

The background of the slide is a complex, abstract composition. It features a dark, muted purple or brownish background. Overlaid on this are several geometric and data-related elements: a network of thin, light-colored lines forming a mesh or web-like structure; numerous small, colored dots (green, blue, yellow) scattered across the field; and a prominent, lighter-colored, semi-transparent geometric shape in the center, resembling a stylized 'V' or a folded piece of paper. In the top-left corner, there is a small, rectangular inset showing a different visualization, possibly a heatmap or a scatter plot with a grid. The overall aesthetic is technical and modern, suggesting a focus on data science or computer graphics.

# **Session 6: Mining Closed Patterns**

# CLOSET+: Mining Closed Itemsets by Pattern-Growth

- ❑ Efficient, *direct* mining of closed itemsets
- ❑ Ex. Itemset merging: If Y appears in every occurrence of X, then Y is merged with X
  - ❑ d-proj. db: {acef, acf} → acfd-proj. db: {e}, thus we get: acfd:2
- ❑ Many other tricks (but not detailed here), such as
  - ❑ Hybrid tree projection
    - ❑ Bottom-up physical tree-projection
    - ❑ Top-down pseudo tree-projection
  - ❑ Sub-itemset pruning
  - ❑ Item skipping
  - ❑ Efficient subset checking
- ❑ For details, see J. Wang, et al., “CLOSET+: .....”, KDD'03

TID	Items
1	acdef
2	abe
3	cefg
4	acdf

Let minsupport = 2

a:3, c:3, d:2, e:3, f:3

F-List: a-c-e-f-d

# Recommended Readings

---

- ❑ R. Agrawal and R. Srikant, “Fast algorithms for mining association rules”, VLDB'94
- ❑ A. Savasere, E. Omiecinski, and S. Navathe, “An efficient algorithm for mining association rules in large databases”, VLDB'95
- ❑ J. S. Park, M. S. Chen, and P. S. Yu, “An effective hash-based algorithm for mining association rules”, SIGMOD'95
- ❑ S. Sarawagi, S. Thomas, and R. Agrawal, “Integrating association rule mining with relational database systems: Alternatives and implications”, SIGMOD'98
- ❑ M. J. Zaki, S. Parthasarathy, M. Ogihara, and W. Li, “Parallel algorithm for discovery of association rules”, Data Mining and Knowledge Discovery, 1997
- ❑ J. Han, J. Pei, and Y. Yin, “Mining frequent patterns without candidate generation”, SIGMOD'00
- ❑ M. J. Zaki and Hsiao, “CHARM: An Efficient Algorithm for Closed Itemset Mining”, SDM'02
- ❑ J. Wang, J. Han, and J. Pei, “CLOSET+: Searching for the Best Strategies for Mining Frequent Closed Itemsets”, KDD'03
- ❑ C. C. Aggarwal, M.A., Bhuiyan, M. A. Hasan, “Frequent Pattern Mining Algorithms: A Survey”, in Aggarwal and Han (eds.): Frequent Pattern Mining, Springer, 2014