The Influence of City Size on Dietary Choices and Food Recommendation

Hao Cheng  
Leibniz Universität Hannover  
Hannover, Germany  
cheng@ikg.uni-hannover.de

Markus Rokicki  
L3S Research Center  
Hannover, Germany  
rokicki@l3s.de

Eelco Herder  
L3S Research Center  
Hannover, Germany  
herder@l3s.de

ABSTRACT
Contextual features have been leveraged by recommender systems in many different domains. Traditional contextual features – such as location and time – have successfully been combined with collaborative filtering or content-based features. However, it is likely that there are other – domain-specific – features that may have even more impact. In this paper, we focus on the influence of city size on food preferences. Apart from location and time, our results show that city size can significantly boost the performance of food recommendation.

KEYWORDS
Online food, city size differences, food recommendation

1 INTRODUCTION
Location-related features are well-explored in recommender systems, including context-aware food or restaurant recommenders. However, in past studies, locational features were mostly extracted on the level of countries, states, or cities [1, 6, 12, 13]. In this paper, we focus on a relatively unexplored feature: the impact of city size on food recommendation. Researchers have found that socio-economic characteristics are largely shaped by city population size [2]. In addition, [9] shows differences in Asian cities with different sizes in the process of urbanization, diet change, and transformation of food supply chains.

In this paper, we first statistically analyze the impact of several spatio-temporal contexts based on rating behaviors in the large German online food community kochbar.de1 and then use them for building various context-aware recommender systems (CARS). Compared with non-context-aware recommender systems, city size turned out to outperform other contexts with an improvement of 48% using a matrix factorization model for item prediction.

2 RELATED WORK
There are many analyses of geographical differences in eating preferences. Ahn et al. built flavor networks to uncover the ingredient preferences of cuisines worldwide [1]. Howell et al. analyzed taste preferences for different countries [6]. Wagner et al. used server logs to reveal ingredient preferences in German-speaking countries [12]. West et al. analyzed web usage logs to discover nutrient patterns of American states [13].

Food recommender systems try to match recipes to user profiles. Freyne and Berkovsky [4] used two hybrid recommendation strategies, recipe-based and ingredient-based. Elahi et al. built a food recommendation application to interact with the user’s long-term and short-term preferences [3]. Research on context-aware recommender systems [5] showed that including context features leads to more robust recommender systems. In this paper, we explore the impact of city size on food recommendation.

3 METHODOLOGY
Our study is based on a large-scale crawl from Kochbar.de [7]. It consists of over 400 thousand recipes published between 2008 and 2014, with more than 7 million ratings. As the ratings are overwhelmingly positive – 99% are 5-star ratings – similar as in [11], we treat the presence of ratings as positive feedback.

We have analyzed and selected two temporal and two locational features for the recommender models. Day-of-the-week (weekday and weekend) and season are derived from the upload time stamp of the recipes. The inner border (north-east, north-west, and south) context conditions are generated by mapping the cooks’ location data to wikipedia data. City size is derived by finding the closest city based on Geonames location data and using Geonames city population data to group cities into five city sizes: metropolis (≥ 1m), big-city (500k, 1m), medium-city (100k, 500k), small-city (50k, 100k), and town (15k, 50k).

For the recommendation task, we employ two baseline recommenders: the unpersonalized most-popular item recommender (MP) and Bayesian Personalized Ranking (BPR), a state-of-the-art matrix factorization model for item ranking [10]. These baseline algorithms are compared with corresponding context-aware recommenders, which are created by filtering users and items according to the relevant context factor. These algorithms are denoted as MP(ui) and BPR(ui). Then, the recommenders only recommend items that are in the same condition to the users. For instance, cosmopolitan users would receive recommendations for cosmopolitan recipes, based on ratings of other cosmopolitan users. We use random partitioning of the user-item matrix to monitor the bias caused by the algorithms with partitioning. The performances of the recommenders are evaluated by mean average precision (MAP).

---

1http://www.kochbar.de

This was partially funded by the German Federal Ministry of Education and Research (BMBF) under project GlycoRec (16SV7172). Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

UMAP’17, July 9–12, 2017, Bratislava, Slovakia
© 2017 Copyright held by the owner/author(s). ACM ISBN 978-1-4503-4635-1/17/07.
DOI: http://dx.doi.org/10.1145/3079628.3079641

2https://en.wikipedia.org/wiki/Inner_German_border

3http://www.geonames.org
MP MP(ui) BPR BPR(ui)
MAP
day-of-the-week
random

Figure 1: Evaluation of Food Recommender Systems.

4 RESULTS

Data Analysis. We start with a brief overview on observed differences. Main dishes in metropolises ($M = 12.6$), big cities ($M = 12.5$), and small cities ($M = 12.3$) contain more ingredients than in medium-cities ($M = 11.52$, e.g. metropolitan vs. medium-city: $W = 11288000, p \ll .001, r = .15$). As expected, spices are more frequently used in big cities ($M = 94.3\%$) and metropolises ($M = 93.4\%$) than in medium ($91.3\%$) cities ($N = 45497, \chi^2 = 35, df = 4, p \ll .001$). Red meat is more eaten in big cities ($M = 38.4\%$) than in cosmopolitan ($M = 34.7\%$) and medium ($M = 36.4\%$) cities ($N = 45497, \chi^2 = 20.34, df = 4, p < .001$). In summary, metropolitan main dishes are more complicated and exotic than medium-city main dishes; in towns, people typically cook more traditional.

Recommendation Experiment. As the user-recipe matrix is very sparse and the ratings are overwhelmingly positive, the recommendation task is item ranking. We compare each context-based recommender with the baseline recommender with the data randomly partitioned using the same number of pseudo-conditions. When a CARS outperforms both its baseline and the random partitioning recommender, it can be concluded that the corresponding contextual feature improves the recommendation performance. This is the case for each of the contexts, as can be seen in Figure 1.

A direct comparison of recommendation performance for different contexts is not valid. Instead, we compare performance improvements of context-aware recommenders to their respective non-context-aware models using the different contextual features. Profound improvements are shown in MP(ui) for the contexts season (184\%) and city size (118\%), and smaller, but significant improvements for day-of-the-week (56\%) and inner border contexts (30\%). The improvements for BPR(ui) are not as large as for MP(ui), but still remarkable for the contexts city size (48\%), season (35\%), and day-of-the-week (17\%). However, the improvement found for the inner border context is fairly small (3\%). The city size aware recommender BPR(ui) (MAP = 0.073) gives the best performance compared with the baseline recommender BPR (MAP = 0.05).

5 CONCLUSIONS

In this paper, we have shown that food preferences depend on the size of city that people are living in. Among others, in Germany, people in cosmopolitan cities eat more foreign food; people in smaller cities and towns eat more traditionally. Using city size has a positive impact on context-aware recipe recommendation – we observed a 48% increase in MAP compared to the baseline.

It is known that there are differences across city sizes in other areas as well, including civic involvement and economic performance [8]. Knowledge on the impact of city size can be effectively translated into measures to reinforce or counteract such effects. Therefore, it is likely that city size has an impact on context-based recommendations in general.

REFERENCES