yıl	31.693	16	halk	15.409	25	of	12.631
8	ile	31.238	17	yer	15.346			
9	de	31.019	18	the	14.973			

Tablo 4'te ise dokuz soruya karşılık, bilgi erişim sistemlerinin erişebildikleri doküman sayıları yer almaktadır. ÜBES, ilgili 119 dokümanın sadece %45'ine (54 doküman), TBES %77'sine (92 doküman), MBES ise %85'ine (101 doküman) erişebilmiştir. Toplamda en fazla ilgili dokümana MBES, en fazla ilgisiz dokümana TBES ve en az ilgisiz dokümana ÜBES erişmiştir. Uygulanan *Kruskal-Wallis H* testinin sonucuna göre, bilgi erişim sistemlerinin erişebildikleri ortalama doküman sayıları arasındaki fark istatistiksel açıdan anlamlı değildir (H (2) = 5,116, p = 0,077).

Tablo 4
Erişim kümelerindeki ilgili ve ilgisiz dokümanlar

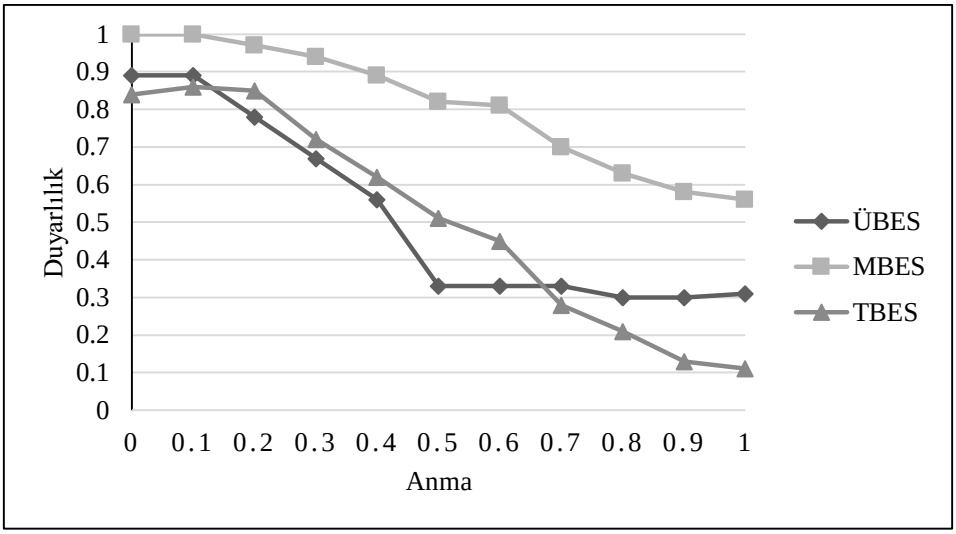
		ÜBES		TBES		MBES	
Soru	Tüm İlgili	Erişilen İlgili	Erişilen İlgisiz	Erişilen İlgili	Erişilen İlgisiz	Erişilen İlgili	Erişilen İlgisiz
1	15	15	0	15	40	15	0
2	14	0	0	14	45	14	0
3	9	4	0	9	8	9	3
4	16	3	0	13	9	13	1
5	21	10	0	14	14	21	14
6	8	8	1	7	29	8	2
7	8	8	1	7	29	8	2
8	14	2	0	5	5	5	5
9	14	4	0	8	16	8	16
Toplam	119	54	2	92	195	101	43

Üç bilgi erişim sisteminin dokuz soruya karşılık 11 anma basamağında sergilediği duyarlılık değerleri Tablo 5'de, interpolasyon uygulanmış ortalama duyarlılık değerleri ise Şekil 2'de yer almaktadır. Ayrıca, her bir bilgi erişim sistemi için ayrıntılı duyarlılık-anma ve R_{norm} ölçümleri Ek 2'de sunulmuştur. Şekil 2'de yer alan üç bilgi erişim sisteminin ortalama duyarlılıkları *Kruskal-Wallis H* ile test edilmiş ve istatistiksel açıdan anlamlı bir fark bulunmuştur (H (2) = 8,595, p = 0,014). Farkın hangi bilgi erişim sistem(ler)inden kaynaklandığını tespit etmek üzere *Mann-Whitney U* testi yapılmış ve farkın MBES'ten kaynaklandığı saptanmıştır. MBES'in ortalama duyarlılığı hem ÜBES'ten (U = 20,5, p = 0,008, Z = -2,635, r = -0,56) hem de TBES'ten (U = 24, p = 0,017, Z = -2,397, r = -0,51) farklıdır. ÜBES ve TBES arasındaki ortalama duyarlılık performans farkı ise istatistiksel açıdan anlamlı değildir (U = 54, p = 0,669, Z = -0,428, r = -0,09). Bir başka ifadeyle, insana dayalı üstveri dizinleme (ÜBES) ile makinaya dayalı otomatik tam-metin dizinleme (TBES) arasında bilgi erişim performansı bakımından anlamlı bir fark yoktur. Öte yandan, insana dayalı dizinleme ile makineye dayalı otomatik dizinleme yaklaşımlarının bir arada kullanılması (MBES) bilgi erişim performansı artmaktadır.

Ayrıca, her bir bilgi erişim sistemi için duyarlılık ve anma arasındaki ilişki test edilmiş, güçlü bir negatif korelasyon saptanmıştır (ÜBES için r(9) = -0,926, p = 0,00 TBES için r(9) = -0,984, p = 0,00 MBES için r(9) = -0,982, p = 0,00). Her bir bilgi erişim sistemi için anma değeri arttığında duyarlılık değeri düşmüştür.

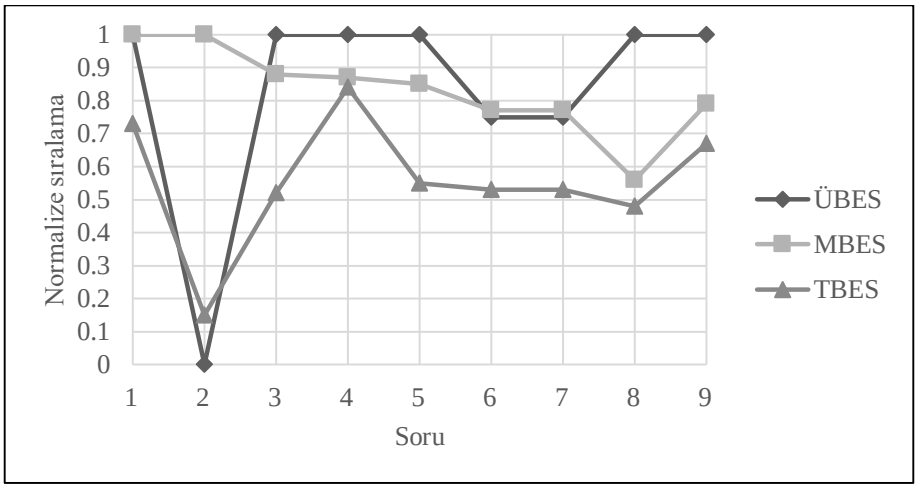
Tablo 5
11 anma basamağında ÜBES (Ü), MBES (M) ve TBES'in (T) ortalama duyarlılığı

	0			0,1			0,2			0,3			0,4			0,5			0,6			0,7			0,8			0,9			1		
Soru	Ü	M	T	Ü	M	T	Ü	M	T	Ü	M	T	Ü	M	T	Ü	M	T	Ü	M	T	Ü	M	T	Ü	M	T	Ü	M	T	Ü	M	T
1	1	1	1	1	1	1	1	1	1	1	1	0,63	1	1	0,67	1	1	0,42	1	1	0,45	1	1	0,45	1	1	0,37	1	1	0,38	1	1	0,27
2	-	1	0,03	-	1	0,06	-	1	0,1	-	1	0,11	-	1	0,13	-	1	0,17	-	1	0,18	-	1	0,2	-	1	0,21	-	1	0,23	-	1	0,24
3	1	1	1	1	1	1	1	1	1	1	1	0,75	1	1	0,8	-	1	0,5	-	1	0,54	-	1	0,58	-	0,88	0,53	-	0,88	0,53	-	0,81	0,52
4	1	1	1	1	1	1	1	1	1	-	1	1	-	1	1	-	1	1	-	1	0,76	-	0,6	0,6	-	0,41	0,41	-	-	-	-	-	-
5	1	1	1	1	1	1	1	1	0,83	1	1	0,53	1	1	0,52	-	1	0,5	-	0,92	0,52	-	0,66	-	-	0,6	-	-	0,62	-	-	0,6	-
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0,8	1	1	0,62	1	1	0,33	0,87	0,87	0,19	0,87	0,87	-	0,88	0,81	-
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0,8	1	1	0,62	1	1	0,33	0,87	0,87	0,19	0,87	0,87	-	0,88	0,8	-
8	1	1	0,5	1	1	0,66	-	0,75	0,75	-	0,5	0,5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9	1	1	1	1	1	1	1	1	1	1	1	1	-	1	0,46	-	0,38	0,38	-	0,33	0,33	-	-	-	-	-	-	-	-	-	-	-	-
Ortal.	0,89	1	0,84	0,89	1	0,86	0,78	0,97	0,85	0,67	0,94	0,72	0,56	0,89	0,62	0,33	0,82	0,51	0,33	0,81	0,45	0,33	0,7	0,28	0,3	0,63	0,21	0,3	0,58	0,13	0,31	0,56	0,11



Şekil 2. Bilgi erişim sistemlerinin ortalama duyarlılık ve anma performansı

Bilgi erişim sistemlerinin normalize sıralama erişim performansları ise Şekil 3'te sunulmuştur. Şekil 3'de yer alan üç bilgi erişim sisteminin R_{norm} performansları *Kruskal-Wallis H* ile test edilmiş ve istatistiksel açıdan anlamlı bir fark bulunmuştur (H (2) = 11,309, p = 0,004). Farkın hangi bilgi erişim sistem(ler)inden kaynaklandığını tespit etmek üzere *Mann-Whitney U* testi yapılmış ve farkın ÜBES ve MBES'ten kaynaklandığı saptanmıştır. Yani, hem ÜBES (U = 11, p = 0,008, Z = -2,656, r = -0,62) hem de MBES'in (U = 6, p = 0,002, Z = -3,051, r = -0,71) ortalama R_{norm} performansı TBES'den farklıdır. ÜBES ve MBES arasındaki fark ise istatistiksel açıdan anlamlı değildir (U = 31, p = 0,380, Z = -0,879, r = -0,21).



Şekil 3. Bilgi erişim sistemlerinin R_{norm} performansı

Tartışma

Bu araştırmada, bilgi erişim sistemleri arasında en fazla "ilgili" dokümana MBES erişmiştir. MBES'te hem üstverinin hem de doküman tam-metninin dizinlenmesi sorgu ve dokümanları temsil eden terimlerin daha fazla çakışmasına neden olmuştur. Bir başka ifadeyle, MBES, üstverideki az ama bilgi erişim açısından yoğunluğa/öneme sahip ve temsil yeteneği yüksek terimlerle birlikte tam-metindeki fazla sayıdaki ayrık terimi işleme avantajlarını kullanarak ilgili dokümanların %85'ine erişmiştir. MBES'in eriştiği ilgili doküman sayısı diğer bilgi erişim sistemlerinden fazla olmasına karşın, eriştiği ilgisiz doküman sayısı (43) TBES'ten (195) çok daha azdır. Bunun sebebi, sorgu terimlerinin dokümanı temsil eden terimlerle kısa üstveri alanları üzerinde çakışmasıdır. Kısa alanlar, doküman uzunluk normalizasyonunun kısa alanlara yüksek skor ataması sebebiyle ilgili dokümanları erişim kümesinin üst sırasına taşımaktadır. Bu durum, hem ortalama duyarlılık hem de R_{norm} performansını olumlu yönde etkilemektedir. Böylece, MBES hem en fazla ilgili dokümana erişebilen hem de en yüksek duyarlılığa sahip bilgi erişim sistemi olmuş ve araştırma hipotezlerinden "MBES'in duyarlılık performansı ÜBES'ten ve TBES'ten yüksektir" hipotezi kabul edilmiştir. Ayrıca, MBES R_{norm} performansı bakımından da ÜBES'e en yakın bilgi erişim sistemi olmuştur.

Araştırmada, ÜBES ve TBES arasındaki ortalama duyarlılık performansında istatistiksel olarak anlamlı bir farka ulaşılamamış ve "ÜBES ve TBES'in ortalama duyarlılık performansları arasında anlamlı bir fark yoktur" hipotezi kabul edilmiştir. Ulaşılan bu sonuç, üç araştırmanın sonucuyla örtüşmemektedir. Hemminger ve diğerlerinin (2007) deney düzeneği ayrıntılandırılmamış araştırmasında, üstveri ve tam-metin bilgi erişim performansları tek bir bilgi ihtiyacından yola çıkılarak, sadece gen isimleri üzerinden ayrı ayrı test edilmiş ve üstveri işleyen bilgi erişim sisteminin duyarlılık performansı tam-metin işleyen bilgi erişim sistemine göre anlamlı olarak yüksek çıkmıştır. Aynı sonuca, McKinin, Sievert, Johnson ve Mitchell'in (1991) farklı bilgi ihtiyaçlarından yola çıkarak deney düzeneğini belirtmeden yaptığı çalışmada da ulaşılmıştır. Lin'in (2009) koordinasyon faktörü değiştirilen Lucene algoritması ile yaptığı çalışmada ise sadece tam-metne dayalı bilgi erişimde ortalama duyarlılık performansı sadece başlık ve öze dayalı performanstan yüksek çıkmıştır. *TK* külliyatındaki dokümanların üstveri kümelerinin yaklaşık %85'inde öz öğesinin bulunmaması, Lucene algoritmasındaki değişiklikler veya diğer araştırmadan faklı olarak *TK* külliyatındaki üstveri kümelerinde anahtar kelimelerin yer alması neticesinde iki ayrı sonuca ulaşılmış olabilir. Diğer taraftan, büyük ölçekli bir külliyat üzerinde yapılan başka bir araştırmada (Kim, Myaeng ve Yoo, 2005), Bayesian ağ modeline göre hem üstveri hem de tam-metin alanları üzerinden otomatik sorgu oluşturan bir bilgi erişim sisteminin sadece üstveri ve sadece tam-metin sorgularından daha iyi sonuç verdiği ortaya konulmuştur. Ayrıca, TK'ya benzer ölçekte olan bir külliyat üzerinde varsayılan Lucene bilgi erişim modelini kullanan bir bilgi erişim sistemine kullanıcıların yönelttiği sorguları işlem günlüğü analiziyle inceleyen başka bir araştırmanın bulgularına göre (Waugh, Tarver, Phillips ve Alemneh, 2015), kullanıcılar gerçek hayatta sorgularının %16'sını üstveri, %9'unu tam-metin ve %75'ini hem üstveri hem de tam-metin alanları üzerinde yapılandırmaktadır. Dolayısıyla, kullanıcıların gerçek hayatta yapılandırdığı sorgulara ilişkin davranışlarını MBES'in özellikleriyle açıklayabilmek mümkün olabilir.

Bilgi erişim sistemleri arasında en az ilgisiz dokümana ÜBES erişmiştir. Bunun sebebi, ÜBES'in az sayıda temsil kabiliyeti yüksek olan ayrık terimi dizinlemesidir. Külliyatta yer alan dokümanların %85'inin öz alanına sahip olmaması, ÜBES'te dizinlenen ayrık terim sayısının az olmasının temel nedenidir. Bu durum, az sayıda ilgisiz dokümana erişilmesine neden olmuştur. Diğer taraftan, ayrık terim sayısının az olması doküman temsilini olumsuz yönde etkilemiş ve ÜBES'in en az sayıda ilgili dokümana erişmesine neden olmuştur. Böylece, ÜBES hem en az sayıda ilgili (tüm dokümanların %45'i) hem de en az sayıda ilgisiz (2 adet) dokümana erişmiştir. Bu durum, ortalama duyarlılık ve R_{norm} değerlerini de etkilemiştir. ÜBES'in en düşük duyarlılık değeri 0,3 olmuştur. ÜBES, R_{norm} değerlerinde de soruların %65'inde maksimum performans (1) sergilemiştir.

Sadece tam-metne/doğal dile dayalı otomatik dizinleme yapan TBES en fazla sayıda ilgisiz dokümana erişmiştir. Bunun sebebi, ayrık terim sayısının fazla olmasına dayanmaktadır. Fazla sayıda ortak terim birçok dokümanda geçtiği için erişilen ilgisiz doküman sayısı artmıştır. Diğer taraftan, ayrık terim sayısının fazla olması doküman temsilini yükseltmiş ve TBES'in ikinci sırada en fazla ilgili dokümana erişilebilmesine neden olmuştur. Böylece, TBES en fazla ilgisiz dokümana (195) ve ikinci sırada en fazla ilgili dokümana (tüm dokümanların %77'si) erişmiştir. Erişilen ilgisiz doküman sayısı TBES'te duyarlılık performansına yansımış, bilgi erişim sistemleri arasında en düşük duyarlılık değeri olan 0,11'e inmiştir. En fazla sayıda ilgisiz dokümana erişmesi nedeniyle R_{norm} bakımından da en düşük performansı sergileyen TBES olmuştur. Böylece, araştırma hipotezlerinden "TBES'in normalize sıralama performansı ÜBES'ten ve MBES'ten düşüktür" hipotezi kabul edilmiştir.

Araştırmada, ÜBES'in erişim noktalarının veya ayrık terim sayısının az olması sebebiyle erişebileceği doküman ortalamasının düşük olması beklenmiştir. Ancak, araştırma sonucunda üç bilgi erişim sisteminin sor(g)ulara karşılık döndürdüğü ilgili doküman ortalamaları aralarında istatistiksel açıdan anlamlı bir farka ulaşılamamış ve araştırma hipotezlerinden "Üç bilgi erişim sisteminin seçilen sorulara karşı eriştikleri ortalama doküman sayısı birbirinden farklıdır" hipotezi reddedilmiştir.

Doğal dille ilgili yapılan çalışmalarda karşılaşılan doküman ve sorgu temsili sorunlarıyla bu çalışmada da karşılaşılmıştır. Örneğin, "OPAC, 'çevrimiçi katalog'" sorgusuyla söz konusu terimler yerine tam-metinde sadece "kütüphane otomasyonu"nun kullanıldığı dokümana TBES erişememiş, dizinleyicinin OPAC veya "çevrimiçi katalog" anahtar kelimelerini atadığı üstveriyi dizinleyen ÜBES ve MBES erişebilmiştir. Buna benzer bir durum geniş veya dar terim kullanılmasıyla da ortaya çıkmıştır. Örneğin, görme engellileri işleyen bir dokümanda "engelli, özürlü" sorgusundaki hiçbir terim tam-metinde geçmemiş, bu geniş terimler yerine daha dar olan "kör" ve "âmâ" terimleri geçmiştir. Ayrıca, kısaltmalarla ve açık terimlerle yapılan sorgularda eş anlamlı terimlerin çakışmaması da erişim performansını olumsuz yönde etkilemiştir (Örneğin, Anglo-American Cataloguing Rules - AACR2, Anglo-Amerikan Kataloglama Kuralları - AAKKII veya AAKK2). Bu tip temsil sorunları bilgi erişimin sistemlerinin zayıf yönleri olarak değerlendirilmektedir (Beall, 2008). Söz konusu zayıf yönleri güçlendirmek üzere sorgu daraltmak veya genişletmek için gömüden (thesaurus), eş anlamlar ve kısaltmalar sözlüğünden faydalanılabilir.

Bu çalışmada, Snowball gövdeleme algoritması kendinden beklenen görevi büyük ölçüde yerine getirmiştir. Bununla birlikte, gövdeleme algoritmasının özel isim ve Türkçe dilbilgisine dayalı (örneğin; yapım-çekim ekleri ve sert sessizlerin benzeşmesi-yumuşaması vb.) terim işleme yeterliliğinin sınırlı olduğu görülmüştür (Örneğin; Girdi: "İrfan Çakın" Çıktı:"irfa çak", Girdi: "OPAC çevrimiçi katalog" Çıktı: "opaç çevrimiç katalogu", Girdi: "kullanıcı okuyucu eğitimi oryantasyon" Çıktı: "kullanıç okuyuç eğit oryantasyo"). Sorgu terimlerinin ve doküman terimlerinin gövdelenmesi durumunda, algoritma büyük ölçüde aynı gövdeye ulaşmış ve sorgu-doküman çakışması gerçekleşmiştir. Diğer taraftan, yapım ve çekim eklerinin iyi işlenemediği durumlarda tek gövdeye indirgenmesi gereken bir terim birden çok gövde ile temsil edilebilmiştir. Bunun sonucunda, sorgu-doküman benzerliği olumsuz etkilenmiştir.

Sonuç ve Öneriler

Bu çalışmada kütüphaneler ve paydaşlar için küçük ölçekli Türkçe üstveri ve tam-metin işlemek üzere BBEM ve VUBEM'e dayalı olarak geliştirebilecek bilgi erişim sistemlerinin etkinlikleri/performansları kullanıcıların yapabileceği muhtemel sorgular üzerinden test edilmiştir. Çalışma sonucunda, ÜBES ile TBES ayrı ayrı ele alındığında, farklı içerikleri işleyen iki bilgi erişim sisteminin terim yönetimi bakımından avantajlı ve dezavantajlı yönlerinin olduğu tespit edilmiştir. Bunun sonucunda, ÜBES ve TBES'in duyarlılık performansı

bakımından birbirlerine üstünlük sağlayamadığı, üstveri ve tam-metin içerik dizinleyen MBES'in duyarlılık performansının anlamlı olarak ÜBES ve TBES'ten yüksek olduğu saptanmıştır. Diğer taraftan, erişim çıktıları kullanıcı merkezli olarak değerlendirildiğinde, ÜBES ve MBES'in R_{norm} performansının TBES'ten yüksek olduğu saptanmıştır. Ayrıca, araştırmadan seçilen tüm sorulara karşılık üç bilgi erişim sisteminin erişebildiği ilgi doküman sayılarının ortalamaları karşılaştırıldığında anlamlı bir fark bulunmamıştır.

Araştırma sonucunda aşağıda yer alan önerilerde bulunulabilir:

- İlgililiği artırmak amacıyla üstveri ve tam-metin dizinleme bir arada kullanılabilir. İnsana dayalı üstveri çıkarma faaliyetlerinde zaman ve maliyet unsurlarının sınırlılık oluşturulduğu durumlarda sadece tam-metin dizinleme yapılabilir.
- Kosinüs uzunluk normalizasyonunu kullanan ve alana dayalı dizinleme yapan bilgi erişim sistemlerinde herhangi bir özel ağırlıklandırma şeması geliştirmeye gerek olmayabilir zira uzunluk normalizasyonun kısa dokümanlara (veya alanlara) yüksek skor atamasından faydalanarak alan uzunluklarına göre dinamik bir ağırlıklandırma elde edebilir.
- Bilgi ihtiyaçlarının ve dokümanların temsilinde kullanılan terimlerin daha sağlıklı çakışabilmesi için gömü, eş anlamlılar sözlüğü, kısaltmalar sözlüğü ve eş anlamlı kısaltmalar sözlüğünün sorgu genişletmede kullanılması veya sorgu genişletme-daraltma seçeneklerinin kullanıcıya sunulması bilgi erişim performansının iyileşmesine neden olabilir.
- Snowball algoritmasının Türkçe dilbilgisi kuralları çerçevesinde morfolojik analiz yapma yeterliliğinin bazı durumlarda sınırlı olduğu görülmektedir. Snowball algoritmasının, Zemberek gibi sözlük de kullanan başka bir doğal dil işleme kütüphanesiyle karşılaştırılıp, bilgi erişim performansına etkisinin test edilmesi faydalı olabilir.
- Bu araştırma küçük ölçekli bir külliyat üzerinde yapılmıştır. Ayrıca, külliyattaki dokümanların yaklaşık %85'inin DC üstveri kümesinde öz öğesi yer almamıştır. Bu sınırlılıklar, araştırma sonuçlarına dayanarak bazı genellemeleri yapmaya engel olabilir. Özellikle, büyük ölçekli sayılabilecek külliyat için (örneğin, kütüphanelerde kullanılan keşif araçları veya elektronik belge yönetim sistemleri) bu araştırmanın tekrarlanması faydalı olabilir.

Teşekkür

Danışmanım Prof. Dr. Nazan Özenç Uçak ve tez jüri üyelerim Prof. Dr. Yaşar Tonta, Prof. Dr. Serap Kurbanoğlu, Prof. Dr. Bülent Yılmaz ve Doç. Dr. Mehmet Toplu'ya katkılarından dolayı teşekkür ederim. Ayrıca, araştırmada kullanılan Türk Kütüphaneciliği külliyatının oluşturulmasını destekleyen EBSCO Information Services (Erol Gökduman) ve Türk Kütüphaneciler Derneği'ne (Ali Fuat Kartal ve Dr. M. Tayfun Gülle) teşekkür ederim.

Kaynakça

Akın, A. A. ve Akın, M. D. (2007). *Zemberek, an open source NLP framework for Turkic languages*. Erişim adresi: http://citeseerx.ist.psu.edu/viewdoc/summary? doi=10.1.1.556.69

Anderson, J. D. ve Pérez-Carballo, J. (2001). The nature of indexing: How humans and machines analyze messages and texts for retrieval. Part II: Machine indexing, and the allocation of human versus machine effort. *Information Processing & Management, 37*(2), 255-277.

Baeza-Yates, R. ve Ribeiro-Neto, B. A. N. (1999). *Modern information retrieval*. New York: ACM Press.

Beall, J. (2008). The weaknesses of full-text searching. *The Journal of Academic Librarianship, 34*(5), 438-444.

Bollmann, P. (1983). The normalized recall and related measures. *Proceedings of the 6th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '83)* içinde (ss. 122-128). New York, NY, USA: ACM.

Brown, E. W., Callan, J. P. ve Croft, W. B. (1994). Fast incremental indexing for full-text information retrieval. Jorge B. Bocca, Matthias Jarke, Carlo Zaniolo (Yay. Haz.). *Proceedings of the 20th International Conference on Very Large Data Bases (VLDB '94)* içinde (ss. 192-202). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Buckley, C. ve Voorhees, E. M. (2004). Retrieval evaluation with incomplete information. *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '04)* içinde (ss. 25-32).

Bush, V. (1945). As we may think. *The Atlantic Monthly*, (Temmuz), 112-124.

Can, F., Kocberber, S., Balcik, E., Kaynak, C., Ocalan, H. C. ve Vursavas, O. M. (2008). Information retrieval on Turkish texts. *Journal of the American Society for Information Science and Technology*, *59*(3), 407-421.

Cleverdon, C. W. (1960). *ASLIB Cranfield research project: Report on the first stage of an investigation into the comparative efficiency of indexing systems* (Teknik Rapor). Erişim adresi: https://dspace.lib.cranfield.ac.uk/handle/1826/1122

Cleverdon, C. W. (1970). *The effect of variations in relevance assessments in comparative experimental tests of index languages* (Teknik Rapor). Erişim adresi: https://dspace.lib.cranfield.ac.uk/handle/1826/967

Cleverdon, C. W. (1977). *A comparative evaluation of searching by controlled language and natural language in experimental N.A.S.A. data base* (Teknik Rapor). Erişim adresi: https://dspace.lib.cranfield.ac.uk/handle/1826/1365

Cleverdon, C. W. ve Keen, M. (1966). *Aslib Cranfield research project: Factors determining the performance of indexing systems; Volume 2, Test results* (Teknik Rapor). Erişim adresi: https://dspace.lib.cranfield.ac.uk/handle/1826/863

Croft, W. B., Metzler, D. ve Strohman, T. (2015). *Search engines: Information retrieval in practice*. Pearson Education. Erişim adresi: http://ciir.cs.umass.edu/irbook/

Cooper, W. S. (1988). Getting beyond Boole. *Information Processing and Management*, *24*(3), 243-248.

Çilden, E. K. (2006). *Stemming Turkish words using Snowball*. Erişim adresi: http://img.eba.gov.tr/542/7b6/2ce/3d5/995/c04/9a5/b2b/041/2a6/8ed/829/046/5ac/002/5427b62ce3d5995c049a5b2b0412a68ed8290465ac002.pdf

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K. ve Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, *41*(6), 391-407.

Dominich, S. (2001). *Mathematical foundations of information retrieval*. Dordrecht: Kluwer Academic Publishers.

Dominich, S. (2008). *The modern algebra of information retrieval*. Berlin: Springer.

Duran, G. (1997). *Gövdebul: Türkçe gövdeleme algoritması*. Yayımlanmamış yüksek mühendislik tezi, Hacettepe Üniversitesi, Ankara.

Eroğlu, M. (2000). *Gövdelemenin ve gömünün Türkçe bir bilgi erişim sistemi üzerindeki etkisinin araştırılması*. Yayımlanmamış yüksek mühendislik tezi, Hacettepe Üniversitesi, Ankara.

Eryiğit, G. ve Adalı, E. (2004). An affix stripping morphological analyzer for Turkish. *Proceedings of the IASTED International Conference on Artificial Intelligence and Applications* içinde (ss. 299-304). Innsbruck, Austria. Erişim adresi: http://web.itu.edu.tr/gulsenc/papers/iasted.pdf

Garfield, E. (1979). *Citation indexing, its theory and application in science, technology, and humanities*. New York: Wiley.

Göker, A. ve Davies, J. (Ed.). (2008). *Information retrieval: Searching in the 21st century*. Chichester: Wiley.

Hemminger, B. M., Saelim, B., Sullivan, P. F. ve Vision, T. J. (2007). Comparison of full-text searching to metadata searching for genes in two biomedical literature cohorts. *Journal of the American Society for Information Science and Technology*, *58*(14), 2341-2352.

Järvelin, K. ve Kekäläinen, J. (2002). Cumulated gain-based evaluation of IR techniques. *ACM Trans. Inf. Syst.*, *20*(4), 422-446.

Kent, A., Berry, M. M., Luehrs, F. U. ve Perry, J. W. (1955). Machine literature searching VIII. Operational criteria for designing information retrieval systems. *American Documentation*, *6*(2), 93-101. doi:10.1002/asi.5090060209

Kim, S. S., Myaeng, S. H. ve Yoo, J. M. (2005). A hybrid information retrieval model using metadata and text. E. A. Fox, E. J. Neuhold, P. Premsmit ve V. Wuwongse (Ed.), *Digital Libraries: Implementing Strategies and Sharing Experiences* içinde, Lecture Notes in Computer Science (ss. 232-241). Springer Berlin Heidelberg.

Lin, J. (2009). Is searching full text more effective than searching abstracts? *BMC Bioinformatics*, *10*, 46. doi:10.1186/1471-2105-10-46

Luhn, H. P. (1957). A statistical approach to mechanised encoding and searching of library information. *IBM Journal of Research and Development, 1,* 309-317.

Manning, C.D., Raghavan, P. ve Schütze, H. (2008). Evaluation of ranked retrieval results. Erişim adresi: http://nlp.stanford.edu/IR-book/html/htmledition/evaluation-of-ranked-retrieval-results-1.html

Marcus, R. (1991). Computer and human understanding in intelligent retrieval assistance. *Proceedings of the 54th American Society for Information Science meeting* içinde (ss. 49-59), Washington: Medford.

Maron, M. E. ve Kuhns, J. L. (1960). On relevance, probabilistic indexing and information retrieval. *J. ACM*, *7*(3), 216-244.

McKinin, E. J., Sievert, M., Johnson, E. D. ve Mitchell, J. A. (1991). The Medline/full-text research project. *Journal of the American Society for Information Science*, *42*(4), 297-307.

Page, L., Brin, S., Motwani, R. ve Winograd, T. (1998). *The PageRank citation ranking: Bringing order to the web*. CA: Stanford University. Erişim adresi: http://ilpubs.stanford.edu:8090/422/

Ponte, J. M. ve Croft, W. B. (1998). A language modeling approach to information retrieval. *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '98)* içinde (ss. 275-281). New York, NY, USA: ACM.

Robertson, S. E. ve Jones, K. S. (1976). Relevance weighting of search terms. *Journal of the American Society for Information Science*, *27*(3), 129-146.

Robertson, S. E., Walker, S., Jones, S., Hancock-Beaulieu, M. M. ve Gatford, M. (1995). Okapi at TREC-3. D. K. Harman (Yay. Haz.). *Proceedings of the Third Text REtrieval Conference (TREC-3)* içinde (ss. 109-126). Gaithersburg, MD: NIST.

Salton, G. (1984). The use of extended Boolean logic in information retrieval. *Proceedings of the 1984 ACM SIGMOD International Conference on Management of Data (SIGMOD '84)* içinde (ss. 277-285). New York, NY, USA: ACM.

Salton, G. (1986). Another look at automatic text-retrieval systems. *Commun. ACM. 29*(7), 648-656.

Salton, G. ve Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Inf. Process. Manage.*, *24*(5), 513-523.

Salton, G., Fox, E. A. ve Wu, H. (1982). *Extended Boolean information retrieval*. Erişim adresi: http://ecommons.library.cornell.edu/handle/1813/6351

Salton, G., Wong, A. ve Yang, C. S. (1975). A Vector Space Model for information retrieval. *Journal of the American Society for Information Science, 18*(11), 613-620.

Sanderson, M. ve Croft, W. B. (2012). The history of information retrieval research. *Proceedings of the IEEE, 100* (Special Centennial Issue), 1444-1451. doi:10.1109/JPROC.2012.2189916

Saracevic, T. (1995). Evaluation of evaluation in information retrieval. Edward A. Fox, Peter Ingwersen, Raya Fidel (Yay. Haz.). *Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '95)* içinde (ss. 138-146). New York, NY, USA: ACM.

Sever, H. ve Tonta, Y. (2006). Arama motorları. *Türkiye Bilişim Ansiklopedisi* içinde (ss. 95-99). İstanbul: Papatya Yayınları. Erişim adresi: http://yunus.hacettepe.edu.tr/~tonta/yayinlar/Turkbilisimansiklopedisi.pdf

Sezer, E. (1999). *Smart Bilgi Erişim Sistemi'nin Türkçe yerelleştirilmesi ve otomatik gömü üretimi*. Yayımlanmamış yüksek mühendislik tezi, Hacettepe Üniversitesi, Ankara.

Shields, G. (2005). *What are the main differences between human indexing and automatic indexing?*. Erişim adresi: http://www.shieldsnetwork.com/LI842_Shields_Automatic_Indexing.pdf

Similarity. (2010). Erişim adresi: http://lucene.apache.org/core/3_0_3/api/core/org/apache/lucene/search/Similarity.html

Singhal , A., Buckley, C. ve Mitra, M. (1996). Pivoted document length normalization. *Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '96)* içinde (ss. 21-29). New York, NY, USA: ACM.

Singhal, A., Salton, G., Mitra, M. ve Buckley, C. (1995). *Document length normalization* (Teknik Rapor). Cornell University. Erişim adresi: http://ecommons.cornell.edu/handle/1813/7186

Spärck-Jones, K. (1972). A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation, 28*(1), 11-21.

Spoerri, A. (1995). *INFOCRYSTAL: A visual tool for information retrieval*. Yayımlanmamış Doktora Tezi, Massachusetts Institute of Technology, Cambridge. Erişim adresi: http://hdl.handle.net/1721.1/36946

Tonta, Y., Bitirim, Y. ve Sever, H. (2002). *Türkçe arama motorlarında performans değerlendirme.* Ankara: Total Bilişim Ltd. Şti.

Turtle, H. ve Croft, W. B. (1989). Inference networks for document retrieval. Jean-Luc Vidick (Yay. Haz.). *Proceedings of the 13th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '90)* içinde (ss. 1-24). New York, NY, USA: ACM.

Turtle, H. ve Croft, W. B. (1991). Evaluation of an inference network-based retrieval model. *ACM Trans. Inf. Syst., 9*(3), 187-222.

Turtle, H. R. ve Croft, W. B. (1997). Uncertainty in information retrieval systems. A. Motro, P. Smets (Yay. Haz.). *Uncertainty management in information systems: From needs to solutions* içinde (ss. 189-224). Boston: Kluwer Academic.

Wang, Y., Wang, L., Li, Y., He, D., Liu, T.-Y. ve Chen, W. (2013). A theoretical analysis of NDCG type ranking measures. Shai Shalev-Shwartz ve Ingo Steinwart (Yay. Haz.). *26th Conference on Learning Theory (COLT)* içinde (ss. 25-54). Erişim adresi: http://www.jmlr.org/proceedings/papers/v30/Wang13.pdf

Waugh, L., Tarver, H., Phillips, M. ve Alemneh, D. (2015). Comparison of full-text versus metadata searching in an institutional repository: Case study of the UNT Scholarly Works. *arXiv:1512.07193 [cs]*. Erişim adresi: http://arxiv.org/abs/1512.07193

Van Rijsbergen, C. J. (1979). Information retrieval: Introduction. Erişim adresi: http://www.dcs.gla.ac.uk/Keith/Chapter.1/Ch.1.html

Yao, Y. Y. (1995). Measuring retrieval effectiveness based on user preference of documents. *Journal of the American Society for Information Science, 46*(2), 133-145.

Yao, Y. Y. (2004). Granular computing for the design of information support systems. W. Wu, H. Xiong, S. Shekhar (Yay. Haz.). *Clustering and Information Retrieval* içinde (ss. 299-329). Dordrecht: Kluwer Academic Publishers.

Summary

Information institutions use text-based information retrieval systems to store, index and retrieve metadata, full-text, or both metadata and full-text (hybrid) contents. The aim of this research was to evaluate impact of these contents on information retrieval performance. For this purpose, metadata (MIR), full-text (FIR) and hybrid (HIR) content information retrieval systems were developed with default Lucene information retrieval model for a small scale Turkish corpus. In this regard, Turkish Librarianship Journal's 2215 text-based documents and their Dublin Core metadata were used as a corpus. Following fields were indexed by each IR systems: for MIR; "title", "author", "description", "subject" and "type" DC elements; for HIR, in addition to MIR

DC elements "full-text" field; for FIR, only the "full-text" field. In order to evaluate performance of these IR systems, the following nine questions were selected and "precision - recall" (Baeza-Yates & Riberio-Neto, 1999, p. 75; Kent, Berry, Luehrs & Perry, 1955) and "normalized recall" (Bollmann, 1983; Yao, 1995) tests were conducted.

1) Which documents were written by İrfan Çakın?
2) Which documents cited İrfan Çakın?
3) "Information seeking behavior"
4) AACR, AACR1 or AACR2
5) OPAC or "Online public access catalog"
6) Disabled people or handicapped people (in Turkish *Engelliler veya özürlüler*)
7) Stemmed and singular form of 6th question's terms (in Turkish *Engelli veya özürlü*)
8) "Patrons' educations", "readers' educations" or orientations (in Turkish "Kullanıcılara eğitimler", "okuyuculara eğitimler" veya oryantasyonlar)
9) Semi-stemmed and singular form of 8th question's terms (in Turkish *"Kullanıcı eğitimi", "okuyucu eğitimi" veya oryantasyon*)

The following results were found:

- HIR, FIR and MIR retrieved maximum number of relevant documents, and maximum number of irrelevant documents were retrieved by FIR, HIR and MIR respectively (Table 1). On the other hand, *Kruskal-Wallis H* test showed that there were no significant differences between systems in relevant retrieved documents means, H (2) = 5,116, p = 0,077.

Table 1
Relevant and irrelevant documents for queries

Query	All Relevant Docs.	MIR		FIR		HIR	
		Retrived Relevant	Retrieved Irrelevent	Retrived Relevant	Retrieved Irrelevent	Retrived Relevant	Retrieved Irrelevent
1	15	15	0	15	40	15	0
2	14	0	0	14	45	14	0
3	9	4	0	9	8	9	3
4	16	3	0	13	9	13	1
5	21	10	0	14	14	21	14
6	8	8	1	7	29	8	2
7	8	8	1	7	29	8	2
8	14	2	0	5	5	5	5
9	14	4	0	8	16	8	16
Sum	119	54	2	92	195	101	43

- Mean average precision (MAP) performance of HIR was significantly higher in comparison to MIR, U = 20,5, p = 0,008, Z = -2,635, r = -0,56, and FIR, U = 24, p = 0,017, Z = -2,397, r = -0,51. There were no significant differences between MIR and FIR, U = 54, p = 0,669, Z = -0,428, r = -0,09 (Figure 1). In each information retrieval system, a strong negative correlation was identified between recall and precision, for MIR r(9) = -0,926, p = 0,00, for FIR r(9) = -0,984, p = 0,00 and for HIR r(9) = -0,982, p = 0,00. It was seen that recall increased while precision decreased.

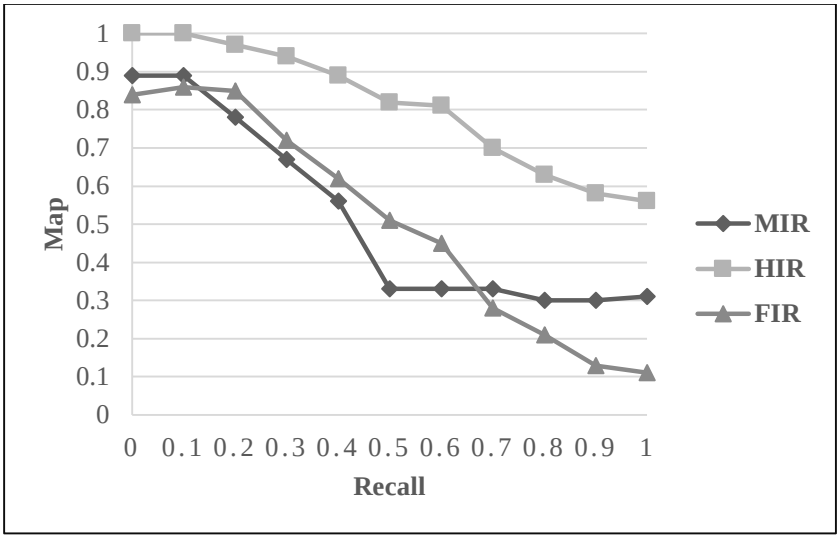


Figure 1. Recall and MAP performances of the IR systems

- Normalized recall performance of MIR, U = 11, p = 0,008, Z = -2,656, r = -0,62 and HIR, U = 6, p = 0,002, Z = -3,051, r = -0,71 was significantly higher in comparison to FIR (Figure 2). There was no significant difference between MIR and HIR, U = 31, p = 0,380, Z = -0,879, r = -0,21.

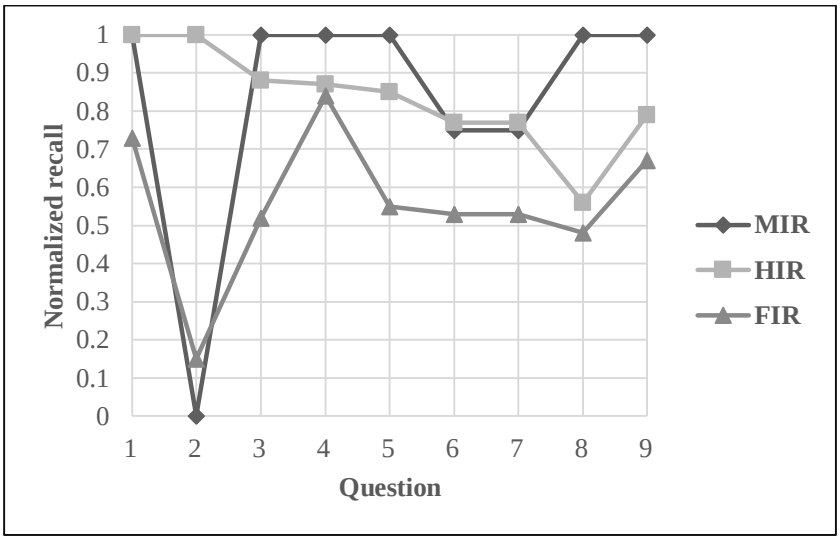


Figure 2. Normalized recall performances of the IR systems

- Processing different types of contents such as metadata and full-text had some advantages and disadvantages for information retrieval systems in terms of term management. The advantages brought together in hybrid content processing and information retrieval performance improved.

Ek 1

Soru 1: İrfan Çakın'ın yazdığı tüm dokümanlar
Formülasyon: İrfan AND Çakın
ÜBES formülasyon çıktısı: + creator:irfan + creator:çakın
TBES formülasyon çıktısı: +fulltext:irfa +fulltext:çak
MBES formülasyon çıktısı: (+ creator:irfan + creator:çakın) OR (+ fulltext:irfa + fulltext:çak)

Soru 2: İrfan Çakın'a bilimsel/hakemli dokümanlarda yapılan atıflar
Formülasyon: "İrfan Çakın" OR "Çakın, İrfan" OR "Çakın, İ."
ÜBES formülasyon çıktısı: Veri olmadığı için formülasyon yapılamamaktadır.
TBES formülasyon çıktısı: fulltext:"irfa çak" OR fulltext:"çak irfa" OR fulltext:"çak i"
MBES formülasyon çıktısı: -(creator:"irfan çakın" OR creator:"çakın, irfan" OR creator:"çakın, i.") + (fulltext:"irfa çak" OR fulltext:"çak irfa" OR fulltext:"çak i") + (type:"Makaleler" OR type:"Hakemli Yazılar")

Soru 3: "Bilgi arama davranışı"
Formülasyon: "bilgi arama davranışı"
ÜBES formülasyon çıktısı: title:"bilgi ara davranış" OR description:"bilgi ara davranış" OR subject:"bilgi ara davranış"
TBES formülasyon çıktısı: fulltext:"bilgi ara davranış"
MBES formülasyon çıktısı: title:"bilgi ara davranış" OR description:"bilgi ara davranış" OR subject:"bilgi ara davranış" OR fulltext:"bilgi ara davranış"

Soru 4: AACR, AACR1 veya AACR2
Formülasyon: AACR OR AACR1 OR AACR2
ÜBES formülasyon çıktısı: title:"aacr" OR title:"aacr1" OR title:"aacr2" OR description:"aacr" OR description:"aacr1" OR description:"aacr2" OR subject:"aacr" OR subject:"aacr1" OR subject:"aacr2"
TBES formülasyon çıktısı: fulltext:"aacr" OR fulltext:"aacr1" OR fulltext:"aacr2"
MBES formülasyon çıktısı: title:"aacr" OR title:"aacr1" OR title:"aacr2" OR description:"aacr" OR description:"aacr1" OR description:"aacr2" OR subject:"aacr" OR subject:"aacr1" OR subject:"aacr2" OR fulltext:"aacr" OR fulltext:"aacr1" OR fulltext:"aacr2"

Soru 5: OPAC veya "çevrimiçi katalog"
Formülasyon: OPAC OR "çevrimiçi katalog"
ÜBES formülasyon çıktısı: title:"opaç" OR title:"çevrimiç katalogu" OR description:"opaç" OR description:"çevrimiç katalogu" OR subject:"opaç" OR subject:"çevrimiç katalogu"
TBES formülasyon çıktısı: fulltext:"opaç" OR fulltext:"çevrimiç katalogu"
MBES formülasyon çıktısı: title:"opaç" OR title:"çevrimiç katalogu" OR description:"opaç" OR description:"çevrimiç katalogu" OR subject:"opaç" OR subject:"çevrimiç katalogu" OR fulltext:"opaç" OR fulltext:"çevrimiç katalogu"

Soru 6: Engelliler veya özürlüler
Formülasyon: engelliler OR özürlüler
ÜBES formülasyon çıktısı: title:"engelli" OR title:"özürlü" OR description:"engelli" OR description:"özürlü" OR subject:"engelli" OR subject:"özürlü"
TBES formülasyon çıktısı: fulltext:"engelli" OR fulltext:"özürlü"
MBES formülasyon çıktısı: title:"engelli" OR title:"özürlü" OR description:"engelli" OR description:"özürlü" OR subject:"engelli" OR subject:"özürlü" OR fulltext:"engelli" OR fulltext:"özürlü"

Soru 7: Engelli veya özürlü
Formülasyon: engelli OR özürlü
ÜBES formülasyon çıktısı: title:"engelli" OR title:"özürlü" OR description:"engelli" OR description:"özürlü" OR subject:"engelli" OR subject:"özürlü"
TBES formülasyon çıktısı: fulltext:"engelli" OR fulltext:"özürlü"
MBES formülasyon çıktısı: title:"engelli" OR title:"özürlü" OR description:"engelli" OR description:"özürlü" OR subject:"engelli" OR subject:"özürlü" OR fulltext:"engelli" OR fulltext:"özürlü"

Soru 8: "Kullanıcılara eğitimler", "okuyuculara eğitimler" veya oryantasyonlar
Formülasyon: "kullanıcılara eğitimler" OR "okuyuculara eğitimler" OR oryantasyonlar
ÜBES formülasyon çıktısı: title:"kullanıcı eğitim" OR title:"okuyucu eğitim" OR title:"oryantasyon" OR description:"kullanıcı eğitim" OR description:"okuyucu eğitim" OR description:"oryantasyon" OR subject:"kullanıcı eğitim" OR subject:"okuyucu eğitim" OR subject:"oryantasyon"
TBES formülasyon çıktısı: fulltext:"kullanıcı eğitim" OR fulltext:"okuyucu eğitim" OR fulltext:"oryantasyon"
MBES formülasyon çıktısı: title:"kullanıcı eğitim" OR title:"okuyucu eğitim" OR title:"oryantasyon" OR description:"kullanıcı eğitim" OR description:"okuyucu eğitim" OR description:"oryantasyon" OR subject:"kullanıcı eğitim" OR subject:"okuyucu eğitim" OR subject:"oryantasyon" OR fulltext:"kullanıcı eğitim" OR fulltext:"okuyucu eğitim" OR fulltext:"oryantasyon"

Soru 9: "Kullanıcı eğitimi", "okuyucu eğitimi" veya oryantasyon.
Formülasyon: "kullanıcı eğitimi" OR "okuyucu eğitimi" OR oryantasyon.
ÜBES formülasyon çıktısı: title:"kullanıç eğit" OR title:"okuyuç eğit" OR title:"oryantasyo" OR description:"kullanıç eğit" OR description:"okuyuç eğit" OR description:"oryantasyo" OR subject:"kullanıç eğit" OR subject:"okuyuç eğit" OR subject:"oryantasyo"
TBES formülasyon çıktısı: fulltext:"kullanıç eğit" OR fulltext:"okuyuç eğit" OR fulltext:"oryantasyo"
MBES formülasyon çıktısı: title:"kullanıç eğit" OR title:"okuyuç eğit" OR title:"oryantasyo" OR description:"kullanıç eğit" OR description:"okuyuç eğit" OR description:"oryantasyo" OR subject:"kullanıç eğit" OR subject:"okuyuç eğit" OR subject:"oryantasyo" OR fulltext:"kullanıç eğit" OR fulltext:"okuyuç eğit" OR fulltext:"oryantasyo"

Ek 2

Tablo 1

TBES'in duyarlılık-anma (D-A) ve R_{norm} performansı

Soru 1		Soru 2		Soru 3		Soru 4		Soru 5		Soru 6		Soru 7		Soru 8		Soru 9	
A	D	A	D	A	D	A	D	A	D	A	D	A	D	A	D	A	D
0,07	1	0,07	0,03	0,11	1	0,06	1	0,05	1	0,13	1	0,13	1	0,08	0,50	0,08	1
0,13	1	0,14	0,06	0,22	1	0,13	1	0,10	1	0,25	1	0,25	1	0,15	0,67	0,15	1
0,20	1	0,21	0,08	0,33	0,75	0,19	1	0,14	1	0,38	1	0,38	1	0,23	0,75	0,23	1
0,27	0,57	0,29	0,11	0,44	0,80	0,25	1	0,19	1	0,50	0,80	0,50	0,80	0,31	0,44	0,31	1
0,33	0,63	0,36	0,12	0,56	0,50	0,31	1	0,24	0,83	0,63	0,63	0,63	0,63	0,39	0,50	0,39	0,46
0,40	0,67	0,43	0,14	0,67	0,55	0,38	1	0,29	0,75	0,75	0,33	0,75	0,33			0,46	0,46
0,47	0,64	0,50	0,16	0,78	0,58	0,44	1	0,33	0,54	0,88	0,19	0,88	0,19			0,54	0,39
0,53	0,42	0,57	0,17	0,89	0,53	0,50	1	0,38	0,50							0,62	0,33
0,60	0,43	0,64	0,18	1	0,53	0,56	1	0,43	0,53								
0,67	0,46	0,71	0,19			0,63	0,77	0,48	0,46								
0,73	0,46	0,79	0,20			0,69	0,58	0,52	0,48								
0,80	0,35	0,86	0,22			0,75	0,60	0,57	0,50								
0,87	0,37	0,93	0,23			0,81	0,42	0,62	0,52								
0,93	0,39	1	0,25					0,67	0,50								
1	0,27																
R_{norm}		R_{norm}		R_{norm}		R_{norm}		R_{norm}		R_{norm}		R_{norm}		R_{norm}		R_{norm}	
0,73		0,15		0,52		0,84		0,55		0,53		0,53		0,48		0,67	

Tablo 2

ÜBES'in duyarlılık-anma (D-A) ve R_{norm} performansı

Soru 1		Soru 2		Soru 3		Soru 4		Soru 5		Soru 6		Soru 7		Soru 8		Soru 9	
A	D	A	D	A	D	A	D	A	D	A	D	A	D	A	D	A	D
0,07	1	0	0	0,11	1	0,06	1	0,05	1	0,13	1	0,13	1	0,08	1	0,08	1
0,13	1	0,1	0	0,22	1	0,13	1	0,09	1	0,25	1	0,25	1	0,15	1	0,15	1
0,20	1	0,2	0	0,33	1	0,19	1	0,14	1	0,38	1	0,38	1			0,23	1
0,27	1	0,3	0	0,44	1			0,18	1	0,50	1	0,50	1			0,31	1
0,33	1	0,4	0					0,23	1	0,63	1	0,63	1				
0,40	1	0,5	0					0,27	1	0,75	1	0,75	1				
0,47	1	0,6	0					0,32	1	0,88	0,88	0,88	0,88				
0,53	1	0,7	0					0,36	1	1,00	0,89	1,00	0,89				
0,60	1	0,8	0					0,41	1								
0,67	1	0,9	0					0,46	1								
0,73	1	1	0														
0,80	1																
0,87	1																
0,93	1																
1,00	1																
R_{norm}		R_{norm}		R_{norm}		R_{norm}		R_{norm}		R_{norm}		R_{norm}		R_{norm}		R_{norm}	
1,00		0,00		1,00		1,00		1,00		0,75		0,75		1,00		1,00	

Tablo 3
MBES'in duyarlılık-anma (D-A) ve R_{norm} performansı

Soru 1		Soru 2		Soru 3		Soru 4		Soru 5		Soru 6		Soru 7		Soru 8		Soru 9	
A	D	A	D	A	D	A	D	A	D	A	D	A	D	A	D	A	D
0,07	1	0,07	1	0,11	1	0,06	1	0,05	1	0,13	1	0,13	1	0,08	1	0,08	1
0,13	1	0,14	1	0,22	1	0,13	1	0,10	1	0,25	1	0,25	1	0,15	1	0,15	1
0,20	1	0,21	1	0,33	1	0,19	1	0,14	1	0,38	1	0,38	1	0,23	0,75	0,23	1
0,27	1	0,29	1	0,44	1	0,25	1	0,19	1	0,50	1	0,50	1	0,31	0,44	0,31	1
0,33	1	0,36	1	0,56	1	0,31	1	0,24	1	0,63	1	0,63	1	0,39	0,50	0,39	1
0,40	1	0,43	1	0,67	1	0,38	1	0,29	1	0,75	1	0,75	1,00			0,46	1
0,47	1	0,50	1	0,78	1	0,44	1	0,33	1	0,88	0,88	0,88	0,88			0,54	0,39
0,53	1	0,57	1	0,89	0,89	0,50	1	0,38	1	1	1	1	0,80			0,62	0,33
0,60	1	0,64	1	1	0,82	0,56	1	0,43	1								
0,67	1	0,71	1			0,63	1	0,48	1								
0,73	1	0,79	1			0,69	0,58	0,52	1								
0,80	1	0,86	1			0,75	0,60	0,57	1								
0,87	1	0,93	1			0,81	0,42	0,62	0,93								
0,93	1	1	1					0,67	0,70								
1	1							0,71	0,65								
								0,76	0,67								
								0,81	0,59								
								0,86	0,60								
								0,91	0,61								
								0,95	0,63								
								1	0,60								
R_{norm}		R_{norm}		R_{norm}		R_{norm}		R_{norm}		R_{norm}		R_{norm}		R_{norm}		R_{norm}	
1,00		1,00		0,88		0,87		0,85		0,77		0,77		0,56		0,79	