Efficiency Improvement of Neutrality-Enhanced Recommendation

Toshihiro Kamishima*, Shotaro Akaho*, Hideki Asoh*, and Jun Sakuma†

**National Institute of Advanced Industrial Science and Technology (AIST), Japan*

†University of Tsukuba, Japan; and Japan Science and Technology Agency

Workshop on Human Decision Making in Recommender Systems

In conjunction with the RecSys 2013 @ Hong Kong, China, Oct. 12, 2013

START

Today, we would like to talk about the enhancement of the neutrality in recommendation.

Overview

Providing neutral information is important in recommendation

- avoidance of biased recommendation
- fair treatment of content suppliers or item providers
- adherence to laws and regulations in recommendation

Information-neutral Recommender System

The absolutely neutral recommendation is intrinsically infeasible, because recommendation is always biased in a sense that it is arranged for a specific user

This system makes recommendation so as to enhance the neutrality from a viewpoint feature specified by a user

Providing neutral information is important in recommendation due to these reasons.
For this purpose, we propose an information neutral recommender system.
Unfortunately, the absolutely neutral recommendation is intrinsically infeasible.
Therefore, this system makes recommendation so as to enhance the neutrality from a viewpoint feature specified by a user.

Outline

Introduction

Recommendation Neutrality

- Viewpoint feature, Intuitive definition

Applications of Recommendation Neutrality

- Avoidance of biased recommendation, Fair treatment of content providers, Adherence to laws and regulations

Information-neutral Recommendation System

- pmf model, information-neutral recommender system, neutrality terms

Experiments

- small data, genre-wise mean differences, larger data

Discussion about the Recommendation Neutrality

- subjectivity & objectivity, recommendation diversity, privacy-preserving data mining, and more

Conclusion

This is an outline of our talk.
After showing definition of the recommendation neutrality and its applications, we introduce an information neutral recommender systems.
And, we show our experimental results, discuss recommendation neutrality, and conclude our talk.

Recommendation Neutrality

We begin with our intuitive definition of the recommendation neutrality.

Viewpoint Feature

As in a case of standard recommendation, we use random variables
X: a user, Y: an item, and R: a rating value

We adopt an additional variable for the recommendation neutrality

V : viewpoint feature

- It is specified by a user depending on his/her purpose
- Recommendation results are neutral from this viewpoint
- Its value is determined depending on a user, an item, and their features

Ex. viewpoint = user's gender / movie's release year

In this talk, a viewpoint feature is restricted to a binary type

As in a case of standard recommendation, we use random variables a user, X, an item, Y, and a rating, R.
We atopt an additional variable for the recommendation neutrality, a viewpoint feature, V.
It is specified by a user, recommendation results are neutral from this viewpoint, its value is determined depending on a user, an item, and their features.
In this talk, a viewpoint feature is restricted to a binary type.

Recommendation Neutrality

Recommendation Neutrality

- Recommendation results are neutral if no information about a given viewpoint feature does not influence the results
- The status of the specified viewpoint feature is explicitly excluded from the inference of the results

Ex. viewpoint = movie's release year

⬇

Whether a movie is new or old does not influence the inference of whether the movie is recommended or not

⬇

If movies A and B are the same except for their release year, the movie A is always recommended when the movie B is recommended, and vice versa

We give an intuitive definition of the recommendation neutrality.
Recommendation results are neutral if no information about a given viewpoint feature does not influence the results.
This implies that the status of the specified viewpoint feature is explicitly excluded from the inference of the results.
For example, movie's release year is specified as a viewpoint feature.
In this case, whether a movie is new or old does not influence the inference of whether the movie is recommended or not.

Applications of the Recommendation Neutrality

We give three example applications of the recommendation neutrality.

Avoidance of Biased Recommendation

Biased Recommendations can lead to inappropriate decisions

Biased Recommendation

- exclude a good candidate from a set of options
- rate relatively inferior options higher

The Filter Bubble Problem

Pariser posed a concern that personalization technologies narrow and bias the topics of information provided to people

http://www.thefilterbubble.com/

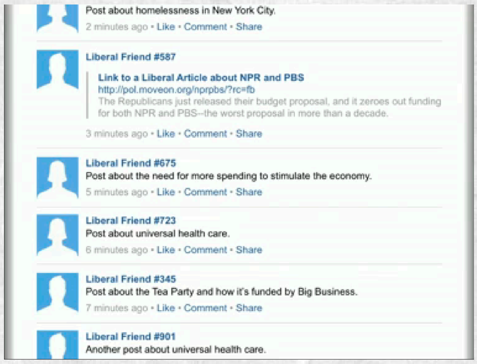
First, biased recommendations may exclude a good candidate from candidates, or may rate relatively inferior option higher.
Consequently, biased recommendations can lead to inappropriate decisions.
Pariser pointed out a problem of such biased recommendations as the filter bubble problem, which is a concern that personalization technologies narrow and bias the topics of information provided to people.

Filter Bubble

[TED Talk by Eli Pariser]

Friend recommendation list in Facebook

To fit for Pariser's preference, conservative people are eliminated from his recommendation list, while this fact is not noticed to him



viewpoint = a political conviction of a friend candidate

Whether a candidate is conservative or progressive does not influence whether he/she is included in a friend list or not

Pariser show an example of a friend recommendation list in Facebook.
To fit for his preference, conservative people are eliminated form his recommendation list, while this fact is not noticed to him.
In this case, a political conviction of a friend candidate is specified as a viewpoint.
Then, whether a candidate is conservative or progressive does not influence whether he/she is included in a friend list or not.

Fair Treatment of Content Providers

System managers should fairly treat their content providers

Ranking in a list retrieved by search engines

[Bloomberg]

The US FTC has been investigating Google to determine whether the search engine ranks its own services higher than those of competitors

Content providers are managers' customers

For marketplace sites, their tenants are customers, and these tenants must be treated fairly when recommending the tenants' products

viewpoint = a content provider of a candidate item

Information about who provides a candidate item is ignored, and providers are treated fairly

The second application is a fair treatment of content providers.
Recommender system managers should fairly treat their content providers.
For example, according to the Blooberg's report, the US FTC has been investigating Google to determine whether the search engine ranks its own services higher than those of competitors.
In this case, a content provider of a candidate item is specified as a viewpoint.
Then, information about who provides a candidate item is ignored, and providers are treated fairly.

Adherence to Laws and Regulations

[Sweeney 13]

Recommendation services must be managed while adhering to laws and regulations

suspicious placement keyword-matching advertisement

Advertisements indicating arrest records were more frequently displayed for names that are more popular among individuals of African descent than those of European descent

Socially discriminative treatments must be avoided

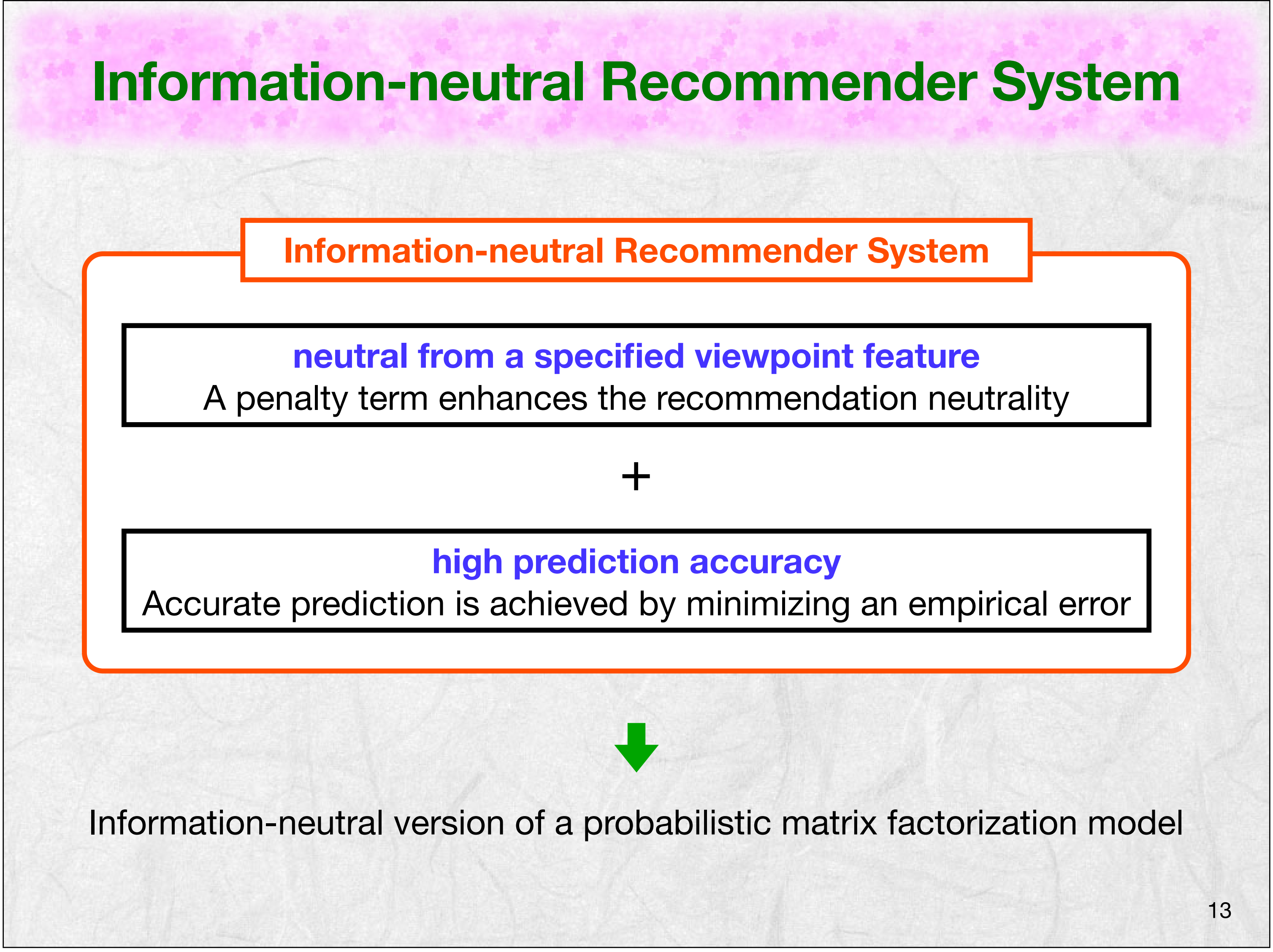
viewpoint = users' socially sensitive demographic information

Legally or socially sensitive information
can be excluded from the inference process of recommendation

Finally, recommendation services must be managed while adhering to laws and regulations.
This is an example of an suspicious placement keyword-matching advertisement reported by Sweeney.
Advertisements indicating arrest records were more frequently displayed for names that are more popular among individuals of African descent than those of European descent.
Socially discriminative treatments must be avoided.
For this purpose, users' socially sensitive demographic information is specified as a viewpoint
Then, legally or socially sensitive information can be excluded from the inference process of recommendation.

Information-neutral Recommender System

Next, we introduce our information-neutral recommender system.



A goal of an information neutral recommender system is to make recommendation that is neutral from a specified viewpoint while keeping high prediction accuracy.
A penalty term enhances the recommendation neutrality, and accurate prediction is achieved by minimizing an empirical error
We introduce an information-neutral version of a probabilistic matrix factorization model.

Probabilistic Matrix Factorization

[Salakhutdinov 08, Koren 08]

Probabilistic Matrix Factorization Model
predict a preference rating of an item y rated by a user x
well-performed and widely used

$$\hat{r}(x,y) = \overbrace{\mu}^{\text{global bias}} + \underbrace{b_x}_{\text{user-dependent bias}} + \underbrace{c_y}_{\text{item-dependent bias}} + \overbrace{\mathbf{p}_x \mathbf{q}_y^\top}^{\text{cross effect of users and items}}$$

For a given training data set (x_i: user, y_i: item, r_i: rating), model parameters are learned by minimizing the squared loss function with a L_2 regularizer

A probabilistic matrix factorization model is designed to predict a preference rating.
A preference rating is modeled by this formula, which consists of three bias terms and one cross term.
For a given training data set, model parameters are learned by minimizing the squared loss function with a L2 regularizer.

Information-neutral PMF Model

information-neutral version of a PMF model

- **adjust ratings according to the state of a viewpoint**
  incorporate dependency on a viewpoint variable
- **enhance the neutrality of a score from a viewpoint**
  add a neutrality function as a constraint term

adjust ratings according to the state of a viewpoint

viewpoint feature

$$\hat{r}(x, y, v) = \mu^{(v)} + b_x^{(v)} + c_y^{(v)} + \mathbf{p}_x^{(v)} {\mathbf{q}_y^{(v)}}^\top$$

- Multiple models are built separately, and each of these models corresponds to the each value of a viewpoint feature
- When predicting ratings, a model is selected according to the value of viewpoint feature

These two points are modified in information neutral version of a PMF model.
First, we modify a PMF model so as to be able to adjust scores according to the state of a viewpoint.
Multiple models are built separately, and each of these models corresponds to the each value of a viewpoint feature.
When predicting ratings, a model is selected according to the state of viewpoint feature.

Neutrality Term and Objective Function

enhance the neutrality of a rating from a viewpoint feature

neutrality term, $\mathrm{neutral}(R, V)$: quantify the degree of neutrality

- It depends on both ratings and view point features
- The larger value of the neutrality term indicates that the higher level of the neutrality

Objective Function of an Information-neutral PMF Model

neutrality parameter to control the balance between the neutrality and accuracy

regularization parameter

$$\sum_{\mathcal{D}} \underbrace{(r_i - \hat{r}(x_i, y_i, v_i))^2}_{\text{squared loss function}} + \eta \underbrace{\mathrm{neutral}(R, V)}_{\text{neutrality term}} + \lambda \underbrace{\|\mathbf{\Theta}\|_2^2}_{\text{L}_2 \text{ regularizer}}$$

Parameters are learned by minimizing this objective function

Second, a PMF model is modified so as to be able to enhance the neutrality of a rating from a viewpoint feature.
For this purpose, we introduce a neutrality term to quantify the degree of neutrality.
This neutrality term is added as a penalty term.
A neutrality parameter η controls the balance between the neutrality and accuracy.
Parameters are learned by minimizing this objective function.

Formal Definition of the Recommendation Neutrality

Recommendation Neutrality

Recommendation results are neutral if no information about a given viewpoint feature does not influence the results

the statistical independence between a rating variable, R, and a viewpoint feature, V

$$\Pr[R|V] = \Pr[R] \equiv R \perp\!\!\!\perp V$$

Two types of neutrality terms

Mutual Information

$$\mathrm{I}(R;V) = \sum_{R,V} \Pr[R,V] \log \frac{\Pr[R|V]}{\Pr[R]}$$

Calders&Verwer's Score

$$\| \Pr[R|V=0] - \Pr[R|V=1] \|$$

We give a formal definition of the recommendation neutrality.
Recall that recommendation results are neutral if no information about a given viewpoint feature does not influence the results.
This statement can be straightforwardly formalized as the the statistical independence between a rating variable and a viewpoint feature.
Given this definition, we developed two types of neutrality terms.

Mutual Information

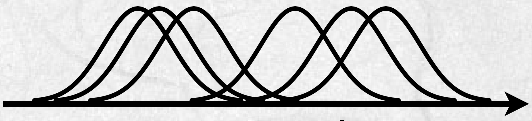
[Kamishima 12]

Neutrality Term in Our Previous Work

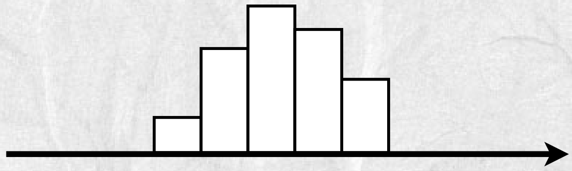
$$-\mathrm{I}(R;V) \approx -\frac{1}{|\mathcal{D}|} \sum_{(r_i,v_i)\in\mathcal{D}} \log \frac{\Pr[\hat{r}_i|v_i]}{\sum_{v\in\{0,1\}} \Pr[\hat{r}_i|v] \Pr[v]}$$

$$\Pr[r|v] = \frac{1}{|\mathcal{D}|} \sum_{(x,y)\in\mathcal{D}} \Pr[r|x,y,v]$$

modeling by a histogram



too complex



mi-hist (mutual information - histogram)

This term cannot differentiate analytically

Optimization is too slow to process even the moderate size of data

The first, mutual information, is a neutrality term in our previous work.
This mi-hist term cannot differentiate analytically.
Therefore, optimization is too slow to process even the moderate size of data.

Calders&Verwer's Score (CV Score)

Our New Neutrality Term

make two distributions of R given $V = 0$ and 1 similar

$$-\|\Pr[R|V=0] - \Pr[R|V=1]\|$$

m-match

$$-\left(\mathrm{Mean}_{\mathcal{D}^{(0)}}[\hat{r}] - \mathrm{Mean}_{\mathcal{D}^{(1)}}[\hat{r}]\right)^2$$

Matching means of predicted ratings when $V = 0$ and $V = 1$

r-match

$$-\sum_{(x,y)\in\mathcal{D}} \left(\hat{r}(x,y,0) - \hat{r}(x,y,1)\right)^2$$

Matching two predicted ratings when $V = 0$ and $V = 1$, regardless of the actual value of V

analytically differentiable and efficient in optimization

The second, CV score, is our new neutrality term.
This is designed to make two distributions of a rating given V equals to 1 and 0 similar.
The m-match term matches means of predicted ratings.
The r-match term matches two predicted ratings.
These two terms are analytically differentiable and efficient in optimization.

Experiments

(Results were improved from those in an proceeding article)

We show experimental results.
Results were improved from those in an proceeding article.

Small Data Set

to compare mi-hist and m-match/r-match terms

General Conditions

- 9,409 use-item pairs are sampled from the Movielens 100k data set (**the mi-hist term cannot process larger than this data set**)
- the number of latent factor $K = 1$ (due to the small size of data)
- regularization parameter $\lambda=1$ (**more finely tuned than that in an article**)
- Evaluation measures are calculated by using five-fold cross validation

Evaluation Measure

- **MAE (mean absolute error)**
  - prediction accuracy
  - Random recommendation: MAE=0.903
  - Original PMF recommendation: MAE=0.759
- **NMI (normalized mutual information)**
  - the neutrality of a predicted ratings from a specified viewpoint

First, we applied our method to small data set to to compare mi-hist and our two new terms, because mi-hist model cannot process larger than this data set.
We used two types of evaluation measure.
MAE, mean absolute error, measures prediction accuracy.
NMI, normalized mutual information, measures the neutrality of a predicted ratings from a specified viewpoint.

Viewpoint Featues

"Year" viewpoint : movie's release year is newer than 1990 or not

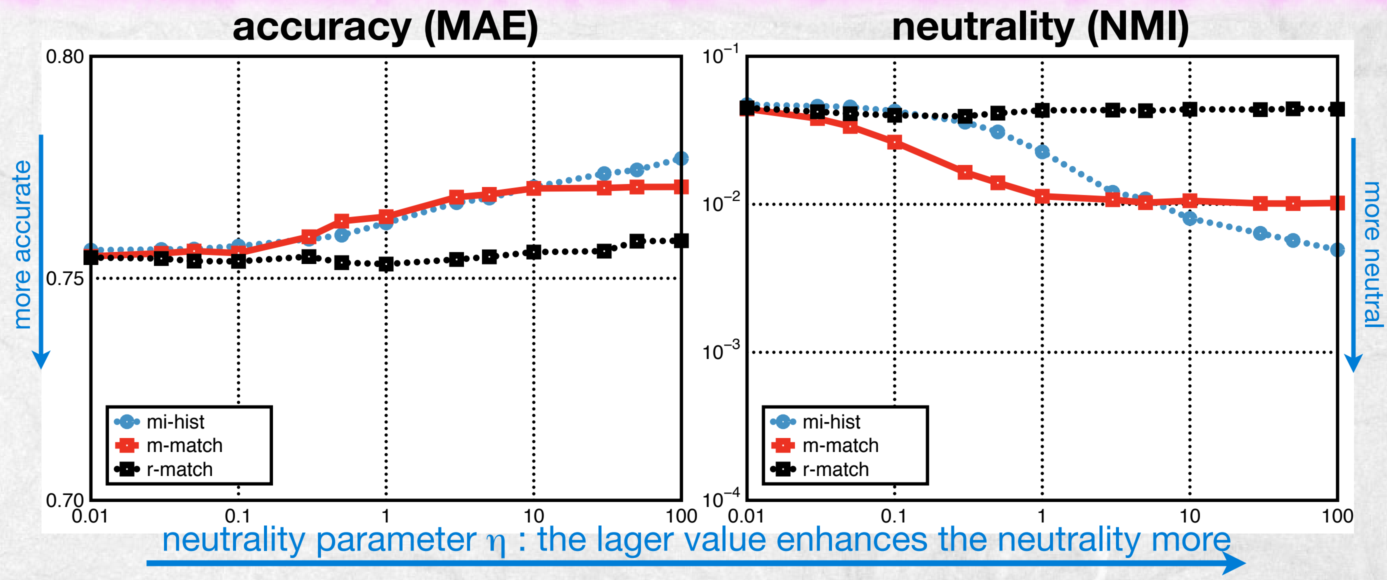
The older movies have a tendency to be rated higher, perhaps because only masterpieces have survived [Koren 2009]

"Gender" viewpoint : a user is male or female

The movie rating would depend on the user's gender

We tested two types of viewpoint variables.
First, a "Year" viewpoint feature represents whether a movie's release year is newer than 1990 or not.
Second, a "Gender" viewpoint feature represents whether a user is male or female.

Small Data & Year Viewpoint



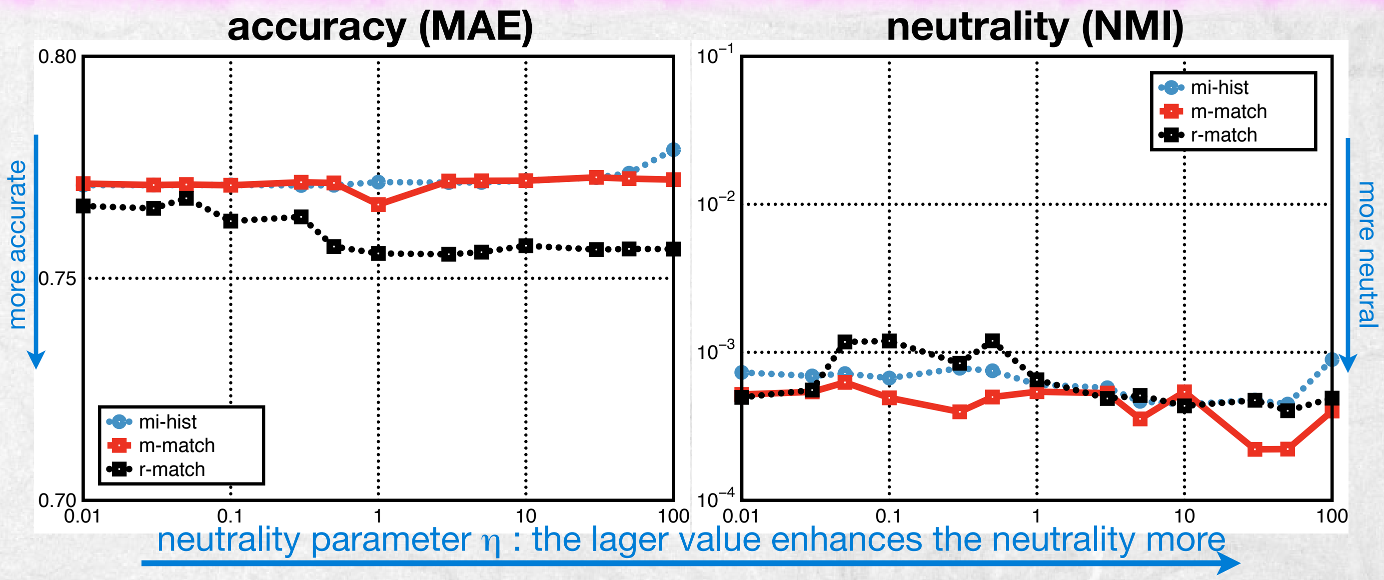
As the increase of a neutrality parameter η,

- prediction accuracies were worsened slightly in all cases,
- the neutralities were enhanced drastically in mi-hist/m-match, cases but not in a r-match case

both m-hist and m-match successfully enhanced the neutrality

These are our experimental results for a Year viewpoint.
X-axes are neutrality parameters, the lager value enhances the neutrality more.
This chart (left) shows the change of the accuracy.
This chart (right) shows the change of the neutrality.
As the increase of a neutrality parameter η, prediction accuracies were worsened slightly in all cases, and the neutralities were enhanced drastically in mi-hist/m-match cases, but not in a r-match case.
Therefore, we can conclude that both m-hist and m-match successfully enhanced the neutrality.

Small Data & Gender Viewpoint



As the increase of a neutrality parameter η,

- both the accuracy and neutrality did not largely change

the neutrality was originally high and failed to improve further

These are our experimental results for a Gender viewpoint.
Unfortunately, both the accuracy and neutrality did not largely change.
This would be because the neutrality was originally high and failed to improve further.

Genre-wise Differences of Mean Ratings

Not provided in a proceeding article

To examine how the recommendation patterns are changed

Computation Procedure

- Data were firstly divided according to their 18 kinds of movies' genres
- Each genre-wise data were further divided into two sets according to their view-point value
- mean ratings were computed for each set, and we showed the differences of between mean ratings for two different values of V

Three Types of Ratings

- the original true ratings in training data
- ratings predicted by the PMF model with the mi-hist and m-match terms, respectively (η = 100)

We then show genre-wise differences of mean ratings to examine how the recommendation patterns are changed.

Genre-wise Differences of Mean Ratings

Gender: males' mean ratings - females' mean ratings

the positive values indicate genres more highly rated by males

	original	mi-hist	m-match
Children's	-0.229	-0.132	-0.139
Animation	-0.224	-0.064	-0.068
Romance	-0.122	-0.073	-0.073
Documentary	0.333	0.103	0.084
Horror	0.479	0.437	0.409
Fantasy	0.783	0.433	0.387

- Differences are basically narrowed by the neutrality enhancement
- INRS didn't simply shift the ratings, and changes were different genre-wise
- NMIs were not changed for a Gender case, but recommendation patterns were surely changed

Predicted ratings under a Gender viewpoint were divided according to movies' genre.
We show the differences that males' mean ratings minus females' mean ratings.
The positive values indicate genres more highly rated by males.
Differences are basically narrowed by the neutrality enhancement.
NMIs were not changed for a Gender case, but recommendation patterns were surely changed.

Large Data Set

to show that m-match and r-match terms are applicable to larger data sets

General Conditions

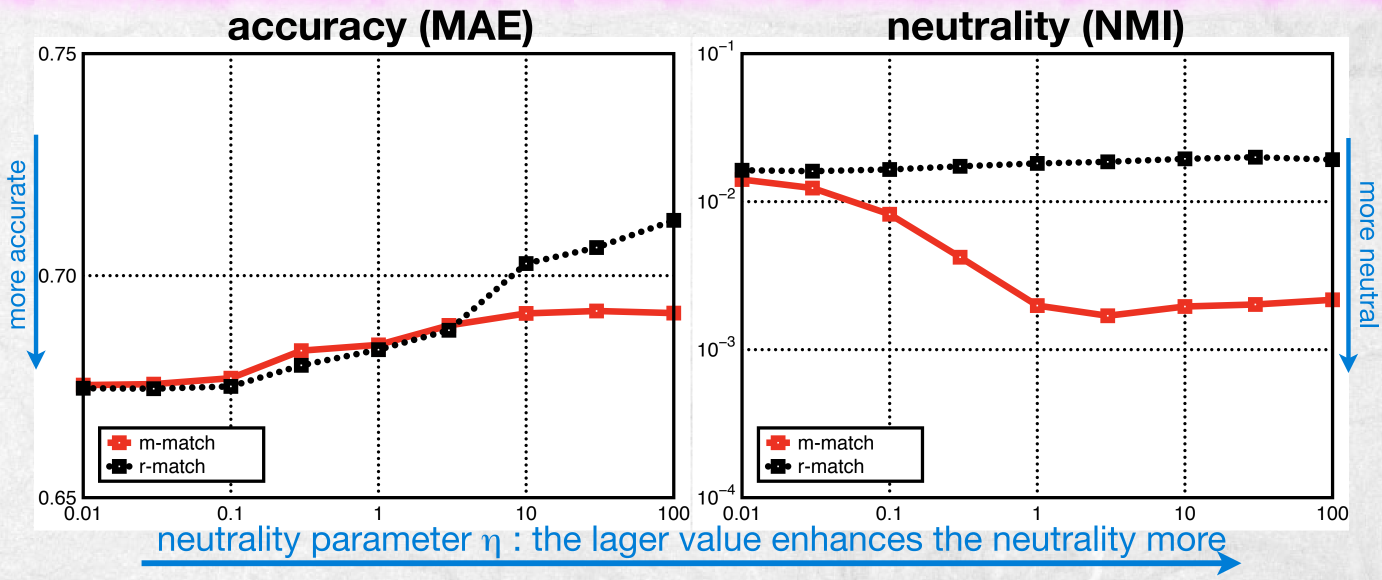
- The Movielens 1M data set (**larger than that in an article**)
- **mi-hist model cannot process this size of data set**
- the number of latent factor $K = 7$
- regularization parameter $\lambda = 1$
- Evaluation measures are calculated by using five-fold cross validation

Baseline Accuracies

- Random recommendation: MAE=0.934
- Original PMF recommendation: MAE=0.685

Finally, to show that m-match and r-match terms are applicable to larger data sets, we applied them to the Movielens 1M data set.

Large Data & Year Viewpoint



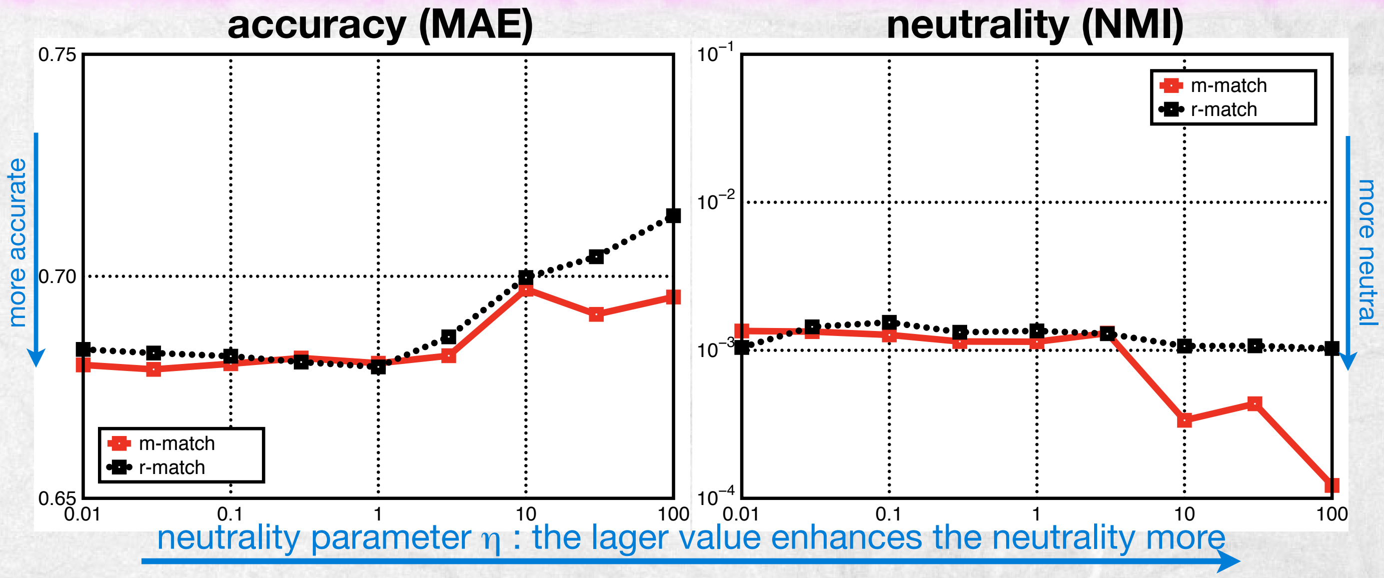
As the increase of a neutrality parameter η,

- accuracies were more steeply worsened in a r-match case
- the neutralities were improved in a m-match, but not in a r-match case

m-match succeed, but r-match failed

These are our experimental results for a Year viewpoint.
As in a case of small data, m-match succeed, but r-match failed.

Large Data & Gender Viewpoint



As the increase of a neutrality parameter η,

- accuracies were similarly worsened in both cases
- neutralities were successfully improved in a m-match case, but not in a r-match case

the m-match term performed slightly better than the small data case

For a Gender viewpoint, the m-match term performed slightly better than the small data case.

Discussion about the Recommendation Neutrality

We finally discuss the recommendation neutrality.

Objectivity and Subjectivity

Objective Neutrality
The neutrality is currently evaluated by objective criteria

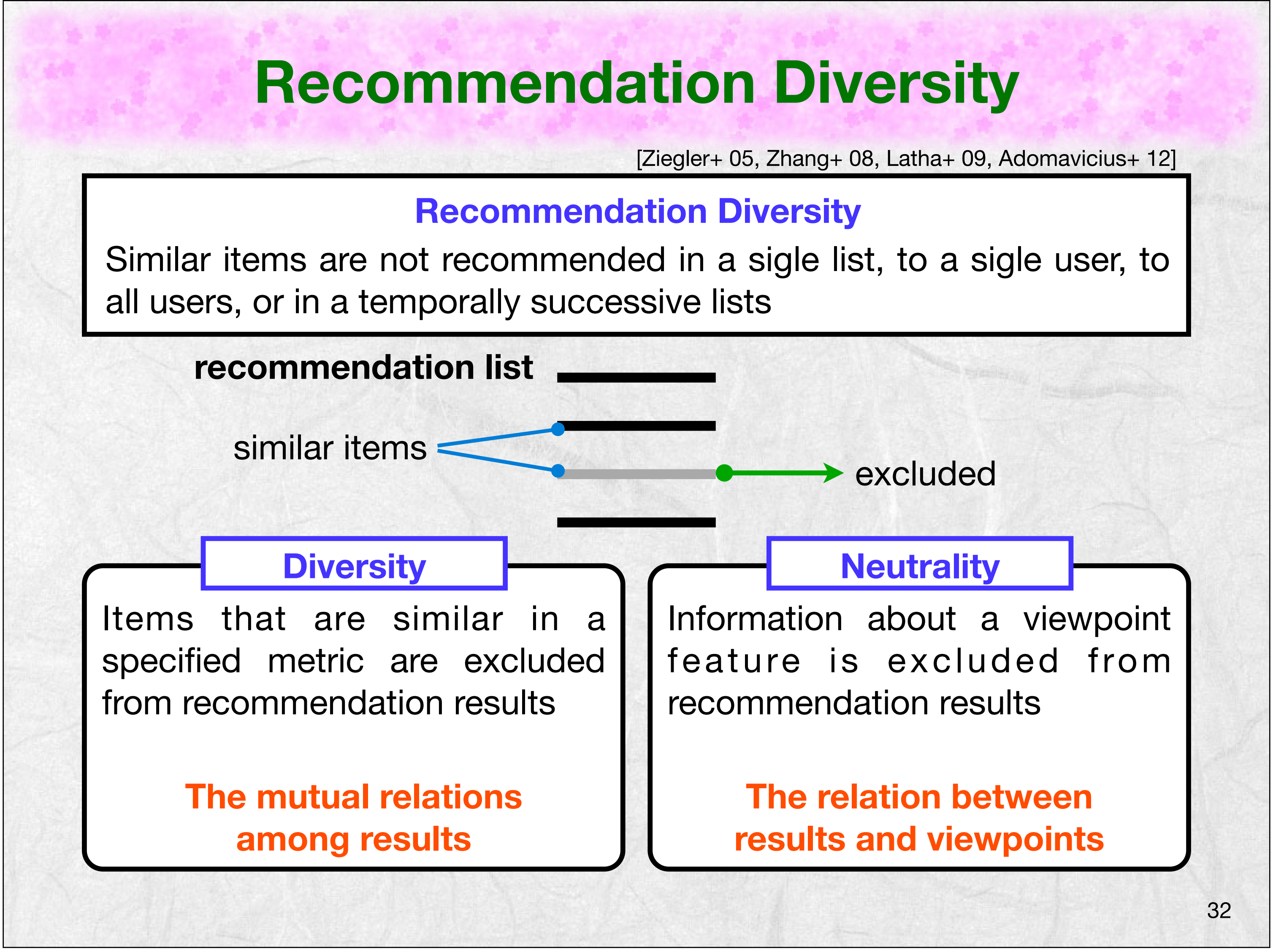
Subjective Neutrality
Reviewers pointed out the importance of testing how users perceive the neutrality

The neutrality should be guaranteed based on objective criteria
If users perceive the neutrality in recommendation, but it is truly biased, such a recommender system would be a very dangerous tool for big brothers to control users

possible tools for displaying the neutrality in recommendation

- showing the neutrality indexes, such as mutual information
- comparing original recommendations and information-neutral ones
- listing items in parallel under the conditions a viewpoint feature is a original and is another value

The neutrality is currently evaluated by objective criteria, but Reviewers pointed out the importance of testing how users perceive the neutrality.
However, in my opinion, the neutrality should be guaranteed based on objective criteria.
This is because if users perceive the neutrality in recommendation, but it is truly biased, such a recommender system would be a very dangerous tool for big brothers to control users.



The recommendation neutrality looks similar to the recommendation diversity, but we consider these two notions are clearly different.
In a case of the diversity, items that are similar in a specified metric are excluded from recommendation results.
In a case of the neutrality, information about a viewpoint feature is excluded from recommendation results.
While the diversity is based on the mutual relations among results, the neutrality is based on the relation between
results and viewpoints.

Privacy-preserving Data Mining

recommendation results, R, and viewpoint features, V,
are statistically independent

mutual information between recommendation results, R,
and viewpoint features, V, is zero
$\mathrm{I}(R; V) = 0$

In a context of privacy-preservation
Even if the information about R is disclosed,
the information about V will not exposed

In particular, a notion of the *t*-closeness has strong connection

The recommendation diversity has connection with privacy-preserving data mining.
The neutrality implies that mutual information between recommendation results and viewpoint feature is zero.
In a context of privacy-preservation, this indicates that even if the information about R is disclosed,
the information about V will not exposed.

More Recommendation Neutrality

Trade-offs between the accuracy and neutrality

Information usable for inferring recommendation decreases in an INRS

$$\mathrm{I}(R; V, F) - \mathrm{I}(R; F) = \mathrm{I}(X; V \mid F) \geq 0$$

where F is all information about other than V

Available information is non-increasing by enhancing the neutrality

red-lining effect

Even if a feature V is eliminated from prediction model, the information of V cannot excluded

the information of V is contained in the other correlated features

Finally, we have two additional comments on the recommendation neutrality.
First, there trade-offs between the accuracy and neutrality, because available information is non-increasing by enhancing the neutrality.
Second, even if a feature V is eliminated from prediction model, the information of V cannot excluded, because the information of V is contained in the other correlated features.

Conclusion

Our Contributions

- We formulate the recommendation neutrality from a specified viewpoint feature
- We developed a recommender system that can enhance the recommendation neutrality
- The efficiency in optimization was drastically improved
- Our experimental results show that the neutrality is successfully enhanced without seriously sacrificing the prediction accuracy

Future Work

- An INRS for non-binary viewpoint features
- Information neutral version of generative recommendation models, such as a pLSA / LDA model

These are our contributions.
We developed a recommender system that can enhance the recommendation neutrality, and the efficiency in optimization was drastically improved.

program codes and data sets

http://www.kamishima.net/inrs/

acknowledgements

- We would like to thank for providing a data set for the Grouplens research lab
- This work is supported by MEXT/JSPS KAKENHI Grant Number 16700157, 21500154, 23240043, 24500194, and 25540094

Not yet updated, but program codes and data sets are available at here.
That’s all I have to say. Thank you for your attention.