Towards Parallel Detection of Moving Flock Patterns in Large Spatiotemporal Datasets

Andres Calderon

October 24, 2016
1 Background

2 Flock Patterns
   - BFE algorithm
   - Finding the disks (a.k.a. The Problem)

3 Proposal (a.k.a. The Solution)
1 Background

2 Flock Patterns
   - BFE algorithm
   - Finding the disks (a.k.a The Problem)

3 Proposal (a.k.a. The Solution)
Trajectory Datasets

- They know where you are...
  - Smart phones, GPS, RFID, WiFi...

(Jing et al, 2013)
A new set of rich and interesting movement patterns...

(Gudmundsson et al, 2008)
Transportation...

(Zhang et al, 2016)
• Weather...

(Turdukulov et al, 2014)
Applications

- Climate change...
Applications

- Ecology...

(Majka, 2016)
Applications

- Ecology...

(Fink et al, 2014)
Outline

1 Background

2 Flock Patterns
   - BFE algorithm
   - Finding the disks (a.k.a The Problem)

3 Proposal (a.k.a. The Solution)
**Flock Patterns**

**Definition ((μ, ε, δ) – flock)**

Sets of at least $\mu$ objects moving close enough (ε) for at least $\delta$ time intervals (Benkert et al, 2008).

(Vieira et al, 2009)
Outline

1 Background

2 Flock Patterns
   - BFE algorithm
   - Finding the disks (a.k.a. The Problem)

3 Proposal (a.k.a. The Solution)
Two steps algorithm:

1. Find sets of disks at each time interval.
2. Join consecutive time intervals detecting disks with same objects.
BFE algorithm

Two steps algorithm:
1. Find sets of disks at each time interval.
2. Join consecutive time intervals detecting disks with same objects.
Two steps algorithm:

1. Find sets of disks at each time interval.
2. Join consecutive time intervals detecting disks with same objects.
BFE algorithm
BFE algorithm
Flock Patterns

BFE algorithm

October 24, 2016
BFE algorithm
BFE algorithm
BFE algorithm
Outline

1. Background

2. Flock Patterns
   - BFE algorithm
   - Finding the disks (a.k.a. The Problem)

3. Proposal (a.k.a. The Solution)
Finding the disks...

- Finding the set of disks is no trivial.
- Vieira et al. (2009) proposed a polynomial solution \( (O(2n^2)) \).

![Diagram](attachment:figure4.png)
Finding the disks...

\[ n = 10 \quad \varepsilon = 1 \]
Outline

1. Background

2. Flock Patterns
   - BFE algorithm
   - Finding the disks (a.k.a The Problem)

3. Proposal (a.k.a. The Solution)
Can be done in parallel?

Partition
Local Pruning
Global
Can be done in parallel?

Partition
Local
Pruning
Global
Can be done in parallel?

Partition
Local
Pruning
Global
Can be done in parallel?

Partition
Local
Pruning
Global
Can be done in parallel?

Partition
Local
Pruning
Global
Can be done in parallel?

Partition
Local
Pruning
Global
Proposal (a.k.a. The Solution)

Can be done in parallel?

Partition
Local
Pruning
Global
Simba is a distributed in-memory spatial analytics engine based on Apache Spark (Xie et al, 2016).
Figure 4: Two-level indexing strategy in Simba.

(Xie et al, 2016)
Goals...

1. Design a parallel program to find the set of disk in a given time interval.
2. Implement the new version in an In-Memory Distributed System (Simba).
3. Test the implementation using different settings and datasets.
Thank you!!!

Do you have any question?