Taking Mobile Broadband for a Drive Run: *Coverage Profiling and Analysis*

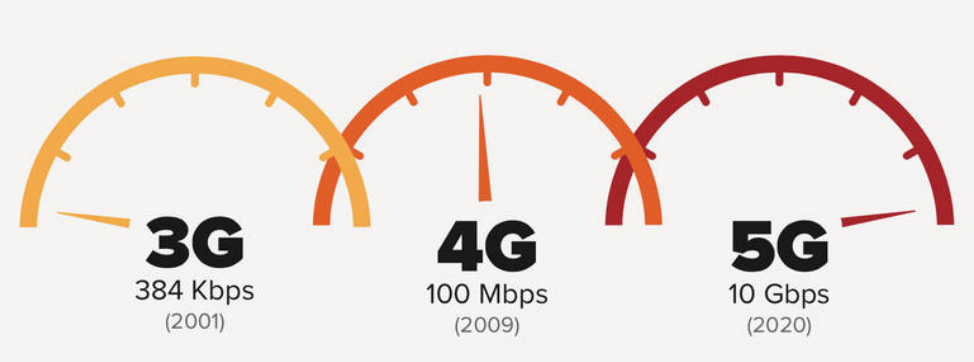
Andra Lutu
Yuba Raj Siwakoti
Özgü Alay
Džiugas Baltrūnas
Ahmed Elmokashfi

Simula Research Laboratory AS

August 28, 2015

Challenges

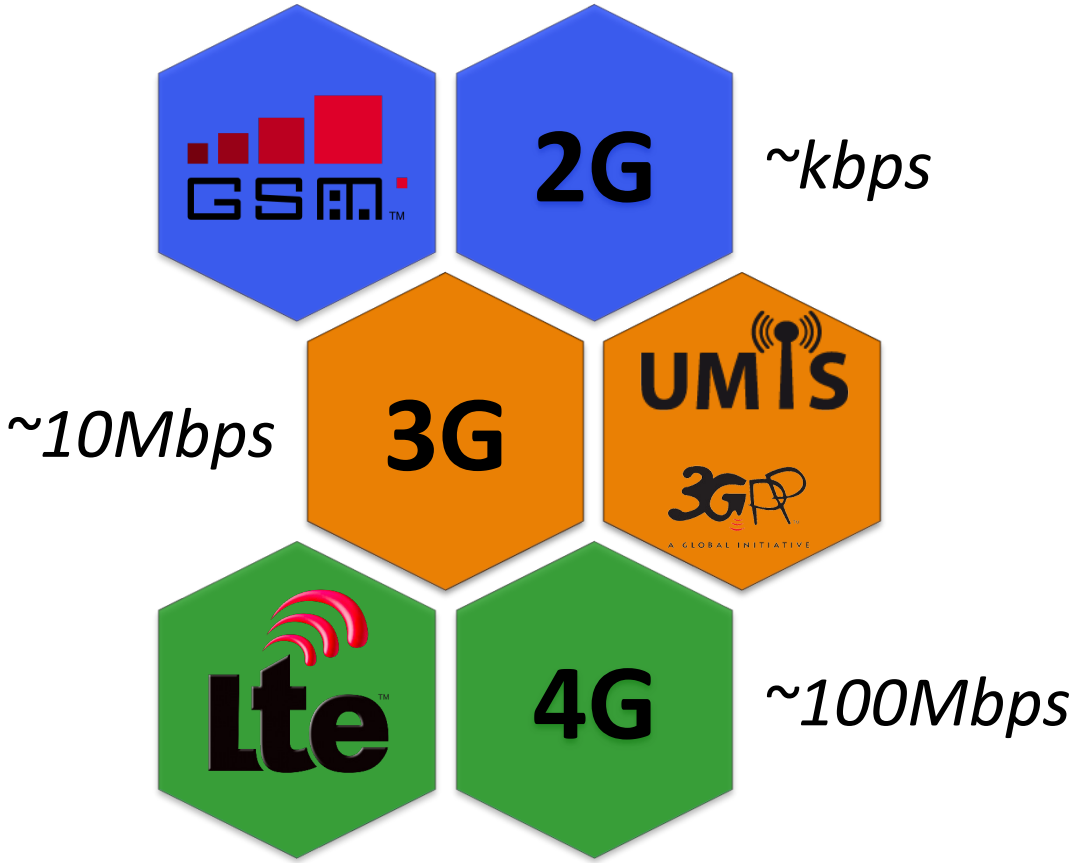
- New use cases, increasing traffic volume and a wide variety of user devices → need for new technologies



- Increasingly complex network infrastructures which incur high operational expenditure
- Need to ensure seamless handovers between existing radio access technologies

Understanding mobile network coverage is no easy task!

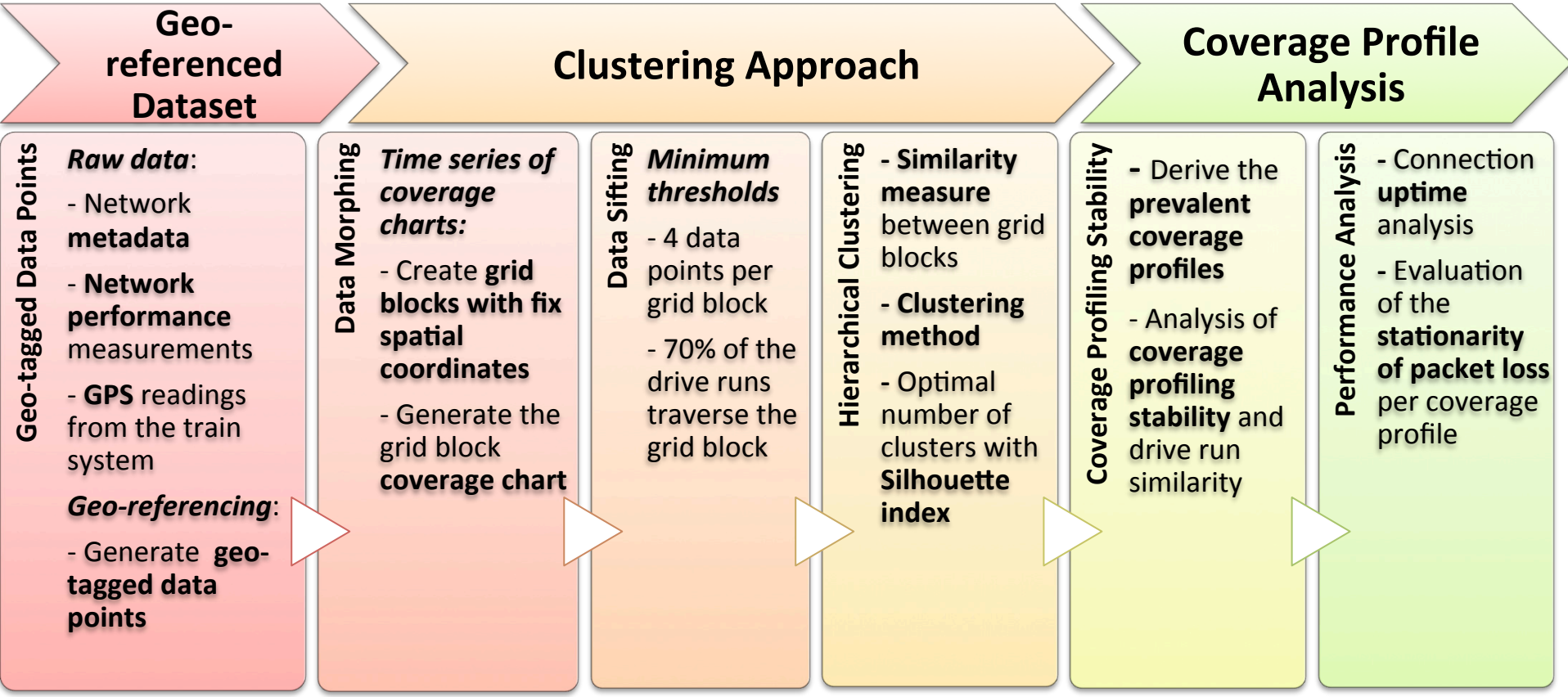
We propose a ***wholistic approach*** to coverage characterization in a certain area, considering the interplay of all available radio access technologies over time



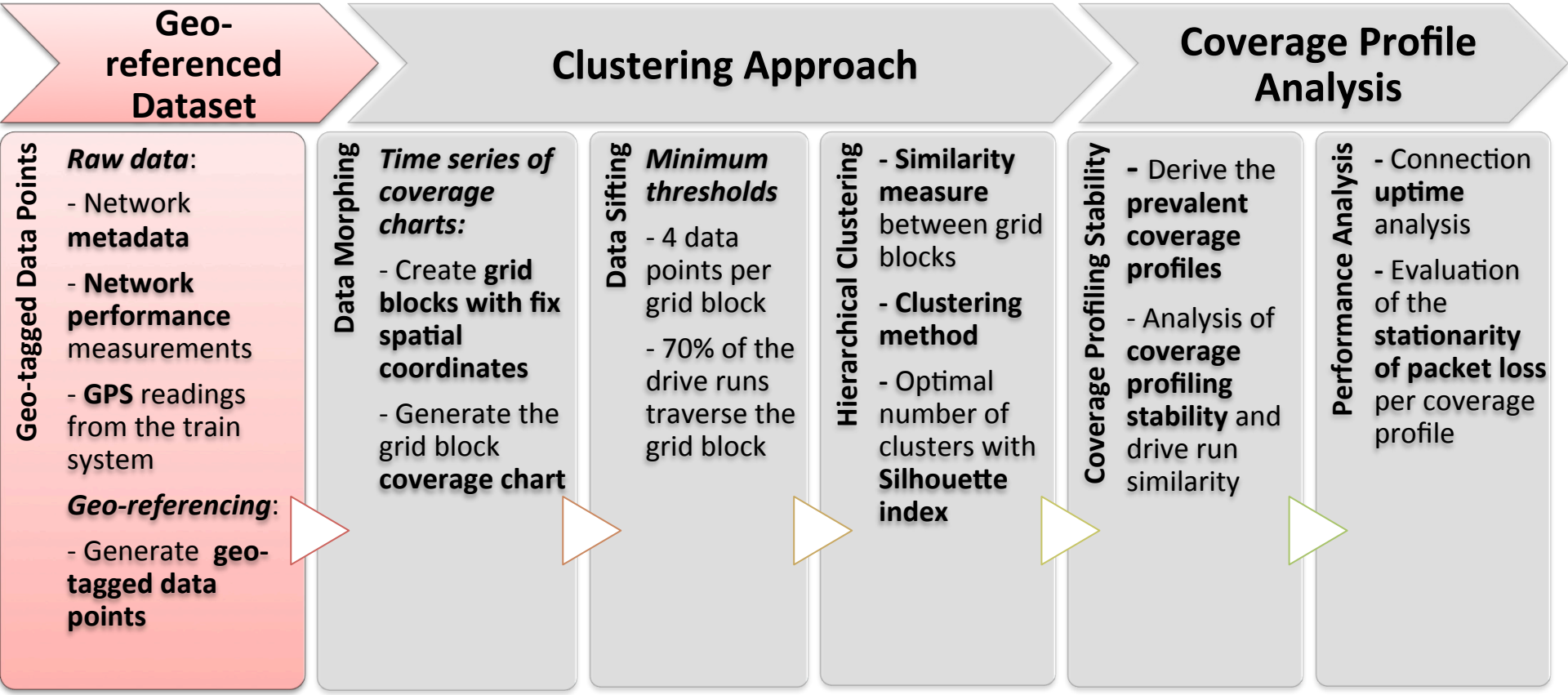
What does this mean?

- Use a vast dataset of periodic coverage measurements performed in certain areas
- Demonstrate clustering approach to draw knowledge in terms of prevalent coverage profiles
- Understand to what extent areas with similar coverage translates into areas with similar performance

Workflow


Geo-referenced Dataset
Clustering Approach
Coverage Profile Analysis
Geo-tagged Data Points
Raw data:
- Network metadata
- Network performance measurements
- GPS readings from the train system
Geo-referencing:
- Generate geo-tagged data points
Data Morphing
Time series of coverage charts:
- Create grid blocks with fix spatial coordinates
- Generate the grid block coverage chart
Data Sifting
Minimum thresholds
- 4 data points per grid block
- 70% of the drive runs traverse the grid block
Hierarchical Clustering
- Similarity measure between grid blocks
- Clustering method
- Optimal number of clusters with Silhouette index
Coverage Profiling Stability
- Derive the prevalent coverage profiles
- Analysis of coverage profiling stability and drive run similarity
Performance Analysis
- Connection uptime analysis
- Evaluation of the stationarity of packet loss per coverage profile

Workflow


Geo-referenced Dataset
Clustering Approach
Coverage Profile Analysis
Geo-tagged Data Points
Raw data:
- Network metadata
- Network performance measurements
- GPS readings from the train system
Geo-referencing:
- Generate geo-tagged data points
Data Morphing
Time series of coverage charts:
- Create grid blocks with fix spatial coordinates
- Generate the grid block coverage chart
Data Sifting
Minimum thresholds
- 4 data points per grid block
- 70% of the drive runs traverse the grid block
Hierarchical Clustering
- Similarity measure between grid blocks
- Clustering method
- Optimal number of clusters with Silhouette index
Coverage Profiling Stability
- Derive the prevalent coverage profiles
- Analysis of coverage profiling stability and drive run similarity
Performance Analysis
- Connection uptime analysis
- Evaluation of the stationarity of packet loss per coverage profile

Common Approaches for Assessing Coverage

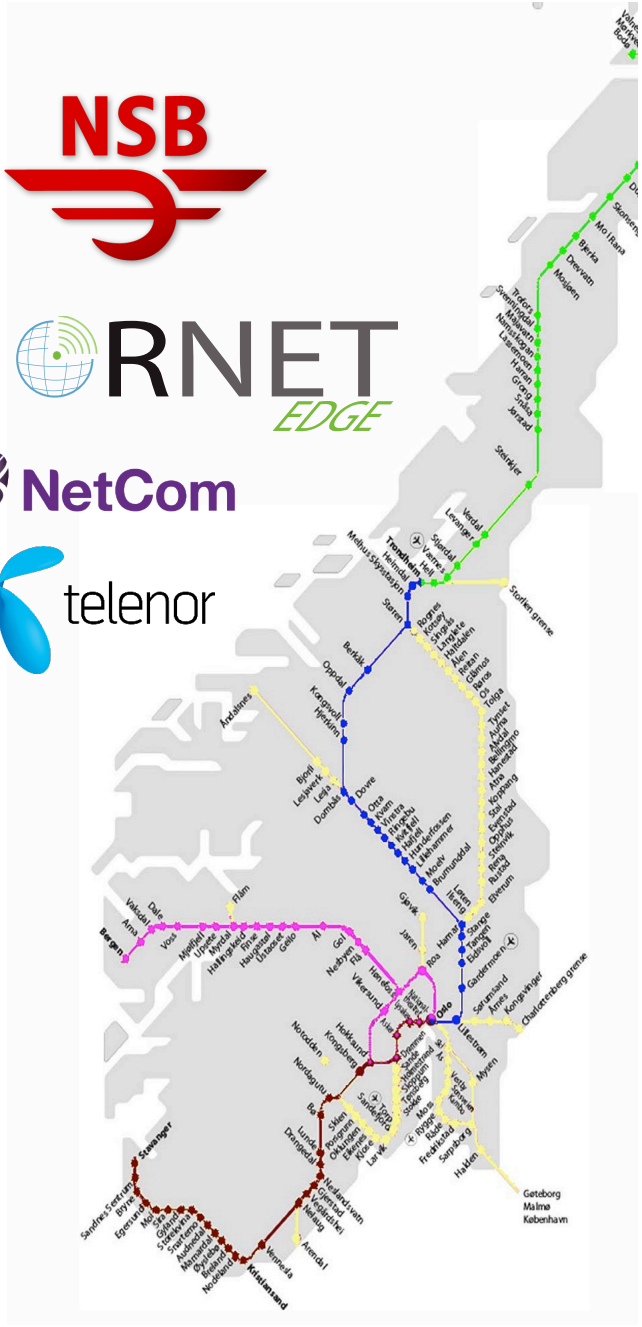
- **Drive testing**
  - High cost, requires custom equipment
  - Small area covered, no time component considered
- **Crowdsourcing**
  - No control over the measurement device
  - Lack of repeatability

Source: http://cinifglobal.in/

Source: opensignal.com

NNE data

- Data collected during **5 months** (Nov. 2014 – Mar. 2015) over 4 routes
  - **ACTIVE MEASUREMENTS:** UDP ping (20bytes packet sent every second to an NNE server)
  - **METADATA**: radio access technology (RAT), signal quality indicators (RSSI, RSRQ), network attachment information (cell ID, LAC), radio resource control state.



Geo-referenced Dataset

- Data collected over **5 months** (Nov. 2014 – Mar. 2015)
  - NNE metadata collected over 4 different routes: *Oslo-Voss, Oslo-Stavanger, Oslo-Trondheim, Trondheim-Bodø*
  - GPS data from the Norwegian railway system with 10-15 sec granularity
  - One run = a train trip on a given route

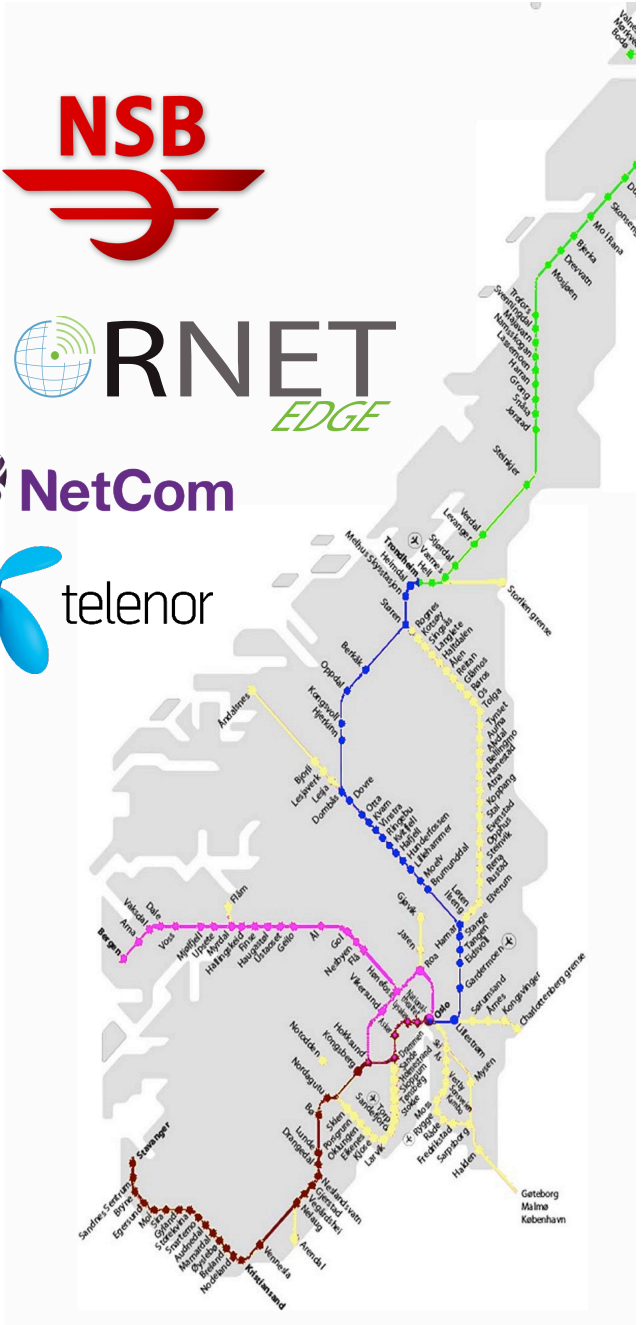
- Merge the ***NorNet Edge measurements*** with the ***GPS dataset*** to generate **geo-tagged data points**

NSB

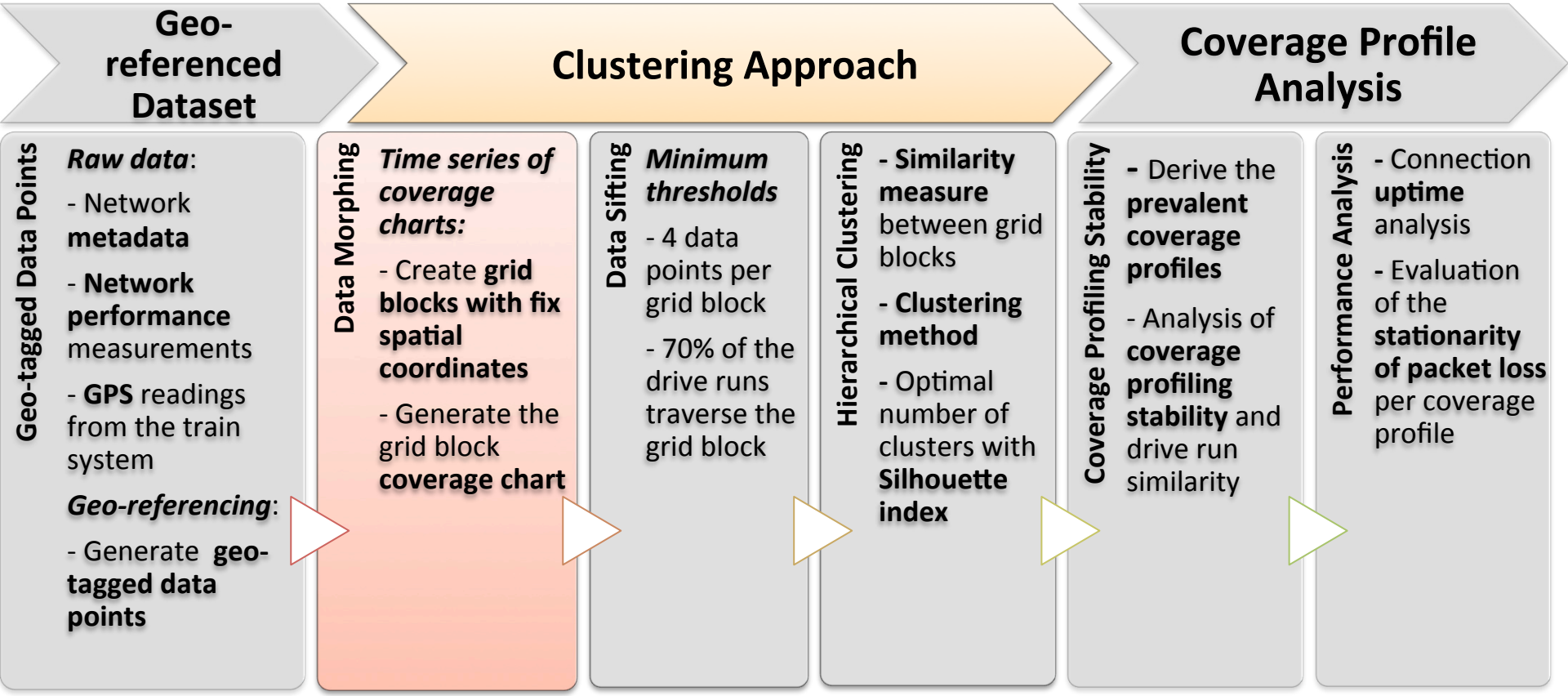
NORNET EDGE

NetCom

telenor



Workflow


Geo-referenced Dataset
Clustering Approach
Coverage Profile Analysis
Geo-tagged Data Points
Raw data:
- Network metadata
- Network performance measurements
- GPS readings from the train system
Geo-referencing:
- Generate geo-tagged data points
Data Morphing
Time series of coverage charts:
- Create grid blocks with fix spatial coordinates
- Generate the grid block coverage chart
Data Sifting
Minimum thresholds
- 4 data points per grid block
- 70% of the drive runs traverse the grid block
Hierarchical Clustering
- Similarity measure between grid blocks
- Clustering method
- Optimal number of clusters with Silhouette index
Coverage Profiling Stability
- Derive the prevalent coverage profiles
- Analysis of coverage profiling stability and drive run similarity
Performance Analysis
- Connection uptime analysis
- Evaluation of the stationarity of packet loss per coverage profile

Data Morphing

- **Geographical data binning**
  - Tackle limitations of the geo-referenced dataset
  - Group the geo-tagged data points into 2km x 2km geographical bins
- **Coverage Chart time series**
  - Distribution of the 4 different levels of RAT within a grid block
  - Time series -> all the trips the train makes on every route over 5 months

Data Morphing

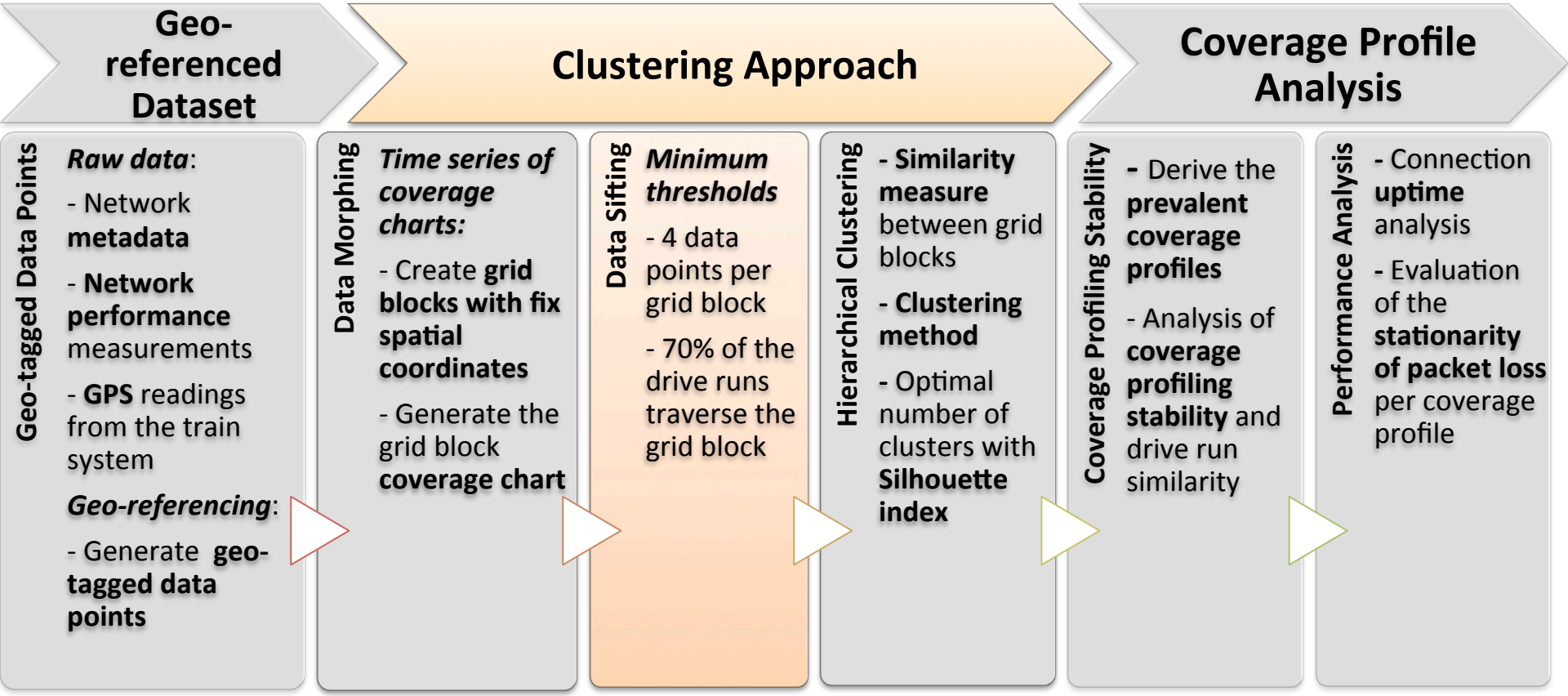
- **Geographical data binning**
  - Tackle limitations of the geo-referenced dataset
  - Group the geo-tagged data points into 2km x 2km geographical bins
- **Coverage Chart time series**
  - Distribution of the 4 different levels of RAT within a grid block
  - Time series -> all the trips the train makes on every route over 5 months

Data Morphing

- **Geographical data binning**
  - Tackle limitations of the geo-referenced dataset
  - Group the geo-tagged data points into 2km x 2km geographical bins
- **Coverage Chart time series**
  - Distribution of the 4 different levels of RAT within a grid block
  - Time series -> all the trips the train makes on every route over 5 months

Coverage Chart: {2G:20%, 3G:40%, 4G:20%, noS:20%}

Workflow

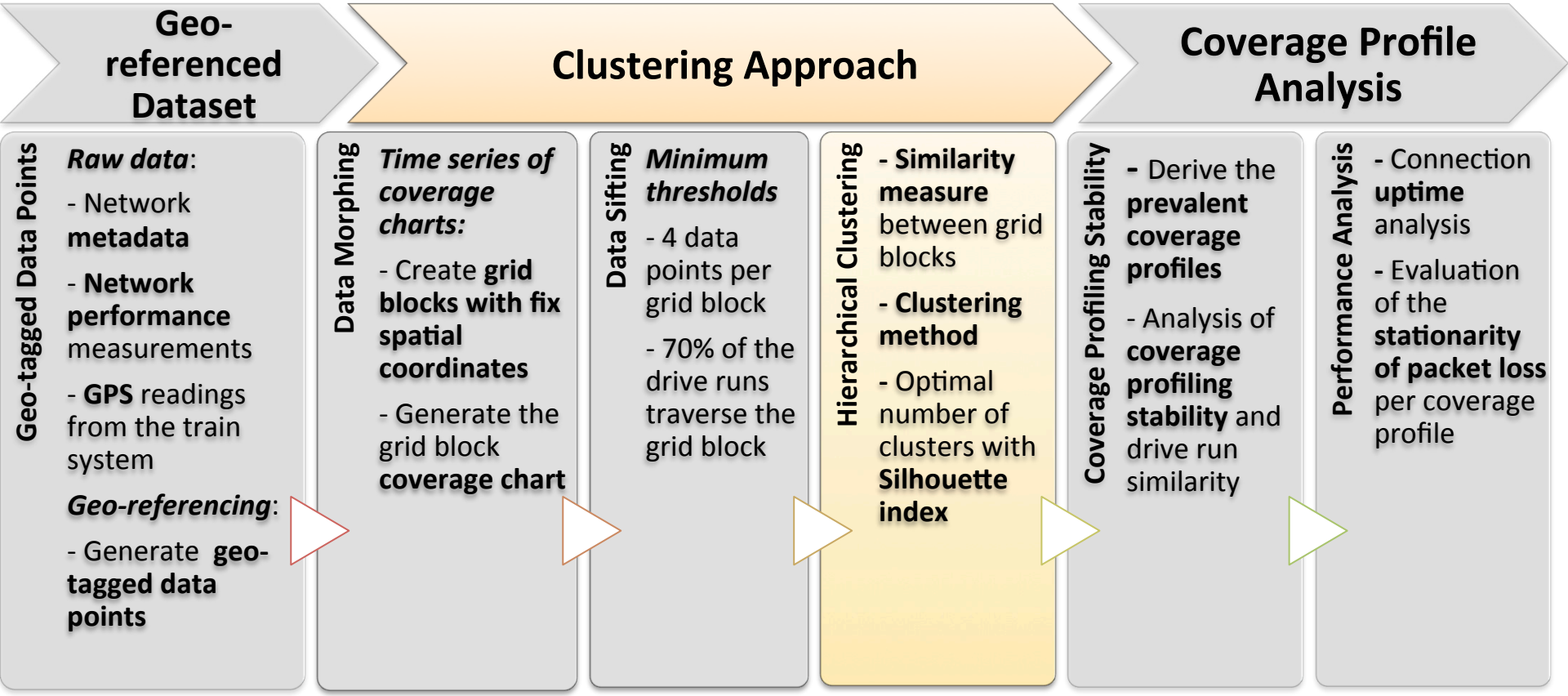

Geo-referenced Dataset
Clustering Approach
Coverage Profile Analysis
Geo-tagged Data Points
Raw data:
- Network metadata
- Network performance measurements
- GPS readings from the train system
Geo-referencing:
- Generate geo-tagged data points
Data Morphing
Time series of coverage charts:
- Create grid blocks with fix spatial coordinates
- Generate the grid block coverage chart
Data Sifting
Minimum thresholds
- 4 data points per grid block
- 70% of the drive runs traverse the grid block
Hierarchical Clustering
- Similarity measure between grid blocks
- Clustering method
- Optimal number of clusters with Silhouette index
Coverage Profiling Stability
- Derive the prevalent coverage profiles
- Analysis of coverage profiling stability and drive run similarity
Performance Analysis
- Connection uptime analysis
- Evaluation of the stationarity of packet loss per coverage profile

Data Sifting

- ***Spatial threshold:*** at least 4 geo-tagged data points per grid block
  - Ensure the measurement ***node spends at least 30-45sec in the grid-block*** to gather enough performance data (UDP ping)
- ***Temporal threshold:*** the grid block is traversed by at least 75% of all the runs in the dataset for that route
  - Ensure we have enough data to ***characterize coverage in the grid block over time***

Route	No threshold		Spatial threshold [data points]		Temporal threshold [drive runs]	
	Telenor	Netcom	Telenor	Netcom	Telenor	Netcom
Oslo - Voss	244	242	207	205	120	92
Oslo - Stavanger	374	374	319	319	182	182
Oslo - Trondheim	365	365	275	276	192	189
Trondheim - Bodø	514	514	380	378	228	157

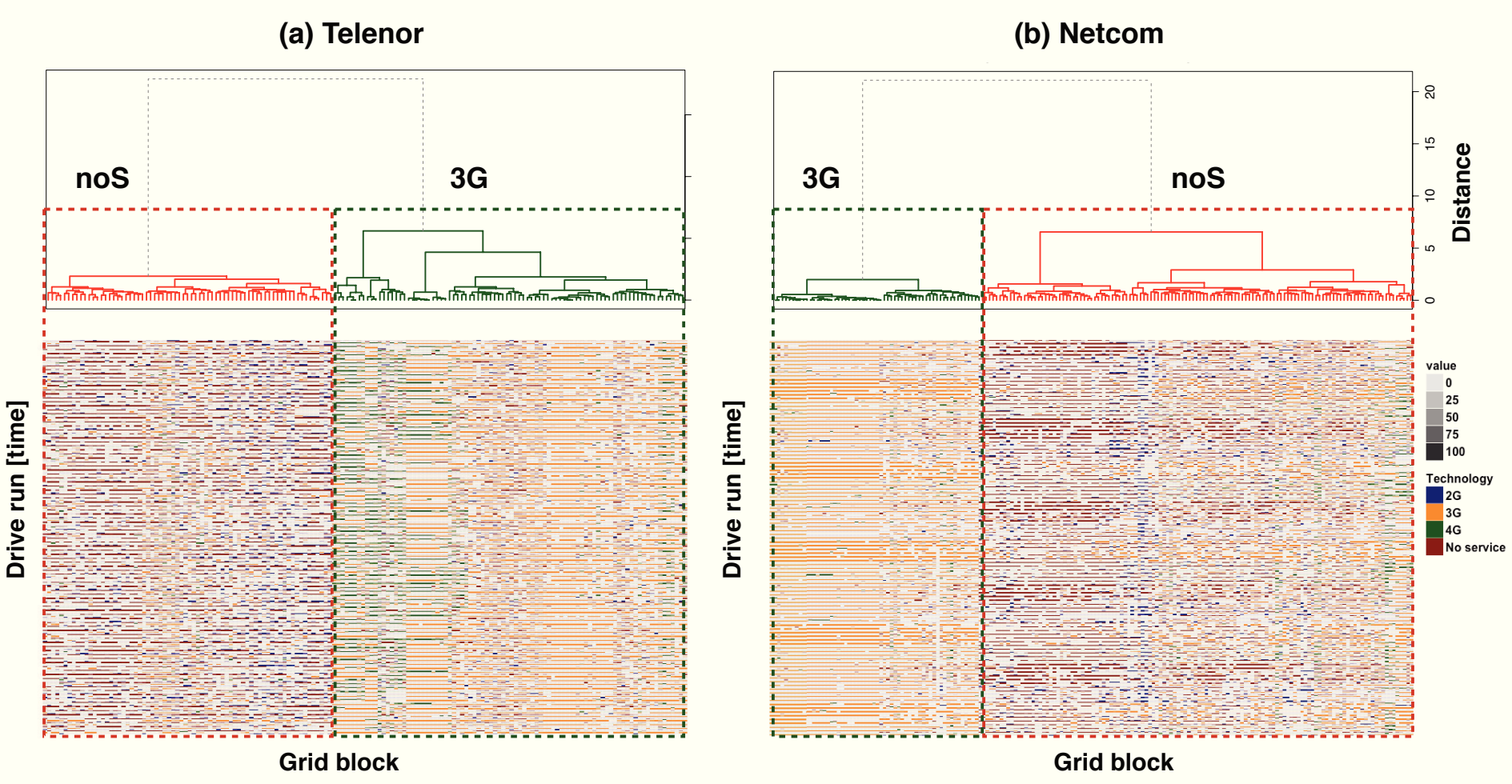
Workflow


Geo-referenced Dataset
Clustering Approach
Coverage Profile Analysis
Geo-tagged Data Points
Raw data:
- Network metadata
- Network performance measurements
- GPS readings from the train system
Geo-referencing:
- Generate geo-tagged data points
Data Morphing
Time series of coverage charts:
- Create grid blocks with fix spatial coordinates
- Generate the grid block coverage chart
Data Sifting
Minimum thresholds
- 4 data points per grid block
- 70% of the drive runs traverse the grid block
Hierarchical Clustering
- Similarity measure between grid blocks
- Clustering method
- Optimal number of clusters with Silhouette index
Coverage Profiling Stability
- Derive the prevalent coverage profiles
- Analysis of coverage profiling stability and drive run similarity
Performance Analysis
- Connection uptime analysis
- Evaluation of the stationarity of packet loss per coverage profile

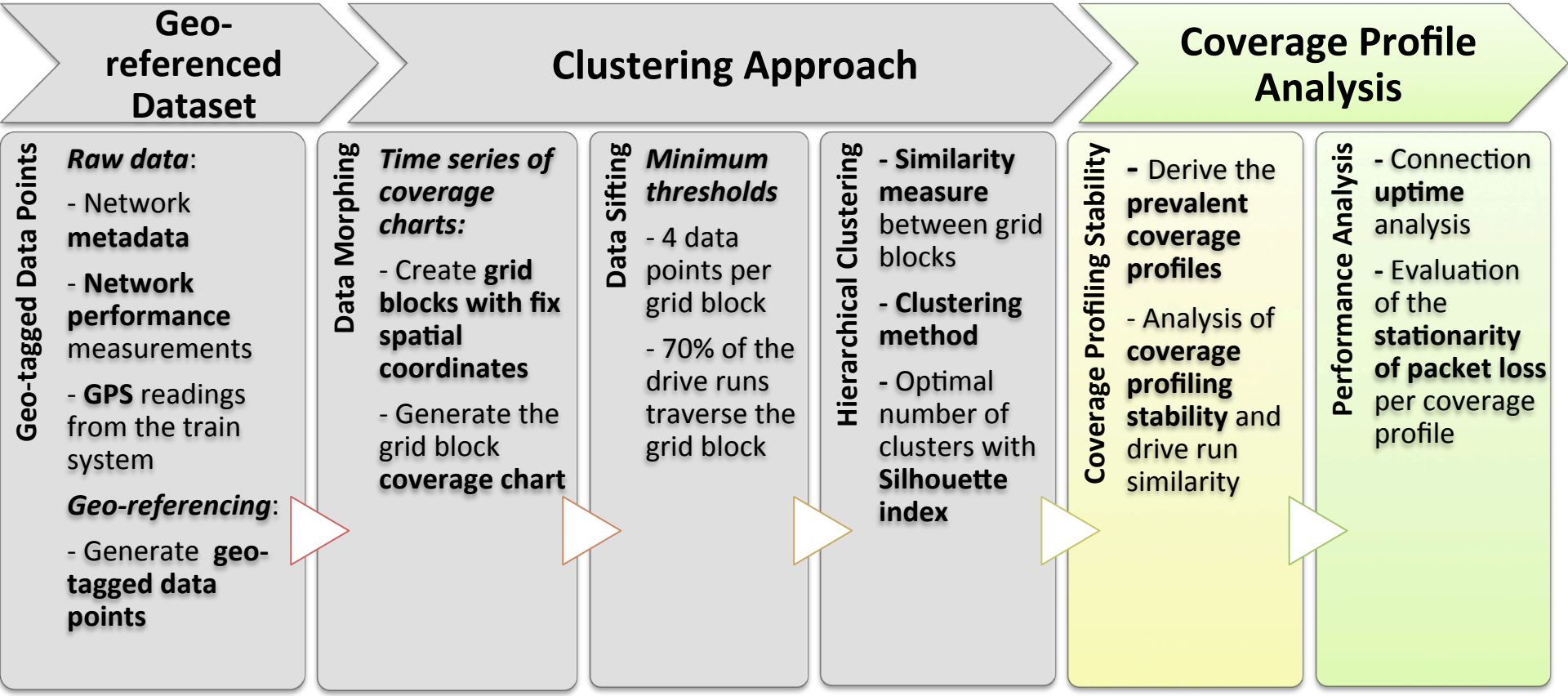
Clustering Approach

- ***Similarity measure:*** extended Jaccard measure
- ***Clustering method***: average-link hierarchical clustering of time series (Ward's minimum variance method)
- ***Optimal number of clusters***: validity index (Silhouette index)

Clustering Approach

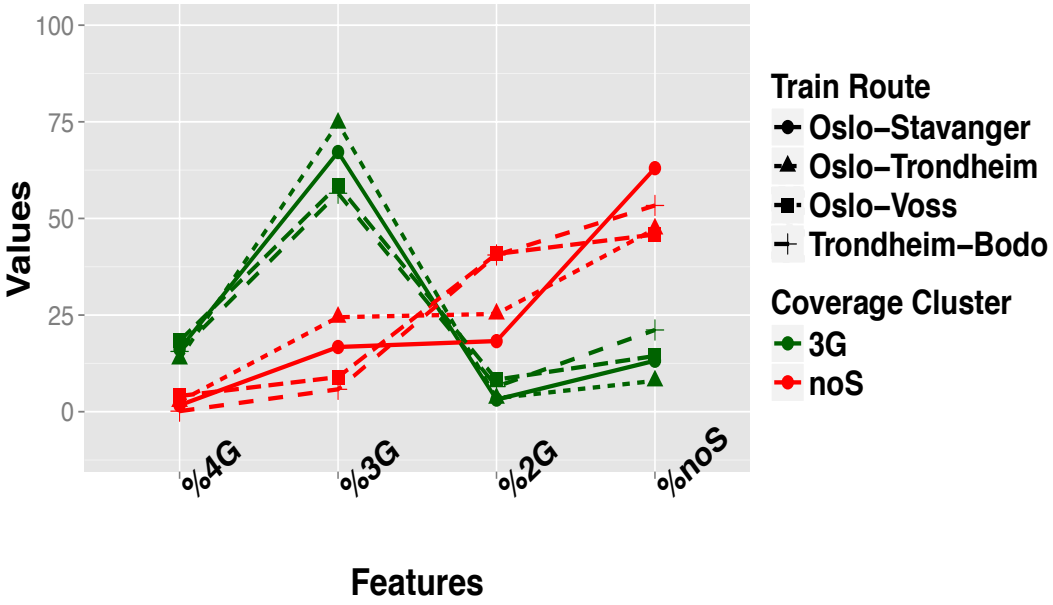


Workflow

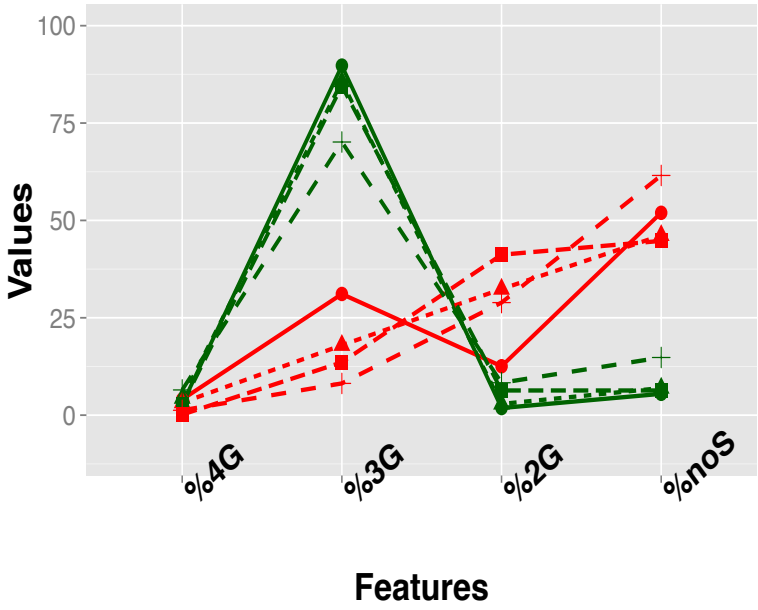

Geo-referenced Dataset
Clustering Approach
Coverage Profile Analysis
Geo-tagged Data Points
Raw data:
- Network metadata
- Network performance measurements
- GPS readings from the train system
Geo-referencing:
- Generate geo-tagged data points
Data Morphing
Time series of coverage charts:
- Create grid blocks with fix spatial coordinates
- Generate the grid block coverage chart
Data Sifting
Minimum thresholds
- 4 data points per grid block
- 70% of the drive runs traverse the grid block
Hierarchical Clustering
- Similarity measure between grid blocks
- Clustering method
- Optimal number of clusters with Silhouette index
Coverage Profiling Stability
- Derive the prevalent coverage profiles
- Analysis of coverage profiling stability and drive run similarity
Performance Analysis
- Connection uptime analysis
- Evaluation of the stationarity of packet loss per coverage profile

Prevalent Coverage Profiles

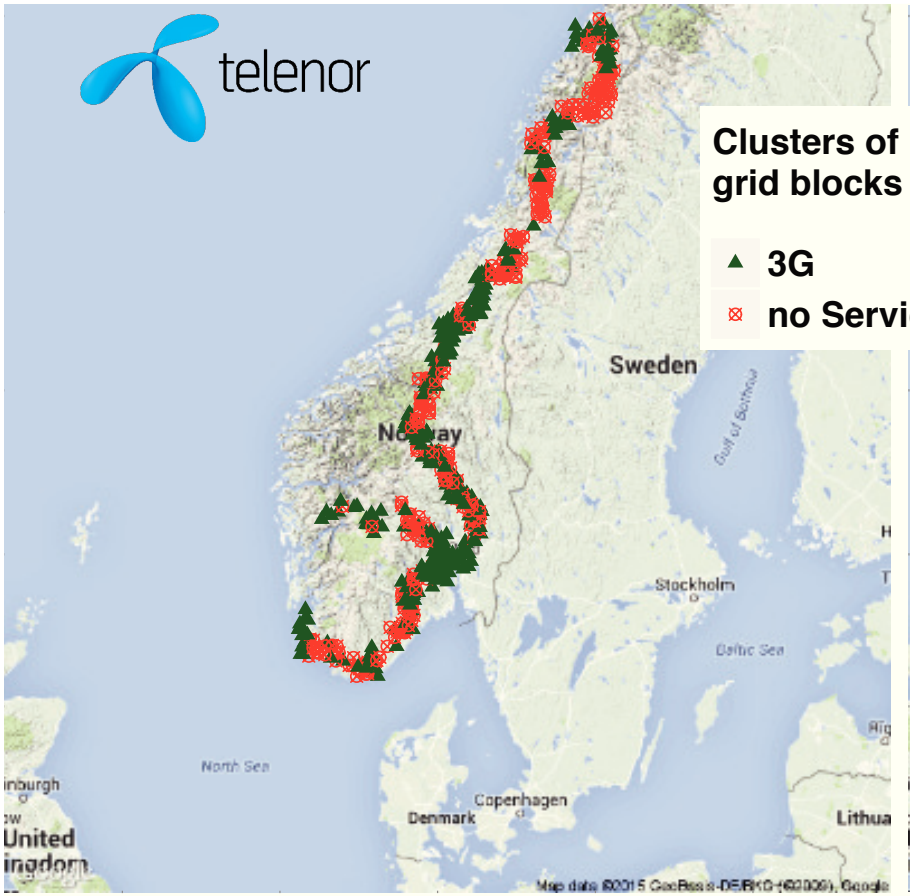
telenor



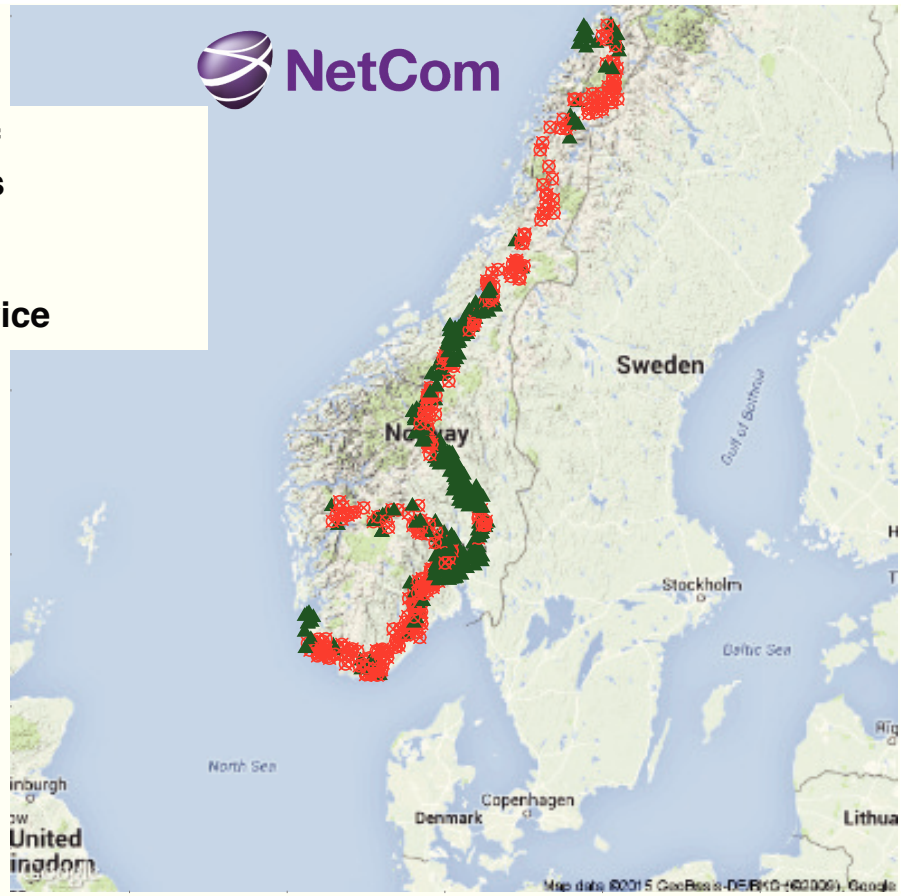
NetCom



Coverage Profiles on the Map



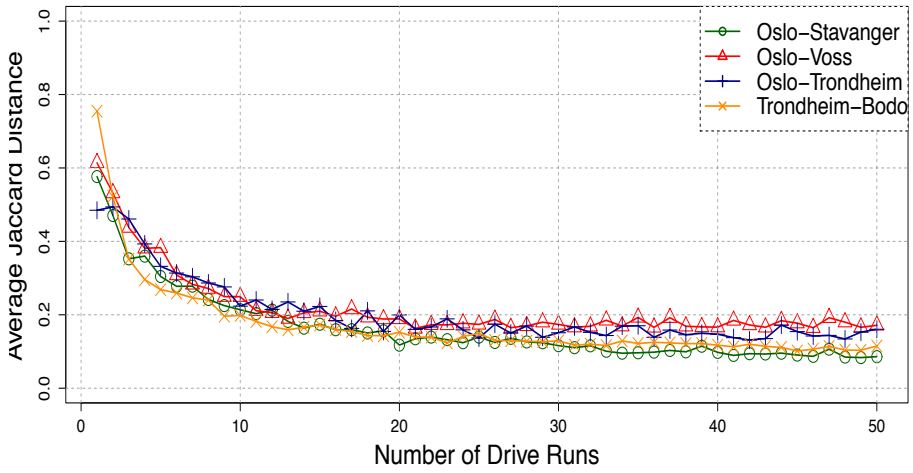
a) Telenor



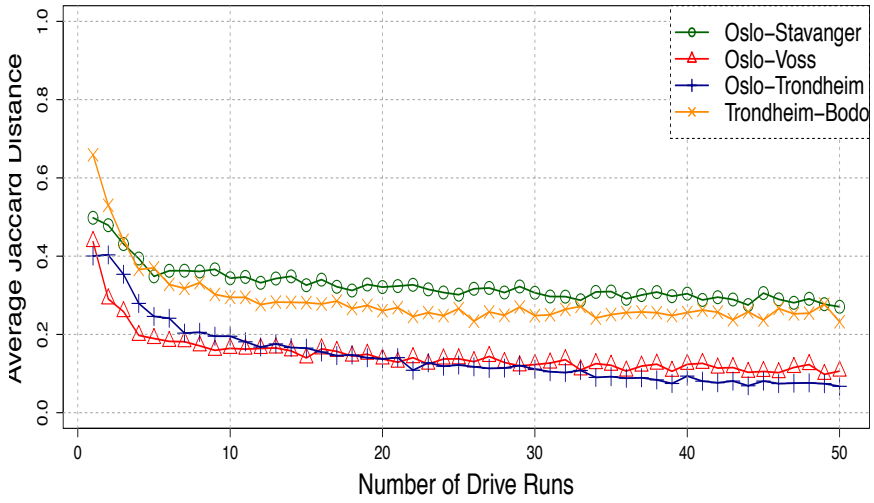
b) Netcom

Stability of Coverage Profiles and Similarity of Runs

telenor



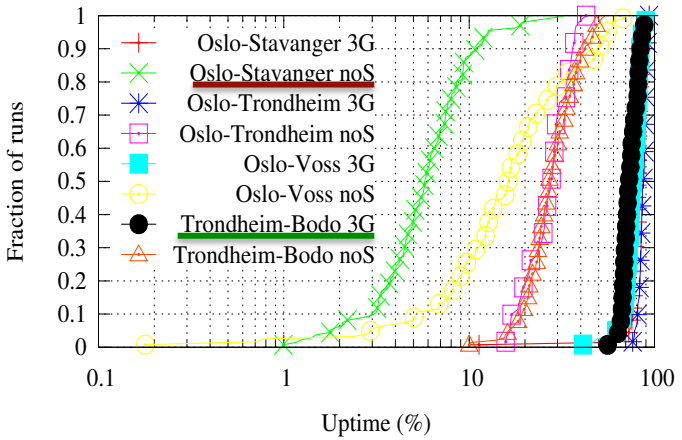
NetCom



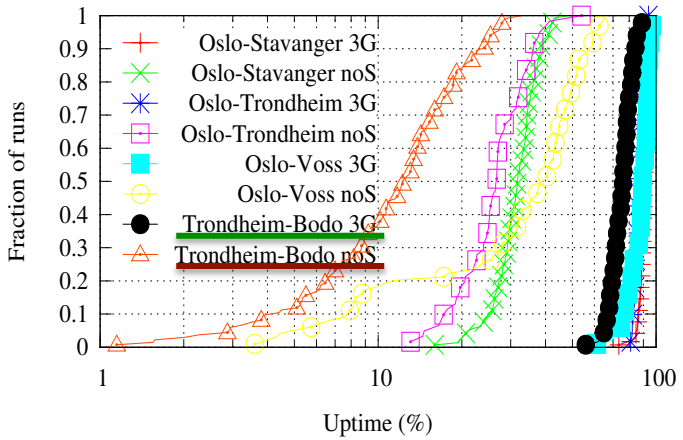
Coverage Implications: Uptime

- Uptime = #UDP_packets/T, T = interval of interest
- Dominant-noS cluster shows low uptime
- Trondheim-Bodø – worst performing “3G” cluster for both operators

telenor



NetCom

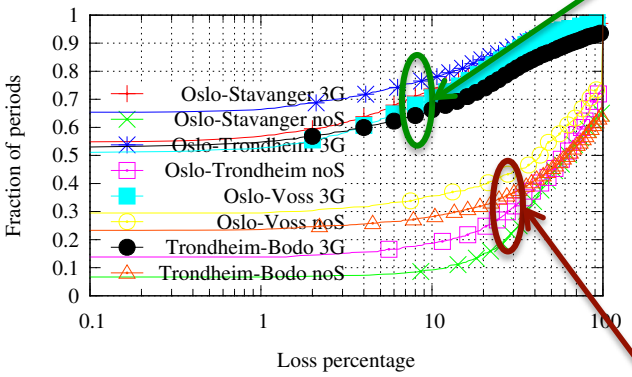


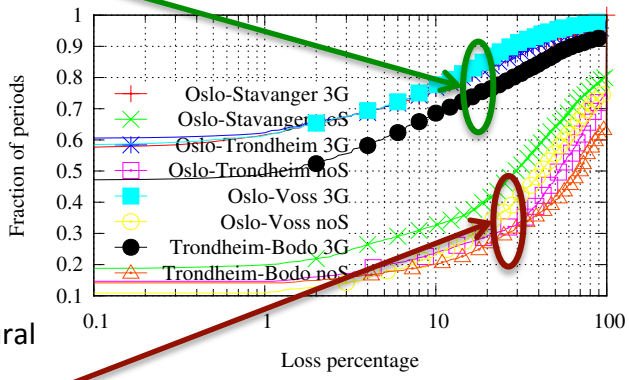
Coverage Implications: Packet Loss

telenor

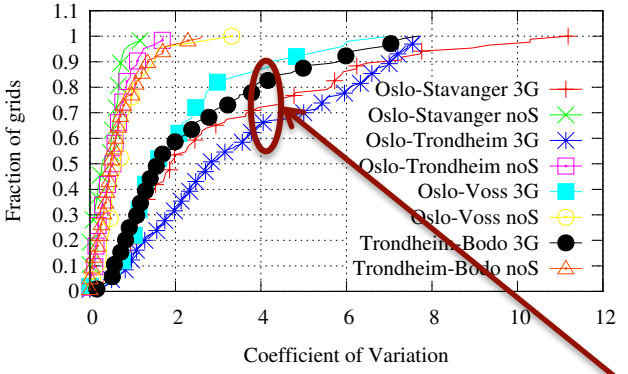
NetCom

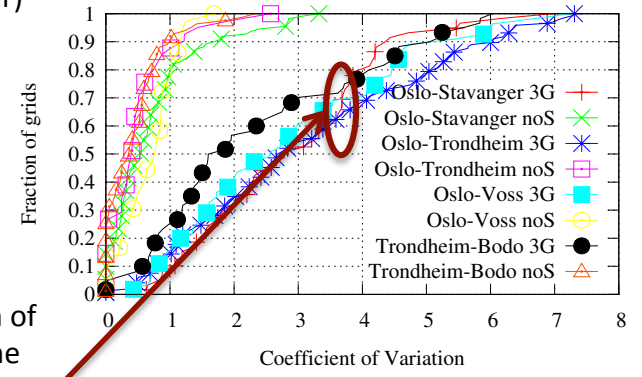
Low loss in urban areas (*3G* cluster)





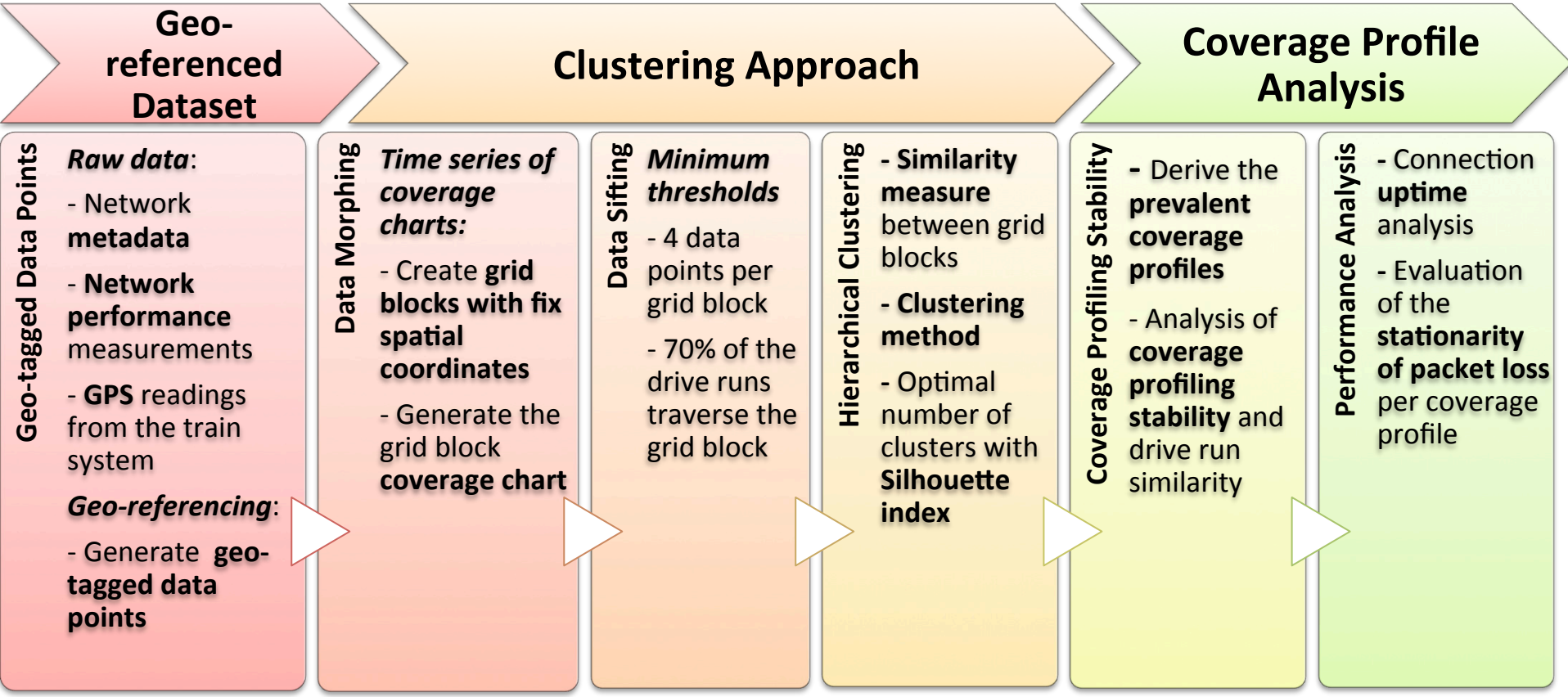
High loss in rural areas with patchy coverage (*noS* cluster)





High dispersion of samples for the *3G* cluster → heterogeneous cluster

Recap


Geo-referenced Dataset
Clustering Approach
Coverage Profile Analysis
Geo-tagged Data Points
Raw data:
- Network metadata
- Network performance measurements
- GPS readings from the train system
Geo-referencing:
- Generate geo-tagged data points
Data Morphing
Time series of coverage charts:
- Create grid blocks with fix spatial coordinates
- Generate the grid block coverage chart
Data Sifting
Minimum thresholds
- 4 data points per grid block
- 70% of the drive runs traverse the grid block
Hierarchical Clustering
- Similarity measure between grid blocks
- Clustering method
- Optimal number of clusters with Silhouette index
Coverage Profiling Stability
- Derive the prevalent coverage profiles
- Analysis of coverage profiling stability and drive run similarity
Performance Analysis
- Connection uptime analysis
- Evaluation of the stationarity of packet loss per coverage profile

For the future…

- Granularity of the coverage map
  - Higher granularity of the geo-tagged data points
  - Smaller size of the grid block
- Accuracy of coverage maps
  - Use spatial interpolation to generate more accurate
- Applications of the coverage maps
  - Dive into the correlation between the coverage profile and end-user experience
- With access to the mobile broadband connection of the railway system, compare effect of being inside the train
- Measurement node deployments in buses in Oslo – better coverage map in urban areas