Ph.D. Thesis Title: OPTIMIZED R+ TREE INDEXING FOR SPATIO TEXTUAL SKYLINE QUERIES

Name: SWATHI SOWMYA BAVIRTHI

Reg. No. 1503PH0620

Department: CSE **Name of the Supervisor:** Dr.K.P.Supreethi

ABSTRACT

In the modern days, usage of portable devices (viz., smart-phones, sensors, smart-wearable devices) is increasing with fast pace, contributing huge amount of spatial data. The spatial data consists of geographical location along with the textually tagged data. Location based applications are dependent on the user's points of interest which are normally added-up with textual descriptions. The insights in retrieving points of interest dependent on both spatial closeness and textual relevance, i.e., searching with local awareness is observed everywhere and has fascinated a lot of research interest. The spatio-textual skyline(STS) querying considers geo-location and evaluates the dominant points according to textual descriptions provided by the user.

The majority of research workspotlight on either textual data query processing or spatial proximity query processing resulting from varied large-sized data and the multifaceted emerging challenges that happen during its exploration. Varied spatial indexing mechanisms for skyline querying are available, supporting either textual or spatial objects. Bitmap index exploits sparsity in both attribute-oriented and tuple-oriented data and suggests enhancements for more effective

query processing. The bitmap indexing is efficient for read-only data or append only data, which is the limitation for the existing indexing technique. To address the issues in spatial mining the techniques like R-Tree, Quad Tree, Z-order curve and variants of these structures are predominant in the context of spatial proximity. To address the issues in searching textual files Inverted file is the popular indexing mechanism apart from signature files and suffix arrays. The STS query focuses at evaluating both spatial dominance and textual relevance and hencethe hybrid indexing technique is vital.

The textual data is pre-processed and a pruned inverted file indexing is proposed. Experiments are conducted on Zomato dataset and Trip Advisor dataset for evaluating the performance of the pruned inverted file indexing. An indexing mechanism for skyline querying is evaluated in this work that supports both textual descriptions and spatial object querying using SSKQR+ tree indexing structure which inserts data quicker than its variant RTree spatial data structure and minimizes evaluation time than the traditional spatial indexing techniques. SkyR+ pruning technique is considered for reducing the comparisons and minimizing search space for the user's query. The proposed indexing structure performs optimized STS query evaluation.