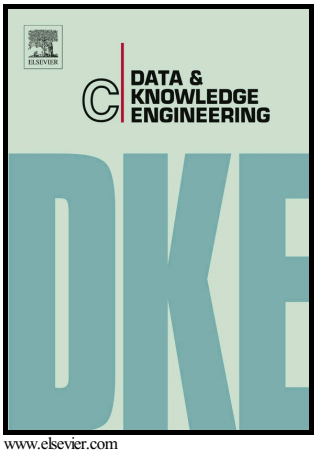

Location disclosure risks of releasing trajectory distances

Emre Kaplan, Mehmet Emre Gursoy, Mehmet Ercan Nergiz, Yucel Saygin

PII: S0169-023X(16)30245-2
DOI: https://doi.org/10.1016/j.datak.2017.10.001
Reference: DATAK1618

To appear in: *Data & Knowledge Engineering*

Received date: 17 October 2016
Revised date: 4 August 2017
Accepted date: 5 October 2017

Cite this article as: Emre Kaplan, Mehmet Emre Gursoy, Mehmet Ercan Nergiz and Yucel Saygin, Location disclosure risks of releasing trajectory distances, *Data & Knowledge Engineering,* https://doi.org/10.1016/j.datak.2017.10.001

Location disclosure risks of releasing trajectory distances

Emre Kaplan[1], Mehmet Emre Gursoy[2], Mehmet Ercan Nergiz[3], Yucel Saygin[1]

Abstract

Location tracking devices enable trajectories to be collected for new services and applications such as vehicle tracking and fleet management. While trajectory data is a lucrative source for data analytics, it also contains sensitive and commercially critical information. This has led to the development of systems that enable privacy-preserving computation over trajectory databases, but many of such systems in fact (directly or indirectly) allow an adversary to compute the distance (or similarity) between two trajectories. We show that the use of such systems raises privacy concerns when the adversary has a set of known trajectories. Specifically, given a set of known trajectories and their distances to a private, unknown trajectory, we devise an attack that yields the locations which the private trajectory has visited, with high confidence. The attack can be used to disclose both positive results (i.e., the victim has visited a certain location) and negative results (i.e., the victim has not visited a certain location). Experiments on real and synthetic datasets demonstrate the accuracy of our attack.

Keywords: Privacy, spatio-temporal data, trajectory data, data mining.

1. Introduction

Location tracking devices such as GPS-equipped vehicles, smartphones and location-based applications have greatly eased the collection of spatio-temporal movement patterns and trajectories. The analysis of this data can help the society, e.g., via traffic management in metropolitan areas, discovery of traffic and passenger flows [1][2], road condition sensing [3] and fleet management. While sharing and mining trajectory data is beneficial for the society, the sensitive nature of location data raises privacy concerns. This has led to substantial research in location privacy [4][5] and privacy-preserving trajectory data management [6][7][8].

Work in the latter can be roughly divided into two camps. In the first camp, data is de-identified and anonymized (e.g., by removing explicit identifiers and

[1]Faculty of Engineering and Natural Sciences, Sabanci University, Istanbul, Turkey.
[2]College of Computing, Georgia Institute of Technology, Atlanta, GA.
[3]Acadsoft Research, Gaziantep, Turkey.