ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

GİZLİLİK KORUYAN k - NN MEKÂNSAL ENTERPOLASYON YÖNTEMİ

Muhammad Rifthy KALIDEEN

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

ANKARA
2018

TEZ ONAYI

Muhammad Rifthy KALIDEEN tarafından hazırlanan **"Gizlilik Koruyan k - NN Mekânsal Enterpolasyon Yöntemi"** adlı tez çalışması 27/09/2018 tarihinde aşağıdaki jüri tarafından oy birliği ile Ankara Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Danışman : Dr. Öğr. Üyesi Bülent TUĞRUL
Ankara Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı

Jüri Üyeleri :

Başkan : Prof. Dr. Fatih Vehbi ÇELEBİ
Ankara Yıldırım Beyazit Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı

Üye : Doç. Dr. Semra GÜNDÜÇ
Ankara Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı

Üye : Dr. Öğr. Üyesi Bülent TUĞRUL
Ankara Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı

Yukarıdaki sonucu onaylarım.

Prof. Dr. Atila YETİŞEMİYEN
Enstitü Müdürü

ETİK

Ankara Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım bu tez içindeki bütün bilgilerin doğru ve tam olduğunu, bilgilerin üretilmesi aşamasında bilimsel etiğe uygun davrandığımı, yararlandığım bütün kaynakları atıf yaparak belirttiğimi beyan ederim.

27/09/2018

Muhammad Rifthy KALIDEEN

ÖZET

Yüksek Lisans Tezi

GİZLİLİK KORUYAN k - NN MEKÂNSAL ENTERPOLASYON YÖNTEMİ

Muhammad Rifthy KALIDEEN

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Bülent TUĞRUL

Son zamanlarda Veri Madenciliğinde Gizlilik Koruması (Privacy – Preserving Data Mining - PPDM), bulut bilişim teknolojisinin ortaya çıkmasıyla daha popüler hale geldi. Bulut bilgi işlem, veri sahiplerinin verilerini depolamak ve işlemek için dış kaynak kullanımına izin verir. Ancak bulutta depolanan veriler gizlilik ve güvenlikten yoksundur. Bunun sonucu olarak, kredi kartı kayıtları, tıbbi kayıtlar gibi hassas verilerle ilgilenen kuruluşlar, verileri bulutlara aktarmak konusunda isteksizdir. Bu nedenle, verilerin bir bulutta dış kaynak oluşturmadan önce şifrelenmesi ve şifrelenmiş veriler üzerinde yapılması gereken işlemlerin yapılması gerekir. Güvenli Çok Taraflı Hesaplama (Secure Multiparty Computation - SMC) ve Homomorphic Şifreleme (Homomorphic Encryption - HE) algoritması, bulutun şifrelenmiş metninler üzerinde sorgular oluşturulmasına izin verir ve bulut servis sağlayıcısından ve diğer taraflara (kendi özel verileri hariç) veri, sorgu ve erişim modelini korur. sorgulama işlemi. k - En Yakın Komşu (k – Nearest Neighbor) algoritması, veri madenciliğinde benzerlik eşleşmesini bulmak için en basit ve en çok kullanılan mekansal enterpolasyon yöntemidir. Önerilen çözüm, en iyi performansı elde etmek için Paillier şifreleme sistemini kullanır. Bunun dışında kd - tree ve R - tree algoritmaları, buluttaki verileri en yakın komşuluğu daha hızlı bulmak için depolamak için kullanıldı. Önerilen çözümler, işlemin her senaryosundaki verilerin gizliliğini ve güvenliğini sağladıkları analiz edilmiştir.

Eylül 2018, 38 Sayfa

Anahtar Kelimeler: gizlilik - korunma, k - en yakın komşu, veri madenciliği, mekansal enterpolasyon, güvenli çoklu parti hesaplaması

ABSTRACT

Master Thesis

PRIVACY – PRESERVING k-NN SPATIAL INTERPOLATION METHOD

Muhammad Rifthy KALIDEEN

Ankara University
Graduate School of Natural and Applied Sciences
Department of Computer Engineering

Supervisor: Dr. Bülent TUĞRUL

In recent times Privacy Preserving Data Mining (PPMD) became more popular with emerge of cloud computing technology. Cloud computing allows data owners to outsource their data to store and process. However, it is lack in providing confidentiality, privacy and security for the stored data. Result of this, organizations which are deal with the sensitive data like credit card records, medical records are reluctant to outsource the data to clouds. Thus, data needs to be encrypted before they outsource to a cloud and processing needs to be done on the encrypted data. Secure Multiparty Computation (SMC) and Homomorphic Encryption (HE) algorithm allows the cloud to process queries on the cipher text and they protect the data, query and access pattern from the cloud service provider and other parties (except their private data) involved in the query processing. k – Nearest Neighbor (k - NN) algorithm is the simplest and mostly used spatial interpolation method to find similarity matching in data mining. Proposed solution employ Paillier cryptosystem to get the best performance. Apart from that kd – tree and R – tree algorithms were used to store the data in the clouds to find the matching nearest neighbor in a quicker way. Proposed solutions were analyzed that they ensure the privacy and security of the data in every scenario of the processing.

September 2018, 38 pages

Key Words: privacy – preserving, k – nearest neighbor, data mining, spatial interpolation, secure multiparty computation

TEŞEKKÜR

Yüksek lisans programımı tamamlamak için bana yardımcı olan herkese teşekkür etmek için bir fırsat olarak görüyorum. Her şeyden önce, bana, bir yüksek lisans programı verme şansını veren yüce Allah'a teşekkür etmek istiyorum. Yüce Allah'tan sonra, danışmanım Dr. Öğr. Üyesi Bülent TUĞRUL'a (Ankara Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı) özveri, öneri, cesaretlendirme ve motivasyonları için teşekkür ediyorum. Tezi hazırlarken farkındalık kazanmamı sağladığı ve zamanında tezimi tamamlamamdaki büyük katkısı için şükranlarımı sunuyorum.

Ankara Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı Başkanı ve tüm akademisyenlere değerli önerileri için teşekkür ederim.

Güzel ülke Türkiyedeki yüksek lisans programı takip edebilmek için Yurtdışı Türkler ve Akraba Topluluklar Başkanlığı'na teşekkür etmek istiyorum. Ayrıca, Sri Lanka Güneydoğu Üniversitesi Rektörü Prof. Dr. MMM. Naajim, Dekan, İslami Çalışmalar ve Arap Dili Fakültesi, Bölüm Başkanı, İslami Çalışmalar Dairesi ve Sri Lanka Güneydoğu Üniversitesi'nin tüm kadrolarına de teşekkürlerimi sunuyorum.

Son olarak, ama en önemlisi babam Kalideen'e, annem Habsha'ya, eşim Aysha Sahani'ye, sevgili çocuklarım Umar Rajeeh'e ve Ameedh Yoosuf ve kız kardeşlerime büyük bir teşekkürü borç biliyorum.

Çalışmayı tüm aile üyelerine adadım.

Hepinize teşekkür ederim.

Muhammad Rifthy KALIDEEN

Ankara, Eylül 2018

İÇİNDEKİLER

KISALTMALAR DİZİNİ

SMC	Secure Multiparty Computation
PPDM	Privacy – Preserving Data Mining
AES	Advanced Encryption Standard
k-NN	k- Nearest Neighbor
ORAM	Oblivious – RAM
HE	Homomorphic Encryption
DO	Data Owner
QO	Query Owner
CSP	Cloud Service Provider
CAD	Computer Aided Design

ŞEKİLLER DİZİNİ

1. GİRİŞ

Son on yılda gelişen bilgisayar teknolojilerinden biri de bulut bilişimidir. Bulut bilişim için en çok kullanılan tanım NIST (Mell ve Grance 2009) tarafından verilmiştir, *"Minimum yönetim çabası veya servis sağlayıcı etkileşimi ile hızlı bir şekilde tedarik edilebilen ve piyasaya sürülebilen, yapılandırılabilir bir hesaplama kaynakları (ör. ağlar, sunucular, depolama, uygulamalar ve hizmetler) havuzuna uygun, isteğe bağlı ağ erişimi sağlayan bir modeldir".*

Bulut bilişim modelleri, dağıtım modellerine göre üçye ayrılır;

1) Özel bulut: Bulut platformu belirli bir şirket veya kuruluşa aittir.
2) Genel bulut: Halk için mevcut bulut platformu. Bu, kendisini kayıt ettikten sonra kamu kullanıcısı tarafından kullanılabilir.
3) Hibrit bulut: Özel mülkiyete ait bir bulutun kaynakları, kamu kullanıcıları tarafından belirli bir ölçüde kullanılabilir.

Üç dağıtım modeli arasında, genel bulut modeli en duyarlı modeldir. Çünkü, tüm kamu kullanıcılarının kendi hizmetlerini sunmalarına izin verir, güvenilmeyen kullanıcılar da hizmetlerini bu modelde sağlayabilir.

Teslimat modeline göre, bulut bilişim dörde ayrılmıştır;

1) Software – as - a – Service (SaaS): Bulutlarda barındırılan uygulamalara, kullanıcılar veya müşteriler tarafından internet veya ağ üzerinden erişilebilir. Uygulamanın istemcinin cihazlarına yüklenmesi gerekmez. SalesForce.com, Google Mail ve Google Dokümanlar, SaaS için örneklerden bazılarıdır.

2) Platform – as - a – Service (PaaS): CSP, müşterilere kendi platformlarını kurmadan platform, araç ve diğer hizmetleri sağlar ve uygulamayı geliştirmeye, dağıtmaya ve yönetmeye olanak tanır. Google Apps, iyi bilinen bir PaaS hizmetidir.

3) Infrastructure - as – a – Service (IaaS): Bu tür bir CSP aşağıdakileri internet yoluyla desteklemektedir; hesaplama kaynakları, depolama ve ağ. IaaS için en iyi örnek Amazon EC2'dir.

4) Data – as – a - Service (DaaS): IaaS'nin özel bir türüdür. Sadece müşterileri tüm veritabanından ziyade gerçekte kullandıkları için ücretlendirir. Ayrıca tablo tarsi veriyi sorgulamayı destekler. Amazon S3, Google BigTable, DaaS sağlayıcılarıdır.

Kullanıcıların veritabanlarını ve işlevlerini herkese açık bir bulutta dışarıdan temin etmelerini sağlar (Mell ve Grance 2010, Hu vd. 2011). Bu bulutlar veritabanlarını ve veri kümesini sorgulamak için erişim mekanizmalarını yönetir. Şirketler ve kuruluşlar, veri setinin dış kaynak kullanımı durumunda yönetim maliyeti, depolama giderleri ve mükemmel hizmetler gibi avantajları elde ederler. Bulutlar yarı - dürüst ve meraklıdır. Sonuç olarak, verilerin gizliliğini ve sorgulamayı korumak, buluttaki depolanmış veriler için bir sorundur. Bulut servis sağlayıcıları tarafından karşı karşıya olan çok fazla güvenlik riski vardır (Hashizume vd. 2013, Rodrigues vd. 2013, Ryan 2013, Wang vd. 2014). Örneğin, bir bulutta düz bir metin olarak saklanan verilerin, bulutta bir kesinti olursa, verilerin saldırganı etkileyebileceğini varsayalım. Verileri üçüncü taraflardan veya bilinmeyen kullanıcılardan korumanın bir yolu, bulutta saklamak için bulut hizmeti sağlayıcılarına göndermeden önce hassas verileri şifrelemektir(Abadi 2009, Pearson vd. 2009, Li vd. 2012). Geleneksel şifreleme tekniği verilerin gizliliğini sağlar. Bununla birlikte, şifreleme tekniği kullanıcının sorgusu için herhangi bir garanti veremez. Sorgu bulut tarafından izlenebilir. Sistem, saklanan verileri ve kullanıcının sorgusunu da korumalıdır. Bu sorunun sonucu, sorguyu eşleşen modelleri aramaya göndermeden önce şifrelemek de gerekiyor. Bulutlar, şifrelenmiş verilerden ve sorgudan sorgu işlemi sırasında faydalı bilgileri alabilir(Williams vd. 2008, di Vimercati vd. 2012). Bulut, verileri şifrelenmiş bile olsa sonuç ilişkilerinden alabilir. Buna veri erişim modeli denir.

Sonuç olarak, sorgulama işlemi sırasında veriler ve kullanıcının sorgusunun korunması gerekir. Ortaya çıkan soru, "Şifreli veriler kullanıcının gizliliğini etkilemeden nasıl işlenecek?" sorusudur. Soruyu cevaplamak için araştırmacılar, şifrelenmiş veriler üzerinde sorgulama işlemi olarak adlandırılan yeni bir araştırma alanı ortaya koymuşlardır.

Güvenli Çok Taraflı Hesaplama (Secure Multiparty Computation - SMC), bu günlerde daha popüler oldu. Gizlilik Koruyan Veri Madenciliği (Privacy – Preserving Data Mining - PPDM). SMC'nin ardındaki fikir, verileri ve diğer işlem bilgilerini diğer taraflara açıklamaksızın, ilgili taraflarla tüm tarafların bir işlevini veya sorgusunu işlemektir. Veriler her bir tarafa aittir. Bu veriler onların özel mülküdür. Bununla birlikte, her bir tarafın çıktısı daha fazla faaliyet için bunların arasında paylaşılabilir. Çok fazla gerçek dünya problemi SMC kullanarak çözüldü. Örneğin, oylama sistemi, çevrimiçi açık artırma, teklif verme vb. Öte yandan, SMC'nin bir sorguyu işlemek için daha fazla iletişim, artıklık ve zamana ihtiyacı vardır.

Araştırmacılar, şifrelenmiş veriler üzerinde sorgulama işleminin başlangıcında, Gelişmiş Şifreleme Standardı (Advanced Encrtpted Standard - AES) kullanır. Bir bulut verilerine dış kaynak kodlamadan önce simetrik anahtar şifreleme yöntemi ile şifrelenir. Başlangıçta, daha iyi bir teknik gibi görünür. Bununla birlikte, AES şifrelenmiş veriler üzerinde herhangi bir işlem gerçekleştirmek konusunda buluta engel olur. Bu nedenle, kullanıcının tüm verileri buluttan indirmesi, şifresini çözmesi ve daha sonra düz metin üzerindeki işlemleri yapması gerekir. Mobil kullanıcılar, büyük veri ve mekansal veriler için etkili değildir.

Şifrelenmiş veriler üzerinde sorgulama işlemi için kullanılan protokollerin aşağıdakileri sağlaması gerekir;

1. Şifrelenmiş verilerin gizliliği
2. Kullanıcı sorgusunun gizliliği
3. Erişim kalıbını gizle

Veri madenciliğinde benzerlik araştırması bulmak için çok fazla algoritma vardır. En yaygın kullanılan algoritmalardan biri k-NN (k- En Yakın Komşu) algoritmasıdır. Çünkü yeni verileri eğitmek ve büyük miktarda veri için etkili olmak açısından anlaşılması kolaydır. Genel olarak noktalar k - NN algoritmasında mesafe ölçümleri olarak temsil edilir. Öklid mesafesi, Manhattan mesafesi ve benzeri gibi çeşitli teknikleri kullanarak belirli bir nokta için en yakın noktaların k sayısını bulur. k - NN algoritması;

- ✓ Her nokta için eğitim noktalarından istenen bir noktaya kadar olan mesafeyi hesaplayın.
- ✓ En yakın k noktalarını seçin
- ✓ Süreci sınıflandır

Burada, eğitim noktalarına bağlı olarak k değişmektedir.

1.1 kd – Tree

kd ağacı (kd - tree), çok boyutlu bir alanda uzamsal verileri temsil etmek için kullanılan ikili ağaçlardan biridir. Puanlar bir kd ağacı oluşturmak için iki şekilde ayrılır. Birincisi medyan x eksenine göre. Dikey bir çizgi noktaları ikiye böler. Puanların yarısı çizginin sol tarafında ve sağ tarafta bir başka yarıda. Noktaları ikiye bölmek için y ekseni tarafından yatay bir çizgi çizilir. Yarısından biri çizginin üstünde ve dibinde kalır.

buildTree(*P,iteration*)

if number of points =1

then return *P*

else if *iteration* is even

then Divide the points into two portion P1 and P2 by median x – axis line l

else Divide the points into two portion P1 and P2 by median y – axis line l

lchild ← buildTree(P_1, *iteration +1*)

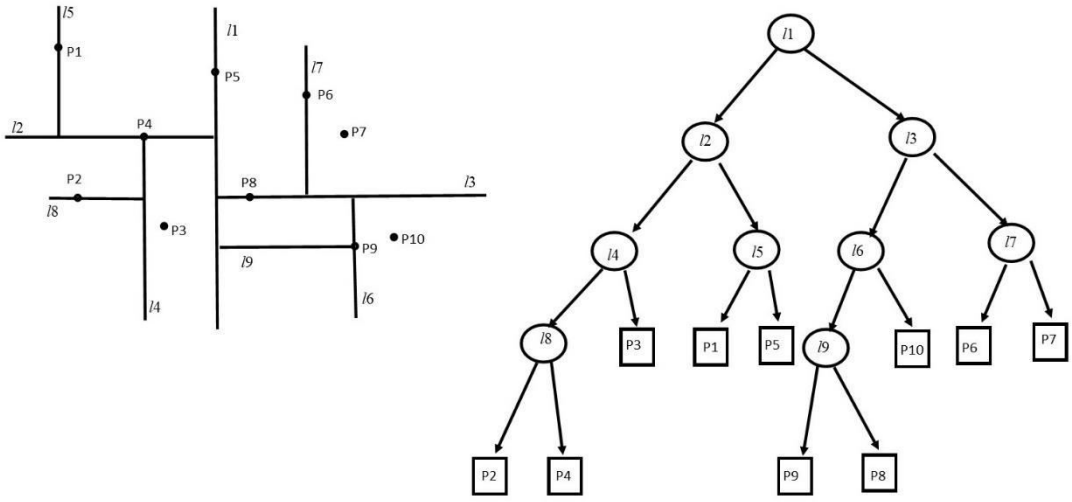
rchild ← buildTree (P_2, *iteration +1*)

Create a node r

Store l in r

make *lchild* as the left child of r, and *rchild* as the right child of r.

return r



Şekil 1.1 Örnek kd - x ekseni tarafından uygulanan ağaç

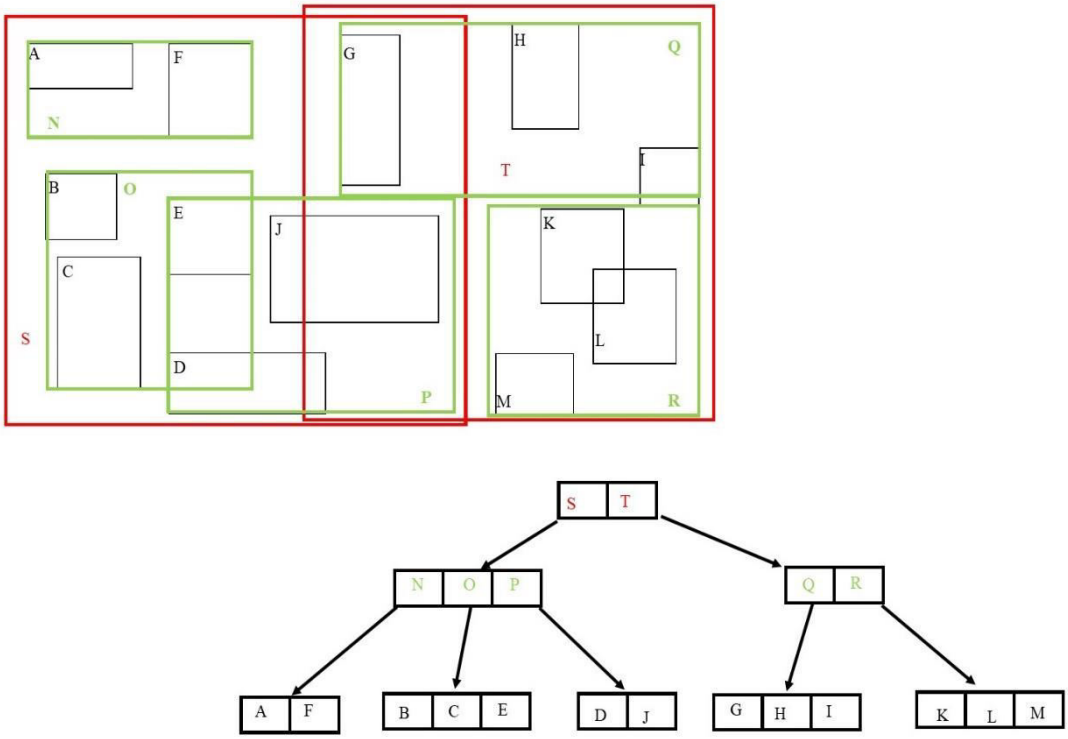
1.2 R – Tree

Mekansal veri, bir soruşturmaya çok hızlı yanıt vermek için çok boyutlu alanlara ihtiyaç duyar. Bu nedenle, geleneksel tek boyutlu veritabanı sistemleri, mekansal veri enterpolasyonu için uygun değildir. Aynı şekilde, Bilgisayar Destekli Tasarım (Computer Aided Design - CAD) işleme de çok boyutlu yapı veritabanı sistemine ihtiyaç duyar. Çok sayıda araştırmacı tarafından önerilen çok boyutlu nokta veri yapısının sayısı artmaktadır. Guttman (Guttman 1984) birden çok boyutta R-Tree (Dikdörtgen Ağaç) adlı bir veri yapısını önerir.

R –Tree yükseklik dengeli ağaçlardan biridir. Dizin oluşturma dinamiktir ve veri nesnelerine işaretçiler yaprak düğümlerinde saklanır. Ağaç, mekansal bir arama için

çıktıları, düğümlere minimum sayıda ziyaretle teslim edebilecek şekilde yapılandırılmıştır.

Ortalama durumda arama verimliliği (Manolopoulos vd. 2010), R - tree ile aralık sorgulamaları için geliştirilebilir (Wang vd. 2014) ve en yakın komşu. Bu avantajdan dolayı, SQL ve NoSQL veritabanlarında R –Tree kullanılır. R - Ağacı Şekil 1.2' deki gibi uygulanır;



Şekil 1.2 Örnek R-Ağacı

- ✓ Komşu nesnelere aynı seviyede grup verin ve ağacın daha yüksek seviyesine bağlı.
- ✓ Yaprak düğümlerinde saklanan noktalar.
- ✓ Yaprak olmayan düğümler dikdörtgen olarak gösterilir

Arama sorgusu aşağıdaki gibi çalışır;

- ✓ Kök düğümden başlar
- ✓ Sorgu dikdörtgeni birbirine bağlarsa, aramayı çocuklarına devam ettirin.
- ✓ Sorgu ve yaprak eşleşiyorsa, yaprak değerini döndürün.

1.3 Oblivious – RAM (ORAM)

Bir istemci - sunucu mimarisi modelinde, bir istemci, istemcinin işlemle ilgili herhangi bir bilgiyi paylaşmak istemediği bir sunucuda depolanan veriler üzerinde bir işlem yapmak istemektedir. Çünkü sunucuya güvenmiyor. Bundan dolayı ne saklanan veri ne de verinin erişim modeli sunucuya geri dönülmez. Golreich ve Ostrovsky (1996), ORAM olarak adlandırılan bu sorunun üstesinden gelmek için bir çözüm önermektedir.

ORAM aşağıdaki gibi çalışır; tüm veriler blok (N blok) olarak saklandı. Bu bloklar sunucuya gönderilmeden önce şifrelenir. İstemci, saklanan bloklardan birine (i^{th} blokla) erişmek istiyorsa, istemci tüm bloklara tek tek erişir ve i^{th} bloğu aldığında, üzerinde çalışır. Sonra tüm bloğu yeniden şifreler ve sunucuya geri gönderir. Sunucu, istemci tarafından birkaç kez eriştiyse, i^{th} bloğunu tanımlamak için herhangi bir ortak kalıp tanımlayamaz.

Goldreich ve Ostrovsky (1996) tarafından sunulan RAM için hiyerarşik veri yapı modellemesi oluşturma ilk fikri. Fikir, kovalarda depolanan ORAM'dir, kovaların büyüklüğü geometrik bir seride artmaktadır. Üstteki en küçük kova ve aşağı doğru yükselir. En kötü senaryoda O ($n\log_2 n$) gerekir. Shi vd. (2011) tarafından önerilen ağaç tabanlı bellek yapısı ORAM.

Ağacın yapısı aşağıdaki gibidir;

- ✓ İstemci bir sunucuda n sayıda dizi bloğunu [M_1, M_2,…, M_n] saklamak ister.
- ✓ Bir bloğun boyutu B = $\log_c$ n burada c sabittir.
- ✓ n yaprakları ile ikili ağaç ve hafızayı saklamak için n logunun yüksekliği oluşturulur.
- ✓ Ağaçtaki düğümler *bucket* olarak kabul edilir ve bir *bucket* Z tane blok içerir.
- ✓ Ağacındaki bir yol *p(i)* bir blok *i* aittir
- ✓ M_i'ye denk gelen her bir kovayı kontrol edin. Eşleşmiyorsa, atılacaktır.

Hafızada tekrar saklanması gereken bloğu okuyun. Bloğu tekrar okurken bir önceki yerde saklanan blok bilgi sızıntısı olasılığı vardır. Bu durumdan kaçınmak için blok yeniden şifrelenir ve en üst düğüme geri gönderilir.

1.4 Homomorfik Şifreleme / Homomorphic Encryption (HE)

Homomorfik şifreleme adı verilen şifreli metinler üzerinde matematiksel işlem sağlayan özel bir kriptografi türüdür. Pek çok araştırmacı Homomorphic Encryption için Paillier kriptosistemi, Goldwasser - Micali şifreleme şeması, ElGamal şifreleme şeması, Boneh - Goh - Nissim şifreleme şeması ve benzeri gibi birçok algoritmayı önermektedir (Yi vd. 2014). Paillier şifreleme sistemi, şifrelenmiş düz metin üzerinde iki cebirsel işlem sağlar. Bunlar, ek işlem ve çoğaltıcı işlemdir. E_{pk}'nin şifreleme işlevi olduğunu ve E_{qk}'nin *pk*'nin ortak anahtar olduğu ve özel anahtarın bu senaryoda *qk* olduğu şifre çözme işlevi olduğunu varsayın. $E_{pk}(x)$'den x'i almak için *qk* hakkında bilmeniz gereken, *qk* hakkında herhangi bir bilgi sahibi olmadan şifreleme işlevini şifreleyemez.

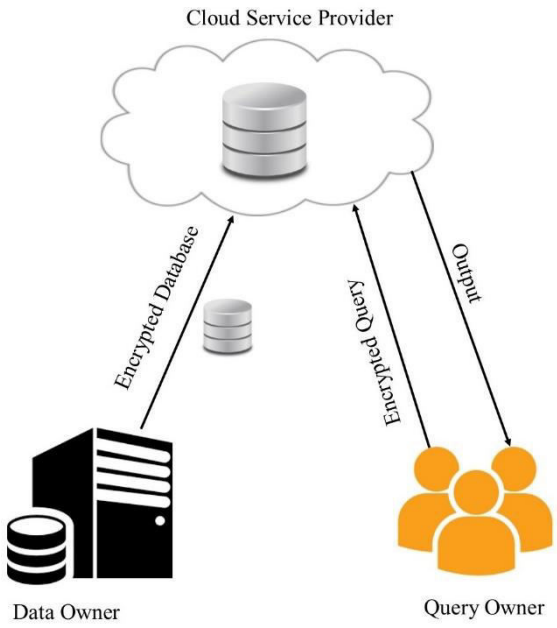
Homomorfik İlaveler - D_{qk} (E_{pk} (a+b)) = $D_{qk}(E_{pk}(a) * E_{pk}(b) \bmod N^2)$
Homomorfik Çarpma - D_{qk} (E_{pk} (a * b)) = $D_{qk}(E_{pk}(a)$ b $\bmod N^2)$, a & b'nin plaintext olduğu yerlerde, N iki büyük asal sayının çarpım sonucudur.

Çalışmamız, yukarıda belirtilen özellikleri sağlamak için bir buluttan dış kaynaklı şifrelenmiş veriler üzerinde bir çözüm sunmaktadır. Bu çalışma esas olarak, tek - tarafı ve iki taraflı durumlarda k - En Yakın Komşu (k-NN) sorgusunun yarı - dürüst modeli için bir çözüm bulmaya odaklanmaktadır.

1.5 Sorunun Tanımlanması

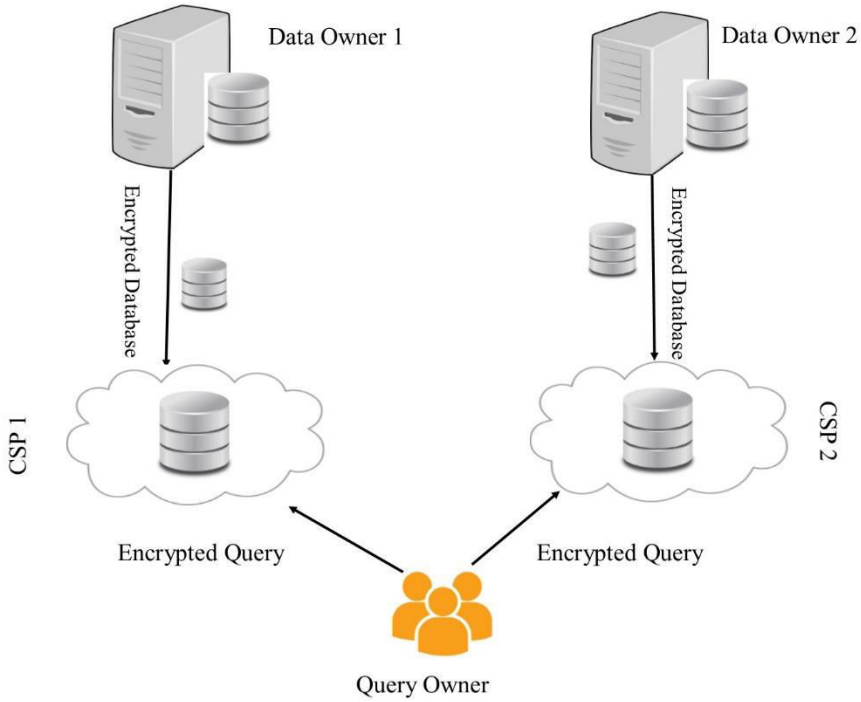
Bu çalışmada iki ana sorun senaryosunu belirledik; 1) Tek - taraflı problemi ve 2) İki - taraflı problemi. Tek - taraflı senaryosunda, önerilen çözümümüze katılan üç katılımcı vardır. Bunlar veri sahibi (bundan sonra DO olarak anılacaktır), bulut servis sağlayıcısı

(bundan böyle CSP olarak anılacaktır) ve sorgu sahibi (bundan böyle QO olarak anılacaktır).



Şekil 1.3 Tek - Taraflı modeli veri sahibini, sorgu sahibi ve bir bulut servis sağlayıcı içerir

İki - taraflı probleminde beş taraf var. Bunlar, iki veri sahibi (iki DO), iki bulut hizmet sağlayıcısı (2 CSP) ve bir sorgu sahibi (QO). Bu senaryoda, her DO'nun kendi CSP'si vardır (örneğin: DO1, veritabanını CSP1 ve CS2'de DO2'de saklar), QO, tek bir bilgisayar, bir şirket veya bir kuruluş olabilir.



Şekil 1.4 İki - Taraflı model, iki veri sahibi, bir sorgu sahibi ve iki bulut servis sağlayıcısı içerir

DO, veritabanını bir üçüncü taraf sunucuya dış kaynaklardan önce düz metin biçiminde alır. DO, veritabanını özel anahtarıyla (Secret Key - SK) şifreler ve şifrelenmiş veritabanını ve sorgu işleme hizmetlerini bir bulut sunucusuna gönderir. DO için anahtarların Paillier şifreleme sistemi tarafından üretildiği varsayılmaktadır (Paillier 1999). Paillier sistemi, bir katkı maddesi homomorfik sistemdir ve semantik olarak güvenli olmasını sağlar. Anlamsal olarak güvenli algoritma, hesaplama açısından ayırt edilemeyen şifre metni sağlar. Bulutlar açısından bakıldığında, şifre metinleri rastgele sayılardır. Anlamsal güvenliğin sonucu olan bulut, şifrelemeyi bir iletiden diğerine iletemez(Graepel vd. 2012).

Ek homomorfik şifreleme şeması, şifre metni üzerinde matematiksel işlem yapılmasına izin verir. Örneğin, DO'nun bulutta saklanan şifrelenmiş veritabanı ve QO'nun verileri buluttan güvenli bir şekilde almasına olanak tanır. QO, buluta bir sorgu gönderdiğinde ve buluttan bir sonuç bilgisi aldığında, ne sorgu ne de şifrelenmiş veritabanındaki veriler bulutta ortaya çıkar ve aynı zamanda veri erişim modelini verilere ve yetkisiz

kullanıcıların erişimini korur. Bu aktivite, şifrelenmiş veriler üzerinde bir Gizlilik - Koruma Sorgusu (Privacy – Preserving Query Processing - PPQP) olarak adlandırılır. QO tarafından yapılan bir sorgu için çıktı yalnızca teslim edilir ve sorguyu başlatan QO'ya bildirilir.

1.6 Amaçlar ve Ana Katkılar

Bu çalışma, veritabanının şifrelenmiş bir biçimde depolandığı bulut hizmeti sağlayıcısına herhangi bir sorgulama yaparken sorunların belirlenmesi ve önerilerde bulunması istenilmektedir. Önerilen çözüm k - en yakın komşu sorgusunu, kullanıcı sorgusunu ve yetkisiz kullanıcıların yetkisiz erişimini ve yetkisiz erişimini sağlamak için erişim düzenini korumaktadır. Süreç bulutta (CSP) olduğunda veri sahibi sorgu işleme etkinliğine katılmaz. Bu nedenle, sorguyu veri sahibi tarafından tahmin etmenin bir yolu yoktur. Bu arada, QO ve CSP, CSP'de saklanan ilgili bilgileri bulamamaktadır. Çünkü, CSP'de saklanan veriler şifreli biçimdedir. Aşağıdaki gizlilik şartları sağlanmıştır;

- Veri gizliliği: CSP'de saklanan veriler ve herhangi bir sonuç sorgu işlemi sırasında CSP'ye açıklanmaz.
- Sorgu gizliliği: QO'nun sorgusu CSP'ye veya DO'ya maruz kalmaz.
- Veri erişim kalıplarını önleme: Sorguyla ilgili kayıtlar, sorgu işlem süresi öncesinde, sırasında ve sonrasında CSP'ye veya QO'ya maruz bırakılmamalıdır.
- Çıkış güvenliği: Sorgunun nihai sonucu yalnızca QO'su görebilir. Ne DO ne de CSP'nin nihai çıktı hakkında bir fikri yoktur.

Veri sahipleri, bulut tarafından üretilen bazı ara çıktıları alır. Ara çıkışlardan DO'ler, QO tarafından istenen sorgu hakkında herhangi bir bilgiyi alamazlar.

1.7 Organizasyon

Bu çalışma şöyle organize edilmiştir: Bölüm 2, önerilen sistemimize çok yakın olan ilgili çalışmaları açıklamaktadır. Bir taraf problemi ve önerilen çözüm 3. Bölümde ifade edilmiştir. Aynı şekilde Bölüm 4'de iki taraflı sorun ve önerilen çözüm açıklanmaktadır. 5. Bölüm, mevcut çözümlerle birlikte önerilen çözümün analizini açıklamaktadır ve son olarak, 6. bölüm sonuç ve gelecekteki çalışmaları önermektedir.

2. İLGİLİ ÇALIŞMALAR

Güvenli k-En Yakın Komşu yöntemleri ile ilgili mevcut çalışmalar (Elmehdwi vd. 2014) ve Gizlilik - Veri Madenciliği Koruma (Samanthula vd. 2015) bu bölümde açıklanmıştır.

2.1 Güvenli k - En Yakın Komşu Yöntemleri

Bir sorgu için kNN bulmak bu günlerde daha popüler oldu. Sosyal medyada (örn. Facebook) en yakın arkadaşı bulma, örüntü tanıma, veri madenciliği vb. Güvenli k – NN'yi çözmek için araştırmacıların önerdiği birçok fikir var. Bu yöntemler iki kategoriye ayrılmıştır. Merkezileştirilmiş ve dağıtılmışlardır. Burada veriler şifreli olabilir veya olmayabilir.

Merkezileşmiş teknik: Merkezileştirilmiş teknikte, veriyi, CSP'nin verileri yönetmek gibi DO adına hareket ettiği, güvenilir kullanıcıların veriyi sorgulamasına izin verdiği, bulut hizmeti sağlayıcı olarak adlandırılan bir üçüncü tarafa veri aktarımı yapar. Verileri güvenilmeyen bir sunucuya dış kaynak kullanmak, birçok güvenlik sorununa da neden olur. Veri gizliliği, yetkisiz kullanıcı erişimi, vb.

Bankacılık bilgileri, tıbbi kayıtlar ve savunma kayıtları gibi verilerin korunmasının geleneksel yollarından biri şifreleme. Veriler bir sunucuya dış kaynak aktarılmadan önce şifrelenir. Şifre metni üzerinde işlem yapmak araştırmacılar için zor bir süreçtir. Bunun sonucu olarak, araştırmacılar tarafından önerilen bazı teknikler vardır. Örneğin, aralık sorguları (Hore vd. 2004, Shi vd. 2007, Hore vd. 2012) ve toplama sorguları (Hacıgümüş vd. 2004, Mykletun ve Tsudik 2006). Bu tekniklerin k-NN sorgusu için güvenliğin değerlendirilmesi konusunda bir sorunu vardır.

Son zamanlarda, araştırmacılar tarafından güvenli k-NN için önerilen yeni yaklaşımlar aşağıda belirtilmiştir. Wong vd. (2009) tarafından önerilen Asimetrik Skaler - Ürün

Korumalı Şifreleme (ASPE). Veri ve sorgu, farklı bir şifreleme teknikleriyle şifrelenmiş, üçüncü bir tarafa dış kaynak aktarılmadan önce, QO her iki şifreleme için şifre çözme anahtarını bilmektedir. Zhu vd. (2013), DO's anahtarının QO'ya maruz kalmayacağı bir yöntem önermiştir. Ancak, her iki senaryoda, sorgu şifreleme sırasında DO yer almıştır. Hu vd. (2011), şifrelenmiş veriler üzerinde toplama, çıkarma ve çarpmaya izin veren başka bir teknik önermişlerdir. Öte yandan, seçilmiş bir risk altındadır - düz metin saldırısı ve yukarıdaki yöntemlerin tümü erişim modelini sunucuya sızdırabilir.

Yao vd. (2013), sorgu için k-NN noktalarından ziyade en yakın şifreli bölümleri seçen güvenli bir k-NN yöntemini önermektedir. k - NN yönteminin güvenliğini sağlar. Ancak, son kullanıcı sorgu işlemenin çoğunu yapar. Dış kaynak kullanımının amacı gerçekleştirilememiştir. Bu arada, bölümün kimliğini buluttan koruyamaz.

Veri dağıtım yöntemi: Veri dağıtım yönteminde bağımsız sunuculara dış kaynak aktarılmadan önce verilerin yatay veya dikey olarak bölümlere ayrıldığı varsayılmaktadır. Söz konusu yöntem, hesaplamaları için Güvenli Çok Taraflı Hesaplama (SMC) kullanır. Her bir taraf kendi hesaplarını diğer tarafları ifşa etmeden kendi hesaplamaları ile yapar. Dağıtılmış bir ortamda gizliliğin korunması için bir çözüm bulmak için birkaç çalışma yapılmıştır. Yatay bölünmüş veriler için gizliliğin korunması, SMC kullanarak k-NN noktalarını bulmak için Shaneck vd. (2009) tarafından gerçekleştirilmiştir. Dikey olarak bölünmüş veriler için, Qi ve Atallah (2008) tek adımlı k-NN araması olarak adlandırılan bir çözüm önermiştir. Yer bazlı problem için Ghinita vd. (2008) Özel Bilgi Erişimi (PIR) diye bir çözüm önermektedir. Önerilen çözümde, sunucudaki depolanmış veriler düz metin içeriyor. Sorgu gizliliğini sağlarken, diğer yandan verilerin gizliliği ve erişim şekli ile ilgili hiçbir şeyi ele almamaktadır.

Genel olarak, yukarıda bahsedilen teknikler, aşağıdaki tezlerde bu tez çalışmasına uygun değildir;

1) Veri ve sorgu şifrelenmiş biçimde olması gerekiyor
2) Saklanan veriler şifreli metin olmalıdır.

2.2 Gizlilik - Veri Madenciliğinde Koruma

Agarwal ve Srikant (2000), Lindell ve Pinkas (2000), Veri Koruma - Veri Madenciliği (PPDM) teknikleri konseptini başlatmışlardır. PPDM, verilerin gizliliğini etkilemeden mevcut verilerden yararlı bilgiler elde etmesini sağlar (Agrawal ve Srikant 2000, Ravu, Neelakandan vd. 2012). (Agrawal ve Srikant 2000, Lindell ve Pinkas 2000) dışında, son zamanlarda araştırmacılar tarafından önerilen birçok teknik var. Bu teknikler ikiye ayrıldı; 1) veri pertürbasyon, 2) veri dağılımı.

Veri pertürbasyon: Bu yöntemin arkasındaki fikir, verilere rastgele bazı sesler eklemek ve verileri farklı bir görünüme dönüştürmektir. Dönüştürülmüş veriler üzerinde yapılan madencilik süreci. Bunun sonucu, veri modeli madenci için gizlidir. Agarwal ve Srikant (2000) çalışmalarında karar ağacı sınıflandırıcısını uygulamak için veri pertürbasyon tekniğini kullanmışlardır. Agarwal ve Srikant'ın dışında, bu teknik üzerinde çalışan birçok araştırmacı vardır (Evfimievski vd. 2004, Fienberg ve McIntyre 2004, Bayardo ve Agrawal 2005, Zhang vd. 2005). Ancak, ek gürültülü veriler eklemek sorguya mükemmel sonuç vermez ve bunlar anlamsal olarak güvenli veriler için uygun değildir.

Veri dağıtım tekniği: Veri dağıtımının ardındaki teknik, verileri yatay veya dikey olarak bölümlere ayırır ve farklı taraflara yayılır. Dağıtılmış veriler birbirine gösterilmez. Lindell ve Pinkas (Lindell ve Pinkas 2000), karar ağacı sınıflandırıcısını uygulamak için veri dağıtım tekniğini tanıtan ilk araştırmacılardır. SMC'de bu teknik üzerinde yapılan çok sayıda çalışma vardır (Clifton vd. 2002, Kantarcıoğlu ve Clifton 2004, Xiong vd. 2006, Aggarwal ve Philip 2008, Hu vd. 2011).

Veri madenciliği bu günlerde daha popüler oldu. Çünkü, bankacılık, sigorta, tıp, sosyal medya gibi pek çok uygulama ve daha pek çok veri madenciliğinin büyük veriden anlamlı bilgi almak için sınıflandırması kullanılmaktadır. Son zamanlarda, bulutta veri madenciliği daha popüler hale geldi. Bulut bilişim, DO bilgilerinin depolanan veriler üzerinde saklanması ve çalıştırılması için veri kümesinin bir buluttan dışarı

aktarılmasını kolaylaştırır. Bununla birlikte, veri gizliliği bulut bilişimde büyük bir sorun haline geldi. Bu nedenle, buluttaki verilerin gizliliğini ve güvenliğini sağlamak için, şifreleme, verileri bir bulutta dış kaynaklara gönderilmeden önce korumak için kullanılır.

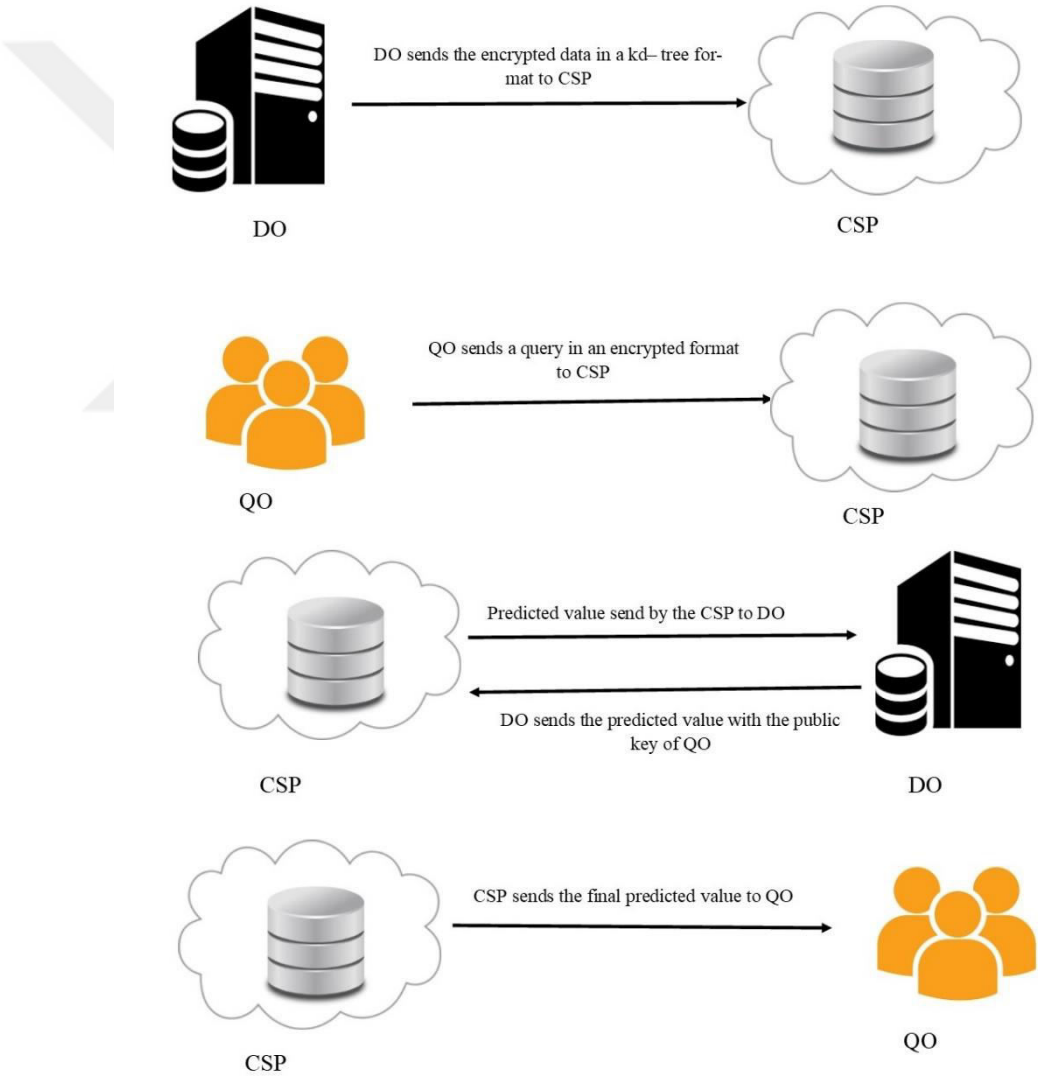
Mevcut gizlilik koruma teknikleri bu çalışma için uygun değildir. Çünkü;

1) Verilerin şifreli bir biçimde saklanması gerekir. Bununla birlikte, mevcut teknikler, bulutların arasında verilerin düz bir metne dağıtılmasına izin verir.
2) Son çıktının doğruluğu. Kullanılabilir yöntemler başarısız. Çünkü orijinal verileri gizlemeye izin veren ek rasgele gürültülü verilere sahiptir.

Veri erişim örüntüsünün, bulutlar tarafından sorgu işlem süresi sırasında türetilmeden korunması gerekir.

3. ÖNERİLEN TEK - TARAFLI ÇÖZÜM

Yukarıda açıklandığı gibi, güvenli k-en yakın komşu interpolasyon yönteminin dış kaynak kullanımına dâhil olan üç kişi vardır: DO, CSP ve QO. Ölçülen tüm verileri depolayan veritabanı DO'ya aittir. CSP depolama ve hesaplama hizmetleri sunar. QO, gelecekteki yatırımlarını planlayabilecekleri sorgulama noktaları ile ilgilenmektedir. Çözümümüz aşağıda açıklanan adımları takip eder;



Şekil 3.1 Tek Taraflı bir Sorun için Önerilen Çözüm Çerçevesi

I. İlk olarak, DO şifrelenmiş formdaki noktalara ve ilgili ölçümlere dayanarak bir kd ağacı oluşturur. Kd ağacını, ORAM sunucusu olarak işlev gören CSP'ye gönderir.

II. QO sorgu noktalarını CSP'ye gönderir. Bununla birlikte, istemcilerin, bir tahmin değerine ihtiyaç duydukları koordinatları saklamak için girdilerini şifreli biçimde göndermeleri gerekir.

III. CSP sorguyu işler ve sorgu koordinatının etrafındaki k - NN noktalarını belirler. CSP, tahmin verileri hesaplanamaz çünkü tüm veriler DO'nun anahtarı ile şifrelenir. Bu nedenle, CSP $\prod_{i=1}^{k} \xi(z_i) = \xi(z_1) * \xi(z_2) * \ldots \xi(z_k)$ formülünü $\sum_{i=1}^{k} z_i$ olarak hesaplar ve DO'ya şifreli bir biçimde gönderir.

IV. DO, gelen değeri CSP'den çözmek için gerekli anahtara sahiptir. Bu, şifresini çözer ve tahmin değerini alır. Tahmin değeri, tek başına sorgu koordinatlarıyla ilgili bilgileri ortaya çıkarmaz. Bu nedenle, DO, QO'nin ilgi duyduğu koordinatı öğrenemez.

V. DO, tahmin değerini CSP'den tahmin değerini gizlemek ve CSP'ye göndermek için istemcinin ortak anahtarı ile şifreler.

VI. CSP, şifre-metin biçimindeki tahmin değerini QO'ya iletir.

VII. QO, şifreli metni açmak için karşılık gelen özel anahtara sahiptir ve bir tahmin gerektirdiği değeri alır.

4. ÖNERİLEN İKİ TARAFLI ÇÖZÜM

Bölümlendirilmiş verilerdeki dış kaynaklı bir k-NN uzaysal enterpolasyon yönteminde gizliliğin korunması için önerilen planımızda yer alan beş taraf vardır:

- Data Owner #1 (DO_1)
- Data Owner #2 (DO_2)
- Cloud Service Provider #1 (CSP_1)
- Cloud Service Provider #2 (CSP_2)
- Query Owner (QO)

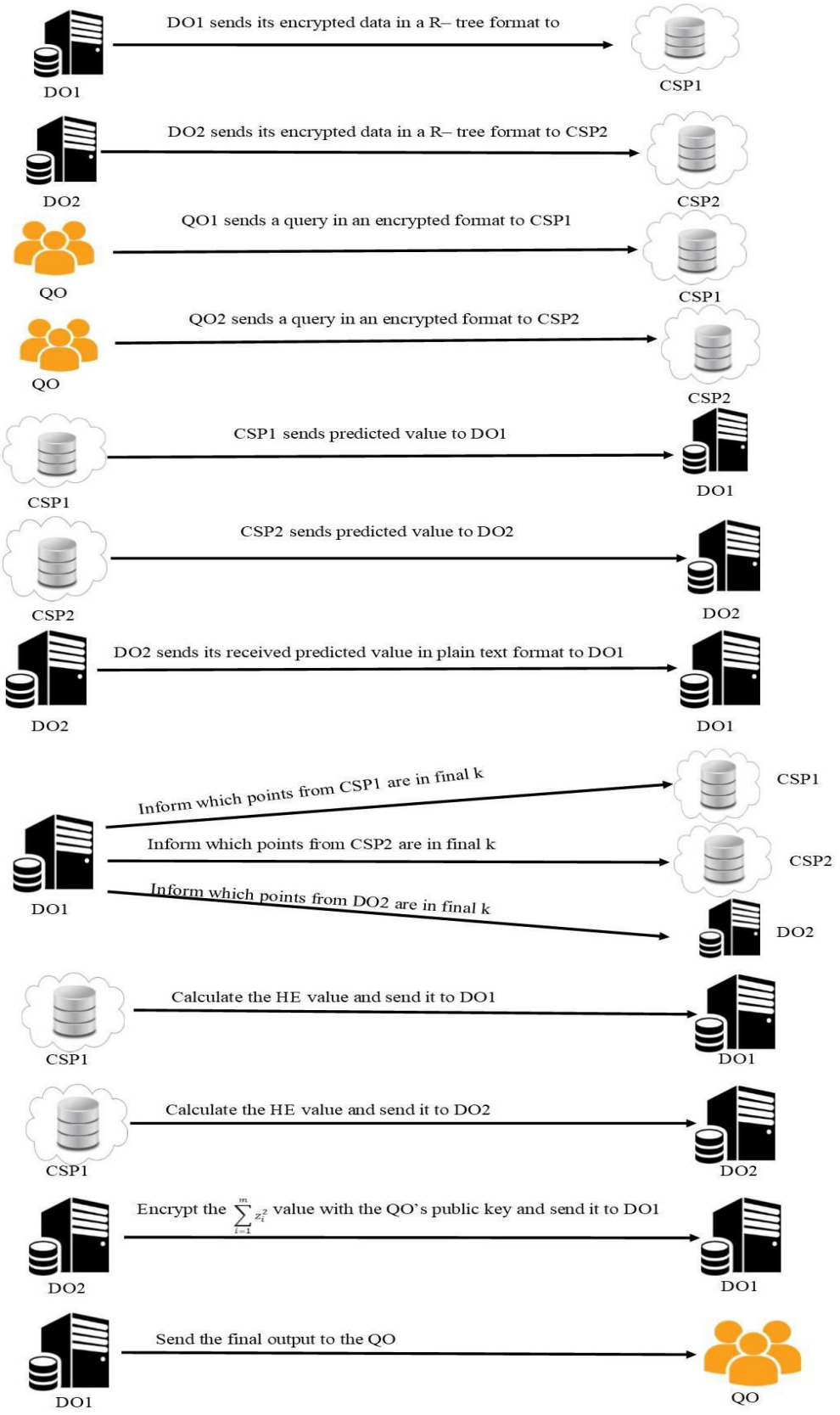
Veri sahipleri, belirli bir bölgeden tüm ilgili mekansal verileri alır. Depolama ve hesaplama hizmetleri CSP'ler tarafından sağlanmaktadır. Sorgu sahipleri, gelecekteki yatırımları için önemli sonuçları olan konumlar için tahmin değerleri talep eder. Çözüm aşağıdaki aşamalara sahiptir;

I. İlk adım olarak, DO'lar kendi R ağaçlarını kendi veri anahtarlarını kullanarak şifreli biçimde kullanarak oluştururlar. Bundan sonra, her DO şifreli R-ağacını karşılık gelen CSP'lerine gönderir (i.e. DO1 CSP1'e gönderir ve DO2, CSP2'ye gönderir).

II. DO1 ve DO2 ortak anahtarlarını QO ile paylaşır.

III. QO, sorgulama noktasını (x_p, y_p) şifrelenmiş bir biçimde, tahmin değerlerini elde etmek için CSP'lere karşılık gelen ortak DO anahtarları ile gönderir. CSP1, DO1'in ortak anahtarıyla şifrelenmiş sorgu noktasını ($\xi_{DO1}(x_p)$, $\xi_{DO1}(y_p)$) alır ve benzer şekilde CSP2, DO2'nin ortak anahtarı tarafından şifrelenmiş sorgu noktasını ($\xi_{DO2}(x_p)$, $\xi_{DO2}(y_p)$) alır.

IV. Her iki CSP, sorgu noktasının etrafında kendi k en yakın noktalarını bulur ve güvenli Euclidean mesafe yöntemini kullanarak sorgu noktasından k en yakın noktalara olan mesafeleri hesaplar. Bundan sonra, CSP'ler ilgili veri sahibine şifreli mesafeler gönderir.



Şekil 3.2 İki Taraflı bir Sorun için Önerilen Çözüm Çerçevesi

V. DO1, CSP1'den hesapladığı mesafeyi ($\xi_{DO1}(d_1^1), \xi_{DO1}(d_2^1), \dots, \xi_{DO1}(d_k^1)$ teslim alır ve DO2, CSP2'den hesapladığı mesafeyi $(\xi_{DO2}(d_1^2), \xi_{DO2}(d_2^2), \dots, \xi_{DO2}(d_k^2)$ teslim alır. Şimdi, her iki veri sahibi de k mesafelerine sahip. Tamamen, gerekli k mesafelerinden daha fazla olan 2 * k mesafeler vardır. Dolayısıyla her DO, düz metinlerdeki uzaklık değerlerine ulaşmak için kendi özel anahtarları ile mesafeleri deşifre eder.

VI. 2*k noktadan k en yakın noktaları bulmak için, DO2, DO1'e düz mesafe değerlerini gönderir (Uzaklık değerleri, herhangi bir sorgu noktası veya DO2'nin verileri hakkında herhangi bir bilgi ortaya çıkarmaz).

VII. DO1 tüm mesafe değerlerini sıralar ve 2 * k'nin en yakın noktalarını bulur. Ayrıca, hesaplama sırasında kullanılması gereken bulut sunucularını ve DO2'yi de bilgilendirir.

VIII. CSP1, homomorfik hesaplama yapar $\xi_{DO1}(z_1^1) \cdot \xi_{DO1}(z_2^1) \cdot \xi_{DO1}(z_n^1)$ hangi noktaların kullanıldığını ve benzer şekilde CSP2'nin hesaplanmasını öğrenerek $\xi_{DO1}(z_1^2) \cdot \xi_{DO1}(z_2^2) \cdot \xi_{DO1}(z_m^2)$ n + m = k burada. Her iki CSP de sonuçları ilgili DO'lere gönderir.

IX. DO1 ve DO2, sonucu çözer ve alır $(z_1^1 + z_2^1, \dots, z_n^1) = \sum_{i=1}^{n} z_i^1$ ve $(z_1^2 + z_2^2, \dots, z_n^2) = \sum_{i=1}^{m} z_i^2$ sırasıyla.

X. DO2, QO'nun ortak anahtarı ile $\sum_{i=1}^{m} z_i^2$ şifreler ve DO1'e gönderir.

XI. DO1 hemen $\sum_{i=1}^{n} z_i^1$ QO'nun genel anahtarı ile şifreler ve $\sum_{i=1}^{n} z_i^1 + \sum_{i=1}^{m} z_i^2$ elde etmek için homomorfik çarpım özelliğini kullanır.

XII. DO1, k-NN hesaplamasında kullanılması gereken tüm z değerlerini QO'ya tutan son sonucu gönderir.

XIII. QO, gelen değeri deşifre eder ve nokta için son tahmin değerini (x_p, y_p) elde etmek için k ile ayırır.

5. ANALİZ

5.1 Tek Taraf İçin Önerilen Çözümün Analizi

5.1.1 Tamamlayıcı maliyet analizi

Geleneksel mekânsal interpolasyon tekniklerinde bir müşteri tarafından yapılan talebin yerini tahmin etmek, verileri depolamak gibi tüm işlevleri sağlar. Öte yandan, gizlilik ve güvenlik konularında eksikler vardır. Ancak gizlilik - korunma ve SMC kullanma fikrimiz gizlilik ve güvenlik konusunda iyidir. Bununla birlikte, çalışmamızın geleneksel olandan biraz daha fazla ek hesaplama, depolama ve iletişim maliyeti olması gerekiyor. Bu ek çalışma, mekansal enterpolasyonun genel performansını etkilemez. Çünkü uzaysal enterpolasyon sistemi gerçek zamanlı bir sistem değildir.

5.1.1.1 Hesaplamalı maliyet analizi

CSP, geleneksel mekansal enterpolasyon yönteminde DO yerine bu çalışmadaki gerekli tüm hesaplamaları sağlar. Çünkü bulut bilişim geleneksel sistemden daha verimli ve ekonomiktir. Normal bulut bilişim ile de gizlilik sorunu var. Ancak bizim fikrimiz, bu senaryodaki tüm katılımcıların mahremiyetini, homomorfik şifreleme ve SMC yardımıyla sağlar. DO, CSP ve C'nin gizliliğini artırmak için ek algoritmalar kullanmak, artan hesaplama gereksinimi ile biter. Hesaplamadaki küçük gecikme miktarı sistemi etkilemez. Çünkü mekansal enterpolasyon sistemleri zor bir gerçek zamanlı sistem değildir.

Verilerin şifreli formda olduğu mekansal verilere dayanarak DO tarafından oluşturulmuş bir kd ağacı. kd ağacını CSP'ye gönderin. CSP, ORAM sunucusu gibi alınan verileri saklar. C, q'nun C'nin ortak anahtarıyla şifrelenmiş olduğu bir CSP'ye q gönderdiğinde. Daha sonra CSP, $\prod_{i=1}^{k} \xi(z_i) = \xi(z_1) * \xi(z_2) * \dots \xi(z_k)$ değerini tahmin değeri yerine hesaplar ve gönderir. Çünkü, saklanan veriler DO'nun ortak

anahtarı ile şifrelenir. Bu yüzden CSP tahmini yapamaz. CSP'den gelen değeri deşifre eder ve tahmin değerini hesaplar. Burada C'nin q değerini tahmin edemez. Bundan sonra, C tarafından paylaşılan anahtar ile DO tarafından şifrelenen öngörülen değer ve CSP'ye gönderilir. CSP, DO'dan C'ye şifrelenmiş değerleri iletir. Son olarak C tahmini değeri deşifre eder. Geleneksel yöntemde C'den bir istek sorgusu alır ve tahmin değerini hesaplar ve sonucu tekrar C'ye gönderir. Ancak bizim senaryoda, genel olarak üç şifreleme, iki şifre çözme ve iki matematiksel hesaplama vardır. Bu aktiviteler geleneksel yöntem yerine matematik maliyetini arttırdı.

5.1.1.2 İletişim maliyeti analizi

Verileri bir CSP'ye dış kaynak aktarırken iletişim maliyeti artacaktır. Şifrelenmiş verileri CSP'ye gönderir. CSP en yakın noktaları hesaplar ve bunları DO'ya gönderir. CSP tarafından gönderilen en yakın noktadan DO tarafından tahmin değeri hesaplanır. DO tahmin değerini şifreler ve CSP'ye geri gönderir. C, CSP'den şifrelenmiş tahmini değeri alır. Bizim önerimiz bir mekansal veri tablosu, k en yakın nokta ve bir tahmin değeri oluşturur. Genel olarak, DO, CSP ve C arasında beş iletişim etkinliği var. Ancak, geleneksel sistemde DO ve C arasında sadece iki iletişim vardır.

5.1.1.3 Depolama maliyeti analizi

Depolama kapasitesi geleneksel olandan iki kat büyüktür. DO tüm verileri kendi içinde geleneksel yöntemlerle depolar. Ancak, önerilen sistemimizde DO, bir düz metin veritabanına sahiptir ve CSP, tüm verilerin şifrelenmiş düz metin biçiminde olduğu bir veritabanına da sahiptir. Bunun sonucu olarak, sistemimiz 2 x (gerçek veritabanı boyutu) gerektirir. Bununla birlikte, önerilen sistem saklanan verilerin gizliliğini ve güvenliğini sağlar.

5.1.2 Doğruluk analizi

Geçmiş gizlilik koruma yöntemlerinin çoğu, doğruluk kaybına yol açmaktadır. Agrawal ve Srikant'ın (2000) çalışmalarında, veri pertürbasyon adı verilen bir teknik önermektedirler. Bu teknik, sahibin verilerinin gizliliğini sağlamak için ek gürültü noktaları ekliyor. Ancak, bu, öngörülen değerin doğruluğu kaybına yol açar. Ancak, Paillier sistemi adı verilen özel bir şifreleme sistemi kullanıldı. Paillier'de, matematiksel hesaplama, şifreleme metninde yapılabilir ve sonuç, sonucu etkilemez. Paillier sisteminin iki özel özelliği vardır; homomorfik ekleme ve çarpma. Böylece, işimiz geleneksel plan gibi aynı tahmin üretir. Bunun ötesinde, tüm katılımcıların özel verilerinin gizliliğini garanti eder.

5.1.3 Gizlilik analizi

Çalışmamızda, verilerin üç katılımcıdan dört farklı yerde korunması gerekmektedir. İlk örnek CSP, şifrelenmiş bir formatta CSP'de saklanan verileri DO anahtarıyla şifreleyemez veya ölçemez. İkinci adım, C'nin kendi isteklerini CSP ve DO'dan gizlemek için şifreli bir biçimde göndermesidir. Üçüncüsü, CSP, istek koordinatının k-NN noktalarını bulmak için sorguyu işler. Şu anda CSP değeri tahmin edemez. Çünkü, tüm veriler DO'nun anahtarı kullanılarak şifrelenmiştir. CSP'nin DO'ya gönderdiği k-NN noktalarını belirledikten sonra. DO şifresini çözer ve tahmin değerini alır. Bu dördüncü adımda DO, sorgu koordinatı hakkında bilgi alamaz. Bu nedenle, DO istenen koordinatın yerini tahmin edemez. Son olarak, DO tahmini değeri C'nin ortak anahtarı ile şifreler ve CSP'ye gönderir. Sadece C kendi özel anahtarı ile şifresini çözebilir. CSP'nin öngörülen değeri bulmak için herhangi bir yolu yoktur. CSP şifre metnini C'ye iletir. C, alınan şifre metninin şifresini çözecek ve tahmini değeri alacaktır. Özet olarak DO, C noktasından sorgu noktasını tahmin edemez. CSP, saklanan verileri DO, C noktasındaki sorgu noktasını ve DO tarafından tahmin edilen değeri tahmin edemez ve / veya ölçemez. Son olarak, C sadece nihai tahmin edilen değeri DO'dan alabilir.

5.2 İki Taraf için Önerilen Çözümün Analizi

5.2.1 Tamamlayıcı maliyet analizi

Geleneksel mekânsal enterpolasyonda, veri sahibi bir müşteri bir tahmin veya talepte bulunma gibi bir tahmin yapmak istediğinde, tahminleri hesaplamak ve müşteri ile iletişimi yapmak için tüm hesaplamaları ve hizmetleri yapar. Tersine, teknik saklanan veriler için mahremiyet ve güvenlik sağlamaz ve ayrıca öngörülen değeri müşteriye transfer ederken gizlilik ve güvenliği garanti eden bir çözüm önerildi. Saklanan veriler için gizlilik ve güvenlik sağlamak ve verileri aktarırken güvenli çoklu-parti hesaplama, R-ağacı ve gizlilik koruma teknikleri kullanıldı. Ek teknikler yüzünden, çözümümüz en erken olandan ziyade biraz daha fazla depolama, hesaplama ve iletişim maliyeti gerektirir. Bu ek ihtiyaçlar tüm sistemi etkilemez. Çünkü mekansal enterpolasyon tekniği gerçek zamanlı bir sistem değildir.

5.2.1.1 Hesaplamalı maliyet analizi

Geleneksel yöntemi ve bizim önerdiğimiz çözümü karşılaştırdığımızda, DO geleneksel sistemdeki tüm çalışmaları yapar, ancak çoğu faaliyetler çözümümüzde CSP tarafından yapılır. Geleneksel yöntemden ziyade bulut bilişim daha verimli ve ekonomiktir. Bu arada, bulut bilişimde de gizlilik ve güvenlik sağlamak da var. Sorunun üstesinden gelebilmek için güvenli çoklu taraf hesaplama (SMC) ve homomorfik şifreleme (HE) kullanıldı. Bu ek teknikler, tüm katılımcılarda ve senaryodaki tüm iletişimde gizliliği ve güvenliği sağlar. Bununla birlikte, geleneksel olandan ziyade biraz daha hesaplamalı güç gerektirir. SMC ve homomorfik gibi ek teknikler nedeniyle. Bu ek hesaplama gücü, sistemimizin genel performansını etkilemez. Çünkü büyük veri analizi zor bir gerçek zamanlı sistem değildir.

Geleneksel yöntemde, veri sahibi istemciden istek sorgusunu alır ve tahmin hesaplamasını yapar ve müşteriye nihai tahmin değerini geri gönderir. Geleneksel yöntemin aksine, matematiksel çalışmanın çoğu CSP tarafından yapılır. İlk olarak, veri

sahipleri (DO1 & DO2) şifrelenmiş uzamsal verilerle kendi R-ağaçlarını yaratır ve R-ağaçlarını ilgili bulut servis sağlayıcısına (CSP1 & CSP2) gönderir. CSP'ler, alınan şifrelenmiş verileri veri sahiplerinden depolar. Müşteri, CSP'lere iki istek gönderir (ör. CSP1 ve diğeri CSP2 için bir tane). İstemci, istek noktasını CSP1'e gönderdiğinde, sorgu noktası DO1'in ortak anahtarı tarafından şifrelenir ve sorgu, CSP2'ye gönderilirken DO2'nin ortak anahtarı tarafından şifrelenir. Her CSP, kendi en yakın noktalarını bulur ve bunları kendi veri sahibine gönderir. Tamamen, veri sahipleri tarafından alınan 2 * k puan. Son k en yakın noktaları hesaplamak için, DO2 k noktalarını DO1'e gönderir. DO2'den k puanları alındıktan sonra, DO1 son k puanlarını hesaplar ve DO2'ye bildirir, hangi noktalar son k altında gelir. DO1 ve DO2, seçilen noktaları ilgili CSP'lerine nihai k değeri altında gönderir. CSP1, homomorfik hesaplamayı yapar $\xi_{DO1}(z_1^1) \cdot \xi_{DO1}(z_2^1) \cdot \xi_{DO1}(z_n^1)$ ve CSP2, $\xi_{DO1}(z_1^2) \cdot \xi_{DO1}(z_2^2) \cdot \xi_{DO1}(z_m^2)$ tahmin değerine alternative değeri hesaplar. Çünkü, CSP'lerde saklanan veriler şifrelenmiş bir formdadır. Sadece, veri sahipleri tahmin değerini bulmak için verilerin şifresini çözebilir. DO, hesaplanan z değerlerini ve DO1 ve DO2 buluntularını alır $(z_1^1 + z_2^1, \ldots, z_n^1) = \sum_{i=1}^{n} z_i^1$ ve $(z_1^2 + z_2^2, \ldots, z_n^2) = \sum_{i=1}^{m} z_i^2$ sırasıyla. DO2, QO'nun genel anahtarı tarafından şifrelenen bir şifre metni olarak $\sum_{i=1}^{m} z_i^2$ değerini gönderir. DO1, DO2'den değeri alır ve $\sum_{i=1}^{n} z_i^1 + \sum_{i=1}^{m} z_i^2$ yi bulmak için homomorfik çarpmayı yapar. Hesaplanan değer QO'ya gönderilir ve QO alınan değeri deşifre eder ve nokta için son tahmin değerini (x_p, y_p) elde etmek için k'ye böler.

Genel olarak, sistemimiz tarafından yapılan sekiz şifreleme, beş şifre çözme ve sekiz matematiksel işlem bulunmaktadır. Geleneksel sistemle karşılaştırıldığında, bazı matematiksel maliyetleri arttırır. Ancak, bu maliyet artışı, mekansal enterpolasyon tekniklerinin performansını etkilemez.

5.2.1.2 İletişim maliyeti analizi

Bir sistem için dış kaynaklı bir veritabanı kullanıldığında iletişim maliyeti artacaktır. Çünkü, dahili veritabanı sisteminden daha fazla iletişime ihtiyaç duyuyor. Önerilen çözümümüzde, öncelikle tüm veri sahipleri veritabanlarını ilgili CSP'lerine gönderir.

İstemci, her CSP için şifreli bir biçimde ayrı bir sorgu gönderir. CSP'ler k en yakın noktaları hesaplar ve bunları ilgili veri sahibine gönderir. Son k en yakın noktaları bulmak için DO2, k noktalarını DO1'e gönderir. DO1, son k değerini 2 * k noktasından hesaplar ve k noktalarından kendi noktalarını seçer ve DO2 ve CSP'lere bilgi verir, noktalar son k'ye aittir. Daha sonra, DO1 ve DO2 seçili noktalarını ilgili CSP'lerine şifreli bir biçimde gönderir. Onlar, z değerini hesaplar ve veri sahiplerine geri gönderir. DO1 ve DO2, puanlarının tahmin değerini almak için alınan değeri şifreler. Daha sonra, DO2, nihai tahminin hesaplanması için tahminini DO1'e gönderir. DO1 son tahmini hesaplar ve tahmin edilen değeri şifrelenmiş biçimde müşteriye gönderir. Son olarak, istemci, sorgu noktası için son tahmin değerini almak için özel anahtarıyla iletinin şifresini çözer. Genel olarak, önerilen çözümümüzde gerekli olan on bir iletişim süreci vardır.

5.2.1.3 Depolama maliyeti analizi

Genel olarak, mekansal enterpolasyon yöntemi tüm verileri kendi içinde saklar. Önerilen sistemimiz, verileri depolamak için bulut bilişim kullanıyor. Veri sahipleri, düz metni kendi başlarına ve bulutlarda saklanan şifrelenmiş verileri depolar. Sistemimiz 2 x (gerçek veritabanı boyutu) gerektirir. Üstelik, sistemimiz her zaman mahremiyet ve güvenliği sağlıyor.

5.2.2 Doğruluk analizi

Agrawal ve Srikant (2000), mahremiyetin veri pertürbasyonu olarak adlandırılması için bir teknik önermektedir. Burada fikir, verilerin gizliliğinden emin olmak için sahibine ait verilere bazı ek gürültülü noktalar eklemektir. Bununla birlikte, bu teknik tahmin doğruluğu kaybına yol açmaktadır. Paillier sistemi (Paillier 1999) nihai çıktıyı etkilemeden şifre metninde bazı matematiksel hesaplamalara izin verir. Önerilen çözümümüzde Paillier sistemi kullanıldı. Böylece, geleneksel olana benzer bir tahmin verir. Dahası, tüm koşullarda verilerin gizliliğini ve güvenliğini sağlar.

5.2.3 Gizlilik analizi

Önerilen çözümümüz, çeşitli yerlerdeki verilerin korunmasını sağlıyor. Öncelikle, verileri CSP'de saklamak için, veriler ilgili veri sahibinin ortak anahtarı ile şifrelenir. CSP'ler için saklanan şifre metninden orijinal verileri şifresini çözmek veya hesaplamak için bir yol yoktur. İstemci sorguları CSP'lere gönderdiğinde, ilgili veri sahibinin ortak anahtarı ile şifrelenir. Bunun sonucu olarak, CSP istenen sorgu noktasının yerini ölçemez. CSP'ler sorgu için en yakın noktaları hesaplar. CSP değeri tahmin edemediğinde. Çünkü saklanan veriler ilgili veri sahibinin ortak anahtarı ile şifrelenir. Veri sahipleri, kendi CSP'lerinden şifrelenmiş bir biçimde k puanları alırlar. Veri sahipleri, sorgulanan değerlerden sorgu noktasını tahmin edemezler. Çünkü alınan noktalar sadece konum değildir. Son k puanlarını bulduktan sonra, DO1 ve DO2, ilgili CSP'lerine göndermeden önce, son k'da bulunan seçtikleri noktaları şifreler. CSP'ler şifrelenmiş değeri alır ve z değerini hesaplar ve bunları sahiplerine geri gönderir. DO1 ve DO2, alınan z değerlerini gizli anahtarlarıyla şifreler ve DO2, nihai tahmini hesaplamak için z değerini DO1'e gönderir. DO1, nihai tahmini hesaplar ve müşterinin genel anahtarı ile şifreleyebilir. Son olarak, müşteri tahmini almak için tahminini gizli anahtarıyla çözer. Tüm durumlarda, çözümümüz verilerin gizliliğini ve güvenliğini sağlar.

6. SONUÇ

Son on yılda en popüler teknolojilerden biri bulut bilişimidir (Buyya vd. 2009, Mell ve Grance 2011). Bulutta dış kaynak kullanımı ile elde edilen verilerin maliyet, esneklik, kullanılabilirlik ve kuruluşlar ve şirketler için yönetimlerden daha az masrafla depolanması, erişilmesi ve işlenmesi gibi hizmetler sunulması nedeniyle daha popüler hale gelmiştir. Öte yandan, bulutların verilerini dışarıya aktarmak için bulutları işe alma konusunda isteksiz olan kuruluşlar, verilerin gizliliği ve güvenliği bulutlarda önemli sorunlardan biridir. Bunun sonucu olarak, araştırmacılar tarafından verileri bir bulutta dış kaynaklar ile etileşimden önce şifrelenmesi önerildi. Ancak, bir sorgu için işlenmeden önce tüm verilerin deşifre edilmesi gerekir. Böylece, geleneksel şifreleme teknikleri, dış kaynaklı verileri bir bulutta korumaya uygun değildir. Burada ortaya çıkan soru, verinin sürece dahil olan tüm taraflardan nasıl korunabileceğiydi.

Bu çalışmada, iki farklı problem türü tespit edilmiştir; 1) tek - taraflı problemi ve 2) iki - taraflı problemidir. Dış kaynaklı mekansal verileri yarı-dürüst bir bulutta korumak için iki farklı çözüm önerdik.

Bölüm 3'te, tek - taraflı problemi için bir çözüm önerdik. Mekansal enterpolasyon, coğrafi bilgi sistemlerinin temel işlemlerinden biridir. Araştırmacılar tarafından kullanılan çeşitli enterpolasyon yöntemleri vardır. Hesaplamadaki basitliği ve verimliliği nedeniyle k - NN interpolasyon yöntemi tercih edilmektedir. Önerilen çözüm, DO, CSP ve QO arasındaki tüm iletimlerde gizliliğin ve güvenliğin sağlanmasını sağlar. Bununla birlikte, kd - ağacı, ORAM ve homomorfik şifreleme gibi ek algoritmaların kullanılması ek hesaplama, depolama ve iletişim maliyetine yol açmaktadır. Bölüm 6, bu terimlerdeki maliyeti analiz etmiş ve tüm katılımcıların bakış açısından bu ek maliyetin tüm çözümü etkilemeyebileceğini ileri sürmektedir.

Bölüm 4'te iki - taraflı bir sorun için çözüm önerilmektedir. Bulut bilişim bilgisayar teknolojisinde ortaya çıktığında verilerin saklanması, erişilmesi ve işlenmesi değişmiştir. Bulut bilişim maliyeti düşürdü, esnek ve idari işi azalttı. Bunun sonucu

olarak, çoğu şirket ve kuruluşun veri kümelerini bulutlara taşındı. Öte yandan, veriler genel bulutta düz metin olarak saklandığında gizlilik ve güvenlik sorunları gözlemlendi. Şifreleme, gizliliği sağlamak için bir yoldur. Ancak, geleneksel şifreleme yöntemiyle oluşturulan şifre metnini işlemek çok zordur. Bu nedenle, bilişim ve iletişim gibi tüm süreçlerde gizliliği ve güvenliği sağlayan bir çözüm k - NN uzamsal interpolasyon tekniğini öneriyoruz. Bizim çözümümüz hesaplama, iletişim ve depolamada küçük bir maliyet artışına sahip olabilir. Bölüm 6'da iletişim, depolama ve hesaplamadaki maliyet artışının, mekânsal enterpolasyonun genel performansını etkilemediğini göstermektedir.

KAYNAKLAR

Abadi, D. J. 2009. "Data management in the cloud: Limitations and opportunities." IEEE Data Eng. Bull. **32**(1): 3-12.

Aggarwal, C. C. and Philip, S. Y. 2008. A general survey of privacy-preserving data mining models and algorithms. Privacy-preserving data mining, Springer**:** 11-52.

Agrawal, R. and Srikant, R. 2000. Privacy-preserving data mining. ACM Sigmod Record, ACM.

Bayardo, R. J. and Agrawal, R. 2005. Data privacy through optimal k-anonymization. Data Engineering, 2005. ICDE 2005. Proceedings. 21st International Conference on, IEEE.

Buyya, R., Yeo, C. S., Venugopal, S., Broberg, J.and Brandic, I. 2009. "Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility." Future Generation computer systems **25**(6): 599-616.

Clifton, C., Kantarcioglu, M., Vaidya, J., Lin, X. and Zhu, M. Y.2002. "Tools for privacy preserving distributed data mining." ACM Sigkdd Explorations Newsletter **4**(2): 28-34.

di Vimercati, S. D. C., Foresti, S. and Samarati, P. 2012. Managing and accessing data in the cloud: Privacy risks and approaches. Risk and Security of Internet and Systems (CRiSIS), 2012 7th International Conference on, IEEE.

Elmehdwi, Y., Samanthula, B. K. and Jiang, W. 2014. Secure k-nearest neighbor query over encrypted data in outsourced environments. Data Engineering (ICDE), 2014 IEEE 30th International Conference on, IEEE.

Evfimievski, A., Srikant, R., Agrawal, R. and Gehrke J. 2004. "Privacy preserving mining of association rules." Information Systems **29**(4): 343-364.

Fienberg, S. E. and McIntyre, J. 2004. Data swapping: Variations on a theme by dalenius and reiss. International Workshop on Privacy in Statistical Databases, Springer.

Ghinita, G., Kalnis, P., Khoshgozaran, A., Shahabi, C. and Tan, K.-L. 2008. Private queries in location based services: anonymizers are not necessary. Proceedings of the 2008 ACM SIGMOD international conference on Management of data, ACM.

Goldreich, O. and Ostrovsky, R. 1996. "Software protection and simulation on oblivious RAMs." Journal of the ACM (JACM) **43**(3): 431-473.

Graepel, T., Lauter, K. and Naehrig, M. 2012. ML confidential: Machine learning on encrypted data. International Conference on Information Security and Cryptology, Springer.

Guttman, A. 1984. R-trees: A dynamic index structure for spatial searching, ACM.

Hacıgümüş, H., Iyer, B. and Mehrotra, S. 2004. Efficient execution of aggregation queries over encrypted relational databases. International Conference on Database Systems for Advanced Applications, Springer.

Hashizume, K., Rosado, D. G., Fernández-Medina, E. and Fernandez, E. B. 2013. "An analysis of security issues for cloud computing." Journal of internet services and applications **4**(1): 5.

Hore, B., Mehrotra, S. and Tsudik, G. 2004. A privacy-preserving index for range queries. Proceedings of the Thirtieth international conference on Very large data bases-Volume 30, VLDB Endowment.

Hore, B., Mehrotra, S., Canim, M. and Kantarcioglu, M. 2012. "Secure multidimensional range queries over outsourced data." The VLDB Journal **21**(3): 333-358.

Hu, H., Xu, J., Ren, C. and Choi, B. 2011. Processing private queries over untrusted data cloud through privacy homomorphism. Data Engineering (ICDE), 2011 IEEE 27th International Conference on, IEEE.

Kantarcıoğlu, M. and Clifton, C. 2004. Privately computing a distributed k-nn classifier. European conference on principles of data mining and knowledge discovery, Springer.

Li, M., Yu, S., Lou, W. and Hou, Y. T. 2012. Toward privacy-assured cloud data services with flexible search functionalities. Distributed Computing Systems Workshops (ICDCSW), 2012 32nd International Conference on, IEEE.

Lindell, Y. and Pinkas, B. 2000. Privacy preserving data mining. Advances in Cryptology—CRYPTO 2000, Springer.

Manolopoulos, Y., Nanopoulos, A., Papadopoulos, A. N. and Theodoridis, Y. 2010. R-trees: Theory and Applications, Springer Science & Business Media.

Mell, P. and Grance, T. 2009. "The NIST definition of cloud computing." National institute of standards and technology **53**(6): 50.

Mell, P. and Grance, T. 2010. "The NIST definition of cloud computing." Communications of the ACM **53**(6): 50.

Mell, P. and Grance, T. 2011. "The NIST definition of cloud computing."

Mykletun, E. and Tsudik, G. 2006. Aggregation queries in the database-as-a-service model. IFIP Annual Conference on Data and Applications Security and Privacy, Springer.

Paillier, P. 1999. Public-key cryptosystems based on composite degree residuosity classes. International Conference on the Theory and Applications of Cryptographic Techniques, Springer.

Pearson, S., Shen, Y. and Mowbray, M. 2009. A privacy manager for cloud computing. IEEE International Conference on Cloud Computing, Springer.

Qi, Y. and Atallah, M. J. 2008. Efficient privacy-preserving k-nearest neighbor search. Distributed Computing Systems, 2008. ICDCS'08. The 28th International Conference on, IEEE.

Ravu, S., Neelakandan, P., Gorai, M., Mukkamala, R. and Baruah, P. 2012. A computationally efficient and scalable approach for privacy preserving kNN classification. IEEE International Conference on High Performance Computing (HiPC).

Rodrigues, J. J., De La Torre, I., Fernández, G. and López-Coronado, M. 2013. "Analysis of the security and privacy requirements of cloud-based electronic health records systems." Journal of medical Internet research **15**(8).

Ryan, M. D. 2013. "Cloud computing security: The scientific challenge, and a survey of solutions." Journal of Systems and Software **86**(9): 2263-2268.

Samanthula, B. K., Elmehdwi, Y. and Jiang, W. 2015. "K-nearest neighbor classification over semantically secure encrypted relational data." IEEE transactions on Knowledge and data engineering **27**(5): 1261-1273.

Shaneck, M., Kim, Y. and Kumar, V. 2009. Privacy preserving nearest neighbor search. Machine Learning in Cyber Trust, Springer**:** 247-276.

Shi, E., Bethencourt, J., Chan, T. H., Song, D. and Perrig, A. 2007. Multi-dimensional range query over encrypted data. Security and Privacy, 2007. SP'07. IEEE Symposium on, IEEE.

Shi, E., Chan, T.H. H., Stefanov, E. and Li, M. 2011. Oblivious RAM with O ((logN) 3) Worst-Case Cost. Asiacrypt, Springer.

Wang, B., Hou, Y., Li, M., Wang, H. and Li, H. 2014. Maple: scalable multi-dimensional range search over encrypted cloud data with tree-based index. Proceedings of the 9th ACM symposium on Information, computer and communications security, ACM.

Williams, P., Sion, R. and Carbunar, B. 2008. Building castles out of mud: practical access pattern privacy and correctness on untrusted storage. Proceedings of the 15th ACM conference on Computer and communications security, ACM.

Wong, W. K., Cheung, D. W.I., Kao, B. and Mamoulis N. 2009. Secure knn computation on encrypted databases. Proceedings of the 2009 ACM SIGMOD International Conference on Management of data, ACM.

Xiong, L., Chitti, S. and Liu, L. 2006. K nearest neighbor classification across multiple private databases. Proceedings of the 15th ACM international conference on Information and knowledge management, ACM.

Yao, B., Li, F. and Xiao, X. 2013. Secure nearest neighbor revisited. Data Engineering (ICDE), 2013 IEEE 29th International Conference on, IEEE.

Yi, X., Paulet, R. and Bertino, E. 2014. Homomorphic encryption and applications, Springer.

Zhang, P., Tong, Y., Tang, S. and Yang, D. 2005. Privacy preserving naive bayes classification. International Conference on Advanced Data Mining and Applications, Springer.

Zhu, Y., Xu, R. and Takagi, T. 2013. Secure k-NN computation on encrypted cloud data without sharing key with query users. Proceedings of the 2013 international workshop on Security in cloud computing, ACM.

ÖZGEÇMİŞ

Adı Soyadı : Muhammad Rifthy KALIDEEN
Doğum Yeri : Kinniya, Sri Lanka
Doğum Tarihi : 23.06.1986
Medeni Hali : Evli
Yabancı Dili : İngilizce, Türkçe

Eğitim Durumu (Kurum ve Yıl)

Lise: T / Kinniya Central College. (2015)
Lisans: Jaffna Üniversitesi Vavuniya Kampüsü, Uygulamalı Bilimler Fakültesi, Fiziksel Bilimler Bölümü (2010)
Yüksek Lisans: Ankara Üniversitesi, Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı (Eylül 2015 – Eylül 2018)

Çalıştığı Kurum/Kurumlar ve Yıl

Öğretim üyesi, Sri Lanka'nın Güney Doğu Üniversitesi, 2012 – Devam ediyor
Yardımcı Öğretim Görevlisi, Jaffna Üniversitesi Vavuniya Kampüsü, Ocak 2012 – Kasım 2012
Araştırma Görevlisi, Jaffna Üniversitesi Vavuniya Kampüsü. 2010 – Ocak 2012

Hakemli Dergiler

Muhammad Rifthy Kalideen and Bulent Tugrul, "Outsourcing of Secure k-Nearest Neighbours Interpolation Method" International Journal of Advanced Computer Science and Applications(IJACSA), 9(4), 2018.

Ulusal Kongre Sunum

Muhammad Rifthy, **K.**, Kirushanth, S., Arthika, S., Nishanthy, S., and Kanaganathan, S. "*Developing a Decision Support System to manage Solid wastes of Local Authority*", Annual Science Research Session (ASRS) - 2012, South Eastern University of Sri Lanka, p13, 2012.

Uluslararası Kongre Sunum

Ilham Safeek, Muhammad Rifthy Kalideen, "Preprocessing On Facebook Data For Sentiment Analysis", 7th International Symposium, South Eastern University of Sri Lanka, Pp 69-78, 2017

Roshan Karunarathna, K.M.M., Rakkitha Samaraweera, M.A.D., and **Muhammad Rifthy, K.** "Developing an e – Tool to do teaching and learning for undergraduates in Vavuniya Campus", 03rd International Symposium - 2013, South Eastern University of Sri Lanka, Vol.2 , Pp 86-89, 2013.