Customized products recommendation based on probabilistic relevance model

(Yue and Mitchell, 2012)

Li. 2019.10.26

How can we utilize the customer and product data?

- **Product configuration optimization or Product Family Design**. optimizing the objection function (cost) under the time or assemblies constraints.

- **Recommendation**. recommending products or design scheme for customer or items historical data.

Product Family Design Review

- **PF:** A product family is a set of products that share common components/ productions, while each variant has its unique specifications to meet demands of certain customers (Meyer and Lehnerd, 1997);
- Product family design can be divided into three stages: product definition; product design and process design; supply chain design

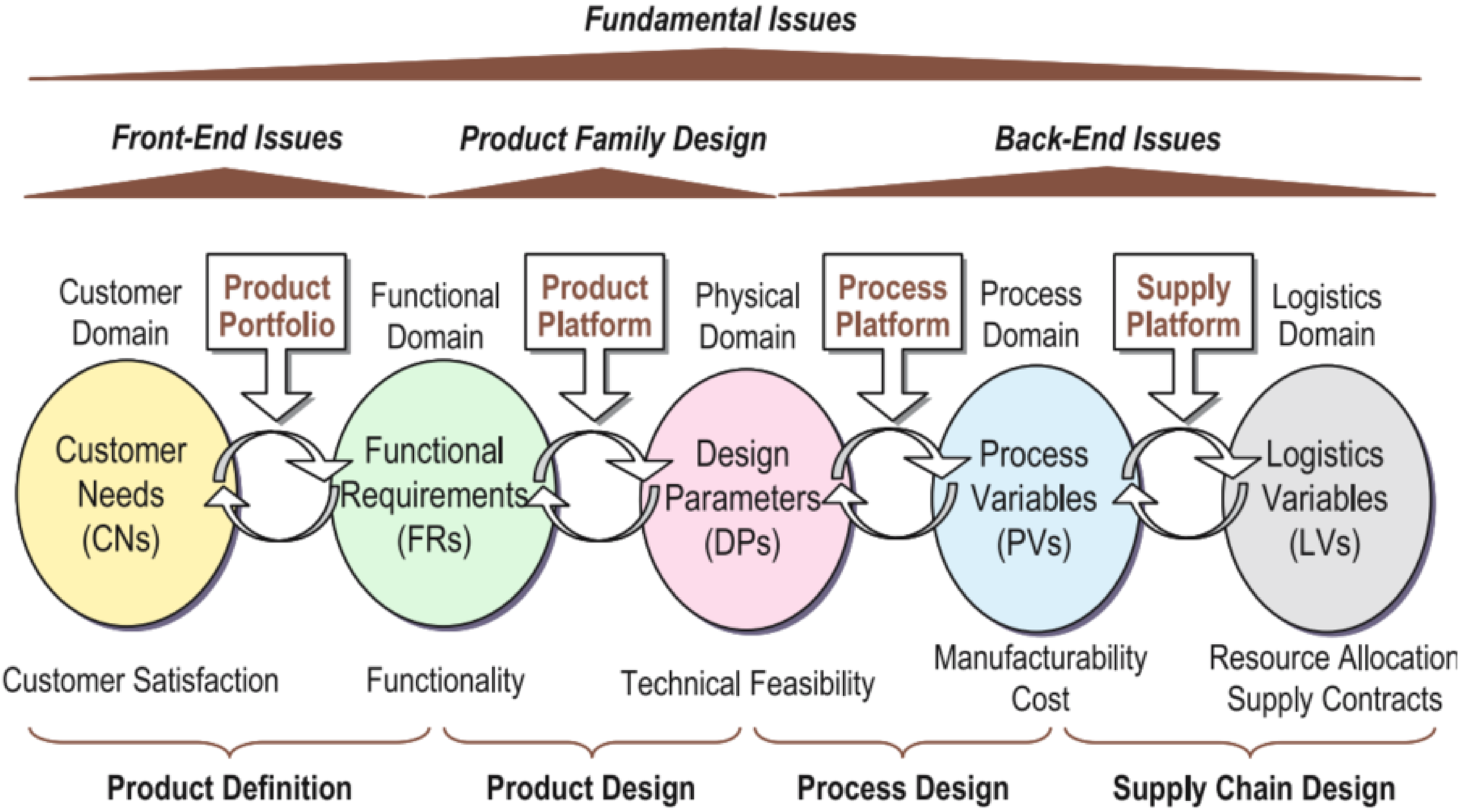
(1) **Product definition**: translate identified customer needs into functional requirements for a product;

(2) **Product design and process design**: mapping those requirements into proper design variables, subject to potential manufacturing constraints;

(3) **Supply chain design and logistics**: process planning and determining process variables.

(Zhila and Gary, 2014; Roger and Timothy, 2007);

Product Family Design Review

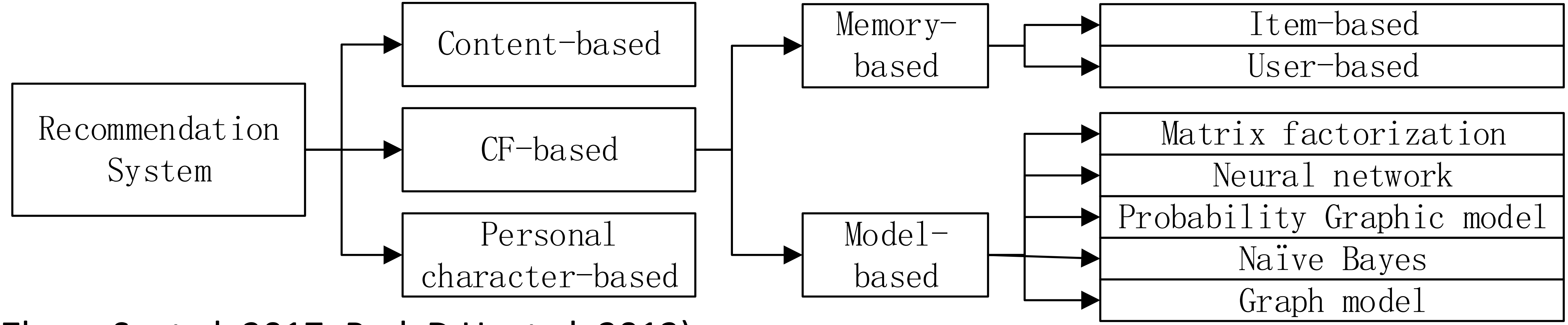


Recommendation System

What's the recommendation?

Generally speaking, recommendation can be regarded as a special kind of information filter to alleviate the information overload problem. Recommendation system trying to meet a customer's requirements. Comparing to the information retrieval, recommendation system is an active process, which focuses more on the exploration and exploitation of users' POI (point of interest).

The basic methods and models:



(Zhang S, et al. 2017; Park D H, et al. 2012)

2B Recommendation & Customized products recommendation

- The limited availability of data – especially for new clients with no previous business interaction;
- Closely related to industry background, different domains will process their own specific features, relying on specialist knowledge;
- Clearly purpose, paying more attention to accuracy and explanation;
- Recommendation diversity. The targeted user of 2B recommendation will be a community due to the complex organization structure, rather than a person. Thus the recommendation will be more various and abundant.
- Time interval or time span will be long, the historical POI will be changed dramatically.
- So as to the customized products recommendation, we are obliged to extract the latent connection between customer preferences and product specification, based on the users' preference to recommend customized product.

Motivation and Contribution

- On the one hand, the vast number of product variants in product customization process often makes it difficult for customers to make purchase decision;
- On the other hand, the customer needs are ambiguous and diversity, which products to recommend and in what order to present the recommendation;
- **New customers**. Memory based models require users' historical record, while as for new customer there is no prior information about his/her preference, "cold start" problem seriously;

Problem definition and Task

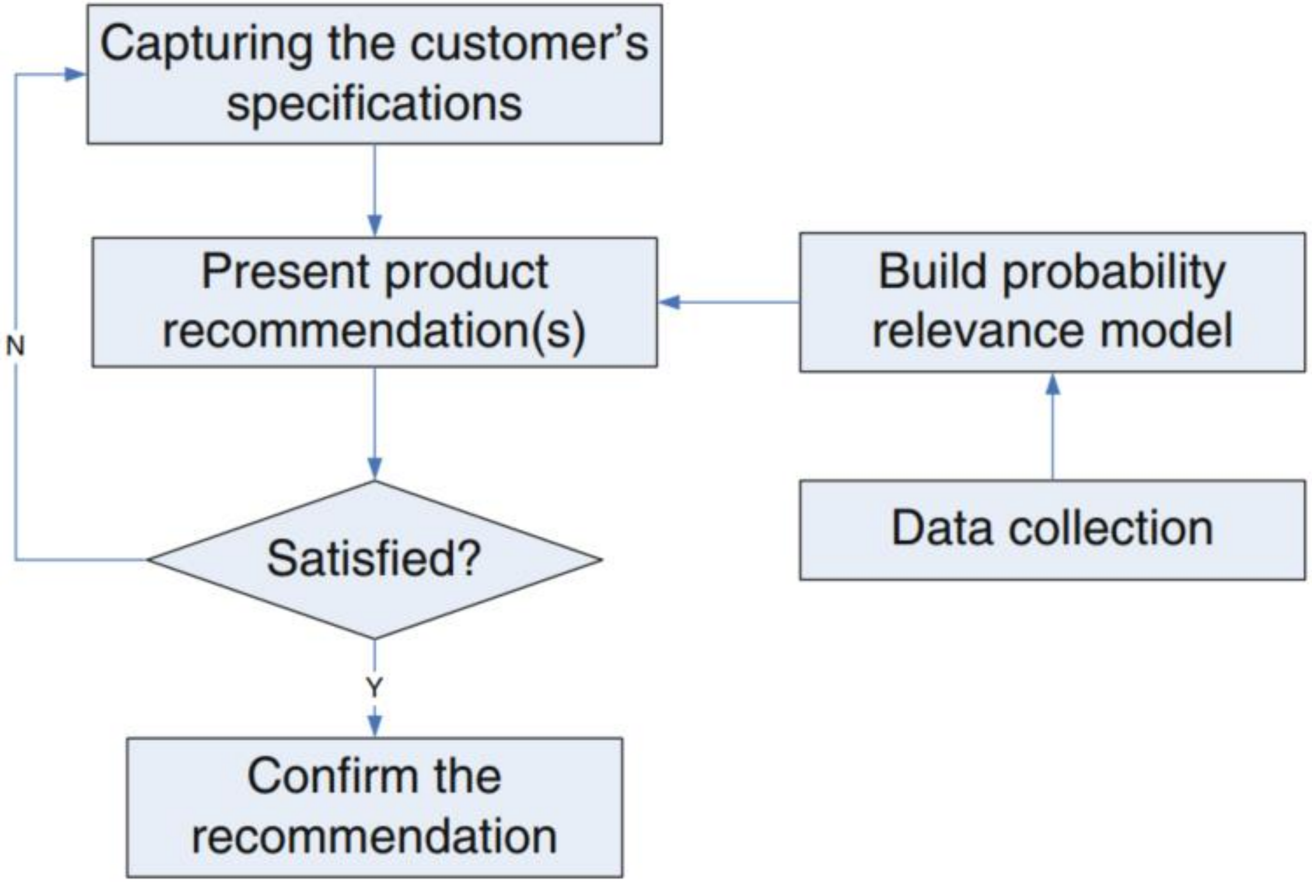
- Table 1 is the components list of a PC, which includes six parts, i.e., a processor, monitor, hard disk, display card, memory and display driver. each component can be alternated or configured.
- According to the limited customer configuration, we have to provide or recommend the complete configuration.
- Maximum likelihood estimation, probability and statistics.

Table 1 List of components and their alternatives for a PC

Component	Code	Description
Processor (A)	A1	Pentium E2160 1.8G
	A2	Pentium E2180 2.0G
	A3	AMD Athlon™ 64 X2
	A4	Intel Core 2 Duo E4300 1.8G
	A5	Intel Core 2 Duo E4500 2.2G
	A6	Intel Core 2 Duo E4600 2.4G
Memory (B)	B1	512 MB DDR2
	B2	1 GB × DDR2
	B3	2 GB × DDR2 dual channel
Monitor (C)	C1	17′ LCD
	C2	19′ LCD
	C3	20′ LCD and above
Hard disk (D)	D1	80 GB
	D2	160 GB
	D3	250 GB
Disk driver (E)	E1	24X CD-RW/DVD* Combo
	E2	48X CD-RW/DVD* Combo
	E3	8X DVD+/−RW*
	E4	16X DVD+/−RW*
Display card (F)	F1	NVIDIA® GeForce® 6150
	F2	Intel® Graphics Media Accelerator 3000
	F3	Intel® Graphics Media Accelerator X3000
	F4	128 MB PCIe™ × 16 ATI Radeon™ X1300
	F5	256 MB PCIe™ × 16 nVidia® GeForce® 7300 LE TurboCache

The framework or pipeline of recommendation in real scenario

- Step1. customer provides a specification;
- Step2. Recommendation System exhibit the complete the configuration, based on the probability relevance model;
- Step3. User confirm whether the recommendation list can meet the needs, if statisfied recommendation over, else turn to Step 4;
- Step 4. customer provides supplementary specification, then turn to Step 2;



Model details

- Supposing we have specification-configuration pair denote as S-C, and let define I as an indicator function with value 1 representing the configuration C will meet the customer's requirements and 0 otherwise. Thus, we want to calculate the value of P(I=1|C,S) for each configuration C. Applying Bayes' rule, we can get:

$$P(I \mid C,S) = \frac{P(C \mid I,S)P(I,S)}{P(C,S)} = \frac{P(C \mid I,S)P(I \mid S)}{P(C \mid S)}$$

- Hence, the denominator does not have a clear meaning and is hard to calculate, then:

$$O(x) = \frac{P(x)}{P(\bar{x})} = \frac{P(x)}{1-P(x)}$$

$$O(I=1 \mid C,S) = \frac{P(I=1 \mid C,S)}{P(I=0 \mid C,S)} = \frac{P(I=1 \mid S)}{P(I=0 \mid S)} \bullet \frac{P(C \mid I=1,S)}{P(C \mid I=0,S)}$$

- The first term in this expression contains no information about a product configuration. Thus it does not affect the recommendation result and can be omitted:

Model details

- By chain rule in probability, we can represent the P(C|I,S) as the product of conditional probabilities:

$$P(C \mid I,S) = P(a_1,\ldots,a_m \mid I,S) = P(a_1,\ldots,a_m \mid I,S)$$

$$= P(a_1 \mid I,S) \bullet P(a_2 \mid I,S,a_1) \ldots P(a_m \mid I,S,a_{1,}\ldots a_{m-1})$$

Where, I is an indicator function, I=1 means this configuration will meet the user's requirements, and 0 otherwise.

- However this is always computationally inefficient and not feasible, we only consider the first order conditional probabilities, then:

$$P(C \mid I,S) = P(a,\ldots,a_m \mid I,S) = \prod_{i=1}^{m} P(a_{i'} \mid I,S,a_{\pi(i^{i'})}), 0 \le \pi(i) < i$$

$$p_{i'} = P(a_{i'} = 1 \mid a_{\pi(i')} = 1, I = 1, S)$$

$$q_{i'} = P(a_{i'} = 1 \mid a_{\pi(i')} = 0, I = 1, S)$$

$$t_{i'} = P(a_{i'} = 0 \mid a_{\pi(i')} = 1, I = 0, S)$$

$$r_{i'} = P(a_{i'} = 1 \mid a_{\pi(i')} = 0, I = 0, S)$$

a_i is an indicator function $a_i = 1$ means the specific component is designated, otherwise is 0; $\pi(i')$ is the first order dependence of i

Model details

- The general expression of $P(a_{i'} \mid a_{\pi(i')}, I=1, S)$ can be stated as:

$$P\left(a_{i'} \middle| a_{\pi(i')}, I = 1, S\right) = [p_i^{a_{i'}}(1-p_{i'})^{1-a_{i'}}]^{a_{\pi(i')}}[q_i^{a_{i'}}(1-q_{i'})^{1-a_{i'}}]^{1-a_{\pi(i')}}$$

- The same:

$$P\left(a_{i'} \middle| a_{\pi(i')}, I = 0, S\right) = [t_i^{a_{i'}}(1-t_{i'})^{1-a_{i'}}]^{a_{\pi(i')}}[r_i^{a_{i'}}(1-r_{i'})^{1-a_{i'}}]^{1-a_{\pi(i')}}$$

- Then we arrive at:

$$\frac{P(C|I=1,S)}{P(C|I=0,S)} = \prod_{i'} \frac{[p_i^{a_{i'}}(1-p_{i'})^{1-a_{i'}}]^{a_{\pi(i')}}[q_i^{a_{i'}}(1-q_{i'})^{1-a_{i'}}]^{1-a_{\pi(i')}}}{[t_i^{a_{i'}}(1-t_{i'})^{1-a_{i'}}]^{a_{\pi(i')}}[r_i^{a_{i'}}(1-r_{i'})^{1-a_{i'}}]^{1-a_{\pi(i')}}}$$

- Take logarithm on both sides of the equation:

$$R = \sum_{i'} \left(a_{i'} log \frac{q_{i'}(1-r_{i'})}{r_{i'}(1-q_{i'})} + a_{\pi(i')} log \frac{(1-p_{i'})(1-r_{i'})}{(1-q_{i'})(1-t_{i'})} + a_{i'} a_{\pi(i')} log \frac{p_{i'} r_{i'}(1-q_{i'})(1-t_{i'})}{q_{i'} t_{i'}(1-p_{i'})(1-r_{i'})} \right) + const$$

- Const is a correction factor.

Model details

- Therefore, the recommendation results are only related to $a_{i'}$ and $a_{\pi(i')}$, counting the corresponding specification-accepted recommendation pairs from the historical records:

I=1

	$a_i = 1$	$a_i = 0$	Sum of row
$a_{\pi(i)} = 1$	m	n	$m + n$
$a_{\pi(i)} = 0$	k	l	$k + l$
Sum of col.	$m + k$	$n + l$	$m + n + k + l$

- Then we have:

$$p_{i'} = P\left(a_{i'} = 1 \middle| a_{\pi(i')} = 1, I = 1, S\right) = \frac{m+\epsilon}{m+n+\varepsilon} \text{ and } q_{i'} = P\left(a_{i'} = 1 \middle| a_{\pi(i')} = 0, I = 1, S\right) = \frac{k+\epsilon}{k+l+\epsilon}$$

- In the same way, when I=0:

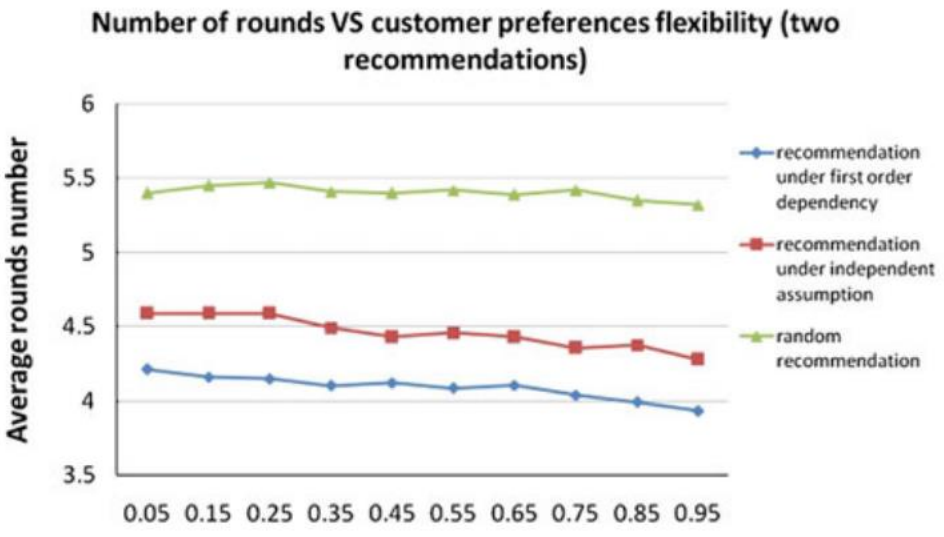
	$a_i = 1$	$a_i = 0$	Sum of row
$a_{\pi(i)} = 1$	m'	n'	$m' + n'$
$a_{\pi(i)}$=0	k'	l'	$k' + l'$
Sum of col.	$m' + k'$	$n' + l'$	$m' + n' + k' + l'$

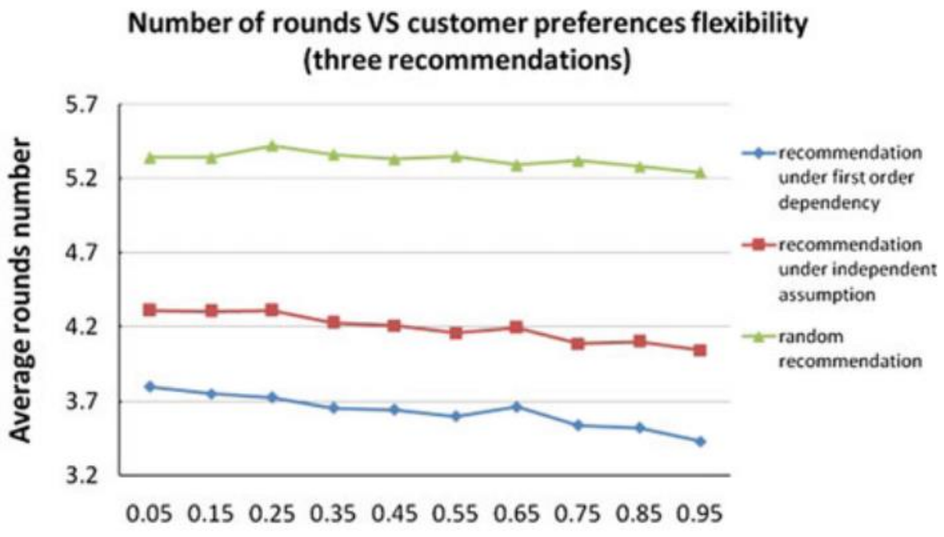
$$t_{i'} = P\left(a_{i'} = 1 \middle| a_{\pi(i')} = 0, I = 1, S\right) = \frac{m'+\epsilon}{m'+n'+\epsilon} \text{ and } r_{i'} = P\left(a_{i'} = 1 \middle| a_{\pi(i')} = 0, I = 0, S\right) = \frac{k'+\epsilon}{k'+l'+\epsilon}$$

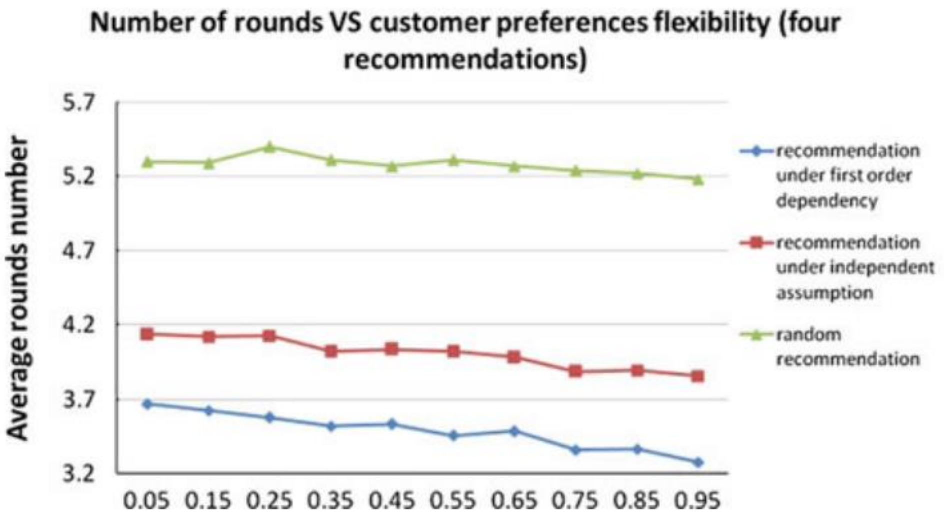
Datasets, evaluation and Experiment setting

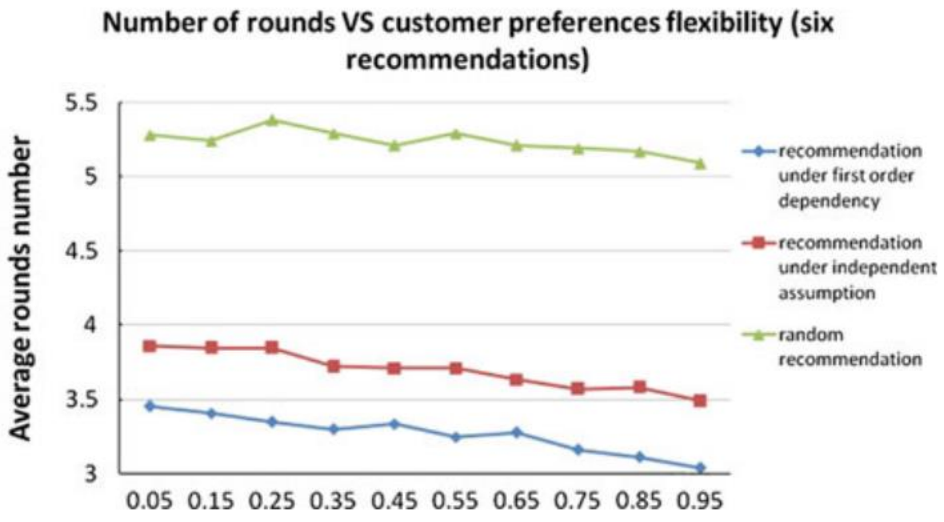
- Collecting 69 sets of customer specification data D1 and the corresponding accept data S1. Then utilizing bootstrap method to generate 1380 specification-recommendation data pairs as the training data and 207 specification-recommendation data pairs as the testing data.
- About the threshold, setting h from 0.05 to 0.95 with step being 0.1 and generate 10 groups of data. If the threshold is bigger, it means the real data and the virtual data will be diversified.
- Choosing top n hit ratio as the evaluation index.
- Random recommendation: each time a customer gives a specification to one component, a random configuration which is consistent with the specification will be proposed.
- Recommendation under independent assumption: there is no dependency among the attributes, thus $p_{i'} = q_{i'}$ and $t_{i'} = r_{i'}$, then $R = \sum_i a_i log \frac{p_i(1-q_i)}{q_i(1-p_i)}$.

Experiment results









Conclusions and Inspirations

- Recommendation under first order dependence has the best performance, and the random recommendation is worst;
- It's safety to say, with the number of recommendations increase, the accuracy will increase gradually;
- The threshold has a scarcely influence to the final results;
- Basically speaking, this paper is to predicted the customer's preference to different configurations based on limited specifications. Hence, the author abstract and reconstruct the problem as a MLE, and only consider the first order conditional probabilities to simplified the calculation, then, counting the historical records to get the final results. Finally, ranking the probability and showing the recommendations.
- Because only related to the customer's specifications, this model can alleviate new customer problem effectively.
- While this method relies on the customer's specification as the input, and the user will not clearly know the literally specification in most situation, there exists a huge gap between customer needs and specifications;
- In addition, in this paper the author didn't consider the correlation between each recommendations, thus there is a sort of redundancy of the whole recommendation list.

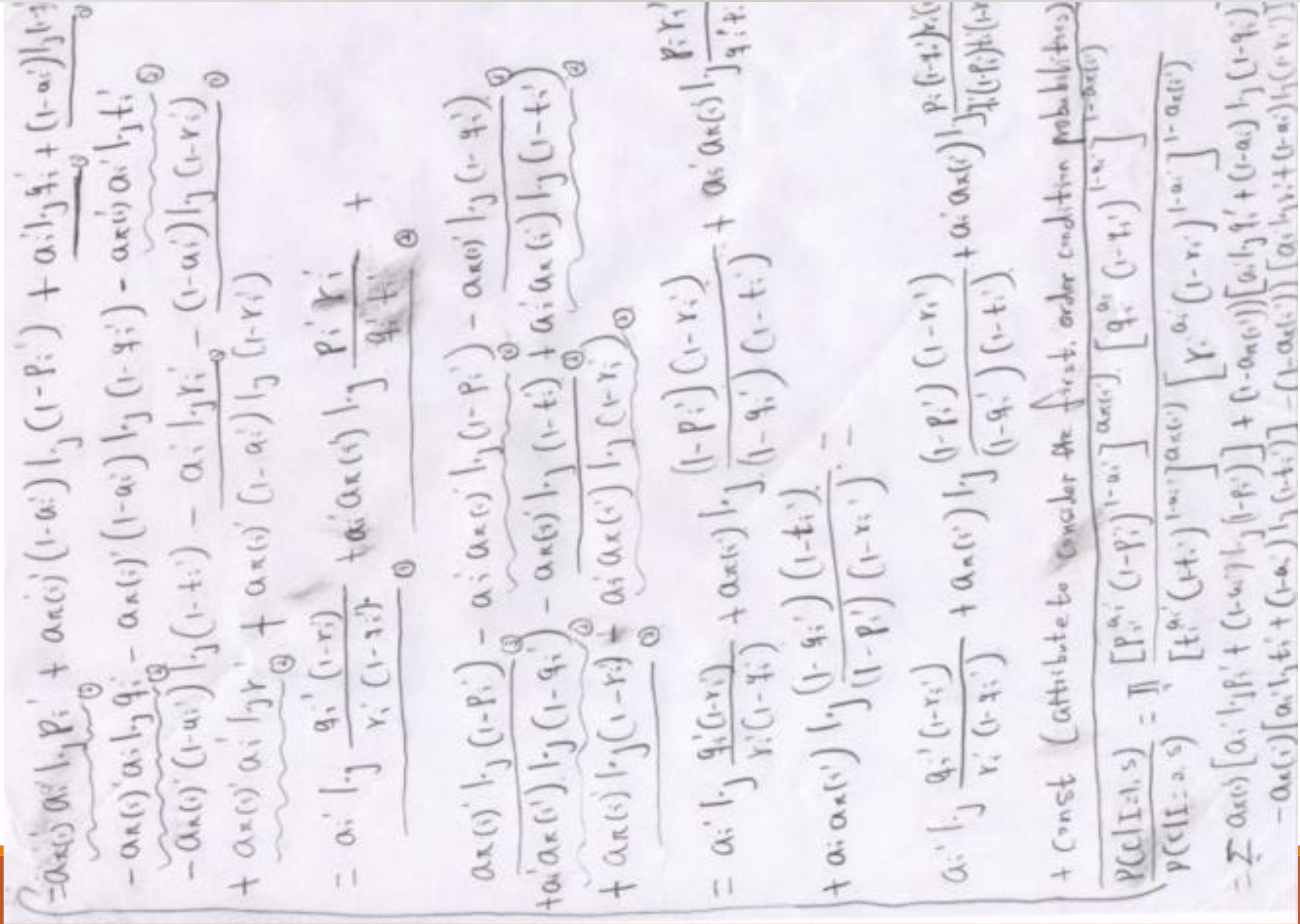
Reference

1. Meyer M, Lehnerd AP (1997) The power of product platform– building value and cost leadship, Free Press: New York.
2. Jiao J R, Simpson T W, Siddique Z. Product family design and platform-based product development: a state-of-the-art review[J]. Journal of intelligent Manufacturing, 2007, 18(1): 5-29.
3. 焦垚. 基于线上数据的客户需求与产品配置智能映射研究[D].重庆大学,2017.
4. Zhang S , Yao L , Sun A , et al. Deep Learning based Recommender System: A Survey and New Perspectives[J]. ACM Computing Surveys, 2017.
5. Park D H, Kim H K, Choi I Y, et al. A literature review and classification of recommender systems research[J]. Expert systems with applications, 2012, 39(11): 10059-10072.
6. Chickering, D. M. (2002). The WinMine toolkit. Microsoft Research Technical Report MSR-TR-2002-103.
7. Wang, Y., & Tseng, M. M. (2009). Recommendation for custom product via probabilistic relevance model. In Proceedings of IEEE conference on engineering management, Hong Kong.

Q&A

Thanks

Appendix



$= a_{x(i)}' a_i' l_{\cdot j} p_i' + a_{x(i)}'(1-a_i') l_{\cdot j}(1-p_i') + a_i' l_{\cdot j} q_i' + (1-a_i') l_{\cdot j}(1-q_i')$

$- a_{x(i)}' a_i' l_{\cdot j} q_i' - a_{x(i)}'(1-a_i') l_{\cdot j}(1-q_i') - a_{x(i)}' a_i' l_{\cdot j} t_i'$

$- a_{x(i)}'(1-a_i') l_{\cdot j}(1-t_i') - a_i' l_{\cdot j} r_i' - (1-a_i') l_{\cdot j}(1-r_i')$

$+ a_{x(i)}' a_i' l_{\cdot j} r_i' + a_{x(i)}'(1-a_i') l_{\cdot j}(1-r_i')$

$= a_i' l_{\cdot j} \frac{q_i'(1-r_i')}{r_i'(1-q_i')} + a_i' a_{x(i)}' l_{\cdot j} \frac{p_i' r_i'}{q_i' t_i'} +$

$a_{x(i)}' l_{\cdot j}(1-p_i') - a_i' a_{x(i)}' l_{\cdot j}(1-p_i') - a_{x(i)}' l_{\cdot j}(1-q_i')$

$+ a_i' a_{x(i)}' l_{\cdot j}(1-q_i') - a_{x(i)}' l_{\cdot j}(1-t_i') + a_i' a_{x(i)}' l_{\cdot j}(1-t_i')$

$+ a_{x(i)}' l_{\cdot j}(1-r_i') - a_i' a_{x(i)}' l_{\cdot j}(1-r_i')$

$= a_i' l_{\cdot j} \frac{q_i'(1-r_i')}{r_i'(1-q_i')} + a_{x(i)}' l_{\cdot j} \frac{(1-p_i')(1-r_i')}{(1-q_i')(1-t_i')} + a_i' a_{x(i)}' l_{\cdot j} \frac{p_i' r_i'}{q_i' t_i'}$

$+ a_i' a_{x(i)}' l_{\cdot j} \frac{(1-q_i')(1-t_i')}{(1-p_i')(1-r_i')} =$

$a_i' l_{\cdot j} \frac{q_i'(1-r_i')}{r_i'(1-q_i')} + a_{x(i)}' l_{\cdot j} \frac{(1-p_i')(1-r_i')}{(1-q_i')(1-t_i')} + a_i' a_{x(i)}' l_{\cdot j} \frac{p_i'(1-q_i') r_i'(1-t_i')}{q_i'(1-p_i') t_i'(1-r_i')}$

+ const (attribute to consider the first order condition probabilities)

$$\frac{P(c|I=1,s)}{P(c|I=2,s)} = \prod_i \frac{[p_i'^{a_i'}(1-p_i')^{1-a_i'}]^{a_{x(i)}'}\,[q_i'^{a_i'}(1-q_i')^{1-a_i'}]^{1-a_{x(i)}'}}{[t_i'^{a_i'}(1-t_i')^{1-a_i'}]^{a_{x(i)}'}\,[r_i'^{a_i'}(1-r_i')^{1-a_i'}]^{1-a_{x(i)}'}}$$

$= \sum_i a_{x(i)}'[a_i' l_{\cdot j} p_i' + (1-a_i') l_{\cdot j}(1-p_i')] + (1-a_{x(i)}')[a_i' l_{\cdot j} q_i' + (1-a_i') l_{\cdot j}(1-q_i')]$

$- a_{x(i)}'[a_i' l_{\cdot j} t_i' + (1-a_i') l_{\cdot j}(1-t_i')] - (1-a_{x(i)}')[a_i' l_{\cdot j} r_i' + (1-a_i') l_{\cdot j}(1-r_i')]$