Nantonac Collaborative Filtering

Recommendation Based on Multiple Order Responses

Toshihiro Kamishima & Shotaro Akaho

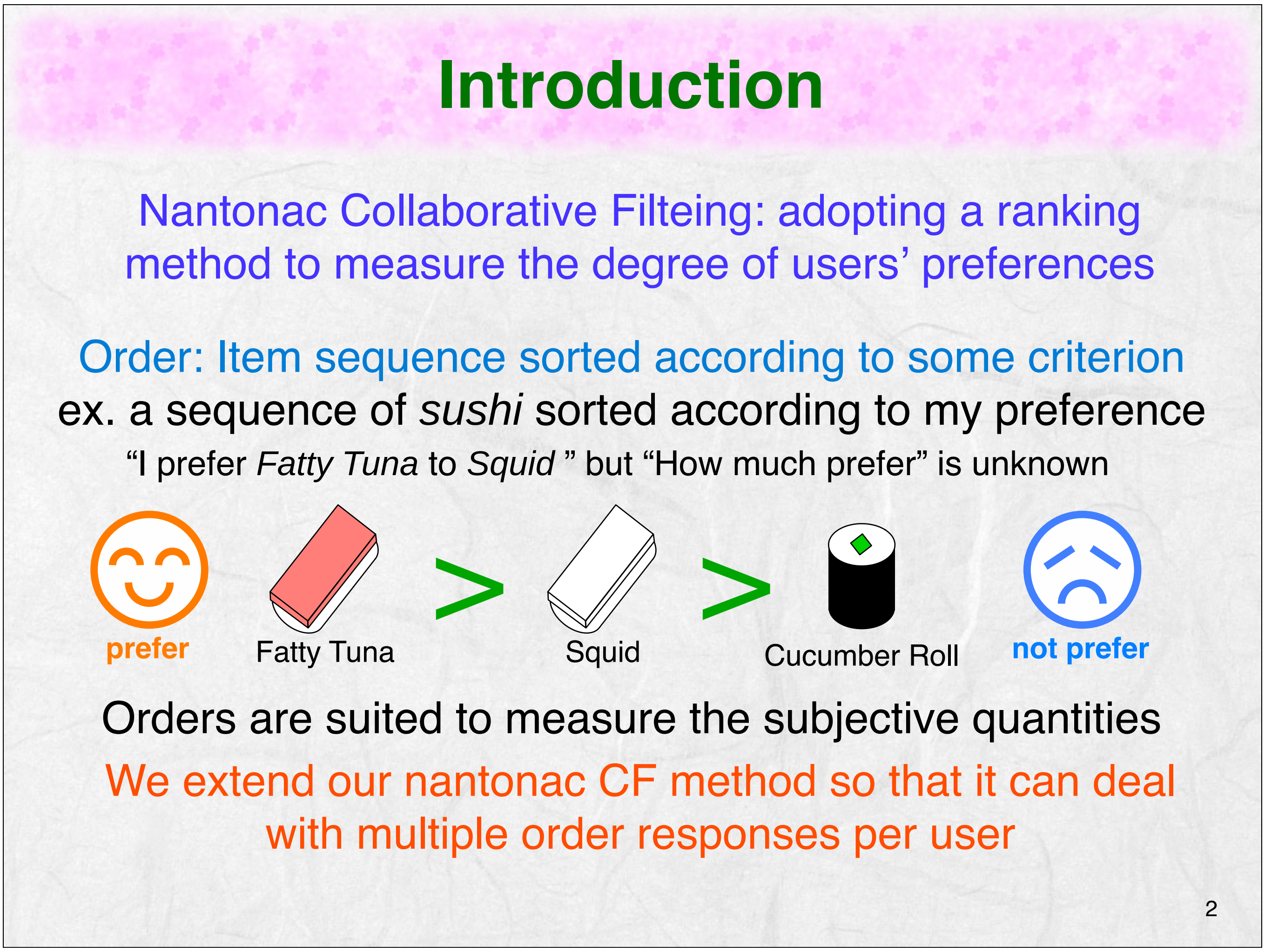
National Institute of Advanced Industrial Science and Technology (AIST)

The International Workshop on Data-Mining and Statistical Science (DMSS2006)

2006.9.25-26 @ Sapporo, Japan

START

We would like to talk about an extension of collaborative filtering adopting a ranking method.



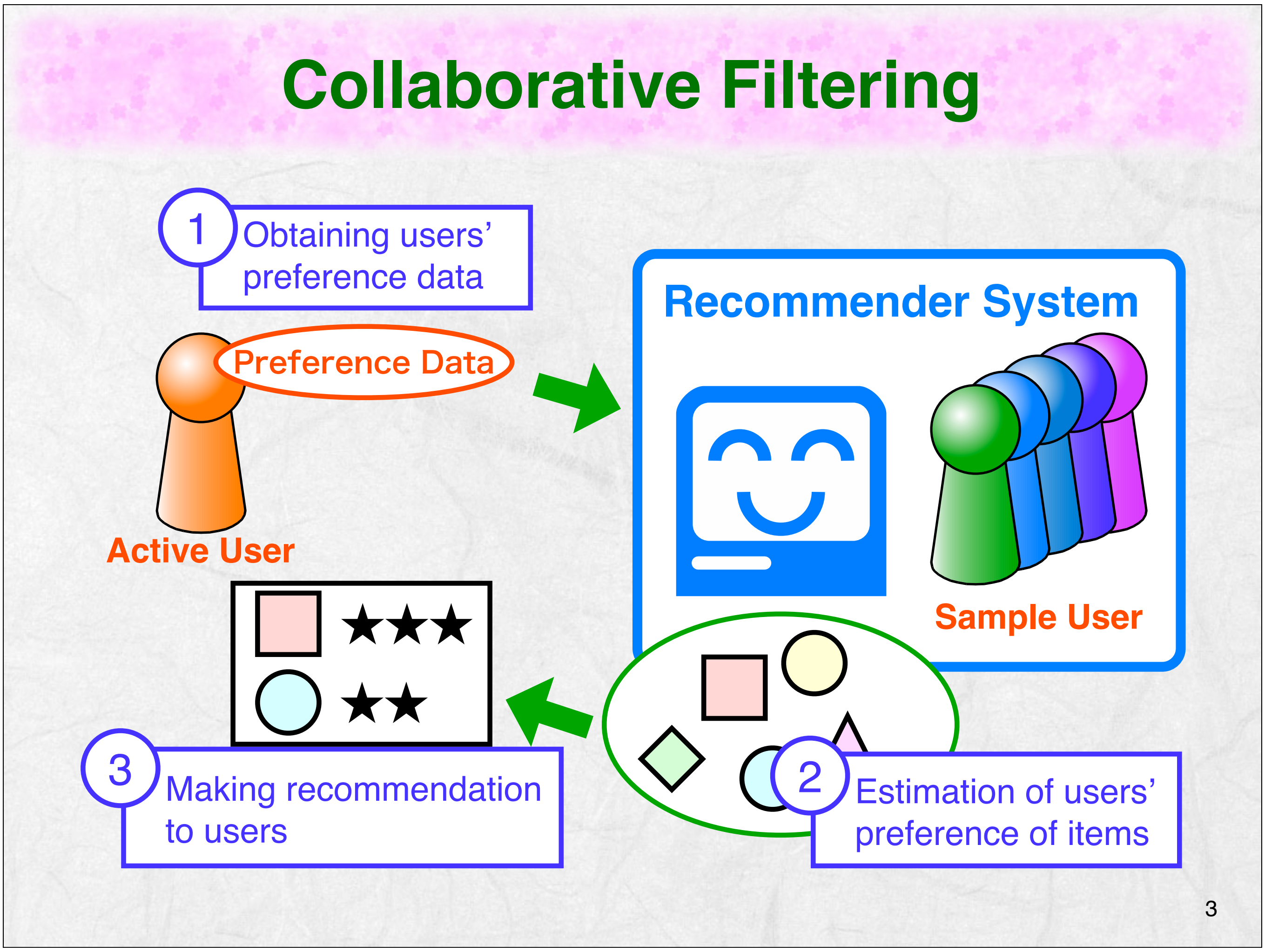
We proposed a framework of nantonac collaborative filtering. In this framework, the degrees of users' preferences are measured by a ranking method.

In a ranking method, users' preference patterns are captured by orders, which are the item sequences sorted according to the users' preference.

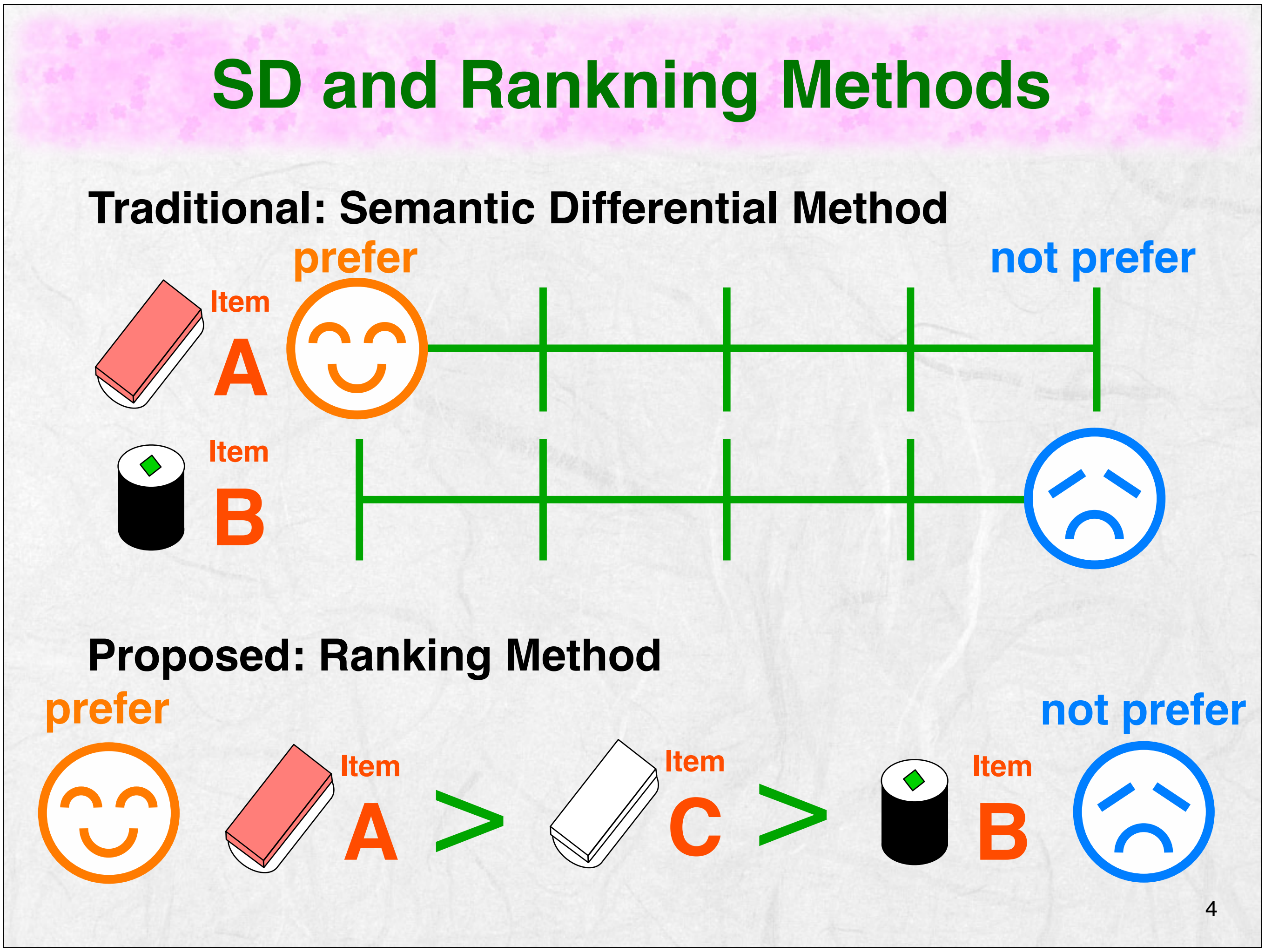
For example, this order indicates my preference in sushi. This order means "I prefer Fatty Tuna to Squid" but "How much prefer" is unknown.

A ranking method is suited to measure the subjective quantities.

In this paper, we extended our nantonac CF method so that it can deal with multiple order responses per user.

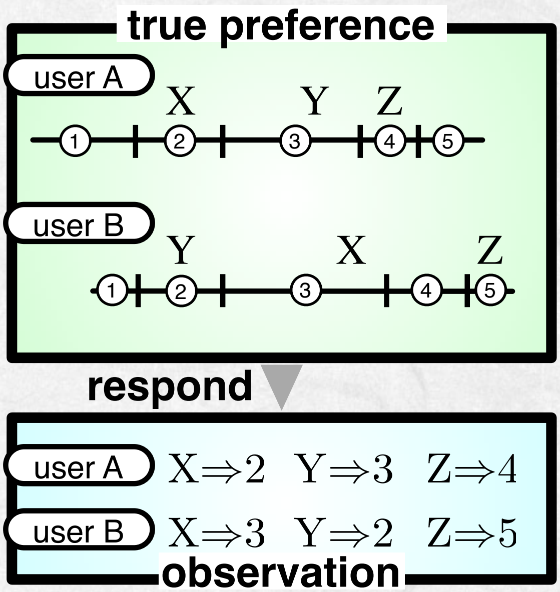


Collaborative filtering is a method to recommend the items that will be preferred by the active user.
The system recommend items preferred by the sample users having similar tastes to the active user.
Collaborative filtering is performed in three steps.
First, the system asks for the active user to input his/her preference pattern.
The active user rates the items shown by the system.
Second, the system estimate the degrees of preference of items which is unknown to the active user.
This is performed based on the databases of preference patterns of sample users.
Finally, the system recommend the items based on the estimation in a suited format.
The second and third steps have been studied well, but the first step have not.



To improve the step of capturing users' preference patterns, we introduced a ranking method.
Traditionally, users show their preferences by pointing on the scales.
For example, if the user prefer item A, select "prefer". if not, select "not prefer"
This method is called a semantic differential method.
Instead, we propose to use a ranking method.
Users sort items according to their degree of preference.
In this example, the user prefer Item A most, and doesn't prefer item B most.

Why Orders? (1)



- The degree of true preference cannot be observed directly
- Each user uses one's own mapping from the degree to rating score
- Ex: The degree of preference on X lies in interval 2 of user A

  ➡ User A replies rating score 2

We show a merit of introducing a ranking method.

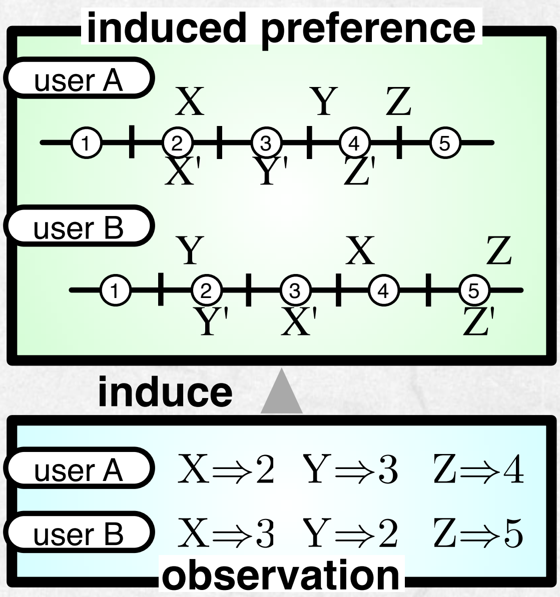
The degrees of true preferences in users' mind cannot be observed directly.

Therefore, each user shows his/her own preference by rating scores on items.

In this case, each user uses one's own mapping from the true preference to rating scores.

For example, the degree of preference on the item X lies in interval 2 of user A; Then, the user A replies rating score 2.

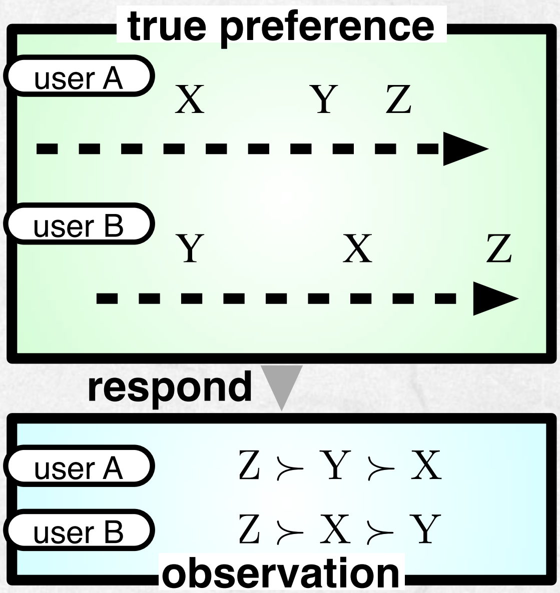
Why Orders? (2)



- We now want to induce the true degree of preference
- The true mapping to rating scores is unknown
- A common idealized mapping scale is of necessity used
- the induced degrees of preferences might not be true
- Ex: The true degrees of X, Y, and Z are changed to X', Y', and Z', respectively

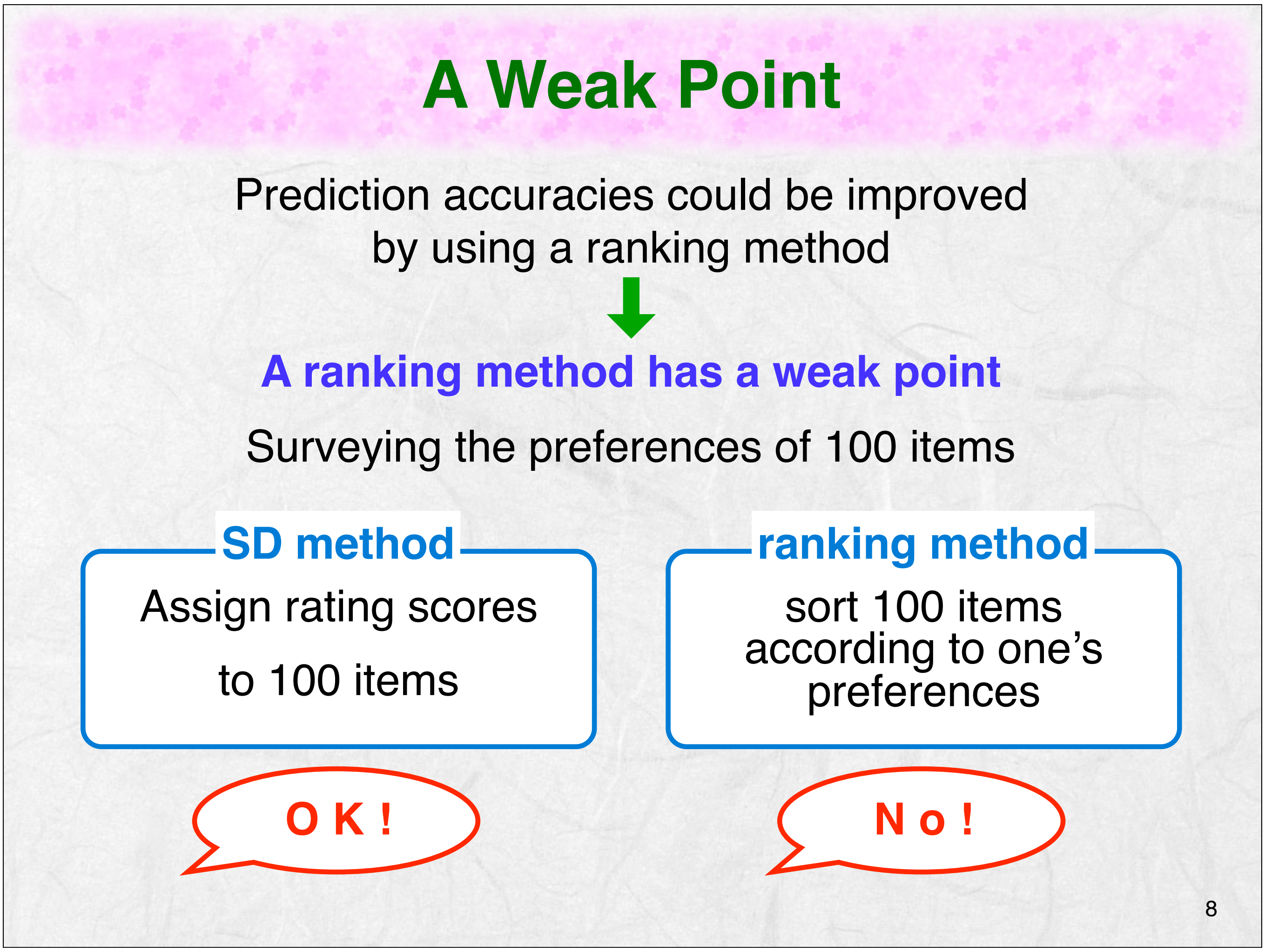
We now want to induce the true degrees of preference from given rating scores.
Unfortunately, we don't know a mapping from rating scores to the true preferences.
Therefore, we of necessity use idealized mapping scales, that is common for all users.
The total lengths of this scale are the same for all users, and the all intervals of rating scores are equal.
Of course, this scale is different from the true mapping of users, so the induced degrees of preferences might not be true.
In this figure, true preference level of X, Y, and Z, are changed to X', Y', and Z', respectively.

Why Orders? (3)

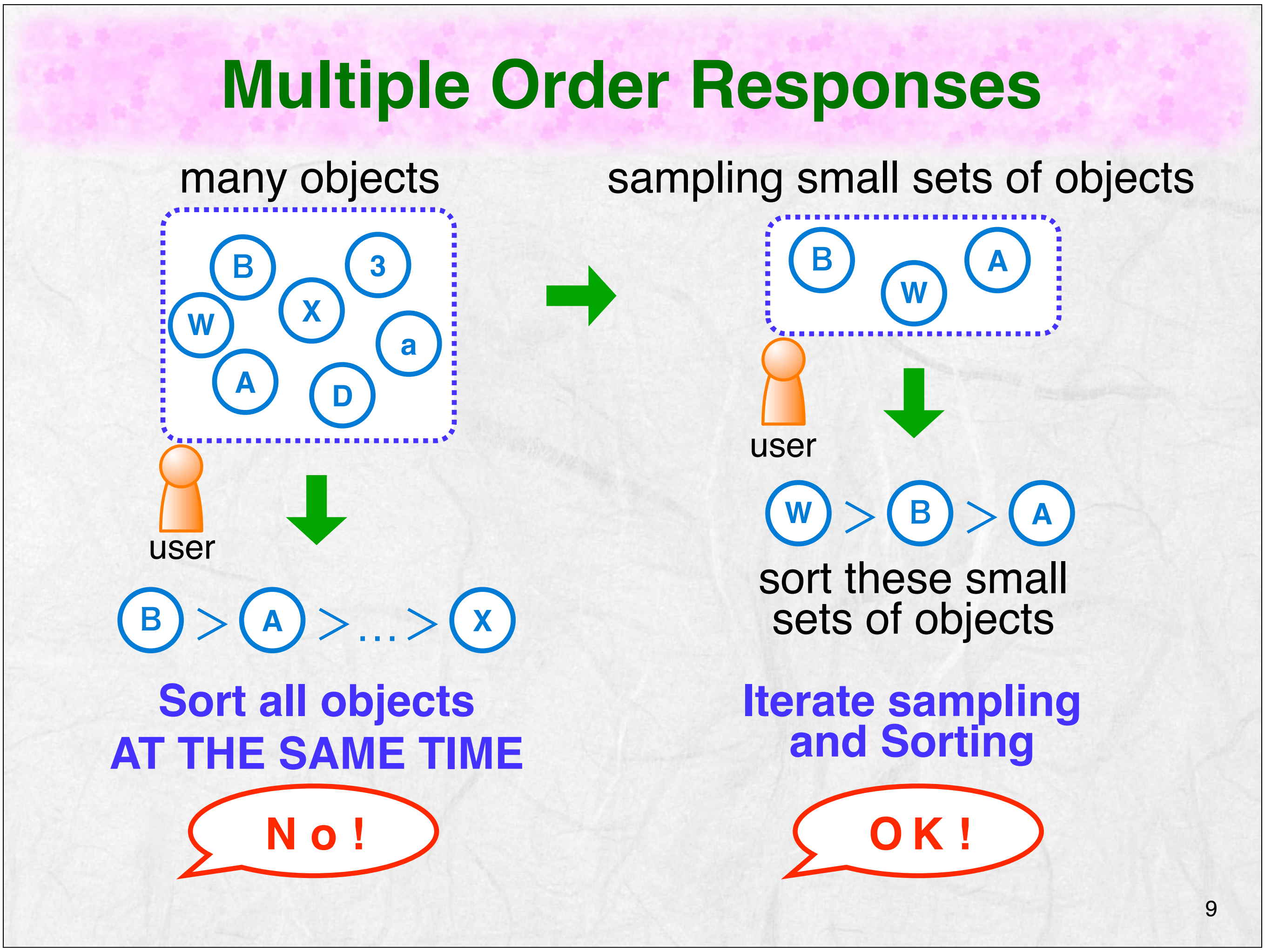


- In a ranking method, the degrees of preferences are relatively specified
- We don’t need to use a unsafe common mapping between the degrees of preference and observed rating scores

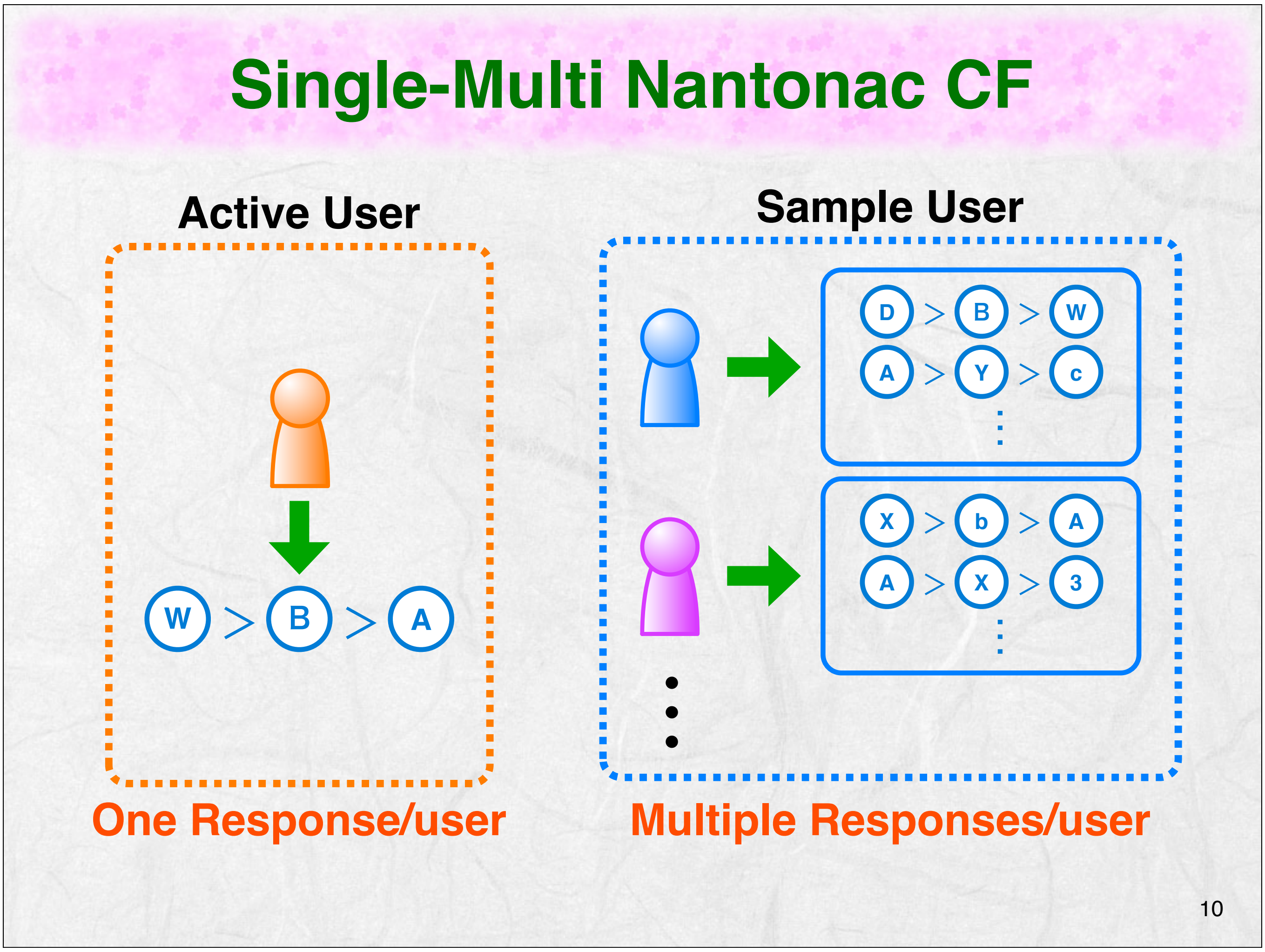
In a case of a raking method, absolute levels of preference are ignored. The degrees of preferences are relatively specified.
Therefore, we don’t require the assumption on the mapping scales, and the order in the true preference is concordant with the observed order.
In summary, by adopting a raking method, the users’ preference patterns can be captured more consistently.



We experimentally showed that the prediction accuracy can be improved by using a ranking method, even if rating scores are appropriately normalized.
However, a ranking method has a weak point.
Assume that we are now trying to survey users' tastes on 100 items.
In an SD method, asking for users to assign rating scores to 100 items is POSSIBLE.
While, in a ranking method, asking for users to sort 100 items according to one's preference is IMPOSSIBLE.
So, the total preference information is limited.



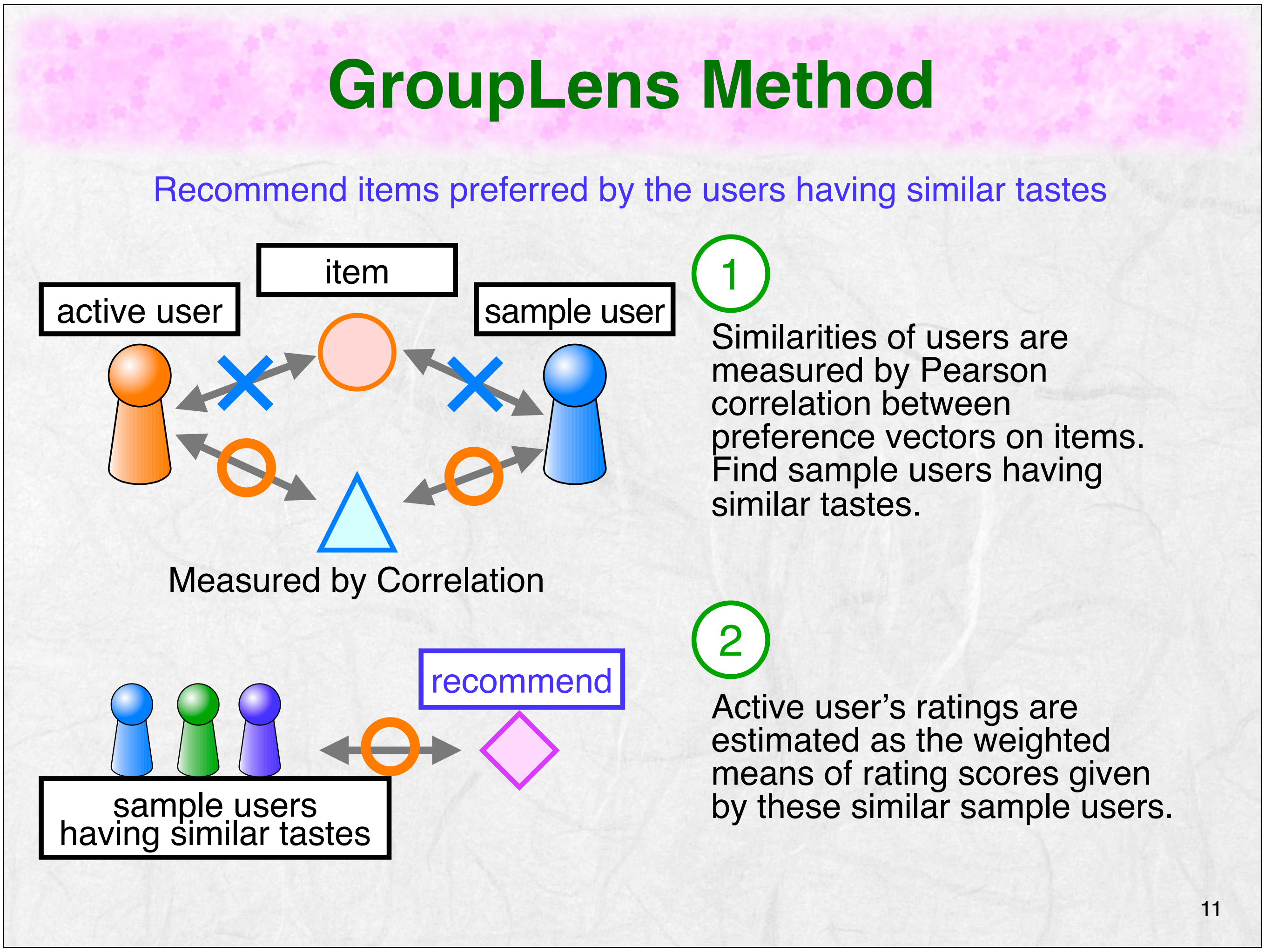
Sorting all objects at the same time is IMPOSSIBLE.
However, the system first samples small item subset from the original entire set.
Then, the system ask for users to sort each sampled set.
By repeating this process, the system can obtain multiple order responses on many items.
Therefore, the preference information is no longer limited.
The difficulty of this method is inconsistency among responses, or loss of information.



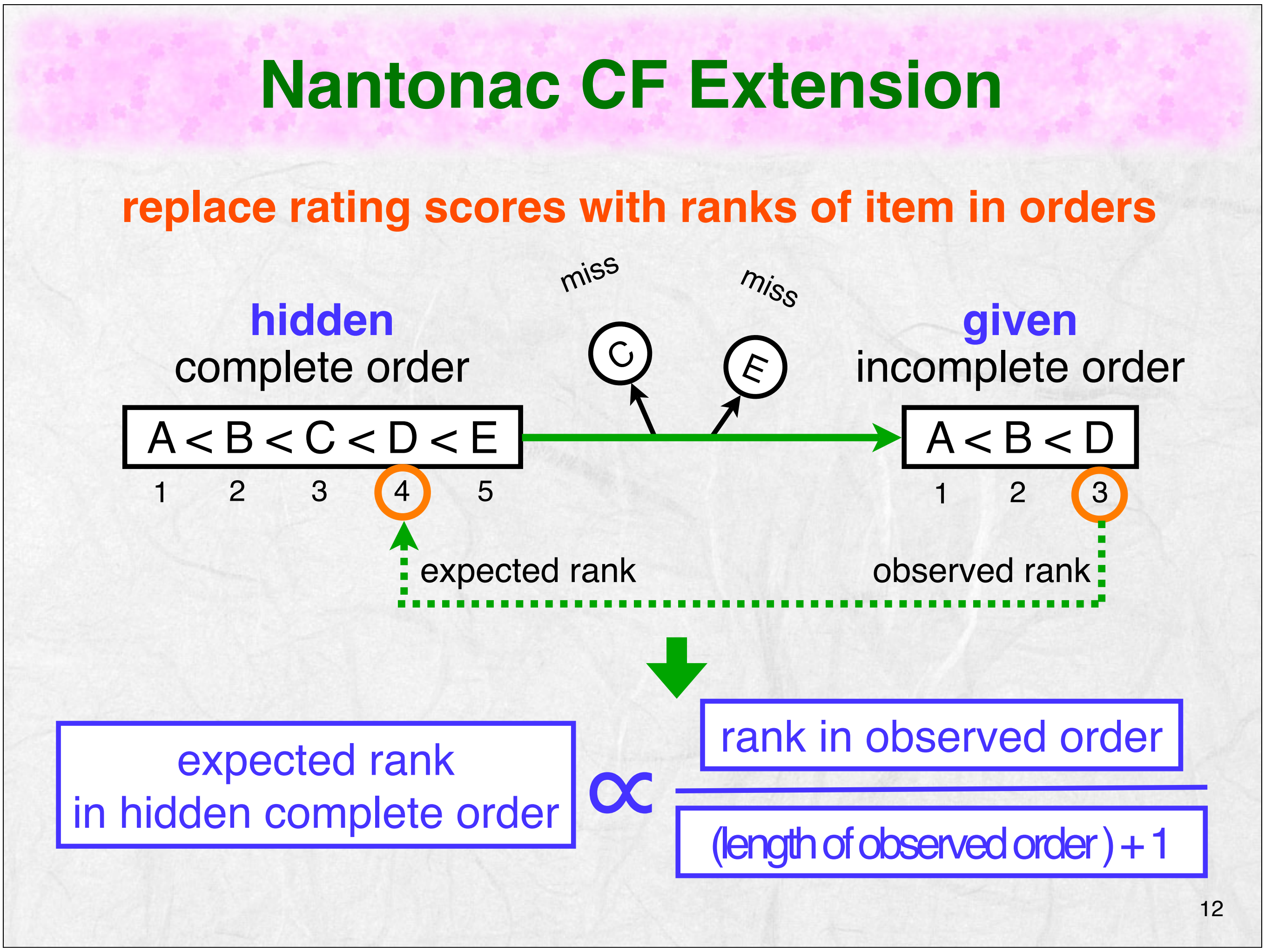
Based on this idea, we extended our original framework.

An active user still returns one response, but sample users return multiple responses per user.

We call this extension a single-multi nantonac CF.

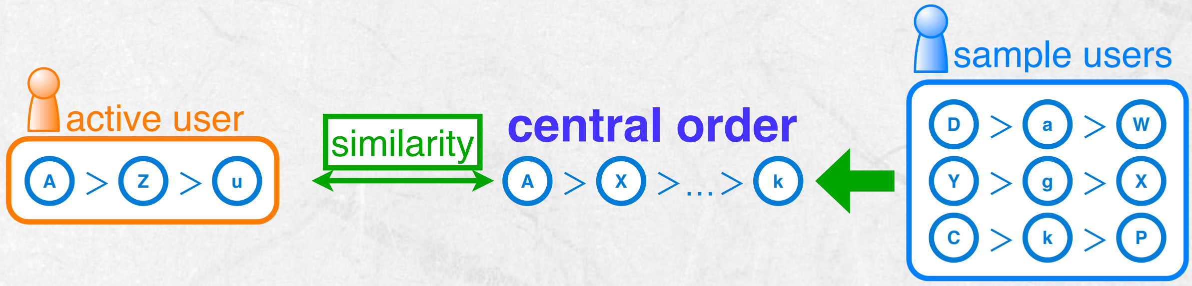


Before showing our extended method, we show the original GroupLens method based on rating scores.
In this GroupLens method, the system recommend items preferred by the users having similar tastes.
In the first step, similarities of users are measured by Preason correlation between preference vectors on items. Then, sample users having similar tastes are found.
In the second step, the active user's ratings on candidate items are estimated as the weighted means of rating scores given by these similar sample users.



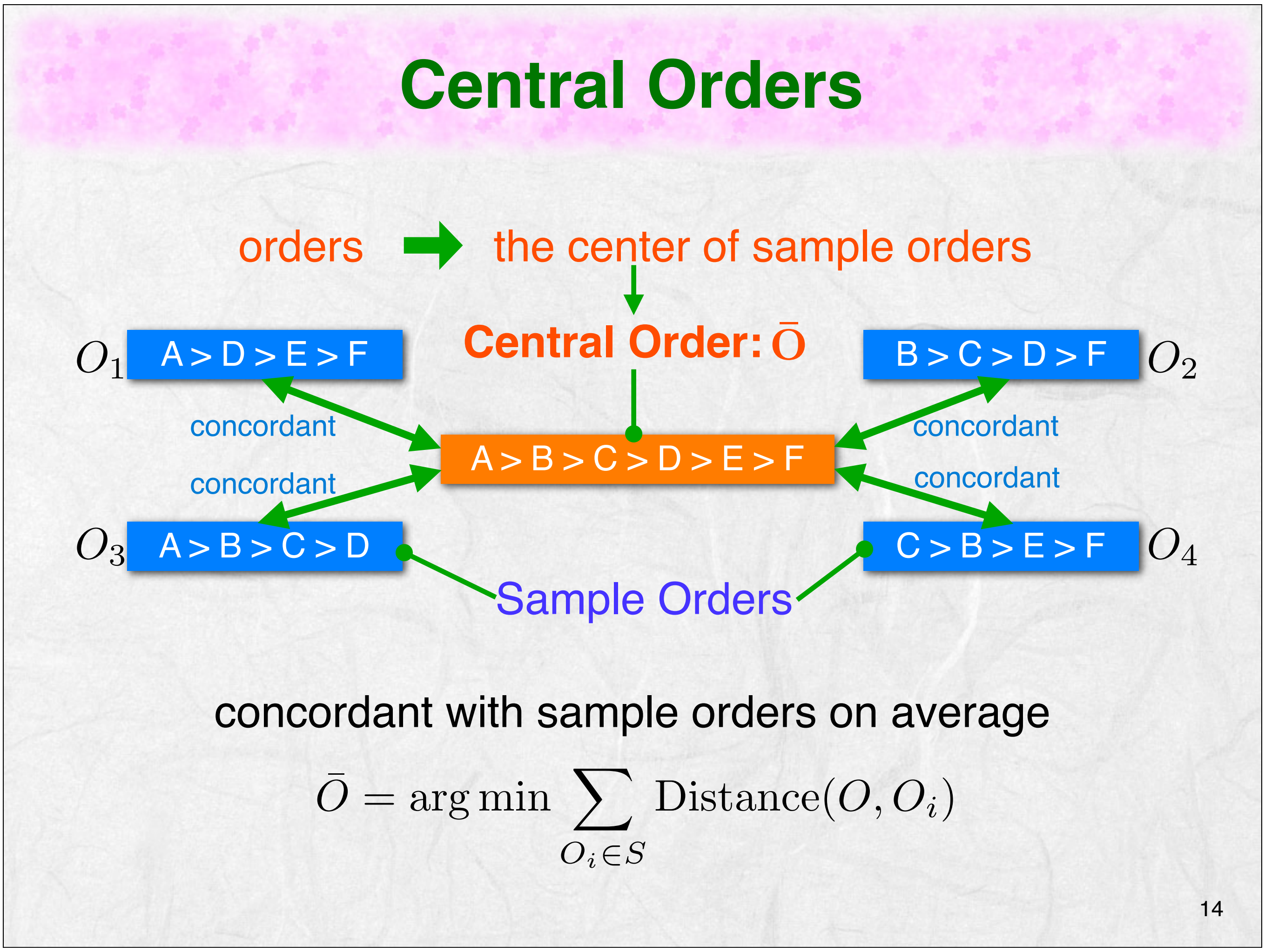
To apply the GroupLens method to a nantonac CF framework, we simply replace rating scores with ranks in order responses.
Here, response orders are incomplete, in which subset of items are ranked. So, ranks in incomplete orders may not be comparable each other. However, this is not the case, under the following assumption.
There is hidden complete order in users' mind.
Items are selected uniformly at random, and these are missed, then incomplete orders are observed as responses.
In this case, according to the theory of order statistics, observed ranks in incomplete orders are proportional to the expected ranks in complete orders.
Therefore, observed ranks can be comparable.

Multiple Responses Extension

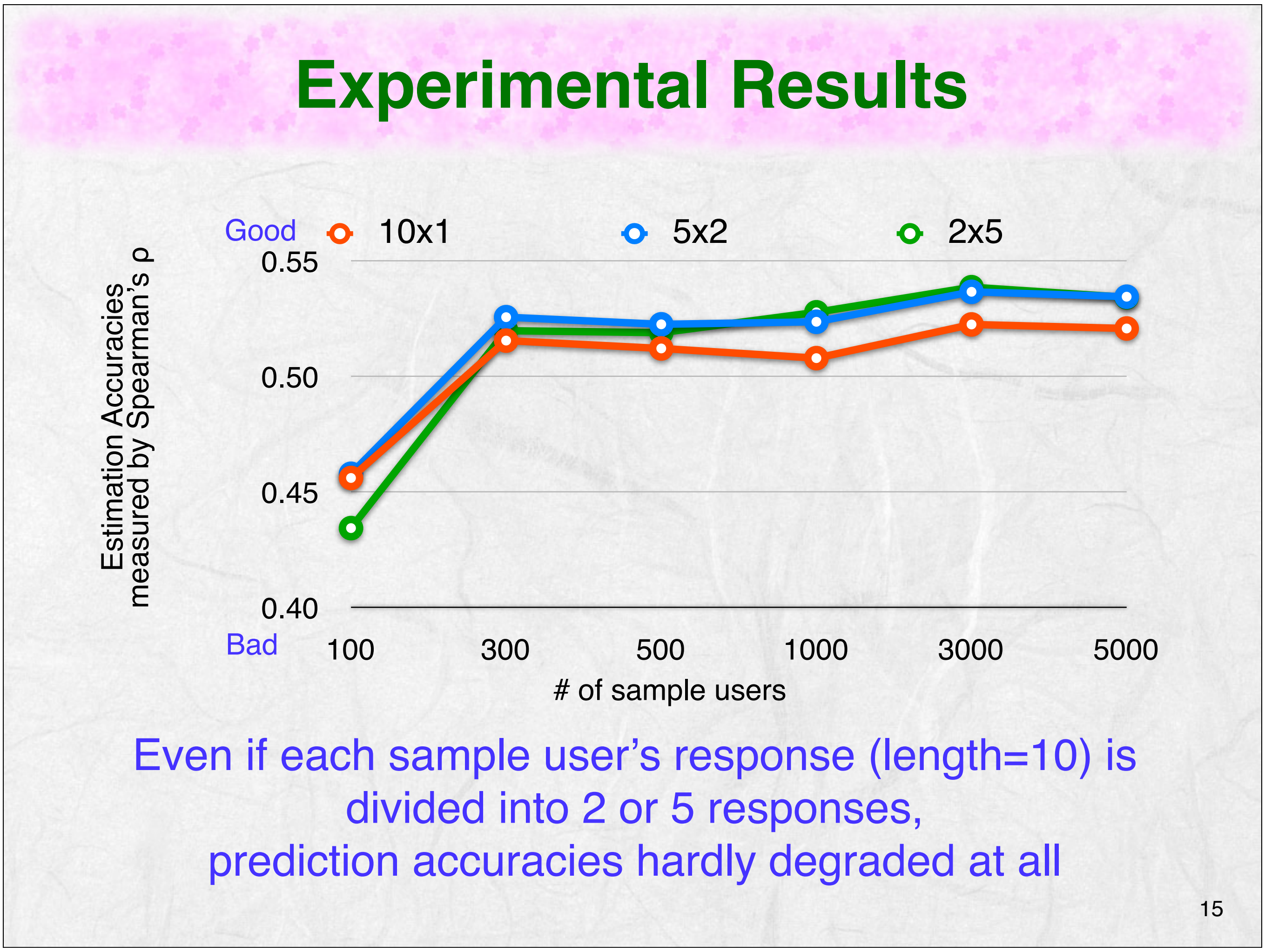


- Multiple orders of sample users are first aggregated into a single central order
- Similarities between users are measured based on ranks in the central order

Here, we extend our original nantonac CF to multiple order setting.
In the GroupLens method, only one response per user is allowed.
To overcome this limitation, we first aggregate multiple orders of sample users into a single central order.
Similarities between users are measured based on ranks in this central order.



A central order is the order that are as concordant with a set of orders as possible.
This is defined as the order that minimizes the sum of distances to a given set of orders.
Examples of distances between orders are Kendall, Spearman, or Ulam distances.
Generally, the problem of deriving central orders is NP-hard.
However, if all sample orders are complete and Spearman's distance is adopted, a central order can be derived by sorting according to mean ranks in linear time.
In the case of incomplete orders, we heuristically use the means of expected ranks.



We performed the experiment to check our method.

To simulate multiple order responses, for example, one response whose length is ten is divided into two orders whose length are five.

The red curve represents original result, and the others represent results of multiple orders.

Even if some information is lost by the division, the prediction accuracies hardly degraded at all.

Conclusion

Summary

- Introducing a ranking method can improve prediction accuracy, even if rating scores are
- Single-Multi case: Division into multiple responses of sample users don't degrade the recommendation quality
- Multi-Multi case: Division into multiple responses of active users rather damage the prediction

More Information

Homepage：http://www.kamishima.net/

おまけ：朱鷺の杜Wiki (機械学習について書き込んでください)

http://www.neurosci.aist.go.jp/ibisforest/

We would like to conclude our talk.
Our contributions are as follows.
That's all we have to say. Thank you.