Unbiased Recommender Learning from Missing-Not-At-Random Implicit Feedback

Yuta Saito, Suguru Yaginuma, Yuta Nishino,
Hayato Sakata, and Kazuhide Nakata
Web Search and Data Mining (WSDM20)
February 4th, 2020 (Tue)
Introduction & Problem Setting
**Objective of Recommendation**

Recommend **Relevant (R) Items** to Each User!!!

example) Top-3 Recommendation

<table>
<thead>
<tr>
<th>Ranking</th>
<th>Recommender A</th>
<th>Recommender B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R=1</td>
<td>R=0</td>
</tr>
<tr>
<td>2</td>
<td>R=1</td>
<td>R=1</td>
</tr>
<tr>
<td>3</td>
<td>R=1</td>
<td>R=0</td>
</tr>
<tr>
<td>9</td>
<td>R=0</td>
<td>R=1</td>
</tr>
<tr>
<td>10</td>
<td>R=0</td>
<td>R=1</td>
</tr>
</tbody>
</table>

**Recommender A** is better than **Recommender B** simply because **Recommender A** recommends more relevant items.
Ideal Loss function of Interest (Pointwise)

To maximize relevance, the following loss should be optimized.

**Definition) Ideal Pointwise Loss Function**

\[
\mathcal{L}_{\text{ideal}}^{\text{point}}(\hat{R}) = \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \left[ R_{u,i} \delta^{(1)}(\hat{R}_{u,i}) + (1 - R_{u,i}) \delta^{(0)}(\hat{R}_{u,i}) \right]
\]

*Binary Relevance Indicator of u and i*
Ideal Loss function of Interest (Pointwise)

To maximize relevance, the following loss should be optimized:

**Definition) Ideal Pointwise Loss Function**

\[
\mathcal{L}_{\text{ideal}}^{\text{point}}(\hat{R}) = \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \left[ R_{u,i} \delta^{(1)}(\hat{R}_{u,i}) + (1 - R_{u,i}) \delta^{(0)}(\hat{R}_{u,i}) \right]
\]

*Prediction for relevance level of u and i*
Ideal Loss function of Interest (Pointwise)

To maximize relevance, the following loss should be optimized

**Definition) Ideal Pointwise Loss Function**

$$
\mathcal{L}_{\text{ideal}}^{\text{point}}(\hat{R}) = \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \left[ R_{u,i} \delta^{(1)}(\hat{R}_{u,i}) + (1 - R_{u,i}) \delta^{(0)}(\hat{R}_{u,i}) \right]
$$

**Example) Cross-entropy loss**

$$
\delta^{(1)}(\hat{R}_{u,i}) = - \log(\hat{R}_{u,i}), \quad \delta^{(0)}(\hat{R}_{u,i}) = - \log(1 - \hat{R}_{u,i})
$$

*Arbitrary loss function (e.g., cross-entropy, squared loss)*
Challenge: Relevance Label is hard to collect

It is desirable to optimize ideal loss function for our objective of relevance maximization
Challenge: Relevance Label is hard to collect

It is desirable to optimize ideal loss function for our objective of relevance maximization.

However, it is often *Expensive* or *Time Consuming* to use relevance information as the label.

- **Explicit Rating Feedback** (Time Consuming)
- **Expert Annotation** (Expensive, Time Consuming)
- **Crowdsourcing** (Time Consuming, Noisy)
Alternative Solution: Implicit Feedback

**Implicit Feedback** is *Cheap* and *Easy to collect* and used as an alternative for the Relevance Label

Implicit Feedback $Y_{u,i}$

- Natural user behaviour (clicks, views, log-in)
- Easily collected in real-world recommender systems
- Used by many Tech companies
Why not use Implicit Feedback as Relevance Label???

One possible way to use implicit feedback is **direct imputation**

\[
\text{ideal loss} \quad \frac{1}{|D|} \sum_{(u,i) \in D} \left[ R_{u,i} \delta_{u,i}^{(1)} + (1 - R_{u,i}) \delta_{u,i}^{(0)} \right]
\]

\[
\text{imputed loss} \quad \frac{1}{|D|} \sum_{(u,i) \in D} \left[ Y_{u,i} \delta_{u,i}^{(1)} + (1 - Y_{u,i}) \delta_{u,i}^{(0)} \right]
\]

**Neural Collaborative Filtering** (He et al.) optimizes the imputed loss function by DNN
Why not use Implicit Feedback as Relevance Label???

One possible way to use implicit feedback is direct imputation.

**ideal loss**
\[
\frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \left[ R_{u,i} \delta_{u,i}^{(1)} + (1 - R_{u,i}) \delta_{u,i}^{(0)} \right]
\]

**imputed loss**
\[
\frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \left[ Y_{u,i} \delta_{u,i}^{(1)} + (1 - Y_{u,i}) \delta_{u,i}^{(0)} \right]
\]

Question: Is this direct imputation valid?
**Implicit Feedback ≠ Relevance**

example) Top-2 recommendation by most-popular policy

<table>
<thead>
<tr>
<th>Item Ranking</th>
<th>Recommended?</th>
<th>Relvance (R)</th>
<th>???</th>
<th>Click (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes!</td>
<td>R=1</td>
<td></td>
<td>Y=1</td>
</tr>
<tr>
<td>2</td>
<td>Yes!</td>
<td>R=0</td>
<td></td>
<td>Y=0</td>
</tr>
<tr>
<td>_____</td>
<td>_____</td>
<td>_____</td>
<td>_____</td>
<td>_____</td>
</tr>
<tr>
<td>99</td>
<td>No...</td>
<td>R=1</td>
<td></td>
<td>Y=0</td>
</tr>
<tr>
<td>100</td>
<td>No...</td>
<td>R=0</td>
<td></td>
<td>Y=0</td>
</tr>
</tbody>
</table>

It seems
**Implicit Feedback**

is **not**
equal to

Relevance Label
Exposure Model (Liang et al., WWW’16)

Exposure model assumes the following connection between implicit feedback and relevance label:

\[ Y_{u,i} = O_{u,i} \cdot R_{u,i} \]

- **Implicit Feedback** (e.g., click)
- **Exposure Variable** *(unobserved)*
- **Relevance Variable** *(unobserved)*

Item is **clicked** = Item is **exposed** & Item is **relevant**
Exposure Model (Liang et al., WWW’16)

Exposure model also assumes the following decomposition

\[
P(Y_{u,i} = 1) = P(O_{u,i} = 1) \cdot P(R_{u,i} = 1)
\]

- \(P(Y_{u,i} = 1)\): click prob
- \(P(O_{u,i} = 1)\): exposure prob
- \(P(R_{u,i} = 1)\): relevance level

\[
= \theta_{u,i} \cdot \gamma_{u,i}
\]

This assumption is equivalent to the **Unconfoundedness** in causal inference.
Implicit Feedback ≠ Relevance

example) Top-2 recommendation by most-popular policy

<table>
<thead>
<tr>
<th>Item Ranking</th>
<th>Recommended?</th>
<th>Relvance (R)</th>
<th>Exposure (O)</th>
<th>Click (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes!</td>
<td>R=1</td>
<td>O=1</td>
<td>Y=1</td>
</tr>
<tr>
<td>2</td>
<td>Yes!</td>
<td>R=0</td>
<td>O=1</td>
<td>Y=0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>No...</td>
<td>R=1</td>
<td>O=0</td>
<td>Y=0</td>
</tr>
<tr>
<td>100</td>
<td>No...</td>
<td>R=0</td>
<td>O=0</td>
<td>Y=0</td>
</tr>
</tbody>
</table>

Exposure Model can clearly explain the situation
Implicit Feedback ≠ Relevance

eexample) Top-2 recommendation by most-popular policy

<table>
<thead>
<tr>
<th>Item Ranking</th>
<th>Recomended?</th>
<th>Relvance (R)</th>
<th>Exposure (O)</th>
<th>Click (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes!</td>
<td></td>
<td></td>
<td>Y=1</td>
</tr>
<tr>
<td>2</td>
<td>Yes!</td>
<td></td>
<td></td>
<td>Y=0</td>
</tr>
<tr>
<td>99</td>
<td>No...</td>
<td></td>
<td></td>
<td>Y=0</td>
</tr>
<tr>
<td>100</td>
<td>No...</td>
<td></td>
<td></td>
<td>Y=0</td>
</tr>
</tbody>
</table>

The problem is how to optimize R using only Y

Exposure Model characterizes the difficulties
Challenge 1: Positive-Unlabeled (PU)

Only positive-side feedback is observed, and the negative feedback is always unobserved.

\[ Y_{u,i} = O_{u,i} \cdot R_{u,i} \]

\[ Y_{u,i} = 0 \quad \Rightarrow \quad R_{u,i} = 0 \]

The lack of implicit feedback doesn’t imply irrelevance between u and i.
Challenge 2: Missing-Not-At-Random (MNAR)

The positive-labels of some items are much more frequently observed (popularity bias)

\[ P(Y_{u,i} = 1) = P(O_{u,i} = 1) \cdot P(R_{u,i} = 1) \]

Exposure probability is not uniform among user-item pairs
In summary,

● We want to maximize *relevance* in recsys using only available *implicit feedback*

● How to define theoretically justified loss function with implicit feedback is the critical problem

● We aimed to *statistically estimate* the ideal loss func using only implicit feedback in our work
Solutions & Experiments
Our Approach: Unbiased Estimation of Ideal Loss Function

We propose the **first unbiased estimator** combining the inverse propensity weighting & positive-unlabeled learning

**ideal loss**

\[
\frac{1}{|D|} \sum_{(u,i) \in D} \left[ R_{u,i} \delta_{u,i}^{(1)} + (1 - R_{u,i}) \delta_{u,i}^{(0)} \right]
\]

**unbiased loss**

\[
\frac{1}{|D|} \sum_{(u,i) \in D} \left[ \frac{Y_{u,i}}{\theta_{u,i}} \delta_{u,i}^{(1)} + \left(1 - \frac{Y_{u,i}}{\theta_{u,i}} \right) \delta_{u,i}^{(0)} \right]
\]
Our Approach: Unbiased Estimation of Ideal Loss Function

We propose the first unbiased estimator combining the inverse propensity weighting & positive-unlabeled learning

\[
\hat{L}_{\text{unbiased}}(\hat{R}) = \frac{1}{|D|} \sum_{(u,i) \in D} \left[ \frac{Y_{u,i}}{\theta_{u,i}} \delta^{(1)}_{u,i} + \left(1 - \frac{Y_{u,i}}{\theta_{u,i}}\right) \delta^{(0)}_{u,i} \right]
\]

The basic idea is to weight each implicit feedback by the inverse of the exposure parameter (the propensity score)
Our Approach: Unbiased Estimation of Ideal Loss Function

This estimator is proved to be theoretically unbiased for the ideal loss function.

\[
\mathbb{E} \left[ \hat{L}_{\text{unbiased}}^{\text{point}} (\hat{R}) \right] = \mathcal{L}_{\text{ideal}}^{\text{point}} (\hat{R})
\]

The proposed loss function \( \mathcal{L}_{\text{unbiased}}^{\text{point}} (\hat{R}) \)

The ideal loss function \( \mathcal{L}_{\text{ideal}}^{\text{point}} (\hat{R}) \)
Summary of Solutions to the Challenges

Our main contribution is to develop the first unbiased loss func for the ideal loss func using only implicit feedback

<table>
<thead>
<tr>
<th>Approach</th>
<th>Unbiased?</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMF (Hu et al., ICDM’08)</td>
<td>No...</td>
</tr>
<tr>
<td>ExpoMF (Liang et al., WWW’16)</td>
<td>No...</td>
</tr>
<tr>
<td>Rel-MF (saito et al., WSDM’20)</td>
<td>Yes!</td>
</tr>
</tbody>
</table>
Real-World Experiment (with Yahoo! R3 dataset)

We conduct performance comparisons using Yahoo data

Yahoo! R3 dataset

- contains *ground-truth relevance label* (5 star-rating)
- contains train-test data with *different item distributions*

This dataset is convenient for the evaluation of Implicit feedback recommenders with MNAR formulation
Real-World Experiment (with Yahoo! R3 dataset)

The unbiased Rel-MF generally outperforms the others

For all items

<table>
<thead>
<tr>
<th></th>
<th>DCG@5</th>
<th>Recall@5</th>
<th>MAP@5</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMF (Hu et al., ICDM’08)</td>
<td>0.363</td>
<td>0.502</td>
<td>0.277</td>
</tr>
<tr>
<td>ExpoMF (Liang et al., WWW’16)</td>
<td>0.402</td>
<td>0.530</td>
<td>0.321</td>
</tr>
<tr>
<td>Rel-MF (saito et al., WSDM’20)</td>
<td><strong>0.485</strong></td>
<td><strong>0.582</strong></td>
<td><strong>0.407</strong></td>
</tr>
</tbody>
</table>
Real-World Experiment (with Yahoo! R3 dataset)

Ours also outperforms for the rare items

For rare items

<table>
<thead>
<tr>
<th>Model</th>
<th>DCG@5</th>
<th>Recall@5</th>
<th>MAP@5</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMF</td>
<td>0.329</td>
<td>0.526</td>
<td>0.242</td>
</tr>
<tr>
<td>(Hu et al., ICDM’08)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ExpoMF</td>
<td>0.382</td>
<td>0.557</td>
<td>0.307</td>
</tr>
<tr>
<td>(Liang et al., WWW’16)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rel-MF</td>
<td>0.428</td>
<td>0.593</td>
<td>0.345</td>
</tr>
<tr>
<td>(saito et al., WSDM’20)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Conclusions

- Implicit feedback is often used but is biased
  *(positive-unlabeled & missing-not-at-random)*
- Previous solutions are **biased** for the ideal loss function
- We proposed the **first unbiased loss function** for
  *unbiasedly learning recsys from biased implicit feedback*

Thank you for Listening & Please Come to the Poster !!!
Appendix
How to estimate the propensity score?

We used the simple relative item popularity as the propensity score

$$\hat{\theta}_{*,i} = \left( \frac{\sum_{u \in U} Y_{u,i}}{\max_{i \in T} \sum_{u \in U} Y_{u,i}} \right)^{\eta}$$

A more sophisticated way of estimating propensities is a future work
## Previous Solutions to the Challenges

**Weighted Matrix Factorization (WMF)** and **Exposure Matrix Factorization (ExpoMF)** are the most basic methods.

<table>
<thead>
<tr>
<th></th>
<th>Approach</th>
<th>Unbiased?</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMF (Hu et al., ICDM’08)</td>
<td>Positive sample weighting</td>
<td>No...</td>
</tr>
<tr>
<td>ExpoMF (Liang et al., WWW’16)</td>
<td>EM Algorithm</td>
<td>No...</td>
</tr>
</tbody>
</table>
Previous Solutions are biased for the ideal loss func

In the paper, the loss function of the previous methods are proved to be *biased*, i.e.,

\[
\mathbb{E} \left[ \hat{\mathcal{L}}_{WMF}(\hat{R}) \right] \neq \mathcal{L}_{ideal}^{\text{point}}(\hat{R})
\]

\[
\mathbb{E} \left[ \hat{\mathcal{L}}_{ExpMF}(\hat{R}) \right]
\]
Future Work

- Propensity score estimation
- Unbiased estimator for the pairwise method
  (e.g., unbiased version of bayesian personalized ranking)
- Theoretical Analysis on the Learnability
- Possible connection with other types of feedback


References


(Liang et al., UAI’16 Causal WS): Dawen Liang, Laurent Charlin, and David M Blei. 2016. In Causation: Foundation to Application, Workshop at UAI.