ABSTRACT:

Extracting and analyzing the interesting patterns from spatio-temporal databases, have drawn a great interest in various fields of research. Recently, a number of experiments have explored the problem of spatial or temporal data mining, and some clustering algorithms have been proposed. However, not many studies have been dealing with the integration of spatial data mining and temporal data mining. Moreover, the data in spatial temporal database can be categorized as high-dimensional data. Current density-based clustering might have difficulties with complex data sets including high-dimensional data. This paper presents Iterative Local Gaussian Clustering (ILGC), an algorithm that combines K-nearest neighbour (KNN) density estimation and Kernel density estimation, to cluster the spatiotemporal data. In this approach, the KNN density estimation is extended and combined with Kernel function, where KNN contributes in determining the best local data iteratively for kernel density estimation. The local best is defined as the set of neighbour data that maximizes the kernel function. Bayesian rule is used to deal with the problem of selecting the best local data. This paper utilized Gaussian kernel which has been proven successful in the clustering. To validate the KNN-kernel based algorithm, we compare its performance against other popular algorithms, such as Self Organizing Maps (SOM) and K-Means, on Crime database. Results show that KNN-kernel based clustering has outperformed others.