

The background features a complex, abstract design. It includes a network of red lines connecting green dots, resembling a graph or a data structure. There are also various geometric shapes, such as triangles and polygons, in shades of purple, blue, and orange. A white banner with a black border runs across the middle of the image, containing the title text. The overall aesthetic is modern and technical.

SPADE: Sequential Pattern Mining in Vertical Data Format

Sequential Pattern Mining in Vertical Data Format: The SPADE Algorithm

- ❑ A sequence database is mapped to: <SID, EID>
- ❑ Grow the subsequences (patterns) one item at a time by Apriori candidate generation

SID	Sequence
1	<a(<u>abc</u>)(a <u>c</u>)d(cf)>
2	<(ad)c(bc)(ae)>
3	<(ef)(<u>ab</u>)(df) <u>cb</u> >
4	<eg(af)cbc>

min_sup = 2

Ref: SPADE (Sequential
Pattern Discovery
using Equivalent Class)
[M. Zaki 2001]

SID	EID	Items
1	1	a
1	2	abc
1	3	ac
1	4	d
1	5	cf
2	1	ad
2	2	c
2	3	bc
2	4	ae
3	1	ef
3	2	ab
3	3	df
3	4	c
3	5	b
4	1	e
4	2	g
4	3	af
4	4	c
4	5	b
4	6	c

a		b		...
SID	EID	SID	EID	...
1	1	1	2	
1	2	2	3	
1	3	3	2	
2	1	3	5	
2	4	4	5	
3	2			
4	3			

ab			ba			...
SID	EID (a)	EID(b)	SID	EID (b)	EID(a)	...
1	1	2	1	2	3	
2	1	3	2	3	4	
3	2	5				
4	3	5				

aba				...
SID	EID (a)	EID(b)	EID(a)	...
1	1	2	3	
2	1	3	4	