

# Konum Önerisi için Zaman Tabanlı Uzman Destekli İş Birliğine Dayalı Filtreleme

## Time Based Expert Supported Collaborative Filtering for Location Recommendation

Başak Melis Öcal (ORCID: 0000-0002-5816-8637)  
Bilkent Üniversitesi Bilgisayar Mühendisliği  
Bölümü  
melisbasakocal@gmail.com

H. Altay Güvenir (ORCID: 0000-0003-2589-316X)  
Bilkent Üniversitesi Bilgisayar Mühendisliği  
Bölümü  
guvenir@cs.bilkent.edu.tr

### Öz

Konuma dayalı sosyal ağlar, son on yılda, kullanıcının konum geçmişlerine dayanarak tercihlerini araştırmamız için bize yeni bir platform sağlayarak önemli ölçüde gelişti. Konuma dayalı sosyal ağların çoğu, kullanıcıların varlıklarını açıklayabilecekleri, yorumlayabilecekleri veya ipucu bırakabilecekleri bir kategori hiyerarşisi altına yerleştirilen çeşitli mekanlar sağlar. Coğrafi bilgili konum önerileri birçok araştırmacının ilgisini çekmesine rağmen, araştırma projelerinin çoğunda zamanın kullanıcının tercihleri üzerindeki etkisi göz ardı edilmiştir. Bir kullanıcı, günün farklı saatlerinde ziyaret etmek için farklı mekanları tercih edebileceğinden, belirli bir kategoride aynı miktarda giriş yapan iki kullanıcı, o mekanda bulunma zamanına bağlı olarak daha az benzer olabilir. Ayrıca, geleneksel iş birliğine dayalı filtreleme teknikleri, tüm kullanıcıların tercihlerini göz önünde bulundururken, yalnızca kategori uzmanlarının tercihlerini göz önünde bulundurarak, o kategorideki bir mekanı önermek, öneri performansını daha da artırabilir. Bu amaçla, mekânları önermek için ölçeklenebilir zamana dayalı yeni bir uzman destekli iş birliğine dayalı filtreleme yaklaşımı önerilmiştir. Bu yeni yaklaşımda öncelikle, belirli bir kategoride yapılan giriş kayıtlarının çeşitliliği göz önünde bulundurularak kategorilere göre kullanıcıların uzmanlığı araştırılır ve her kategori için uzmanların en üst m tanesi seçilir. Daha sonra, her bir kullanıcı-mekan çifti için tercih puanı, günün zaman aralığı ile ilgili olarak

hesaplanır. Giriş kayıt saati bilgilerini farklı aşamalarda dikkate alan üç algoritma geliştirilmiştir. İlk algoritma, önceden tanımlanmış her bir zaman aralığı için kullanıcılar arasındaki benzerlik değerlerini hesaplar. İkinci algoritma, bu belirli bir zaman aralığında kullanıcı-mekan giriş frekans matrisini dikkate alır. Üçüncü algoritma hem benzerlik değerlerinden hem de zaman aralığının kullanıcı-mekan giriş frekans matrisinden faydalanır. Son olarak, en üst-k sıradaki konular kullanıcıya öneri olarak sunulur. Önerilen algoritmalar iki büyük ölçekli Foursquare veri kümesi ile değerlendirilmiş ve temel yaklaşımlarla karşılaştırılmıştır.

**Anahtar Sözcükler:** İşbirliğine Dayalı Filtreleme, Konum bazlı önerici Sistemleri, Sosyal Ağlar

### Abstract

Location based social networks have evolved significantly during the last decade providing us a new platform to investigate user's preferences based on their location histories. Most of the location based social networks provide a variety of venues placed under a hierarchy of categories where users are able to declare their presence, comment or leave tips. Though, location recommendations with geographical information have been a subject of interest to many researchers, the effect of time on user's preferences has been omitted in most of the research projects. As a user may prefer different venues to visit during different hours of a day, two users having the same number of check-ins at a particular category may be less similar depending on the time of their presence. Moreover, while

Gönderme ve kabul tarihi: 31.10.2019 - 19.12.2019

Makale türü: Araştırma

*traditional collaborative filtering techniques consider preferences of all users, considering only preferences of category experts while recommending a venue of that particular category can further improve the recommendation performance. To this end, a novel scalable time-based expert-supported collaborative filtering approach to recommend venues is proposed. In this new approach, first, expertise of users is investigated with respect to categories by considering the variety of check-ins made at that particular category and top-m of the experts are selected for each category. Later, preference score for each user-venue pair is calculated regarding the time interval of the day. Three algorithms are developed that consider check-in time information at different stages. First algorithm calculates similarity values between users for each predefined time interval. Second algorithm considers user-venue check-in frequency matrix at that particular time interval. Third algorithm utilizes both from the similarity values and user-venue check-in frequency matrix of the time interval. Finally, the top-k ranked locations are provided to the user as recommendations. Proposed algorithms are evaluated with two large-scale Foursquare datasets and compared with baseline approaches.*

**Keywords:** Collaborative filtering, Location based recommendation systems, Social Networks

## 1. Giriş

Mobil cihazların ve kablosuz iletişim ağlarının hızla gelişmesiyle insanlar, bireylerin yaşam bilgilerini tüm dünyayla paylaşabildikleri örneğin Foursquare, Loopt gibi konum bazlı sosyal ağların (*Location Based Social Networks: LBSN*) gelişmesine yardım ederek sosyal ağlardaki faaliyetlerine konum bilgisi ekleyebilmektedirler [1]. Kullanıcılar, LBSN'leri kullanarak mevcudiyetlerini belli bir mekanda giriş (check-in) yaparak bildirebilmekte, mekanlar hakkında yorum ve ipuçları bırakabilmektedirler. Kullanıcının konum geçmişini ve kullanıcı profilinin karakterize eden bileşenini analiz ederek, kullanıcının tercihlerine, davranışlarına, sosyal yapısına ve sosyal rolüne dayanarak kapsamlı bilgi edinilebilir. Örneğin, aynı İtalyan restoranını tercih eden kullanıcılar benzer olabilir ve birbirlerinin hoşlandığı diğer mekanları beğenebilirler. Ayrıca, farklı dış mekanlarda giriş yapan kullanıcılar aynı kategoriyi tercih ettikleri için benzer olabilirler.

Bundan farklı olarak, spor salonlarında aynı sayıda giriş yaptıran iki kullanıcıdan, biri günün saatlerine yayararak spor salonuna giriş yapmışken, diğeri 12:00-13:00 veya 18:00'dan sonra spor salonunda giriş işlemlerine sahipse, benzer olmayabilir, çünkü kullanıcıların farklı sosyal rolleri olabilir. Örneğin, ilk kullanıcı bir öğrenci veya emekli olabilirken, ikincisi çalışıyor olabilir ve sadece öğle tatilinde veya işten sonra egzersiz yapabilir. Sosyal rollerde böyle bir farklılaşma, kullanıcıların diğer tercihlerini de etkileyebilir.

Artan konum sayısı ilgi alanlarına göre konumları keşfetme ve ziyaret etme arzusuyla paralel hale geldiğinden konum bazlı öneri sistemleri (*Location Based Recommendation Systems: LBRS*) popüler bir araştırma konusu haline gelmiştir. İşbirliğine dayalı filtrelemeyle ilgili önceki araştırmalar, kullanıcılar arasındaki benzerliği hesaplamak için kullanıcıların mekanlardaki giriş sıklığından veya kategorilerden sıklığından yararlanmaktadır. Daha sonra, her mekanda kullanıcı benzerliği ve giriş sıklığı kullanılarak öneriler yapılmıştır. Bununla birlikte, yaklaşımlardan hiçbiri, kullanıcı profilini daha da karakterize ederek, giriş zamanı ile ilgili tam yerinde, isabetli mekan önerilerinde bulunmaz. İş saat 09.00'da başlayan ve sabah 10.00'da bir kahve dükkanını sık sık ziyaret eden bir kullanıcı, akşam yemeğinden sonra saat 23.00 civarında kahve dükkanında giriş yapan bir kullanıcıdan farklı profilde düşünülmemelidir. Saat 10.30 civarında kahve dükkanı giriş işlemleri yapan üçüncü bir kullanıcının, ilk kullanıcıya ikinci kullanıcıdan daha fazla benzer olması daha olasıdır. Ayrıca, geceleri kahve içmeyi tercih ettiği için kahve dükkanlarını ikinci kullanıcıya 10.00 civarında önermek yetersiz olabilir. Geleneksel iş birliğine dayalı filtreleme yaklaşımları tüm kullanıcıların tercihlerini göz önünde bulundurur. Ancak, kategori uzmanlarının sadece tercihlerini göz önünde bulundurarak o belirli kategorideki bir mekanı önermek, öneri performansını daha da artırabilir. Örneğin, bir kullanıcı aynı restoranda 100 giriş yaparken, ikinci kullanıcı çeşitli restoranlarda aynı sayıda giriş yaptıysa, ikinci kullanıcının gıda kategorisindeki uzmanlığı birinciden daha yüksektir. Dolayısıyla, önerileri etkileyen diğer tercihler aynı ise, ikinci kullanıcıdan tavsiye almak daha doğru olacaktır. Ayrıca, iş birliğine dayalı filtreleme yaklaşımları, kullanıcının sınırlı sayıda yeri fiziksel olarak ziyaret edebileceği için seyrek kullanıcı-konum tercihi

matrisinden muzdariptir [2]. Ziyaret edilmeyen mekanlar, kullanıcının kendilerine karşı olumsuz tercihi olduğu anlamına gelmediğinden, ziyaret edilmeyen mekanları olumsuz olarak değerlendirmek yerine bunlara daha düşük bir güven değeri verilmelidir.

Konum önerisinde bahsedilen sorunları çözmeyi amaçlayan bu makalede, belirli bir kullanıcıya, kullanıcının konum geçmişi ve belirlenen tercihlerine bağlı olarak günün zaman aralığı ile ilgili bir dizi mekan öneren, ölçeklenebilir zaman bazlı uzman destekli iş birliğine dayalı filtreleme yaklaşımı (*Time-Based Expert supported Collaborative Filtering: TECF*) önerilmiştir. Öncelikle, uzmanlık, çok sayıda farklı mekan ziyareti gerektiren belirli bir kategorideki mekan tecrübesine ve kapsamlı bilgisine sahip olmak olarak tanımlanmıştır. Bu nedenle, makalede önerilen öneri sistemi, her bir kategorideki uzmanları, belirli bir kategorideki farklı mekanlarda yapılan giriş sayısına göre aday uzmanları sıralayarak çıkarır ve aday uzmanların en üst  $m$  tanesini seçer. Seçilmiş uzmanlar, çeşitli yerleri ziyaret ettikleri ve tercihleri daha değerli olduğu için doğru öneriler sunabilirler. Daha sonra, her bir kullanıcı-mekan çifti için tercih skoru, geçerli zaman aralığı için hesaplanır ve en üst  $k$  sıralarında yer alan mekanlar, tavsiye olarak kullanıcıya iade edilir. Tercih puanlarını hesaplamayı hedefleyen önerilen öneri sistemi üç ana bileşenden oluşur: 1) Kullanıcı profilini daha ayrıntılı ifade etmek için giriş zamanı, 2) farklı kategorilerde kullanıcının giriş sıklığı ve 3) farklı mekanlarda giriş yapma sıklığı. Giriş zamanı bilgisinin kullanımına göre üç algoritma geliştirilmiştir. İlk algoritma, önceden tanımlanmış her bir zaman aralığı için kullanıcılar ve uzmanlar arasındaki benzerlik değerlerini hesaplar. Bu belirli aralıktaki kullanıcı-kategorisi giriş sıklığı matrisi, ana kullanıcı kategorisinde giriş sıklığı matrisinden ziyade benzerlik değerleri elde etmek için kullanılır. İkinci algoritma benzerliği hesaplamak için ana kullanıcı-kategorisi giriş sıklığı matrisini kullanır, ancak her bir uzmanın bu aralıktaki kullanıcı-mekan skoruna katkısını hesaplar, belirli bir zaman aralığında kullanıcı-mekan giriş sıklığı matrisinden yararlanır. Üçüncü algoritma, o zamanki kullanıcı-mekan skorlarını hesaplamak için, o kullanıcı-mekan giriş sıklığından ve kullanıcı-mekan giriş sıklığı matrislerinden, söz konusu aralıktaki kullanıcı-mekan skorlarını hesaplamak için yararlanır.

Bu makalede, zaman bilgisinin kullanımına göre geliştirilen üç algoritma ve zaman bileşeni olmayan yöntemimiz, yaklaşık 10 ay boyunca toplanan iki Foursquare giriş veri kümesi ile değerlendirilmiştir. Önceden işlenmiş Tokyo veri kümesi, 203.575 giriş işlemiyle 2291 aktif kullanıcıyı içermektedir ve 35 kategoriye sahiptir. İkinci veri kümesi, 1083 aktif kullanıcı, 111.973 giriş ve 35 kategoriden oluşan New York girişlerinden oluşmaktadır. Algoritmaların performansı ve sonuçların doğruluğu birbirleriyle ve temel yaklaşımlarla karşılaştırılmıştır.

Spesifik olarak, önerilen yaklaşımın mevcut çalışmaya katkıları şöyle özetlenebilir:

- Önerilen yaklaşım, mekan önerilerinin doğruluğunu artıran kullanıcıların genel kategori giriş sıklıklarından faydalanan her bir kategorideki uzmanları önceden hesaplar.
- Önerilen yaklaşım, zaman bilgisinin kullanımına göre üç modelden faydalanan uzmanların benzerlik ve mekan giriş sıklığını dikkate alarak, belirli bir zaman aralığında bir kullanıcı mekan çiftine tercih skorunu verir. Ardından, bu zaman aralığının önerileri olarak kullanıcıya mekanların en üst  $k$  tanesini sunar. Bu yöntem kullanıcı profilini daha ayrıntılı karakterize etmeye olanak sağlar.
- Zaman bileşeni olmayan yöntemimize ek olarak önerilen üç model, Tokyo ve New York'tan yapılan giriş işlemleri dahil olmak üzere Foursquare'den toplanan iki gerçek veri kümesi ile değerlendirildi. İlk 2291 kullanıcı, 203.575 giriş ve 35 mekan kategorisi, ikincisi ise 1083 kullanıcı, 111.973 giriş ve 35 kategori içermektedir.

Bu makalenin geri kalanı şu şekilde düzenlenmiştir: İlk önce ilgili çalışmaları ele alıyoruz. Daha sonra, uzman destekli iş birliğine dayalı filtreleme yaklaşımı sunulmuştur. Ardından önerilen zaman bazlı uzman destekli iş birliğine dayalı filtreleme algoritmalarını ayrıntılı olarak açıklanmıştır. Deneysel bölümü, bazı tartışmalarla iki büyük ölçekli Foursquare veri setine dayanan kapsamlı deneysel sonuçları sağlamaktadır. Son olarak, gelecek için planlanan çalışmalar özetlenmiştir.

## 2. İlgili Çalışmalar

Konum önerisi üzerine önceki araştırmalar iki kategoride özetlenebilir: 1) genel konum önerileri 2) kişiselleştirilmiş konum önerileri.

### 2.1 Genel Konum Önerileri

Kullanıcının konum geçmişinden çıkardığı tercihlerden bağımsız olarak, genel konum önerme sistemleri, kullanıcılara, belirli bir mekansal bölgede kamuoyuna göre en popüler mekanları sunar. Örneğin, ağaç tabanlı bir hiyerarşik çizge, çok sayıda GPS yürüncesinden ilginç konumlar ve seyahat dizileri çıkartarak kişilerin konum geçmişini temsil etmek için kullanılır [3]. Daha sonra, kullanıcı-konum matrisinden elde edilen çıkarım modeli, kullanıcının deneyimleri, kullanıcı deneyimi ile konum ilgisi arasındaki karşılıklı pekiştirme ilişkisi ile ilgili bir konumun ilgisini çıkarmak için önerilir. Başka bir yaklaşımda, anlam çoğaltan bir çizge üzerindeki rastgele gezintileri kullanarak konumlara anlam verilir [4]. Kullanıcı tercihlerini dikkate almamak, kullanıcılara isabetli teklifleri sunmayı engelleyen tüm kullanıcılara aynı mekanları sunma sonucunu doğurur.

### 2.2 Kişiselleştirilmiş Konum

#### Önerileri

Bireyin mekan tercihlerini göz önünde bulundurarak, kişiselleştirilmiş konum öneri sistemleri, kullanıcılara en isabetli mekan tekliflerini sunmayı amaçlamaktadır. Kişiselleştirilmiş öneriler sağlamak için takip edilen birkaç yaklaşım vardır. Spesifik olarak, basit öneri sistemleri bir kullanıcının tercihlerini manuel olarak ekleyerek bir profil sağlamasını ister [5]. Daha sonra, kullanıcı tarafından seçilen kategoriler ve mevcut konum bilgileri, mekanları önermek için kullanılır. Kategori tercihleri ikili değerlerden ziyade ağırlıklandırıldığından, bu yaklaşım, seçilen her kategorinin eşit olarak gerçek yaşam tercihlerini yansıtmaması ile sonuçlanmaktadır. Bayes Ağı kullanımı yaklaşımında, bireylerin tercihlerini modelleyen harita tabanlı bir önerici sistem tanıtılmıştır [6]. Daha önce ziyaret edilmemiş mekanlar için bir önerici sistem kurma sorunu da sunulmuştur [7]. Yazarlar, yönlendirilmemiş bir kullanıcı-mekan grafiği üzerinde gezintinin kararlı durum olasılıklarını hesaplamak için yeniden

başlatmalı rastgele gezintiler önermektedir. Çizgenin ağırlıklı ve yönlendirilmiş bir versiyonunu sunarak bu yaklaşımı daha da genişletirler. Matris faktörizasyonu performansından dolayı oldukça tercih edilen bir yaklaşımdır [8, 9, 10]. Kullanıcıların tercihlerine göre ilgi alanlarının ilginçliğini tahmin etmek amacıyla normalize edilmiş giriş alanındaki regresyon şeması bazlı bir öngösterge de tanıtılmıştır [11]. Konum önerisinden farklı olarak, Derin Etkinlik Katılım Tahmini çerçevesinde, belirli kişilerin bir etkinliğe birlikte katılıp katılmayacağını öngörmeyi amaçlayan, sıralı, bağlamsal, özel, üç boyutlu modelleme tercihleri önerilmektedir [12]. Sıralı tercihler farklı konulara olan ilgiyi temsil ederken, bağlamsal tercihler olayın yeri ve saati gibi mekansal-zamansal özellikleri modellemektedir. Özel tercih boyutu, olaylar arasındaki etkileri belirtmek için tanımlanır.

Türleriyle birlikte işbirliğine dayalı filtreleme, kullanıcı-mekan puanlarını tahmin etmek için çoğunlukla kullanılan başka bir yaklaşımdır. Yaklaşımın bu kategoriye girdiğinden, iş birliğine dayalı filtreleme tekniklerini kullanan önceki çalışmalar kısaca incelenecektir. Bir kullanıcı-konum matrisi ile kullanıcıların konum geçmişini yansıtmak, konumun bazı yerlerde mekan ve diğerlerinde kategori olduğu durumlarda önerilmektedir [13, 14, 15]. Her satır-sütun girişi, belirli bir kullanıcının o kategoriye veya mekanına ziyaret sıklığını gösterir. Daha sonra, kullanıcı-konum matrisindeki giriş vektörlerine dayanarak kullanıcılar arasındaki benzerlikleri hesaplamak için kosinüs benzerliği kullanılır. Son olarak, kullanıcı-mekan skorları, hedef-mekandaki kullanıcı-kullanıcı benzerliği ve kullanıcıların giriş sıklıkları kullanılarak tahmin edilir. Bu temel iş birliğine dayalı filtreleme yaklaşımları, seyrek kullanıcı-konum matrisini ele almak ve yeni bir konumdaki mekanları tavsiye etmek konusunda yetersiz kalmaktadır. Kullanıcı-mekan skorlarını daha iyi tahmin etmek için iki aşamalı bir öneri sistemi önerilmiştir [2]. Çevrimdışı kısım, her şehir için kullanıcı kategorisinde giriş sıklığı matrisleri hazırlayarak, en fazla giriş sayısına sahip kullanıcıları seçerek, ağırlıklı kategori hiyerarşisine sahip kullanıcıların tercihlerini modelleyerek sosyal bilgi öğrenimi ve tercihlerin keşfedilmesinden sorumludur. Her şehir için giriş matrislerine sahip olmak, daha önce ziyaret edilen şehirlerdeki giriş geçmişinden bir kullanıcıya ziyaret edilmemiş bir

şehirdeki mekanların tavsiye edilmesini sağlar. Daha sonra, her şehirdeki yerel uzmanlar merkez-otorite skorları ile gösterilir ve çevrimiçi adımda iş birliğine dayalı filtreleme uygulanır.

Ne yazık ki, iş birliğine dayalı filtreleme kullanan yaklaşımlarda mevcut çalışmalar, kullanıcı-konum matrisinin seyrekliğini ele almak ve kullanıcı profilini daha ayrıntılı karakterize etmek için yarıdır. Kullanıcı tercihleri bir günün zaman aralığına göre şekillenebildiğinden, yalnızca genel giriş sıklıklarını kullanmak, önerilerin kalitesini sınırlar. Bu çalışmada önerilen yaklaşım, yukarıda belirtilen çalışmalarda tartışılan sorunları ele alan önemli özelliklere sahiptir:

1) Giriş zamanını kullanıcının tercihlerine göre farklı bir faktör olarak değerlendirerek, zaman bilgisini farklı adımlarda kullanan üç algoritma geliştirilmiştir. Bu yaklaşımla, giriş zamanının kullanıcılar ve mekan tercihleri arasındaki benzerlik üzerindeki etkisi daha iyi gözlenmektedir.

2) Daha önceki iş birliğine dayalı filtreleme yöntemleri, her kullanıcıyı, kullanıcı-mekan skor tahminine dahil etmektedir. O belirli kategorideki bir mekanı önerirken yalnızca kategori uzmanlarının tercihlerini göz önünde bulundurmanın, öneri performansını daha da artırabileceği iddia edilmektedir. Dolayısıyla, önerilen öneri sistemi ilk önce kategori uzmanlarını çıkarmaktadır. Daha sonra, mekan skorlarını tahmin etmek için bunların benzerlik değerleri ile orantılı olarak her bir uzmanın mekan giriş sıklığını dahil etmektedir.

3) Seyrek matris problemi ile başa çıkmak için onları olumsuz olarak değerlendirmek yerine, görülmemiş yerlere daha düşük güven skorları verilmektedir.

### 3. Uzman Destekli İşbirliğine Dayalı Filtreleme

Zamanı kullanıcı profilinin karakteristik bir bileşeni olarak görmek, öneri kalitesini iyileştirir, çünkü kullanıcının tercihlerine, davranışlarına, sosyal yapısına ve rollerine dayalı kapsamlı bilgi, konum geçmişi zamana göre analiz edildiğinde elde edilebilir. Zamanın kullanıcının tercihleri üzerindeki etkisini değerlendirmek amacıyla, ilk olarak, giriş zamanı bilgisini reddeden Uzman Destekli İşbirliğine Dayalı Filtreleme (*Expert-supported Collaborative Filtering, ECF*) algoritması

geliştirilmiştir. ECF hala uzman seçimi ve ziyaret edilmeyen yerlere daha az güven vermeyi içermektedir.

Bu bölümde sunulan ECF algoritması şu bileşenlere sahiptir:

- 1) kullanıcı kategorisini ve kullanıcı-mekan matrislerini oluşturan ve daha sonra her bir kategorinin en üst  $m$  uzmanlarını seçen *Giriş analizi ve uzman seçimi*, ve
- 2) kullanıcı-mekan skorlarını tahmin eden *Benzerlik hesaplama ve mekan skoru tahmini*.

#### 3.1 Giriş Analizi ve Uzman Seçimi

Önerilen yöntemde, her kullanıcının konum geçmişini kategorilere ve mekanlara ayırılır. Girişler her veri kümesinde tek bir şehirden toplandığından, kullanıcıların konum geçmişlerini yansıtmak için yalnızca bir ana kullanıcı kategorisi matrisi yeterlidir. Tokyo veri seti kullanıldığında, matris her satırın bir Foursquare kullanıcıyı temsil ettiği ve sütunun bir kategoriye temsil ettiği 2291 satır ve 35 sütun içerir. New York veri kümesi kullanıldığında, aynı yaklaşımla 1083 satır ve 35 sütunla matris oluşturulur. Ana kullanıcı-kategori matrisindeki her değer, belirli bir kullanıcının belirli bir kategorideki giriş sıklığını temsil eder. Benzer şekilde, yalnızca bir ana kullanıcı-mekan matrisi tanımlanmıştır. Yine matris, her kullanıcı için bir satır ve her kategori için bir sütun içerir. Matrisin her girişi, kullanıcının o belirli kategorideki giriş yaptığı mekanları ve giriş sıklıklarını kaydeden bir listedir.

Matrisler doldurulduktan sonra, her kategori için uzman seçim aşaması başlar. Her bir kategori için kullanıcıları ve bunlara ait giriş sayısını, o kategorinin farklı mekanlarında saklamak için geçici bir liste tanımlanır. Daha sonra, giriş-in sayılarına göre liste sıralanır. Son adım olarak, en üst- $m$  uzmanlar sıralanan listelerden çıkarılır ve kategori-uzman tablosunda saklanır.

## 3.2 Benzerlik Hesaplama ve Mekan

### Skoru Tahmini

İki kullanıcı arasındaki benzerlik, ana kullanıcı-kategori matrisinden elde edilen iki kullanıcının giriş vektörleri arasındaki kosinüs benzerliği ile gösterilmektedir. Her bir kullanıcı-kullanıcı çifti için benzerlik hesaplaması yapılır. Elde edilen değerler benzerlik matrisinde saklanır.

$v_u =$  kullanıcı  $u_u$ 'nun giriş vektörü

$v_v =$  kullanıcı  $u_v$ 'nin giriş vektörü

$ben(u_u, u_v)$   
= kullanıcı  $u_u$  ve  $u_v$  arasındaki benzerlik

$$ben(u_u, u_v) = \frac{v_u \cdot v_v}{|v_u| |v_v|} \quad (1)$$

Uzmanlar ve kullanıcı-mekan skorları öngörülen kullanıcı arasındaki benzerlik, birkaç deneyden sonra 0,6'ya ayarlanan benzerlik eşik değeri ile karşılaştırılmıştır. Benzerlik değeri benzerlik eşiklerinin altındaysa, kullanıcı-mekan skor tahminine uzman dahil edilmez. Daha sonra benzerlik eşiklerini karşılayan uzmanların ziyaret ettiği mekanlar, ana kullanıcı-mekan matrisinden toplanır. Bir uzman tarafından ziyaret edilmeyen mekanlar için 0 atama, önyargılı önerilere neden olur, zira uzman mekanı tercih etmemektense tanımlıyor olabilir. Bu konuyu ele almanın basit bir yolu, her ziyaret edilmemiş yer için [16]'da önerilen şekilde daha düşük güven değerleri belirlemektir. Denk. (2) ana kullanıcı-mekan matrisinin nasıl değiştirildiğini göstermektedir.

$u_u =$  kullanıcı,  $p_i =$  mekan

$c(u_u, p_i) = p_i$ 'de  $u_u$ 'nun gerçek giriş sayısı

$$c_{yeni}(u_u, p_i) = \begin{cases} c(u_u, p_i) + 1, & \text{eğer } c(u_u, p_i) > 0 \\ 1, & \text{değilse} \end{cases} \quad (2)$$

Daha sonra, her bir uzmanın mekanı beğenme yüzdesi, Denk. (3) tarafından temsil edildiği şekilde hesaplanır.

$$beğenme(e_m, p_i) = \frac{c_{yeni}(e_m, p_i) \times 100}{e_m \text{'nin toplam giriş sayısı} + \text{toplam mekan sayısı}} \quad (3)$$

Daha sonra, her uzmanın katkısı kullanıcı-mekan skoruna eklenir. Kullanıcı-mekan skorunu tahmin

eden ana iş birliğine dayalı filtreleme formülü Denk. (4)'te verilmiştir.

$u_u =$  öneri sahibi

$p_i = u$  için skor hesaplanan mekan

$e_m =$  mekanın kategorisinin uzmanı

$s(u_u, p_i) =$  hesaplanan kullanıcı – mekan skoru

$$s(u_u, p_i) = \frac{\sum ben(e_m, u_u) \times beğeni(e_m, p_i)}{\sum ben(e_m, u_u)} \quad (4)$$

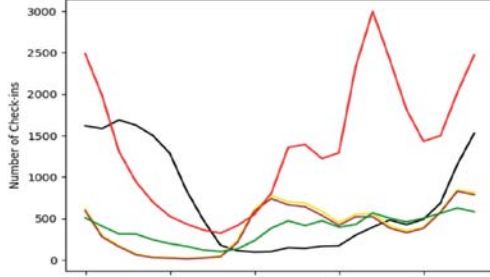
Her mekan için kullanıcı-mekan skoru tahmini tamamlandığında, mekanlar skorlarına göre sıralanır. Son olarak, en üst- $k$  sırasında mekanlar, tavsiye olarak kullanıcıya iade edilir. ECF için algoritma Kod Listesi-1'de verilmiştir.

## 4. Zaman Bazlı Uzman Destekli

### İşbirliğine Dayalı Filtreleme

ECF'den farklı olarak, girişlerin zaman bilgileri, Zaman Bazlı Uzman Destekli İşbirliğine Dayalı Filtreleme yaklaşımında (*Time-based Expert-supported Collaborative Filtering, TECF*) göz önünde bulundurulur. Günün 24 saatini birer saatlik aralıklara bölmek, konum geçmişinin ve seyrekliğin aşırı bölünmesine neden olabileceğinden, beş aralık tanımlanır. Aralıklara Foursquare kategorilerini gruplandırarak ve gruplandırılmış kategorilerin giriş düzenlerini 24 saatlik aralıklarla gözlemleyerek karar verilir. Foursquare kategorilerini birleştirerek elde edilen beş kategori grubu Çizelge-1'de, grupların giriş dağılımları Şekil-1'de verilmiştir. Tanımlanan zaman aralıkları şöyledir: a) 06.00 – 10.59, b) 11.00 – 14.59, c) 15.00 – 18.59, d) 19.00 – 22.59, e) 23.00 – 05.59.

Performansı en iyi algoritmayı bulmayı hedefleyen, zaman bilgisini farklı adımlarda kullanan üç algoritma geliştirilmiştir. İlk algoritma, önceden tanımlanmış her bir zaman aralığı için kullanıcılar ve uzmanlar arasındaki benzerlik değerlerini hesaplayarak beş benzerlik matrisi oluşturur. Bu belirli aralıktaki kullanıcı-kategori matrisi, zamana bakılmaksızın girişleri tahmin eden ana kullanıcı-kategori matrisi yerine benzerlik değerlerini bulmak için kullanılır. İkinci algoritma, benzerliği



**Şekil-1:** Beş kategori grubunun gözlemlenen giriş modelleri. Temsil ettikleri renkler ve kategoriler şöyledir: 1) Siyah – Gece Hayatı Noktaları 2) Kırmızı – Aristokrat Yemek Düzeni Restoranları 3) Sarı – Sabah Noktaları 4) Kahverengi – Kahve Saati Noktaları 5) Yeşil – Doğal Manzara Yerleri.

hesaplamak için ana kullanıcı-kategori matrisinden faydalanır, ancak her bir uzmanın bu aralıktaki kullanıcı-mekan skoruna katkısını hesaplarken, bu belirli zaman aralığındaki kullanıcı-mekan matrisini kullanır. Üçüncü algoritma, o zamandaki kullanıcı-mekan skorlarını hesaplamak için bu belirli aralıktaki hem kullanıcı-kategori hem de kullanıcı-mekan matrislerini kullanır. Bu bölümde TECF algoritmaları arka arkaya sunulacaktır.

```

uzman  $e_m$  'sinin toplam giriş sayısı =  $ToplamE_m$ 
mekan  $p_i$  'in toplam giriş sayısı =  $ToplamP_i$ 
1 kullanıcı_listesindeki her bir  $u_u$  için:
2 kategori_listesindeki her bir kategori için:
3 mekan_listesindeki her bir  $p_i$  için:
4 toplam = 0, benzerlik = 0
5 her bir kategori  $e_m$  'si için:
6 Eğer  $ben(e_m, u_u) >= 0,6$  ise:
7  $c(e_m, p_i) = 0$  ise:
8  $c_{yeni}(e_m, p_i) = 1$ 
9 Değil ise:
10  $c_{yeni}(e_m, p_i) = c(e_m, p_i) + 1$ 
11  $beğenme = (c_{yeni}(e_m, p_i) * 100) /$ 
 $(ToplamE_m + ToplamP_i)$ 
12  $benzerlik = (sim(e_m, u_u) * beğeni) +$ 
 $benzerlik$ 
13 toplam =  $ben(e_m, u_u) + toplam$ 
14 Eğer toplam != 0 ise:
15  $s(u_u, p_i) = benzerlik / toplam$ 
16 Değil ise:
17  $s(u_u, p_i) = 0$ 
18 öneri_listesi.ek(( $p_i$ , skor))

```

#### Kod Listesi-1: ECF Algoritması

**Çizelge-1:** İlgili Foursquare kategorilerini birleştirerek elde edilen beş kategori grubu.

Sabah Noktaları	Kahve Saati Noktaları	Aristokrat Düzeni Restoranları	Yemek
Kahvaltı Noktası		Asya Restoranı	
Spor Salonu & Doğal Manzara Yerleri	Kahve Dükkanları	İngiliz Restoranı	
		Meksika Restoranı	
		Orta Doğu Restoranı	
		Türk Restoranı	
		Avrupa Restoranı	
		Hint Restoranı	
		Akdeniz Restoranı	
		Deniz Ürünleri Restoranı	
		Amerika Restoranı	
Gece Hayatı Noktaları	Doğal Manzara Yerleri		
Bar	Plaj		
Şarap/Yiyecek Barı	Doğal Manzaralar		
Gece Hayatı Noktası	Liman & Marina		
	Dış Mekanlar & Eğlence		

#### 4.1 TECF Algoritması 1

ECF algoritmasının giriş analizi adımına benzer şekilde, ana kullanıcı-kategori ve ana kullanıcı-mekan matrisleri ve kategori-mekan tablosu kendi içerikleriyle oluşturulmaktadır. ECF'de tanımlanan yapıardan farklı olarak, zaman bazlı bir kullanıcı-kategori listesi tanımlanır. Listedeki her giriş, giriş sıklıklarını önceden tanımlanmış zaman aralıklarından birinde yansıtan bir kullanıcı-kategori matrisi saklar. Ana kullanıcı-kategori matrisini kullanarak uzman seçimi ECF yaklaşımında olduğu gibi yapılır.

ECF yaklaşımından farklı olarak, zaman bazlı kullanıcı-kategori listesi kullanılarak benzerlik hesaplaması yapılır. Bir ila beş arasındaki her zaman aralığı için, listedeki ilgili kullanıcı-kategori matrisi dikkate alınır. Son olarak, aralıkların her biri için beş benzerlik matrisi oluşturulur. Denk. (5) her aralığın benzerlik değerlerinin nasıl hesaplandığını göstermektedir.

$v_{w,j}$  = zaman aralığı  $j$ 'deki kullanıcı  $u_u$ 'nin giriş vektörü

$v_{v,j}$  = zaman aralığı  $j$ 'deki kullanıcı  $u_v$ 'nin giriş vektörü  
 $ben(u_u, u_v)$  = zaman aralığı  $j$ 'de  $u_u$  ve  $u_v$  arasındaki benzerlik

$$ben(u_u, u_v, i) = \frac{v_{ui} \cdot v_{vi}}{|v_{ui}|x |v_{vi}|} \quad (5)$$

Algoritma 1'in mekan skor tahmini ECF yaklaşımına benzer. Her aralık farklı öneriler gerektirdiğinden, Kod Listesi-1'de verilen algoritmaya bir dış döngü eklenir. Algoritmanın ilk satırından sonra, algoritmanın geri kalanını içeren tek özyinelemesi bir aralığı temsil eden bir döngü yerleştirilir. Yeni aralık döngüsünün her bir özyinelemesinde, bu aralığın benzerlik matrisi dikkate alınır. Benzerlik eşik değeri olarak 0,6 burada da kullanılmaktadır.

#### 4.2 TECF Algoritması 2

Ana kullanıcı-kategori matrisi ve kategori-mekan tablosu, Algoritma 1 bölümünde sunulan algoritma ile aynı şekilde üretilir. Bununla birlikte, TECF Algoritması 1'in aksine kullanıcı-mekan matrisi zaman aralığı olan başka bir boyuta sahiptir. Her bir matris girişi, her bir kullanıcının, o aralıkta giriş yaptığı mekanları saklayan bir zaman aralığını temsil eden beş listenin bir listesidir. Kullanıcı-mekan matrisinin biçimi Denk. (6)'da verilmiştir.

$kullanici\_mekan\_matris[kullanici][kategori][zamanaraligi][mekan]$  (6)

Ana kullanıcı kategorisi matrisi kullanılarak uzman seçimi ECF'deki gibi yapılır. TECF Algoritması 2, benzerlik matrisi oluşturmak için ECF ile aynı yöntemi izler ve hesaplamalarında ana kullanıcı-kategori matrisini dikkate alır. Zamana göre öneriler üretilmesi sayesinde, her bir aralığı temsil eden ekstra bir dış döngü, TECF Algoritmasında olduğu gibi ECF algoritmasının ilk satırından sonra eklenir. Döngünün her bir özyinelemesinde, bu zaman aralığındaki mekanların giriş sıklıkları tahmine dahil edilir. Benzerlik eşik değeri burada da uygulanmaktadır.

#### 4.3 TECF Algoritması 3

Algoritma 3, yukarıda açıklanan iki algoritmanın melez bir versiyonudur. Daha önceki zaman bazlı yaklaşımlarda açıklandığı gibi, ana kullanıcı-kategori matrisi ve kategori-mekan tablosu bunların içerikleriyle doldurulur. Zaman bazlı kullanıcı-kategori listesi TECF Algoritması 1'deki gibi üretilirken kullanıcı-mekan matrisi TECF Algoritması 2'deki gibi üretilir. Benzerlik matrisini doldururken, zamana göre kullanıcı-kategori listesi kullanılır. Belirli bir aralık için listedeki ilgili kullanıcı-kategori matrisi dikkate alınır. Benzerlik hesaplama işlevi, her aralık için bir tane olmak üzere beş benzerlik matrisinin bir listesini ifade eder. Diğer iki zaman tabanlı algoritmalarda olduğu gibi, her aralık için bir kez yinelenen bir dış döngü kullanılır.

Döngünün her yinelenmesinde, aralığın benzerlik matrisine ek olarak, o zaman aralığındaki mekanların giriş sıklıkları göz önünde bulundurulur. Benzerlik eşik değeri hala uygulanmaktadır.

#### 4.4 Zaman Karmaşıklık Analizi

ECF algoritması her iterasyonda farklı bir kullanıcı için kullanıcı-mekan skoru tahmin etmektedir. Bu işlem yapılırken kullanıcı-kategori matrisindeki her bir kullanıcının, farklı kategorilerdeki mekanlara giriş sayılarına ulaşılması gerekmektedir. Ayrıca uzmanların, uzman oldukları kategorilerdeki mekanlara giriş sayıları benzerlik hesaplaması için kullanılmaktadır. Bu durumda  $u$  kullanıcı sayısı,  $c$  kategori sayısı,  $p$  mekan sayısı ve  $e$  uzman sayısı olmak üzere; ECF algoritmasının zaman karmaşıklığı  $O(ucp \cdot e)$  olarak ifade edilir. Uzman sayısı önceden belirlenen sabit bir değer olduğundan, karmaşıklık  $O(ucp)$ 'dir. TECF algoritmaları kullanıcı-mekan skoru tahminini her zaman aralığı için yapmaktadır. Bu durumda,  $i$  zaman aralığı olmak üzere TECF algoritmalarının zaman karmaşıklığı  $O(ucp \cdot i)$ 'dir. Zaman aralığı sayısı sabit olduğundan bu karmaşıklık da  $O(ucp)$  olarak ifade edilir.

### 5. Deneyler

Bu bölümde, ilk deney ayarları ayrıntılı olarak açıklanmış, daha sonra, deneylerin sonuçları tartışılmaktadır.

#### 5.1 Deney Ayarları

Bu bölümde, deney ayarları sunulmuş ve veri setleri, temel yaklaşım ve değerlendirme ölçütleri hakkında kapsamlı bilgi verilmiştir.

**Veri seti** Önerilen algoritmalar, [17, 18]'in yazarları tarafından yaklaşık 10 ay boyunca toplanan iki büyük ölçekli Foursquare veri kümesi ile değerlendirilmiştir. Foursquare'in gizlilik politikası nedeniyle, kişisel giriş bilgilerine erişilememektedir. Öte yandan, bazı kullanıcılar girişlerini Twitter'da erişilebilen kamuya açık bir gönderi olarak paylaşmayı tercih etmektedir. Bu nedenle, girişler Twitter Streaming API'si aracılığıyla Foursquare etiketli tweet'leri tarayarak yakalanır. Toplanan veri kümeleri gürültülü ve ani hareket girişlerinden filtrelendir [17, 18]. İlk olarak, ilk veri kümesi 2291



kullanıcı tarafından yapılan Tokyo'dan gelen 573.303 adet girişten oluşurken, ikinci veri kümesi 1083 kullanıcı tarafından gönderilen New York'tan gelen 227.428 adet girişten oluşmaktadır. Veri kümeleri, hiyerarşik olarak düzenlenmiş ve 900'den fazla olan gerçek Foursquare kategorilerini içerir. Gerçek Foursquare kategorilerinden oluşturulan bir kategori giriş sıklığı matrisi çok seyrek olacağından, ilgili kategorileri birleştirdik ve gereksiz olanları ortadan kaldırarak kategori sayısını 35'e indirdik. Ayrıca, toplam giriş sayısı Tokyo'da 203.575'e, New York veri kümesinde ise 111.973'e inmiştir. Kullanılan nihai kategoriler Çizelge-2'de verilmiştir.

**Çizelge-2:** Önerilen yaklaşımda kullanılan nihai kategorilerin listesi.

Sanat & Eğlence	Akdeniz Restoranı
Sinema Salonu	Meksika Restoranı
Müze	Orta Doğu Restoranı
Müzik Mekanı	Deniz Ürünleri Restoranı
Sahne Sanatları Mekanı	Vejetaryen Restoranı
Kamu Sanatı	Gece Hayatı Noktası
Yemek	Bar
Amerika Restoranı	Dış Mekan & Dinlenme
Asya Restoranı	Atletizm & Spor
Kahvaltı Noktası	Spor Salonu & Fitness Merkezi
Kafe	Plaj
Karayıpler Restoranı	Doğal Manzara Yerleri
Kahve Dükkanı	Liman & Marina
Tatlı Dükkanı	Kütüphane
Hazır Yemek Restoranı	Ruhani Merkez
Şarap/Yiyecek Barı	Alışveriş Merkezi
Hint Restoranı	Şaraphane
Avrupa Restoranı	

Temel İşbirliğine Dayalı Filtreleme Yaklaşımı Algoritmalarımızı, giriş zamanı bilgilerini ve uzman bilgisini göz ardı eden temel iş birliğine dayalı filtreleme algoritması ile karşılaştırdık. Temel

algoritma (base-line), öncelikle ana kullanıcı-kategori ve ana kullanıcı-mekan matrisleri oluşturularak kullanıcıların konum geçmişlerini mekanlara ve kategorilere ayırır. Daha sonra, Kosinüs benzerliği ile kullanıcılar arasındaki benzerliği temsil eden kullanıcı-kullanıcı benzerliği matrisini oluşturur. Son olarak, kullanıcı bazlı işbirlikçi bir filtreleme modeli uygular ve kullanıcı-mekan skorlarını tahmin eder. TECF algoritmalarımızdan farklı olarak, temel yaklaşımı, belirli bir zaman aralığında benzerlik veya kullanıcı-mekan giriş sıklığı değerlerinden faydalanmamaktadır. Her bir yaklaşım arasındaki farkları daha iyi göstermek için, algoritmaları karşılaştıran Çizelge-3 verilmiştir.

**Değerlendirme Metriği** Değerlendirilmekte olan algoritmalar, her aday mekan için bir kullanıcı-mekan skoru tahmin etmekte ve en yüksek puan alan mekanların en üst-k'sini tavsiye olarak geri getirmektedir. Tahmin doğruluğunu değerlendirmek için, kullanıcının konum geçmişini iki bölüme ayırmaktayız: 1) Her kullanıcı tarafından oluşturulan konum geçmişinin %30'unu rastgele bir test grubu olarak seçtik 2) Kullanıcıların konum geçmişinin geri kalanını eğitim seti olarak kullanıyoruz. Test setindeki mekanlar temel gerçekler olarak kabul edilir ve önerilen yerlerle eşleştirilir. Önerilen yerler gerçekten ziyaret edildiğinde daha yüksek bir doğruluk elde edilir. En üst  $k$  mekana ilgi duyduğumuzdan, önerilen mekanların kategori sıralamalarını, gerçekten ziyaret edilen mekanların kategori sıralamalarıyla eşleştiriyoruz. Daha özel olarak, *ortalama ölçüm kesinliğini* değerlendirme metriğimiz olarak inceliyoruz

**Çizelge-3: Temel yaklaşım, ECF ve TECF algoritmaları arasındaki karşılaştırma.**

Yöntem	Mekan Kategorisi	Uzman Seçimi	Giriş Bazlı Benzerlik	Giriş + Zaman Bazlı Benzerlik	Giriş Bazlı Kullanıcı -Mekan Matrisi	Giriş + Zaman Bazlı Kullanıcı-Mekan Matrisi
Temel CF	✓		✓		✓	
ECF	✓	✓	✓		✓	
TECF Algoritması 1	✓	✓		✓	✓	
TECF Algoritması 2	✓	✓	✓			✓
TECF Algoritması 3	✓	✓		✓		✓

Temel olarak, bir dizi sorguda ortalama kesinliğinin ortalamasını tahmin ediyoruz ve tavsiye edilen mekan kategorisi sıralamasının değerlendirilmesinde bu ölçüyü kullanıyoruz. Tavsiye edilen en üst  $n$  kategorilerdeki ve ilgili kategorilerdeki kesinliğe dayanarak, Denk. (7)'de verilen ortalama kesinliği hesaplayabiliriz.

$|R|$  = ilgili kategorilerin toplam sayısı, en üst  $R$  kategorileri  
 $kes(i) =$  en üst  $i$  önerilen kategorilerindeki kesinlik

$$ilgililik(i) = \begin{cases} 1, & \text{ilgili ise} \\ 0, & \text{değil ise} \end{cases}$$

$$AP = \frac{1}{|R|} \sum_{i=1}^n kes(i) \times ilgililik(i) \quad (7)$$

Her bir kullanıcı için her bir öneri kümesini ayrı ayrı sorgular olarak görerek, ortalama kesinlik Denk. (8)'de olduğu gibi hesaplanır.

$$|U| = \text{toplam sorgu sayısı,} \\ \text{(bizim durumumuzda toplam kullanıcı sayısı)} \\ AP(Q_i) = \text{tek bir } Q \text{ sorgusu için ortalama kesinlik} \\ MAP = \frac{1}{|U|} \sum_{Q_i \in Q} AP(Q_i) \quad (8)$$

## 5.2 Deneysel Sonuçlar

Bu bölümde, Tokyo ve New York veri kümelerinden elde edilen deney sonuçları ayrıntılı olarak sunulmaktadır.

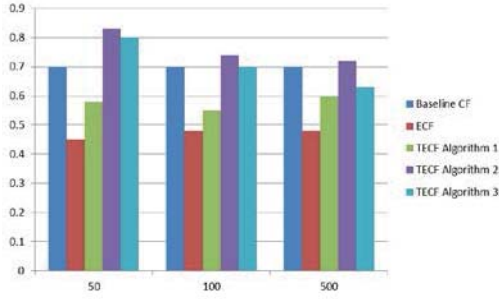
Tokyo Veri Seti Sonuçları Çizelge-4, Tokyo veri kümesi kullanılarak önerilen uzman sayısına ve üst kategori sayısına göre değişen farklı algoritmaların

haritadaki  $map@N$  sonuçlarını göstermektedir.  $map@N$  sonuçları, Şekil-2(a), Şekil-2(b) ve Şekil-2(c)'de grafik içine yerleştirilmiştir. Açıkça görülüyor ki, TECF Algoritma 2, tüm  $N$  ve uzman sayımları değerleri kapsamında geliştirilen temel ve diğer algoritmalarından daha iyi performans gösterir. Bu durum aralık başına bir kullanıcı-mekan giriş sıklığı matrisi kullanmanın gücünü gösterir. TECF Algoritması 3, çoğu durumda yine zaman bilgisinden faydalanmanın olumlu etkisini göstererek temel değerleri aşarken, TECF Algoritması 1 bazı durumlarda temelin biraz gerisinde kalmaktadır. Uzman seçimi içeren ve zamanı göz ardı eden ECF algoritmamız diğer tüm algoritmaların gerisinde kalmaktadır. Genel olarak, deneysel sonuçlar Algoritma 2 ve Algoritma 3'te sunulan zaman bazlı yaklaşımların önerilerin etkinliğini arttırdığını ve diğer iş birliğine dayalı filtreleme tekniklerine kıyasla çok rekabetçi olduğunu göstermektedir.

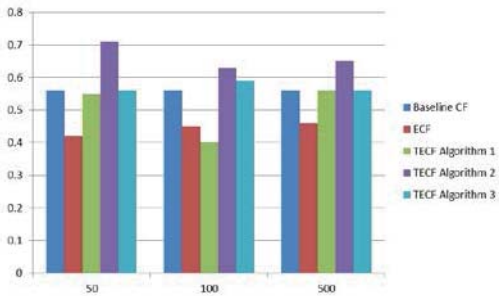
TECF Algoritması 1'de, kategori giriş bilgisi ve TECF Algoritması 2'de, mekan giriş bilgisi beş zaman aralığına bölünmüştür. Algoritma 3'te, hem kategori giriş hem de mekan giriş bilgisi zaman aralıklarına bölünmüştür. Bu durumda, veri kümesini birçok matrise bölmek, TECF Algoritması 2 ile karşılaştırıldığında seyreklikle sonuçlanır. Algoritma 3'te elde edilen matrisler, kullanıcıların zamana göre gerçek tercihlerini yakalamada daha az temsil edici ve daha az etkili olur. Bu nedenle, bazı durumlarda Algoritma 3'ün biraz daha düşük ölçülen  $map@N$ 'i veriler en fazla bölümlendiği için değerlidir.

**Çizelge-4:** Tokyo veri setini kullanarak temel yaklaşım, ECF ve TECF algoritmaları arasındaki deneysel sonuç karşılaştırması.

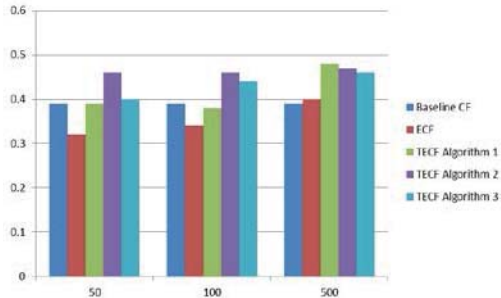
Uzman sayısı	MAP@3			MAP@5			MAP@10		
	50	100	500	50	100	500	50	100	500
Yöntem									
Temel CF	0,39			0,56			0,70		
ECF	0,32	0,34	0,40	0,42	0,45	0,46	0,45	0,48	0,48
TECF Algoritması 1	0,39	0,38	0,48	0,55	0,40	0,56	0,58	0,55	0,60
TECF Algoritması 2	0,46	0,46	0,47	0,71	0,63	0,65	0,83	0,74	0,72
TECF Algoritması 3	0,40	0,44	0,46	0,56	0,59	0,56	0,80	0,70	0,63



Şekil-2 (a) map@10



Şekil-2 (b) map@5



Şekil-2 (c) map@3

**Şekil-2:** Tokyo veri seti kullanılarak elde edilen, sırasıyla 50, 100, 500 uzman sayımları ile temel, ECF, TECF algoritmalarının map@N skorları

Temel iş birliğine dayalı filtreleme algoritması TECF Algoritması 2 - 3 ile karşılaştırıldığında, zaman bazlı algoritmalar önemli ölçüde iyileştirilmiş sonuçlar verir. Temel yaklaşım, kullanıcıların bir günün farklı zaman aralıklarında farklı tercihleri

olabileceğini varsaymadan, genel olarak kullanıcıların tercihlerini modellemektedir. Bir kullanıcı kahve dükkanlarında genel olarak en fazla sayıda giriş ve geceleri kahve dükkanlarından daha fazla sayıda barlarda giriş yaptığında, temel algoritma hala kahve dükkanı mekanlarını gece için artırır. Bununla birlikte, zaman aralıklarına bölündüğü için, TECF algoritmaları genel girişlerle ilgilenmez ve kullanıcının beşinci aralıkta barlarda en fazla giriş yaptığını dikkate alır. Bu nedenle, kullanıcının zaman aralıklarına göre değişen tercihleri daha iyi yakalanır.

Önerilerimizde zamanın etkinliğini daha ileri düzeyde keşfetmek için, 1 ve 2 numaralı algoritmalar detaylı bir şekilde karşılaştırılmalıdır. Algoritma 2, zaman aralığı başına bir kullanıcı-mekan giriş sıklığı matrisi kullanır ve Algoritma 1, aralık başına bir kullanıcı-kategori giriş matrisi kullanır. Algoritma 2, N arttıkça Algoritma 1'i in  $map@N$  sonuçlarında geçmektedir. Bu, kullanıcı başına aralık içi giriş sıklığı matrisinin aralık başına benzerlik değerlerini hesaplamaktan daha fazla öneri doğruluğunu artırdığını ortaya koymaktadır. Her aralık için kullanıcı-mekan giriş matrisi kullanıldığında, sadece bu aralıktaki tüm kullanıcılar tarafından giriş yapılan mekanlar, mekan set büyüklüğünü azaltarak ve bu aralıkta ziyaret edilen mekanı önerme olasılığını artırarak, kullanıcı-mekan skoru tahminine dahil edilir. Bununla birlikte, aralık başına benzerlik değerlerinin hesaplanması, mekan setinin boyutu hala aynı olduğundan, bu kadar gelişmiş bir etkiye sahip değildir.

$map@3$  bir değerlendirme metriği olarak kullanıldığında, uzman sayısı arttıkça ortalama keskinlik artmıştır. Bununla birlikte,  $map@5$  ve  $map@10$  kullanıldığında, aksi durumlar gözlenmiştir.

New York Veri Seti Sonuçları Çizelge-5, New York veri seti kullanılarak önerilen uzman sayısına ve üst kategori sayısına göre değişen farklı algoritmaların  $map@N$  sonuçlarını göstermektedir. TECF Algoritması 2 ve Algoritması 3, temel iş birliğine dayalı filtreleme algoritmasını geçse de, genel olarak, TECF algoritmalarının ortalama keskinlik değerleri azalır. Ayrıca, uzman seçimini içeren ve giriş zamanını göz ardı eden ECF algoritmamız, diğer tüm algoritmaların gerisinde kalmaktadır.

1083 aktif kullanıcı ve 111.973 giriş içeren New York veri kümesi göz önüne alındığında, bu makul bir sonuç olarak kabul edilebilir. Bu veri kümesi Tokyo veri setinden daha küçük olduğundan ve veriler bol olduğunda algoritmalarımız daha iyi performans gösterdiğinden, bu model daha düşük hassasiyetli değerler elde eder.

## 6. Sonuç

Bu makalede, mevcut sistemlerden bazı ana yönlerden farklılaşan zaman bazlı uzman destekli iş birliğine dayalı filtreleme yaklaşımı sunulmuştur:

Zaman, diğer yer tercihlerini etkileyebilecek kullanıcıların tercihlerini, davranışlarını ve sosyal yapılarını anlamak için kullanılmıştır.

Bu belirli bir zaman aralığında bir kullanıcı-mekan çifti için tercih skoru, TECF algoritmalarından faydalanan uzmanların zaman bilgisi kullanımlarına göre benzerlik ve mekan giriş sıklığı dikkate alınarak çıkarılmaktadır. Ardından, mekanların en üst  $k$  tanesi kullanıcıya bu zaman aralığının önerileri olarak sunulur.

Kategori uzmanlarının tercihleri, öneri performansını daha da arttırmayı amaçlamaktadır. Konum geçmişi önceden hesaplanır ve her bir kategorideki uzmanlar, kullanıcıların genel kategori giriş sıklıklarından gösterilir.

Ziyaret edilmemiş konumlar için önyargılı önerileri engellemek amacıyla, düşük güven değerleri, onları olumsuz olarak değerlendirmek yerine ziyaret edilmeyen değerlere tayin edilir.

Temel algoritma, ECF ve TECF algoritmaları, Foursquare'den toplanan iki gerçek dünya veri seti ile değerlendirilmektedir. Tokyo veri kümesi 2291 kullanıcı, 203.575 giriş, New York veri kümesi 1083 kullanıcı ve 111.973 giriş içermektedir. Her iki veri kümesinde de 35 kategori vardır. Tokyo veri kümesini kullanarak yapılan deneyin sonuçları, Algoritma 2 ve Algoritma 3'te sunulan zaman bazlı yaklaşımların öneri kalitesini önemli ölçüde artırdığını ve diğer iş birliğine dayalı filtreleme tekniklerine kıyasla çok rekabetçi olduğunu göstermektedir. Bu arada, TECF Algoritması 3, verinin seyrek matrislere bölünmesi nedeniyle Algoritma 2'nin biraz gerisinde kalmaktadır. Algoritmalarımızın performansları, her aralıkta az

miktarda veri bulunduğundan New York veri kümesi kullanıldığında düşmektedir. Gelecekteki çalışmalar için, kullanıcı profilini daha iyi temsil etmek için giriş verilerin alınan kullanıcıların coğrafi bilgilerinden yararlanmayı planlıyoruz. Ayrıca, önerici sistemimizin etkinliğini daha da artırmak için mekanların coğrafi bilgilerini dahil etmeyi düşünüyoruz.

## Kaynaklar

- [1] Yu Zheng. 2011. Location-based social networks: Users. *Computing with Spatial Trajectories* 243–276.
- [2] Jie Bao, Yu Zheng and Mohamed F. Mokbel. 2012. Location based and preference-aware recommendation using sparse geo-social networking data. In *Proceedings of the 20th International Conference on Advances in Geographic Information Systems*, 199–208. ACM
- [3] Yu Zheng, Xing Xie and Wei-Ying Ma. 2009. Mining interesting locations and travel sequences from GPS trajectories. In *Proceedings of the 18th International Conference on World Wide Web - WWW 09*.
- [4] Xin Cao, Gao Cong and Christian Søndergaard Jensen. 2010. Mining significant semantic locations from gps data. *Proceedings of the VLDB Endowment* 3, 1-2, 1009–1020.
- [5] Kazuki Kodama, Yuichi Iijima, Xi Guo and Yoshiharu Ishikawa. 2009. Skyline queries based on user locations and preferences for making location-based recommendations. In *Proceedings of the 2009 International Workshop on Location Based Social Networks*, 9-16. ACM.
- [6] Moon-Hee Park, Jin-Hyuk Hong and Sung-Bae Cho. 2007. Location-based recommendation system using bayesian user's preference model in mobile devices. In *International Conference on Ubiquitous Intelligence and Computing* 1130-1139.
- [7] Anastasios Noulas, Salvatore Scellato, Neal Lathia and Cecilia Mascolo. 2012. A random walk around the city: New venue recommendation in location-based social networks. In *International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing (socialcom)*.
- [8] Chen Cheng, Haiqin Yang, Irwin King and Michael R. Lyu. 2012. Fused matrix factorization with geographical and social influence in location-based social networks. In *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*.

- [9] Dingqi Yang, Daqing Zhang, Zhiyong Yu and Zhu Wang. 2013. A sentiment-enhanced personalized location recommendation system. In *Proceedings of the 24th ACM Conference on Hypertext and Social Media - HT 13*.
- [10] Betim Berjani and Thorsten Strufe. 2011. A recommendation system for spots in location-based online social networks. In *Proceedings of the 4th Workshop on Social Network Systems*. ACM.
- [11] Josh Jia-Ching Ying, Eric Hsueh-Chan Lu, Wen-Ning Kuo and Vincent S. Tseng. 2012. Urban point-of-interest recommendation by mining user check-in behaviors. In *Proceedings of the ACM SIGKDD International Workshop on Urban Computing* 63-70. ACM.
- [12] Xian Wu, Yuxiao Dong, Baoxu Shi, Ananthram Swami and Nitesh V. Chawla. 2018. Who will Attend This Event Together? Event Attendance Prediction via Deep LSTM Networks. In *Proceedings of the 2018 SIAM International Conference on Data Mining* 180-188. Society for Industrial and Applied Mathematics.
- [13] Mohamed Sarwat, Justin J. Levandoski, Ahmed Eldawy and Mohamed F. Mokbel. 2014. LARS\*: An efficient and scalable location-aware recommender system. *IEEE Transactions on Knowledge and Data Engineering*, 26, 6, 1384-1399.
- [14] Chi-Yin Chow, Jie Bao and Mohamed F. Mokbel. 2010. Towards locationbased social networking services. In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks* 31-38. ACM.
- [15] Pavlos Kosmides, Chara Remoundou, Konstantinos Demestichas, Ioannis Loumiotis, Evgenia Adamopoulou and Michael Theologou. 2014. A location recommender system for location-based social networks. In *Proceedings of International Conference on Mathematics and Computers in Sciences and in Industry (MCSI)* 277-280. IEEE.
- [16] Yifan Hu, Yehuda Koren and Chris Volinsky. 2008. Collaborative filtering for implicit feedback datasets. In *ICDM* 263-272.
- [17] Dingqi Yang, Daqing Zhang, Zhiyong Yu, Zhiwen Yu. 2013. Fine-grained preference-aware location search leveraging crowdsourced digital footprints from LBSNs. *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp 13* 479-488. ACM.
- [18] Dingqi Yang ; Daqing Zhang ; Vincent W. Zheng ; Zhiyong Yu. 2015. Modeling user activity preference by leveraging user spatial temporal characteristics in lbsns. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*.
- [19] Mao Ye1, Peifeng Yin, Wang-Chien Lee and Dik-Lun Lee. 2011. Exploiting geographical influence for collaborative point-of-interest recommendation. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information - SIGIR 11*, 325-334.