Overview

In the task scenario, there are multiple objects in motion being detected by a network of sensors, with each object potentially being detected by more than one sensor. The goal is to create a unified dataset that provides a comprehensive view of object movements and positions as detected by a network of sensors.

Input Data Format

The input data is in the format of csv containing the following columns:

timestamp_id: The timestamp when the detection was made.
sensor_id: The identifier of the sensor that made the detection.
id: A sensor-specific identifier assigned to each detected object.
x_position: The x-coordinate of the detected object's position.
y_position: The y-coordinate of the detected object's position.
unique_id: A universal identifier for the object. A value of 0 indicates the unique ID is not known.

Processing Steps

1. Reads the CSV file and processes the data sequentially by timestamp.
2. Clusters detections based on their **(x, y)** positions using the DBSCAN algorithm with a 2-meter threshold for proximity. The goal was to group detections that were within 2 meters of each other, considering this distance as a threshold for determining whether detections are of the same object.
3. As part of the clustering, the process ensures that all detections within a cluster are assigned the same ***unique_id*** if at least one detection in the cluster has a known ***unique_id***.
4. Generates a fused ID ***(f_id)*** randomly for each cluster, along with a fused timestamp ***(f_timestamp)*** and clusters the data ***(cluster_data)*** accordingly.
5. Outputs the processed data to a new CSV file with columns for ***f_timestamp, f_id, cluster_data, and f_u_id***.

Output Data Format

The output CSV file contains the following columns:

- ***f_timestamp:*** The fused timestamp for the cluster, representing the time of the detections.
- ***f_id:*** A randomly assigned identifier for the cluster.
- ***cluster_data:*** A list of lists, where each inner list contains the x and y positions and the sensor ID of a detection in the cluster.
- ***f_u_id:*** The unique ID assigned to the cluster, based on the ***unique_id*** of the detections.