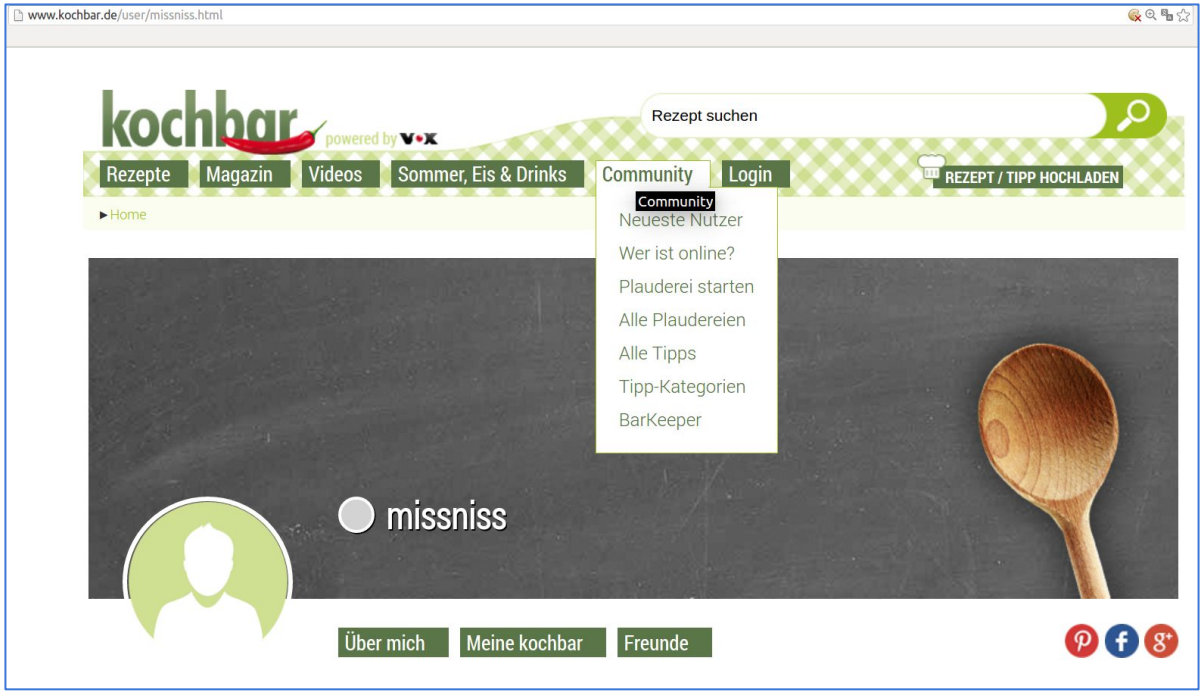
Understanding and Predicting Online Food Recipe Production Patterns

Tomasz Kuśmierczyk, Christoph Trattner, Kjetil Nørvåg

NTNU – Trondheim
Norwegian University of
Science and Technology

Context: On-line food communities



Project

NTNU – Trondheim
Norwegian University of Science and Technology

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: **Temporal Patterns in Online Food Innovation**. TempWeb 2015 (WWW Proceedings).
- T. Kusmierczyk, C. Trattner, K. Nørvåg: **Understanding and Predicting Online Food Recipe Production Patterns**. HT 2016.
- M. Rokicki, E. Herder, T. Kusmierczyk, C. Trattner: **Plate and Prejudice: Gender Differences in Online Cooking.** UMAP 2016.

current

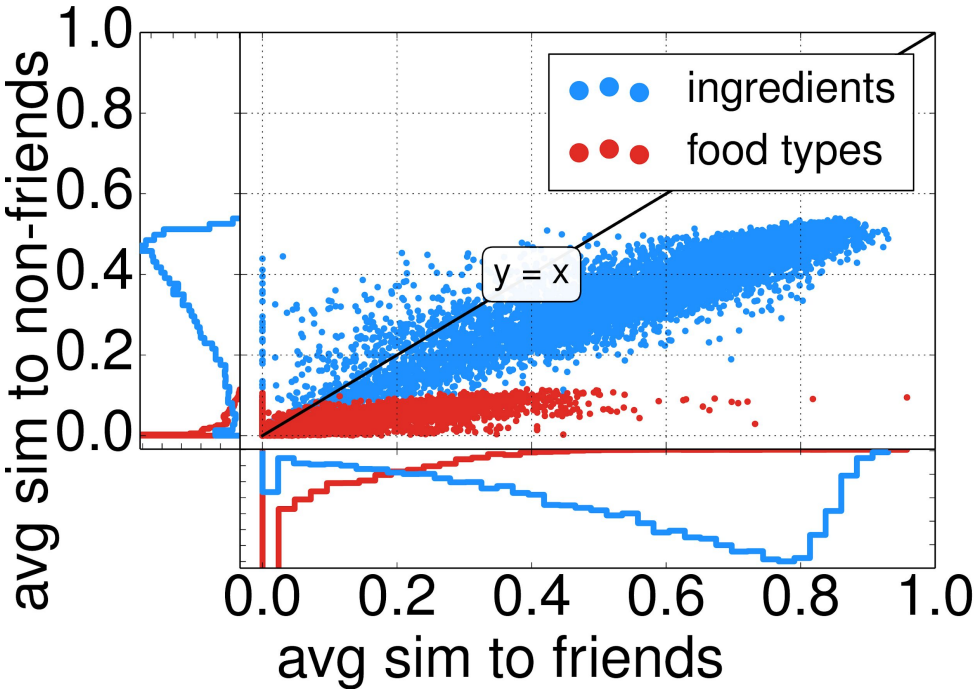
Goal: study factors influencing production

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

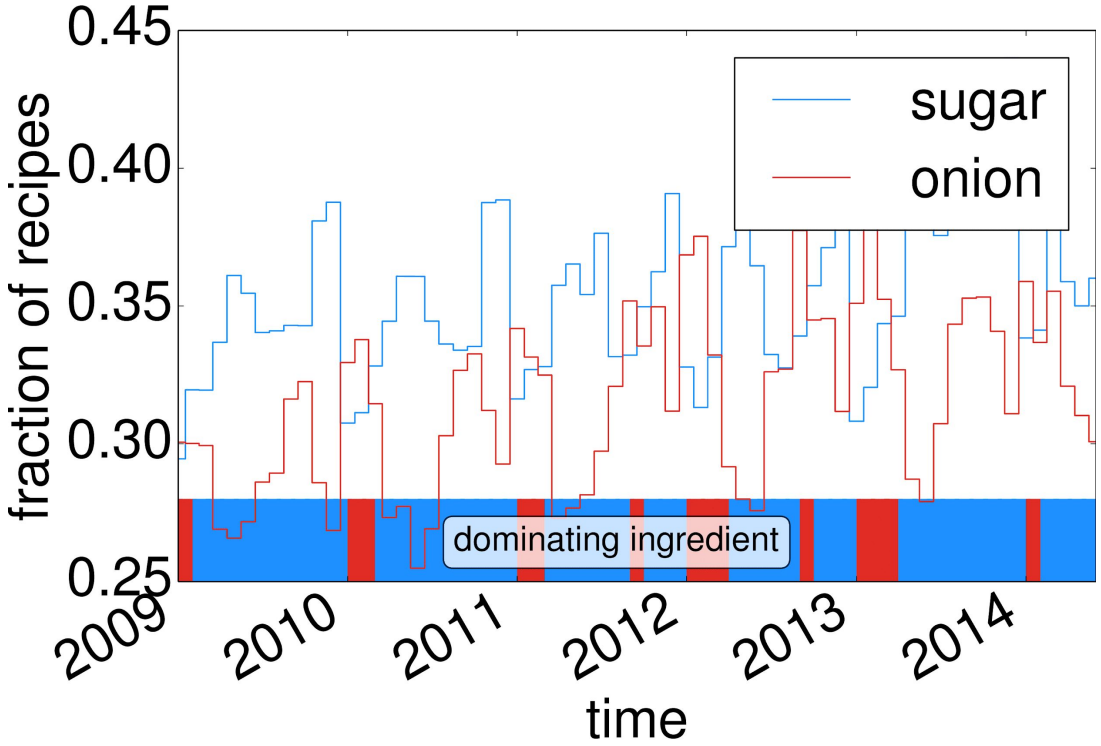
- food type (e.g., pizza)
- ingredients (e.g., tomatoes)

Sample patterns and observations

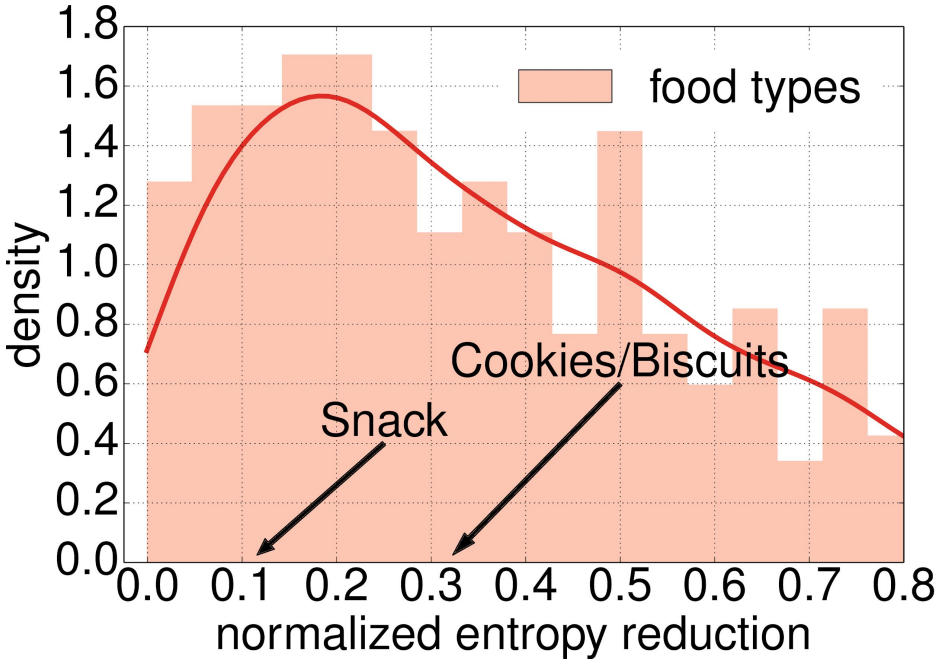
Bias towards friends



Popularity in time

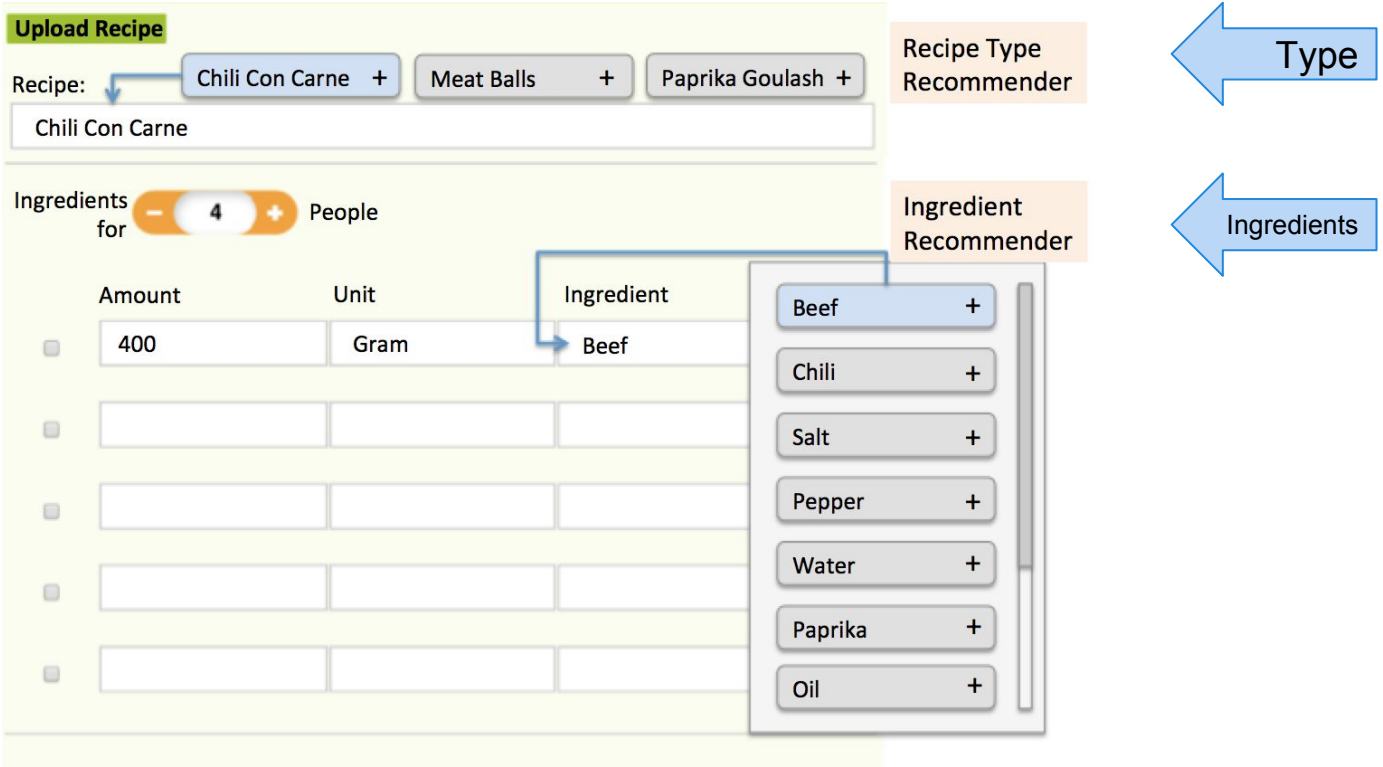


How categories can help with data sparsity



Application: recommend food types and ingredients

Sample Web app & Problem formulation



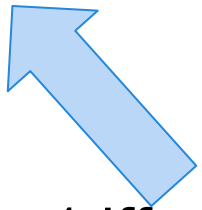
Solution: recommend list of items

- observe patterns
- design a personalized scoring function: $S(u,e)$
- rank items (ingredients, types) e according to scores

Sample scoring function

Item popularity in uploads by friends:

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

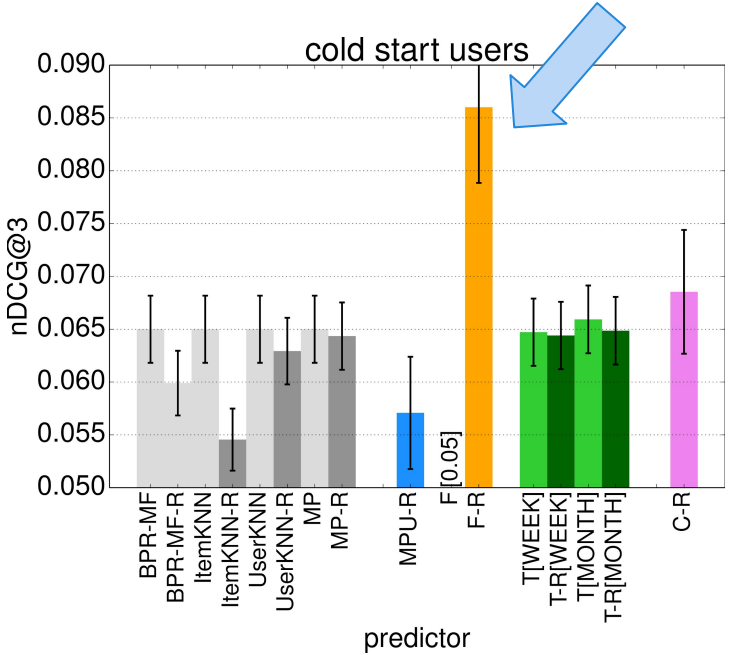


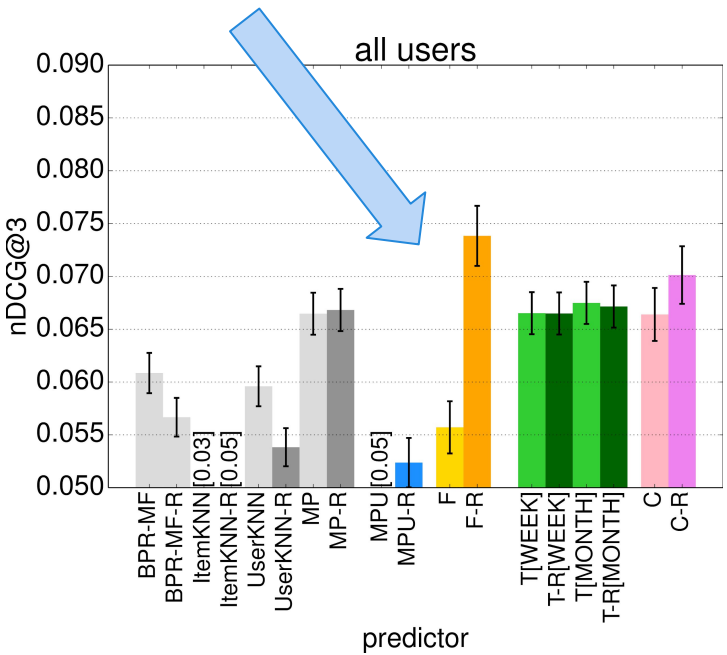
1 iff recipe r contains e

Results

Evaluation: Food types prediction

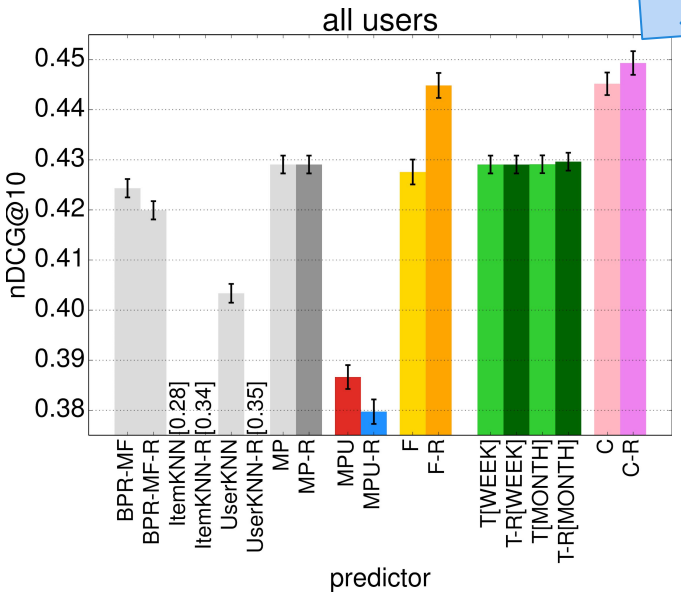
RATINGS FROM FRIENDS



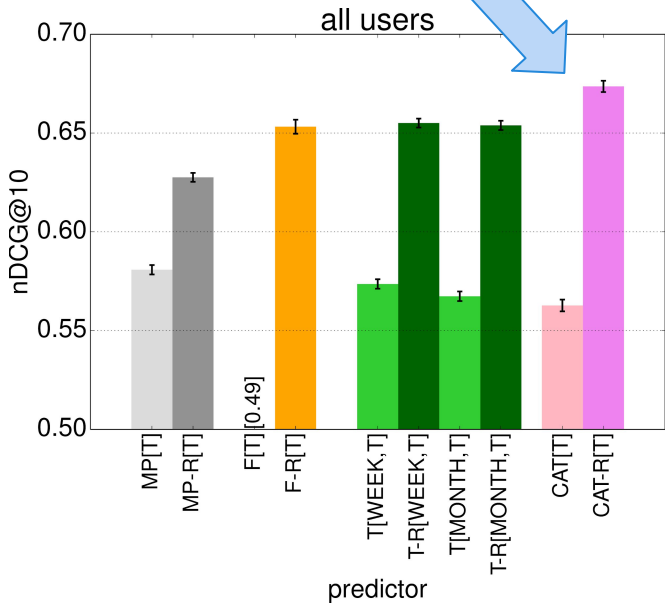


Evaluation: Ingredients prediction

CATEGORIES OF RATINGS



FOOD TYPE UNKNOWN



FOOD TYPE KNOWN

(Selected) Conclusions

- 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 ($\mathrm{F} \rightarrow \mathrm{F\text{-}R}$) of friends:

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

Item popularity in categories of uploads/ratings ($\mathrm{C} \rightarrow \mathrm{C\text{-}R}$):

$$\mathrm{C}(u,e) = \sum_{c \in \mathscr{C}} w(u,c) \cdot \left(\sum_{r \in \mathscr{R}} [c \in r \wedge e \in r] \right)$$