

# A Hybrid Content-Collaborative Reciprocal Recommender for Online Dating

Joshua Akehurst

School of Information Technologies  
Faculty of Engineering & Information Technologies

## Problem

Online dating sites are used by millions of people. Our overall goal is to improve matching between people in online dating.



## Motivation

Traditional recommender systems are *items-to-people*, i.e. they recommend items such as books, movies and restaurants to people. They are not successful in *people-to-people* or *reciprocal* domains, such as online dating, matching job applicants with jobs and matching mentors with mentees, which require satisfying the preferences of both people involved.

Reciprocal recommenders [1] are a novel type of recommender systems which uses the preferences of both people.

## Contribution

We propose a new reciprocal recommender [2] which combines content-based and collaborative filtering approaches. It uses both user profiles and user interaction to recommend potential matches for a user of an online dating service.

## Dataset

The recommender was experimentally tested using a large dataset provided by a major Australian online dating website. The dataset consists of all users who have send an *expression of interest (EOI)* to at least one other user or responded to an EOI within a one month period (March 2010). This dataset was further restrained to users from Sydney.

## Approach

**Do similar people like/dislike similar people and are liked/disliked by similar people?**

Group	Computed similarity		Baseline similarity	
	Female	Male	Female	Male
U Likes	0.496	0.545	0.031	0.039
U Dislikes	0.447	0.404	0.021	0.015
U is Liked By	0.439	0.527	0.019	0.028
U is Disliked By	0.462	0.519	0.031	0.052

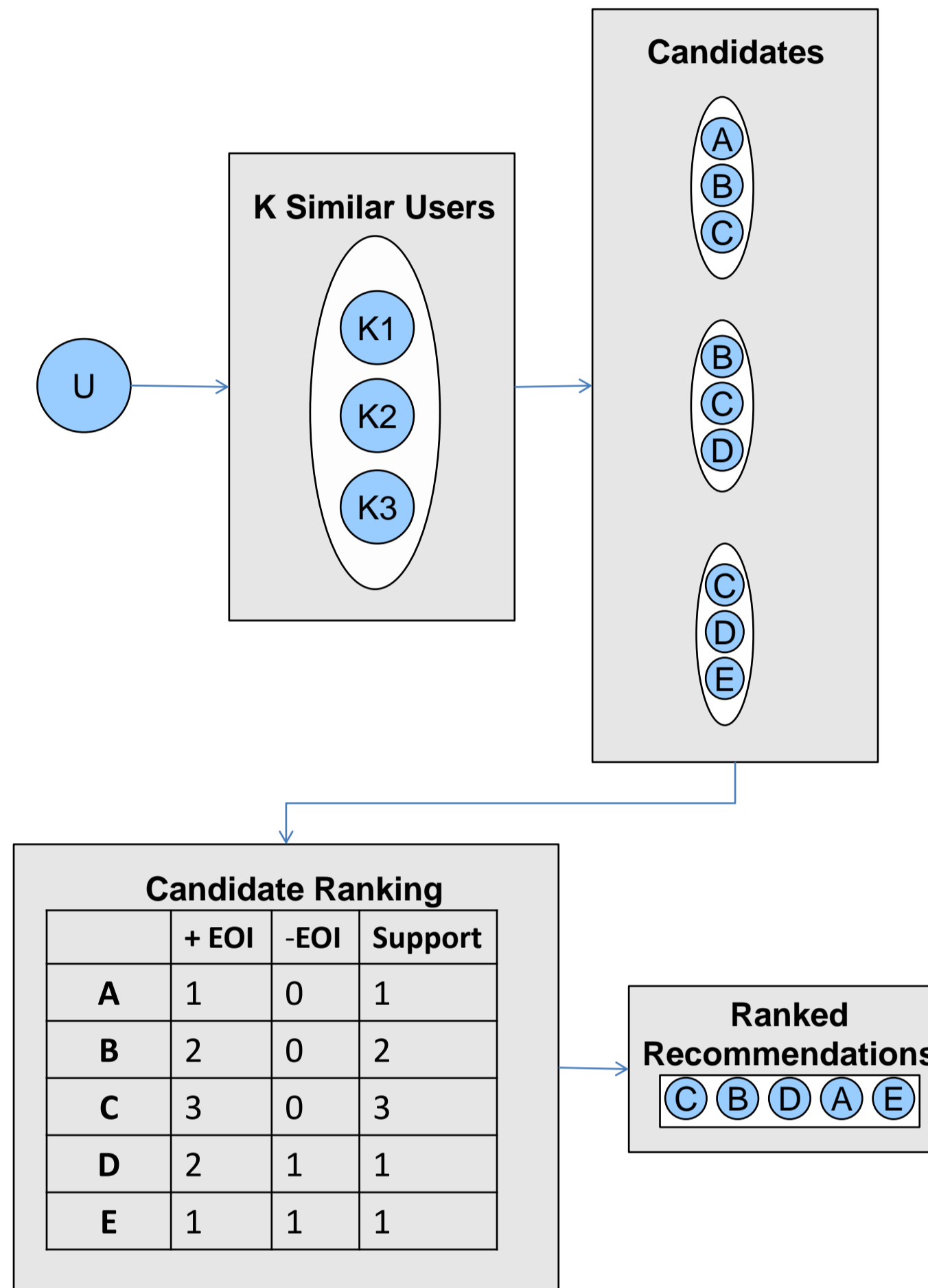
**Yes!**

For every pair of users, using a selected subset of profile attributes and distance measure, we computed the correlation between the: a) distance between the users and b) distance between the centroid of the group of users they liked/disliked and are liked/disliked by.

## Recommendation Algorithm

To generate potential matches for a user U:

- Find the K most similar users based on user profile
- For each of these K similar users, compute a candidate set of users with reciprocal interest in them using interaction data
- Rank the candidates by calculating support



## Efficient Generation of the Set of Similar Users Using a Double Constraint

- Start with a minimum distance threshold
- Consider users from training set in random order and compute the distance between them and the test user. If the distance is below the threshold, add them to the set of similar users.
- Once all users in the training set have been considered and the K-C constraint is not satisfied, increment the threshold and repeat. The distances do not need to be recomputed.
- Continue until at least K users have been added to the set of similar users and the size of the candidate list is at least C. This process is efficient as most users satisfy the double constraint at low thresholds, especially when the training sets have a large number of users.

## Evaluation Metrics

Success rate, over all users in test set:

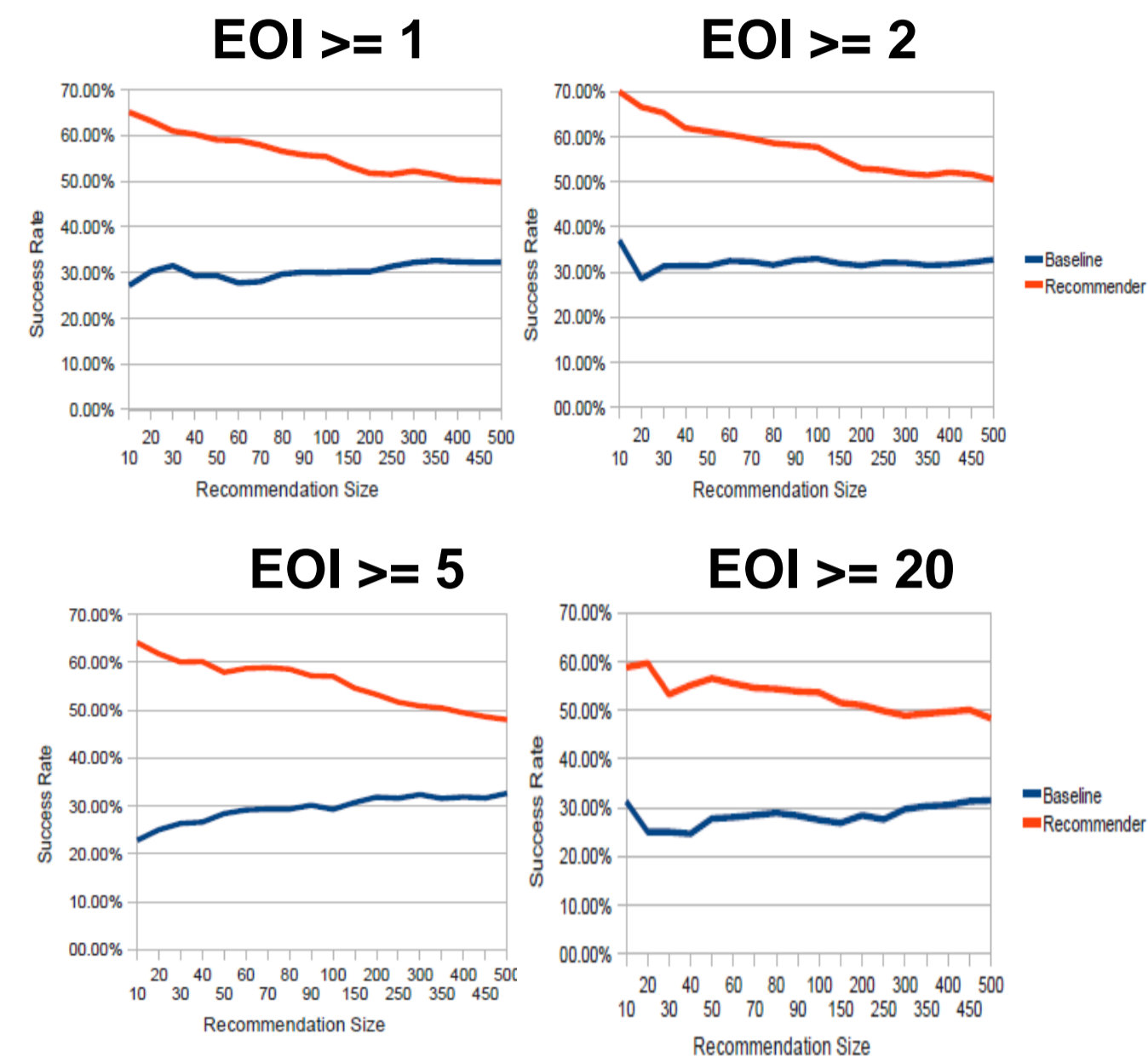
$$\text{Success rate} = \frac{\text{positive\_EOI\_sent/recv}}{\text{interacted\_recommendations}}$$

•  $\text{positive\_EOI\_sent/recv} = \text{EOI sent by U to other users which were replied positively and EOI received by U from other users which were replied positively by U.}$

•  $\text{interacted\_recommendations} = \text{all EOI sent by U or received by U.}$

**Baseline success rate:** Using a random set of K users for each user instead of the K most similar users.

## Results



• Our recommender is much more accurate than the baseline in all cases. E.g. when the top 10 recommendations are considered, the success rate of the recommender is between 60.16% and 69.26% compared with a baseline of 35.19% and 30.30%.

• As the number of recommendations shown to the user increases from 10 to 500, the success rate decreases by 10-20%. Thus, the successful recommendations are shown at the top of the recommendation list and our ranking criterion is effective.

• The success rate trends are very similar for different EOI\_sent.

## Conclusion

• Correlation analysis shows that people with similar profiles like and dislike similar people and are liked and disliked by similar people.

• This provided the basis for our novel hybrid recommendation approach, which combines content-based and collaborative filtering, and utilizes data from both user profiles and user interactions.

• Our recommender significantly outperforms the baseline; the success rate of our recommender is [69.26% - 60.16] for different number of EOI, compared with the baseline of [35.19% - 30.30%], respectively.

• Our recommender addresses the cold start problem of new users joining the site by being able to provide recommendations immediately, based on the new user profile.

## References

- [1] L. Pizzato, T. Rej, T. Chung, I. Koprinska, J. Kay: RECON: A Reciprocal Recommender for Online Dating. *ACM Conference Recommender Systems (RecSys '2010)*.  
[1] J. Akehurst, K. Yacef, I. Koprinska, L. Pizzato, J. Kay, T. Chung, T. Rej: A Hybrid Content-Collaborative Reciprocal Recommender for Online Dating, PAKDD'2010 (submitted)

## Acknowledgement

This research was supported by the Smart Services Co-operative Research Centre.