Evolving System for Anomaly Detection

Andressa Stéfany Oliveira Leandro Rochink Costa

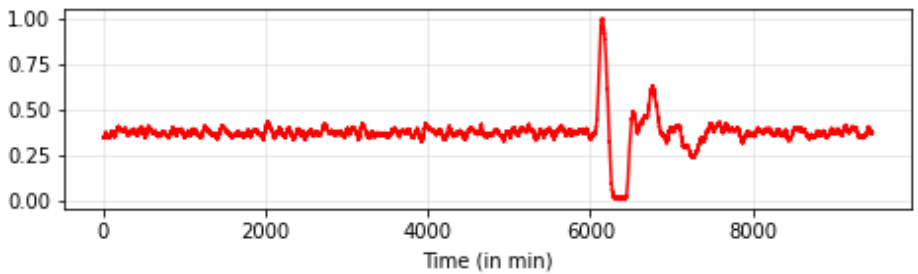
Polytechnique Montréal
DORSAL Laboratory

Outline

Introduction

- Real problems:
  - Detecting fraudulent use of credit cards
  - Data interoperability in IoT environments
  - Detecting and diagnosing faults in industrial processes, like a delayed response from a service

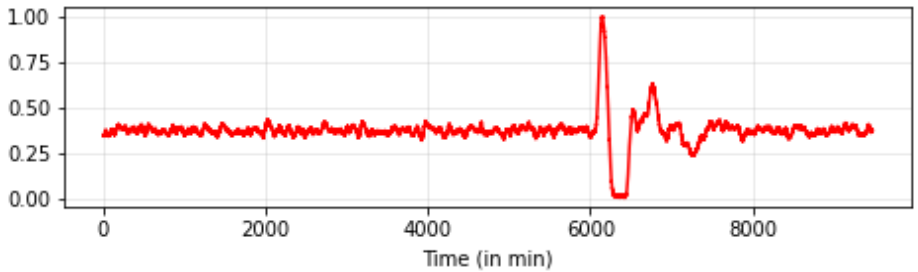
- Data stream
- Anomaly



Introduction

- Real problems:
  - Detecting fraudulent use of credit cards
  - Data interoperability in IoT environments
  - Detecting and diagnosing faults in industrial processes, like a delayed response from a service

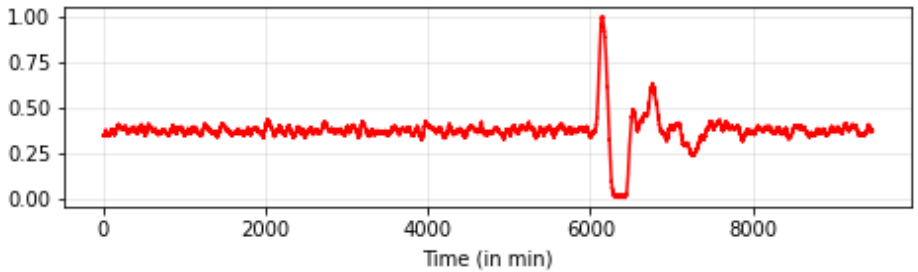
- Data stream
- Anomaly



Introduction

- Real problems:
  - Detecting fraudulent use of credit cards
  - Data interoperability in IoT environments
  - Detecting and diagnosing faults in industrial processes, like a delayed response from a service

- Data stream
- Anomaly



Objective

- Introduce the Macro SOStream algorithm
- Propose framework employing the Macro SOStream

Objective

- Introduce the Macro SOStream algorithm
- Propose framework employing the Macro SOStream

Evolving system

- Concept evolution
- Shift and concept drift
- Single data pass
- The Macro SOStream algorithm
  - Clustering
  - Online and density
  - Microclusters and Macroclusters

Figure: Four types of patterns of changes over time.

Evolving system

- Concept evolution
- Shift and concept drift
- Single data pass
- The Macro SOStream algorithm
  - Clustering
  - Online and density
  - Microclusters and Macroclusters

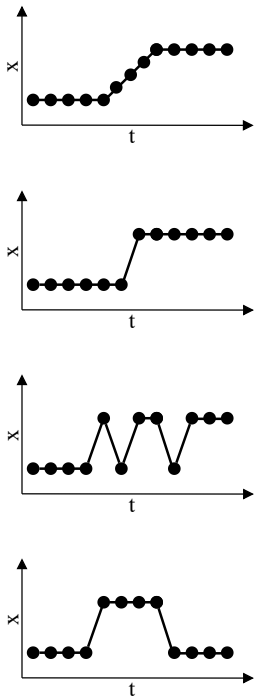


Figure: Four types of patterns of changes over time.

Evolving system

- Concept evolution
- Shift and concept drift
- Single data pass
- The Macro SOStream algorithm
  - Clustering
  - Online and density
  - Microclusters and Macroclusters

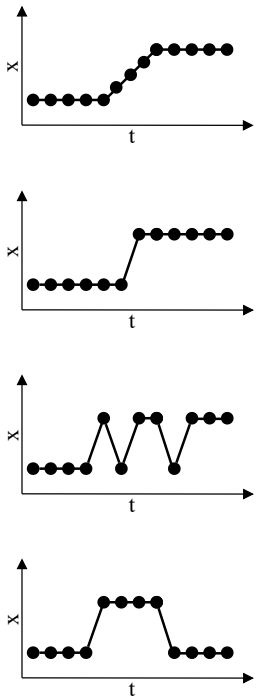


Figure: Four types of patterns of changes over time.

Evolving system

- Concept evolution
- Shift and concept drift
- Single data pass
- The Macro SOStream algorithm
  - Clustering
  - Online and density
  - Microclusters and Macroclusters

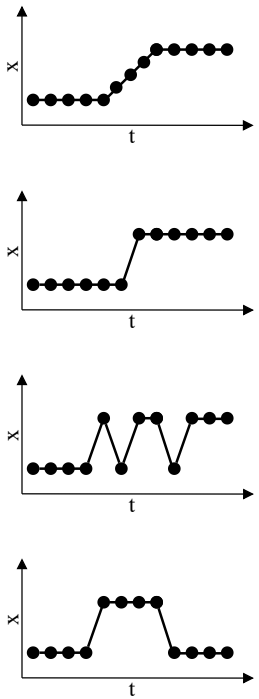
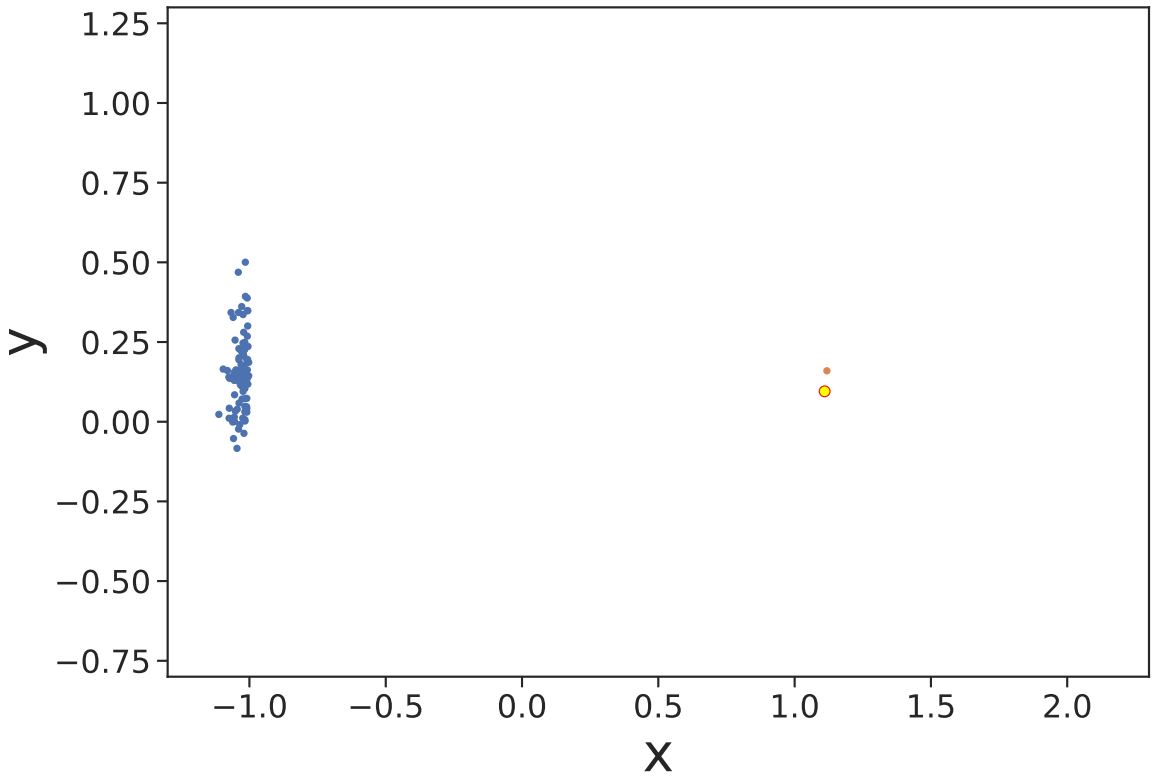
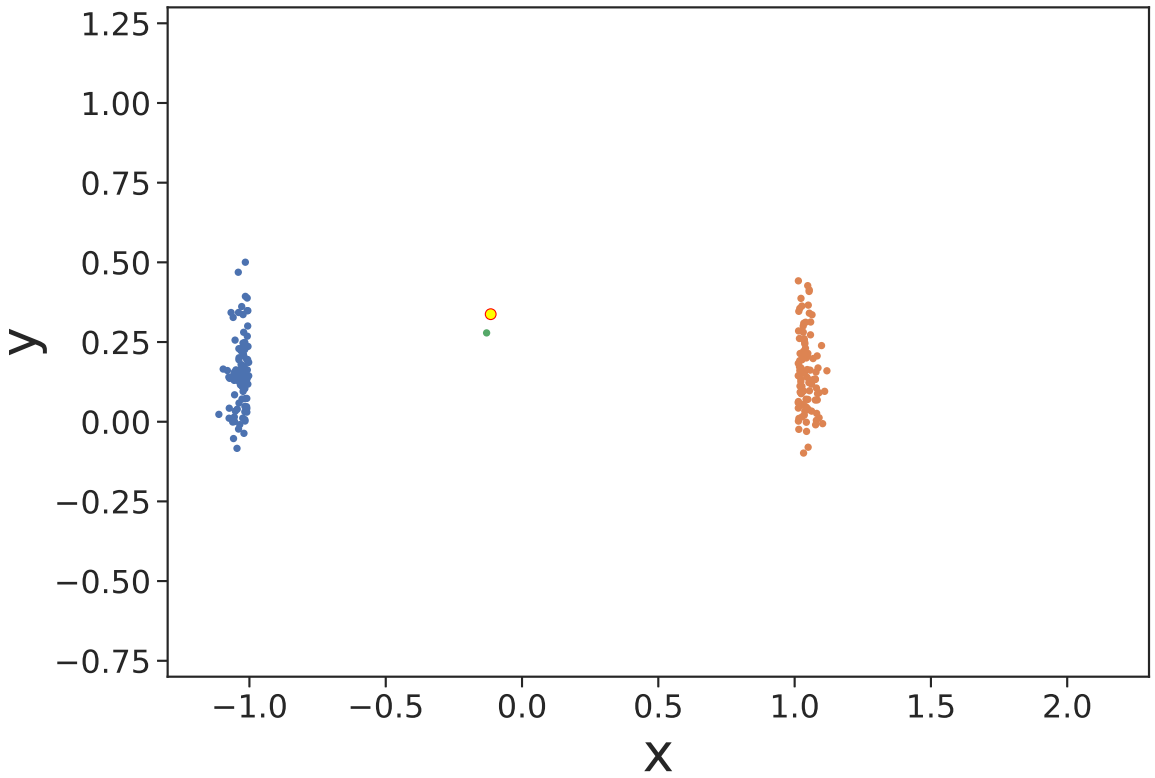


Figure: Four types of patterns of changes over time.

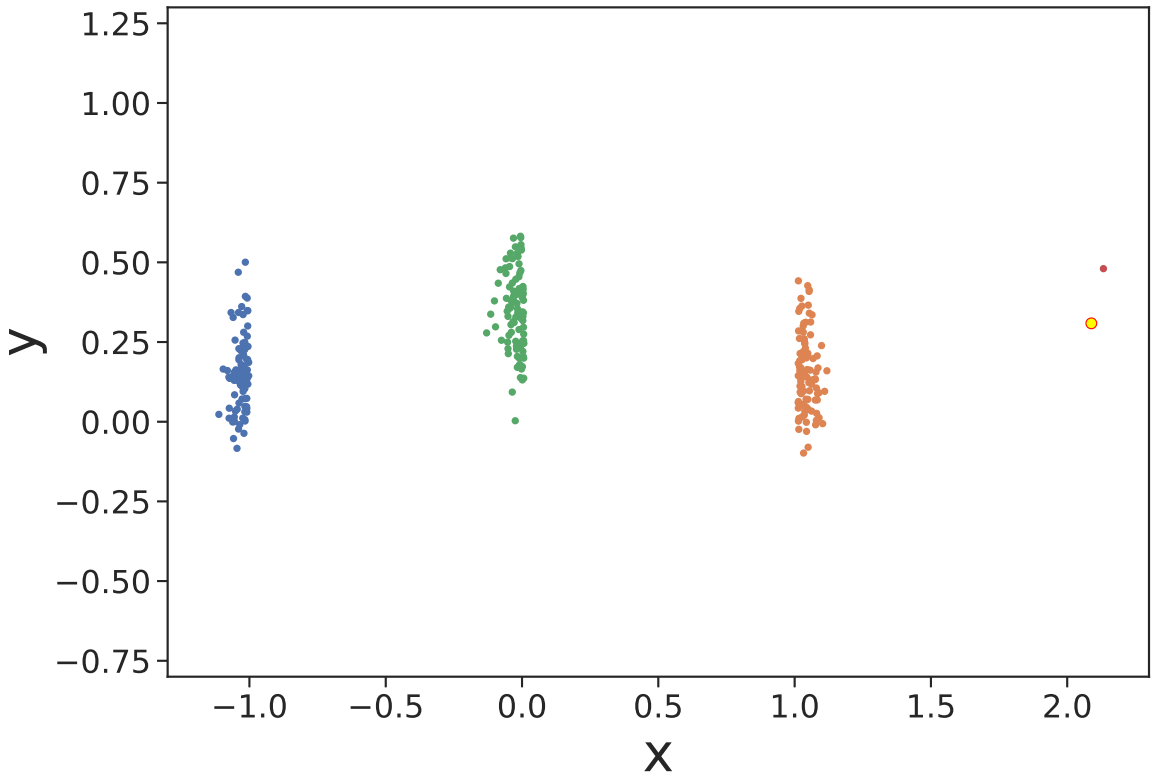
Macro SOStream



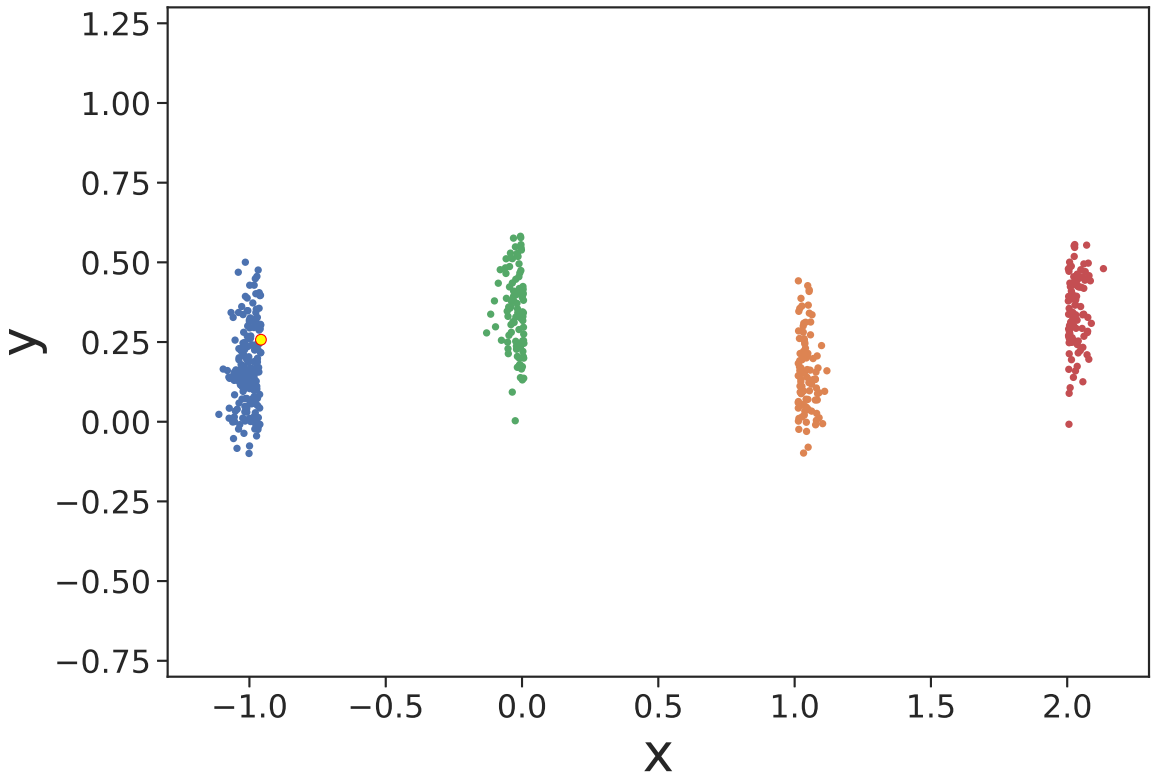
Macro SOStream



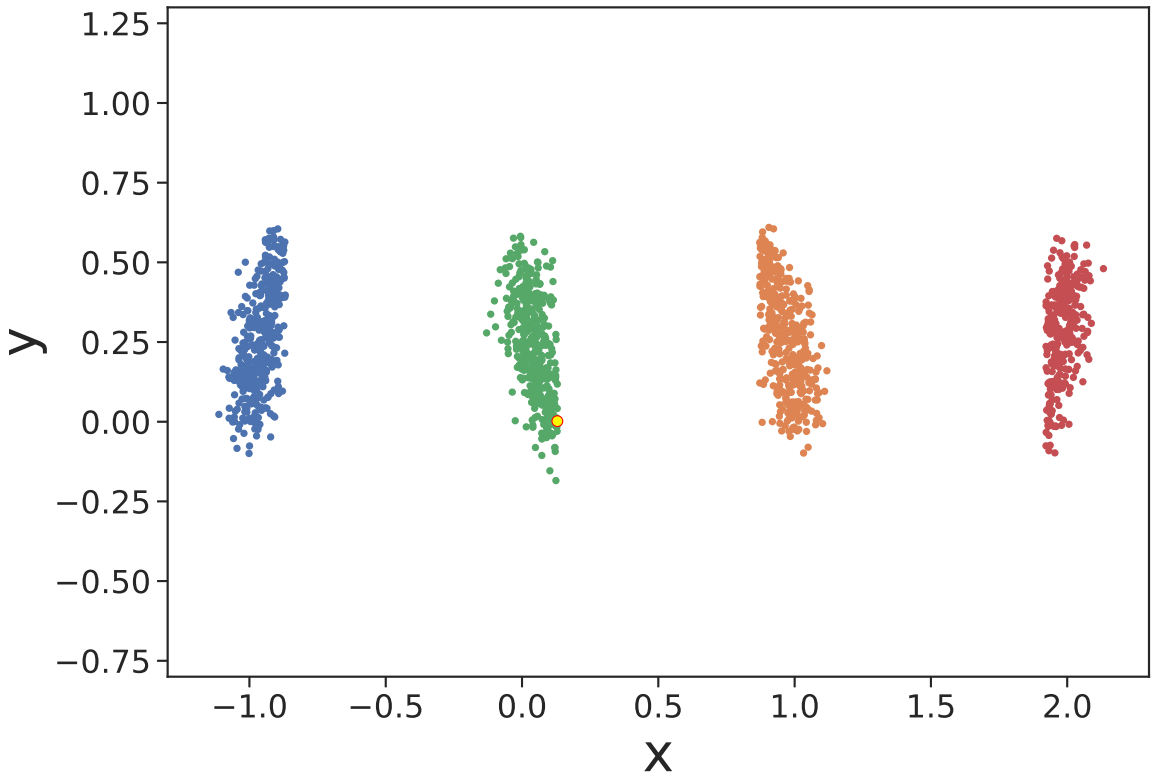
Macro SOStream



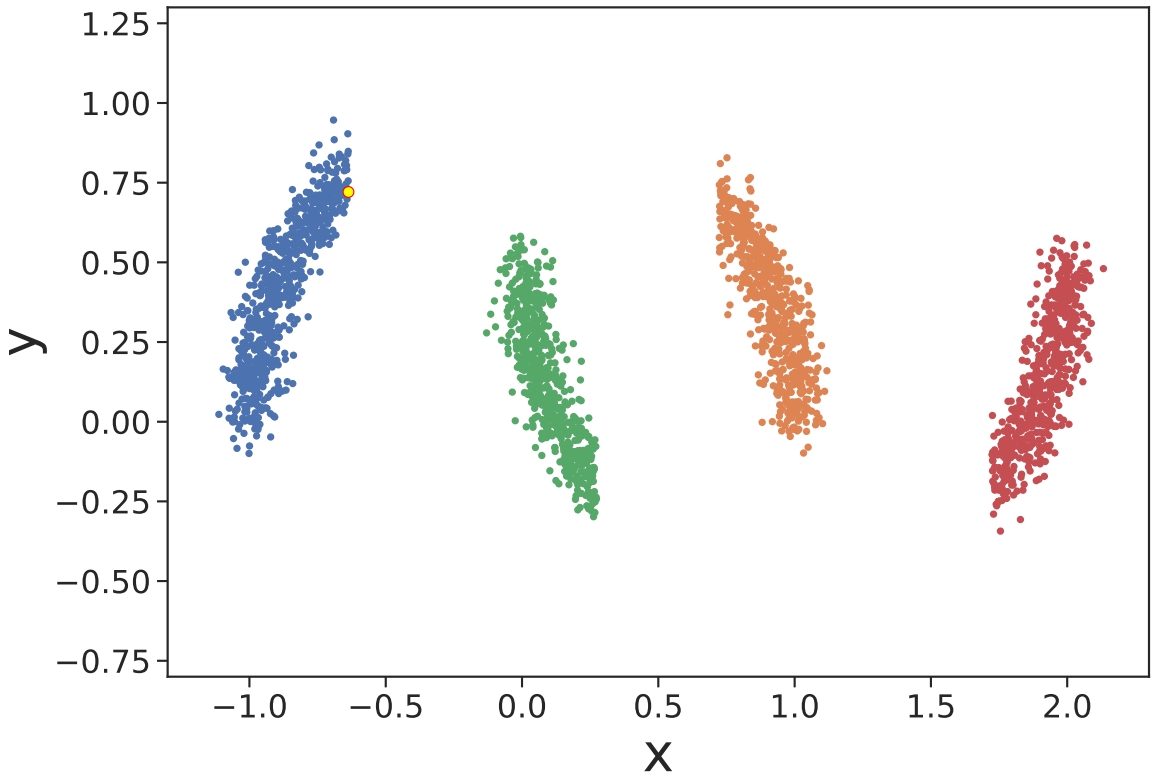
Macro SOStream



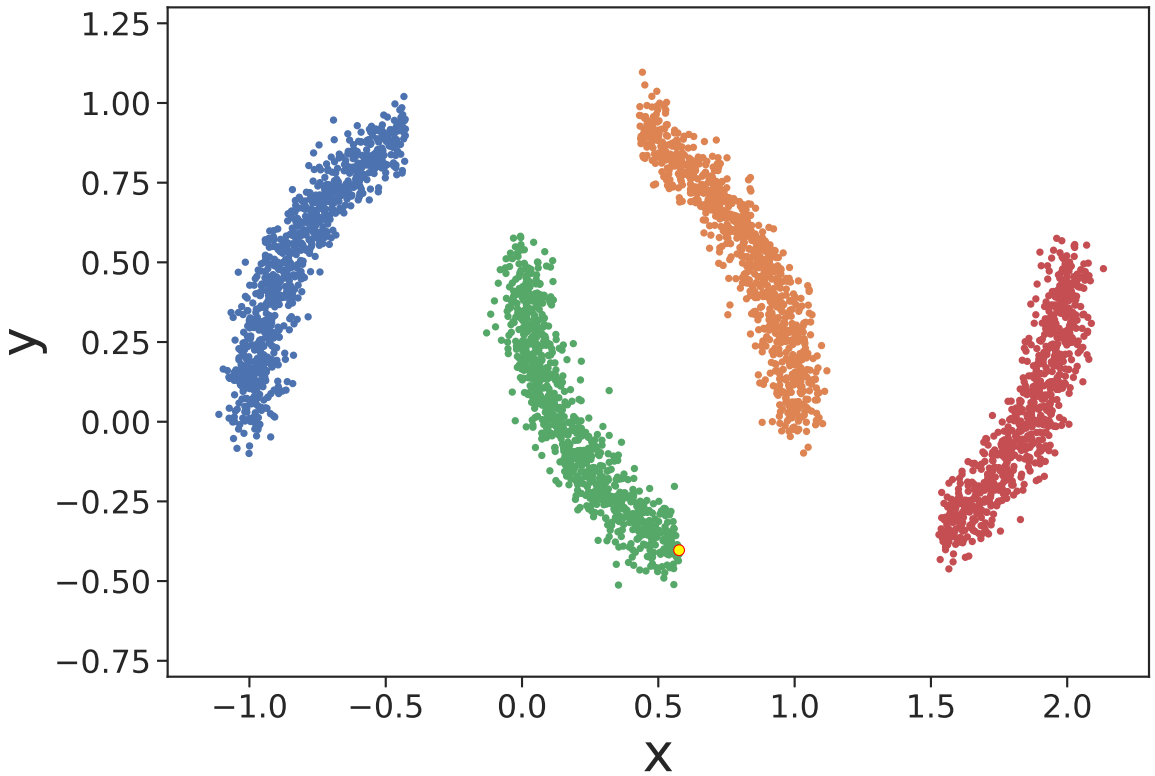
Macro SOStream



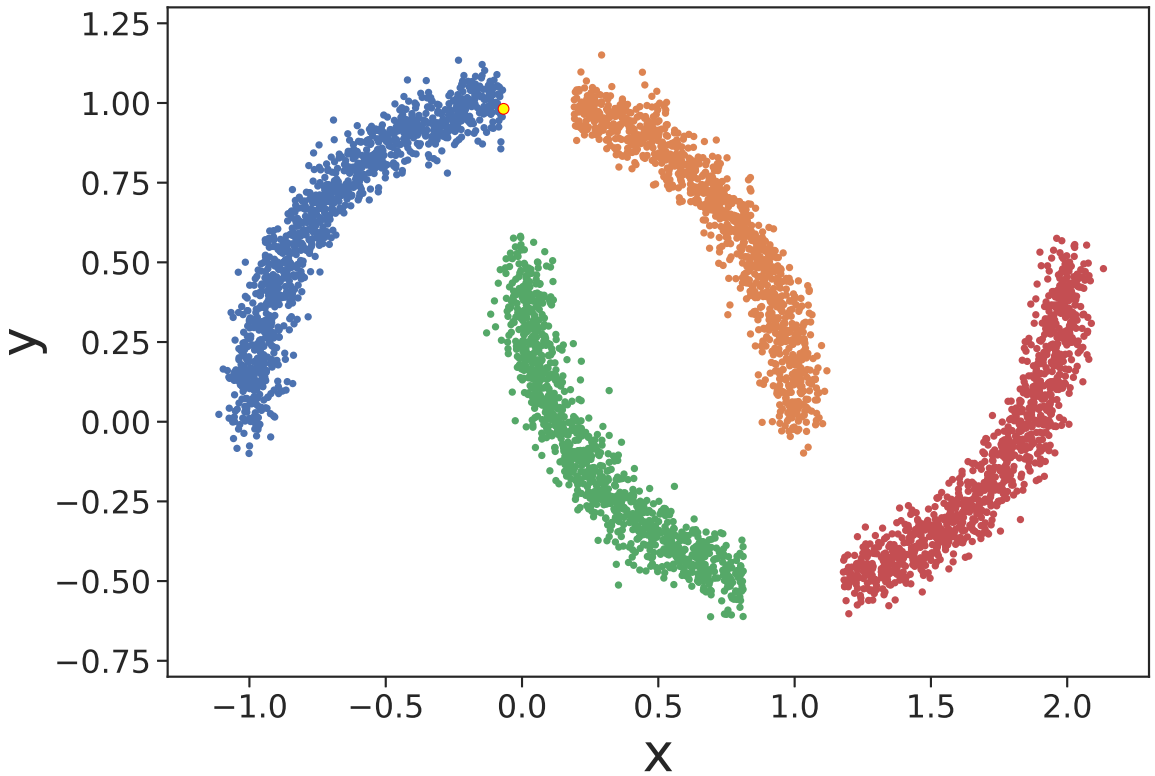
Macro SOStream



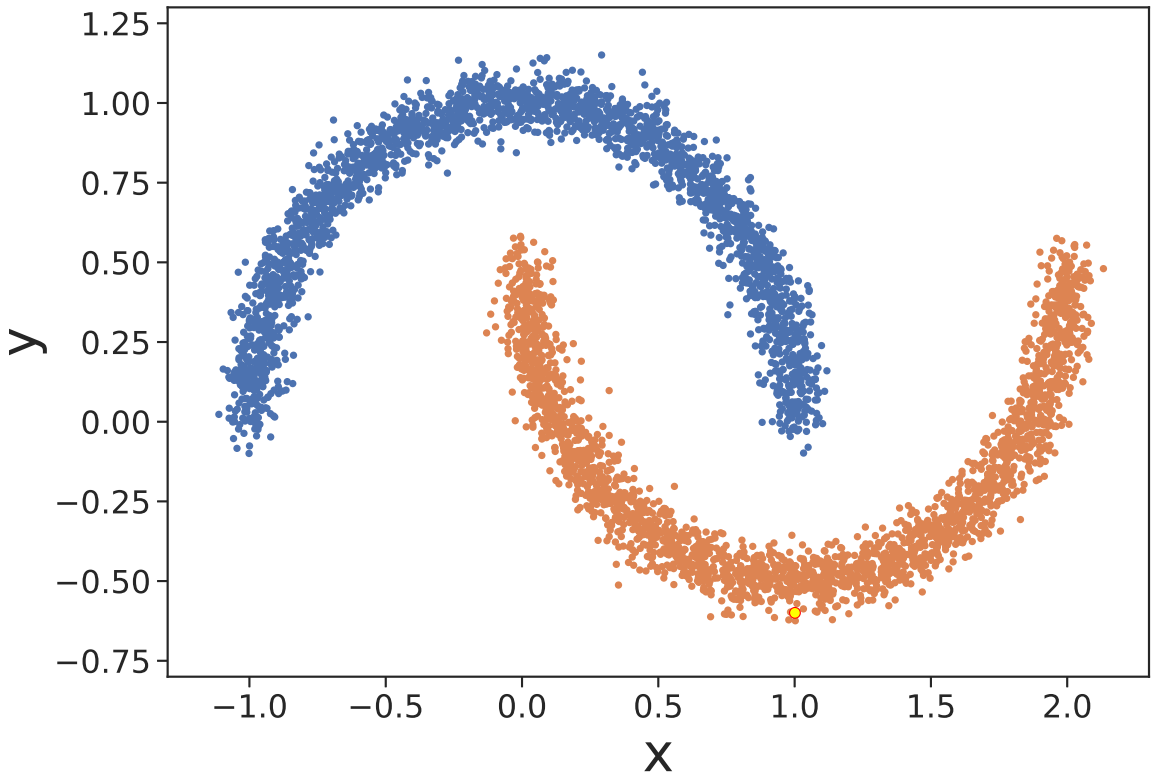
Macro SOStream



Macro SOStream



Macro SOStream



Pipeline

- Leverage Macro SOStream's strengths:
  - Unsupervised learning
  - Single pass online method
  - Robust to distribution shifts
- Design an anomaly detection framework
- Provide intelligible insights to users

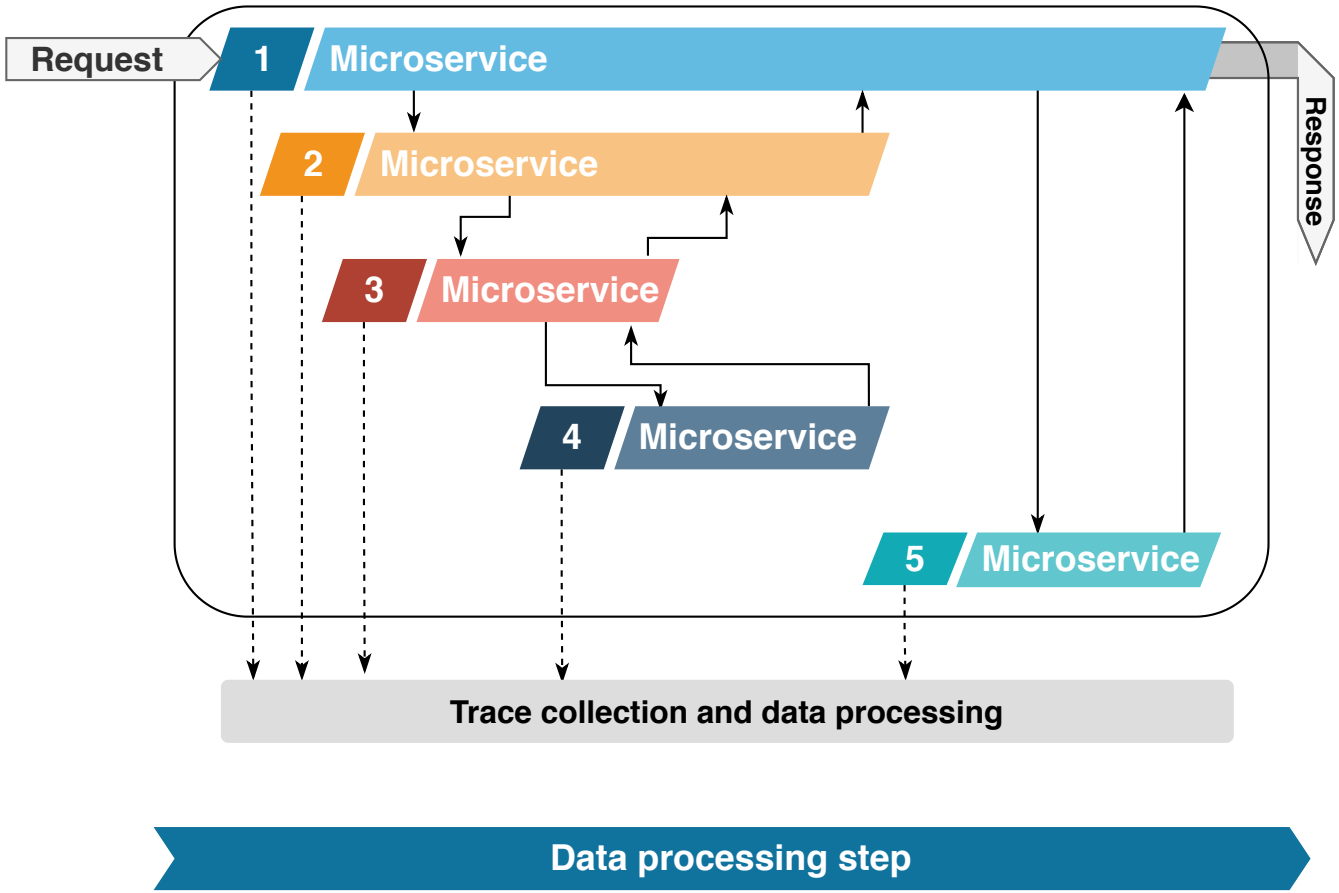
Pipeline

- Leverage Macro SOStream's strengths:
  - Unsupervised learning
  - Single pass online method
  - Robust to distribution shifts
- Design an anomaly detection framework
- Provide intelligible insights to users

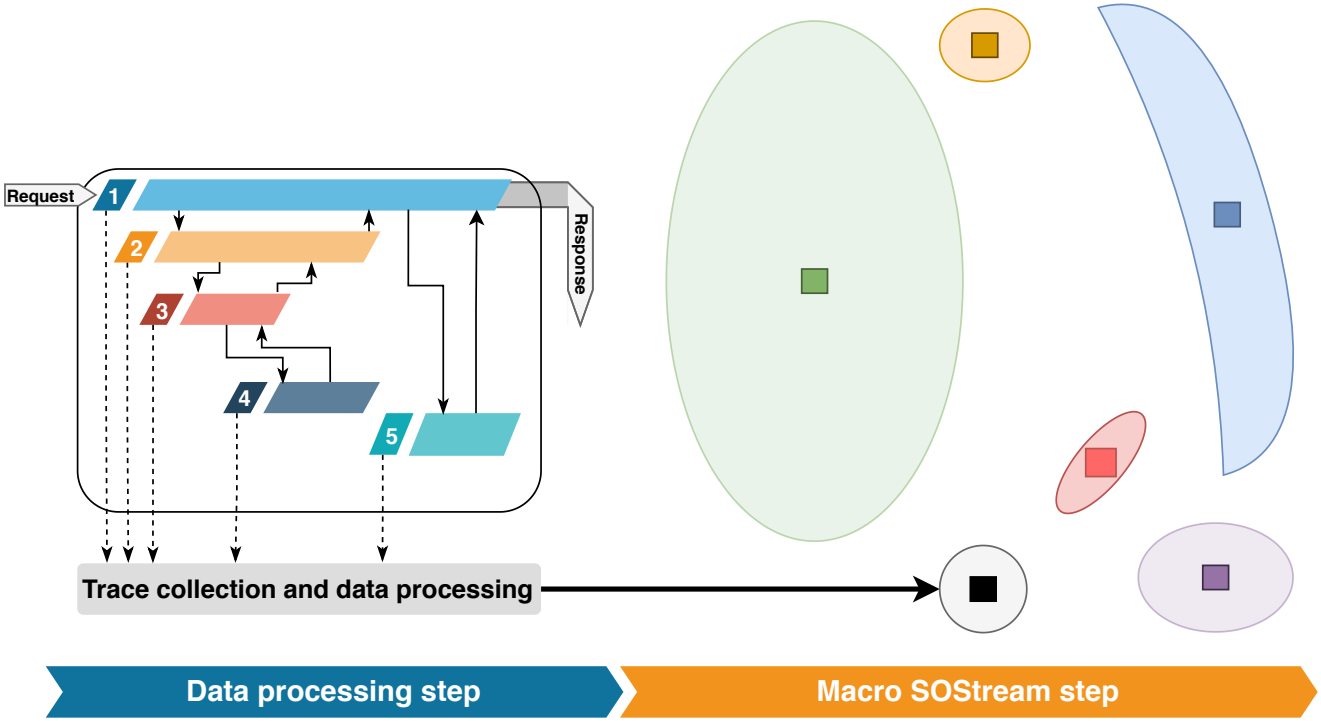
Pipeline

- Leverage Macro SOStream's strengths:
  - Unsupervised learning
  - Single pass online method
  - Robust to distribution shifts
- Design an anomaly detection framework
- Provide intelligible insights to users

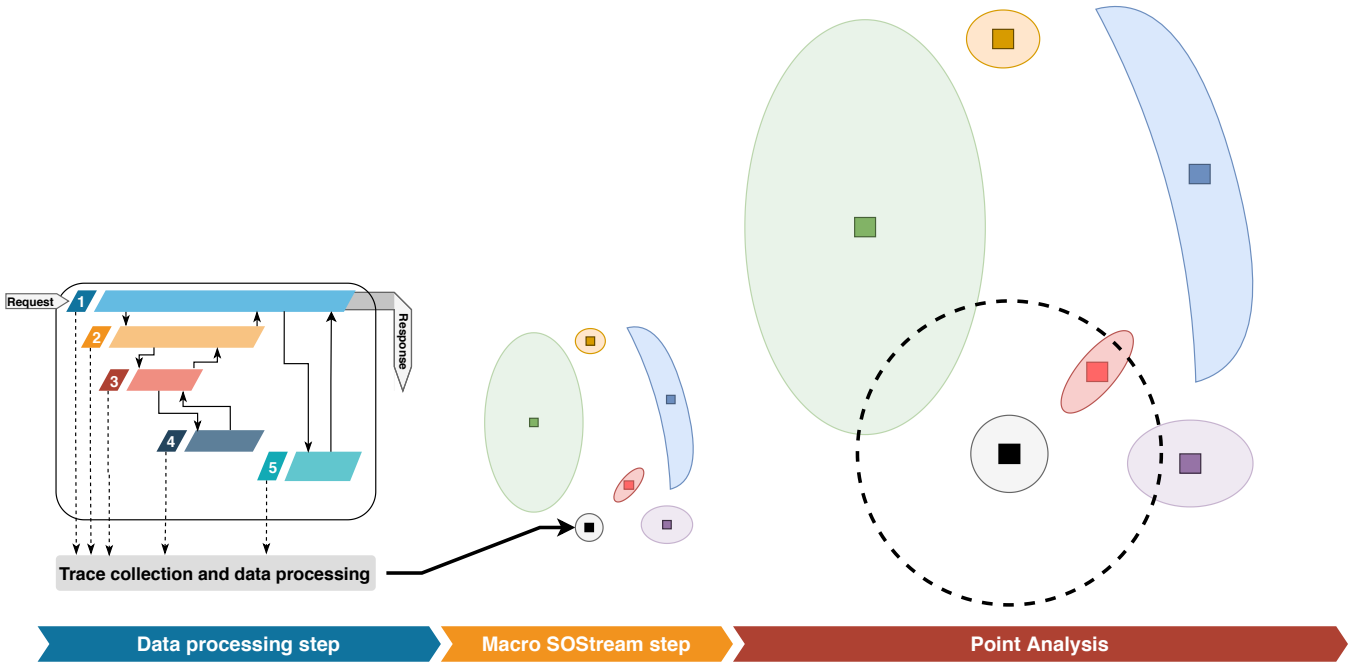
Pipeline


Request
1
Microservice
2
Microservice
3
Microservice
4
Microservice
5
Microservice
Response
Trace collection and data processing
Data processing step

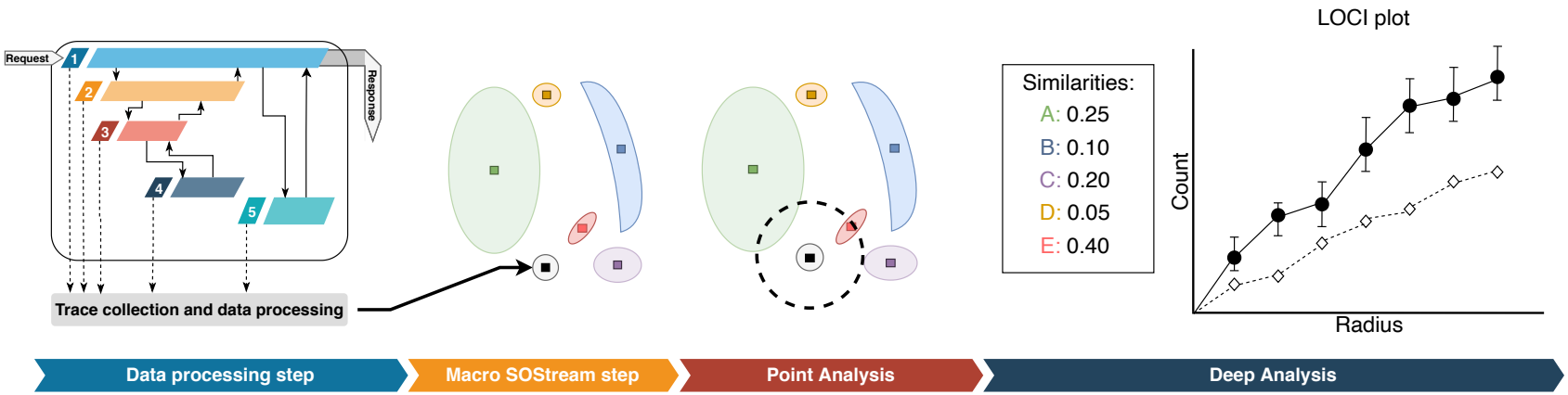
Pipeline


Request
1
2
3
4
5
Response
Trace collection and data processing
Data processing step
Macro SOStream step

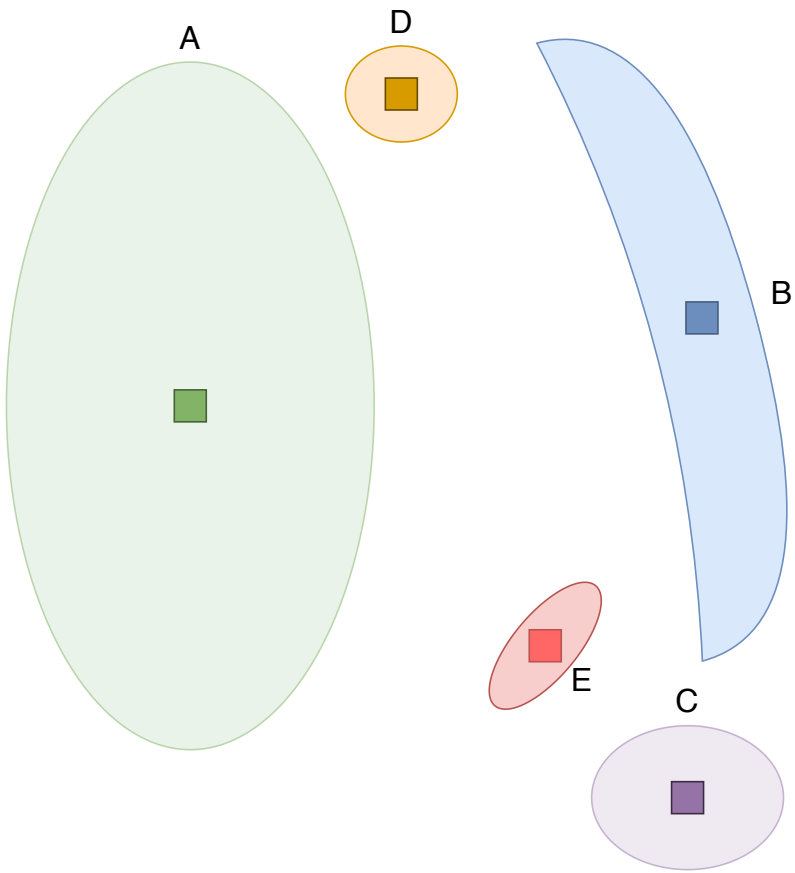
Pipeline


Request
1
2
3
4
5
Response
Trace collection and data processing
Data processing step
Macro SOStream step
Point Analysis

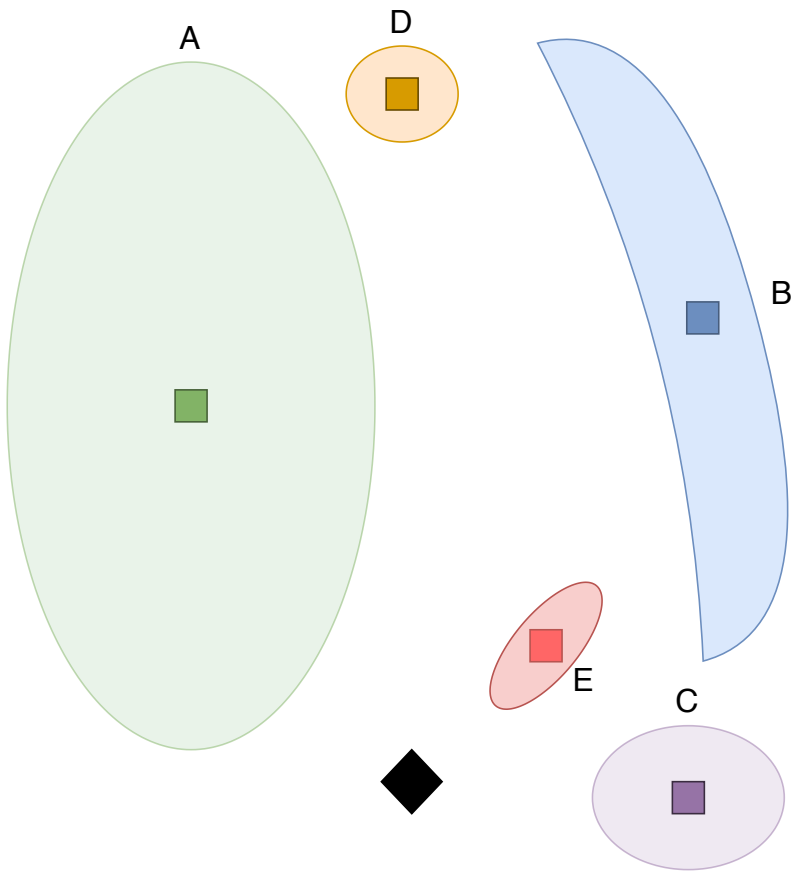
Pipeline


Request
1
2
3
4
5
Response
Trace collection and data processing
Similarities:
A: 0.25
B: 0.10
C: 0.20
D: 0.05
E: 0.40
LOCI plot
Count
Radius
Data processing step
Macro SOStream step
Point Analysis
Deep Analysis

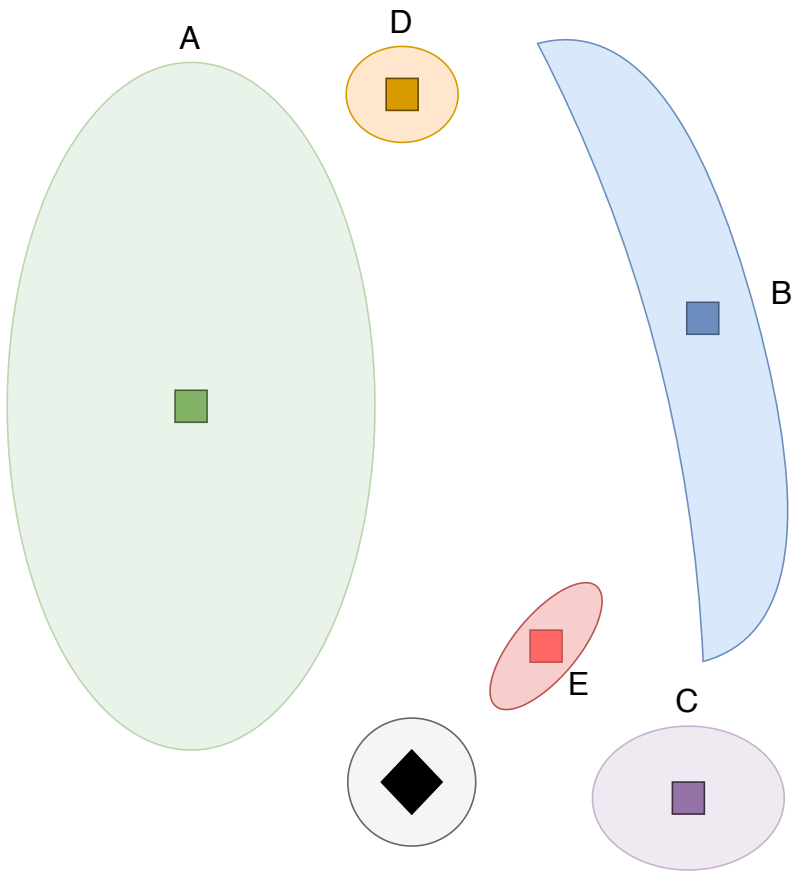
Pipeline



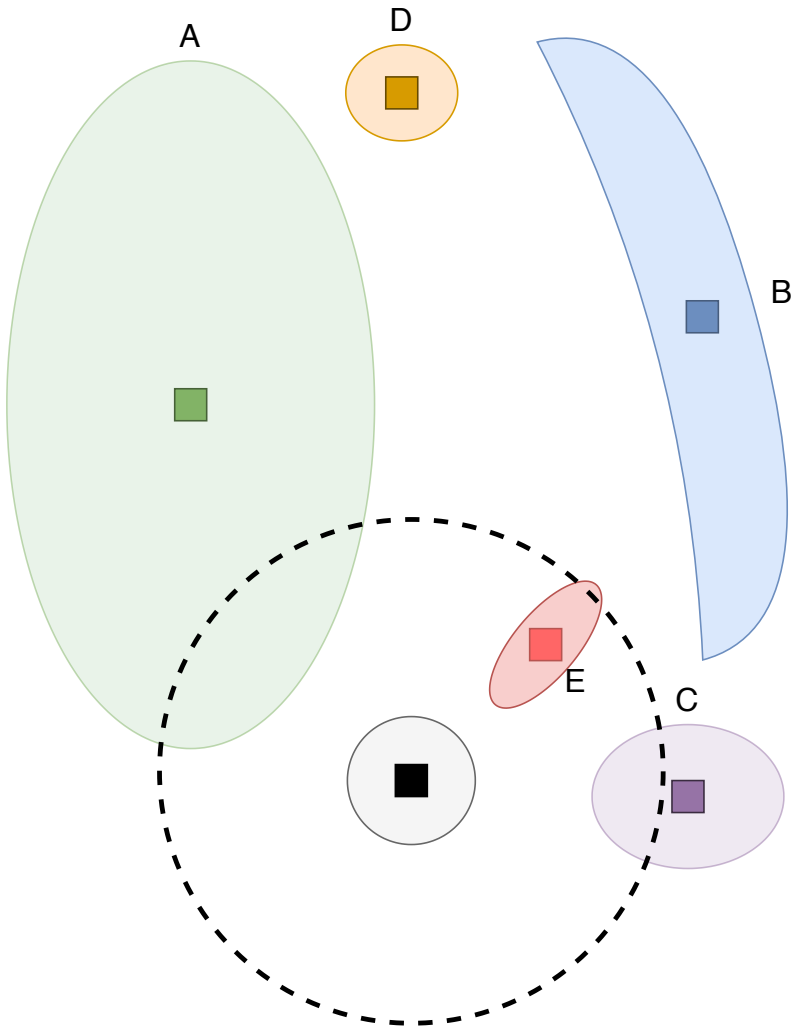
Pipeline



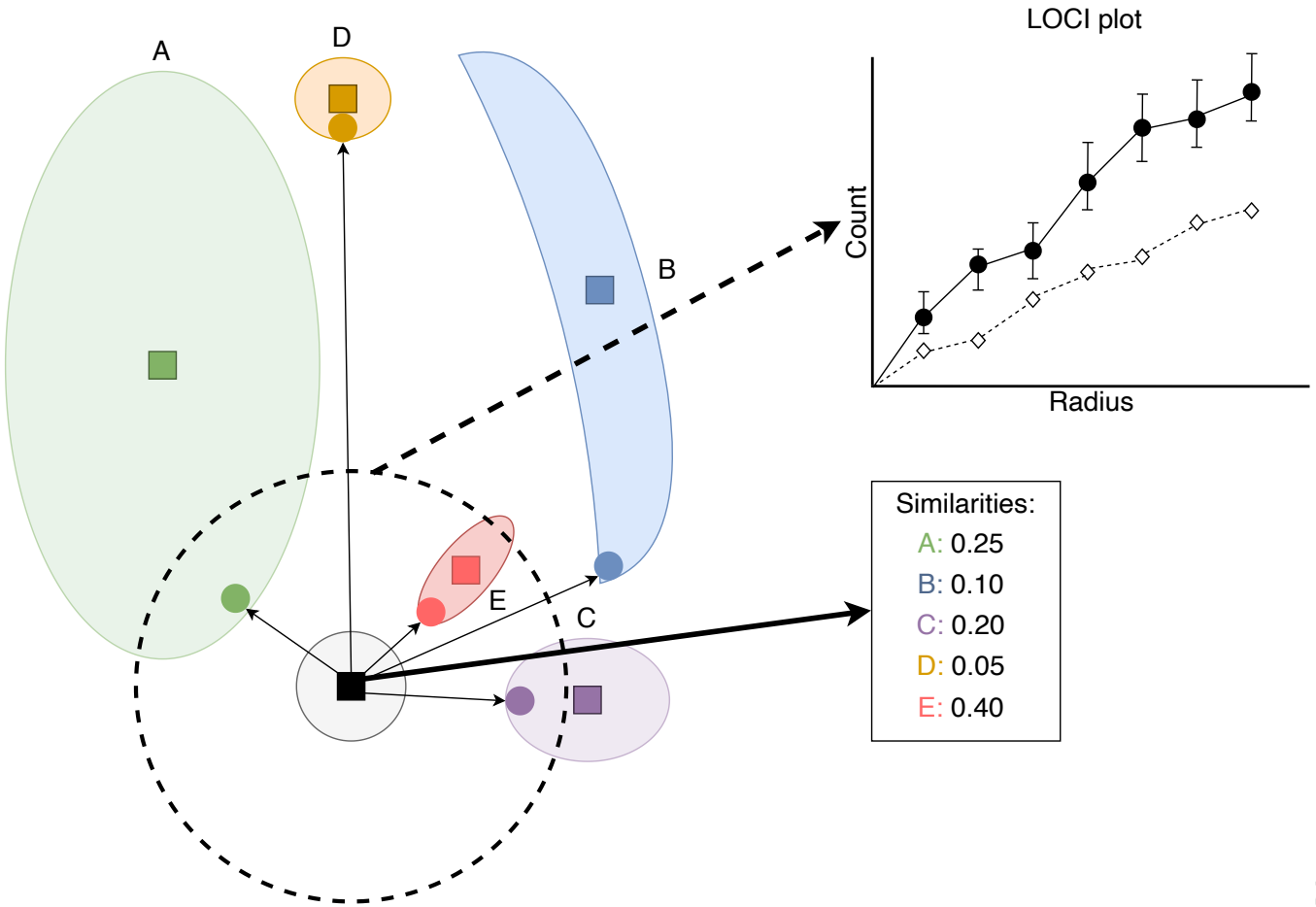
Pipeline



Pipeline



Pipeline


A
D
B
E
C
LOCI plot
Count
Radius
Similarities:
A: 0.25
B: 0.10
C: 0.20
D: 0.05
E: 0.40

Applications

- Experiment with a broad range of applications
- Microservices architecture
  - Performance issues
  - Networking problems
  - Intrusion attacks
  - Benchmark with four use cases
- Real-world use case with kernel traces

Applications

- Experiment with a broad range of applications
- Microservices architecture
  - Performance issues
  - Networking problems
  - Intrusion attacks
  - Benchmark with four use cases
- Real-world use case with kernel traces

Applications

- Experiment with a broad range of applications
- Microservices architecture
  - Performance issues
  - Networking problems
  - Intrusion attacks
  - Benchmark with four use cases
- Real-world use case with kernel traces

Final Remarks

- Macro SOStream
- Intelligible anomaly detection framework
- Current state: Optimizing Macro SOStream
- Future work:
  - Implement analysis steps
  - Tackle the Microservices benchmark
  - Prospect a real-world use case

Final Remarks

- Macro SOStream
- Intelligible anomaly detection framework
- Current state: Optimizing Macro SOStream
- Future work:
  - Implement analysis steps
  - Tackle the Microservices benchmark
  - Prospect a real-world use case

Final Remarks

- Macro SOStream
- Intelligible anomaly detection framework
- Current state: Optimizing Macro SOStream
- Future work:
  - Implement analysis steps
  - Tackle the Microservices benchmark
  - Prospect a real-world use case

Final Remarks

- Macro SOStream
- Intelligible anomaly detection framework
- Current state: Optimizing Macro SOStream
- Future work:
  - Implement analysis steps
  - Tackle the Microservices benchmark
  - Prospect a real-world use case

Thank You!

Andressa Stéfany Oliveira
andressa-stefany.silva-de-oliveira@polymtl.ca
GitHub - Macro SOStream

Leandro Rochink Costa
leandro.costa@polymtl.ca
GitHub - Dynamic VP-tree