Range Search on Uncertain Trajectories

LIMING ZHANG¹, YING ZHANG², WENJIE ZHANG¹, XIAOYANG WANG¹, XUEMIN LIN¹

1. UNIVERSITY OF NEW SOUTH WALES, AUSTRALIA
2. UNIVERSITY OF TECHNOLOGY, SYDNEY, AUSTRALIA
Outline

- Problem Definition
- Application
- General Framework
- Statistics Based Filtering
- Partition Based Filtering
- Experiments
- Conclusion
Uncertain trajectory for an object (Markov Chain model)
Uncertain trajectory for an object (Markov Chain model)
Uncertain Trajectory

- Uncertain trajectory for an object (Markov Chain model)
Uncertain Trajectory

- Uncertain trajectory for an object (Markov Chain model)
Uncertain Trajectory

- Uncertain trajectory for an object (Markov Chain model)
Range Search on Uncertain Trajectory

- Range Search

Range Search Query
- Query region $R$
- Time interval $(t_1 \sim t_2)$
- Probability threshold $\theta$
- Duration threshold $\alpha$

If $\theta = 0.4$, $\alpha = 2$, then the object should be returned
Applications

Possible Routines

Ocean Current

Iceberg ($t_1$)

Iceberg ($t_{10}$)

R

Island
General Framework

- Minimum Bounding Box (MBB) of Segment
MBB based pruning and validation

If Q and MBB have no overlapping, then the segment can be pruned.
MBB based pruning and validation

If Q contains the MBB, then the duration of object $d = d + ds$.

If Q overlaps with MBB, then the duration of object $d \leq d + ds$. 
General Framework

- Segments Summaries Tree based Search
• $P(A_i)$ denotes the probability of object that is contained by $Q.R$

• $P(B_i)$ records the probability of object that is not contained by $Q.R$

• If $\min\{P^+(A_1), P^+(A_2)\} < \theta$, time $t$ can be pruned

• If $1 - (P^-(B_1) + P^-(B_2)) \geq \theta$, time $t$ can be validated
Statistics Based Filtering

- $P(A_i)$ denotes the probability of object that is contained by Q.R
- $P(B_i)$ records the probability of object that is not contained by Q.R
- $P^+(A_i), P^-(B_i)$ can be approximated by maintaining the expectation and variance along dimension $i$ using Cantelli’s inequality
- To reduce the summary size, we can maintain the statistic information of a time interval
Partition Based Filtering

\[ P^{-}(o(t), q.R) = \sum_{c \in S(g) \land q.R \text{ contains } c} P(o(t), c) \]

\[ P^{+}(o(t), q.R) = \sum_{c \in S(g) \land q.R \text{ overlaps } c} P(o(t), c) \]
Partition Based Filtering

\[ P^-(o(t), q) = \sum_{c \in S(g) \land q.R \text{ contains } c} b(t, c).p^- \]

\[ P^+(o(t), q) = \sum_{c \in S(g) \land q.R \text{ overlaps } c} b(t, c).p^+ \]
Experiments

- **Algorithms**
  - **UST**: The range search techniques proposed in [1] where sub-diamonds based filtering technique is employed.
  - **STA**: General framework by using statistics based filtering.
  - **GRID**: General framework by using partition based filtering.

- **Parameter Settings**

<table>
<thead>
<tr>
<th>Notation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of trajectories ($N$)</td>
<td>2500, <strong>5000</strong>, 7500, 10000</td>
</tr>
<tr>
<td>segment duration $\Delta_t(q)$</td>
<td>[10, 15], [15, 20], [20, 25], [25, 30]</td>
</tr>
<tr>
<td>probabilistic threshold ($\theta$)</td>
<td>0.1, 0.3, <strong>0.5</strong>, 0.7, 0.9, 1.0</td>
</tr>
<tr>
<td>duration threshold ($\eta$)</td>
<td>1, 4, <strong>6</strong>, 8, 10</td>
</tr>
<tr>
<td>query extent (area of $q.R$)</td>
<td>0.05, <strong>0.1</strong>, 0.15, 0.20, 0.25</td>
</tr>
<tr>
<td>query duration $\Delta_t(q)$</td>
<td><strong>10</strong>, 15, 20, 25</td>
</tr>
</tbody>
</table>

Impact of # Trajectories

(a) Response Time

(b) # Candidate Segments

Impact of # Trajectories
Experiments

Impact of Segment Duration

(a) Response Time

(b) # Candidate Segments
Conclusion

- We formally define the problem of range search on uncertain trajectories.

- We introduce an indexing structure as well as a general framework to support range search on uncertain trajectories.

- We develop effective statistics based and partition based filtering techniques.
Thanks!

Q&A