

1. Recommendation systems/Collaborative filtering

Recommendation systems have immense utility in commercial applications. Numerous systems involving users and products, for example movie and music-streaming applications such as Netflix and Pandora and product/consumer systems like Amazon, employ recommender systems. These systems attempt to recommend new products to users based on their purchase history or some other calculation of preferences.

Many recommender systems with a large number of users and products use collaborative filtering methods, which often utilize information from many users to calculate preferences for product-item pairs of a different user. These systems are normally equipped with a rating system, where users explicitly rate items either positively or negatively.

2. Implicit feedback algorithm

However, not all currently developed algorithms require explicitly defined user ratings. One method developed by [Hu et al.](#) (2008), is designed specifically for implicit feedback. Unlike explicit feedback such as ratings, implicit feedback is an indirect measure of user preferences based on behaviors. For example, abundant implicit feedback can be gathered from a system involving users and television programs. Normally users have no way of explicitly rating a television program. However, we can infer from their behavioral patterns that a user may prefer a certain program if he or she watches it several times. [Hu et al.'s](#) method was specifically developed for a TV shows recommender engine.

Julia Jones 8/16/14 9:27 PM

Deleted: Hu, Koren, and Volinsky's

Although implicit feedback systems are similar to explicit systems, there exist important distinctions. Because users never explicitly rate products, it's difficult to infer which items a user does not like. For example, a user may not watch a certain show for several reasons, including time conflicts or lack of knowledge about the program. Even shows that the user did watch do not necessarily imply that the user enjoys that show. Perhaps the user was away from the television at that time, or was watching the show preceding a favorite program.

Julia Jones 8/16/14 9:27 PM

Deleted: Hu, Koren, and Volinsky's

For this reason, unlike explicit feedback systems where the numerical value of a user-item interaction is an indication of preference, the numerical value in implicit feedback systems indicates a confidence of the interaction. For implicit systems, a large value doesn't imply a higher preference, only a higher frequency of interaction. While it would difficult to say whether a user prefers a show seen very few times, we could confidently say a user prefers a show seen several times.

3. Application to plant-pollinator network

The implicit feedback system developed by [Hu et al. \(2008\)](#) can be applied to plant-pollinator networks. Like the TV show scenario, we have no way of observing any explicit rating of plant and flower preferences. However, a high-frequency interaction grants a degree of confidence for the preference of that interaction. Also, just like with seasonality in TV programming, plants are subject to seasonality and [Hu et al.](#)'s core model is robust enough to handle this seasonality.

Julia Jones 8/16/14 9:28 PM
Deleted: Hu, Koren, and Volinsky

4. Algorithm implementation

I have implemented the algorithm detailed in Collaborative Filtering for Implicit Feedback Datasets ([Hu et al. 2008](#)), called IFMF, using the R programming language. This method utilizes matrix factorization, which decomposes the original matrix into two latent factors. Besides the matrix of interactions, the method also generates a confidence matrix, which is a logarithmic reduction of the interaction matrix, and an original preference matrix, which is a binary representation of the interaction matrix.

Julia Jones 8/16/14 9:28 PM
Deleted: Hu, Koren, and Volinsky

Julia Jones 8/16/14 9:28 PM
Deleted: - ... [1]

Julia Jones 8/16/14 9:28 PM
Deleted: Hu, Koren, Volinsky

Julia Jones 8/16/14 9:30 PM
Comment: What is the interpretation of the two "latent factors"?

5. Testing/Training

Currently, the most apparent complication of using IFMF is the lack of complete and accurate testing and training data sets. For the TV show recommender engine, several months of complete data were used to train the parameters of IFMF, and several more months of complete data were used to validate that training.

Julia Jones 8/16/14 9:28 PM
Deleted: -

However, with current observation methods we suspect it is impossible to collect complete data of the plant-pollinator interactions. For example, if IFMF predicts that a certain unobserved interaction is likely to occur, it can be difficult to validate the occurrence of that interaction.

Julia Jones 8/16/14 9:28 PM
Deleted: -

Rebecca and I have discussed two possible solutions to this complication. First, we can divide the data by subplots, using some for training and some for testing. If our algorithm can predict interactions that occurred in the data but not in the subset of plots used for training, we can conclude with some confidence that our parameters are well tuned.

Julia Jones 8/16/14 9:30 PM
Deleted: -

Alternatively, we could intentionally remove a number of known interactions in the training set. We could then tune our algorithm to predict the removed interactions based on the remaining interactions. Rebecca has suggested that a combination of these two approaches may be the most appropriate and robust.

Julia Jones 8/16/14 9:30 PM
Deleted: -

6. Analysis of results

The TV show recommender engine recommends those shows not previously watched with the highest computed preferences to the user. In our plant-pollinator scenario, we could predict which observations we may have missed in our data-collection.

With generated predictions/preferences, there are several comparisons of results that could be interesting. First, we could see the effect of different parameter tunings on the predictions. Also, we could compare the results to a prediction method that only recommend the most popular plants and flowers. The distinction between these two prediction methods could have some interesting ecological implications.

Also, we could compare the results of IFMF with the method Rebecca has been working on. Her method also takes into account flower abundances and negative implicit feedback. Basically, her method works on the assumption that the lack of an interaction with an abundant plant implies a negative preference.

Finally, we can compare our results with Andy Moldenke's preference matrix, which he derived from ecological intuition and experience.

6. Goals for this week

- Produce some initial results and visualization figures
- Clearly restate the ideas written here along with (a) in a detailed paper
- Develop an informative and general-audience presentation for this Thursday

7. Future work (math thesis project?)

- Perform the comparisons mentioned in (5).
- Recommendation explanations: We can algorithmically discover which interactions collaboratively lead to a certain prediction. For newly discovered interactions, it would be interesting to see which past interactions led to the prediction of that interaction.
- Weight the preference matrix by insect/plant size: I hypothesize that the frequency of missing an interaction in data collection is correlated with insect and plant size. Therefore to accurately predict which interactions we've missed, we should weight our results by insect/plant size.
- Further account for plant seasonality

Julia Jones 8/16/14 9:31 PM

Deleted: -

•
•
•

Julia Jones 8/16/14 9:31 PM

Comment: interactions

Julia Jones 8/16/14 9:31 PM

Deleted: -

•
•
•

Julia Jones 8/16/14 9:32 PM

Comment: of a prediction including rare interactions to?

Julia Jones 8/16/14 9:33 PM

Comment: do you mean plants and pollinators?

Julia Jones 8/16/14 9:33 PM

Comment: because one focuses on all interactions including rare ones, whereas the other focuses only on common plants and pollinators?

Julia Jones 8/16/14 9:32 PM

Deleted: -

•
•
•

Julia Jones 8/16/14 9:33 PM

Deleted: -

•
•
•

Julia Jones 8/16/14 9:33 PM

Comment: did you find a copy of this? I would like to see it.

Julia Jones 8/16/14 9:33 PM

Deleted: -

•
•
•

Unknown

Deleted: -

Julia Jones 8/16/14 9:34 PM

Deleted: -

•
•
•