Practical Algorithms for Mining Flock Patterns from Trajectories

Xiaoliang GENG¹, Hiroki ARIMURA¹ and Takeaki UNO²
1. Graduate School of Information Science and Technology, Hokkaido University
2. National Institute of Informatics
Trajectory Data

- By the growth of mobile devices and GPS sensors, many kinds and amount of trajectory data have appeared. There are demands for analyzing them.
Research Goal

- To find the places where moving objects that travel together.

Applications
- Optimization of traffic system in civil engineering.
- Recommendation and guidance for sightseeing and shopping.
Related Research

- Flock pattern mining
We define the class of pattern in a trajectory database.

Problem of mining patterns from trajectory database.

Algorithm for mining flock patterns.
Definition

Point and Trajectory Flock pattern
Basic Definition: Point and Trajectory

Point (School)
Time stamp: 11:00 a.m.
\[ x \text{ (Latitude)}: 5.873943\ldots \]
\[ y \text{ (Longitude)}: 3.98732\ldots \]

Point (Home)
Time stamp: 10:00 a.m.
\[ x \text{ (Latitude)}: 5.873943\ldots \]
\[ y \text{ (Longitude)}: 3.98732\ldots \]

Trajectory:
... 
Point (Home), 10:00 a.m.
Point (School), 11:00 a.m.
Point (Friend home), 15:00 p.m.
... 

Point (Friend home)
Time stamp: 15:00 p.m.
\[ x \text{ (Latitude)}: 5.873943\ldots \]
\[ y \text{ (Longitude)}: 3.98732\ldots \]
Trajectories in our study

length

time stamp
Flock Pattern

Collection place: post office; Time: 10:00 a.m.

Finishing place: south door; Time: 15:00 p.m.
Flock Pattern (FP)

Flock Pattern $P = (X, [b, e]), X = \{1, 2, 3\}$
Given parameters $k$ as minimum length, $r$ as maximum width, a FP $P$ is $(r, k)$-flock pattern if it satisfies the following conditions:

- Length of $P \geq k$,
- Width of $P \leq r$. 
A flock pattern $P = (X, [b, e])$ in $S$ is a rightward length-maximal flock pattern (RFP, for short), if there is no other flock pattern $P_0 = (X, [b, e_0])$ in $S$ such that

- $P_0$ has the same ID set $X$ as $P$;
- the right end of $P_0$ is strictly more larger than that of $P$.

RFP($r$, $k$)
Algorithm for Finding Flock Patterns

» Basic Flock Pattern Miner
» Rightward length-maximal Flock Pattern Miner
» Grid-based Flock Pattern Miner
Algorithm: Basic Flock Pattern Miner (FPM) for FP and FPP

0. Input database $S$, min-length $k$, max-width $r$.

1. Enumerate one subset $X$ from given set $S$ in depth-first order.

2. Check subset $X$.

3. Call $FP(X, r, k, m)$ recursively; Output flock pattern.

Finish algorithm.

$S = \{1, 2, 3, 4, 5\}$

Input database $S$, min-length $k$, max-width $r$.

Time Complexity: $O(knT^2)$
Space Complexity: $O(k^2)$
FPM Check for FP and RFP

Consider with mining FP, mining RFP will generate less redundancy!

Check for FP

$\max - width \ r$

Check for RFP

$\text{length } \geq k$

and as long as possible
Grid-based Flock Pattern Miner (GFPM)

- **Purpose:** To reduce the number of the input trajectory data, we use geometric index.
Implementation and Experiments
Implementation Environment

- We implement our algorithm in C++ and compiled by g++ of GNU, version 4.6.3.
- We use a PC with a Intel(R) Xeon(R) CPU, E5-1620, 3.60GHz with 32GB of memory, OS Ubuntu Linux, version 12.04.
We use a set of synthesis trajectory datasets using our data generator implemented in C++. 
- Range of x, y-coordinate [1,10],
- Default length of trajectory 20 points,
- Default time interval 1 sec,
- Default size of DB 500.

Our data set is a collection of random trajectories in which copies of random patterns are implanted.
- Default patterns number is 6,
- Default size of patterns is 3,
- Default length of patterns is 12 points.
Performance

![Graph showing Performance](image-url)
Conclusion

- Background and Research Goal
- Basic Definition
  - Point and trajectories data,
  - Flock Pattern.
- Algorithm for Mining FP and RFP Patterns.
- Implementation and Experiments.
Thank you for listening
ご清聴ありがとうございます
感谢您的倾听