Data & Knowledge Engineering 104 (2016) 32–44

Contents lists available at ScienceDirect

Data & Knowledge Engineering

journal homepage: www.elsevier.com/locate/datak

Editorial

Hilbert curve-based cryptographic transformation scheme for spatial query processing on outsourced private data

Hyeong-Il Kim, Seungtae Hong, Jae-Woo Chang *

Dept. of Computer Engineering, Chonbuk National University, Jeonju, Republic of Korea

ARTICLE INFO

Available online 8 May 2015

Keywords:
Database outsourcing
Data privacy
Spatial data cryptographic scheme
Query processing
Hilbert curve

ABSTRACT

Research on preserving location data privacy in outsourced databases has been spotlighted with the development of cloud computing. However, the existing spatial transformation schemes are vulnerable to various attack models. The existing cryptographic transformation scheme provides good data privacy, but it has a high query processing cost. To improve privacy and reduce cost, we propose a Hilbert curve-based cryptographic transformation scheme to preserve the privacy of the spatial data from various attacks on outsourced databases. We also provide efficient range and *k*-NN query processing algorithms using a Hilbert-order index. A performance analysis confirms that the proposed scheme is robust against attack models and achieves better query processing performance than the existing cryptographic transformation scheme.

1. Introduction

With the development of cloud computing, research on outsourced databases has been actively pursued. In outsourced databases, a data owner (an individual or a company) outsources to a service provider system resources such as data, hardware, and software as well as their management in order to focus on his/her core business area. The service provider (SP) manages the resources and allows authorized users (AUs) to access the outsourced database for issuing queries. The outsourced database has the advantage of allowing the data owner (DO) to easily manage the data at a low cost even if the DO lacks experience related to IT services.

Meanwhile, location-based services (LBSs) have been popular with the rapid spread of mobile devices equipped with GPS [1]. LBSs include navigational systems, friend finder, e-commerce, searching a Point of Interest (POI), such as restaurants, hospitals, shopping malls, etc. Accordingly, the amount of spatial data has been explosively increasing [2]. However, traditional LBS systems with a single server cannot handle the large amount of spatial data that is common in real application scenarios. As the amount of spatial data increases, more storage and more computational costs are required to manage it. A single server with limited resources has a low performance for handling a large amount of spatial data. Therefore, the need to outsource spatial databases to a cloud system is increasing.

However, when the DO outsources the original spatial database, some privacy threats occur because the spatial data contains private information of the DO [3]. Another reason is that an AU needs to send queries with his/her exact location to take advantage of LBSs. If the location information has been disclosed to an attacker, or a service provider abuses the information for a malicious purpose, the attacker can find places where users have frequently visited and the dates they have done so. In this case, the visit pattern and lifestyle of individuals can be revealed.

To solve this problem, research on protecting data privacy of outsourced spatial databases in the cloud system is being actively carried out. The existing methods for protecting spatial data privacy in outsourced databases fall into two categories: spatial

* Corresponding author.
E-mail address: jwchang@jbnu.ac.kr (J.-W. Chang).

http://dx.doi.org/10.1016/j.datak.2015.05.002
0169-023X/

transformation schemes and encryption schemes. First, there are several spatial transformation schemes [4–12] that convert original data into other data. However, these methods are vulnerable to attack because they use simple transformation equations to convert the data. For example, an attacker can discover the transformation equation with some pairs of original data and its corresponding transformed data. In the worst case, the attacker can find all the original spatial data from the transformed data by using the revealed transformation equation. Secondly, encryption schemes [10,13–16] solve the problems of spatial transformation schemes by encrypting the original data points. However, most encryption schemes [14–16] cannot provide range and *k*-NN query processing, which are popular query types for the spatial databases. The encryption schemes that can support range and *k*-NN query processing have a high query processing cost [10,13].

For outsourcing spatial databases, the transformation method should preserve data privacy and support efficient query processing. However, to the best of our knowledge, there is no work that balances the trade-off between data privacy and query processing performance. To solve these problems, in this paper we propose a Hilbert curve-based cryptographic transformation scheme (HCT) that preserves data privacy of the outsourced spatial databases in the cloud system. In our method, we design a Hilbert aggregation index to enhance the efficiency of the query processing. The proposed method reduces the number of message transmissions for query processing and minimizes the size of the communication message by performing local data grouping based on the Hilbert-curve order [17]. Our contributions can be summarized as follows:

- We present a framework for providing the confidentiality of spatial data outsourced to a cloud system.
- We propose a Hilbert curve-based cryptographic transformation scheme (HCT) that encrypts the spatial data using an AES algorithm [18] to provide data privacy preservation.
- We design a Hilbert-curve index that satisfies both data privacy preservation and efficient query processing properties. To the best of our knowledge, our work is the first work that balances the trade-off between data privacy and query processing performance.
- We present range and *k*-NN query processing algorithms for transformed data using our HCT. The algorithms provide good performance by making use of the Hilbert-order index.
- We also present an extensive experimental evaluation of our scheme by comparing it with the existing studies. The evaluation confirms that our scheme provides good query processing performance while protecting the privacy of the spatial data against various attacks.

The rest of the paper is organized as follows. In Section 2, we introduce the existing methods for spatial data privacy preservation, especially in the outsourced spatial databases. Section 3 presents overall system architecture and details of our Hilbert-curve cryptographic transformation scheme. In Section 4, we propose a range query processing algorithm and a *k*-NN query processing algorithm for encrypted spatial databases. In Section 5, we compare the performance of our proposed method with that of the existing methods. Finally, we conclude this paper with future work in Section 6.

2. Related work

In this section, we first review preservation schemes for spatial data privacy in LBSs where the original spatial databases are published to a service provider. Next, we describe existing spatial transformation schemes and encryption schemes for preservation of spatial data privacy in outsourced databases.

2.1. Location privacy in LBSs

In LBSs, a mobile user issues spatial queries that are processed by an SP. However, the user does not want to reveal his/her exact location information to the SP. This is because if an attacker or a malicious LBS provider abuses this information, the user's private information like lifestyle and health issues can be disclosed. To protect the users location privacy, there are several schemes that blur the location of the user by using the k-anonymity property. Mokbel et al. [19] proposed the New Casper scheme which creates a cloaking area using a grid-based data structure. The cloaking area satisfies the *k*-anonymity property by including not only the user who requests the query, but also the *k*-1 other users nearby him/her. Lee et al. [20] proposed the GCC scheme which creates a cloaking area by calculating the privacy protection level of the cloaking area using entropy. Kim et al. [21] proposed a cloaking area creation scheme where a query issuer generates a cloaking area in a distributed manner by collaborating with other mobile users. Wang et al. [22] proposed the XStar scheme which creates a cloaking area by considering the road networks. Kim et al. [23] proposed a *k*-NN query processing algorithm in road networks. For this, a set of road segments generated by XStar is considered as a query region to hide the exact location of a user. To process a query for the region, they designed an Island-index [24] where each node stores its neighboring POIs. However, these schemes just hide the users exact location information at the query time. Thus, an attacker or malicious SP can abuse the original spatial databases which are published to the SP. Therefore, the original spatial databases should be transformed into another form before the databases are outsourced to the SP.

2.2. Spatial transformation scheme

The typical spatial transformation schemes that protect the spatial data in the outsourced databases are as follows. Hacigumus et al. first introduced the idea of outsourcing a database to a third-party service provider in [4] and addressed the confidentiality of the outsourced database in [5]. Evfimievski et al. [6] proposed a random noise addition approach. Gutscher et al. [7] proposed a spatial transformation scheme that converts the spatial data points by performing a parallel translation of the plane coordinate. However,