CS6120: Intelligent Media Systems

Dr. Derek Bridge
School of Computer Science & Information Technology
UCC

Motivations

- Why do we want an alternative to content-based approaches?
  - Suppose items do not have readily available descriptions or their descriptions fail to capture the subjective experience of consuming the items
    - e.g. music, video, poems, art, photographs, jokes
  - Content-based recommenders rarely make recommendations that extend our tastes
    - we would like *serendipitous* recommendations

Intuitions: Sharing Opinions

- Ann asks her friends whether she should see the latest Hollywood release
- Ben recommends it
  - but he seems to recommend everything
- Clare doesn't think much of it
  - and she has a habit of recommending things Ann likes
- Dan hated it
  - but he hates all Hollywood movies
- …
- Over time, Ann learns whose opinions can be applied to help her determine the quality of items

Intuitions: Sharing Opinions

- In 'word-of-mouth' recommendations
  - we take into account how similar the other person's tastes are to our own
  - item descriptions are not needed
- *Collaborative recommendations*: evaluating items using the opinions of other people
  - automates word-of-mouth
  - but, through the web, we can access the opinions of thousands of people
  - recommendations are based on the opinions of many similar users rather than a small group of friends

Ratings Matrix

	Alien	Brazil	Crash	Dumbo	E.T.	Fargo
Ben		2	5	3	1	2
Clare	5	5		3	4	
Dan					3	
Edd	5	4	2	4	3	3
Flo	2	5	4	4		
Ann	**2**			**4**	**3**	**5**

Notation

$\mathbb{U}$	the set of all users
$\mathbb{I}$	the set of all items
$\mathbb{U}_i$	the set of users who have rated item i
$\mathbb{U}_{i,j}$	the set of users who have rated both item i and item j
$\mathbb{I}_u$	the set of items that have been rated by user u
$\mathbb{I}_{u,v}$	the set of items that have been rated by both user u and user j
$r_{u,i}$	user u's *actual* rating for item i
$\widehat{r_{u,i}}$	user u's *predicted* rating for item i

Collaborative Recommenders

- Typically, collaborative recommendation is regression:
  - predict a rating for each candidate item
- Recommend the items with highest predicted ratings (in descending order of predicted rating)
- Many ways to construct the regression system
  - again we look at *k*-nearest-neighbours (*kNN*)

k-Nearest Neighbours for User-Based Collaborative Recommending

for each candidate item, i
    for each user v in $\mathbb{U}_i$ except for active user u, i.e. other users who have rated i
        compute $sim(u, v)$
    let NN be the set of k nearest neighbours, i.e. the k users who have rated i and who are most similar to u and whose similarity to u is positive
    let $\widehat{r_{u,i}}$ be the *weighted average* of the ratings for i in NN
recommend the candidates in descending order of predicted rating

How Many Neighbours?

- What should k be?
  - if too small, predictions overly influenced by a few users
  - if too large,
    - predictions are based on the opinions of ever less similar users
- Typically, $k = 50$
  - or some value between say 20 and 100 – determined from experiments
- Note you may not always find k neighbours

User-User Similarity, $sim(u, v)$

- Many systems use Pearson Correlation:

$$sim(u,v) = \frac{\sum_{i \in \mathbb{I}_{u,v}} (r_{u,i} - \bar{r_u})(r_{v,i} - \bar{r_v})}{\sqrt{\sum_{i \in \mathbb{I}_{u,v}} (r_{u,i} - \bar{r_u})^2} \sqrt{\sum_{i \in \mathbb{I}_{u,v}} (r_{v,i} - \bar{r_v})^2}}$$

- Notice how it is computed over the co-rated items, $\mathbb{I}_{u,v}$
  - $\bar{r_u}$ is u's average rating, *over the co-rated items*
  - $\bar{r_v}$ similarly for v
- Pearson correlation will be 1.0 for users in perfect agreement and -1.0 for users in perfect disagreement

Example: $sim(Ann, Ben)$

$$sim(u,v) = \frac{\sum_{i \in \mathbb{I}_{u,v}} (r_{u,i} - \bar{r_u})(r_{v,i} - \bar{r_v})}{\sqrt{\sum_{i \in \mathbb{I}_{u,v}} (r_{u,i} - \bar{r_u})^2} \sqrt{\sum_{i \in \mathbb{I}_{u,v}} (r_{v,i} - \bar{r_v})^2}}$$

	Alien	Brazil	Crash	Dumbo	E.T.	Fargo
Ben		2	5	3	1	2
Ann	**2**			**4**	**3**	**5**

Exercise: $sim(Ann, Clare)$

$$sim(u,v) = \frac{\sum_{i \in \mathbb{I}_{u,v}} (r_{u,i} - \bar{r_u})(r_{v,i} - \bar{r_v})}{\sqrt{\sum_{i \in \mathbb{I}_{u,v}} (r_{u,i} - \bar{r_u})^2} \sqrt{\sum_{i \in \mathbb{I}_{u,v}} (r_{v,i} - \bar{r_v})^2}}$$

	Alien	Brazil	Crash	Dumbo	E.T.	Fargo
Clare	5	5		3	4	
Ann	**2**			**4**	**3**	**5**

Similarity: Niceties

- People rate differently
  - can normalize their ratings, e.g. z-scores
- Note how Pearson correlation does something similar already
  - subtracts averages, divides by standard deviations
- But it is different
  - it uses averages and standard deviations over co-rated items only

Similarity: Niceties

- If one user gives the same rating to all the co-rated items
  - his/her standard deviation is zero
  - a special case of this is when there is only one co-rated item
- To avoid a division by zero problem, we simply take Pearson Correlation to be zero in such cases
  - e.g. $sim(Ann, Dan)$

Similarity: Niceties

- What if the number of co-rated items is small?
  - it's easy for two users to agree on a small number of items
  - but we can't trust such Pearson Correlation values
- Significance weighting
  - decrease the similarity of two users with few co-rated items
  - E.g. if $\mathbb{I}_{u,v}$ contains n items and $n < 50$, multiply $sim(u, v)$ by $^{n}/_{50}$

u's Predicted Rating for i, $\widehat{r_{u,i}}$

- We could just take an average:
$$\widehat{r_{u,i}} = \frac{\sum_{v \in NN} r_{v,i}}{k}$$
- But, as before, we want to take into account that some neighbours are more similar to u than others
  - their ratings should contribute more to the prediction
  - so take a weighted average
$$\widehat{r_{u,i}} = \frac{\sum_{v \in NN} sim(u,v) \times r_{v,i}}{\sum_{v \in NN} sim(u,v)}$$

u's Predicted Rating for i, $\widehat{r_{u,i}}$

- But people use the rating scale differently
  - e.g. a rating of 4 from one user means the same as a 5 from another
  - so include an adjustment for users' average ratings
$$\widehat{r_{u,i}} = \overline{r_u} + \frac{\sum_{v \in NN} sim(u,v) \times (r_{v,i} - \overline{r_v})}{\sum_{v \in NN} sim(u,v)}$$
  - here $\overline{r_u}$ is u's average rating, over all his/her ratings; similarly $\overline{r_v}$

Predictions: Niceties

- The adjustment is similar to using z-scores
  - but it is different because it ignores standard deviations
- Note that this formula may produce non-integer results
  - so you may want to round the predictions
- Note that this formula may produce results that fall off the rating scale (bigger than 5, less than 1)
  - if so, round down or up to the nearest end-point

Example, part I

	Alien	Brazil	Crash	Dumbo	E.T.	Fargo
Ben		2	5	3	1	2
Clare	5	5		3	4	
Dan					3	
Edd	5	4	2	4	3	3
Flo	2	5	4	4		

	Alien	Brazil	Crash	Dumbo	E.T.	Fargo
Ann	**2**			**4**	**3**	**5**

- We'll predict Ann's rating for Brazil
- Question: We must compute the similarity between Ann and…who? (Why?)

Example, part II

- The similarities
  - $sim(Ann, Ben) = 0.5$
  - $sim(Ann, Clare) = -1.0$
  - $sim(Ann, Edd) = -0.6742$
  - $sim(Ann, Flo) = 1.0$
- Suppose $k = 3$
- Ann's 3 nearest neighbours are Flo, Ben and Edd
- But we ignore Edd. Why?

Example, part III

	Alien	Brazil	Crash	Dumbo	E.T.	Fargo
Ben		2	5	3	1	2
Flo	2	5	4	4		

	Alien	Brazil	Crash	Dumbo	E.T.	Fargo
Ann	**2**			**4**	**3**	**5**

$$\widehat{r_{u,i}} = \overline{r_u} + \frac{\sum_{v \in NN} sim(u,v) \times (r_{v,i} - \overline{r_v})}{\sum_{v \in NN} sim(u,v)}$$

Variations of the kNN User-Based Collaborative Recommender

- Normalization
  - z-scores or not
- Measuring similarity
  - Pearson Correlation with or without Significance Weighting
  - Cosine Similarity (or Adjusted Cosine Similarity) (with 'blanks' set to zero)
    - found to be as accurate as using Pearson Correlation without the need for Significance Weighting
- Finding neighbours
  - use a value k, or a threshold θ
- Making predictions
  - similarity-weighted or not
  - mean-centered or not

Discussion

- The big advantage is that we don't need item descriptions, just user-item ratings
  - we may even obtain these implicitly

- This means that collaborative recommenders work in 'subjective' domains
  - books, movies, music, pieces of art,…

Discussion

New items

- When a new item becomes available

  - it *cannot* be recommended immediately
  - we must wait for people to rate it
  - this is the *cold-start problem*
  - contrast with content-based recommenders

New users

- When a new user joins,
  - no recommendations can be made to him/her until s/he has built a user profile
    - Either rating some items when s/he registers
    - or rating items while using the system
  - similar to content-based recommenders

Discussion

- Collaborative recommenders can extend our tastes
  - If our neighbours have taste combinations that are new and pleasing to us
  - *serendipitous recommendations*

Explanations

- Often worth providing a facility for obtaining explanations of recommendations
  - especially for high risk items
    - high purchase cost/high consumption cost
  - in all of the kinds of recommenders
    - user-based collaborative, item-based collaborative, content-based, etc.
- In effect, the user is asking "Why?"
  - the explanation provides an answer

The Goal of an Explanation

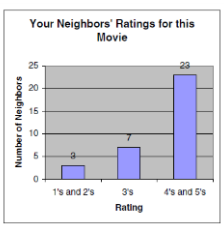
Transparency	Explain how the system works
Scrutability	Allow users to make corrections
Trust	Increase users' confidence in the system
Effectiveness	Help users make good decisions
Persuasiveness	Convince users to purchase/consume
Efficiency	Help users make decisions faster
Satisfaction	Increase the ease of use or enjoyment

Based on Table 15.1 in
N. Tintarev and J. Masthoff: "Designing and Evaluating Explanations for Recommender Systems", in F. Ricci et al. (eds.), *Recommender Systems Handbook*, Springer, pp.479-510, 2011

Explanations of kNN User-Based Collaborative Recommendations

- Explanations can exploit the fact that the algorithm works in an easily-understood way
  - automates word-of-mouth
- E.g. Herlocker et al. compared 21 explanations
  - evaluated primarily for persuasiveness
  - found that simple visualizations of supporting data were the most persuasive

Herlocker et al: Two Persuasive Explanations

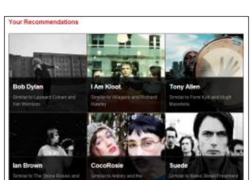


J. L. Herlocker, J. A. Konstan and J. Riedl: "Explaining collaborative filtering r ecommendations", in *Proceedings of the 2000 ACM conference on Computer Supported Cooperative Work*, ACM, pp.241-250., 2000

Other Explanations

- Amazon:
  - item-based collaborative recommender

- Last.fm:
  - either item-based collaborative or content-based recommender (using tags)



Attacks

- Hackers may breach security
  - for personal gain, e.g. to steal private data or disrupt the service of a competitor
  - as a protest
  - 'non-maliciously', e.g. to reveal a security flaw
  - …
- But the attacks we are interested in are ones that 'game' the system
  - *push attacks*
    - inject fake user profiles to promote your own product
  - *nuke attacks*
    - inject fake user profiles to demote your competitor's products

Example: Before Attack Predict Ann's Rating for *Gravity*

	Alien	Brazil	Crash	Dumbo	E.T.	Fargo	Gravity
Ben	+	-		+	+		+
Clare	-	+	+	-	-		-
Dan	+	-	+		-	-	-
Edd	-	+	+	-			
Flo	-		-	-	-		-
Guy	+	-	+	+	+		+
Helen		-	+	+	-	-	+
Ann	+	-	+	+	+		

Based on Fig 25.2 in
R. Burke, M. O. O'Mahony & N. J. Hurley: "Robust Collaborative Recommendation"
In F. Ricci et al (eds.), Recommender Systems Handbook, Springer, pp.805-835, 2011

Example: After Attack Predict Ann's Rating for *Gravity*

	Alien	Brazil	Crash	Dumbo	E.T.	Fargo	Gravity
Ben	+	-		+	+		+
Clare	-	+	+	-	-		-
Dan	+	-	+		-	-	-
Edd	-	+	+	-			
Flo	-		-	-	-		-
Guy	+	-	+	+	+		+
Helen		-	+	+	-	-	+
$Fake_1$	+	-	+		-	-	-
$Fake_2$	-	+	+	-			-
$Fake_3$	-		-	-	-		-
$Fake_4$	+	-	+	+	+		-
$Fake_5$		-	+	+	-	-	-
Ann	+	-	+	+	+		

It's an Arms Race

Attacking

- Effort
  - creating fake accounts
  - populating them with ratings
- Knowledge
  - must discover how the recommender works
  - must discover what kinds of profiles are associated with the item you want to attack

Defence

- Increase the effort
  - for account creation: captchas, email verification,…
- Detect unusual patterns of activity, e.g.:
  - many accounts being created from same IP address
  - an unusual shift in an item's distribution of ratings

Fake Reviews

- Estimates that 20-30% of user-submitted reviews on sites such as Amazon, Yelp, TripAdvisor may be fake
  - written by friends, family, employees
  - or written in exchange for payment
- Sites such as Yelp claim to have filters to identify fake reviews using signals such as
  - the reviewer has few or no other reviews
  - the review is short or vague
  - the star rating is the min or max
  - the wording duplicates that of other reviews

Truth or Spoof?

- "I want to make this review in order to comment on the excellent service that my mother and I received on the Serenade of the Seas, a cruise line for Royal Caribbean. There was a lot of things to do in the morning and afternoon portion for the 7 days that we were on the ship. We went to 6 different islands and saw some amazing sites! It was definitely worth the effort of planning beforehand. The dinner service was 5 star for sure. One of our main waiters, Muhammad was one of the nicest people I have ever met. However, I am not one for clubbing, drinking, or gambling, so the nights were pretty slow for me because there was not much else to do. Either than that, I recommend the Serenade to anyone who is looking for excellent service, excellent food, and a week full of amazing day-activities!"

http://www.cs.uic.edu/~liub/FBS/fake-reviews.html

Truth or Spoof?

- "This movie starring big names - Tom Hanks, Sandra Bullock, Viola Davis, and John Goodman - is one of the most emotionally endearing films of 2012. While some might argue that this film was "too Hollywood" and others might see the film solely because of the cast, it is Thomas Horn's performance as young Oskar that is deserving of awards. The story is about a 9-year-old boy on a journey to make sense of his father's tragic death in the 9/11 attacks on the World Trade Center. Oskar is a bright and nervous adventurer calmed only by the rattle of a tambourine in his ear. "I got tested once to see if I had Asperger's disease," the boy offers in explain of his odd behavior. "The tests weren't definitive." One year after the tragedy, Oskar finds a key in his father's closest and thus begins a quest to find the missing lock. Oskar's battle to control his emotional anxiety and form and mend relationships proves difficult, even with his mother. "If the sun were to explode, you wouldn't even know about it for eight minutes," Oskar narrates. "For eight minutes, the world would still be bright and it would still feel warm." Those fleeting eight minutes Oskar has left of his father make for two hours and nine minutes of Extremely Emotional and Incredibly Inspiring film. Leaving the theatre, emotionally drained, it is a wonder where a movie like this has been. We saw Fahrenheit 9/11 and United 93, but finally here is the story of a New York family's struggle to understand why on "the worst day" innocent people would die. I highly recommend this movie as a must see."

http://www.cs.uic.edu/~liub/FBS/fake-reviews.html

Truth or Spoof?

- "High Points: Guacamole burger was quite tall; clam chowder was tasty. The decor was pretty good, but not worth the downsides. Low Points: Noisy, noisy, noisy. The appetizers weren't very good at all. And the service kind of lagged. A cross between Las Vegas and Disney world, but on the cheesy side. This Cafe is a place where you eat inside a plastic rain forest. The walls are lined with fake trees, plants, and wildlife, including animatronic animals. A flowing waterfall makes sure that you won't hear the conversations of your neighbors without yelling. I could see it being fun for a child's birthday party (there were several that occurred during our meal), but not a place to go if you're looking for a good meal."

http://www.cs.uic.edu/~liub/FBS/fake-reviews.html