

# Towards Parallel Detection of Moving Flock Patterns in Large Spatiotemporal Datasets

Andres Calderon

October 21, 2016

# Agenda

1 Background

2 Flock Patterns

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

3 Proposal

# Agenda

## 1 Background

## 2 Flock Patterns

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

## 3 Proposal

# Trajectory Datasets

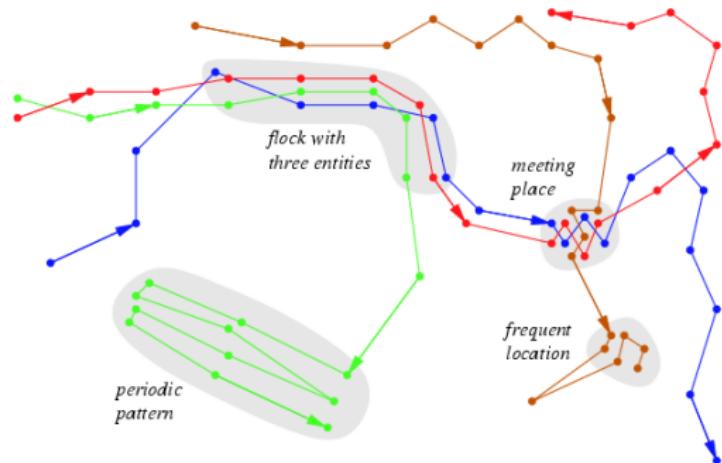
- They know where you are...
  - Smart phones, GPS, RFID, WiFi...



(Jing et al, 2013)

# Movement patterns

- A new set of rich and interesting movement patterns...



(Gudmundsson et al, 2008)

# Applications

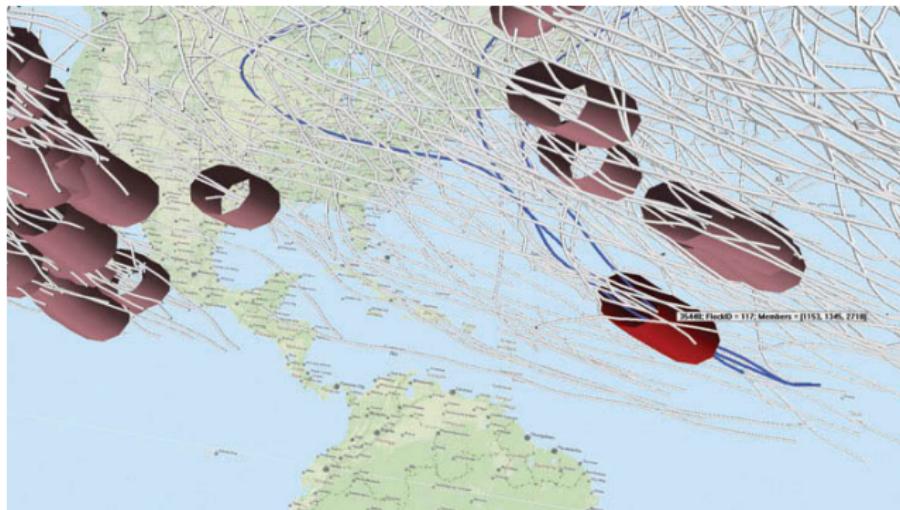
- Transportation...



(Zhang et al, 2016)

# Applications

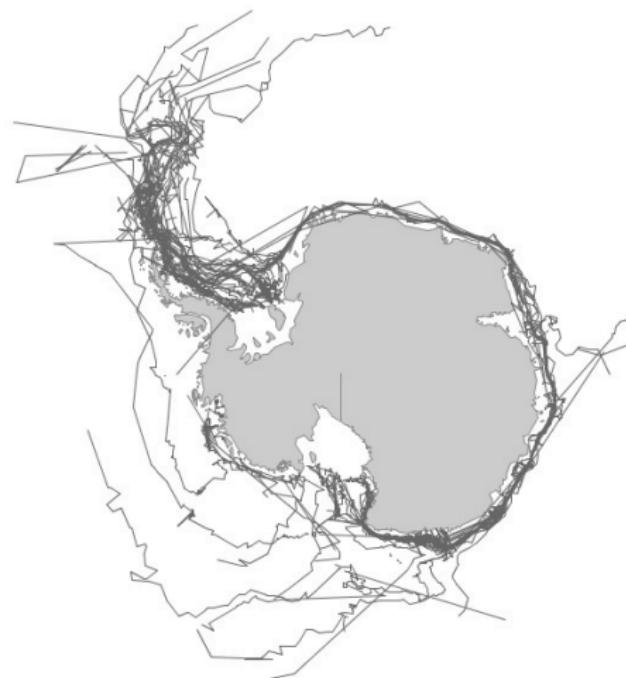
- Weather...



(Turdukulov et al, 2014)

# Applications

- Climate change...



(Calderon, 2011)

# Applications

- Ecology...



(Majka, 2016)

# Applications

- Ecology...

(Fink et al, 2014)

# Agenda

1 Background

2 Flock Patterns

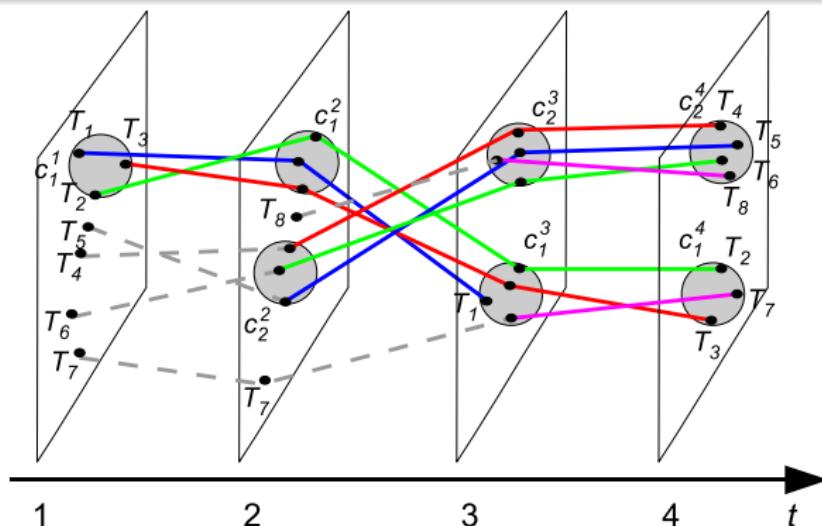
- BFE algorithm
- Finding the disks (a.k.a The Problem)

3 Proposal

# Flock Patterns

Definition  $((\mu, \varepsilon, \delta) - flock)$

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



(Vieira et al, 2009)

# Agenda

## 1 Background

## 2 Flock Patterns

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

## 3 Proposal

# BFE algorithm

- Two steps algorithm:

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

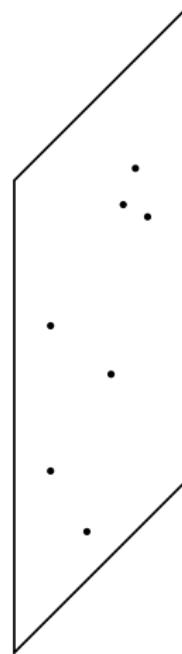
# BFE algorithm

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

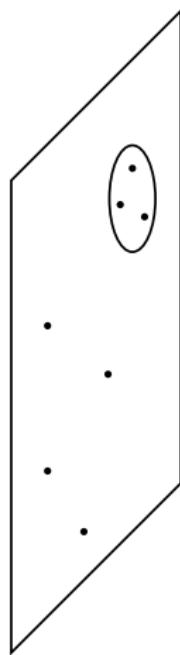
# BFE algorithm

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

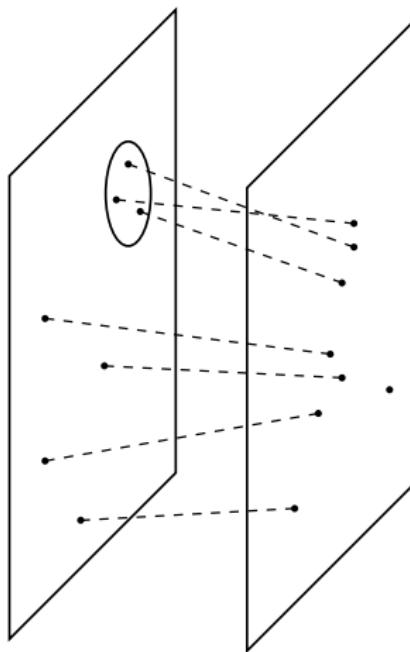
# BFE algorithm



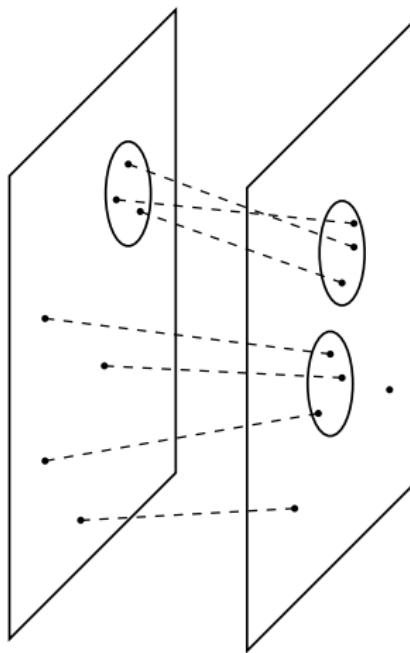
# BFE algorithm



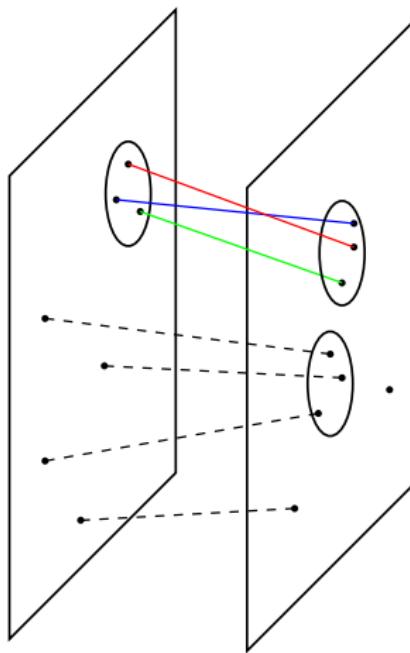
# BFE algorithm



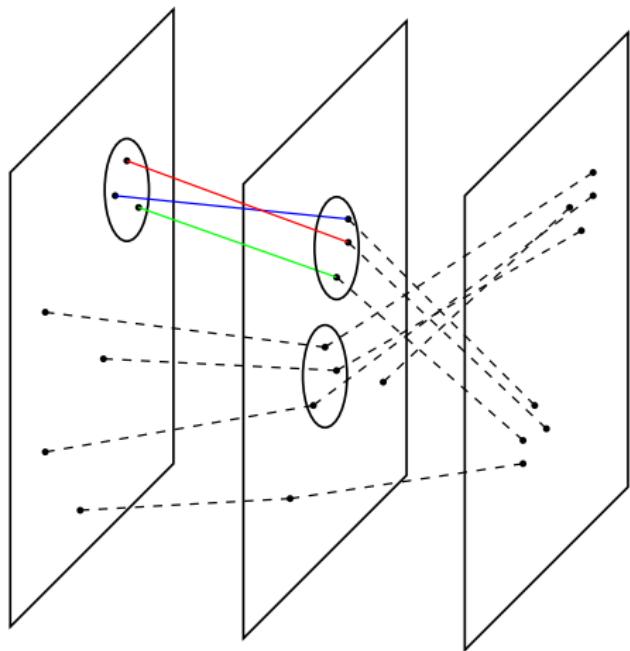
# BFE algorithm



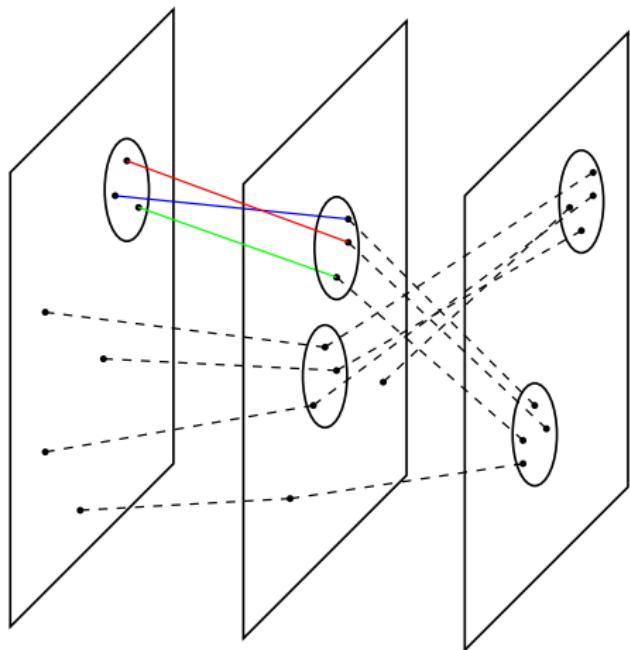
# BFE algorithm



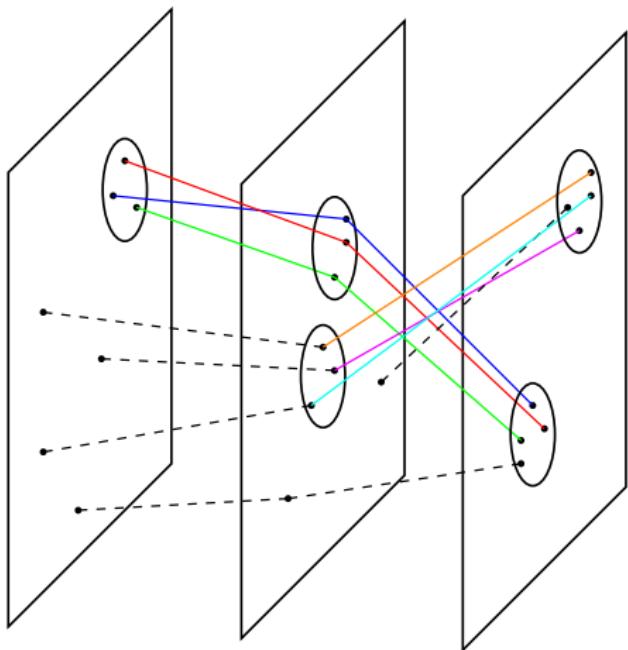
# BFE algorithm



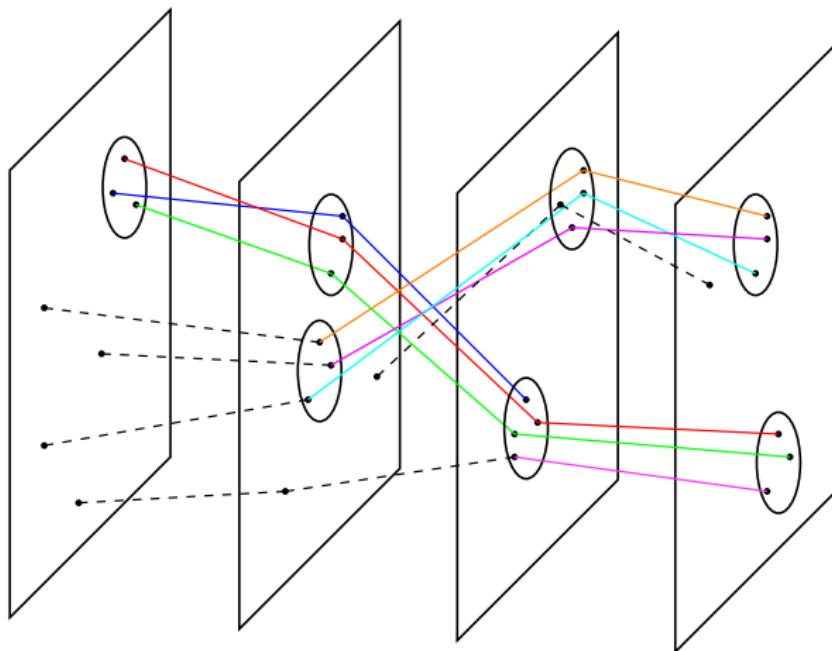
# BFE algorithm



# BFE algorithm



# BFE algorithm



# Agenda

## 1 Background

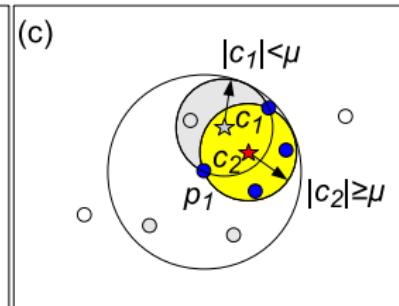
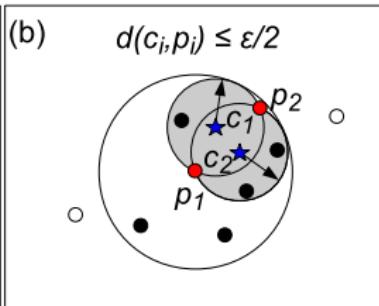
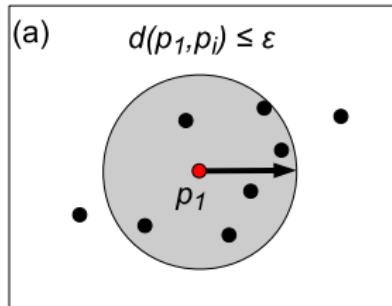
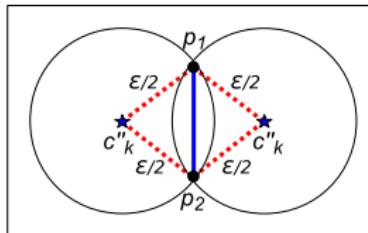
## 2 Flock Patterns

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

## 3 Proposal

# Finding the disks...

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



# Finding the disks...

# Agenda

1 Background

2 Flock Patterns

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

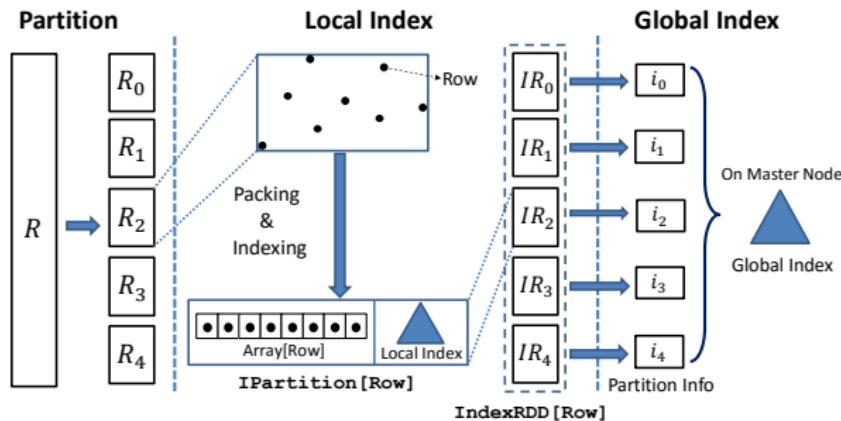
3 Proposal

# Spatial In Memory Big data Analytics



- Simba is a distributed in-memory spatial analytics engine based on Apache Spark (Xie et al, 2016).
  - Extends the SQL and DataFrame query interfaces of Spark SQL.
  - Supports building native (spatial) indexes over RDDs.
  - Implements efficient algorithms for different spatial operators.
  - Introduces spatial and index-aware optimizers to Spark SQL.
  - Open source [<http://www.cs.utah.edu/~dongx/simba/>]

# Spatial In Memory Big data Analytics



**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?