Abstract

Traffic accidents are an important concern of today's governments and societies, due to the high cost of human and economical resources involved. Data mining has been proven able to significantly help in improving traffic safety. Among several data mining tasks, clustering technique is mostly applied on spatio-temporal data, especially for the traffic data. A number of traffic related works proposed different clustering techniques for mining the spatio-temporal of traffic accident. However, some difficulties appeared when analyzing these datasets, such as the size of data, the lack of statistical evaluation methods, and interpreting the valuable patterns. With regard to solving this problem, this paper proposes a clustering approach for mining spatio-temporal data of fatal accident using local triangular kernel clustering (LTKC) algorithm. LTKC is kernel-density-based clustering algorithm that has the ability to determine the number of clusters automatically. We also propose three visualization techniques for use to interpret and present the optimal clustering result in an easy-understanding form. From the experimental results, LTKC approach was found to be able to discover responsible clusters within fatal accident data, which had proven by silhouette and Dunn index values close to 1. In addition, using visual techniques, we can state that the clustering results were well-separated and compact clusters.
References

- Tan, P. N., M. Steinbach, and V. Kumar, Introduction to Data Mining. 2006: Addison Wesley.
Index Terms

Computer Science Pattern Recognition

Keywords

Fatal accident spatio-temporal data cluster analysis data mining