

The background of the slide is a complex, abstract composition. It features a network graph with numerous green nodes and red edges, overlaid on a grid of small grey plus signs. The background is divided into several geometric sections by white lines, including a large white triangle on the right and a smaller one on the left. A small inset image in the top left corner shows a cluster of orange and red dots with a grid overlay. The title text is centered in a large, bold, black font.

Session 4: External Measures I: Matching-Based Measures

In-video Quiz

❑ The following table summarizes the clustering result of a clustering algorithm. What is the maximum matching schema for maximum matching, i.e., which partition should each cluster match?

❑ $(C_1-T_3), (C_2-T_3), (C_3-T_1)$

❑ $(C_1-T_2), (C_2-T_3), (C_3-T_1)$

❑ $(C_1-T_3), (C_2-T_1), (C_3-T_1)$

❑ $(C_1-T_2), (C_2-T_2), (C_3-T_1)$

❑ Answer: $(C_1-T_3), (C_2-T_1), (C_3-T_1)$

❑ Explanation: First off, based on the definition of **Maximum matching**, we have that one cluster can match one partition.

❑ $(C_1-T_3), (C_2-T_3), (C_3-T_1)$: two clusters C_1 and C_2 match the same partition T_3 .

❑ $(C_1-T_2), (C_2-T_2), (C_3-T_1)$: two clusters C_1 and C_2 match the same partition T_2 .

❑ $(C_1-T_2), (C_2-T_3), (C_3-T_1)$: The weight matching is $15 + 30 + 45 = 90$;

❑ $(C_1-T_3), (C_2-T_1), (C_3-T_1)$: The weight matching is $45 + 25 + 45 = 115 > 90$.

❑ Thus the maximum matching schema should be $(C_1-T_3), (C_2-T_1), (C_3-T_1)$.

$C \backslash T$	T_1	T_2	T_3	Sum
C_1	5	15	45	65
C_2	10	25	30	65
C_3	45	15	10	70
m_j	60	55	85	200