Compressed Spatial Hierarchical Bitmap (cSHB) Indexes for Efficiently Processing Spatial Range Query Workloads

Parth Nagarkar, K. Selçuk Candan, Aneesha Bhat
School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ 85281, USA

Introduction
- Traditional databases are optimized for single query execution.
- In many cloud environments, queries in a query workload are correlated. Traditional databases are unable to leverage shared resources between queries leading to wasted resources.
- Spatial range queries have become an important class of queries due to the popularity of spatial and mobile applications.
- We present our index structure, Compressed Spatial Hierarchical Bitmap (cSHB), that can be used for efficiently executing spatial range query workloads.

Experimental Data Characteristics

<table>
<thead>
<tr>
<th></th>
<th># of Points</th>
<th># of Points per (non-empty) cell (for height = 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum</td>
<td>Average</td>
</tr>
<tr>
<td>Uniform Data (Synthetic)</td>
<td>100M</td>
<td>54</td>
</tr>
<tr>
<td>Clustered Data (Gowalla)</td>
<td>6.4M</td>
<td>1</td>
</tr>
<tr>
<td>Clustered Data (OSM)</td>
<td>688M</td>
<td>1</td>
</tr>
</tbody>
</table>

Hierarchical Bitmap Indexes

Column Domains are often hierarchical in nature. E.g. Biological taxonomies, Geographical data, etc.

CShB System Architecture

Experimental Evaluation

* 0.5%, 1%, and 5% indicate the query range sizes relative to the size of the considered 2D space.

*This work is supported by NSF grant #1116394 “RanKloud: Data Partitioning and Resource Allocation Strategies for Scalable Multimedia and Social Media Analysis” and by NSF grant #1430144 “PFI:BIC Fraud Detection via Visual Analytics: An Infrastructure to Support Complex Financial Patterns (CFP)-based Real-Time Services Delivery”