Relevance Feedback on Association Rules

Finding the most interesting rules
[additional slides]

Georg Ruß

Otto-von-Guericke Universität Magdeburg

British Telecommunications plc

28 September 2006

Additional slide - parameter overview

parameter	constraints	impact
δ	$\in [0, 1]$ exponential/linear	decay factor of relevance decisions, "memory" of the relevance feedback
OWA	weight vector W	emphasise certain similarities, aggregation of similarities
$\omega_1 \ldots \omega_6$	$\sum_i \omega_i = 1$ and $\omega_i \geq 0$	allocation of interesting feature combinations
$w_{\mathrm{rel}}, w_{\mathrm{nrel}}$	$w_{\mathrm{rel}} + w_{\mathrm{nrel}} = 1$ and $w_{\mathrm{rel}}, w_{\mathrm{nrel}} \geq 0$	balance between the relevance and non-relevance decisions

- Relevance selection:

5	TEC=ADSL, AGE=36−50	WCC=NO			
6	AGE=51−65, VOL=10GB	SAT=VSAT			
7	AGE=66+	VOL=10GB			

select as: relevant / non−relevant

- Collected rule sets

relevant rules	
AGE=51−65, VOL=10GB AGE=66+	SAT=VSAT VOL=10GB
non−relevant rules	
TEC=ADSL, AGE=36−50	WCC=NO

- Parameter settings for expert/average user

decay parameters

0.5 factor ○ exp ● lin

OWA operator parameters

○ min ○ max ○ average ● other

interesting combinations

| 30% | 30% | 0 | 0 | 20% | 20% |

balance between rule sets

0.8 relevant 0.2 non−relevant

validity of relevance decisions

short long

similarity influence

similar features dissimilar features

interesting combination

1 6

balance between rule sets

relevant non−relevant

(a) expert user

(b) average user

Additional slide - OWA operator [1/2]

1	TEC=ADSL	WCC=YES	stable	stable
3	AGE=66+	VOL=NONE	down	down
5	TEC=ADSL, AGE=36-50	WCC=NO	stable	stable
7	AGE=66+	VOL=10GB	up	up

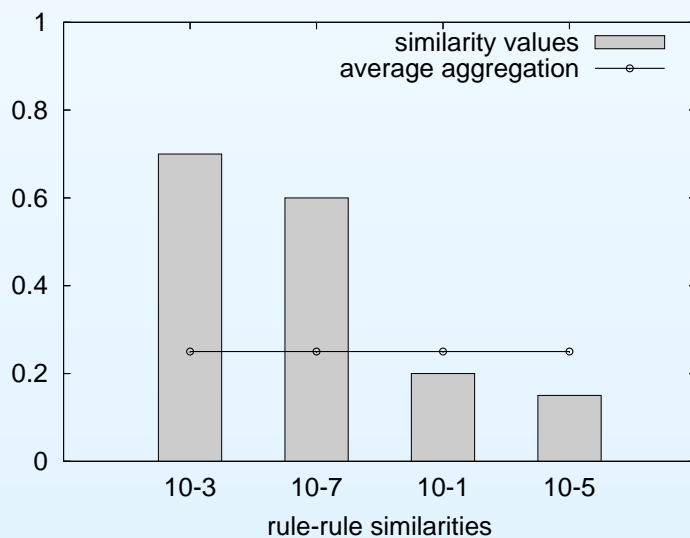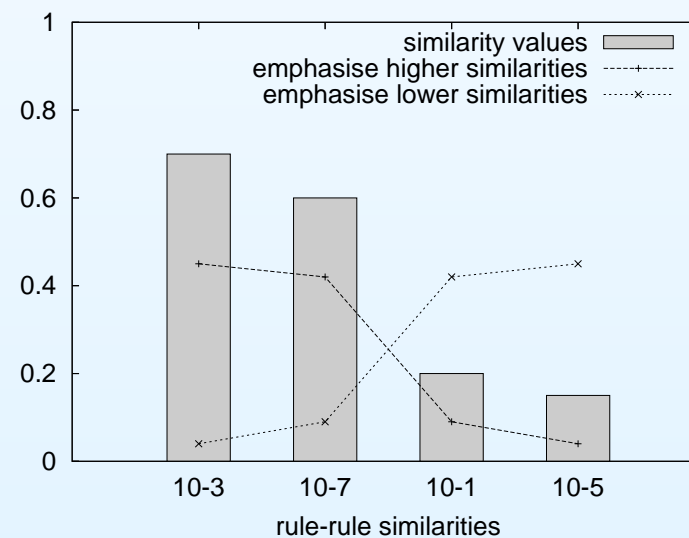| 10	AGE=66+, TEC=CABLE	SAT=SAT	stable	stable
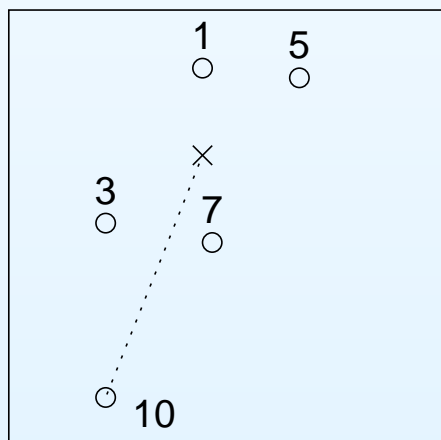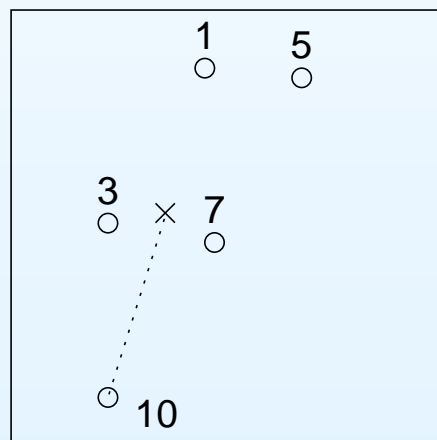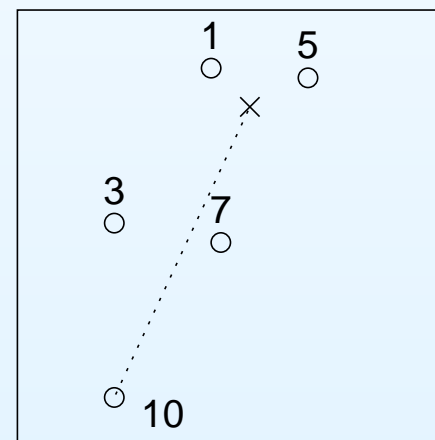
(a)   (b)   (c)

Additional slide - OWA operator [2/2]

- Set of similarities is calculated
- Similarity values are ordered
- Weights are applied
- Weighted similarities are accumulated

(a)

(b)

Additional slide - Initial ranking

- Initial ranking for relevance feedback required

- Use existing orders on the association rules (supp, conf, …)
  - this approach was used here

- Pre-aggregate rules, e.g., cluster them based on similarity of symbolic or time series
  - Likely to be computationally heavy
  - Can be added as an additional step, due to the feature vector representation