

## Textual Frequent Pattern Mining:

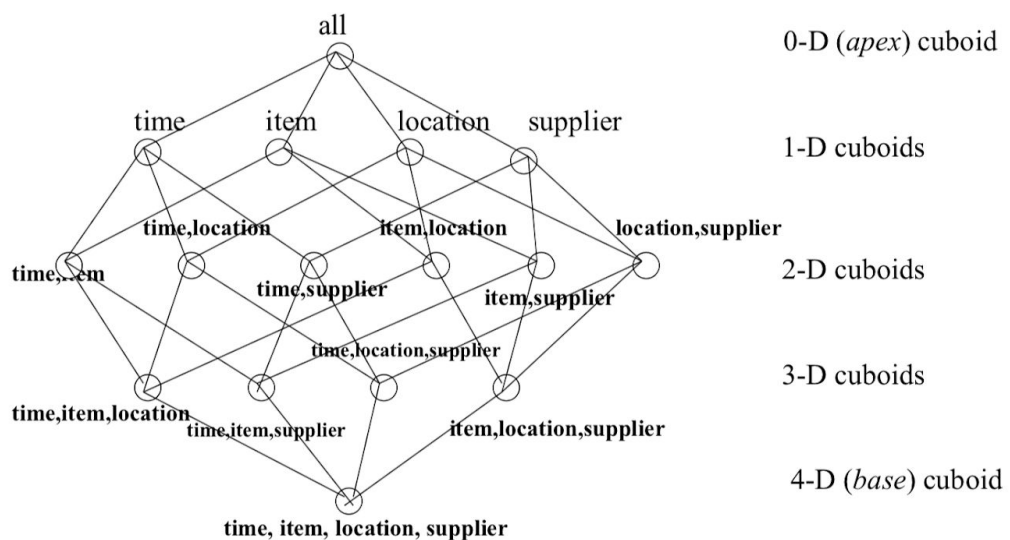
### 1. Data cube:

Data cube, a lattice of cuboids, is a multidimensional (n-d) array of data.

An n-d based cube is the base cuboid.

In computer programming contexts, a data cube (or datacube) is a [multi-dimensional \("n-D"\) array](#) of values.

## Data Cube: A Lattice of Cuboids



Data cube can be helpful for **OLAP** (Online Analytical Processing) in data warehousing.

Related concepts:

RDBMS (Relational database management system): Spreadsheet

Unrelational: {Attri1: value1, }: MongoDB, Spark, NoSQL,

### 2. Clustering based pattern detection.

Top-down: easier

Bottom-up: hard & not recommended.

### 3. Textual frequent pattern application

- 3 ways of definition on definition.
- 3 characters for each candidate/ transaction.
- Potential application of frequent pattern (phases)

#### 4. Basic tasks in textual frequent mining.

1. Select context units and design a strength weight for each unit to model the contexts of frequent patterns.
2. Design similarity measures for the contexts of two patterns, and for a transaction and a pattern context.
3. For a given frequent pattern, extract the most significant context indicators, representative transactions and semantically similar patterns to construct a structured annotation.

Ref:

<https://docs.google.com/document/d/1OCs-yDqr12Rn14ZKDZw1rnS7crY-Y8tdgUzjMbynUql/edit?usp=sharing>

1. 😊
2. 找书 (Chap8.3)