Moving Together Pattern -- an overview

PRESENTED BY ZHAOYI, UESTC

xiaoFine@live.com
Outline

1. Moving Together
   - Categories
   - Application
   - Earlier Models

2. Gathering Pattern
   - Improvement
   - Implementation

3. Extension Models
   - Urban black holes
Introduction

Part 1

Group Moving Patterns

- Company Patterns
- Aggregation Patterns
- Divergence Patterns
- Leadership Patterns
- Popular Patterns
- Mutant Patterns

So.......where’re Moving Together Patterns?
Relative Motion Patterns

• To identify similar movements in a collection of MOPs (moving point objects)
• REMO analysis
  • A transformation of lifeline data to a REMO matrix featuring motion attributes (i.e. speed, acceleration or motion azimuth)
  • Match of formalized patterns on the matrix
Part 1  Relative Motion Patterns

An example:
Basic Motion:

**Constance**: sequence of equal motion attributes for $r$ consecutive timestamps

**Concurrence**: incident of $n$ MPOs showing the same motion attributes at time $t$

**Trend-setter**: one trend-setting MPO anticipates the motion of $n$ others
Spatial Motion Patterns

Basic Motion + Spatial Constraints (proximity measure)

- The maximal length of the cumulated distances to the mean or median center
- The average length of the Delaunay edges of the group
- MBB (i.e. a ellipse)
- The indication of a maximal border length of the convex hull
Part 1  Relative Motion Patterns

Flock:

Concurrence + Spatial constraints
Leadership:

Trend-setter + Spatial constraints
Aggregation/Disaggregation Motion Patterns

- **Convergence**: Set of $m$ MPOs at interval $i$ with motion azimuth vectors intersecting within a range $R$ of radius $r$
- **Encounter**: Actually meeting within $R$ extrapolating the current motion
- **Divergence**: The opposite of the Convergence
- **Breakup**: The opposite of the Encounter
An example: *Convergence without cluster*
Drawbacks:
• Hard to define an absolute distance between two objects
• Hard to define $r$ (i.e. Lossy-flock problem)
• A single $r$ is unrealistic
 Density-Based Motion Patterns

Allow the capture of trajectories of arbitrary shape

- **Convoy**: Density-Based *Flock*
- **Swarm**: Time-Relaxed *Convoy*
- Moving Cluster: A sequence of spatial cluster
Moving Cluster:
A set of objects that move close to each other for a time duration
Part 1  Density-Based Motion Patterns

**Flock:**
- A disc of rigid size
- K consecutive timestamps

**Convoy:**
- Dense-based clustering

**Swarm:**
- K (non-consecutive) timestamps
Part 1
Density-Based Motion Patterns

Dense Area Detection: Drawbacks
Part 2 Gathering Patterns

- Key Attributes
- Definitions
- How does it work
• Scale: A gathering typically involves a relatively large number of individuals
• Density: Those individuals forms a dense group
• Durability: It should last for a certain time period continuously
• Stationariness: The geometric properties of the group is relatively stable
• Commitment: At any time of the gathering, there exist several dedicated members who stick to the group for a certain time (possibly non-consecutive)
Part 2 Definitions

• The trajectory of a moving object
  \[ o = \langle (p_1, t_1), (p_2, t_2), \ldots, (p_n, t_n) \rangle \]
  where \( p_i \in \mathbb{R}^2 \) is the geo – spatial position sample at \( t_i \in \mathcal{T}_{DB} \)

• Directly density-reachable
  A point \( p \) is directly density reachable from a point \( q \) w.r.t a given distance threshold \( \epsilon \) and a integer \( m \), if
  \[ p \in N_\epsilon (q) \text{ and } |N_\epsilon (q)| \geq m \]
  where \( N_\epsilon (p) = \{ q \in S | D(p, q) < \epsilon \} \)
• Snapshot cluster

The snapshot cluster $c_t$ is
- a non-empty subset of objects $\mathcal{O} \in \mathcal{O}_{DB}$
- $\forall o_p, o_q \in \mathcal{O}, o_p(t)$ is density-connected to $o_q(t)$
- $\mathcal{O}$ is maximal
• Crowd

A crowd $C_r$ is

- A sequence of snapshot cluster at consecutive timestamps
- The lifetime of $C_r$ is no less than $k_c$
- There should be at least $m_c$ objects at any time
- The distance between any consecutive pair of clusters is not greater than $\delta$
• Gathering

A crowd $C_r$ is called a gathering iff there exists at least $m_p$ participators in each snapshot cluster of $C_r$.

• Participator

An object $o$ is called a participator iff it appears in at least $k_p$ snapshot cluster.
### Definitions

$k_p = 2, m_p = 3$

<table>
<thead>
<tr>
<th>object</th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$c_4$</th>
<th>#</th>
<th>object</th>
<th>$c_1$</th>
<th>$c_3$</th>
<th>$c_4$</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>$o_1$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>$o_1$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>$o_2$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>$o_2$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>$o_3$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>$o_3$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>$o_4$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>$o_4$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>$o_5$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>$o_5$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>$o_6$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>$o_6$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td># Par.</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
<td># Par.</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

Data Mining Lab | UESTC
Part 2

How does it work

1. Snapshot cluster
2. Crowd discovery
   - Indexing clusters with R-tree/grid
3. Gathering detection
   - TAD
4. Updating
Crowd Discovery:

- $C_r$ is said to be closed if it has no super-crowd
- Longer gathering can exist in super-crowd if the crowd is not closed
- Computing Hausdorff distance is high-cost!
Indexing cluster with R-tree:

- $d_{\min}(M(c_i), M(c_j)) \leq d_H(c_i, c_j)$
- Index the MBRs of the cluster in $C$ by a R-tree

$O(MN) \rightarrow O(\log_M N)$
Indexing cluster with R-tree: Drawbacks

- R-tree still costs a lot in construction and maintain
- MBRs may not capture the distribution of clusters
Indexing cluster with Grid:

- Partition the space into by a grid
- The side length of each cell equals to $\frac{1}{\sqrt{2}} \delta$
- Maintain a cell list for each cluster and a inverted list for each cell
- **Affect Region**: A cell $g_{ab}$’s AF is the set of cells whose minimum distance with $g_{ab}$ less than $\delta$
Gathering Detection

The downward closure property doesn’t hold anymore

• TAD
• BVS
• Discovering gathering incrementally
Part 2 Gathering Detection

TAD (Test-and-Divide)

- The gathering output by TAD are closed
Part 2  Gathering Detection

TAD(Test-and-Divide)  \[ k_p = k_c = 3, \quad m_p = m_c = 3 \]

<table>
<thead>
<tr>
<th></th>
<th>(c_1)</th>
<th>(c_2)</th>
<th>(c_3)</th>
<th>(c_4)</th>
<th>(c_5)</th>
<th>(c_6)</th>
<th>(c_7)</th>
<th>(c_8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(o_1)</td>
<td>01</td>
<td>01</td>
<td></td>
<td>01</td>
<td>01</td>
<td></td>
<td>02</td>
<td>02</td>
</tr>
<tr>
<td>(o_2)</td>
<td>02</td>
<td>02</td>
<td>02</td>
<td>02</td>
<td></td>
<td>03</td>
<td>03</td>
<td>03</td>
</tr>
<tr>
<td>(o_3)</td>
<td>03</td>
<td>03</td>
<td>03</td>
<td>04</td>
<td>04</td>
<td>04</td>
<td>03</td>
<td>03</td>
</tr>
<tr>
<td>(o_4)</td>
<td>04</td>
<td>04</td>
<td>04</td>
<td>04</td>
<td>04</td>
<td>04</td>
<td>04</td>
<td>04</td>
</tr>
<tr>
<td>(o_5)</td>
<td>05</td>
<td>05</td>
<td>05</td>
<td></td>
<td>06</td>
<td></td>
<td>06</td>
<td></td>
</tr>
</tbody>
</table>
Part 2  How does it work

BVS (Bit Vector Signature)

<table>
<thead>
<tr>
<th>$B(o_1)$</th>
<th>01101100</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B(o_2)$</td>
<td>11110011</td>
</tr>
<tr>
<td>$B(o_3)$</td>
<td>11010111</td>
</tr>
<tr>
<td>$B(o_4)$</td>
<td>10111111</td>
</tr>
<tr>
<td>$B(o_5)$</td>
<td>01110000</td>
</tr>
<tr>
<td>$B(o_6)$</td>
<td>00001100</td>
</tr>
</tbody>
</table>
Part 2  How does it work

TAD & BVS
• Test Step
  • Count the 1 bits in $B(o)$ with bit operation

1) Let $m_1 = 01010101$,
   $$x = (x \& m_1) + ((x \gg 1) \& m_1) = 01011000$$
2) Let $m_2 = 00110011$,
   $$x = (x \& m_2) + ((x \gg 1) \& m_2) = 00100010$$
3) Let $m_4 = 00001111$,
   $$x = (x \& m_4) + ((x \gg 1) \& m_4) = 00000100$$

$m_1$, $m_2$ and $m_3$ are called masks
Part 2  How does it work

TAD & BVS

• Divide Step
  • No need to process BVSs of non-participators
  • Extract clusters by AND operation and *masks for clusters* i.e. 11110000
Part 2

Discovering gathering incrementally

- New database

\[ \mathcal{O}'_{DB} = \mathcal{O}_{DB} \cup \mathcal{O}_{new} \]

- New time domain

\[ \mathcal{T}'_{DB} = \mathcal{T}_{DB} \cup \mathcal{T}_{new} \]
Part 2  How does it work

Discovering gathering incrementally

• Crowd Extension:

Given a closed crowd $C_r = \{c_i, ..., c_j\}$ in $O_{DB}$, if its last cluster is not at the most recent time point of $T_{DB}$, then $C_r$ cannot be extended into $O_{DB}'$.
Discovering gathering incrementally

- Crowd Extension:

<table>
<thead>
<tr>
<th></th>
<th>$t_1$</th>
<th>$t_2$</th>
<th>$t_3$</th>
<th>$t_4$</th>
<th>$t_5$</th>
<th>$t_6$</th>
<th>$t_7$</th>
<th>$t_8$</th>
<th>$t_9$</th>
<th>$t_{10}$</th>
<th>$t_{11}$</th>
<th>$t_{12}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_1$</td>
<td>$c_2$</td>
<td>$c_3$</td>
<td>$c_4$</td>
<td>$c_5$</td>
<td>$c_6$</td>
<td>$c_7$</td>
<td>$c_8$</td>
<td>$c_9$</td>
<td>$c_{10}$</td>
<td>$c_{11}$</td>
<td>$c_{12}$</td>
<td></td>
</tr>
<tr>
<td>$c_1$</td>
<td>$c_2$</td>
<td>$c_3$</td>
<td>$c_4$</td>
<td>$c_5$</td>
<td>$c_6$</td>
<td>$c_7$</td>
<td>$c_8$</td>
<td>$c_9$</td>
<td>$c_{10}$</td>
<td>$c_{11}$</td>
<td>$c_{12}$</td>
<td></td>
</tr>
</tbody>
</table>
Part 2  How does it work

Discovering gathering incrementally

• Gathering Update:
  • \( IC(C_{r_{new}}) \cap C_{r_{old}} \subseteq IC(C_{r_{old}}) \)
    Invalid cluster \( C_{r_{old}} \) can be valid in \( C_{r_{new}} \)
  • Given an invalid cluster \( c_j \in IC(C_{r_{new}}) \) with \( j \leq n + 1 \), then any closed gathering \( G_r \subset \langle c_i, \ldots, c_{j-1} \rangle \) remains closed in \( C_{r_{new}} \)
    Closed gathering remain closed
Part 3  Extension Models

Urban Black Holes: STG(spatial-temporal Graph)
Urban Black Holes: STG(spatial-temporal Graph)
THANK YOU

PRESENTED BY ZHAOYI