Understanding and Predicting Online Food Recipe Production Patterns

Tomasz Kuśmierczyk, Christoph Trattner, Kjetil Nørvåg

NTNU – Trondheim
Norwegian University of Science and Technology
Context: On-line food communities
Project

Studying Food Consumption and Production Patterns on the Web:

○ T. Kusmierczyk, C. Trattner, K. Nørvåg: Temporality in Online Food Recipe Consumption and Production. WWW 2015.
○ T. Kusmierczyk, C. Trattner, K. Nørvåg: Understanding and Predicting Online Food Recipe Production Patterns. HT 2016.
Goal: study factors influencing production

- consumed information
- social connections
- historical factors
- time
- etc.

- food type (e.g., pizza)
- ingredients (e.g., tomatoes)
Bias towards friends
Popularity in time

![Graph showing the fraction of recipes containing sugar and onion over time from 2009 to 2014.

- Blue line represents sugar, while the red line represents onion.
- The dominating ingredient is highlighted for each year.

The graph indicates a fluctuation in the popularity of recipes containing sugar and onion over the years, with certain periods showing a dominance of one ingredient over the other.]

7
How categories can help with data sparsity
Application: recommend food types and ingredients
Sample Web app & Problem formulation

Upload Recipe

Recipe:
- Chili Con Carne
- Meat Balls
- Paprika Goulash

Ingredients for 4 People

<table>
<thead>
<tr>
<th>Amount</th>
<th>Unit</th>
<th>Ingredient</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>Gram</td>
<td>Beef</td>
</tr>
</tbody>
</table>

Recipe Type Recommender

Ingredients Recommender
- Beef
- Chili
- Salt
- Pepper
- Water
- Paprika
- Oil
Solution: recommend list of items

- **observe patterns**

- design a personalized **scoring function**: $S(u,e)$

- **rank items** (ingredients, types) according to scores
Sample scoring function

Item popularity in uploads by friends:

\[ F(u, e) = \sum_{f \in F_u} \sum_{r \in U_f} [e \in r] \]

1 iff recipe \( r \) contains \( e \)
Results
Evaluation: Food types prediction
Evaluation: **Ingredients prediction**

**CATEGORIES OF RATINGS**

**FOOD TYPE UNKNOWN**

**FOOD TYPE KNOWN**

![Graph showing nDCG@10 for different predictors and categories of ratings for food type unknown and known.](image-url)
• only some fraction of users have **strong preferences** when producing on-line content

• **consumption** (ratings), **social factors** (friendship) correlate strongly with production (uploads)

• **categorical information** ‘smoothing’ improves significantly prediction quality
Not covered in the presentation

- dependencies between food types and ingredients
- similarities between uploads and ratings
- historical uploads correlation
- etc.
Thank you! Questions?
Two sample scoring functions

Item popularity in uploads/ratings (\(F \rightarrow F-R\)) of friends:

\[
F(u, e) = \sum_{f \in F_u} \sum_{r \in U_f} [e \in r]
\]

Item popularity in categories of uploads/ratings (\(C \rightarrow C-R\)):

\[
C(u, e) = \sum_{c \in \mathcal{C}} w(u, c) \cdot \left( \sum_{r \in \mathcal{R}} [c \in r \land e \in r] \right)
\]