University of Maribor

Faculty of Electrical Engineering and Computer Science

TIME-AWARE ANALYSIS OF OVERLAPPING CYCLING SEGMENTS FROM ACTIVITY TRACKER DATA

IEEE 23rd International Symposium on Intelligent Systems and Informatics (SISY 2025)

Authors:

Tilen Hliš (tilen.hlis@um.si), Iztok Fister (iztok.fister@proton.me), Patrik Rek (patrik.rek@um.si) and Iztok Fister Jr. (iztok.fister1@um.si)

PRESENTATION AGENDA

1. Motivation and purpose of the article
2. TCX data and preprocessing
3. Property graphs
4. Segment overlap detection algorithm
5. Experimental results
6. Discussion and future work
7. Conclusion

MOTIVATION

- **GPS tracker data:** Modern cycling trackers record detailed metrics (location, speed, heart rate, etc.) for every ride, producing large real-world datasets to analyze.

- **Need for deeper insights:**
  Detecting overlapping segments in cycling sessions can reveal **behavioral patterns** and **performance trends**.

- **Temporal dimension matters:** A rider’s performance on a given route can vary by **time of day**, **day of week**, and over multiple rides, so analyzing segments with time context is crucial for uncovering true performance patterns.

PURPOSE

- **Efficient detection pipeline (TCX2Graph.jl):** Developed a pipeline combining candidate segment generation, **KD-tree spatial filtering**, and **discrete Fréchet distance validation** for accurate segment matching.

- **Contextual data enrichment:** Integrated **OpenStreetMap data** to enrich each detected segment with road conditions and environmental context, providing deeper insights into performance variations.

- **Temporal performance study:** Conducted a **case study** on a frequently ridden segment, examining how speed, heart rate, and segment times differ between morning vs. afternoon, weekday vs. weekend, and over time.

TCX DATA STRUCTURE

What is a TCX File?

- **Training Center XML (TCX):**
  Developed to store data from GPS-enabled fitness devices.
- **Stores detailed workout data:**
  - **GPS Coordinates** (latitude, longitude, altitude).
  - **Performance Metrics** (speed, heart rate, cadence, power).
  - **Time-based session details.**

TCX File Hierarchy

- **Activity:**
  Overall session data (sport type, session metadata).
- **Lap:**
  Segments of the session (time, distance, heart rate).
- **Trackpoint:**
  Individual data points (timestamp, GPS, metrics).

TCX FILE


```xml
<TrainingCenterDatabase>
 <Activities>
  <Activity Sport="Biking">
   <Id>2023-10-19T10:38:47.000Z</Id>
   <Lap StartTime="2023-10-19T10:38:47.000Z">
    <TotalTimeSeconds>5050.412</TotalTimeSeconds>
    <DistanceMeters>29305.5</DistanceMeters>
    <MaximumSpeed>10.039999961853027</MaximumSpeed>
    <Calories>571</Calories>
    <AverageHeartRateBpm>
     <Value>121</Value>
    </AverageHeartRateBpm>
    <MaximumHeartRateBpm>
     <Value>160</Value>
    </MaximumHeartRateBpm>
    <Intensity>Active</Intensity>
    <TriggerMethod>Manual</TriggerMethod>
    <Track>
     <Trackpoint>
       <Time>2023-10-19T10:38:47.000Z</Time>
       <Position>
        <LatitudeDegrees>46.3886827044189</LatitudeDegrees>
        <LongitudeDegrees>15.728210052475333</LongitudeDegrees>
       </Position>
       <AltitudeMeters>237.60000610351562</AltitudeMeters>
       <DistanceMeters>0.0</DistanceMeters>
       <HeartRateBpm>
        <Value>91</Value>
       </HeartRateBpm>
       <Extensions>
        <ns3:TPX>
         <ns3:Speed>0.0</ns3:Speed>
        </ns3:TPX>
       </Extensions>
     </Trackpoint>
```

TCX FILES READING AND OSM FUSION

TCXReader.jl Library (https://github.com/firefly-cpp/TCXReader.jl)

- **Custom-built Julia library** for efficient reading and processing of TCX files.

Data Fusion with OpenStreetMap

- GPS alone isn’t enough — we add **context** with OSM:
  - Batch queries via **Overpass API**, linked to each trackpoint.

TCX		**OSM**	
Altitude	Cadence	Lane markings	Smoothness
Longitude	Speed	Lit	Surface
Latitude	Max speed	Crossing	Land use
Heart rate	Time	Width	–
Power	Distance	Incline	–
File name	–	Barrier	–

CONSTRUCTING THE PROPERTY GRAPH

Purpose of Property Graphs

- Provides a **structured representation** of TCX data for efficient storage and retrieval.
- Organizes **GPS track points** and related **performance metrics** for analysis.
- Facilitates the enrichment of detected overlapping segments with performance data.

What is a Property Graph?

- A **Directed Multigraph** with nodes and edges storing rich data.
- **Nodes (Vertices):** Represent **GPS track points** from TCX files.
- **Edges (Arcs):** Connect consecutive points, forming the athlete's path.
- **Properties:** Store key-value pairs (e.g., speed, heart rate, cadence).

Formal Definition: $G=\langle N,A,K,V,\alpha,\kappa,\pi\rangle$

CONSTRUCTING THE PROPERTY GRAPH

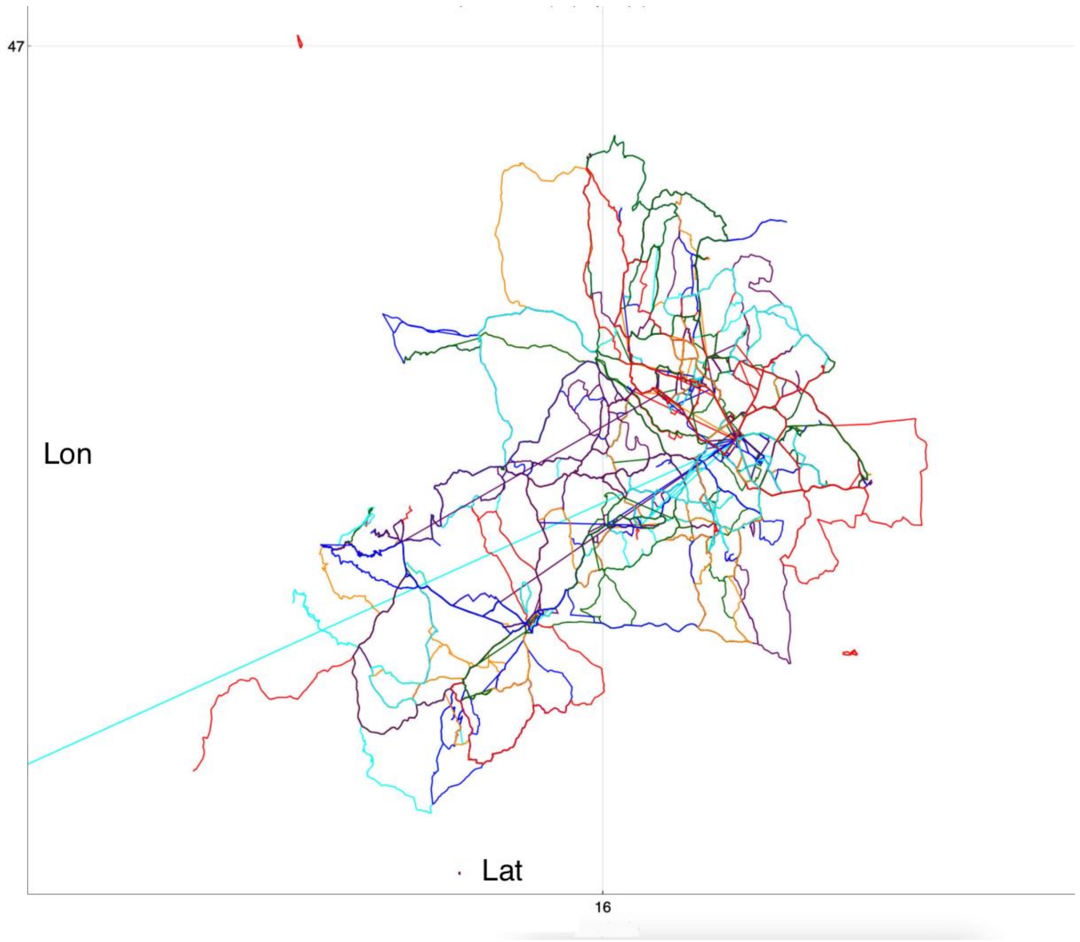
Graph Construction Workflow

1. **Parse TCX Files:** using TCXReader.jl.

2. **Add Nodes:** Each GPS track point becomes a graph vertex.

3. **Connect Nodes with Edges:** Sequential points are linked to form a path.

4. **Attach Properties:** Key metrics (speed, heart rate, cadence,...) are stored as properties of nodes and edges.



SEGMENTS OVERLAP DETECTION

Motivation for segment overlap detection

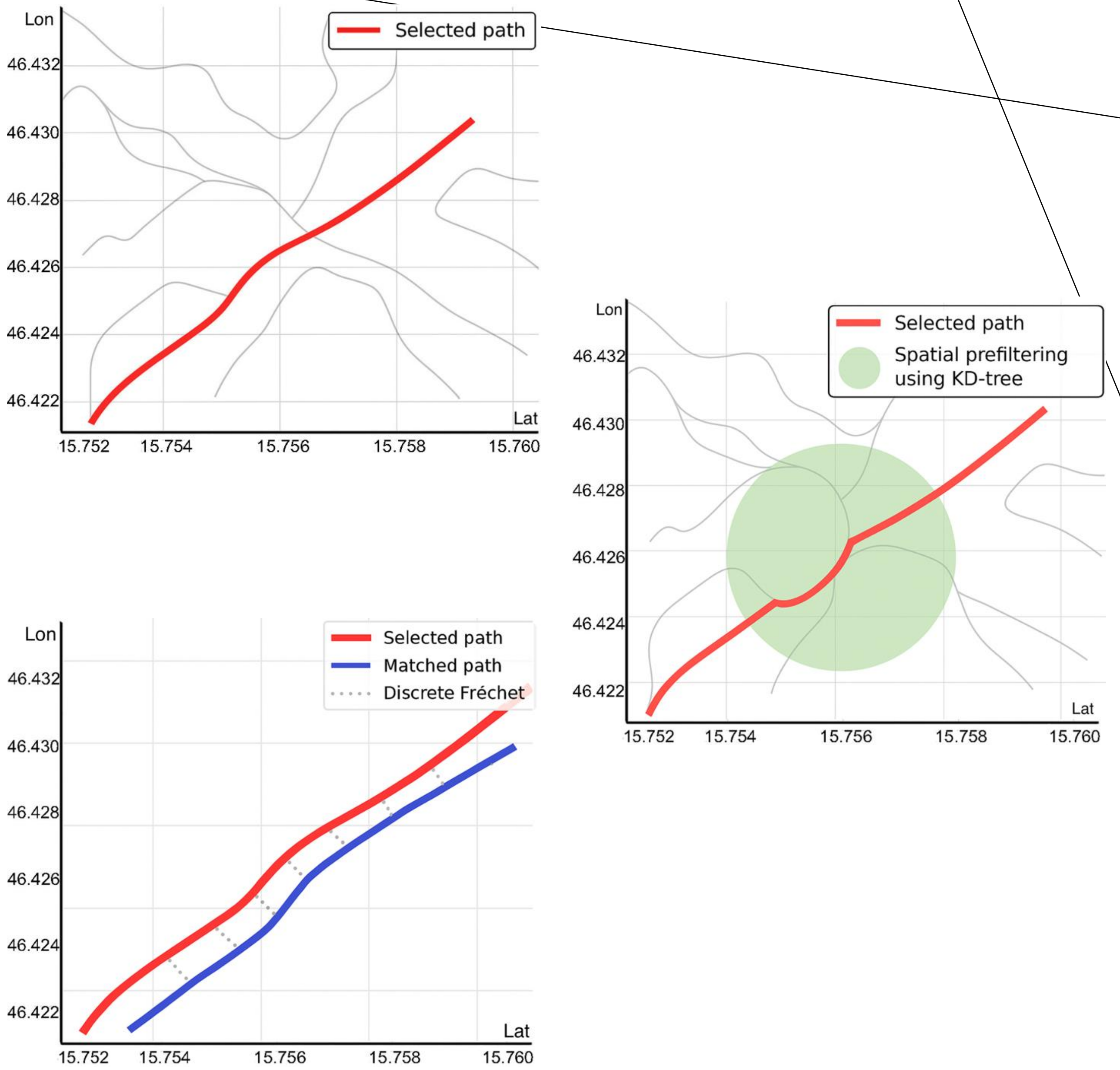
- **Cyclists often repeat routes**, leading to overlapping segments in TCX data.
- **Detecting these overlaps** helps analyze how performance changes across sessions.
- Understanding repeated segments provides insights into **training behavior** and **physiological responses**.

Challenges in detection

- **GPS inaccuracies:** Slight deviations in recorded data due to device errors.
- **Path variations:** Small changes in route due to cyclist behavior or environmental factors.
- **Efficient processing:** Large datasets require optimized algorithms for real-time analysis.

SEGMENTS OVERLAP DETECTION

1. **Reference ride:** Select one ride and generate candidate segments.
2. **KD-Tree filter:** Use KD-tree to find nearby rides within a bounding region.
3. **Fréchet check:** Compare shapes using discrete Fréchet distance.
4. **Validation:** Mark as overlap if distance ≤ *tolerance*.



EXPERIMENTAL SETUP

Objective

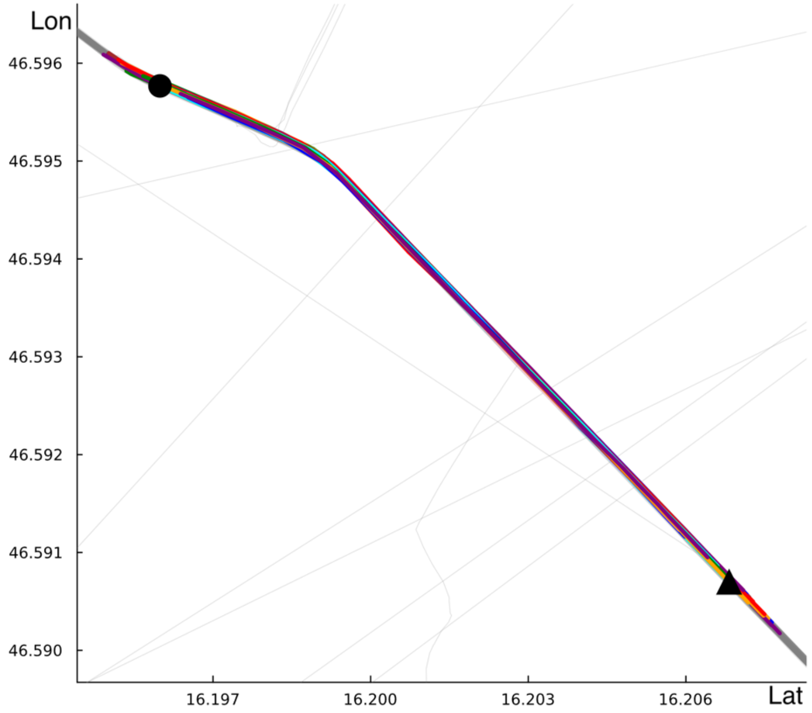
- ***Time-aware performance analyses*** *by isolating the effect of temporal factors on cycling performance using a controlled, consistent road segment.*

Data sources:

- TCX files + OSM geospatial context fused into one enriched dataset.

Dataset:

- **462 TCX rides**, single cyclist, repeated sessions.



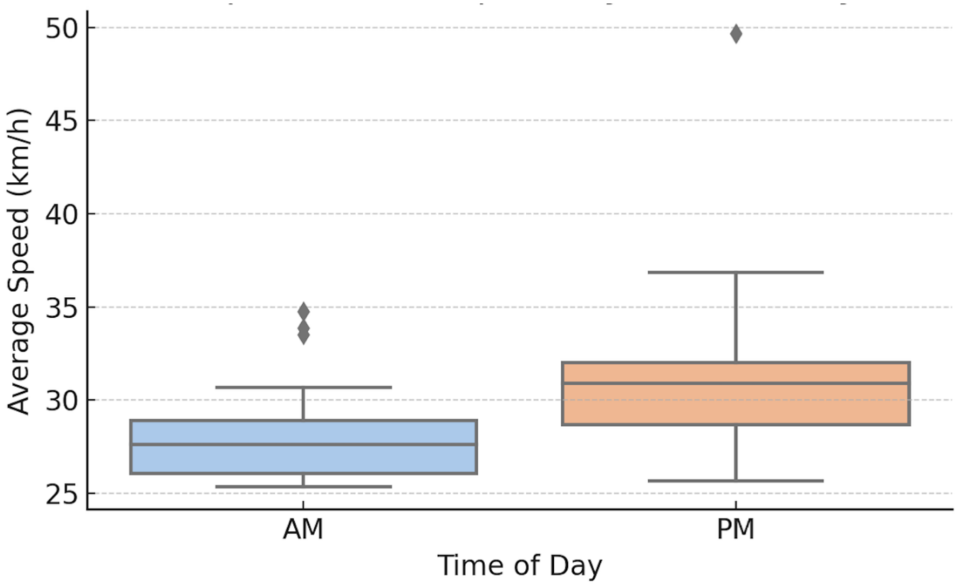
Selected segment for case studies:

- Length **1.027 m, 52 rides**, asphalt, flat terrain (avg. gain 1.85 m).
- Consistently appears **early in each ride (~first 40% of path)** → minimal fatigue confounding.
- Stable characteristics ideal for temporal comparisons.

CASE STUDY 1 – MORNING VS AFTERNOON

Research Question (RQ1):

- *Does the cyclist's performance on the segment differ between morning and afternoon rides?*



Grouping:

- **Morning rides:** Start before 12:00 (***n = 21***)
- **Afternoon rides:** Start at or after 12:00 (***n = 31***)

Results:

- **Afternoon** rides **significantly faster and shorter.**

Interpretation:

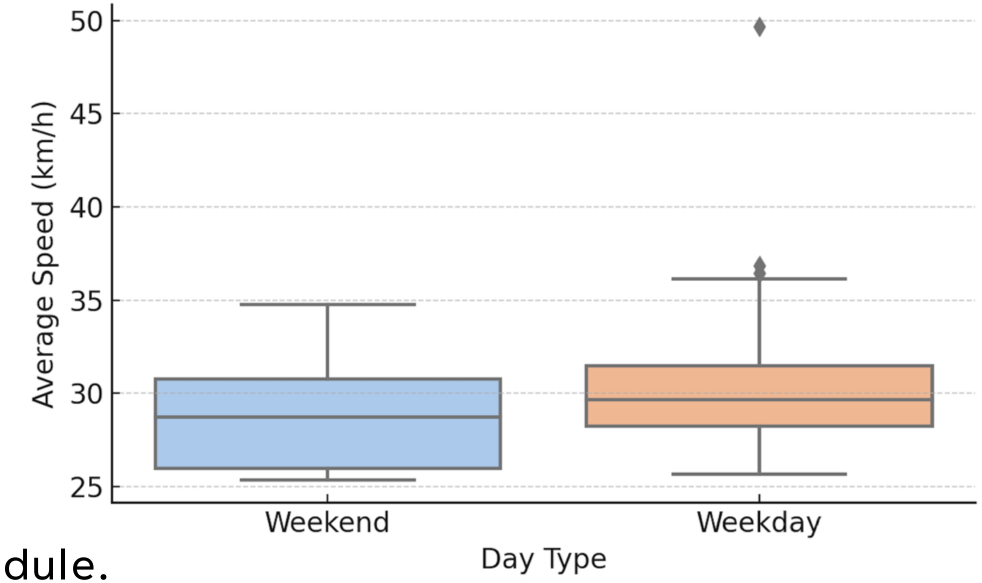
- Clear **time-of-day effect**, consistent with known **circadian peaks** in sports performance.

Metric	AM (n = 21)	PM (n = 31)	p-value
Avg. Speed (km/h)	28.29 ± 2.82	**31.36 ± 4.49**	**0.004**
Avg. Heart Rate (bpm)	136.21 ± 11.35	141.13 ± 11.80	0.139
Duration (s)	129.33 ± 13.81	**116.87 ± 17.89**	**0.007**

CASE STUDY 2 – WEEKDAY VS WEEKEND

Research Question (RQ2):

- *Does performance on the segment differ between weekdays and weekends?*



Grouping:

- **Weekdays:** Monday-Friday (***n = 38***)
- **Weekends:** Saturday-Sunday (***n = 14***)

Results:

- **No statistically** significant differences.

Interpretation:

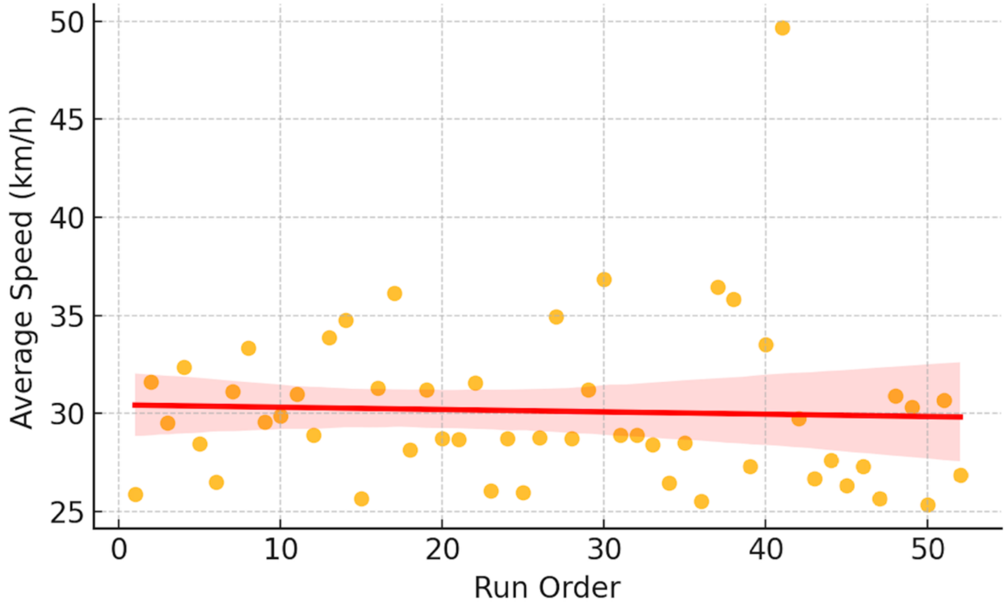
- Performance remained **consistent** regardless of weekly schedule.
- Slight weekday speed edge likely random or due to minor external factors.

Metric	Weekday (n = 38)	Weekend (n = 14)	p-value
Avg. Speed (km/h)	30.58 ± 4.40	28.87 ± 3.23	0.135
Avg. Heart Rate (bpm)	140.36 ± 11.86	135.85 ± 11.24	0.218
Duration (s)	120.39 ± 18.34	126.00 ± 14.12	0.253

CASE STUDY 3 – PERFORMANCE OVER TIME

Research Question (RQ2):

- *Does performance improve, decline, or remain stable over repeated rides?*



Approach:

- Chronologically ordered all **52 rides**.
- Applied **linear regression** to:
  - Speed (km/h), Heart rate (bpm), Duration (s).

Results:

- **No statistically** significant differences.

Interpretation:

- **Stable performance over time**, no measurable training improvement or decline.

Metric	**Slope (per ride)**	**p-value**
Avg. Speed (km/h)	−0.005	0.891
Avg. Heart Rate (bpm)	+0.051	0.641
Duration (s)	+0.077	0.636

DISCUSSION AND FUTURE WORK

- **Segment-based temporal analysis** enables precise, repeatable insights into performance on **specific road segments**, supporting personalized training and terrain-specific performance evaluation.

- Ongoing integration of **weather data** will soon allow detailed analysis of how environmental conditions influence performance.

- Development of **association rule mining (ARM)** capabilities is in progress to discover complex patterns and relationships across multiple overlapping segments and rides.

- Broader potential beyond sports: applications in **urban cycling planning**, **route optimization**, and **transportation research**.

- Future capability to **generate entirely new paths** by intelligently combining discovered overlapping segments, opening new possibilities.

- Expanding datasets and refining algorithms will extend the potential of **TCX2Graph.jl** to other domains.

THANK YOU FOR YOUR ATTENTION