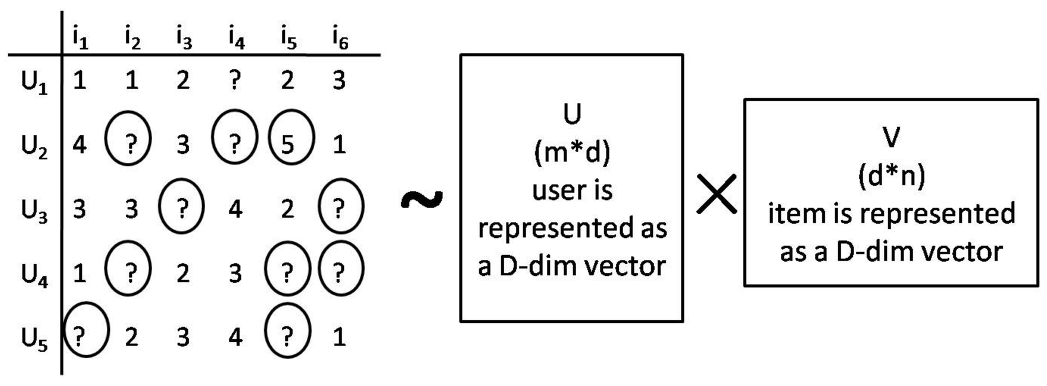
Modeling Retail Transaction Data for Personalized Shopping Recommendation

Pengfei Wang, Jiafeng Guo, Yanyan Lan, Xueqi Cheng
CAS-NDST Lab, Institute of Computing Technology, Chinese Academy of Sciences
wangpengfei@software.ict.ac.cn, {guojiafeng, lanyanyan, cxq}@ict.ac.cn

1.BACKGROUND AND MOTIVATION

Massive transaction data has been routinely recorded in offline retails, which convey rich preference information on brands and goods from customers. Personalized Recommendation based on these valuable transaction data is a critical task.



collaborative filtering

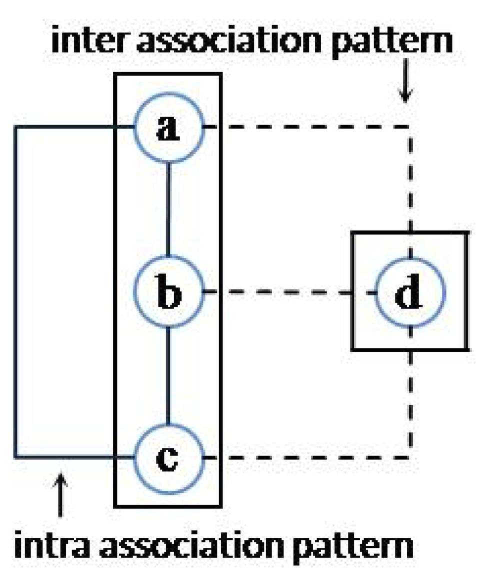
{onions, potatoes} ➡ {burger}
{bread, butter} ➡ {milk}
{beer} ➡ {napkin}
{tennis shoes} ➡ {tennis}
...

association rules

Collaborative filtering directly models the user-item matrix ,ignoring transaction information.
Association rule mining discovers valuable relations among products, but lacks personalization which is important to customers.

2. DEFINITION

Inspired by association rules, we introduce association patterns as basic units to capture the correlation between products



Definition 1 *Inter Association Pattern(Weak Relation)* *Given the transaction set* $T^u = \{t_1^u, t_2^u, \ldots, t_{|T|}^u\}$ *of user* u, *where* $|T|$ *is the count of transactions belonging to* u, *an inter association pattern is defined as a weighted pair of products* $< I_i, I_j, w_{ij} >$, *where* $I_i \in t_m^u$, $I_j \in t_n^u$, $m \neq n$, *and* w_{ij} *denotes the weight of the pattern.*

Definition 2 *Intra Association Pattern(Strong Relation)* *Given the transaction set* $T^u = \{t_1^u, t_2^u, \ldots, t_{|T|}^u\}$ *of user* u, *where* $|T|$ *is the count of transactions belonging to* u, *an intra association pattern is defined as a weighted pair of products* $< I_i, I_j, w_{ij} >$, *where* $I_i, I_j \in t_m^u$ *and* w_{ij} *denotes the weight of the pattern.*

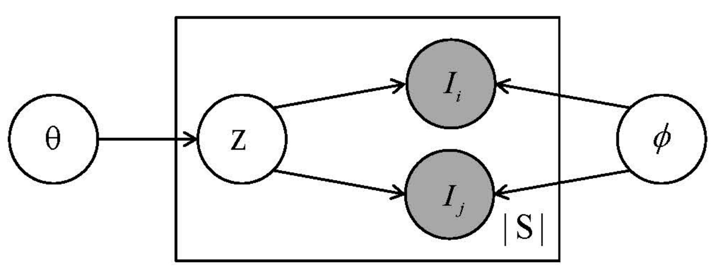
weight definition :

$$w_{ij} = \exp\{-\frac{|time(t_m^u) - time(t_n^u)|}{K}\}$$

3. PAP MODEL

Key idea :reduce association patterns into a low-dimensional shopping interest space, inference the interest of each individual, and then provide personalized shopping recommendations.

Probabilistic model over ***A***ssociation ***P***atterns:



Algorithm 1 The generative process of PAP

1: sample a distribution of shopping interests:
   $\Theta \sim \text{Dirichlet}(\alpha)$
2: for each shopping interest z
   draw a distribution $\Phi_z \sim \text{Dirichlet}(\beta)$
3: for each pattern $< I_i, I_j > \in S$
   draw a latent shopping interest:
   $z \sim \text{Multinomial}(\Theta)$
   draw a pattern $< I_i, I_j > \sim \text{Multinomial}(\Phi_z)$

$$P(< I_i, I_j > |\Theta, \Phi) = \sum_z P(z)P(I_i|z)P(I_j|z) = \sum_k \theta_k \Phi_{k,i} \Phi_{k,j}$$

4. RECOMMENDATION

Inference of User Preference

Probability of the k-th shopping interest:

$$\theta_k^u = P(z = k|u) = \sum_{<I_i, I_j> \in S^u} P(z = k| < I_i, I_j >)P(< I_i, I_j > |u)$$

$$P(z = k| < I_i, I_j >) = \frac{P(< I_i, I_j > |z = k)P(z = k)}{\sum_z P(< I_i, I_j > |z)P(z)} = \frac{P(z = k)P(I_i|z = k)P(I_j|z = k)}{\sum_z P(I_i|z)P(I_j|z)P(z)} = \frac{\theta_k \phi_{k,i} \phi_{k,j}}{\sum_k \theta_k \phi_{k,i} \phi_{k,j}}$$

$$P(< I_i, I_j > |u) = \frac{w_{ij}}{\sum_{<I_i, I_j> \in S^u} w_{ij}}$$

Personalized Recommendation

By sorting the products according to $P(I_i|u)$, we can recommend top-k products to the user:

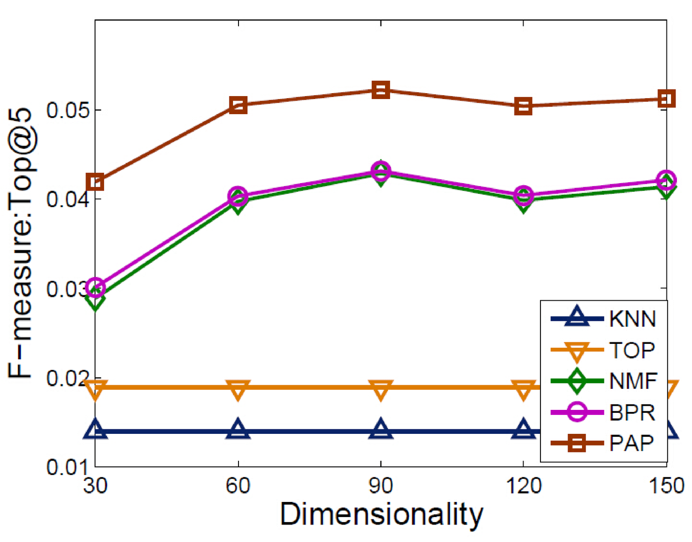
$$P(I_i|u) = \sum_z P(I_i|z)P(z|u) = \sum_k \theta_k^u \phi_{k,i}$$

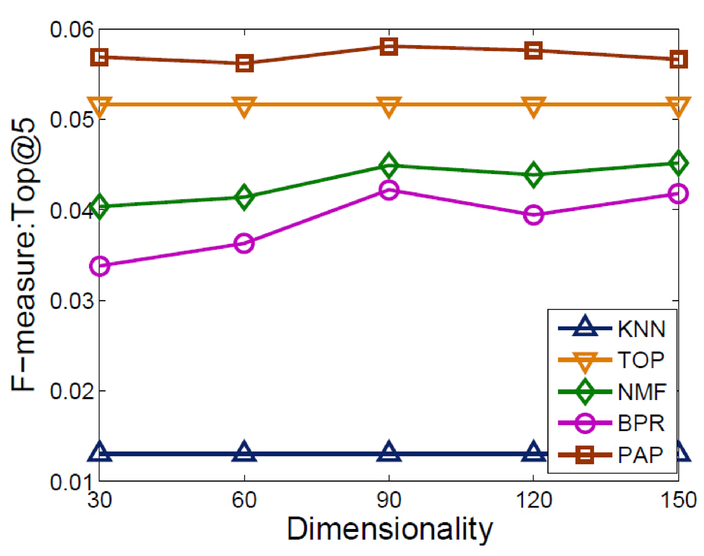
5. EXPERIMENT

1.Statistics of datasets:

id	name	# users	# products	# transactions
1	*BeiRen*	18315	1442	242894
2	*Tafeng*	7141	6894	37269

2.Results on *BeiRen and Tafeng*:





Two datasets show that our *PAP* model perform better than the state-of-the-art recommendation methods.

6. CONCLUSION

Contributions

- We introduce association patterns as basic units to capture the correlation between products.
- We proposed a novel ***P***robabilistic model over the ***A***ssociation ***P***atterns for personalized recommendation.
- Experiments showed that our method outperformed than state-of-the-art recommendation methods.

Future Work

- In the future we will try to explore long-term and short-term association patterns , and analyze the impact of two patterns to personalized recommendation.