**Privacy Preserving Location Data Publishing: A Machine Learning Approach**

**Abstract**

Publishing datasets plays an essential role in open data research and promoting transparency of government agencies. However, such data publication might reveal users’ private information. One of the most sensitive sources of data is spatiotemporal trajectory datasets. Unfortunately, merely removing unique identifiers cannot preserve the privacy of users. Adversaries may know parts of the trajectories or be able to link the published dataset to other sources for the purpose of user identification. Therefore, it is crucial to apply privacy preserving techniques before the publication of spatiotemporal trajectory datasets. In this paper, we propose a robust framework for the anonymization of spatiotemporal trajectory datasets termed as machine learning based anonymization (MLA). By introducing a new formulation of the problem, we are able to apply machine learning algorithms for clustering the trajectories and propose to use k-means algorithm for this purpose. A variation of k-means algorithm is also proposed to preserve the privacy in overly sensitive datasets. Moreover, we improve the alignment process by considering multiple sequence alignment as part of the MLA. The framework and all the proposed algorithms are applied to TDrive and Geolife location datasets. The experimental results indicate a significantly higher utility of datasets by anonymization based on MLA framework.

**Existing System**

Publication of data by different organizations and institutes is crucial for open research and transparency of government agencies. Just in Australia, since 2013, over 7000 additional datasets have been published on ’data.gov.au,’ a dedicated website for the publication of datasets by the Australian government. Moreover, the new Australian government data sharing legislation encourage government agencies to publish their data, and as early as 2019, many of them will have to do so. Unfortunately, the process of data publication can be highly risky as it may disclose individuals’ sensitive information

.

**Disadvantages**

1.Less accuracy.

**PROPOSED SYSTEM**

Now-a-days due to mobile all online applications are recording user locations and then storing them in their app and this location details they can use to track users. Sometime some malicious users can track the user location by knowing their home address and then they can match the home location with other location details to know where user is travelling like bank, hospital or any other locations. To overcome from this problem author has introduce Machine Learning based data privacy preserving technique

**Advantages**

1.High accuracy

**SYSTEM REQUIREMENTS**

**HARDWARE REQUIREMENTS**:

Processor - Intel i3 or i4

Speed - 1.1 GHz

RAM - 4 GB (min)

Hard Disk - 500 GB (min)

Key Board - Standard Windows Keyboard

Mouse - Two or Three Button Mouse

Monitor - SVGA

**SOFTWARE REQUIREMENTS:**

Operating System - Windows 10 or above

Programming Language - Python