Toward mining of spatiotemporal maximal frequent patterns

Luboš Popelínský and Jan Blaťák
Knowledge Discovery Lab, Faculty of Informatics
Masaryk University in Brno
{popel,xblatak}@fi.muni.cz

spatiotemporal logics

inductive inference in a spatiotemporal logic

mining of first-order maximal frequent patterns in spatiotemporal logic

Example 1:

temporal logic, $\Box$ (always in future), $\Diamond$ (sometimes in future)

windstorms data
PLACE, TIME, TYPE (characteristics of the windstorm), DAMAGE (no, moderate, strong), CHARACTERISTICS (a kind of damage)

"after a wind K, in the period 1971-72 all the winds was strong"

year(K,C),$1971 \leq C \leq 1972$,
$\Box$(K,K1,type(strong)&char(house_destruction))

adding the literal $\Diamond$(K,K2,X) only if X refines the term type(strong) & char(house_destruction).

Example 2:

For **spatial predicates** po(X,Y) (X overlaps Y) and tpp(X,Y) (X is tangential partial part of Y) an appearance of one of them prevents from use of the other.

Table of contents

Data model

spatiotemporal data = a sequence of events

event has a unique identifier, is connected with an explicit time instant

there is no limit on the number of events with the equal time stamp

At least one attribute has to be spatial

attributes of complex type

domain knowledge = a logic program (sequence of predicate definitions)

spatiotemporal pattern = conjunction of non-spatiotemporal atoms and at least one spatiotemporal atom

spatiotemporal atom

temporal - $\Diamond(X)$, $\bigcirc^n(X)$ or $\Box(X)$ –

spatial atom - e.g. from *RCC-8*

Mining spatiotemporal maximal frequent patterns

finding, for a given M (M is usually called a *minimal support*),

all frequent spatiotemporal patterns, i.e. those that cover at least M examples,

and that cannot be further refined without decreasing support under M.

GRAPE

extension of RAP a system for mining of first-order maximal frequent patterns that employs different search strategies for mining long patterns

The downward refinement operator in RAP:

1. Add a most general literal into the pattern;
2. Bind two distinct variables of the same type;
3. Split the range of a numeric variable.

GRAPE: new refinement operations:

for descending in user-defined is-a hierarchy (not necessarily temporal or spatial)

for specialization of temporal formulas.

Minimal support in GRAPE can be defined as global (ignoring granularity level), separately for each level in a hierarchy (like in $SPADA$) or as a user defined predicate.

Refinement operator for spatiotemporal logic

Is-a hierarchy

for each non-temporal attribute there is maximally one hierarchy

Temporal predicates

5. $\rho(P) = (P, \Diamond(X))$ where X is a new variable and there is no other temporal predicate in P with a free variable (i.e. unused in P).
6. $\rho((P, \Diamond(T), S)) = (P, \Box(T), S)$ if there is no term T_1 θ-equivalent (in terms of θ-subsumption) with T in the rest of the pattern.
7. $\rho((P, \Diamond(T), S)) = (P, \bigcirc^n(T), S)$ if there is no term T_1 θ-equivalent with T in the rest of the pattern.
8. $\rho(P) = (P, \Diamond(T_1))$ if P contains $\Box(T)$, where T, T_1 are terms and T_1 is a proper specialization of T and T_1 does not appear elsewhere in the pattern.
9. $\rho(\star(X)) = \star(\rho(X))$ for $\star \in \{\Diamond, \bigcirc^k, \Box\}$

Experiments

looking for **emerging patterns**

for classified data, a pattern is *emerging* if coverage on different classes differs significantly

here: if difference between maximal and minimal coverage is greater or equal to 60%.

frequent patterns as new features

Example: Keystroke dynamics data

a set of keystroke sequences, together 14483 records

Six persons (described with AGE, SEX, LEVEL of writing), have written repeatedly three different texts.

keystroke record = TIME-STAMP (the moment of the event), TYPE (release or press) and CODE of the key pressed or released

For each key there are coordinates (its layout on the keyboard) and also its membership into spatial hierarchy on keys – FINGER-TO-WRITE (left thumb, right thumb, left forefinger, right forefinger etc.), and HAND-TO-WRITE (left, right).

class attribute = LEVEL of a user - `NON-EXPERIENCED, ADVANCED`

Example 1

$$\Diamond(\bigcirc(press(x_1) \wedge i(x_1,'v') \wedge \Diamond(press(x_2) \wedge I(x_2,'backspace')$$

"it is always true that, as the second event, the key 'v' was pressed and always in future the key 'bspace' was pressed"

frequent for non-experienced users

Example 2

d(Key1,Key2,Delay) = delay between press of two keys Key1 and Key2.

$$\Diamond(\bigcirc(P(x_1) \wedge I(x_1,'v') \wedge \Diamond(P(x_2) \wedge I(x_2,'h') \wedge$$
$$\bigcirc(P(x_3) \wedge I(x_3,'a') \wedge d(x_2, x_3, z) \wedge 162 \leq z \wedge z \leq 191))))$$

"always the second key was 'v' and after there were always the sequence of 'h' and 'a' and the delay between pressing these two keys was in the interval [162,191]"

frequent for advanced users.

Frequent patterns as new features

windstorm data

Naive Bayes, decision tree classifier J48 and support vector machines SMO

10-cross validation

	Naive Bayes	J4.8	SMO
original	58.8	80.2	79.7
only frequent	81.3	81.1	84.1
orig+frequent	73.1	82.1	84.1
orig+max	64.3	76.2	84.1

use of frequent patterns results in an accuracy increase

adding original attributes has not affected accuracy significantly

Future work

is the refinement complete for ST_0 logic?

refinement operators for ST_1 and ST_2 logics?

mining in annotated texts, e.g. for morphlogical disambiguation