ANLAMSAL KONUM BİLGİLERİNİN YAYINLANMASINDA MAHREMİYETİN SAĞLANMASI

EMRE YİĞİTOĞLU

YÜKSEK LİSANS TEZİ

BİLGİSAYAR MÜHENDİSLİĞİ

TOBB EKONOMİ VE TEKNOLOJİ ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

EYLÜL 2012

ANKARA

Fen Bilimleri Enstitü onayı

_____

Prof. Dr. Ünver KAYNAK

Müdür

Bu tezin Yüksek Lisans derecesinin tüm gereksinimlerini sağladığını onaylarım.

_____

Doç. Dr. Erdoğan DOĞDU

Anabilim Dalı Başkanı

Emre YİĞİTOĞLU tarafından hazırlanan ANLAMSAL KONUM BİLGİLERİNİN YAYINLANMASINDA MAHREMİYETİN SAĞLANMASI adlı bu tezin Yüksek Lisans tezi olarak uygun olduğunu onaylarım.

_____

Yrd. Doç. Dr. Osman ABUL

Tez Danışmanı

Tez Jüri Üyeleri

Başkan : Prof. Dr. Faruk POLAT _____

Üye : Doç. Dr. Erdoğan DOĞDU _____

Üye : Yrd. Doç. Dr. Osman ABUL _____

TEZ BİLDİRİMİ

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada orijinal olmayan her türlü kaynağa eksiksiz atıf yapıldığını bildiririm.

Emre YİĞİTOĞLU

Üniversitesi : **TOBB Ekonomi ve Teknoloji Üniversitesi**
Enstitüsü : **Fen Bilimleri**
Anabilim Dalı : **Bilgisayar Mühendisliği**
Tez Danışmanı : **Yrd.Doç. Dr. Osman ABUL**
Tez Türü ve Tarihi : **Yüksek Lisans – Eylül 2012**

Emre YİĞİTOĞLU

ANLAMSAL KONUM BİLGİLERİNİN YAYINLANMASINDA MAHREMİYETİN SAĞLANMASI

ÖZET

Mobil cihaz teknolojilerindeki gelişme ile birlikte insanların konumları içerde veya dışarıda birkaç metrelik kesinlikle hesaplanabilmektedir. Bu ilerleme ile birlikte konum tabanlı servislerin kullanımı önemli boyutlara ulaşmış ve bu servislerle yapılan konum paylaşımı, mahremiyet sorunlarını beraberinde getirmiştir. Konum paylaşımı sırasında gönderilen koordinat değerlerinden kolaylıkla kişinin bulunduğu anlamsal konum bilgisinin çıkarılabilmesi kişinin davranışları hakkında bilgi vermesi sebebiyle her zaman istemeyeceği bir durumdur. Bu tezde konum mahremiyeti sağlama tekniklerinin çözmeye çalıştığı iki problem için çözüm önerilmiştir. Birinci problemde, kullanıcıların konum tabanlı servislerle yapmış oldukları konum paylaşımı sırasında mahremiyeti sağlamak amacıyla, insanların farklı mahremiyet gereksinimlerine göre yeni bir gizlenmiş bölge hesaplama yöntemi geliştirilmiştir. Bu yöntemde gizlenmiş bölgeler anlamsal konumlar, yol ağı kısıtı ve hız tabanlı atakları dikkate alarak oluşturulmuştur. İkinci problem konum-zamansal veri tabanlarının yayınlanması sırasında ortaya çıkmaktadır. Kullanıcıların belli bir süre içerisinde bulundukları anlamsal konum serilerinin yayınlanması sırasında, kullanıcı kimliklerinin gizlenmesi mahremiyeti sağlamada yetersiz kalmaktadır. Bu sebeple genel olarak kullanılan genelleştirme yönteminin yanı sıra sırasal verilere özel sıra esnekleştirme tekniğinin beraber kullanılması önerilmiştir.

Anahtar Kelimeler: Anlamsal konum mahremiyeti, Konum tabanlı servisler, Konum paylaşımı

University : **TOBB Economics and Technology University**
Institute : **Institute of Natural and Applied Sciences**
Science Programme : **Computer Engineering**
Supervisor : **Assistant Professor Dr. Osman ABUL**
Degree Awarded and Date : **M.Sc. – September 2012**

Emre YİĞİTOĞLU

PRESERVING PRIVACY WHEN SHARING SENSITIVE SEMANTIC LOCATIONS

ABSTRACT

With recent mobile technological advances, people location can be tracked both indoor and outdoor spaces with a spatial accuracy of a few meters. Along with this advance, location based services usage has increased remarkably and location sharing through these services brought versatile privacy issues. From the shared coordinates one can easily infer the respective semantic location such as hospital, nigh club and restaurant. This thesis investigates two privacy problems concerning the sharing of semantic sensitive locations. In the former, user identity is assumed to be known by service provider but the semantic location is asked to be obfuscated with respect to user privacy profile specification under road network and velocity based-attack constraints. Algorithms solving the problem in a few settings have been developed and extensive experimental evaluations are performed. Unlike the former where privacy is service-centric issue, the latter problem addresses data-centric privacy where collected semantic location trajectory database needs to be published in a privacy-preserving manner. The problem asks providing anonymity given semantic location traces of users. The solution approach uses generalization and order relaxation techniques to achieve anonymity.

Keywords: Semantic location privacy, Location based services, Location sharing

TEŞEKKÜR

Çalışmalarım boyunca değerli yardım ve katkılarıyla beni yönlendiren hocam Yrd. Doç. Dr. Osman ABUL ile birlikte Yrd. Doç. Dr. Maria Luisa DAMIANI, Yrd. Doç. Dr. Claudio SILVESTRI, kıymetli tecrübelerinden faydalandığım TOBB Ekonomi ve Teknoloji Üniversitesi Bilgisayar Mühendisliği Bölümü öğretim üyeleri, her konuda desteklerini esirgemeyen ailem, arkadaşlarım ve nişanlım Aşkın GÜLER'e teşekkürü bir borç bilirim.

İÇİNDEKİLER

ÇİZELGELERİN LİSTESİ

ŞEKİLLERİN LİSTESİ

Şekil **Sayfa**

KISALTMALAR

Kısaltmalar	Açıklama
KTS	Konum tabanlı servisler
CR	Gizlenmiş bölge
CRs	Gizlenmiş bölge kümesi
QS_T	Zaman hatası
QS_S	Mesafe hatası
FR	Hata oranı
PCS	Konum gizleme servisi
MCS	Gizlenmiş bölge hazırlama servisi
UCS	Sabit maliyet araması
SAWLnet	Yol ağı altında anlamsal konuma dayalı gizleme
LCP	Yerel bilgi kaybı
GCP	Genel bilgi kaybı

SEMBOL LİSTESİ

Bu çalışmada kullanılmış olan simgeler açıklamaları ile birlikte aşağıda sunulmuştur.

Simgeler	Açıklama
k	Anonimlik derecesi
pt	Yerleşke türü
p	Yerleşke
G	Annotated şehir ağı çizgesi
$pop(.)$	Popülerlik derecesi
tt	Zamansal yol ağırlığı
r	Bölge
d	Uzaklık fonksiyonu
τ	Mahremiyet eşik değeri
cr	Gizlenmiş bölge
K	Hesaplanacak maksimum duyarlı yerleşke sayısı
d	Saldırgan bilgi ölçütü
I	Sıra esnekleştirme derecesi
$g()$	Genelleştirme fonksiyonu
$p()$	Permütasyon sayısı

1. GİRİŞ

1.1. Konum Tabanlı Servisler

Mobil cihaz teknolojilerindeki ilerlemeler ile birlikte sadece iletişim kurmak için kullanılan düşük seviye cihazlardan yüksek hızlı akıllı telefonlara geçilmiştir. Bu geçiş ile birlikte kullanıcılar istedikleri her an internet erişimi ile bilgiye ulaşabilmektedir. Şu anki akıllı telefon kullanıcı sayısı yüz milyonlarla ifade edilmektedir. Bununla birlikte kullanıcıların konumları Global konum belirleme sistemi (GPS), şebeke sinyali ve ağ tabanlı konum belirleme sistemleri (örn: Skyhook Wireless) gibi araçlarla açık veya kapalı alanlarda birkaç metrelik kesinlikle belirlenebilmektedir. Bunların sonucu olarak kişilerin konum-zamansal verilerini kullanarak gerçek zamanlı olarak bilgi sağlayan Konum Tabanlı Servisler (Location Based Services) ortaya çıkmıştır [1]. Girdileri kullanıcının konumu ve isteği, çıktısı ise konumu ve hareketine göre isteğin sonucu olan bu servisler insanların bilgiye çok hızlı bir şekilde ulaşmasını sağlamaktadır. Örneğin, bulunduğu konuma yakın restoranları arayan bir kişi internette yapacağı çeşitli kriterlerde sorgulardan sonra restoranların tek tek web sitelerine ulaşması gerekmektedir. Bunun yerine çeşitli KTS uygulamalarında sadece restoran aradığını belirterek, kendisine en yakın restoranları, bu restoranın özelliklerini, diğer kullanıcıların restoran hakkındaki yorumlarını ve restorana nasıl gidebileceğini kolaylıkla öğrenebilmektedir. Sağladığı bu kolaylıklarla, çok büyük ve her gün artmakta olan bir pazar olan mobil uygulamalarda KTS'lerin sayısıda her geçen gün artmaktadır. Juniper Research [2010] mobil konum tabanlı servis piyasasının 2014 yılında 12 milyar dolar düzeyinde olacağı tahmininde bulunmuştur [2].

1.1.KTS'lerin Uygulama Alanları

Konum tabanlı servisler kullanıcıların ihtiyaçlarına göre farklı uygulama alanlarında özelleşmişlerdir [6].

Acil Durum Uygulamaları: Kullanıcıların yerlerini tam olarak bilemediği veya herhangi bir tehlike anında bulunduğu konumu bildiremediği durumlarda bu servisler büyük önem taşımaktadır. Örneğin bir safari sırasında araç bozulduğunda kullanıcı bu servisi kullanarak yardım isteyip bulunduğu konumu bilmese bile yardım ekibinin kullanıcıya hızlı bir şekilde ulaşması mümkün olacaktır. Bir diğer örnek ise ABD'de yapılan 911 acil aramalarında zor durumda bulunan bir kişinin bulunduğu yeri söylemesine ihtiyaç duymayan Enhanced911 projesi ile yer tespiti yapılarak gerekli yardım sağlanmaktadır [7].

Navigasyon Uygulamaları: Kullanıcının bulunduğu noktadan belirlediği bir noktaya zaman, trafik gibi parametreleri de kullanarak en uygun şekilde izleyeceği yolu anlık olarak sunan sistemlerdir.

Bilgi Alma Uygulamaları: Kullanıcının bulunduğu nokta veya çevresi hakkında bilgi alabildiği uygulamalardır. En yakındakini bulma, trafik hakkında bilgi alma, seyahat rehberliği, hava durumu gibi birçok alanda bilgi sağlayan servislerdir. Örnek olarak en yakındaki restoran nerde, bulunulan konumda önümüzdeki yirmi dört saatteki hava durumu nasıl gibi sorulara cevap veren servisler verilebilir. Kullanıcı bu tür servislere üye olarak hareket sırasında da anlık bilgiler alabilmektedir. Örneğin turistin gezisi sırasında yakınında bulunan bir tarihi bina hakkında bilgi anında kişiye sunulabilmektedir.

İzleme ve Yönetim Uygulamaları: Bu tip uygulamalar kişisel bazda kullanımının yanı sıra şirket bazında büyük faydalar sağlamaktadırlar. Örneğin posta gönderilerinin o an nerde olduğu birçok kargo şirketi tarafından kullanıcılarına sunulmaktadır. Ayrıca araç takibi ile her bir aracın konumu, hızı gibi birçok bilgiye anlık olarak erişebilmektedir ve herhangi bir olumsuzluğa anında müdahale etme imkanı sunmaktadır. Örnek olarak kullanıcı sadece taksiye ihtiyacı olduğunu belirttiğinde en yakın taksinin oraya yönlendirilmesi sağlanmaktadır.

Sosyal Ağ Uygulamaları: Kullanıcıların bulundukları konumları arkadaşlarıyla paylaştıkları, arkadaşlarının nerede olduklarını öğrendikleri uygulamalardır.

Reklam Uygulamaları: Reklam uygulamaları KTS içerisinde en yaygın uygulama alanlarından bir tanesidir. Örneğin, alışveriş firmaları satışlarını arttırmak için kendilerine belli bir uzaklıktaki kullanıcılara reklamlarını göndermekte veya orada olduklarını arkadaşlarıyla paylaşan kullanıcılara indirim yapmaktadırlar. Bununla birlikte kullanıcılarda kendisine yakın nerelerde indirim olduğunu kolaylıkla öğrenebilmektedir.

1.2.Konum Paylaşımında Mahremiyet

KTS'lerin kullanımıyla birlikte kullanıcılar güvenilirliği garanti olmayan bu servislere konum bilgilerini paylaşmaktadır. Bu servislerin giderek artması konum mahremiyeti konusunun da daha dikkatli ele alınması gerekliliğini ortaya çıkartmaktadır. İnsan hayatını çok çeşitli alanlarda kolaylaştıran bu servislerden faydalanırken konum bilgisinin istenmeyen amaçlarla kullanılmayacağının sağlanması gerekmektedir. Konum mahremiyetinin sağlanamamasının doğurduğu bazı problemleri şu şekilde sıralayabiliriz [3]:

- Konum tabanlı istenmeyen posta: Kullanıcıların konumlarına bağlı olarak bazı market etiğine uymayan firmalar ürün veya servislerini tanıtmak için kullanıcıların istememelerine rağmen istenmeyen posta olarak reklam yollayabilirler.

- Kişisel güvenlik tehditleri: Konum bilgisinin kötü niyetli kişiler tarafından bilinmesi bazı durumlarda kullanıcının güvenliğini tehdit edebilmektedir.

- İstenmeyen çıkarsamalar: Konum bilgisinin paylaşılması o anki içinde bulunulan veya yol boyunca bulunduğu mekanlar göz önüne alınarak kişinin politik görüşü, sağlık durumu ve kişisel tercihleri hakkında çıkarsama yapılmasını sağlamaktadır. Örneğin, KTS kullanıcısının yaptığı "en yakın restoran nerede?" gibi masum görünen sorguları sırasında hastanede bulunması kişinin sağlık durumu hakkında bilgi edinilmesini sağlayabilir.

Konum paylaşımında mahremiyetin sağlanması amacıyla birçok yöntem ortaya atılmıştır. Bu yöntemler çözmeye çalıştığı problemlere göre kategorilere ayrılmaktadır. Buna göre kullanıcının KTS ile olan etkileşimi sırasında kimlik bilgisine ihtiyaç duyup duymaması ve kullanıcıların tek bir konum yerine belli bir zaman aralığındaki konumlarının paylaşılmasına göre konum mahremiyeti sağlama problemleri gruplanmıştır [8].

Kimlik mahremiyeti: Kullanıcının kimlik bilgisinin, konum bilgisiyle ilişkilendirilerek çıkarılamamasını sağlamak problemin temelini oluşturmaktadır. Buna göre sağlanan mahremiyette paylaşılan konum bilgisinin kesinliği yüksek iken kişinin kimliği gizlenmektedir.

Konum mahremiyeti: Servis sağlayıcının kimlik bilgisine ihtiyaç duyduğu durumlardır. Kullanıcın bulunduğu konumun kesin olarak paylaşılmasının yerine farklı veya daha geniş bir bölge içerisinde bulunduğunun paylaşılmasını gerektiren problemlerdir. Burada dikkat edilmesi gerek konu paylaşılan bu yeni bölgenin servis kalitesini olumsuz yönde etkileyeceğidir.

Yol Mahremiyeti: Konum bilgilerinin anlık yerine belli bir zaman aralığında kullanıcıların bulunduğu noktaların paylaşılması sırasında doğan mahremiyet problemidir. Bu mahremiyet problemi kullanıcıların kimlikleri saklanmasına rağmen ev, işyeri gibi bilgiler dikkate alınarak hangi yolun hangi kişiye ait olduğunun tespit edilebilmesinden kaynaklanmaktadır.

Bu problemleri çözmek için kullanılan yöntemlerde kendi aralarında üçe ayrılmaktadırlar. Bu yöntemlerin bazıları tek bir problemin çözümü olarak karşımıza çıkarken, bazıları birden fazla problem için uygulanabilmektedir.

1.2.1. Anonimlik Tabanlı Yöntemler

Anonimlik kavramı, kullanıcın kimliğinin paylaşılan konum bilgisinden çıkarılamaması temeline dayanmaktadır. Kimlik mahremiyeti ve yol mahremiyeti problemlerinin çözümü için önerilmiş yöntemler vardır.

Bu konuda yapılmış çalışmalardan Mix-Zones [9] kullanıcının izlediği yolun daha önce tanımlanmış bir bölge içerisinde bulunduğunda izlediği yolun takip edilememesiyle anonimlik sağlanmasını amaçlar. Böyle bir bölge içerisine giren kullanıcıların takma adları değiştirilerek bölgeye giriş ve çıkışları arasında bir ilişki kurulması engellenerek kimliklerin fark edilmemesi sağlanır. Üçüncü bir güvenilir katman kullanıcılar ve servisler arasında bulunarak güvenilirliği garanti olmayan uygulamaların konum bilgilerine ulaşması engellenmiştir. [10], [11], [12], [20] Mix-Zones yaklaşımını kullanan diğer çalışmalardır.

Bir diğer önemli yaklaşım ise *k*-Anonimlik yaklaşımıdır [13], [14]. *K*-Anonimlik veri tabanındaki her bir kayıt için aynı özellikteki *k*-1 tane farklı kayıt bulunduğunu garanti ederek kayıtların birbirinden ayırt edilememesiyle hangi kayıtın kime ait olduğunun bulunamaması mantığı üzerine kurulmuştur. Benzer şekilde konumsal *k*-Anonimlik kavramı da herhangi bir kullanıcının konumu *k*-1 farklı kullanıcıdan ayırt edilememesidir. Servis ve kullanıcı arasında bulunacak bir mahremiyet sağlayıcı güvenilir kaynak ile kullanıcıların servis kullanmadan önce bu katmana istekleri gönderilip, servislere *k* farklı kullanıcının bulunduğu bir bölge yollanılarak mahremiyet sağlanır. Açıkça görülmektedir böyle bir mahremiyet sağlama kullanıcının kimliği gerekli bir servis için uygun değildir.

Anonimlik tekniklerinin bir diğer kullanıldığı alan olan yol mahremiyeti problemlerinde ise belli bir zaman aralığında kullanıcıların ziyaret ettiği noktaların yayınlanması sırasında bazı çıkarımsal bilgilerle kayıtın kime ait olduğunun belirlenebilmesi engellenmektedir [21], [23].

1.2.2. Konum Gizleme Tabanlı Yöntemler

Anonimlik tabanlı yöntemlerde kişilerin konum-kimlik eşlemesinin saldırgan tarafından yapılabilmesi engellenmeye çalışılırken, konum gizleme tabanlı yöntemlerde kişinin tam olarak nerede olduğunun bilinmesi engellenmektedir. Kimlik bilgisinin yayınlanması kişiye özgü sonuçlar üreten servislerin kullanılmasını mümkün kılmaktadır.

Gizleme tabanlı yöntemlerden bir tanesi belirsizlik konsepti üzerine kurulmuştur [24]. Buna göre kullanıcının gerçek pozisyonuna ek olarak bulunabilme olasılığı aynı olan *n* adet başka noktanın gerçek nokta ile birlikte servis sağlayıcıya sunulmaktadır. Böylelikle saldırgan kişinin bu noktaların herhangi birinde olduğunu bilecek ama tam olarak hangisinde olduğunun kararını veremeyecektir.

Bir diğer yöntem Gruteser [22] tarafından sunulmuştur. Yoğunluğuna göre alanlar hassas veya önemsiz olarak ele alınıp, hassas alanlar için oluşturulan gizlenmiş bölgenin ek olarak *k*-1 farklı alanı da içermesi gerekmektedir. Kullanıcı bu bölgede bulunduğu sürece konum paylaşmamaktadır.

Probe [35] ise kullanıcı tarafından belirlenen duyarlı bölgeler, duyarsız bölgelerle birleştirilerek bir gizlenmiş bölge hesaplanır. Bölgelerin alanına göre kullanıcın bulunduğu alanın tahmin edilme olasılığının kullanıcının belirlediği değerden yüksek olmayacak şekilde bölge genişletilir.

1.2.3. Kural Tabanlı Teknikler

Mahremiyet gereksiniminin kişiden kişiye veya uygulamadan uygulamaya farklılık göstermesi sebebiyle konum bilgisine erişimin belli kurallarla sınırlandırılması temeline dayanan tekniklerdir. Bu teknikteki anahtar nokta kimin hangi konum bilgisine hangi şartlarda erişebileceğinin kararının verilmesidir ve bu noktadan hareketle KTS ile yapılacak konum paylaşımından önce ön tanımlı kuralların uygulanarak erişilecek bilgi kısıtlandırılır. Tüm mahremiyet problemleri için uygulamaları mevcuttur [25, 26, 27, 28, 29, 30, 31, 32, 33, 34].

1.3. Döküman Yapısı

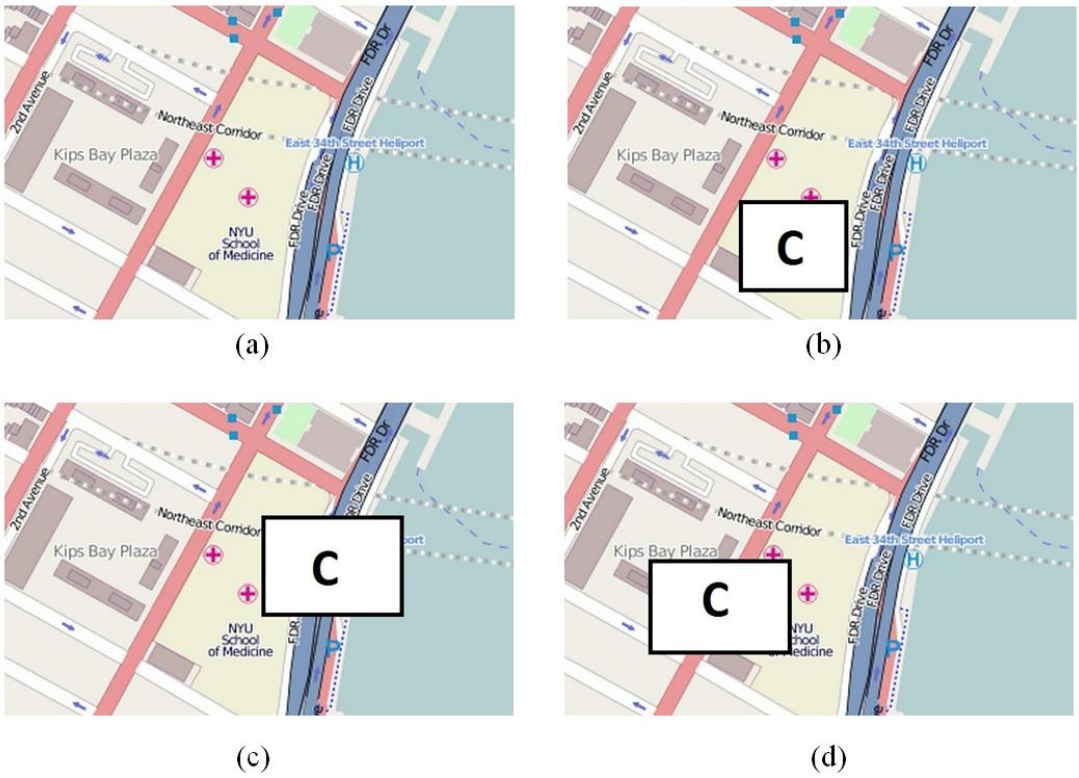
Tezin bundan sonraki kısımları şu şekilde sıralanmaktadır: Bölüm 2'de konum paylaşımı sırasında doğan mahremiyeti önlemek için öne sürülen algoritma ve Bölüm 3'de geliştirilen algoritmanın örnek bir mobil uygulaması olan SAWLnet sunulmuştur. Bölüm 4'de anlamsal konum veri tabanlarının yayınlanması sırasında doğacak mahremiyeti önlemek için önerilen yaklaşım verilmiş, Bölüm 5'de ise tezin sonuç ve ileriki çalışmalar kısmı yer almıştır. [44] ve [45] numaralı referanslar tez kapsamından yapılan çalışmalar ile ilgilidir.

2. ANLAMSAL HASSAS KONUM BİLGİLERİNİN PAYLAŞIMINDA YOL AĞI KISITI ALTINDA MAHREMİYETİN SAĞLANMASI

2.1.Motivasyon

Konum tabanlı servislerin sayısının her geçen gün artması ve bunların güvenilirliklerinin garanti olmamasıyla birlikte konum paylaşımda mahremiyet sağlamanın önemi artmıştır. Bu çalışmada KTS'lerin kullanıcıların kimliklerine ihtiyaç duyduğu problemler göz önüne alınarak, hassas anlamsal konum bilgilerini ve yol ağı kısıdını aynı anda ele alan yeni bir konum gizleme tabanlı mahremiyet sağlama tekniği ortaya atılmıştır.

Anlamsal konum; kişinin bulunduğu coğrafi koordinat noktasına karşılık gelen hastane, market, vb. gibi bölgeleri ifade eder. Kullanıcıların herhangi bir anlamsal konumda bulunduğunun başkaları tarafından öğrenilmesi istenmeyebilir. Bu durumda bu konum kişiye göre hassas anlamsal konumdur. Konum paylaşımı sırasında kullanıcın hassas anlamsal bir konumda bulunduğunun gizlenmesi gerekmektedir [3], [35]. Buna göre anlamsal konum bilgilerinin dikkate alınmadan yapıldığı konum gizleme uygulamalarda [24], [22] saldırganının daha kesin konum bilgisine ulaşması mümkün olacaktır. Örneğin [35], Şekil 2.1(a)'da New York şehrinden bir kesit görülmektedir. Buna göre bu bölgede bir hastane, sağ tarafında nehir ve sol tarafında da yerleşke bulunmaktadır. Ayrıca bu anlamsal konumları bağlayan yollarda görülmektedir. Ek olarak nehir üzerinde yolculuk yapılamadığı varsayılmıştır.



Şekil 2.1: Anlamsal konum

Anlamsal konum bilgisi göz önüne alınmadan belirlenen gizlenmiş bölgelerde ortaya çıkabilecek ilk bilgi sızıntısı Şekil 2.1(b)'de karşımıza çıkmaktadır. Belirlenen gizlenmiş bölge halen belli bir anlamsal konum içinde kalmış olabilir. Bu gizlenmiş bölgenin (cloaking region, CR) paylaşılması sonucu saldırgan kullanıcın CR içerisinde tam olarak nerede olduğunu bilememesine rağmen hastanede olduğu bilgisine ulaşabilmektedir.

Şekil 2.1(c)'de ise oluşturulan CR'nin bir kısmı hastanede iken bir kısmı nehirdedir. Anlamsal konumların göz önüne alınmadan oluşturulmuş CR, sınırların ve nehir üzerinde ulaşımın olmayacağı bilgisine sahip saldırganın kullanıcının kesin olarak hastanede olduğu bilgisine kolayca ulaşabilmesine sebep olacaktır.

Bir diğer problem de Şekil 2.1(d)'de karşımıza çıkmaktadır. Oluşturulan CR hem yerleşkede, hem de hastanede bulunmaktadır. İlk olarak, herhangi bir anlamsal

konumun hassas olup olmaması kişiden kişiye değişmektedir. Örneğin bir doktor için hastane bilgisi hassas değilken, bir politikacı için hassas bir konum olabilmektedir. Oluşturulan CR'ların servis kalitesini düşürdüğü düşünüldüğünde herkes için aynı CR oluşturulması gereksiz gizleme ile servisten sağlanan faydayı azaltacaktır. Bu sebeple oluşturulacak gizlenmiş bölgenin kişinin tercihlerine göre farklılık göstermesi gerekmektedir. Bir diğer sorun ise saldırganın anlamsal bölgelerin sınırlarına ek olarak yerleşim yerindeki popülasyon dağılımı bilgisine sahip olduğu durumda karşımıza çıkmaktadır. Buna göre saldırgan yerleşke ve hastanedeki kişi sayısını karşılaştırarak, kullanıcının daha yüksek olasılıkla hastanede olduğu bilgisine ulaşabilecektir.

Hassas anlamsal konum bilgilerinin paylaşılmasını engellemek için ilk akla gelebilecek yollardan bir tanesi olan kural-tabanlı mahremiyet sağlama beklenilen faydayı sağlamamaktadır. Örneğin kullanıcının hastanede bulunduğu sırada konum paylaşımının yapılmaması kuralı uygulanması, saldırganın kullanıcı izlerinden çıkarsama yaparak hassas anlamsal konumun paylaşılmamasına rağmen kullanıcının burada olduğu sonucuna ulaşmasını engellemeyecektir. Ayrıca servis sağlayıcının kimlik bilgisine ihtiyaç duyduğu varsayımıyla anonimlik tabanlı yöntemlerin kullanılamayacağı açıktır.

Bu çalışmanın temelindeki, anlamsal konumlara ve kişisel mahremiyet gereksinimlerine göre gizlenmiş bölgeleri oluşturan Probe [35], yerleşim yeri üzerinde yapılacak hareketin sınırsız olduğuna göre çalışmaktadır. Oysa bir yerleşim yerinde yapılacak hareket yol ağı kısıtı ile sınırlandırılmıştır. Yol ağı gözetilmeden yapılacak gizleme tehlikelere açıktır. Örneğin, böyle bir uygulamada kuş uçuşu olarak birbirine yakın gözüken iki alan aynı gizlenmiş bölge içerisine alınırken, yol ağı kısıtını göz önünde bulundurulduğunda bu alanların birbiri arasındaki yolun çok uzun olduğu görülebilir. Bir diğer eksiklik de [4]'de belirtildiği gibi hız tabanlı ataklara karşı korumasız olmasıdır. Bir saldırgan yol ağı ve yol üzerinde bulunan hız limiti bilgilerini kullanarak kullanıcının yayınlanan gizlenmiş bölgenin bir kısmında bulunamayacağını belirleyerek gizlenmiş bölgenin etkinliğini azaltabilmektedir.

2.2.Model

Yerleşim merkezini bir model üzerine oturtmak amacıyla *PT* yerleşke türü (ör: hastane, ibadethane) ve *P* yerleşkeleri ifade etmektedir. Ayrıca yerleşim alanı yolların kenar, yerleşke ve kavşak noktalarının ise düğümleri oluşturduğu bir çizge olarak tanımlanmıştır. Yerleşkeler insanların bulunabileceği anlamsal alanlar olarak kabul edilmektedir.

Tanım 1 (Annotated şehir ağı) : Annotated şehir ağı, bağlı ve yönsüz bir *G* çizgesidir ve $G = (V , E , pop , pt , tt)$ şeklinde 5 öğeli olarak tanımlanır, öyleki:

- $V = V_p \cup V_j$ çizgedeki düğümler kümesidir. Bu düğümler $v \in V_p$ ifadesiyle yerleşkeleri, $v \in V_j$ ifadesiyle de kavşak noktalarını ifade etmektedir.

- $E \subseteq V \times V$ çizgedeki kenarlar kümesine karşılık gelmektedir. $(v , v) \in E$ iki kavşak noktası veya bir kavşak noktası ile bir yerleşke arasındaki yol parçasını ifade etmektedir. Herhangi iki yerleşke arasında doğrudan bir yol parçası yoktur ve en az bir tane kavşak noktası içermesi gerekmektedir. Ayrıca yerleşim merkezinin bir bağlı çizge olarak ifade edilmesinin sonucu olarak tüm yerleşkeler çizgeye doğrudan bağlıdır ve tüm düğümler arasında en az bir tane yol bulunmaktadır.

- Her bir yerleşke için popülerlik metriği tanımlanmıştır. *pop* ile gösterilen bu metrik $pop = V_p \rightarrow (0,1)$ fonksiyonuyla ifade edilir ve $pop(v)$, $v \in V_p$ için herhangi bir kullanıcının v yerleşkesinde bulunma olasılığına (bir başka deyişle yerleşkenin yoğunluğu) karşılık gelir. Bir yerleşke için bu değerin 0 olması o yerleşkenin kullanıcılar tarafından ulaşılabilir olmadığı anlamına gelmektedir.

- $pt : V_p \rightarrow PT$ Her yerleşkenin ait olduğu bir yerleşke türü vardır. $v \in V_p$ olmak üzere $pt(v)$, v yerleşkesinin türünü ifade etmektedir.
- Çizge üzerinde bulunan bütün $e = (\upsilon , v) \in E$ kenarlarına tt ağırlığı atanmaktadır. Bu $tt : E \rightarrow \mathbb{R}$ ağırlığı υ düğümünden v düğümüne ulaşmak için gerekli olan minimum zamandır.

Koordinat sistemindeki herhangi bir noktanın çizge üzerindeki en yakın düğüm veya kenara eşleme mekanizmasının olduğu varsayılmıştır. Böylelikle kullanıcının bulunduğu (x , y) noktası önerilen model üzerinde bir kenar veya düğümle eşleştirilecektir.

Bu modelde, *bölge* şehir ağı içerisindeki bir bağlı altçizgedir ve $V' \subseteq V$ ve $E' \subseteq E$ iken $G' = (V' , E')$ olarak ifade edilir. Bir bölge tek bir düğüm veya kenardan oluşabileceği gibi tüm şehir ağına da karşılık geliyor olabilir. Bir r bölgesinin içerisindeki pt yerleşke türündeki yerleşkelerin popülerliği $pop_r(pt)$ olarak gösterilmektedir ve o bölgedeki pt yerleşke türüne sahip tüm yerleşkelerin popülerliklerinin toplamına karşılık gelmektedir. Daha genel bir ifadeyle bölgenin popülerliği $pop_r(.)$ ile gösterilir ve

$$\sum_{pt_i \in PT} pop_r(pt_i) \qquad (2.1)$$

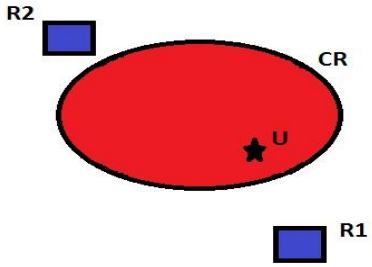
toplamına eşittir.

Kullanılan modelde bir diğer ifade olan *gerçek zamanlı hareket izleri* kullanıcının şehir ağı içerisinde hareketi sırasında zamana bağlı konum hareketliliğini belirtmektedir ve $t_i < t_{i+1}$ iken $T = \{ (r_1,t_1),(r_2,t_2),(r_3,t_3),\ldots,(r_n,t_n)\}$ olarak tanımlanır. Kullanıcın belli zamanlarda bulunduğu alanı gösteren (r_i,t_i) ifadesinde t_i o anki zamanı ve r_i de kullanıcın bulunduğu bölgeyi ifade etmektedir. Mahremiyet sağlamak amacıyla kullanıcının KTS ile iletişimi sırasında KTS kullanıcının (x , y)

coğrafik koordinat bilgileri yerine bu r bölgesinin içerisinde olduğu bilgisini öğrenebilecektir.

2.2.1. Mahremiyet Gereksinimleri

Önerilen model gizleme tabanlı bir mahremiyet sağlama yöntemi üzerine kurulmuştur [3]. Buna göre kullanıcının belirleyeceği duyarlı yerleşkeler baz alınarak gizlenmiş bölge kümeleri (CRs) oluşturulacak, daha sonra ise kullanıcının o anki bulunduğu nokta herhangi bir gizlenmiş bölge (CR) içerisinde olup olmadığı test edilecek ve eğer gerekliyse dönüşüm işlemi gerçekleştirilecektir. Önerilen modelde bu CR'ler kişilerin mahremiyet gereksinimlerini karşılayan birer altçizge olarak ifade edilmektedir. Kullanıcıların kesin konumlarını göndermek yerine daha geniş bir alan göndermeleri KTS'ler den sağlanacak bilginin kalitesinde azalmaya yol açacacaktır. Örneğin, Şekil 2.2'de kullanıcı U'nun yapacağı "en yakın restoran nerede?" sorgusuna KTS kişinin konumu bir nokta olarak değerlendirdiğinde döneceği sonuç R1 olacakken gizlenmiş bölge olarak değerlendirdiğinde R2 sonucu üretecektir.



Şekil 2.2 : CR'lerın servis kalitesine etkisi

Buna göre KTS'lerden sağlanacak bilginin kalitesi CR'lerin büyüklüğüyle doğru orantılıdır. Buradan yola çıkarak gizlenmiş bölgeye bağlı servis kalite metriği oluşturulan CR'lerin çapı olarak tanımlanmıştır. $\{G_1 , G_2 , ... G_n\}$ altçizgelerinin ortalama çapı;

$$QS_{CR} = \frac{1}{n}\sum_{i=1}^{n} cap(G_i) \quad (2.2)$$

şeklinde hesaplanmaktadır. Bir CR'nın çapı ifadesi gizlenmiş bölge içerisindeki bir düğümden diğer bir düğüme ulaşırken gerekli olan minimum zamanların maksimumu olarak tanımlanır.

Bunun yanı sırada [4] de belirtilen zaman hatası ve mesafe hatası metrikleri kalite ölçümünde kullanılan diğer metriklerdir ve hız tabanlı ataklara karşı mahremiyetin sağlanmasında ortaya çıkan hataların servis kalitesine olan etkisini belirtmektedirler. Zaman hatası kullanıcının KTS'ye yapmış olduğu isteğin gerçek zamanı t_i yerine gecikmeli olarak t_i' anında KTS'ye iletilmesi sonucu oluşan hatadır ve $t_i' - t_i > 0$ süre gecikmeden doğan zaman hatası QT şu şekilde hesaplanır;

$$QS_T = \frac{1}{n}\sum_{i=1}^{n}(t_i' - t_i) \quad (2.3)$$

Kullanıcın bulunduğu noktanın veya içinde bulunduğu gizlenmiş bölgenin yayınlanmasının mahremiyeti zafiyete uğratacağı durumlarda kullanılan bir diğer yöntemde kullanıcı konumunun başka bir nokta veya gizlenmiş bölge olarak göstermektir. Bu da kullanıcının KTS'den alacağı servisin kalitesinin düşmesine neden olacaktır. Mesafe hatası olarak ifade edilen QS_S;

$$QS_S = \frac{1}{n}\sum_{i=1}^{n} d(p_i, p_i') \quad (2.4)$$

eşitliği ile bulunur. Buradaki d uzaklık fonksiyonudur ve kullanıcın gerçek noktasıyla, KTS ile paylaşılan noktası arasındaki uzaklıktır. Bu uzaklık hesabı iki nokta arasının aşılabileceği minimum zamandır. Görüldüğü üzere çizgenin kenarlarının *tt* zaman ağırlığı atanması, kullanılan tüm kalite metriklerinin zamana bağlı metrikler olarak hesaplanmasını sağlamıştır. Buda kullanılan yöntemlerde karşılaştırmanın daha sağlıklı yapılmasına neden olmaktadır.

Kullanıcının yayınlanacak nokta veya gizlenmiş bölgenin hesaplanabilmesi mahremiyet gereksinimleri altında her zaman mümkün olmayabilir. Bu durumda kullanıcının isteği yok sayılacaktır. Bu hatanın tüm isteklere oranı da bir başka metrik olan hata oranını (FR) verir ve şu şekilde hesaplanır;

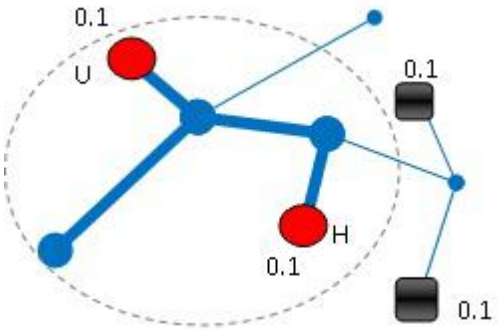
$$FR = \frac{\text{Düşürülen istek sayısı}}{\text{Toplam istek sayısı}} \tag{2.5}$$

Kullanıcıların mahremiyet sağlama servisinden beklentileri kişiden kişiye farklılık göstermektedir. Bunun için kişiye özel mahremiyet profillerinin belirlenmesi gerekir. Bu mahremiyet profillerinde her bir yerleşke türü $pt_i \in PT$ için kullanıcın belirleyeceği eşik değeri τ_i değeri tanımlanmalıdır. Bu τ_i değeri kullanıcının herhangi bir yerleşkede bulunduğunun saldırgan tarafından ne kadarlık bir olasılıkla tahmin edebileceğine izin verileceğinin ölçütüdür. $0 < \tau_i < 1$ aralığında tanımlı olan bu değerin $\tau_i = 1$ olması pt_i yerleşke türünün kullanıcıya göre duyarlı bir tür olmadığını belirtir ve mahremiyet profilinde bu değerlere yer verilmez. Aynı şekilde eşik değerinin 0 eşit olması karşılanamayacak bir mahremiyet gereksinimine karşılık gelir ki bu da tanımsızdır. τ_i değeri 0'a ne kadar yaklaşırsa yerleşke türünün duyarlılığının aynı oranda artmasına karşılık gelmektedir. Kullanıcının mahremiyet gereksinimlerini ifade ederken tüm duyarlı yerleşke türleri için (pt_i , τ_i) ikilileri tanımlanır. Buna göre herhangi bir CR'in mahremiyet gereksinimi karşılaması kullanıcın $p \in pt_i$ duyarlı yerleşke içerisinde bulunduğunun daha önceden izin verdiği τ_i eşik değerinden daha yüksek bir olasılıkla belirlenemeyeceği olarak tanımlanır. Buna göre bir CR'nin mahremiyet gereksinimi şu şekilde ifade edilir:

$$\frac{pop_r(pt_i)}{pop_r(.)} \leq \tau i \tag{2.6}$$

Örnek olarak Şekil 2.3 ele alındığında kırmızı daireler ile üniversite (U) ve hastane (H), siyah dikdörtgenlerle de duyarsız yerleşkeler gösterilmiştir, ayrıca mavi noktalar ile kavşaklar, mavi çizgiler ile de yollar belirtilmiştir. Tüm yerleşkelerin popülerliği

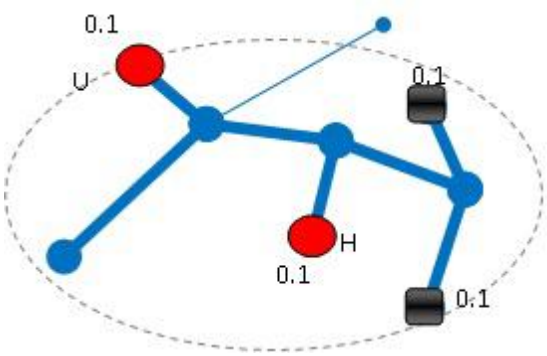
0.1 olarak belirlendiği ve mahremiyet profilinin üniversite ve hastane duyarlı yerleşkeleri için $PP = \{(U, 0.5), (H, 0.5)\}$ şeklinde tanımlandığı varsayılmıştır. Buna göre şekilde noktalarla sınırlandırılmış gizlenmiş bölge mahremiyet gereksinimlerini karşılamaktadır. Çünkü saldırgan kullanıcının bu gizlenmiş bölge içerisinde hastanede olduğunu 0.5 olasılıkla tahmin edebilecektir ve buda kişinin hastane için belirlediği eşik değerini aşmamaktadır. Ayrıca aynı durum üniversite için de geçerlidir. Fakat açıkça görülmektedir ki, kullanıcının hangi yerleşkede olduğunun bilinme olasılığına göre mahremiyet sağlanmış olsa da saldırgan tarafından kişinin bir duyarlı bölge içerisinde olduğu bilgisine ulaşabilmiştir. Bu sebeple her bir duyarlı yerleşkenin ayrı ayrı mahremiyet gereksinimi karşıladığı bu modelde gizlenmiş bölge kişinin mahremiyetini sağlamakta yetersiz kalmıştır.



Şekil 2.3 : Zayıf mahremiyet

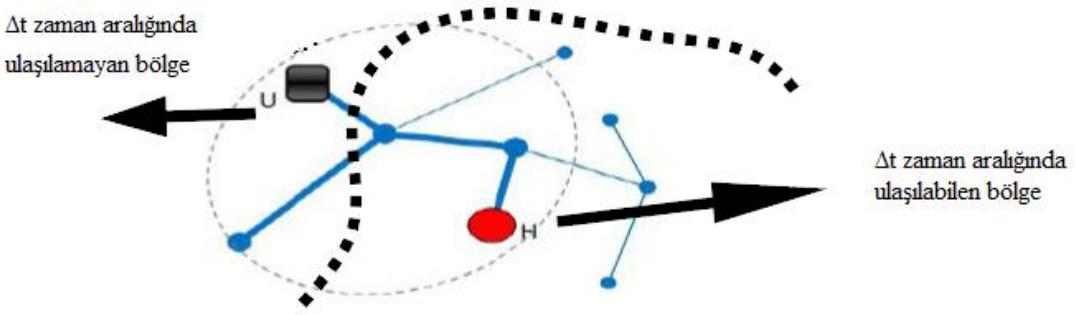
Bunun yerine gizlenmiş bölge için daha güçlü bir mahremiyet gereksinimi tanımlanmıştır. Buna göre her bir duyarlı yerleşkenin mahremiyet gereksinimi sağlayacak toplam ihtiyaç duyduğu popülerlik değeri toplanıp , gizlenmiş bölgenin olması gereken minimum popülerlik değeri bulunur. Bir önceki örneğin güçlü mahremiyet ile oluşturulmuş gizlenmiş bölgesi Şekil 2.4'de verilmiştir.

$$\sum_{pt_i \in PT_s} \frac{pop_r(pt_i)}{\tau i} \geq pop_r(.) \qquad (2.7)$$



Şekil 2.4 : Güçlü mahremiyet

Bir diğer gereksinim de [4]'de belirtilen hız tabanlı atağın ele alınmasıdır. Şekil 2.5'de hız tabanlı atak örnek olarak gösterilmiştir. H duyarlı ve U duyarsız yerleşkeleri içeren bir gizlenmiş bölgenin yayınlanması sonucu tek başına ele alındığında mahremiyet gereksinimini karşılıyor olmasına rağmen, yol ve maksimum hız bilgisine sahip saldırgan kullanıcın bir önceki paylaşılan noktadan iki paylaşım arasındaki sürede ulaşabileceği noktaları belirleyebilir. Bunun sonucu olarak gizlenmiş bölge içersindeki $\Delta t = (t_c - t_p)$ süresinde ulaşılamayacak noktaları eleyebilir. Örnekte olduğu gibi gizlenmiş bölge içerisinde U yerleşkesi zaten erişilemez olduğu için saldırgan kullanıcının H içerisinde olduğu bilgisine ulaşabilecektir.



Şekil 2.5 : Hız tabanlı atak

Hız tabanlı atakları önlemek için, o anki ve bir önceki paylaşılan nokta veya gizlenmiş bölgeler arası uzaklığın $d_{pp}(G_1, G_2)$ iki istek arası zaman farkından daha az olmaması gerekir. $d_{pp}(G_1, G_2)$, $G_1 = (V_1, E_1)$ içerisindeki herhangi bir nokta ile $G_2 = (V_2, E_2)$ içerisindeki herhangi bir nokta arasında çizilebilecek en kısa yolun alabileceği en büyük değerdir, hız tabanlı mahremiyet gereksinimi şu şekilde tanımlanır:

$$d_{pp}(G_1, G_2) \leq \Delta t \tag{2.8}$$

2.3. Mimari

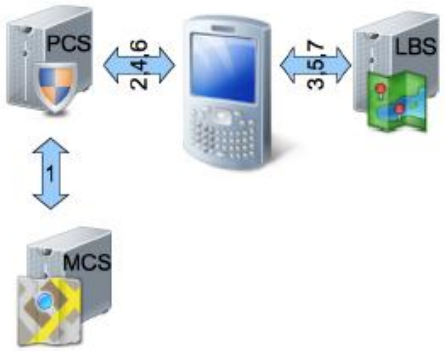
Mobil cihazlar sahip oldukları CPU, RAM, ikincil bellek gibi özelliklerine göre farklılık göstermektedir. Bu farklılıklar göz önüne alınarak önerilen sisteminde mimarisi değişiklik göstermektedir. Sistem dizaynı sırasında üç farklı türde cihaz ele alınmıştır; düşük kaynaklı akıllı telefonlar, tablet PC'ler ve notebooklar.

Cihazların kaynaklarına göre üçüncü parti mahremiyet sağlama servislerine olan bağımlılık şekillenecektir. Buna göre kaynaklarına göre sınıflandırılan üç farklı sistem mimarisi önerilmiştir.

2.3.1. Sunucu Tabanlı Mimari

Düşük çalışma hızı ve hafızaya sahip akıllı telefonlar için önerilen mimaridir. Bu cihazların gizlenmiş bölgeleri ne hesaplayacak ne de depolayacak kaynakları yoktur. Bu sebeple bu tür cihazlar üçüncü parti konum gizleme servislerine (PCS) ihtiyaç duyarlar. Sunucu tabanlı mimarinin örnek görüntüsü Şekil 2.6'da verilmiştir. Buna göre kullanıcı güvenirliği kesin olmayan konum tabanlı servislerden yararlanmadan önce PCS ile iletişime geçecektir. Mahremiyet profiline göre gizlenmiş bölge hazırlama servisi (MCS) tarafından hazırlanan gizlenmiş bölgeler bir kere hesaplanıp PCS içerisinde saklanacaktır. Daha sonra PCS kullanıcının her isteği sırasında bu

bölgeler ile hız tabanlı atakları göz önünde bulundurarak bulunduğu nokta yerine yayınlanacak bölgeyi hesaplayacaktır. Kullanıcı PCS'den dönen yayınlanmasında sakınca olmayan bölgeyi KTS ile paylaşarak isteğine cevap alacaktır.



Şekil 2.6: Sunucu tabanlı mimari

2.3.2. Melez Mimari

Tablet PC gibi yüksek depolama alanına sahip fakat gizlenmiş bölgeleri hesaplayacak kadar yeterli CPU ve RAM'e sahip olmayan kullanıcılar için önerilen mimaridir. Bu mimaride kullanıcı dışarıdan gizlenmiş bölgeleri hesaplayan MCS'den yararlanacak fakat gizlenmiş bölgeler bir kere hazırlandıktan sonra kendi içerisinde depolayacak ve bulunduğu konum için gerekli dönüştürme işlemini kendisi yapacaktır. Herhangi başka bir bölgeye veya mahremiyet profilinde bir değişiklik olmadığı sürece kendisinde saklanan CR'ler kullanılacaktır. Örnek mimari Şekil 2.7'de gösterilmiştir.

Şekil 2.7: Melez Mimari

2.3.3. Kullanıcı Tabanlı Mimari

Yüksek depolama alanına ve hızlı işlemciye sahip notebook gibi araçlara sahip kullanıcılar için önerilen mimaridir. Görsel olarak Şekil 2.8'de verilen bu mimariye göre kullanıcı mahremiyet profiline uygun gizlenmiş bölgeleri kendi hazırlayıp bunları depolayabilecektir. Dışarıdan sadece bulunan bölgenin haritasına ihtiyaç duyacaktır. Bölge içerisinde bulunan yollar, yerleşkeler ve bunların türleri gibi bilgileri OpenStreetMap[5] üzerinden sağlayacaktır. Buradan alacağı bilgiye sadece yeni bir bölgeye girdiği zaman ihtiyaç duyacaktır.

Şekil 2.8: Kullanıcı Tabanlı Mimari

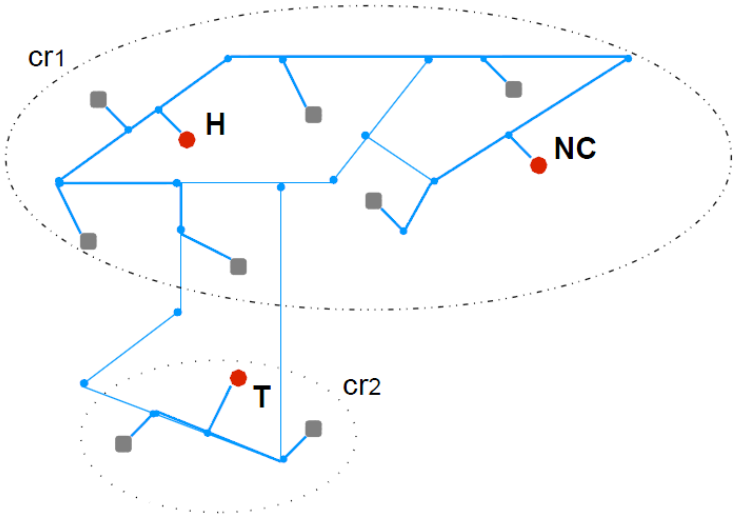
2.4.Algoritmalar

Bu bölümde KTS kullanıcılarının mahremiyet gereksinimlerinin karşılanması amacıyla kullanılan algoritmalar açıklanmıştır. Kullanıcının konum paylaşımı sırasında mahremiyet profiline göre hesaplanan gizlenmiş bölgeler baz alınarak konum bilgisinde dönüşüm işlemi uygulanacaktır. Bu dönüşüm işlemi sırasında ayrıca hız tabanlı ataklarda göz önüne alınacaktır. Gizlenmiş bölgelerin başta veya istek sırasında hesaplanmasına göre algoritmalar sırasıyla *offline* ve *online* olmak üzere iki ana başlık altında toplanmıştır.

2.4.1. Offline Gizleme

Bu algoritmada mahremiyet sağlama iki farklı aşamada yapılmaktadır. İlk olarak kullanıcının mahremiyet profiline göre tüm duyarlı yerleşkelerin gizlenmiş bölgeleri hesaplanacaktır. İkinci aşama ise kullanıcının KTS ile kurdukları iletişim sırasında konumlarının mahremiyet gereksinimlerini karşılayacak şekilde dönüşümden geçmesi aşamasıdır.

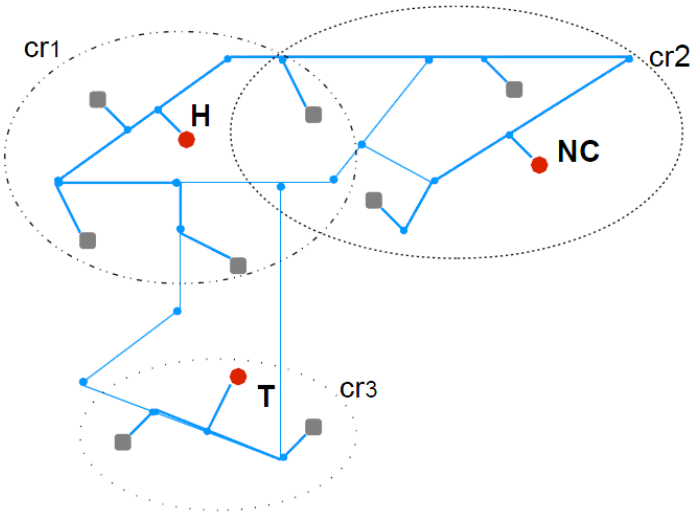
Gizlenmiş bölgelerin hesaplanması için iki farklı metot kullanılmıştır. Bunlardan ilki olan *ayrık gizleme* oluşturulan gizlenmiş bölgelerin birbiri ile kesişmemesini gerektirir. Şekil 2.9 kırmızı dairlerle belirtilen hastane, gece kulübü ve tapınak duyarlı yerleşkeleri için örnek bir ayrık gizleme algoritmasının uygulanışını göstermektedir. Görüldüğü üzere bir gizlenmiş bölge içerisinde birden çok duyarlı yerleşke bulunabilmektedir ve CR'ler birbirleriyle herhangi ortak düğüm veya kenar paylaşmamaktadır.



Şekil 2.9 : Ayrık Gizleme

Bir diğer yöntem olan *kesişimli gizleme* algoritmasında ise herhangi bir gizlenmiş bölge içersinde sadece bir tane duyarlı yerleşke bulunacağını garanti eder. Yukarıdaki örnekte belirtilen duyarlı yerleşkeler için kesişimli gizleme algoritmasıyla oluşturulmuş gizlenmiş bölgeler Şekil 2.10'da görülmektedir. Ayrık gizleme ve kesişimli gizleme arasındaki fark gizlenmiş bölgelerin alanıdır. Gizlenmiş bölgelerin çapı servis kalitesini etkileyen bir metrik olması açısından kesişimli gizleme algoritmasının daha iyi sonuçlar üreteceği açıktır. Fakat, bir noktanın birden fazla gizlenmiş bölge içerisinde bulunabilmesi dönüşüm sırasında mahremiyet gereksinimi sağlayan tüm CR'ler arasından birinin rastgele seçilerek belirlenmesi bazı durumlarda mahremiyet ihlaline sebep olmaktadır. Örneğin, kesişim bölgesinde bulunan bir kullanıcın belli zaman aralıklarında yapacağı konum paylaşımı sırasında gizlenmiş bölgelerin rastgele seçimiyle paylaşılan bölgenin sürekli değişmesi saldırgan tarafından kullanıcının yerinin kesişim bölgesinde olduğu fikrine ulaşmasına sebep olacaktır. Bir diğer örnek ise kullanıcının duyarlı bir yerleşke içerisindeyken yapacağı belli zaman aralıklarındaki konum paylaşımı sırasında saldırganın her seferinde aynı gizlenmiş bölgenin yayınlanması sonucu kullanıcının kesişim bölgesinde olamayacağı sonucunu çıkartabileceği durumdur. Böyle durumları önlemek rastgele seçim yerine bir önceki paylaşılan gizlenmiş bölgenin

diğer seferlerde de yeniden paylaşılarak mümkün olabilmektedir. Ancak bu durumda da kullanıcın diğer CR'ye geçmesi durumunda saldırgan kesişim bölgesinde olmadığı sonucuna ulaşabilecektir. Görüldüğü üzere kesişimli gizleme algoritmasıyla ayrık gizleme algoritmasına göre KTS'den daha yüksek fayda sağlanırken, bu algoritmanın kullanımı sırasında ortaya çıkabilecek özel mahremiyet aşımlarının dikkatli ele alınması gerekmektedir.



Şekil 2.10: Kesişimli Gizleme

Gizlenmiş bölgeler, kullanıcının belirlediği mahremiyet profiline göre duyarlı yerleşkelerin mahremiyet gereksinimini sağlayacak şekilde diğer yerleşkeler eklenerek oluşturulurlar. Bu altçizgelerin servis kalitesi açısından mümkün olan minimum çapta olması gerekmektedir. Bu amaçla bir duyarlı yerleşkeye ait CR hesaplanırken, bu duyarlı yerleşkeye en yakın yerleşke ve arasındaki en kısa yol gizlenmiş bölge altçizgesine eklenecektir. Eğer CR mahremiyet gereksinimi sağlamıyorsa bir sonraki en yakın yerleşke ve arasındaki yol altçizgeye eklenecektir. Bu işlem CR'nin mahremiyet gereksinimi karşılayıncaya kadar devam eder. En yakındaki yerleşkenin bulunması ağırlıklı çizgelerde arama algoritmalarından Sabit Maliyet Araması (Uniform Cost Search) kullanılmıştır. Etkili ve hızlı bir algoritma olan UCS ile en yakın yerleşke ve arasındaki en kısa yol bulunarak kompakt

gizlenmiş bölgelerin oluşturulması sağlanmıştır. Gizlenmiş bölgeler oluşturulurken başka bir duyarlı yerleşke veya CR ile karşılaşıldığında *kesişimli gizleme* ve *ayrık gizleme* algoritmaları farklı şekilde davranmaktadırlar.

Kesişimli CR Algoritması: Bu algoritma ile herhangi bir düğüm veya kenar birden fazla CR içerisinde yer alırken, bir CR içerisinde sadece bir tane duyarlı yerleşke bulunacaktır. Bunun için gizlenmiş bölgesi hesaplanacak duyarlı yerleşkeden başlanarak UCS algoritmasıyla en yakın yerleşkeler sırasıyla bulunacak eğer bulunan bu yerleşke başka bir duyarlı yerleşke ise göz ardı edilecektir. Bu algoritmanın sözde kodu Algoritma 1'de verilmiştir.

Algoritma 1 Kesişimli Gizleme Algoritması

Girdi Annotated şehir ağı $G = (V , E , pop , pt , tt)$,
mahremiyet profili $PP = \{(pt_i , \tau_i)\}_{i \in [1,n]}$
Çıktı CR'ye eşleme (m)

1. m ← ∅
2. **for all** $u \in V$ ve $u.pt \in PT_S$ **do**
3. $cr \leftarrow \emptyset$
4. $toplamPop \leftarrow u.pop$
5. **while** (true) **do**
6. $v \leftarrow$ *UCS(u)* ile bir sonraki düğüme ilerle
7. **if** $v.pt \in PT_{NS}$ **then**
8. *cr* çizgesine *u* ile *v* arasındaki en kısa yolu ekle
9. $toplamPop \leftarrow toplamPop + v.pop$
10. **if** $u.pt = pt_i$ için $\frac{u.pop}{pop_{cr}(.)} \leq \tau i$ **then**
11. **break**
12. m ← m ∪ *cr*

Ayrık CR Algoritması: Gizlenmiş bölgelerin birbiriyle kesişmediği ve bir kullanıcın en fazla bir CR içerisinde olacak şekilde hazırlanırken kullanılan algoritma ayrık gizleme algoritmasıdır. Bu algoritmanın sözde kodu Algoritma 2'de verilmiştir. Ayrık gizleme algoritmasında UCS algoritması ile en yakın yerleşkeler bulunurken eğer başka bir CR veya duyarlı yerleşke ile karşılaşıldığında bunlarda söz konusu altçizgeye eklenecektir. Yeni oluşan bu altçizgeye eklenmesi gereken

duyarsız yerleşke popülaritesi, altçizge içerisinde bulunan tüm duyarlı yerleşkelerin mahremiyet gereksinimi karşılayacak duyarsız yerleşke popülaritelerinin toplamıdır.

Algoritma 2 Ayrık Gizleme Algoritması

Girdi Annotated şehir ağı $G = (V , E , pop , pt , tt)$,
mahremiyet profili $PP = \{(pt_i , \tau_i)\}_{i \in [1,n]}$
Çıktı CR 'ye eşleme (m)

1. m ← ∅
2. **for all** $u \in V$ ve $u.pt \in PT_S$ **do**
3. $cr \leftarrow \emptyset$
4. $gerekliNSPop \leftarrow 0$
5. **while** (true) **do**
6. $v \leftarrow UCS(u)$ ile bir sonraki düğüme ilerle
7. **if** $v.pt \in PT_{NS}$ **then**
8. *cr* çizgesine *u* ile *v* arasındaki en kısa yolu ekle
9. $cv \leftarrow$ *v'nin içinde bulunduğu CR*
10. $v.pt \in pt_i$ için $gerekliNSPop \mathrel{+}= v.pop\frac{(1-\tau_i)}{\tau_i}$
11. **if** $cv \neq \perp$ **then**
12. *cr.birleştir(cv)*
13. **for all** $w \in cv.V$ ve $w.pt \in PT_S$ **do**
14. $w.pt \in pt_i$ için $gerekliNSPop \mathrel{+}= w.pop\frac{(1-\tau_i)}{\tau_i}$
15. **if** $pop_{cr}(PT_{NS}) \geq gerekliNSPop$ **then**
16. **break**
17. m ← m ∪ *cr*

Kesişimli gizleme ve ayrık gizleme algoritmalarının sonucunda kişinin mahremiyet profiline göre gizlenmiş bölgeler hazırlanmıştır. Offline gizleme ile bu gizlenmiş bölgeler kullanıcın KTS'den yararlanmaya başlamasından önce hazır olarak tutulacaktır. İkinci aşama olan dönüştürme aşamasında kullanıcının bulunduğu konum ve daha önceki yayınlanan bölge veya noktaya bağlı ortaya çıkabilecek hız tabanlı ataklar göz önüne alınarak yeni paylaşılacak bölge veya nokta belirlenmektedir. Dönüştürme algoritmasının sözde kodu Algoritma 3'de verilmiştir. Buna göre ilk olarak kullanıcın hangi gizlenmiş bölgeler içerisinde olduğu belirlenecektir. Ayrık gizleme algoritması kullanılmışsa bu değer en fazla bir CR olacaktır. Eğer herhangi bir CR içerisinde değilse bulunduğu noktada bir altçizge olarak göz önünde bulundurulup gizlenmiş bölge olarak ele alınmaktadır. Daha sonra bu gizlenmiş bölgelerden hız atağına karşı hangilerinin paylaşılmasında bir sakınca

olmadığı hesaplanır. Bu hesap bir önceki paylaşılan CR içerisindeki tüm düğümler ile şu an içinde bulunulan CR arasındaki tüm düğümler arasında zamana göre en kısa mesafeler hesaplanır. Buna göre en büyük zaman hız atağına karşı CR'nin güvenli olarak yayınlanabilmesi için gerekli olan zamandır. Eğer bir tane nokta veya CR güvenliyse bu yayınlanacaktır. Birden fazla gizlenmiş bölgenin güvenli olması durumunda ise aralarından biri rastgele seçilecektir. Fakat hiçbir güvenli CR yoksa bu durumda *zaman gecikmesi* ve *sahte konum* yaklaşımları uygulanacaktır. İlk olarak zaman gecikmesi yaklaşımı ile içinde bulunulan CR'ın yayınlanabilmesi için gerekli olan zamana belli bir zaman gecikmesi ile ulaşılabiliyorsa bu zaman gecikmesinin en az olacağı CR seçilir. Bazı durumlarda bu zaman gecikmesi çok fazla olabilmektedir. Örneğin içinde bulunan gizlenmiş bölgenin çapının çok büyük olması gerekli zaman gecikmesinin kullanıcının beklemek istemeyeceği boyutlara ulaşmasına sebep olabilecektir. Böyle bir durumda kullanıcının servisten yararlanırken bekleyeceği maksimum süre belirlenir. Gerekli zaman gecikmesinin tanımlı maksimum bekleme süresinden fazla olması durumunda sahte konum yaklaşımı devreye girecektir. Bu yaklaşıma göre o anki bulunulan konumdan bir önce yayınlanan konuma en kısa yol hesaplanmaktadır. En kısa yolun bulunması Floyd-Warshall algoritması ile yapılmaktadır. Daha sonra o an bulunulan noktadan başlanılarak en kısa yol üzerinde ilerlenilir ve her bir düğüm için yukarıdaki algoritma sanki gerçek bulunulan nokta gibi ele alınır. Eğer bir düğümün içinde bulunduğu CR mahremiyet gereksinimini karşılıyorsa bu bölge yayınlanacaktır. En kötü durumda bir önceki bölge yayınlanması gerekli maksimum zamanın sıfır olması sebebiyle bu bölge sonuç olarak döndürülecektedir.

Algoritma 3 Dönüştürme Algoritması

Girdi Annotated şehir ağı $G = (V , E , pop , pt , tt)$, CR 'ye eşleme (m) , t_q istek zamanı , U kullanıcısının bulunduğu nokta loc , bir önce yayınlanan CR/nokta P , bir önceki konumun yayınlanma zamanı t_p

Çıktı CR/nokta ve yayınlanma zamanı

1. $CRsU \leftarrow \{cr \in m : loc \in cr\}$ // kullanıcın içinde bulunduğu gizlenmiş bölgeler
2. **if** $CRsU = \emptyset$ **then**
3. $CRsU \leftarrow loc$
4. $\overline{CRsU} \leftarrow \{ cr \in CRsU :$ cr hız tabanlı ataklara karşı güvenli $\}$
5. **if** $\overline{CRsU} \neq \emptyset$ **then**
6. **return** rastgele $cr \in \overline{CRsU}$ ve t_q
7. $minZamanGecikmesi \leftarrow min_{cr \in CRsU}\{$ cr 'nin güvenli olabilmesi için gerekli zaman gecikmesi$\}$
8. **if** $minZamanGecikmesi \leq MaxİzinVerilenZamanGecikmesi$ **then**
9. // zaman gecikmesi
10. $cr_{min} \leftarrow argmin_{cr \in CRsU}\{$ cr 'nin güvenli olabilmesi için gerekli
11. zaman gecikmesi$\}$
12. **return** cr_{min} ve $(tq + minZamanGecikmesi)$
13. **else**
14. // *sahte konum*
15. $cr_f \leftarrow$ *enKısaYol(loc , P) üzerinde karşılaşılan ilk güvenli*
16. *bölge veya nokta*
17. **return** cr_f ve t_q

2.4.2. Online Gizleme

Offline gizleme sırasında yerleşkelerin popülerlikleri belli bir zamanı göz önüne almadan ortalama yoğunluklarına göre hesaplanmaktadır. Fakat bir yerleşkenin yoğunluğu belli zamanlarda farklılık göstermektedir. Örneğin bir gece kulübü gece saatlerinde çok yoğunken gündüz saatlerinde oldukça seyrektir. Aynı şekilde okulların eğitim dönemindeki yoğunluğu ile tatil dönemindeki yoğunluğu oldukça farklıdır. Bu sebeple ortalama popülerlik değerlerine göre hesaplanmış gizlenmiş bölgeler saldırgan tarafından yerleşkelerin o anki popülerlik değerlerini göz önüne aldığında mahremiyet aşımlarına sebep olacaktır.

Yerleşkelerdeki yoğunluğun değişiminden kaynaklanan mahremiyet sorunu göz önüne alarak gizlenmiş bölgelerin hesabını kullanıcının KTS'den yapacağı istek

sırasında hesaplanarak dönüşüm işleminin gerçekleştirilmesini öngören *online gizleme* algoritması önerilmiştir. Bu algoritmanın sözde kodu Algoritma 4'de verilmiştir. 1. Satırda yer alan *altçizgeBul* fonksiyonu ile bir önce yayınlanan P gizlenmiş bölge veya noktasından başlanarak iki istek arasındaki zaman olan $t_q - t_p$ süresinde ulaşılabilinen tüm noktalardan oluşan altçizge hesaplanacaktır. Böylelikle bir sonraki adım olan gizlenmiş bölgelerin hesabında ulaşılması mümkün olmayan duyarlı yerleşkeler için gereksiz gizlenmiş bölge hesabından kurtulunmaktadır. *gizlemişBölgeleriHesapla* fonksiyonunda Algoritma 1 ve Algoritma 2'de verilen *ayrık gizleme* veya *kesişimli gizleme* algoritmalarından biri kullanılarak G' altçizgesi içerisinde bulunan duyarlı yerleşkelere ait gizlenmiş bölgeler hesaplanmaktadır. Daha sonra ise bu gizlenmiş bölgeler kullanılarak Algoritma 3'de verilen dönüştürme işlemi gerçekleştirilmektedir.

Algoritma 4 Online Gizleme Algoritması

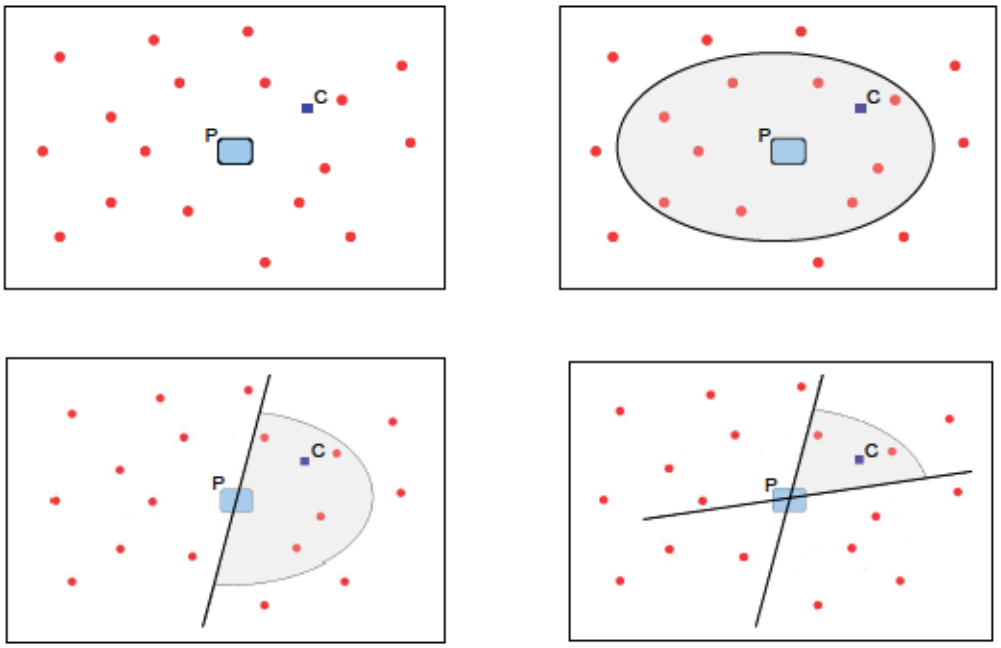
Girdi Annotated şehir ağı $G = (V, E, pop, pt, tt)$, mahremiyet profili $PP = \{(pt_i, \tau_i)\}_{i \in [1,n]}$, t_q istek zamanı, U kullanıcısının bulunduğu nokta *loc*, bir önce yayınlanan CR/nokta P, bir önceki konumun yayınlanma zamanı t_p

Çıktı CR/nokta ve yayınlanma zamanı

1. $G' \leftarrow altçizgeBul(G, P, t_q - t_p)$
2. $m' \leftarrow gizlemişBölgeleriHesapla(G', PP)$
3. **return** *Dönüştürme(G' , m' , tq , loc)*

Online algoritmada, offline algoritmadaki gibi gizlenmiş bölgelerin daha önceden hazır olmaması ve bu bölgelerin konum paylaşımı sırasında hesaplanması mahremiyet sağlama servisinin kullanıcıya döneceği cevap süresi açısından daha kötü sonuç vereceği açıktır. Öyle ki en kötü durumda bir önceki istek ile o anki istek arasında çok fazla zaman olması offline algoritmada yapılan tüm işlemlerin istek sırasında yapılması anlamına gelecektir. Ele alınan bölge içerisindeki gizlenmiş bölgelerin sayısı algoritmanın çalışma zamanını etkileyen en önemli faktördür. Bu sebeple gizlenmiş bölgeleri hesaplanacak duyarlı yerleşke sayısının azaltılması gerekir. Fakat bu sayının azaltılması kullanıcının o anki bulunan konumuna göre yapıldığında ister belli bir yöntemle isterse de rastgele olarak seçilsin saldırganın yayınlanan gizlenmiş bölgenin bir kısmında kullanıcının olamayacağı sonucuna ulaşarak bu bölgeyi daraltabilecektir. Buna karşın bir önceki yayınlanan bölge P göz önüne alınarak gizlenmiş bölgeleri hesaplanacak duyarlı yerleşkelerin

belirlenmesi önerilmiştir. Bu yaklaşıma göre ilk olarak P üzerinden geçen rastgele seçilmiş bir çizgiyle $t_q - t_p$ sürede erişilebilecek alan ikiye bölünecektir. Bu bölme işleminden sonra kullanıcın içinde bulunduğu kesim göz önününe alınacak ve buradaki duyarlı yerleşkelerin sayısı daha önceden belirlenmiş gizlenmiş bölgesi hesaplanabilecek maksimum duyarlı yerleşke sayısı K ile karşılaştırılacaktır. Bu sayı K değerinden fazla ise P bölgesinden gecen başka bir rastgele çizgi ile bu sayı K sayısından daha düşük veya eşit oluncaya kadar devam edilecektir. Bu işlem Algoritma 4 de belirtilen *gizlemişBölgeleriHesapla* fonksiyonunun ilk adımı olarak uygulanmaktadır. Burada anlatılan adımlar Şekil 2.11'de görsel olarak ifade edilmiştir. Şeklin birinci kısmında bir önce yayınlanan bölge P, kullanıcının o an bulunduğu konum c ve kırmızı noktalarla da duyarlı yerleşkeler gösterilmiştir. İkinci kısımda ise elips bölge $t_q - t_p$ zamanda erişilebilecek alanı göstermektedir. K değerinin 3 olarak belirlendiği varsayımına göre ilk olarak üçüncü kısımda P üzerinden geçen bir çizgiyle erişilebilinecek bölge ikiye bölünmüş ve kullanıcının içinde bulunduğu bölgedeki duyarlı yerleşke sayısı K değerinden büyük olduğu için dördüncü kısımda gösterildiği gibi ikinci bir kesme işlemiyle istenilen sayıya ulaşılmıştır.



Şekil 2.11: Gizlenmiş bölgeleri hesaplanacak duyarlı yerleşkelerin belirlenmesi

2.5.Deneysel Çalışmalar

Bu bölümde OpenStreetMap'den[5] elde edilmiş Milan (İtalya) haritası üzerinde rastgele oluşturulmuş kullanıcı istekleri ile önerilen algoritmaların çalışma performansları ele alınmıştır.

2.5.1. Kullanılan Veri Kümesi

Önerilen yöntemlerin daha sağlıklı test edilmesi açısından gerçek bir harita olan OpenStreetMap'den faydalanılmıştır. Yerleşke alanı olarak bu haritadan indirilen Milan şehir merkezinin 9 km^2 (3 km x 3 km) alanı kullanılmıştır. Bu harita ile yollar, yollarda gidilebilecek maksimum hız, yerleşkelerin sınırlandırdığı bölgeler ve yerleşkelerin türleri bilgileri alınarak bazı önişlemlemlerden geçirerek annotated şehir ağı oluşturulmuştur. Ayrıca veri temizleme işlemleri de uygulanarak yollara bağlantısı olmayan bölgeler silinmiş ve elde edilen çizgenin bağlı olması sağlanmıştır. Çizelge 1'de kullanılan veriler sunulmuştur.

Çizelge 2.1 : Kullanılan veri seti

Özellik	Değer
Alan	Milan şehir merkezinin 3km x 3km lik kısmı
Düğüm sayısı	8263 (yerleşke ve kavşak toplamı)
Yol sayısı	34000
Yerleşke sayısı	3800
Yerleşke türleri	eğitim (22) , sağlık (27) , ibadethane (64) , sosyal aktivite alanları (20) , eğlence merkezi (30) , alışveriş merkezleri (40) , spor (6)

Elde edilen yerleşke türlerinin popülerlik dereceleri yerleşkelere atanan varsayımsal populasyonlar sonucu belirlenmiştir. Bu atamalara göre yerleşkelerin popülerlik değerleri Çizelge 2.2'de görülmektedir. Bu popülerlik değeri bir kişinin herhangi bir yerleşkede bulunma olasılığını göstermektedir. Bu olasılık çizelgede da görüldüğü gibi yerleşke türüne göre farklı değerler almaktadır. Yerleşkelere atanan populasyon değerleri belli türdeki yerleşkeler için aynı olarak verilmiştir.

Çizelge 2.2: Yerleşke & Popülerlik değeri

Yerleşke Türü	Popülerlik Değeri
Eğitim	0.003351
Sağlık	0.001675
İbadethane	0.000502
Sosyal Aktivite Alanları	0.000335
Eğlence Merkezi	0.000837
Alışveriş Merkezleri	0.000279
Spor	0.000558
Diğer	0.000223

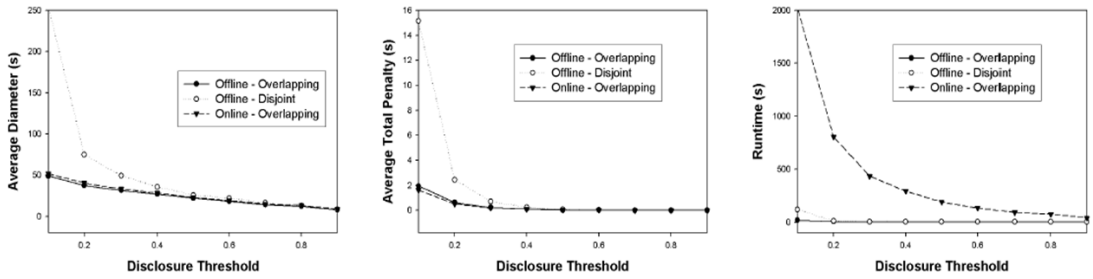
Oluşturulan annotated şehir ağı içerisinde KTS kullanan 1000 farklı kullanıcı göz önüne alınarak bunlar için rastgele oluşturulmuş 100 nokta için yöntemler test edilmiştir. Bir noktadan diğer bir nokta arasında geçen süre, bu iki nokta arasındaki en kısa yoldan gidildiği ve hızının da yollardaki maksimum hız miktarı ile maksimum hızın %80 i arasında rastgele atanan bir hız olmasına göre hesaplanmıştır. Örneğin 70 km/sa hız limiti bulunan bir yoldan 56 ila 70 km/sa arasında rastgele seçilmiş bir hızla gidildiği varsayılmıştır. Buna göre her bir kullanıcı 100 noktayı ortalama 7 saatlik bir gezintisi sırasında paylaşmıştır.

2.5.2. Yapılan Testler

Online algoritmada cevap süresinin offline algoritmaya göre daha kritik olması sebebiyle, online gizleme algoritmasında sadece kesişimli gizleme yaklaşımı kullanılmıştır. Bunun sebebi ayrık gizleme algoritmasının kesişimli gizlemeye göre dönüşüm aşamasında daha yavaş çalışmasıdır.

Yapılan testler sonucunda şu metrikler göz önüne alınarak ölçümler yapılmıştır; gizlenmiş bölgelerin çapı , ortalama toplam ceza ve çalışma zamanı. Ortalama toplam ceza, zaman gecikmesi ve sahte konum yaklaşımlarından doğan zaman hatası ile mesafe hatası metiklerinin toplamıdır. Her ikisininde zaman olarak ölçülmesi bu şekilde toplama yaparak oluşan toplam hatayı görebilmeyi mümkün kılmaktadır.

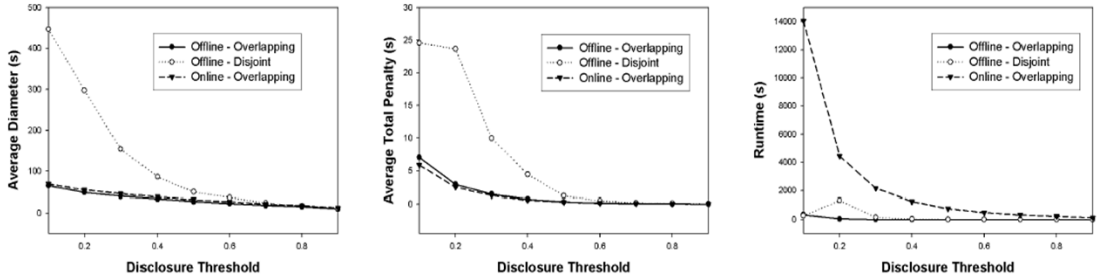
Farklı eşik değerlerinin performansa etkisi Şekil 2.12 ve Şekil 2.13'de görülmektedir. Bunlardan ilkinde mahremiyet profili $PP = \{(ibadethane\ ,\ \tau = x)\}$ iken ikincisinde mahremiyet profili $PP = \{(ibadethane\ ,\ \tau = x)\ ,\ (sağlık\ \ ,\ \tau = x)\ ,\ (eğlence\ \ ,\ \tau = x)\}$ olarak belirlenmiştir. Buradaki x değeri grafikleride verilen kordinat düzleminin x eksenine karşılık gelmektedir. Beklenildiği gibi ayrık gizleme kesişimli gizlemeye göre daha kötü sonuçlar üretmiştir. Bununla beraber ortalama çap ve ortalama toplam ceza değerlerinin offline-kesişimli ve online-kesişimli arasında farkın olmaması iki durumda da üretilecek gizlenmiş bölgelerin eşit olacağındandır. İki yöntem arasındaki fark çalışma zamanında ortaya çıkmaktadır. Sonuçlardan da anlaşılacağı üzere online metot daha yavaş çalışmaktadır.



(a) (b) (c)

$PP = \{(ibadethane\ ,\ \tau = x)\}$

Şekil 2.12: Eşik değerinin performans üzerine etkisi 1

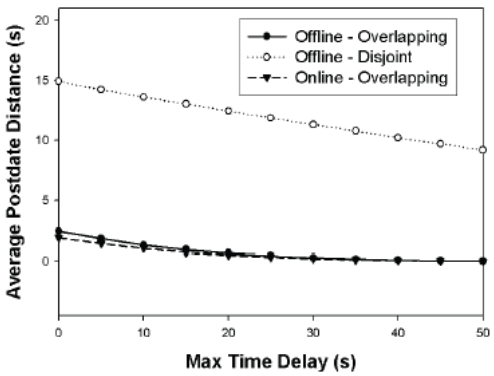


(a) (b) (c)

$PP = \{(ibadethane\ ,\ \tau = x)\ ,\ (sağlık\ ,\ \tau = x)\ ,\ (eğlence\ ,\ \tau = x)\}$

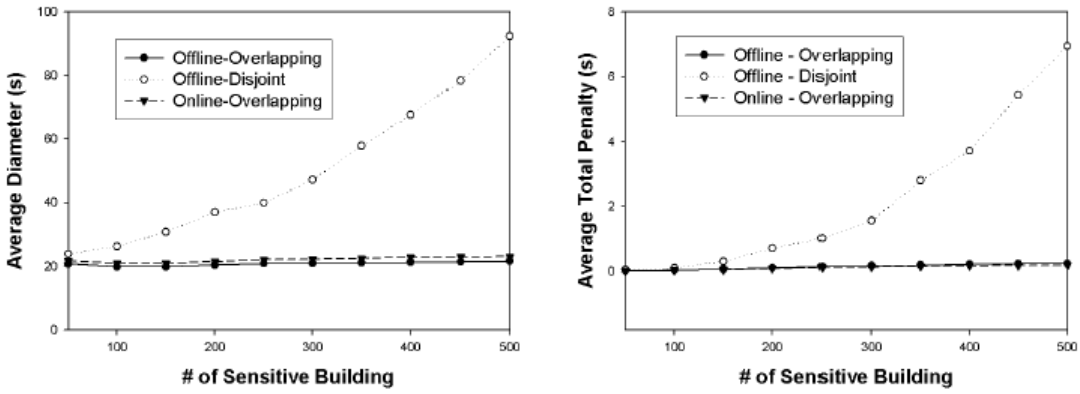
Şekil 2.13: Eşik değerinin performans üzerine etkisi 2

Dönüştürme algoritmasında hız tabanlı ataklara karşı ilk olarak zaman gecikmesi, eğer yeterli olmazsa sahte konum yaklaşımları uygulanmaktadır. Buradaki yeterlilik ölçütü belirlenen maksimum zaman gecikmesi ile belirlenmektedir. Dolayısıyla maksimum zaman gecikmesinin az veya çok olması sahte konum belirtmeyle ortaya çıkacak mesafe hatasınında az veya çok olmasını etkileyecektir. Bu etki Şekil 2.14 deki grafikte açıkça görülmektedir.

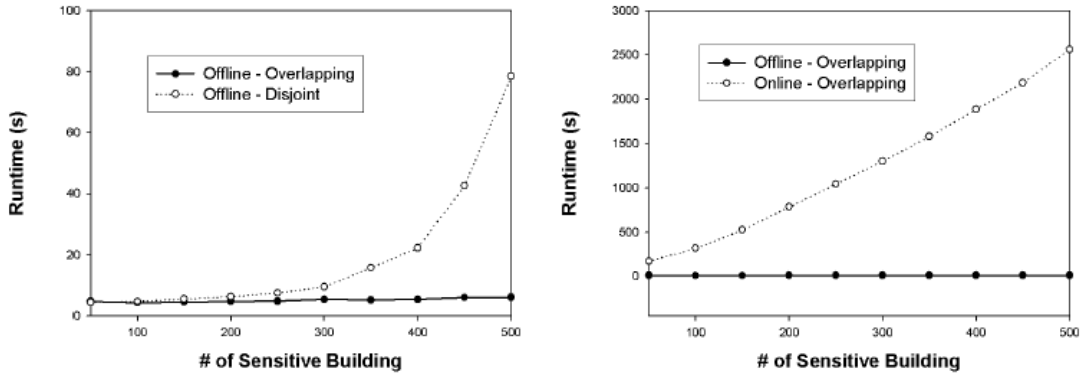


Şekil 2.14 : Maksimum zaman gecikmesinin mesafe hatasına etkisi

Duyarlı yerleşkelerin çok olması ölçülebilirliği etkileyen en önemli faktördür. Bunu test edebilmek için popüleriteleri 0.1 olan 10 farklı yerleşke türü belirlenmiştir. Herbir yerleşke türü için rasgele 50 tane yerleşke seçilmiştir. Test sırasında her seferinde yeni bir yerleşke türü daha duyarlı olacak şekilde mahremiyet profili belirlenerek aynı popülerliğe sahip sırasıyla 50, 100 … 500 yerleşke için ortalama çap, toplam ceza ve çalışma zamanı hesaplanmıştır. Böylelikle diğer tüm faktörler sabitken sadece yerleşke sayısının performansa etkisi gözlemlenebilmiştir. Buna göre Şekil 2.15'de duyarlı yerleşke sayısının ortalama gizlenmiş bölge çapı ve ortalama toplam ceza değerlerine etkisi kullanılan üç algoritma için verilmiştir. Şekil 2.16'da ise anlaşılırlığı kolaylaştırmak adına duyarlı yerleşke sayısının çalışma zamanına etkisi ilk olarak offline algoritmaların kendi arasında farkları gözlemlenmiş, ikinci kısımda ise offline ile online algoritmaların çalışma zamanındaki fark ortaya konulmuştur. Bu ölçümlerde göstermektedir ki online algoritma sağladığı yüksek mahremiyete karşın duyarlı yerleşkelerin artmasıyla performans metriklerinden düşük değerler almaktadır. Aynı durum ayrık gizleme yaklaşımında da söz konusudur.

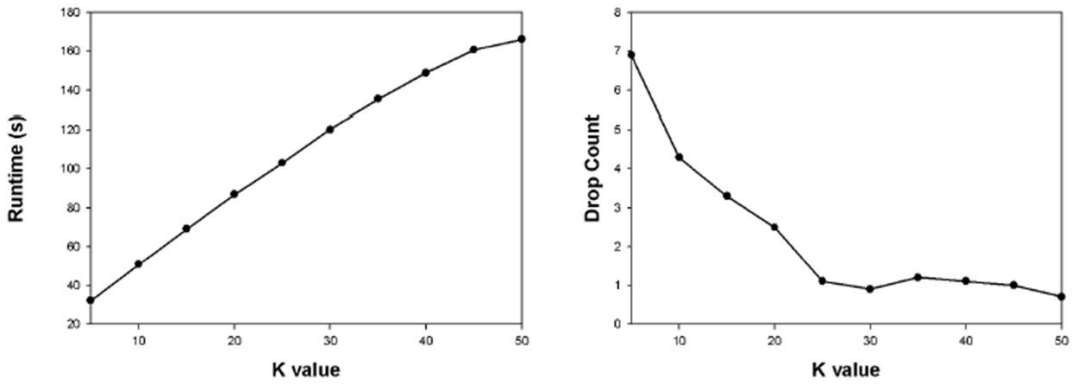


Şekil 2.15: Duyarlı yerleşke sayısının ortalama çap ve toplam ceza değerlerine etkisi



Şekil 2.16: Duyarlı yerleşke sayısının çalışma zamanına etkisi

Bir diğer değerlendirmede online algoritmada *K* değerinin etkisi üzerine olmuştur. Şekil 2.17(a) ve (b) de sırasıyla *K* değerinin değişimi ile çalışma zamanı ve istek düşürme sayısının değişimi gösterilmektedir. İstek düşürme durumu *K* sayında duyarlı bölge bulunurken bir önce yayınlanan gizlenmiş bölge içerisindeki duyarlı yerleşkenin dahil edilmemesinden kaynaklanmaktadır ve bir önceki gizlenmiş bölgenin tüm düğümlerinden bölge içerisinde bulunan bir noktaya verilen zaman aralığında gidilememesi durumunda ortaya çıkmaktadır. Grafikte de görüldüğü üzere *K* değerinin artması bir önceki CR içerisindeki duyarlı bölgenin seçilme ihtimalini arttırdığından istek düşürme sayısının azaldığı görülmektedir. Çalışma zamanı ise buna zıt olarak *K* değerindeki artışla doğru orantılı olarak artmaktadır.



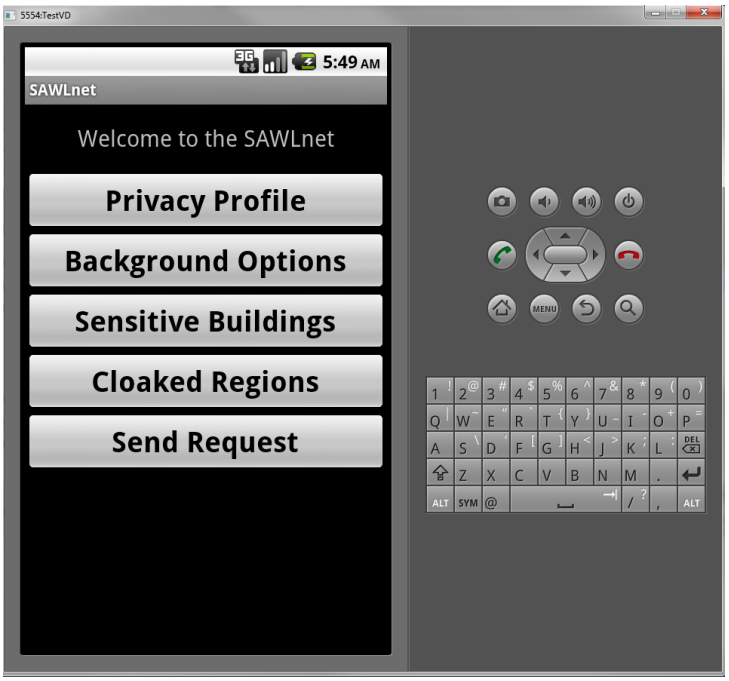
(a) (b)

Şekil 2.17: *K* değerinin etkisi

3. SAWLnet

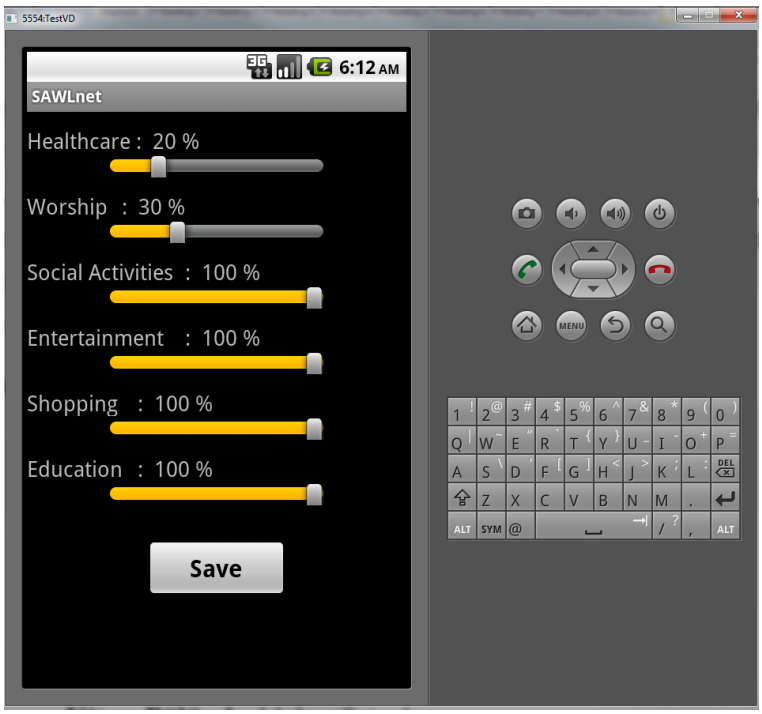
SAWLnet ile kullanıcının KTS ile yapacağı konum paylaşımından önce mahremiyet sağlama servisi ile arasındaki etkileşimin nasıl olacağı gösterilmiştir. Ayrıca kullanıcın mahremiyet profili ve kullanılan yöntemlerde yapılan değişikliklerden nasıl etkileneceği gösterimiştir. Andorid 2.2 ve daha yüksek seviye mobil cihazlar için hazırlanmış bu demo, sunucu tabanlı ve kullanıcı tabanlı mimarilere uygun olarak hazırlanmıştır. Ayrıca birden fazla kullanıcıyada eş zamanlı olarak hizmet sağlayabilmektedir. Deneysel çalışmada kullanılan veri seti demo içinde kullanılmıştır. Buna göre Milan'a ait annotated şehir bilgisi demo başlangıcında hazır olacaktır. Bu tezde SAWLnet demoya ait ekran görüntüleri simulasyondan alınan ekran görüntüleridir.

İlk olarak basit bir ana ekran görüntüsü Şekil 3.1'de gösterilmiştir. Bu ekranda kullanıcı butonları kullanarak yapmak istediği işlemi seçebilecektir. Bu ekran ile seçilen işlemler sırayla anlatılmıştır.



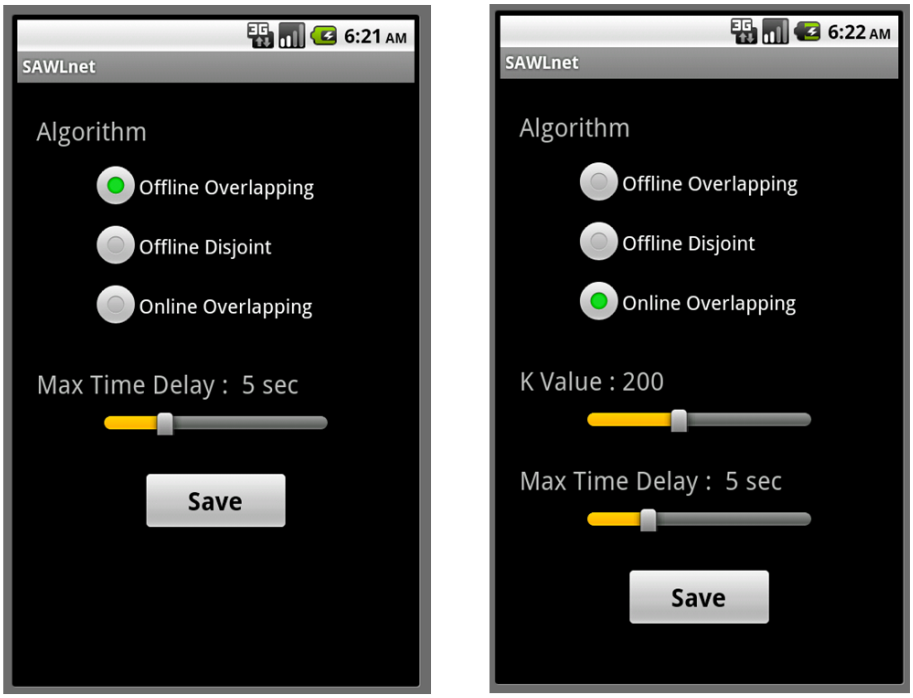
Şekil 3.1: SAWLnet ana ekran görüntüsü

Mahremiyet profili seçim ekranında bazı yerleşke türleri ve bu türlere ait eşik değerlerinin seçilebileceği bar bulunmaktadır. Bu eşik değerinin %100 olması yerleşke türünün duyarlı olmadığı anlamına gelecektir. Farklı bir değer olması ise kullanıcının o yerleşke türünde olduğunun başka bir kişi tarafından ne kadarlık bir yüzde olasılıkla tahmin edilebileceğini ifade eder. Mahremiyet seçim ekranı Şekil 3.2'de verilmiştir.



Şekil 3.2: Mahremiyet profili seçim ekranı

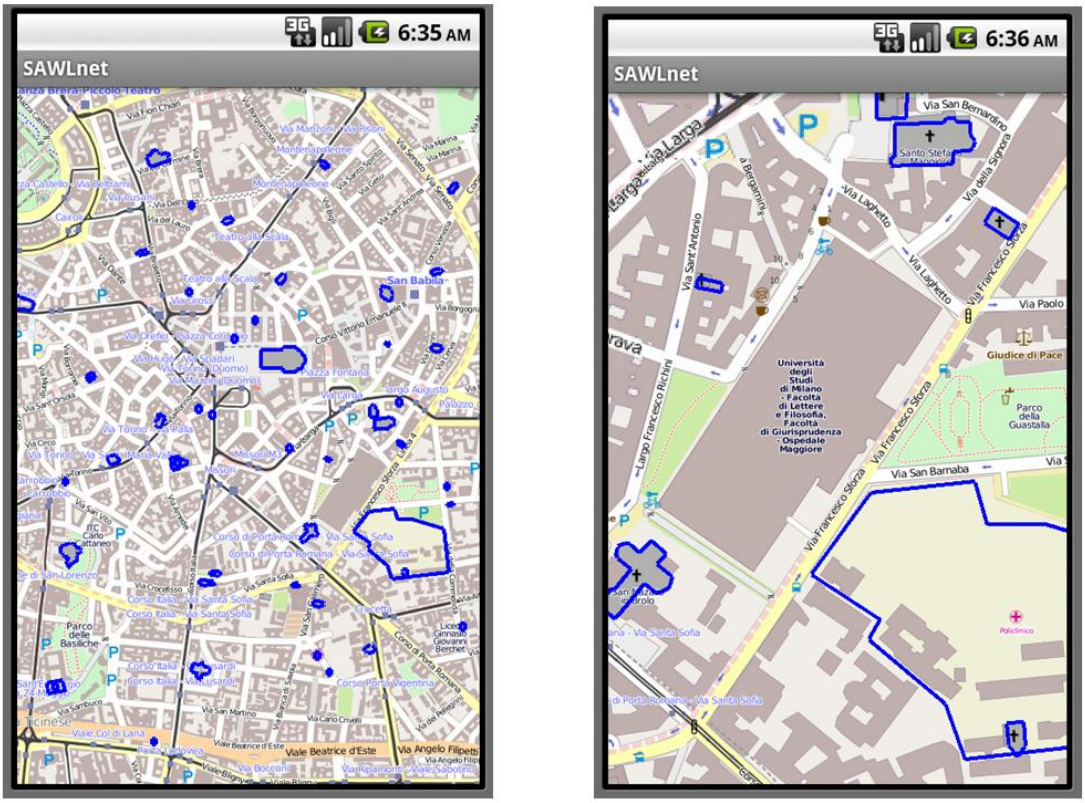
Bir diğer ekran görüntüsü kullanılan algortimaların seçim ekranıdır. Şekil 3.3'de görüldüğü üzere kullanıcı kullanılan metot ve izin verilen maksimum zaman gecikmesini seçebilecektir. Şekil 3.3(b)'de ise online algoritmaya özel olan K value seçimide eklenmiştir.



(a) (b)

Şekil 3.3: Çalışma seçenekleri seçim ekranı

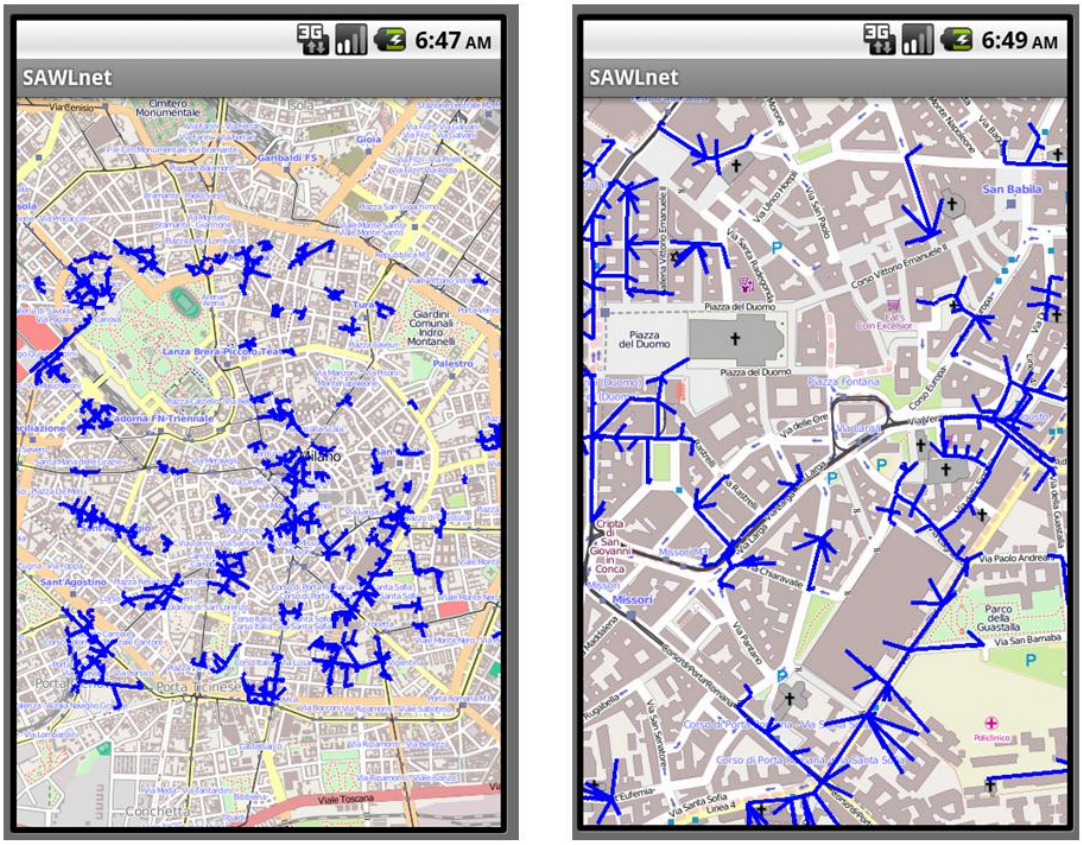
Kullanıcı duyarlı yerleşkeleri görüntüleyerek, kendi belirlediği mahremiyet profiline göre duyarlı yerleşkelerin harita üzerinde nerede olduğunu görebilmektedir. Şekil 3.4'de gösterilen ekranlar farklı yakınlık seviyesinde ibadethane ve sağlık türündeki duyarlı bölgeler gösterilmiştir.



(a) (b)

Şekil 3.4: Mahremiyet profiline göre duyarlı yerleşkelerin haritada gösterimi

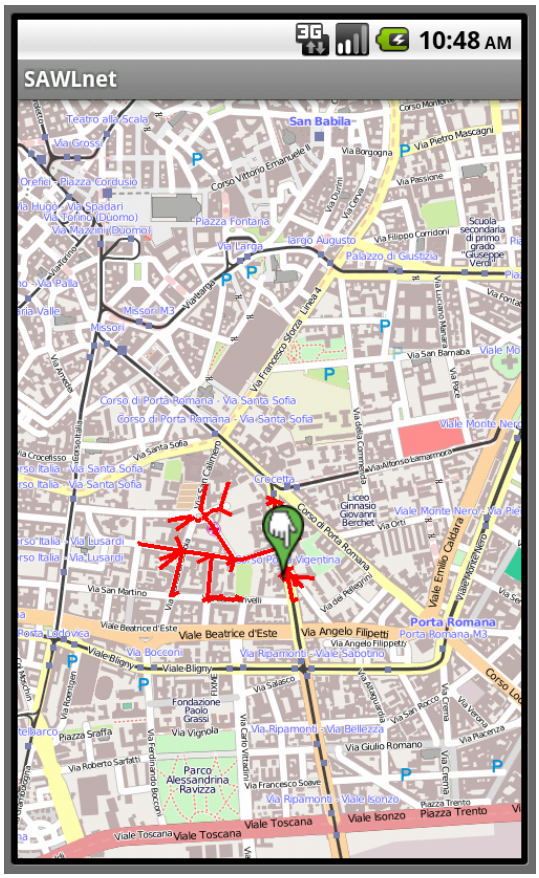
Bir diğer ekran görüntüsü Şekil 3.5'de belirtilen gizlenmiş bölgelerin harita üzerinde gösterimidir. Burada gizlenmiş bölgeler bir altçizge olarak gösterilmiş yerleşkeler karmaşıklığı azaltmak için nokta olarak ele alınmıştır. Seçilen profile göre oluşturulan tüm gizlenmiş bölgeler bu sayfada görüntülenmektedir. Eğer online algoritma seçilmişse bu sayfa açılmayacaktır, çünkü online algoritmada gizlenmiş bölgeler istek sırasında hesaplanacaktır.



(a) (b)

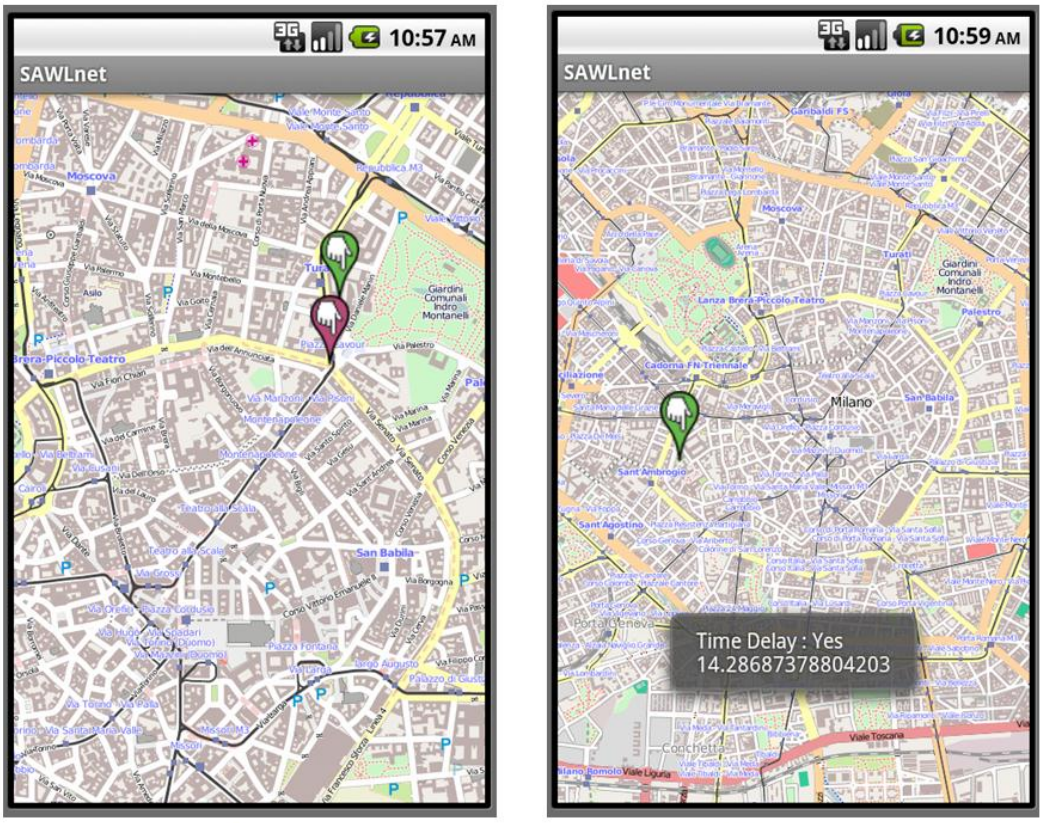
Şekil 3.5: Mahremiyet profiline göre duyarlı yerleşkelerin CR'leri

Kullanıcı son olarak istek gönder bölümünde mahremiyet sağlama servisinin konum paylaşımında nasıl bir değişiklik yaptığını gözlemleyebilecektir. Bir kullanıcı her bir istek göndermesi sırasında rastgele belirlenmiş bir noktadan diğer bir noktaya hareket ettiği varsayılacaktır. Buna göre Şekil 3.6'da görüldüğü gibi yeşil simge ile kullanıcın o anki konumu, kırmızı altçizge ise bulunduğu nokta yerine mahremiyeti açısından güvenli olan CR'nin paylaşılacağını göstermektedir. Saldırgan kullanıcının bu kırmızı bölge içerisinde olduğunu bilecek fakat tam olarak neresinde olduğu sonucuna ulaşamayacaktır.



Şekil 3.6: Gerçek pozisyon ve paylaşılan gizli bölge

Kullanıcı menu üzerinden yeni istekler gönderdiği zaman bulunduğu noktaya göre hız tabanlı ataklarda göz önüne alınarak yeni yayınlanacak bölge hesaplanacaktır. Buna göre Şekil 3.7'de farklı isteklere göre hesaplanmış paylaşılacak güvenli bölgeler gösterilmiş. (a) kısmında sahte konum metotu görünmektedir. Yeşil simge ile kullanıcının o anki konumunu ifade ederken bu noktanın paylaşılması mahremiyet gereksinimlerini karşılamamaktadır. Bu sebeple kırmızı simgeyle belirtilen noktanın kalite metriklerine göre en düşük ceza ile paylaşılacak nokta olarak belirlenmiştir. (b) kısmında ise tek bir yeşil simge görülmektedir. Bu kullanıcının bulunduğu noktanın belirtilen zaman gecikmesi içerisinde paylaşılmasının mahremiyeti sağladığı anlamına gelmektedir. Ne kadarlık bir zaman gecikmesi olduğu yeşil simge üzerine tıklanarak öğrenilebilmektedir.



(a) (b)

Şekil 3.7: Zaman gecikmesi ve sahte konum metotlarının gözlemlenmesi

4. ANLAMSAL KONUM SERİLERİNİN PAYLAŞIMINDA MAHREMİYETİN GENELLEŞTİRME VE SIRA ESNEKLEŞTİRME İLE SAĞLANMASI

4.1. Motivasyon

Günümüzde insanların konumlarının etkin bir şekilde hesaplanabilmesi ve bu konumlarının konum tabanlı servisler veya telefon şebekesi yardımıyla toplanması sonucu insanların yapmış oldukları konum değişiklikleri ile ilgili çok fazla veri toplanmaktadır. Ayrıca bu konum bilgilerinin okul, alışveriş merkezi, istasyon gibi anlamsal konumlar ile kolaylıkla ilişkilendirilmesi yapılabilecek araştırmaları daha da çeşitlendirmektedir. Bu veriler şehir planlama, trafik yönetimi, sosyal analiz gibi çok sayıda veri madenciliği uygulamasında kullanılmaktadır [36], [37], [38]. Fakat bu durum mahremiyet kavramını beraberinde getirmektedir. Konum bilgilerinin yayınlanmasından önce kimlik bilgilerinin çıkartılması çoğu zaman yeterli bir sonuç olmamaktadır. Örneğin Çizelge 4.1'de dört kişiye ait anlamsal konum serileri verilmiştir. Aranan kişinin bu veri tabanında olduğu düşünüldüğünde ve kişinin ziyaret ettiğini bildiği yerleri kullanarak hangi serinin kime ait olduğu sonucuna ulaşabilir. Örnek olarak saldırgan kafe ve tiyatroya gittiğini bildiği bir kişinin veri tabanında sadece u3 kişisine karşılık geldiğini görecek ve aynı kişinin hastaneye de gittiği sonucuna ulaşacaktır.

Çizelge 4.1: Anlamsal konum serileri veritabanı örneği

u_1	kafe banka sinema okul market
u_2	park restoran salon kafe müze
u_3	kafe market okul tiyatro metro hastane
u_4	hastane müze havaalanı otel

Bu sebeple veri yayınlanmadan önce veri üzerinde bazı değişiklikler yaparak hangi serinin kime ait olduğunun kesin olarak belirlenememesi amaçlanmıştır. Eldeki bilgi

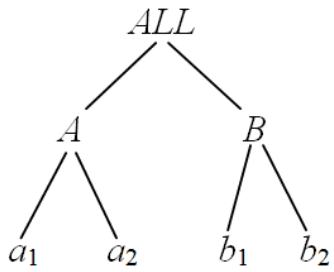
ile *k* tane seriye ulaşılmasını amaçlayan *k*-Anonimlik genel olarak genelleştirme [39], [40], [41] ve yok etme [42] üzerine kurulmuştur. Genelleştirme verinin kesinliğini bozmasına rağmen veri hakkında yinede bilgi vermektedir. Bunun yanında yok etme veri hakkında bilgi sağlamayı tamamen engellemektedir. Fakat sırasal verilerde sadece genelleştirme kullanımı daha fazla genelleştirme gerektireceği için küme verilerinde olduğu kadar iyi sonuçlar vermemektedir. Bu sebeple bazı verilerin sırasını belirsizleştirerek daha az genelleştirme ile verinin daha az bozulması önerilmiştir.

4.2. Model

(k,d)-Anonimlik: d tane konum bilgisine sahip bir kişi yayınlanan veri tabanı içerisinde en az k tane farklı seri ile karşılaşacağının ifadesidir.

4.2.1. Genelleştirme

Genelleştirme işlemi anlamsal hiyerarşi temeline dayanmaktır. Örnek bir genelleştirme hiyerarşisi Şekil 4.1'de görülmektedir. Bu hiyerarşi kuralına göre $\{a_1, a_2\} \rightarrow A$, $\{b1 , b2\} \rightarrow B$ ve $\{A, B\} \rightarrow ALL$ şeklindedir.



Şekil 4.1: Örnek hiyerarşi

Örnek olarak Çizelge 4.2(a)'da örnek bir D veri tabanı verilmiştir. Buna göre $\{a_1, a_2\} \rightarrow A$ kuralı bu veritabanına uygulandığında sonuç veritabanı D' Çizelge 4.2(b)'de verilmiştir. Buna göre tüm veri tabanındaki a_1 ve a_2 genelleştirilerek A ile değiştirilmiştir.

Çizelge 4.2 Örnek veritabanı dönüşümü

id	contents
S_1	$a_1\ a_1$
S_2	$a_2\ a_2$
S_3	$a_2\ b_1\ b_2\ b_1$
S_4	$b_2\ a_2\ b_1\ b_1$

(a) Orijinal Veritabanı (D)

id	contents
S_1	$A\ A$
S_2	$A\ A$
S_3	$\{A\ b_1\ b_2\}\ b_1$
S_4	$\{A\ b_1\ b_2\}\ b_1$

(b) Dönüştürülmüş Veritabanı (D')

4.2.2. Sıra Esnekleştirme

Kullanıcıların bulundukları konum bilgileri, yol veritabanı D içerisinde zamana göre sıralı bir şekilde yer almaktadır. Bu da saldırganın daha az bir bilgiyle kayıdın kime ait olduğu bilgisine ulaşmasına imkan sağlayacaktır. Eğer Çizelge 4.2'deki örnekte sadece genelleştirme ile mahremiyet sağlanmaya çalışılsaydı tüm konumların ALL ile genelleştirilmesi gerekecek ve yayınlanacak verinin kalitesi çok düşük olacaktır. Bu noktadan hareketle veritabanın bazı I aralıklarının konum sırası belirsizleştirilerek saldırganın belli sayıdaki konumun sırasının gerçekte olmasa bile birden fazla seride karşılaşması sağlanarak anonimlik sağlanacaktır. Örneğin Çizelge 4.2'de S'_3 ve S'_4 serilerinin ilk üç konumun $I = [1,3]$ sırası belirsizleştirilerek bu aralıktaki konumların sırası üç noktanın herhangi bir sırada olabileceği anlamına gelmektedir. Böylelikle bir kullanıcıya ait $b_1\ b_2$ konum ve sıra bilgisine sahip bir saldırgan orijinal veritabanında bu kullanıcın S'_3 serisine karşılık geldiği bilgisine ulaşabilirken sıra esnekleştirme tekniği ile dönüştürülmüş veritabanında kullanıcıya karşılık gelen serinin S'_3 veya S'_4 olduğuna karar veremeyecektir. Bunun yanı

sıra dönüştürülmüş veritabanı halen araştırmalarda kullanılabilecek veriler sunmaktadır.

4.2.3. Bilgi Kaybı Metriği

Uygulanan genelleştirme ve sıra esnekleştirme teknikleri orijinal veritabanında bazı bilgi kayıplarına ve dolayısıyla verinin kalitesinde azalmaya sebep olacaktır. Bu iki metriğin birlikte kullanımıyla ortaya çıkacak bilgi kaybını ölçmek için yeni bir metrik ortaya konulması gerekmektedir. Verilen bir $S = s_1 s_2 \ldots s_{|S|}$ serisinin S' serisine dönüştürülürken s_i konumunun genelleştirilmesi $g(s_i)$ ile, uygulanan sıra esnekleştirme kümesi de $\{ I_1, I_2, \ldots, I_r \}$ ile ifade edilmektedir. Buna göre tek bir seri için bilgi kaybı metriği LCP şu şekilde ifade edilmektedir:

$$LCP(S') = \begin{cases} \frac{\prod_{i=1}^{|S|} l(g(s_i)) \cdot \prod_{j=1}^{r} p(I_j)}{l(\mathbb{I})^{|S|}} & , \ \exists l\big(g(s_i)\big) > 1 \ or \ r > 0 \\ 0 & , \ otherwise. \end{cases} \quad (4.1)$$

$l(g(s_i))$, hiyerarşi ağacında $g(s_i)$ düğümünün altındaki yaprak sayısına karşılık gelmektedir. Eğer s_i bir yaprak düğüm ise $l(g(s_i)) = 1$ olacaktır. $p(\ Ij\ )$ ise Ij aralığındaki noktaların permütasyon sayısını ifade etmektedir. Örnek olarak $S = a_1 b_1 b_2 a_1 b_1 a_2$ serisinin $S' = A\{\ Ab_1 b_1 b_2\} A$ serisine dönüştürülmesi ile oluşacak bilgi kaybı şu şekilde hesaplanır; ilk olarak $\{\ a_1, a_2\} \rightarrow A$ genelleştirmesiyle ortaya çıkan 2.1.1.2.1.2 = 8 farklı olasılık $l(g(s_i))$ ve $I_1 = [2,5]$ sıra esnekleştirmesinden kaynaklanan $\frac{4!}{1!.2!.1!} = 12$ farklı permutasyon $p(\ Ij\ )$ hesaplanır. Daha sonra bunların çarpımının olabilecek maksimum permütasyon olan $l(\mathbb{I})^{|S|} = 4^6 = 4096$ ile bölümü ile S' serisinin belirsizlik cezası olan $CP(S') = \frac{8.12}{4096} = 0.023$ hesaplanır. Görüldüğü üzere belirsizlik cezası genelleştirme ve sıra esnekleştirme tekniklerinin eş zamanlı kullanımı sırasında ortaya çıkan belirsizliği ölçmektedir. Tek bir seri için hesaplanan bu belirsizlik tüm veritabanı D' için hesaplanırken şu şekilde hesaplanır:

$$GCP = \frac{\sum_{s' \in D'} |S'|.CP(S')}{\sum_{s' \in D'} |S'|} \quad (4.2)$$

Böylelikle bilgi kaybı metriğinin sonucu normalleştirilerek 0 ile 1 arasında bir değer olacaktır.

4.3. Algoritma

Saldırganın kullanıcıya ait bulunduğu d tane konumu ve sırasını bildiği düşünüldüğünde, dönüştürülmüş veritabanında gidilebilecek tüm konumların d permütasyonlarının herbirinin k kişide bulunması gerekmektedir. Bunun her permütasyon için test edilmesi ve eğer koşulu sağlanmıyor ise seri esnekleştirme veya genelleştirme metodlarından hangisinin daha az bilgi kaybına yol açacağı hesaplanıp buna göre bir dönüştürme yapılması çok fazla süre almaktadır. Bu sebeple bu işlemleri gerçekleştirirken bazı yöntemler uygulanarak hesaplama süresi azaltılmıştır.

Bunlardan ilki olan *arttırımsal yaklaşım;* veri tabanının ilk olarak birli permütasyonlarının, daha sonra ikili permütasyonların olacak şekilde d'li permütasyonlara varıncaya kadar her seviye için test edilmesi temeline dayanmaktadır. Bunun sebebi daha düşük seviyelerdeki permütasyonlarda bile k sayısına ulaşılamıyor ise genelleştirme veya sıra esnekleştirmenin daha önce yapılarak bir üst seviyedeki permütasyon sayısının azaltılması amaçlanmıştır. Çizelge 4.2'deki örnek olarak ele alındığında $\{a_1 b_1\}$, $\{a_2 b_1\}$, $\{b_1 a_2\}$, gibi ikili permütasyonlarının $k=2$ kadar bulunup bulunmadığı araştırılmasından önce birinci seviyede a_1 konumunun sadece bir kullanıcıda olduğu görüldüğünde $\{a_1, a_2\} \rightarrow A$ genelleştirmesi yapılacağı görülecektir. Bu sebeple ikili permütasyonlara geçildiğinde a_1, a_2 konumlarının ayrı ayrı yer aldığı permütasyonlar A yer değiştirilerek permütasyon sayısı azalacaktır.

Bir diğer yöntem *Apriori* yaklaşımıdır. Bu yöntem sayesinde bir önceki seviyede bulunamayan permütasyonların bir sonraki seviyedeki permütasyonların içerisinde bulunmamasını sağlayarak permütasyon sayısı azaltılmaktadır.

Kullanılan bir diğer yöntem de permütasyonların veri tabanında hangi serilerde bulunduğunun belirlenmesini hızlandıran *bit bazlı* operasyonlardır. Buna göre Çizelge 4.3'de olduğu gibi ilk seviyede hangi konumun hangi seride bulunduğu bilgisi 0 ve 1 mantık kuralıyla oluşturulan sayılarda tutulur. Çizelge 4.3 (a)'da verilen verilen veri tabanındaki konumların hangi serilerde olduğu bilgisinin ikili tabanda nasıl tutulduğu Çizelge 4.3 (b) ' de verilmiştir. Bir sonraki aşamaya geçilirken bu değerlerin mantıksal VE işlemi uygulanarak yeni permütasyonun hangi seride olduğu bilgisi kolaylıkla ulaşılacaktır. Çizelge 4.3 (c)'de de örnek olarak verildiği üzere $\{a_1,a_2\}$ ikilisinin hangi serilerde bulunduğunu bulmak için a_1 ve a_2 noktalarının bulundukları serileri belirten ikili sayılar üzerinde mantıksal VE işlemi uygulandığında ikisinin beraber bulunduğu bir seri olmadığı sonucu ulaşılmış olacaktır. Fakat burada dikkat edilmesi gerek konu {a1,a1} gibi aynı değerlerin permütasyonda olması ikili işlem sonucunda bulunduğu bilgisine ulaşılan serilerin yeniden test edilmesi gerekmektedir.

Çizelge 4.3: Bit Bazlı Operasyonlar

S_1	a_1 a_1
S_2	a_2 a_2
S_3	a_2 b_1 b_2 b_1
S_4	b_2 a_2 b_1 b_1

(a)

a_1	1 0 0 0
a_2	0 1 1 1
b_1	0 0 1 1
b_2	0 0 1 1

(b)

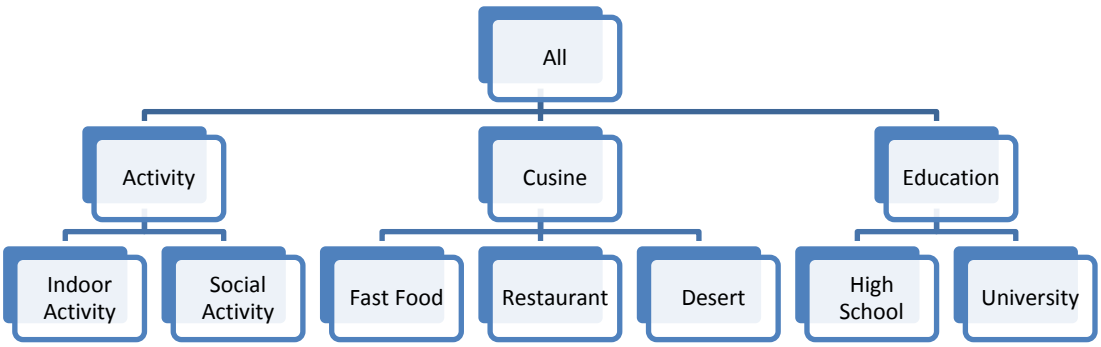
a_1 a_2	0 0 0 0
a_2 b_1	0 0 1 1

(c)

Bir diğer kullanılan performans arttırıcı yöntem ise genelleştirme ve sıra esnekleştirmeden doğacak bilgi kaybının herseferinde hesaplanmasının yerine sıra esnekleştirmenin genel olarak daha az bilgi kaybına sebep olduğu varsayımından yola çıkarak eğer sıra esnekleştirme mahremiyet gereksinimini karşılıyor ise bunun uygulanması, aksi takdirde genelleştirmeye başvurulmasıdır.

4.4.Deneysel Çalışmalar

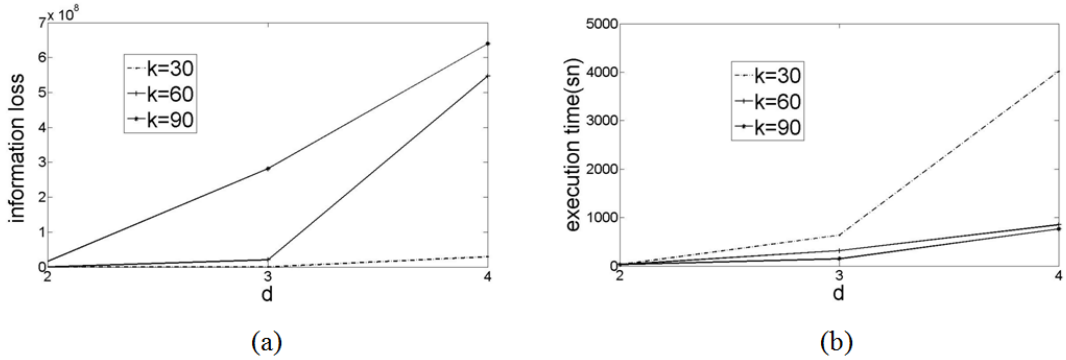
Deneysel çalışmalar için Gowalla [43] veri seti kullanılmıştır. Bu veri seti insanların bulundukları noktaları paylaşması sonucu hazırlanmış bir veri setidir. New York şehrindeki kullanıcıların hareketlerini içeren bu veri setinden 1000 kullanıcının 6343 farklı noktaya yapmış olduğu 21045 konum paylaşımı kullanılmıştır. Bu 6343 farklı nokta anlamsal olarak 30 farklı grup altında toplanmıştır. Ayrıca bu gruplarda belli bir anlamsal hiyerarşi içerisinde yer almaktadır. Genelleştirme aşamasında kullanılacak bu hiyerarşinin küçük bir kesimi Şekil 4.2'de verilmiştir.



Şekil 4.2: Kullanılan hiyerarşiden bir kesit

Kullanılan yöntem için k ve d değerlerinin değişimine bağlı olarak çalışma zamanı ve bilgi kaybı değerleri test edilmiştir. İlk olarak d değerinin artımına paralel olarak bilgi kaybı ve çalışma zamanı değerleri de artmaktadır. Şekil 4.3 (b)'de görüldüğü gibi artan her seviye için test edilecek permütasyon

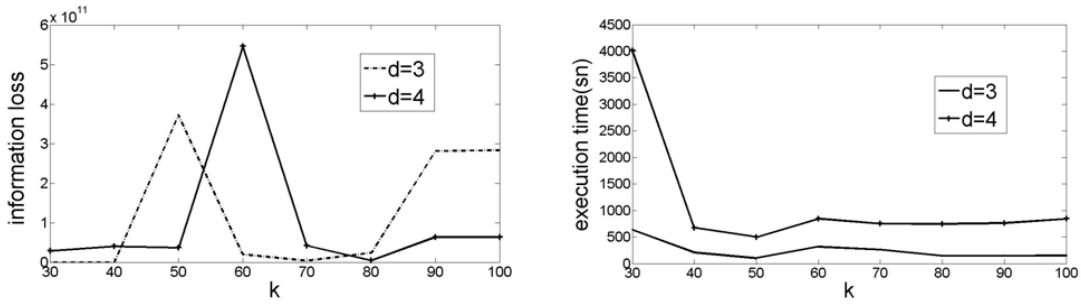
sayısının artması sonucu çalışma zamanı da artmaktadır. Aynı şekilde Şekil 4.3 (a)'da artan her seviye için mahremiyet gereksinimi artacağından bilgi kaybının arttığı görülmektedir.



(a) (b)

Şekil 4.3. *d* değerinin çalışma performansına etkisi

Şekil 4.4 (a) ve Şekil 4.4 (b) 'de ise sırasıyla *k* değerine bağlı olarak bilgi kaybı ve çalışma zamanı değerlerinin grafiği gözükmektedir. Burdaki dalgalanmaların temel sebebi bilgi kaybının karşılaştırılmadan ilk olarak sıra esnekleştirmenin yapılmasıdır.



(a) (b)

Şekil 4.4. *k* değerinin çalışma performansına etkisi

5. SONUÇ

Günümüzde konum tabanlı servislerin çok fazla artması, yapmış oldukları konum paylaşımlarından dolayı kullanıcılarının mahremiyeti konusunda endişelerin artmasına sebep olmuştur. Kişiye ait konum verilerinin güvenilmeyen kişilerin eline geçmesi istenmeyen sonuçlar doğurmaktadır. Konum tabanlı servislerin çeşitliliği konum mahremiyeti sağlama yaklaşımlarının da çeşitlenmesine sebep olmuştur. Bu tezde iki farklı problemin çözümü için mahremiyet sağlama yaklaşımı önerilmektedir.

İlk olarak KTS kullanıcılarının konumlarına bağlı servis alırlarken yapmış oldukları konum paylaşımı sırasında konum bilgisinin istenmeyen kişilere geçmesini önlemeye yönelik yaklaşım sunulmuştur. Burada arkadaş bulma gibi servisler göz önünde bulundurularak kullanıcının kimlik bilgisinin de paylaşılması, konum bazında bir gizleme yapmayı gerektirmiştir. Yapılan bu konum gizleme sırasında park, market, okul gibi anlamsal konum bilgilerini de değerlendirilerek, kişinin beklediği mahremiyet sağlanmış olmaktadır. Ayrıca bir yerleşim yeri içerisinde bulunan yol bilgisinin herkes tarafından bilinebiliyor olması sonucu gizlenmiş bölgelerin saldırgan tarafından hız tabanlı ataklar gibi ataklarla bilgi sızdırması sebebiyle yerleşim yeri yollar ve alanlardan oluşan bir çizge olarak ele alınmış ve daha güvenli bir mahremiyet sağlanmıştır. Tez kapsamında bu yaklaşım temel alınarak çeşitli gizlenmiş bölge oluşturma ve konum paylaşımı sırasında dönüştürme algoritmaları geliştirilmiştir. Bununla birlikte bir Android uygulaması ile yaklaşımın gerçek yaşamda nasıl kullanılacağı gösterilmiştir.

Kullanıcının bir noktada uzun bir süre bulunması gizlenmiş bölge içersindeki bazı bölgelerde olamayacağı bilgisine ulaşılmasına sebep olabilir. İleriki çalışmalarda bulunma zamanını da katarak daha güçlü bir mahremiyet sağlanmış olacaktır.

Tezin ikinci kısmında ise insanların anlamsal konum bilgilerinin araştırma amaçlı yayınlanması sırasında ortaya çıkan mahremiyet problemi bir öneri getirmektedir. Kimlik bilgisi çıkartılmış veriler üzerinde belli bir bilgiye saldırganın veri

tabanındaki hangi seriye ait olduğunun anlaşılması engellenmedir. Bu sebeple ortaya atılan öneri ile anlamsal konum bilgilerinin belli bir hiyerarşiye göre genelleştirilmesi ve konumların bir kesiminin sırasının belirsizleştirmesi yaklaşımlarının beraber kullanılarak minimum bilgi kaybıyla mahremiyet sağlanmak amaçlanmıştır.

Ortaya atılan algoritma mahremiyeti sağlamasına rağmen çözüm uzayının çok fazla olması sebebiyle veri tabanı tarama işlemi zaman almaktadır. Bunun için geliştirilecek yeni sezgisel yaklaşımla bunların sayısı azaltılmaya çalışılacaktır. Ayrıca tüm veri tabanının hesaplanması uzun sürdüğü durumlarda serilerin benzerlikleri hesaplanarak gruplandırılıp daha küçük veri kümelerinde çalışması ileriki yapılacaklar arasındadır.

KAYNAKLAR

[1] Virrantaus, K., Veijalainen, J., Markkula, J., Katasonov, A., Garmash, A., Tirri, H., Terziyan, V., Developing GIS-Supported Location-Based Services. In: Proc. of WGIS'2001 – First International Workshop on Web Geographical Information Systems. Kyoto, Japan, 423–432, 2001.

[2] Shek, S., Next-generation Location-Based Services for mobile devices, CSC Grants, 2011.

[3] M. L. Damiani, C. Silvestri, and E. Bertino, "Fine-Grained Cloaking of Sensitive Positions in Location-Sharing Applications. IEEE Pervasive Computing," IEEE Pervasive Computing, vol. 10(4), pp. 64–72, 2011.

[4] G. Ghinita, M. L. Damiani, C. Silvestri, and E. Bertino, "Preventing velocity-based linkage attacks in location-aware applications," in Proc. 17th ACM GIS, 2009.

[5] "Milano Street Map Dataset" erişim adresi: http://www.openstreetmap.com,

[6] S. Steiniger, M. Neun, and A. Edwardes. Foundations of Location Based Services. Lecture Notes on LBS, Department of Geography, University of Zürich, 2006.

[7] Werbach, K., Location-Based Computing: Wherever You Go, There You Are, Esther Dyson's Monthly Report, 18(6), 1-31, 2000**.**

[8] C. A. Ardagna, M. Cremonini, S. De Capitani di Vimercati, and P. Samarati, "Location Privacy in Pervasive Computing,", Security and Privacy in Mobile and Wireless Networking, 2008.

[9] A. Beresford and F. Stajano. Mix zones: User privacy in location-aware services. In IEEE Workshop on Pervasive Computing and Communication Security (PerSec), 2004.

[10] J. Freudiger, M. Raya, and M. Feleghhazi, "Mix zones for location privacy in vehicular networks," in Proc. Int. Workshop WiN-ITS, Vancouver, BC, Canada, Aug. 2007.

[11] J. Freudiger, R. Shokri, and J.-P. Hubaux. On the optimal placement of mix zones. PETS, 2009.

[12] Butty´an L., Holczer T., Vajda I., On the Effectiveness of Changing Pseudonyms to Provide Location Privacy in VANETs, ESAS, 2007

[13] Samatri P., Sweeney L., Protecting privacy when disclosing information_ k-anonymity and its enforcement through generalization and suppression, Technical Report, 1998.

[14] Sweeney L., k-Anonymity: a Model for Protecting Privacy, IEEE Security and Privacy, 1998

[15] Gruteser M., Grunwald D., Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking, The First International Conference on Mobile Systems, Applications, and Services, 2002.

[16] B. Gedik and L. Liu, Location Privacy in Mobile Systems: A Personalized Anonymization Model, Proc. 25th Int'l Conf. Distributed Computing Systems (ICDCS '05), pp. 620-629, 2005.

[17] Bamba B., Liu L., Pesti P., Wang T., Supporting Anonymous Location Queries in Mobile Environments with PrivacyGrid, World Wide Web Conference, 2008

[18] Mokbel M.F., Walid C.C., Aref w.G., The New Casper: Query Processing for Location Services without Compromising Privacy, Very Large Data Base Endowment, 2006.
[19] Wang T., Liu L., PrivacyAware Mobile Services over Road Networks, Very Large Data Base Endowment, 2009
[20] B. Palanisamy and L. Liu, Mobimix: Protecting Location Privacy with Mix-zones Over Road Networks, IEEE 27th International Conference on Data Engineering, 2011.
[21] M. Gruteser, J. Bredin, and D. Grunwald. Path Privacy in Locationaware Computing. Workshop on Context Awareness, MobiSys 2004 , Boston, US, 2004.
[22] Gruteser M., Liu X., Protecting Privacy in Continuous Location-Tracking Applications. IEEE Security & Privacy 2(2): 28-34, 2004
[23] Hoh, B., Gruteser, M., Protecting Location Privacy Through Path Confusion. SECURECOMM, 194-205, 2005
[24] Duckham M., Kulik L., A Formal Model of Obfuscation and Negotiation for Location Privacy, In Pervasive Computing, Vol. 3468, 152-170, 2005
[25] Myles G., Friday A., Davies N., Preserving Privacy in Environments with Location-based Applications, Pervasive Computing, IEEE, 56-64, 2003
[26] Youssef M., Atluri V., Adam N.R., Preserving Mobile Customer Privacy: An Access Control System for Moving Objects and Customer Profiles, IEEE MDM, 67-76, Cyprus, 2005
[27] Hengartner U., Steenkiste P., Access Control to People Location Information, ACM Transactions on Information and System Security (TISSEC), 424-456, 2005
[28] Sneekkenes E., Concepts for Personal Location Privacy Policies, ACM Conference on Electronic Commerce (EC'01), Florida, USA, 2001
[29] Leonhardt U., Magee J., Security Considerations for a Distributed Location Service, Journal of Network and System Management, 51-70, March 1998
[30] Poolsappasit N., Ray I., Towards Achieving Personalized Privacy for Location-Based Service, Transaction on Data Privacy, 77-99, 2009
[31] Hauser C., Kabatnik M., Towards Privacy Support in a Global Location Service, IFIP Workshop on IP and ATM Traffic Management, 2001
[32] Hengartner U., Steenkiste P., Protecting Access to People Location Information, International Conference on Security in Pervasive Computing – SPC , 25-38, 2003
[33] Hong d., Yuan M., Shen V.Y., Dynamic Privacy Management: A Plug-in Service for the Middleware in Pervasive Computing, 7th International Conference on Human Computer Interaction with Mobile Devices & Services, 1-8, 2005
[34] Langheinrich m., A Privacy Awareness System for Ubiquitous Computing Environments, Ubiquitous Computing International Conference, Göteborg, Sweden, 2002
[35] Damiani M.L., Bertino E., Silvestri C., The PROBE Framework for the Personalized Cloaking of Private Locations, Transactions on Data Privacy 3(2), 123-148, 2010
[36] Ying J.J.C., Lu E.H.C., Lee W.C., Weng T.C., Tseng V.S., Mining User Similarity from Semantic Trajectories, 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks , 19-26, San Jose, California, 2010

[37] Ying J.J.C., Lee W.C., Weng T.C., Tseng V.S., Semantic Trajectory Mining for Location Prediction, 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Chicago, Illinois, 2011

[38] Ying J.J.C., Lu E.H.C., Kuo W.N., Tseng V.S., Urban Point-of-Interest Recommendation by Mining User Check-in Behaviors, ACM SIGKDD International Workshop on Urban Computing, Beijing, China, 2012

[39] Terrovits M., Mamoulis N., Kalnis P., Privacy-Preserving Anonymization of Set-Valued Data, VLDB Endowment, 115-125, 2008

[40] He Y., Naughton J.F., Anonymization of Set-valued Data via top-down Local Generalization, Proceedings of the VLDB Endowment, 934-945, 2009

[41] Monreale A., Trasarti R., Pedreschi D., Renso C., Bogorny V., C-Safety: A Framework for the Anonymization of Semantic trajectories, Transactions on Data Privacy, 73-101, 2011

[42] Xu Y., Wang K., Fu A.W.C., Yu P.S., Anonymizing Transaction Databases for Publication, 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, 2008

[43] "Gowalla veri seti" erişim adresi: http://code.google.com/p/locrec/

[44] Yigitoglu, E., Damiani, M.L., Abul, O., Silvestri, C., Privacy-Preserving Sharing of Sensitive Semantic Locations Under Road-Network Constraints. 13th International IEEE Mobile Data Management, Bengaluru, India, July 2012.

[45] Silvestri, C., Yigitoglu, E., Damiani, M.L., Abul, O., SAWLnet: Sensitivity AWare Location cloaking on road-NETworks. 13th International IEEE Mobile Data Management, Bengaluru, India, July 2012.

ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı, adı : YİĞİTOĞLU, Emre

Uyruğu : T.C.

Doğum tarihi ve yeri : 01.11.1986 Karaman

Medeni hali : Bekar

Telefon : 0 (312) 292 42 95

e-mail : eyigitoglu@etu.edu.tr

Eğitim

Derece	Eğitim Birimi	Mezuniyet tarihi
Lisans	Hacettepe Üniversitesi/Bilgisayar	2009

İş Deneyimi

Yıl	Yer	Görev
2010-2012	TOBB Ekonomi ve Teknoloji Üniversitesi	Araştırma Görevlisi

Yabancı Dil

İngilizce

Yayınlar