Ludwig-Maximilians-Universität München
Institut für Informatik
Prof. Dr. Peer Kröger
Yifeng Lu

Knowledge Discovery in Databases II
SS 2019

## Exercise 7: Data Stream Clustering

Exercise 7-1 Damped Window Model

Assume a special microcluster decaying mechanism, where all microclusters are fading out after each time stamp according to $f(t) = b^{-\lambda t}$, and only the weight of the microcluster that is currently updated (hit by a point in the current timestamp) increases by 1.

(a) What is the maximum weight of a microcluster?

(b) What is the minimum time needed for a newly created microcluster to become potential (weight larger than τ)?

(c) What is the minimum time needed for a potential microcluster of weight w to became an outlier (weight less than τ)?

Exercise 7-2 Cluster Features

Given the following dataset:

ObjID	Cluster	X	Y	t
1	A	3	3	1.7
2	A	4	3	3.5
3	A	3	2	1.2
4	B	9	6	4.1
5	B	7	5	5.0
6	B	8	4	1.2
7	C	5	7	4.7
8	C	4	7	2.3
9	C	3	7	2.2
10	C	4	8	2.2
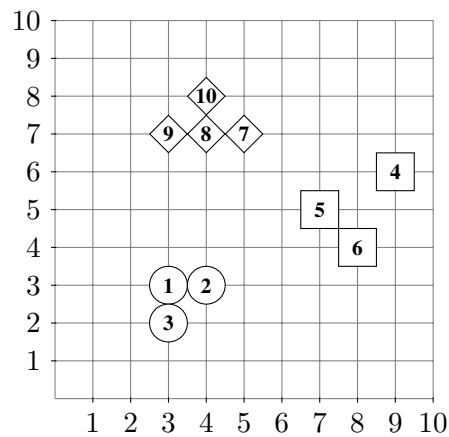
Compute the CluStream cluster features CFT for each of these three clusters.

A new observation in the stream is $p = (X = 8, Y = 5, t = 6.1)$.

Run the "online micro-cluster maintainance" of CluStream for this Point p.