

The background features a complex, abstract design. It includes a network of thin, reddish-brown lines forming a web-like structure. Scattered throughout are numerous small, green circular dots. On the left side, there is a vertical strip with a grid of small, light-colored squares. In the upper left corner, there is a small inset image showing a cluster of orange and red dots. The overall color palette is muted, with earthy tones and a soft, hazy atmosphere.

Mining Spatial Colocation Patterns

Spatial Colocation Patterns

□ **Colocation pattern:** A group of spatial features or events that are frequently co-located in the same region

□ Ex. West Nile Virus often occur in regions with poor mosquito control and the presence of birds

□ Figure: Neighborhood instances are connected by edges

□ Ex. {3, 6, 17}, {4, 7, 10, 16}, {2, 8, 11, 14, 15}, {2, 9}, ...

□ Rowset(C) if every feature in patter C appears as a feature of an instance in the neighbor-set L, e.g.,

□ $\text{rowset}(\{A, B, C, D\}) = \{\{4, 7, 10, 16\}, \{2, 11, 14, 15\}, \{8, 11, 14, 15\}\}$

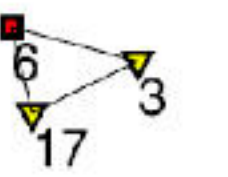
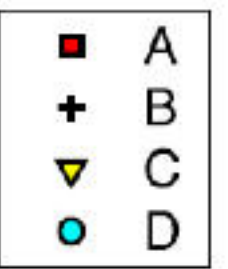
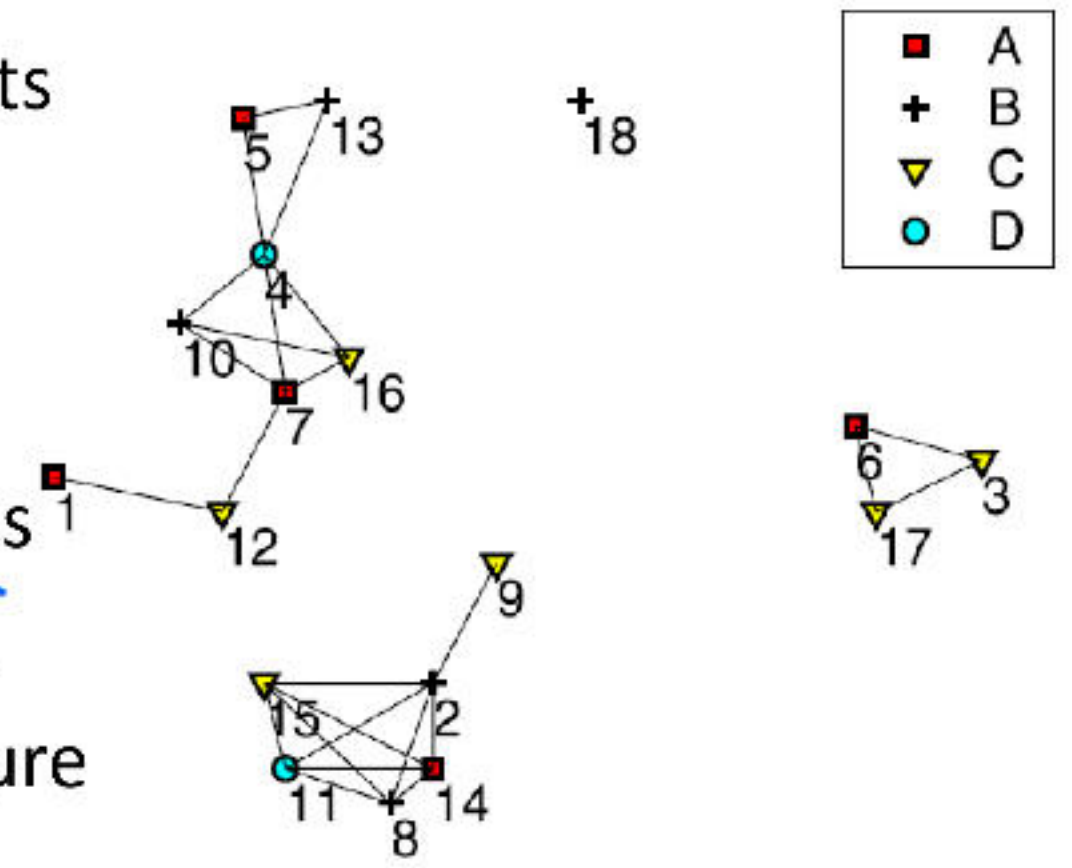
□ $\text{rowset}(\{A, B\}) = \{\{5, 13\}, \{7, 10\}, \{2, 14\}, \{8, 14\}\}$

□ A colocation rule $R: A \rightarrow B$, conditional probability $\text{cp}(R)$ is defined as

$$\text{cp}(R) = \frac{|\{L \in \text{rowset}(A) \mid \exists L' \text{ s.t. } (L \subseteq L') \wedge (L' \in \text{rowset}(A \cup B))\}|}{|\text{rowset}(A)|}$$

□ $\text{cp}(\{A, B\} \rightarrow \{C, D\}) = |\text{rowset}(\{A, B, C, D\})| / |\text{rowset}(\{A, B\})| = \frac{3}{4} = 75\%$

即 A, B 为 collocated, 求 C, D 为 collocated 的概率。



字面意思理解

Mining Spatial Colocation Patterns

- Participation ratio $pr(C, f)$: probability that C is observed in a neighbor-set wherever feature f is observed

$$pr(C, f) = \frac{||\{r | (r \in S) \wedge (r.f = f) \wedge (r \text{ is in a row instance of } C)\}||}{|\{r | (r \in S) \wedge (r.f = f)\}|}$$

- Ex. $pr(\{A, B, C, D\}, A) = 2/5, \dots, pr(\{A, B, C, D\}, D) = 2/2 = 1$

单调 Monotonicity of participation ratio MPR定理

- Let C, C' be two co-location patterns such that $C' \subset C$

- Then, for each feature $f \in C', pr(C', f) \geq pr(C, f)$

- An Apriori-like algorithm can be derived for efficient mining colocation patterns

- Ex: Let min-feature-support = σ , min-pr = ρ

- Start with a set of single feature pattern $\{p_1\}$ with support $\geq \sigma$

- Grow to size k , in Apriori way (i.e., stop growing if the pattern is infrequent)

- For each such p , mine its super-pattern P , s.t., $pr(P, p) \geq \rho$, in Apriori way

$p \subset P$

MPR定理

