Mining various kinds of Association Rules

T.R.Lekhaa

AP/IT

Various kinds of Association Rules

- Multi level association rules

  Involves concept at different levels of abstraction.

- Multidimensional association rules

  involves more than one dimensions or predicate.

- Quantitative association rules

  involves numeric attributes that have an implicit ordering among values.

Multilevel association rules

- It is difficult to find strong associations among data items at low or primitive levels of abstraction due to the sparsity of data at those levels.
- Strong associations discovered at high levels of abstraction may represent commonsense knowledge.
- Data mining systems should provide capabilities for mining association rules at multiple levels of abstraction, with sufficient flexibility for easy traversal among different abstraction spaces.

Contd…

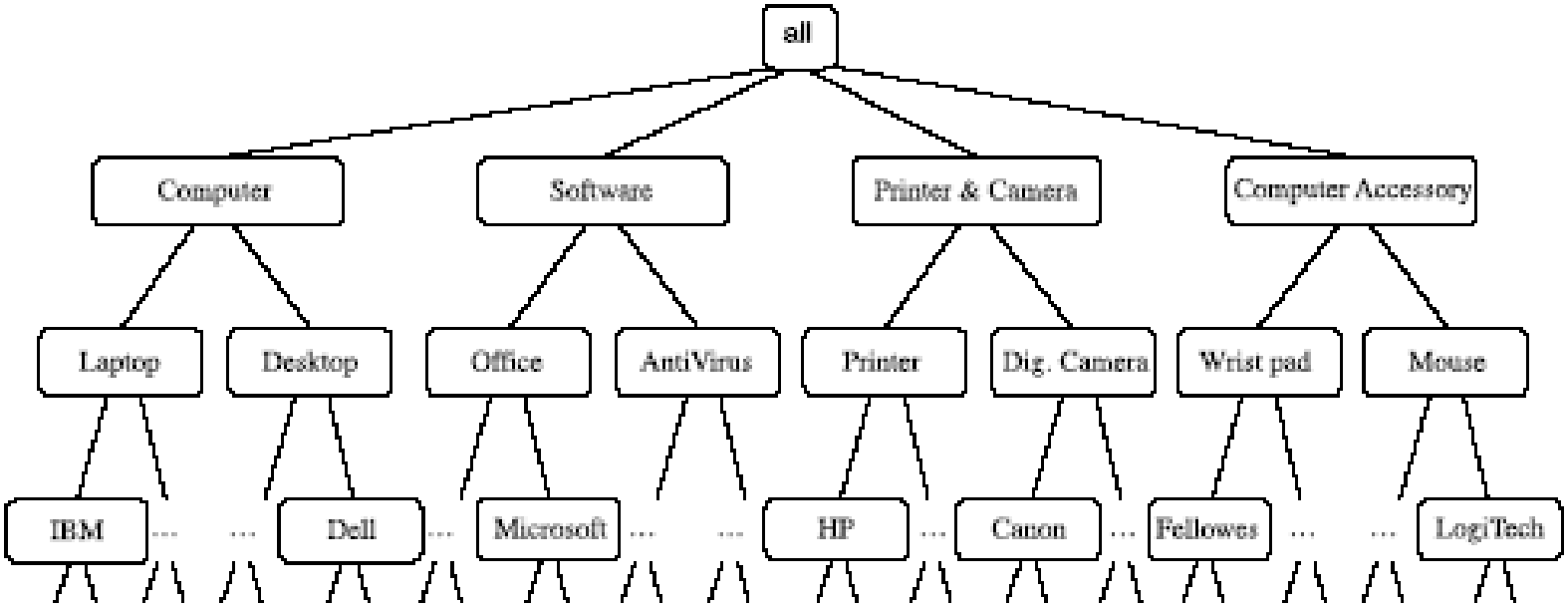
- Association rules generated from mining data at multiple levels of abstraction are called multiple-level or multilevel association rules.
- Multilevel association rules can be mined efficiently using concept hierarchies under a support-confidence framework.
- A top-down strategy is employed, starting at the concept level 1 and working downward in the hierarchy toward the more specific concept levels, until no more frequent item sets can be found.
- For each level, any algorithm for discovering frequent item sets may be used, such as Apriori or its variations.

Contd…

- A concept hierarchy defines a sequence of mappings from a set of low-level concepts to higher level, more general concepts.
- Data can be generalized by replacing low-level concepts within the data by their higher-level concepts, or *ancestors*, from a concept hierarchy.

Contd…

- The concept hierarchy for the items is shown in Figure

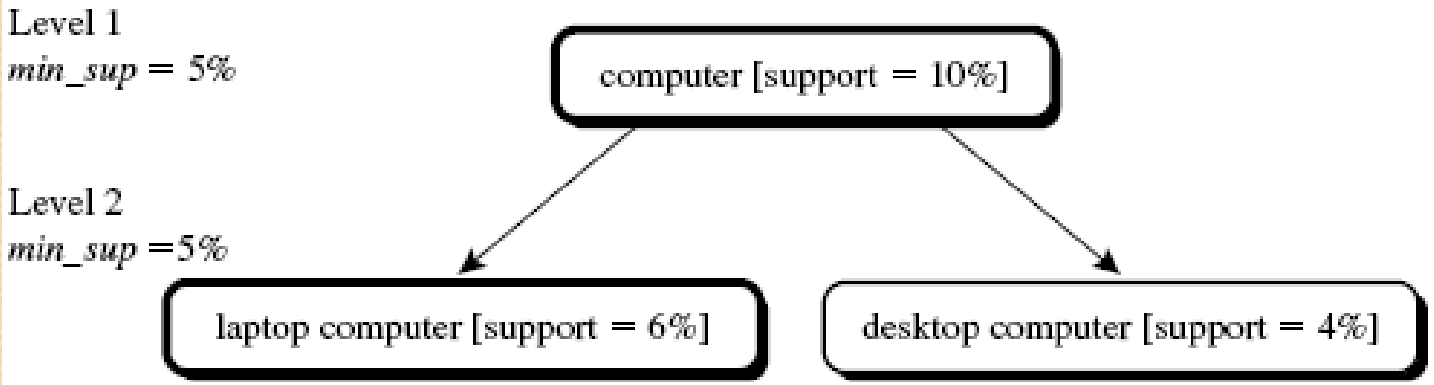


Contd…

- The given concept hierarchy has 5 levels, referred to as levels 0 to level 4.
- Starting with level 0 at the root node for all.
- Here, level 1 includes computer, software, printer and camera so on….
- Level 2 includes laptop computer, desktop computer and so on….
- Level 3 includes IBM desktop computers….
- Level 4 is the most specific abstraction level of this hierarchy which includes raw data.

Contd…

- **Using uniform minimum support for all levels (referred to as uniform support):** The same minimum support threshold is used when mining at each level of abstraction.



Multilevel mining with uniform support.

Contd…

- For example, in the given Figure , a minimum support threshold of 5% is used throughout (e.g., for mining from *"computer"* down to *"laptop computer"*). Both *"computer"* and *"laptop computer"* are found to be frequent, while *"desktop computer"* is not.
- When a uniform minimum support threshold is used, the search procedure is simplified.

Contd…

- **Using reduced minimum support at lower levels (referred to as reduced support):** Each level of abstraction has its own minimum support threshold. The deeper the level of abstraction, the smaller the corresponding threshold is.
- For example, in the given Figure, the minimum support thresholds for levels 1 and 2 are 5% and 3%, respectively. In this way, *"computer," "laptop computer,"* and *"desktop computer"* are all considered frequent.

Contd…

- **Using item or group-based minimum support (referred to as group-based support):** Because users or experts often have insight as to which groups are more important than others, it is sometimes more desirable to set up user-specific, item, or group based minimal support thresholds when mining multilevel rules.

Thank You…