

Mining Web Usage Patterns

Tomáš Kramár

supervisor Ing. Michal Barla

S T U . .

 F I I T .

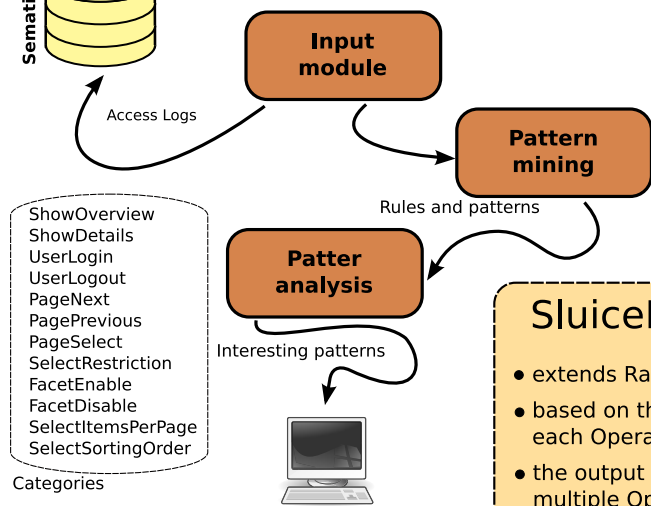
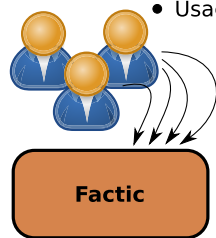
Goal

- Detecting the behavior model
- Services like Google Analytics provide only statistical overview - what was clicked and how often
- Usage patterns capture the users' behavior

/about.html -> /contact.html [sup=0.5, conf=0.6]

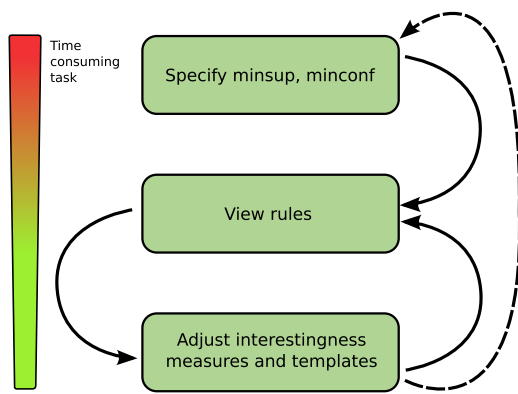
50% visitors clicked on /about.html and /contact.html together and those who clicked /about.html also clicked /contact.html 60% of the time

Transactions				about + contact	about -> contact	contact -> about
/about.html	/contact.html	/search.htm	/services.html	✓	✓	✓
/index.html	/about.html	/contact.html	/products.html	✓	✓	✓
/index.html	/about.html	/services.html			✗	
/about.html	/index.html				✗	
/index.html	/products.html	/contact.html				✗
/index.html	/about.html	/contact.html		✓	✓	✓
Summary:				3/6	3/5	3/4



- ### Concepts
- a Association rules
 - b Sequential rules
 - c Multiple minimum support
 - d Class association rules
 - e Interestingness measures
 - f Templates

- ### SluiceBox
- extends RapidMiner
 - based on the Operator concept each Operator represents exactly one task
 - the output is produced by chaining multiple Operators



- ShowOverview
 ShowDetails
 UserLogin
 UserLogout
 PageNext
 PagePrevious
 PageSelect
 SelectRestriction
 FacetEnable
 FacetDisable
 SelectItemsPerPage
 SelectSortingOrder
- Categories

Input processing

Parse URI

a `http://nazou.fiit.stuba.sk/nazou/ontologies/v0.6.17/region#World`
=> restriction: region#World

Replace items with categories

b `offer-job-inst#Offer_jakubik_f650b3e70bd54d92df0a6edf5958756b`
=> showOfferDetails

Filter items

c `login`

Resolving categories

Ask the input module

Prefix items with categories, in a form:

category:subcategory#item
restriction:offer-job#bCar

- ✓ faster
- ✓ more informative items
- ✓ input from SemanticLog already has the correct form

Template filtering

Types:

- inclusive
- restrictive

Each item must fall into one of the categories.

Examples:

Inclusive: Any* → FacetDisable
Restrictive: SelectRestriction, Any* → Any

Root Process

- DatabaseExampleSource** Load transactions as comma delimited data
- DelimitedStringExploder** Explode the delimited data and create RapidMiner Examples
- DefaultMISMapper** Map *minsup* to each item
- MSApriori** Generate frequent itemsets
- MISAssociationRulesGenerator** Generate patterns from the itemsets
- RulesWriter** Write the patterns to the database