

Chapter XI

Privacy in Trajectory Data

Aris Gkoulalas-Divanis

University of Thessaly, Greece

Vassilios S. Verykios

University of Thessaly, Greece

ABSTRACT

In this era of significant advances in telecommunications and GPS sensors technology, a person can be tracked down to proximity of less than 5 meters. This remarkable progress enabled the offering of services that depend on user location (the so-called location-based services—LBSs), as well as the existence of applications that analyze movement data for various purposes. However, without strict safeguards, both the deployment of LBSs and the mining of movement data come at a cost of privacy for the users, whose movement is recorded. This chapter studies privacy in both online and offline movement data. After introducing the reader to this field of study, we review state-of-the-art work for location and trajectory privacy both in LBSs and in trajectory databases. Then, we present a qualitative evaluation of these works, pointing out their strengths and weaknesses. We conclude the chapter by providing our point of view regarding the future trends in trajectory data privacy.

INTRODUCTION

Privacy is a fundamental right for every human being. As well stated in Alderman & Kennedy (1997), “Privacy covers many things. It protects the solitude necessary for creative thought. It allows us the independence that is part of raising a family. It protects our right to be secure in our own homes and possessions, assured that the government cannot come barging in. Privacy

also encompasses our right to self-determination and to define who we are. Although we live in a world of noisy self-confession, privacy allows us to keep certain facts to ourselves if we so choose. The right to privacy, it seems, is what makes us civilized”. However, as technology increasingly permeates society, new means are becoming available that in the wrong hands can lead to the breach of people’s privacy. The currently available technological equipment along with the recent

achievements in computational analysis, enable the collection, storage and in-depth analysis of huge piles of personal data in a matter of a few minutes. As an effect, it has become now more important than it was ever before, to identify new ways to protect the individuals' right to privacy. This chapter deals with a specific form of privacy that involves the collection and analysis of location and trajectory data, and discusses privacy preservation techniques by means of modern technology. The location/trajectory privacy, as is called, refers to the right of individuals to keep their whereabouts unknown to untrusted third entities. In what follows, we discuss the currently available means for the collection and the analysis of both location and trajectory data, we indicate the severe threats that these can pose to user privacy and we motivate the rest of this work.

Database management systems (DBMSs) are typically capable of collecting, storing and analyzing millions of data in an efficient manner. In the majority of the cases, the involved data is relational, consisting of transactions that are building upon a set of basic data types (e.g., integer/floating point numbers and variable length strings) and organized in tables with references among them. To efficiently handle this data, the database engine contains various functions that can be applied either to manipulate or to query the stored information. Soon enough, the necessity for storing and analyzing more complex information, than the one adhering to the basic types, became apparent and led the DBMS vendors to upgrade the functionality of their products. A first major extension regarded the handling of temporal data, such that the stored transactions can carry timestamps denoting the date and/or time of an event. This extension made possible e.g., the issuing of queries regarding events that took place over a given period of time or calculations involving date/time data. It also made possible the storing and efficient manipulation of time series data, such as information related to weather forecasting, to banking transactions or to

financial data. A second, and much more recent, extension of DBMSs regarded the appropriate handling of spatial data. This was made possible through the use of spatial geometries that support either Euclidean (flat-earth) data, or geography, ellipsoidal (round-earth) data, involving GPS latitude and longitude coordinates. To allow for the storing and manipulation of spatial data, a new data type, known as (*spatial*) *geometry*, was employed in the database products of vendors such as Oracle, IBM and Microsoft. Through the use of this data type, a transaction can contain information regarding e.g., the location of a store, or the extent of a football field. Furthermore, the handling of the spatial data through the DBMS functions, allows for creating queries that return the nearest neighbors of a given geometry, depict the spatial relations that exist among a set of geometries (e.g., overlap, disjoint, contains, etc) or calculate non-spatial information, such as the perimeter or the area of a geometry. The most recent extension of DBMSs is in academic stage and regards the handling of spatio-temporal data, where an entity is referenced both in space and time (Pelekis & Theodoridis 2006). This extension allows a DBMS to manipulate movement data, such as moving vehicles, and is expected to be officially supported in the future releases of all the major commercial database products.

Apart from the advances in the database industry, significant advances have also taken place in telecommunications and in GPS sensors technology. The offering of personalized Location Based Services (LBSs) is one of the most remarkable evidence of this progress by allowing mobile subscribers, equipped with positioning devices, to access a set of high-end spatially-aware services. In the typical use of such a service, the mobile subscriber communicates his or her position to the telecom provider and asks for a location based service, such as the nearest pharmacy or routing instructions to a certain place (e.g., a restaurant). Upon receiving this information, the telecom provider performs a query to the database to

12 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/privacy-trajectory-data/29151

Related Content

A Survey of Open Source Tools for Business Intelligence

Christian Thomsen and Torben Bach Pedersen (2009). *International Journal of Data Warehousing and Mining* (pp. 56-75).

www.irma-international.org/article/survey-open-source-tools-business/3896

A Novel Hybrid Correlation Measure for Query Expansion-Based Information Retrieval

Ilyes Khennak and Habiba Drias (2020). *Critical Approaches to Information Retrieval Research* (pp. 1-19).

www.irma-international.org/chapter/a-novel-hybrid-correlation-measure-for-query-expansion-based-information-retrieval/237637

Advances in Classification of Sequence Data

Pradeep Kumar, P. Radha Krishna, Raju S. Bapi and T. M. Padmaja (2008). *Data Mining and Knowledge Discovery Technologies* (pp. 143-174).

www.irma-international.org/chapter/advances-classification-sequence-data/7517

Variations on Associative Classifiers and Classification Results Analyses

Maria-Luiza Antonie, David Chodos and Osmar Zaiane (2009). *Post-Mining of Association Rules: Techniques for Effective Knowledge Extraction* (pp. 150-172).

www.irma-international.org/chapter/variations-associative-classifiers-classification-results/8442

Recommender Systems Based on Resonance Relationship of Criteria With Choquet Operation

Hiep Xuan Huynh, Le Hoang Son, Giap Nguyen Cu, Tri Minh Huynh and Huong Hoang Luong (2020). *International Journal of Data Warehousing and Mining* (pp. 44-62).

www.irma-international.org/article/recommender-systems-based-on-resonance-relationship-of-criteria-with-choquet-operation/265256