RECOMMENDATION FOR NEW USERS WITH PARTIAL PREFERENCES BY INTEGRATING PRODUCT REVIEWS WITH STATIC SPECIFICATIONS

Feng Wang, Weike Pan, and Li Chen

Department of Computer Science

Hong Kong Baptist University

MOTIVATION AND OUR IDEA

New user

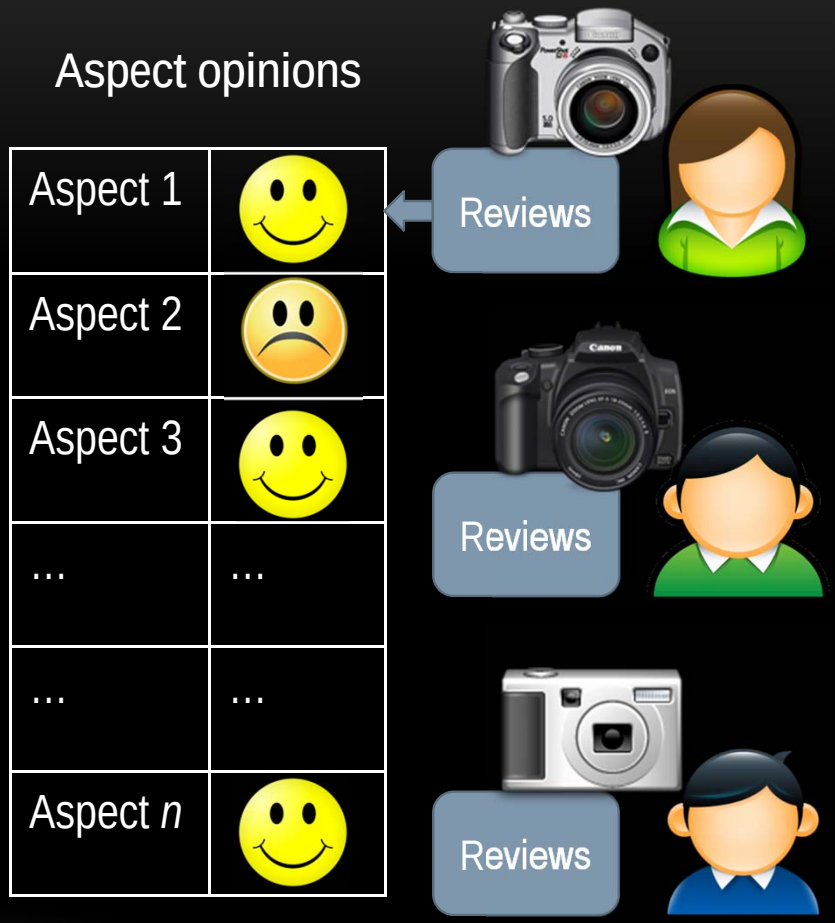
Preference elicitation

Related work

Attributes	Value preferences
Brand	Sony
Sensor resolution	?
Removable flash	?
Thickness	?
Price	<$300
Optical zoom	?
Screen size	?
Weight	<200g

How to estimate the user's preferences on un-stated attributes?

MOTIVATION AND OUR IDEA



Related work: few have fully exploited reviews for helping new users with partial preferences

MOTIVATION AND OUR IDEA

New user

Preferences elicitation

Attributes	Value preferences
Brand	Sony
Sensor resolution	?
Removable flash	?
Thickness	
Price	
Optical zoom	?
Screen size	?
Weight	<200g

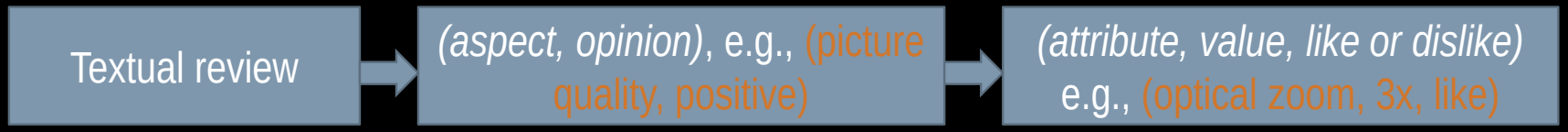
Aspect opinions	
Aspect 1	
Aspect 2	
…	…
Aspect *n*	

To bring reviewers' value preferences on attributes, for helping the new user

How to estimate the user's preferences on un-stated attributes?

RESEARCH QUESTIONS

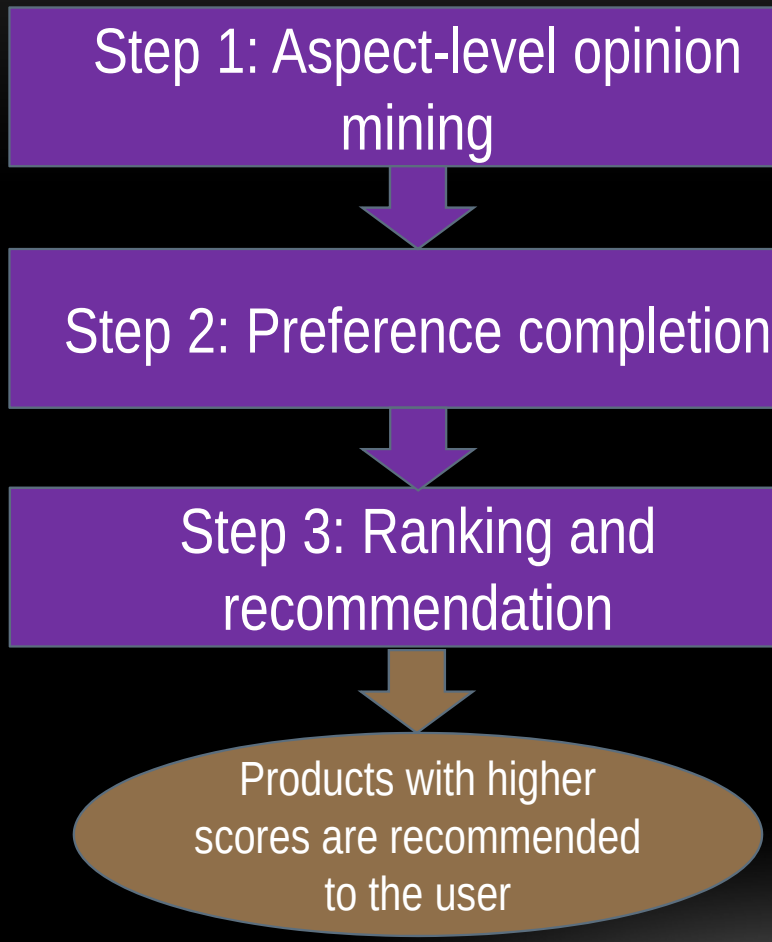
- *How to derive a reviewer's value preferences?*



- *How to incorporate reviewers' value preferences into completing the new user's preferences?*

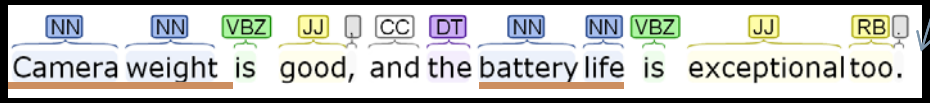
- CompleteRank*:* preference completion and ranking

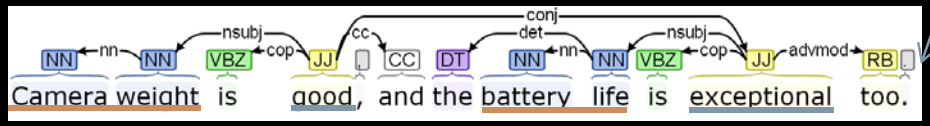
METHODOLOGY - COMPLETERANK



METHODOLOGY

Step 1: Aspect-level opinion mining





Good => *positive*
Exceptional => positive

Extract frequent nouns and noun phrases as aspect candidates via Part-Of-Speech (POS) tagger

Identify opinion words via syntactic dependency parser

Group synonymous aspects via WordNet

Identify the opinion's sentiment polarity by SentiWordNet

Map the opinion to the attribute's static specification => *(attribute, value, like or dislike)*

METHODOLOGY

Step 1: Aspect-level opinion mining

Aspect	Attribute
Picture quality	Optical zoom, sensor resolution
Body	Thickness, height, width, screen size
Money, Price, Cost, …	Price
Weight, Pocket, Heavy, …	Weight
Battery life, Endurance, …	Battery life
…	…

(weight, 570g, like)
(battery life, 15 hours, like)
(price, $370, like)
(resolution, 20.3 megapixels, dislike)

Extract frequent nouns and noun phrases as aspect candidates via Part-Of-Speech (POS) tagger

Identify opinion words via syntactic dependency parser

Group synonymous aspects via WordNet

Identify the opinion's sentiment polarity by SentiWordNet

Map the opinion to the attribute's static specification => *(attribute, value, like or dislike)*

METHODOLOGY

Step 1: Aspect-level opinion mining

Step 2: Preference completion

User *u*'s preference on attribute *a*

Reviewer $\tilde{u}$'s preference on attribute *a*

Preference similarity

$$\overrightarrow{\bar{\phi}_{ua}} = \begin{cases} (\overrightarrow{\phi_{ua}} + \sum_{\tilde{u} \in N_u} \bar{s}_{u\tilde{u}} \overrightarrow{\tilde{\phi}_{\tilde{u}a}})/2, & \text{if } \phi_{ua} \text{ is not missing} \\ \sum_{\tilde{u} \in N_u} \bar{s}_{u\tilde{u}} \overrightarrow{\tilde{\phi}_{\tilde{u}a}}, & \text{otherwise} \end{cases},$$

John (a new buyer)

Partial preferences on attributes

Aspect opinions

Attributes' value preferences

Rebecca

...

Peter

Reviewers

Like-minded reviewers

John's stated preference on attribute sensor resolution → To adjust John's stated preference by Rebecca's and Peter's preferences on sensor resolution

John's missing preference on attribute screen → To complete John's missing preference by Rebecca's and Peter's preferences on screen

METHODOLOGY

Step 1: Aspect-level opinion mining

Step 2: Preference completion

Step 3: Ranking and recommendation

Products with higher scores are recommended to the user

The matching score of product *p* according to the buyer *u*'s preferences

$$M_{up} = \frac{1}{k} \sum_{a=1}^{k} match_w(\bar{\phi}_{ua}, x_{pa})$$

Inner product of the buyer *u*'s preference vector and the product *p*'s vector on the attribute *a*

EXPERIMENT SETUP

A previous user study

57 users

Full preferences on 8 attributes

?

Target choice

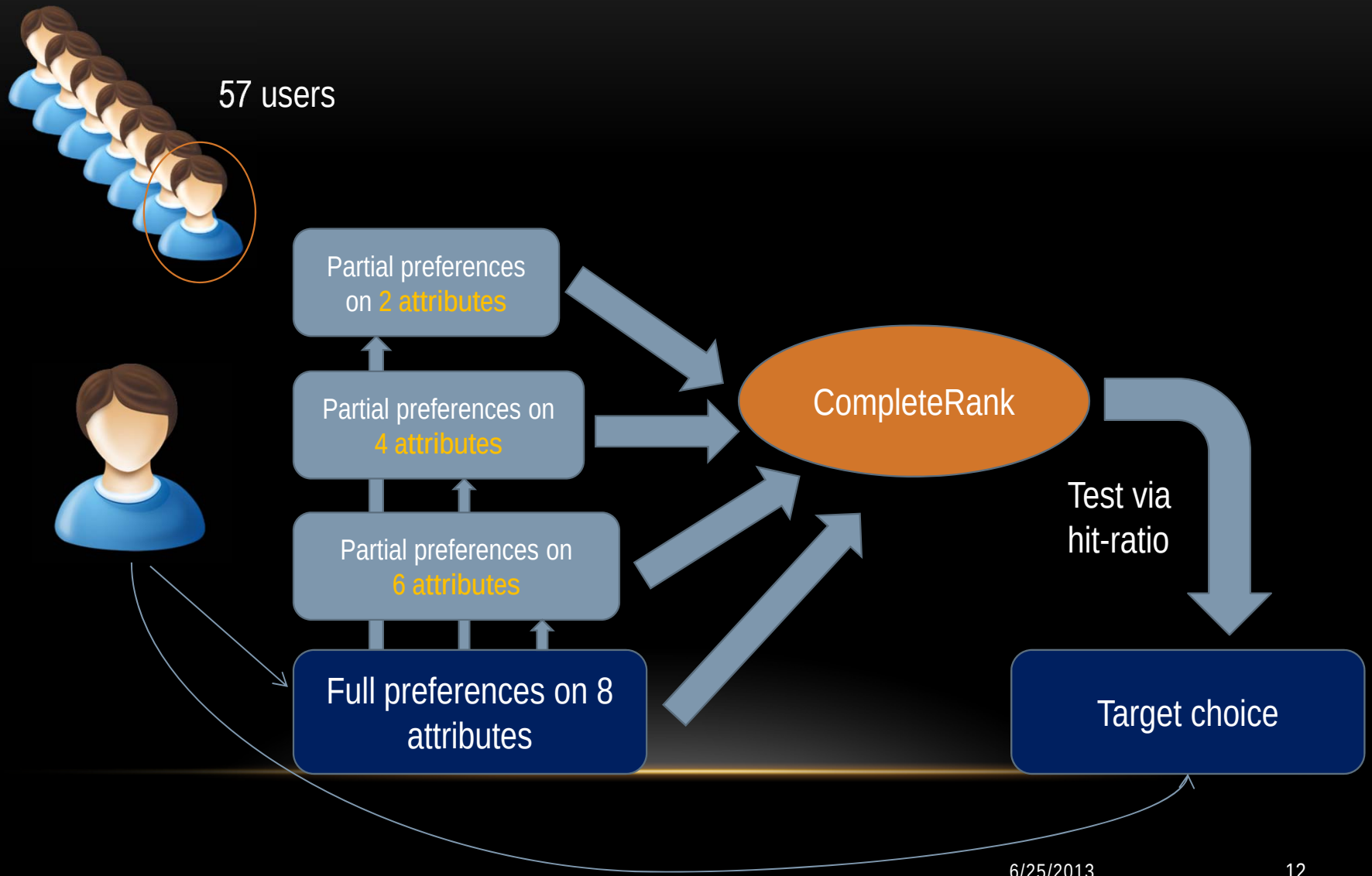
64 digital cameras

Reviewers' value preferences

4904 reviews

amazon.com and you're done.™

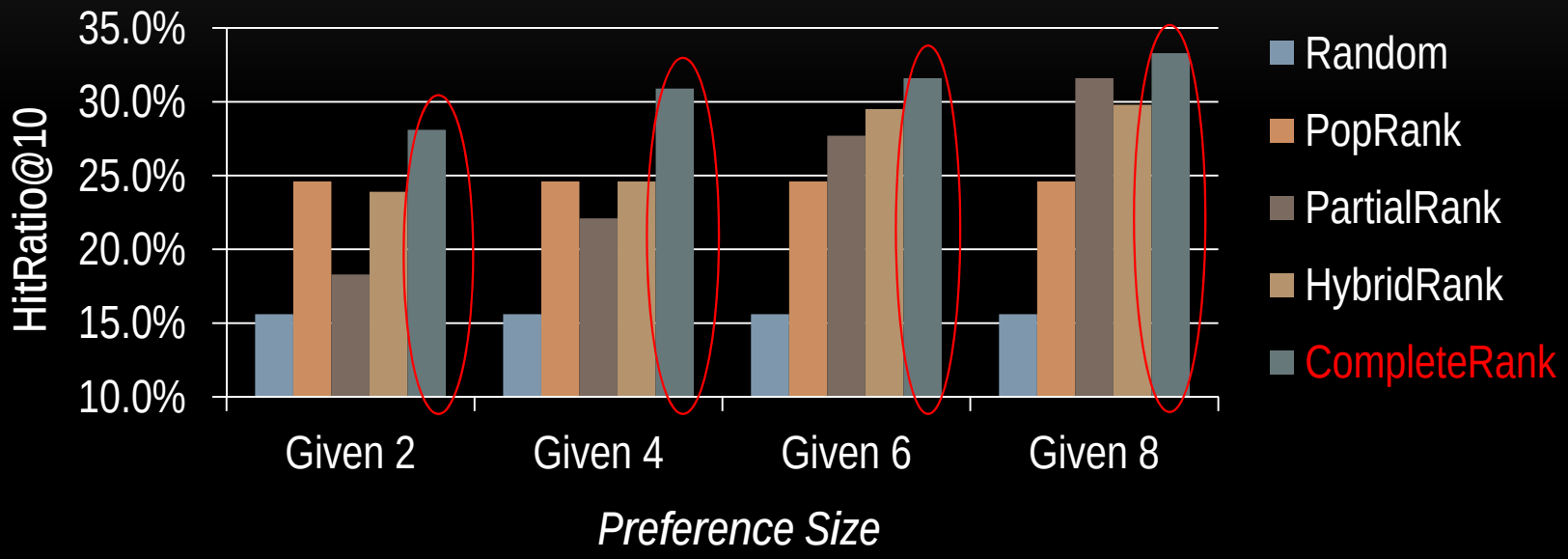
LEAVE-ONE-OUT EVALUATION



COMPARED METHODS

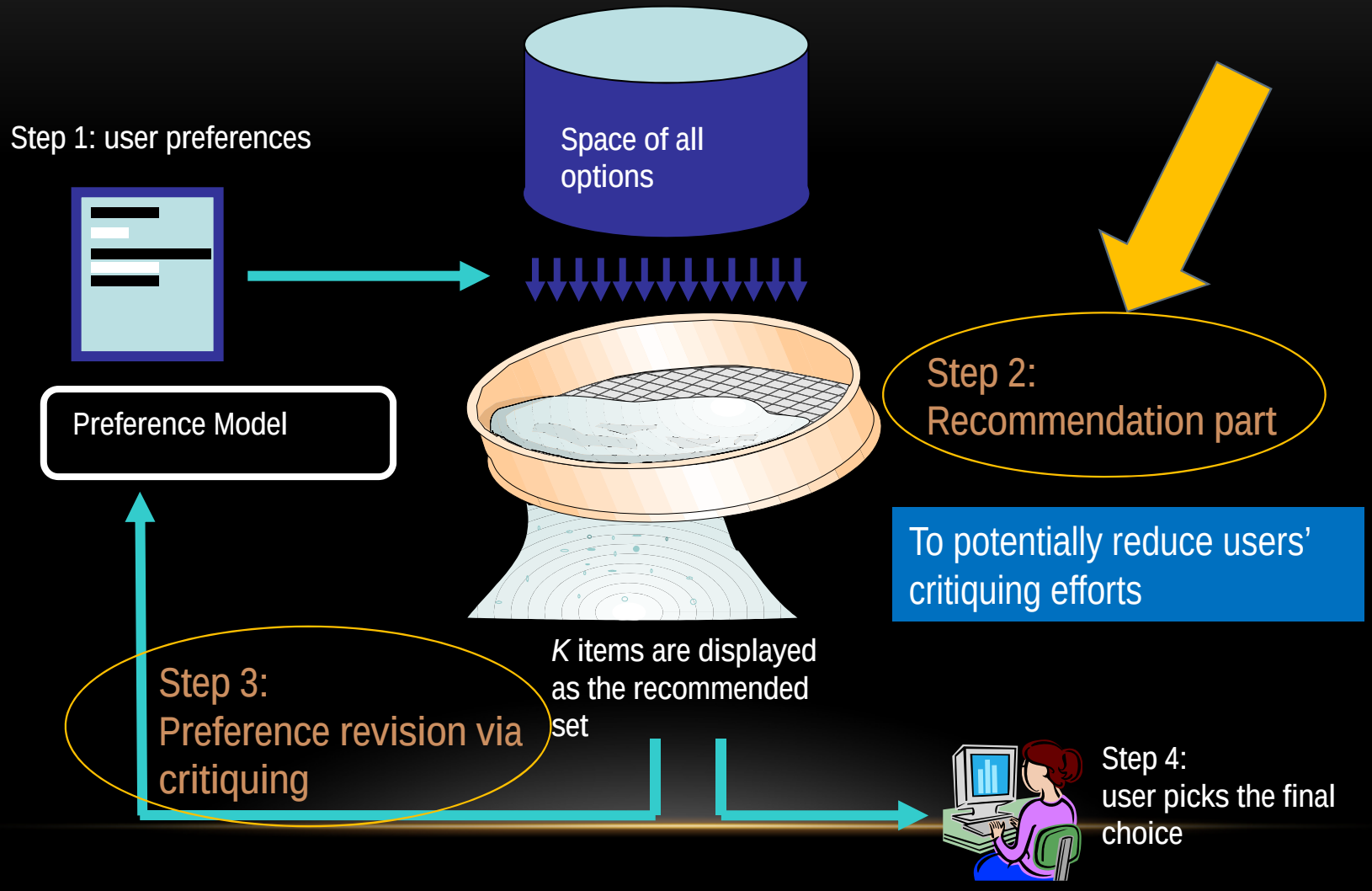
- Random: the probability that the user's target choice appears in top 10 is 0.1563 (10/64)
- PopRank: rank all products by their popularity among reviewers
- PartialRank: match a product to the user's stated (partial) preferences on attributes
- HybridRank: rank products by combining reviewers' opinions and products' matching scores to the user's stated preferences

RESULTS



- CompleteRank is better than all baselines, especially for the buyers with partial preferences
- The usefulness of incorporating online review data for augmenting new-user recommendation
- The usefulness of deriving reviewers' value preferences on the product's attributes

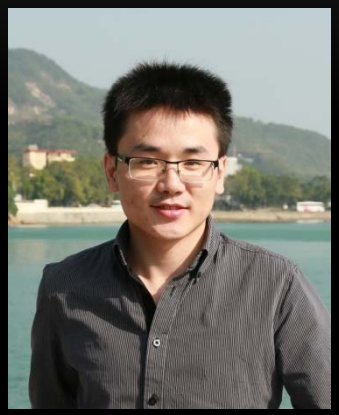
PRACTICAL IMPLICATIONS – CRITIQUING-BASED RECOMMENDER SYSTEMS



FUTURE WORK

1. To learn reviewers' weight preferences on attributes
2. To automatically map aspects to attributes
3. To integrate the algorithm into critiquing-based recommender systems and conduct user evaluations

THANKS!



Feng Wang

Weike Pan

Li Chen

http://www.comp.hkbu.edu.hk/~lichen/