Understanding the Impact of Weather for POI Recommendations

Christoph Trattner*
Know-Center, Austria

Alexander Oberegger
TUG, Austria

Lukas Eberhard
TUG, Austria

Denis Parra
PUC, Chile

Leandro Marinho
UFCG, Brasil

ABSTRACT

POI (point of interest) recommender systems for location-based social network services, such as Foursquare or Yelp, have gained tremendous popularity in the past few years. Much work has been dedicated into improving recommendation services in such systems by integrating different features that are assumed to have an impact on people's preferences for POIs, such as time and geolocation. Yet, little attention has been paid to the impact of weather on the users' final decision to visit a recommended POI. In this paper we contribute to this area of research by presenting the first results of a study that aims to predict the POIs that users will visit based on weather data. To this end, we extend the state-of-the-art Rank-GeoFM POI recommender algorithm with additional weather-related features, such as temperature, cloud cover, humidity and precipitation intensity. We show that using weather data not only significantly increases the recommendation accuracy in comparison to the original algorithm, but also outperforms its time-based variant. Furthermore, we present the magnitude of impact of each feature on the recommendation quality, showing the need to study the weather context in more detail in the light of POI recommendation systems.

Keywords

POI Recommender Systems; Location-based services; Weather-Context

1. INTRODUCTION

Location-based social networks (LBSN) enable users to check-in and share places and relevant content, such as photos, tips and comments that help other users in exploring novel and interesting places in which they might not have been before. Foursquare[1], for example, is a popular LBSN with millions of subscribers doing millions of check-ins everyday all over the world[2]. This vast amount of check-in data, publicly available through Foursquare's data access APIs, has recently inspired many researchers to investigate human mobility patterns and behaviors with the aim of assisting users by means of personalized POI (points of interest) recommendation services [15, 16].

Problem Statement. The problem we address in this paper is the POI recommendation problem. Hence, given a user u and her check-in history L^u, i.e., the POIs that she has visited in the past, and current weather conditions $C = \{c_1, \ldots, c_{|C|}\}$, where c_i are weather features such as temperature, wind speed, pressure, etc., we want to predict the POIs $\hat{L}^u = \{l_1, \ldots, l_{|L|}\}$ that she will likely visit in the future that are not in L^u.

Objective. Most of the existing approaches on POI recommendation exploit three main factors (aka contexts) of the data, namely, social, time and geolocation [5, 10, 15]. While these approaches work reasonably well, little attention has been paid to weather, a factor that may potentially have a major impact on users' decisions about visiting a POI or not. For example, if it is raining in a certain period of time and place, the user may prefer to check-in indoor POIs.

In this paper we contribute to this area of research by presenting the first results of a recently started project that exploits weather data to predict, for a given user within a given city, the POIs that she will likely visit in the future. To this end, we extract several weather features based on data collected from forecast.io such as temperature, cloud cover, humidity or precipitation intensity, and feed it into a state-of-the-art POI recommender algorithm called Rank-GeoFM [10].

Research Questions. To drive our research the following three research question were defined:

- **RQ1.** Do weather conditions have a relation with the check-in behaviour of Foursquare users?
- **RQ2.** Is it possible to improve current POI recommendation quality using these weather features?
- **RQ3.** Which weather features provide the highest impact on the recommendations?

Contributions. To the best of our knowledge, this is the first paper that investigates in detail the extent to which

*Corresponding author: ctrattner@know-center.at

RecSys'16, September 15–19, 2016, Boston, USA.

DOI: xx.xxx/xxx_x

[1]https://foursquare.com/
[2]https://foursquare.com/infographics/10million

City	#Check-Ins	#Venues	#Users	Sparsity
Minneapolis	37,737	797	436	89.1%
Boston	42,956	1141	637	94.3%
Miami	29,222	796	410	91.0%
Honolulu	16,042	410	173	77.4%

Table 1: Basic statistics of the dataset.

weather features such as temperature, cloud cover, humidity or precipitation intensity have an impact on users' check-in behaviors and how these features perform in the context of POI recommender systems. Although there is literature showing that POI recommender systems can be improved by using some kind of weather context such as e.g. temperature, it is not clear yet, how much they add or what type of weather feature is the most useful or maybe least useful one. Another contribution of this paper is the introduction of a weather-aware recommender method that builds upon a very strong state-of-the-art POI recommender system called Rank-GeoFM. The method is implemented and embedded into the very popular recommender framework MyMediaLite [7] and can be downloaded for free from our GitHub repository, details in Section 8.

Outline. The structure of this paper is as follows: In Section 2 we highlight relevant related work in the field. Section 3 describes how we enriched Rank-GeoFM with weather data. Section 4 describes the experimental setup and presents results from our empirical analysis. Section 5 presents insights on the results obtained with our weather-aware recommender approach. Finally, Sections 6 and 7 conclude the paper, with a summary of our main findings and future directions of the work.

2. RELATED WORK

With the advent of LBSNs, POI recommendation rapidly became an active area of research within the recommender systems, machine learning and GIS research communities [2]. Most of the existing research works in this area exploit some sort of combination between (some or all) of the following data sources: check-in history, social (e.g. friendship relations), time and geolocations [1,5,6,8,10,13,15]. While these different sources of data (aka contexts) affect the user's decision on visiting a POI in different ways, weather data, which according to common sense may have a great influence on this decision, they are still rarely used.

Martin et al. [11] proposed a mobile application which architecture considered the use of weather data to personalize a geocoding mobile service, but no implementation or evaluation was presented. A similar contribution was done by Meehan et al. [12], who proposed a hybrid recommender system based on time, weather and media sentiment when introducing the VISIT mobile tourism recommender, but they neither implemented nor evaluated it.

Among the few works that have actually used weather into the recommendation pipeline, Braunhofer et al. [3] introduced a recommender system designed to run in mobile applications for recommending touristic POIs in Italy. The authors conducted an online study with 54 users and found out that recommendations that take into consideration weather information were indeed able to increase the user satisfaction. Compared to this work, our implementation is based in a more recent and state-of-the-art algorithm, and we also provide details of which weather features contribute the most to the recommender performance. In an extension of their initial work, Braunhofer et al. [4] implemented and evaluated a context-aware recommender system which uses weather data. They find that the model which leverages the weather context outperformed the version without it. Although more similar to our current work, they did not provide a detailed feature analysis as the present article.

In summary, compared to previous works which have used weather as a contextual factor for recommendation systems, we provide detailed information about our recommendation algorithm and we contribute an implementation extending a state-of-the-art matrix factorization model exploiting rich weather data. Moreover, we also provide details on how the weather features were exploited by it, as well as a detailed analysis about the impact of the features on the recommendation quality.

Sym.	Description
$\mathcal{U}$	set of users $u_1, u_2, ..., u_{\|U\|}$
$\mathcal{L}$	set of POIs $l_1, l_2, ..., l_{\|L\|}$
FC_f	set of classes for feature f
F	set of weather feature classes $f_1, f_2, ..., f_{\|FC_f\|}$
Θ	latent model parameters containing the learned weights $\{L^{(1)}, L^{(2)}, L^{(3)}, U^{(1)}, U^{(2)}, F^{(1)}\}$ for locations, users and weather features.
X_{ul}	$\|U\| \times \|L\|$ matrix containing the check-ins of users at POIs.
X_{ulc}	$\|U\| \times \|L\| \times \|FC_f\|$ matrix containing the check-ins of users at POIs at a specific feature class c.
D_1	user-POI pairs: $(u, l)\|x_{ul} > 0$.
D_2	user-POI-feature class triples: $(u, l, c)\|x_{ulc} > 0$.
W	geographical probability matrix of size $\|L\|x\|L\|$ where $w_{ll'}$ contains the probability of l' being visited after l has been visited according to their geographical distance. $w_{ll'} = (0.5 + d(l, l'))^-1)$ where $d(l, l')$ is the geographical distance between the latitude and longitude of l and l'.
WI	probability that a weather feature class c is influenced by feature class c'. $wi_{cc'} = cos\_sim(c, c')$.
$N_k(l)$	set of k nearest neighbors of POI l.
y_{ul}	the recommendation score of user u and POI l.
y_{ulc}	the recommendation score of user u, POI l and weather feature class c.
$I(\cdot)$	indicator function returning $I(a) = 1$ when a is true and 0 otherwise.
ϵ	margin to soften ranking incompatibility.
γ_w	learning rate for updates on weather latent parameters.
γ_g	learning rate for updates on latent parameters from base approach.
$E(\cdot)$	a function that turns the rating incompatibility $Incomp(y_{ulc}, \epsilon)$, that counts the number of locations $l' \in \mathcal{L}$ that should be ranked lower than l at the current weather context c and user u but are ranked higher by the model, into a loss $E(r) = \sum_{i=1}^{r} \frac{1}{i}$.
$\delta_{ucll'}$	function to approximate the indicator function with a continuous sigmoid function $s(a) = \frac{1}{1+exp(-a)}$. $\delta_{ucll'} = s(y_{ul'c} + \epsilon - y_{ulc})(1 - s(y_{ul'c} + \epsilon - y_{ulc}))$
$\lfloor\frac{\|L\|}{n}\rfloor$	if the n^{th} location l' was ranked incorrect by the model the expactation is that overall $\lfloor\frac{\|L\|}{n}\rfloor$ locations are ranked incorrect.
g, μ	auxiliary variable that save partial results of the calculation of the stochastic gradient.

Table 2: The notations used to describe Rank-GeoFM and the incorporation of the weather context.

3. RECOMMENDATION APPROACH

Our recommendation approach is built upon a state-of-the-art POI recommender algorithm named Rank-GeoFM [10], a personalized ranking based matrix factorization method. We have selected Rank-GeoFM over other alternatives because it has been shown to be a very strong POI recommender method compared to other approaches often cited

Algorithm 1: Rank-GeoFM with weather context

Input: check-in data D_1, D_2, geographical influence matrix W, weather influence matrix WI, hyperparameters ϵ, C, α, β and learning rate γ_g and γ_w
Output: parameters of the model $\Theta = \{L^{(1)}, L^{(2)}, L^{(3)}, U^{(1)}, U^{(2)}, F\}$

1 **init:** Initialize Θ with $\mathcal{N}(0, 0.01)$; Shuffle D_1 and D_2 randomly
2 **repeat**
3   **for** $(u, l) \in D_1$ **do**
4   approach from Li et al. [10]
5   **end**
6   **for** $(u, l, c) \in D_2$ **do**
7   Compute y_{ulc} as Equation 3 and set n = 0
8   **repeat**
9   Sample l' and c', Compute $y_{ul'c'}$ as Equation 3
10   n++
11   **until** $I(x_{ulc} > x_{ul'c'})I(y_{ulc} < y_{ul'c'} + \epsilon) = 1$ *or* $n > |L|$
12   **if** $I(x_{ulc} > x_{ul'c'})I(y_{ulc} < y_{ul'c'} + \epsilon) = 1$ **then**
13       $\eta = E\left(\left\lfloor \frac{|L|}{n} \right\rfloor\right) \delta_{ucll'}$
14       $g = \left(\sum_{c^* \in FC_f} wi_{c'c^*} f^{(1)}_{c*} - \sum_{c^+ \in FC_f} wi_{cc^+} f^{(1)}_{c+}\right)$
15       $f^{(1)}_c \leftarrow f^{(1)}_c - \gamma_w \eta (l^{(2)}_{l'} - l^{(2)}_l)$
16       $l^{(3)}_l \leftarrow l^{(3)}_l - \gamma_w \eta g$
17       $l^{(2)}_{l'} \leftarrow l^{(2)}_{l'} - \gamma_w \eta f_c$
18       $l^{(2)}_l \leftarrow l^{(2)}_l + \gamma_w \eta f_c$
19   **end**
20   Project updated factors to accomplish constraints
21   **end**
22 **until** *convergence*
23 **return** $\Theta = \{L^{(1)}, L^{(2)}, L^{(3)}, U^{(1)}, U^{(2)}, F^{(1)}\}$

in the literature. In Li et al. [10] the authors compared Rank-GeoFM against twelve other recommender methods, showing that Rank-GeoFM significantly outperforms strong generic baselines, such as user-KNN, item-KNN CF, WRMF, BPR-MF [7] as well as specialized POI recommender methods, such as BPP [17]. Another reason for choosing Rank-GeoFM is related to its ability to easily accommodate additional features that we plan to use in this work. The aim of Rank-GeoFM is to learn latent parameters that model the relationship between the context of interest (in our case weather conditions) and the user/POI.

Table 2 describes the symbols used in the recommender algorithm. For each type of contextual data considered, latent model parameters are introduced. The prediction of a <*user, POI, context*> triple is then made based on this learned latent parameters. The parameters are trained using a fast learning scheme introduced by the authors that is based on Stochastic Gradient Descent (SGD).

To add the weather context into Rank-GeoFM, the weather features' values needed to be discretized. This was done to reduce data sparsity. For example, if we considered temperature as a real number, most of the check-ins concerning specific temperature values would probably be zero. Thus, transforming continuous values of weather features (e.g., temperature) into intervals might alleviate this problem. Hence, a mapping function is introduced (see Equation 1) that converts the weather features into interval bins. $|FC_f|$ defines the size of the bin for the current weather feature. We will refer to these bins as feature classes. Best results were obtained with $|FC_f| = 20$ (validated on held-out data).

$$c_f(value) = \left\lfloor \frac{(value - min(f)) \cdot (|FC_f| - 1)}{(max(f) - min(f))} \right\rfloor \quad (1)$$

To extend the original Rank-GeoFM approach with weather context, three additional latent factors are introduced that are represented by matrices in a K-dimensional space. The first one is for incorporating the *weather-popularity-score* that models whether or not a location is popular in a specific weather feature class and is named $L^{(2)} \in \mathbb{R}^{|L| \times K}$, where K denotes the size of the latent parameter space. Furthermore, a matrix $L^{(3)} \in \mathbb{R}^{|L| \times K}$ is introduced to model the *influence between two feature classes*. In other words, $L^{(3)}$ softens the borders between the particular feature classes. The third latent parameter $F^{(1)} \in \mathbb{R}^{|FC_f| \times K}$ is then used to parametrize the feature classes of the specific weather feature. In addition to the latent parameters, a Matrix $WI \in \mathbb{R}^{|FC_f| \times |FC_f|}$ is introduced for storing the probability that a weather feature class c is influenced by feature class c'. Denoting x_{ulc} as the frequency that a user u checked in at POI l with the current weather context c, this probability is calculated as follows:

$$wi_{cc'} = \frac{\sum_{u \in U} \sum_{l \in L} x_{ulc} x_{ulc'}}{\sqrt{\sum_{u \in U} \sum_{l \in L} x^2_{ulc}} \sqrt{\sum_{u \in U} \sum_{l \in L} x^2_{ulc'}}} \quad (2)$$

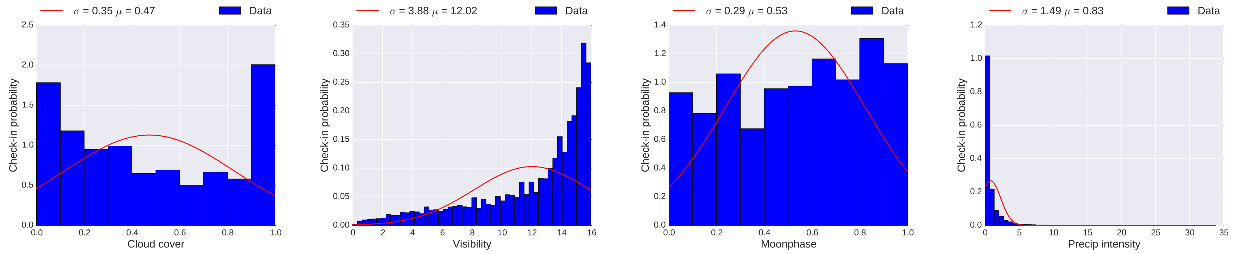
To calculate the recommendation score for a given user u, POI l and weather feature class c, Equation 3 is introduced, where y_{ul} denotes the recommendation score as computed in Li et al. [10].

$$\begin{aligned} y_{ul} &= u^{(1)}_u \cdot l^{(1)}_l + u^{(2)}_u \cdot \sum_{l^* \in N_k(l)} w_{ll^*} l^{(1)}_{l^*} \\ y_{ulc} &= y_{ul} + f^{(1)}_c \cdot l^{(2)}_l + l^{(3)}_l \cdot \sum_{c^* \in FC} wi_{cc^*} f^{(1)}_{c^*} \end{aligned} \quad (3)$$

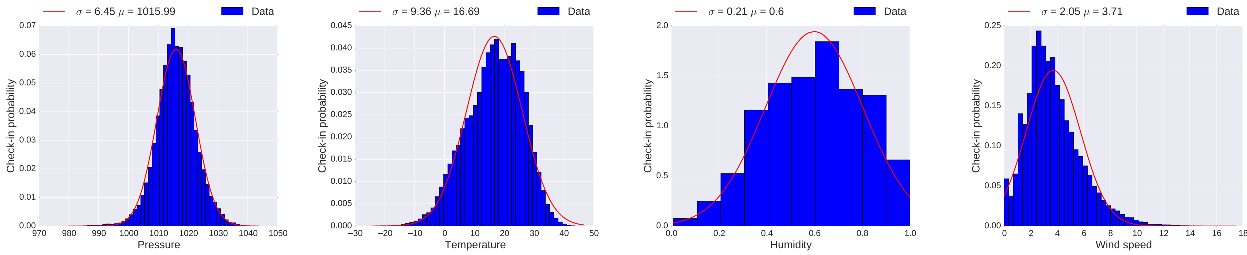
Algorithm 1 shows how we incorporated the weather context features into the base Rank-GeoFM approach. Taking the initialization and the hyperparameters from the original approach we first iterate over all pairs of users and POIs $(u, l) \in D_1$, where D_1 is the set of all check-ins and do the adjustments of the latent parameters as described in Li et al. [10].

We then introduce an iteration over all <*user, venue, feature-class*> triples $(u, l, c) \in D_2$ in order to adjust the latent parameters on the incorrect ranked venues according to the specific weather context. This adjustment is necessary because the algorithm might rank a triple (u, l, c) correctly where on the other hand (u, l, c') might be ranked incorrectly. The adjustments are then done accordingly to the base algorithm in lines 6-20.

During our studies we found that with a learning rate of $\gamma_g = .0001$, as used in Li et al. [10], the algorithm did not converge. The reason for that is that the adjustments are done on a higher granularity for each (u, l, c) triple and not just on the (u, l) level. Henceforth, we introduce a new learning rate parameter $\gamma_w = .00001$ for the weather context, for which stable results could be observed (validation on hold-out data). Similarly to Li et al. [10], we found in our experiments that the best values of the hyperparameters are as follows (validated on hold-out data): $\epsilon = .3$, $C = 1.0$,

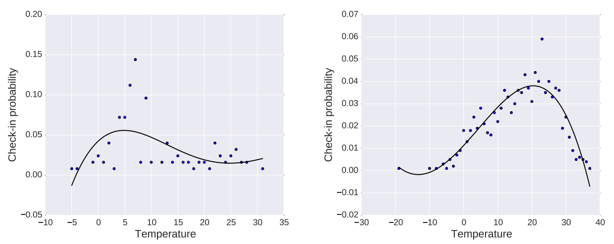


(a) Cloud cover (b) Visibility (c) Moonphase (d) Precipitation intensity

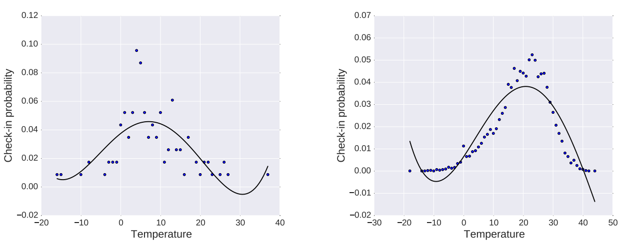


(e) Pressure (f) Temperature (g) Humidity (h) Windspeed

Figure 1: Check-in distributions over the 8 weather features.



(a) "Austrian Restaurant" (b) "Farm"



(c) "Ski Area" (d) "Ice Cream Shop"

Figure 2: Examples of check-in distributions over different types of places in Foursquare. On the left hand side, places where people check-in at lower temperatures are shown and on the right higher temperature places are featured.

$\alpha = \beta = .2$, and $K = 100$ as used for the dimensions of the matrices $L^{(1)}$, $L^{(2)}$ and $L^{(3)}$.

4. EXPERIMENTAL SETUP

In this section we describe in detail our experimental setup, i.e., the datasets we used, a brief characterization of this datasets concerning the weather features used, and the evaluation protocol we have chosen to conduct our study.

4.1 Datasets

The dataset we used in this study was obtained from the work of Yang et al. [14]. It is a Foursquare crawl comprising user check-in data from April 2012 to September 2013. The original dataset contains more than 33 million check-ins from 415 cities in 77 countries.However, before dealing with our problem on such a large scale, we decided to first concentrate our investigation on a small set of US cities. We selected four cities that could represent some weather variety in order to investigate whether our is resilient to such variety of weather conditions (see Figure 3). Table 1 provides an overview of the check-in statistics of the four target cities chosen for our experiments: Minneapolis, Boston, Miami and Honolulu.

Concerning the weather information, we have used the API of forecast.io[3] to collect, for each <*time, place*> tuple present in our dataset, their corresponding weather information. For that, we need to pass the following request to the API:

https://api.forecast.io/forecast/APIKEY/LAT,LON,TIME

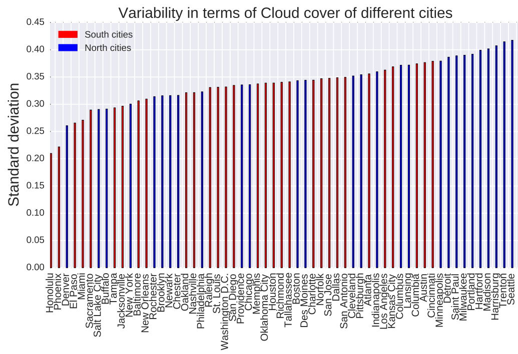
For the purposes of our analysis, we obtained eight weather features, namely, cloud cover, visibility, moon phase, precipitation intensity, pressure, temperature, humidity and wind speed, for all places and time-stamps in our dataset that are provided by forecast.io.

4.2 Data Analysis

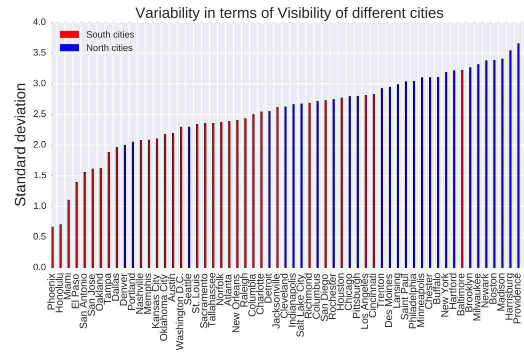
Figure 1 shows the probability distributions of check-ins for each of the eight weather features used. Notice that the distributions of pressure, temperature, humidity and wind speed resemble a normal distribution (see the colored approximation curve). Moreover, while moon phase seems to follow a uniform distribution, which indicates that it will likely not help the recommendation model, the distribution of precipitation is very skewed, showing that users have a strong preference to check-in places when there is low precipitation intensity (i.e., not raining), indicating that this feature might have a good discriminative power.

In addition to this, Figure 2 illustrates the check-in distribution as a function of temperature in four different POI categories. As highlighted in this Figure, different patterns occur depending on the category chosen. While people prefer to check-in in e.g., "Austrian Restaurants" or "Ski Areas" when the temperature is low, "Ice Cream Shops" or "Farms" are preferred when temperatures are higher.

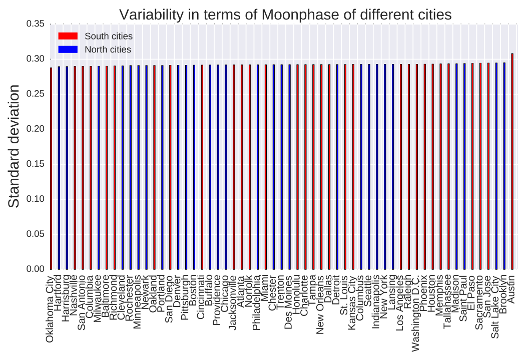
[3]https://developer.forecast.io/docs/v2



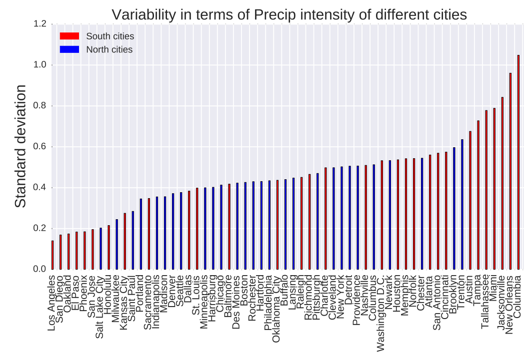
(a) Cloud cover



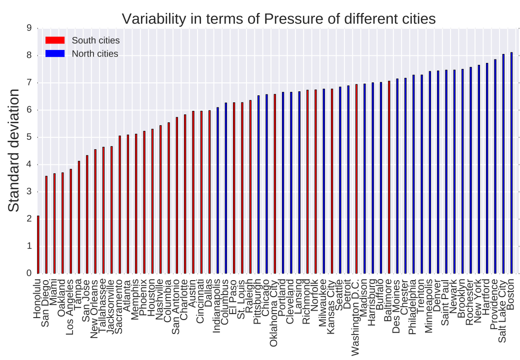
(b) Visibility



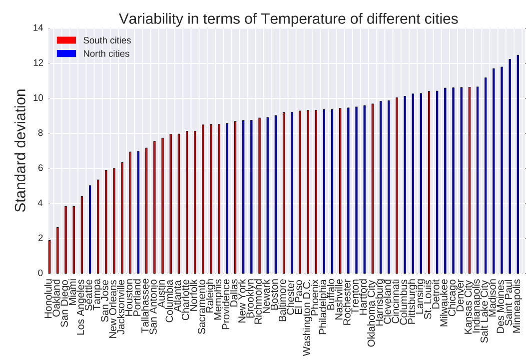
(c) Moonphase



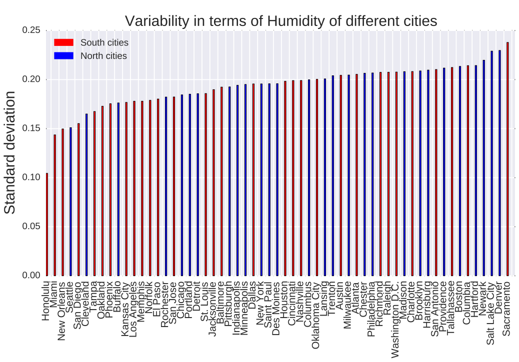
(d) Precipitation intensity



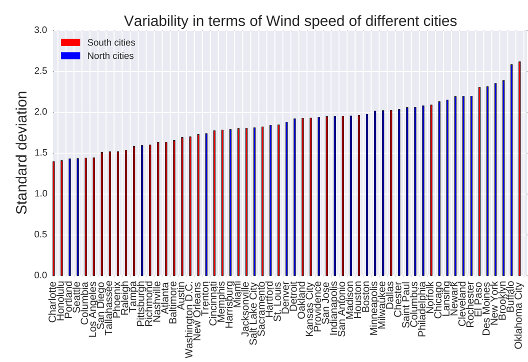
(e) Pressure



(f) Temperature

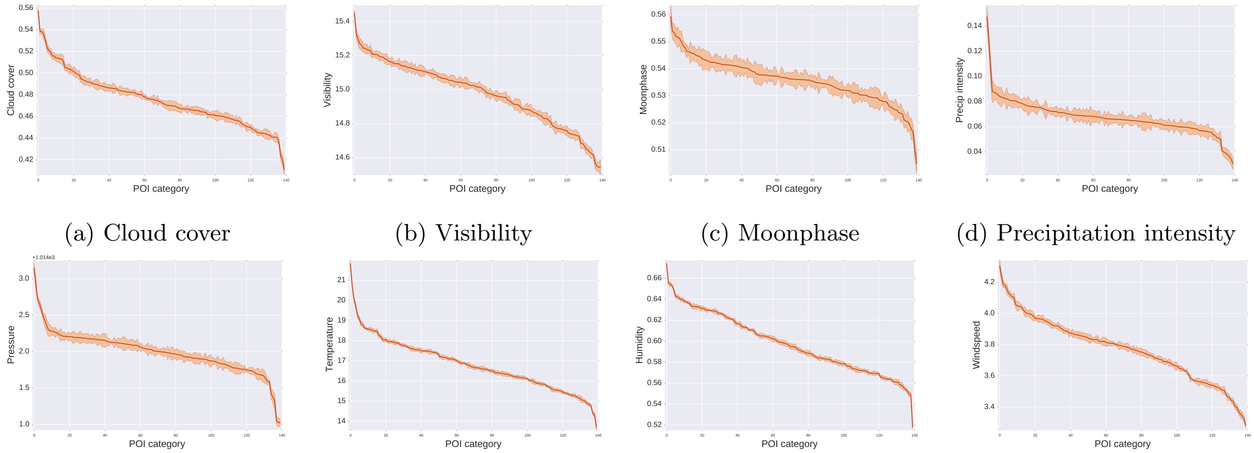


(g) Humidity



(h) Windspeed

Figure 3: Weather feature variability (sorted) measured via standard deviation over cities. Left: cities with lowest variability. Right: cities with highest variability.



(a) Cloud cover (b) Visibility (c) Moonphase (d) Precipitation intensity

(e) Pressure (f) Temperature (g) Humidity (h) Windspeed

Figure 4: Mean weather feature values for POI categories with standard errors.

Figure 3 shows how the weather features vary in each city of the original Foursquare dataset. Notice that with the exception of moonphase, all the features present a dependency regarding the city where they are measured, indicating that a different recommendation model should probably be trained for each different city. Moreover, in general, weather shows a higher variability in the north of the US and a very low variability in the south that peaks in the island Honolulu which shows almost no variability in terms of weather. Figure 4 shows the different mean values of the eight weather features over the POI categories. With the small overlapping of the standard error of the means it's revealed that indeed categories have a distinct popularity across various weather feature values. Even moonphase shows a divergent category popularity at its tails.

After this analysis we can confidently state that there is indeed a relation between the weather conditions and the check-in behavior of Foursquare users, which answer our first research question (RQ1) stated at Section 1.

4.3 Evaluation

Protocol. To evaluate the performance of our algorithm, we have chosen the same evaluation protocol as described in the original Rank-GeoFM paper [10]. Hence, we split the dataset (according to the time line) into training, validation and testing sets for each city by adding the first 70% of the check-ins of each user to the training set, the following 20% to the test set and the rest to the validation set (=10%). The training set was then used to learn the latent model parameters. During the training phase of the algorithm, the validation set was used to tune the algorithm convergence. When convergence was observed (typically around 3,000-5,000 iterations with fast learning scheme enabled), the training was stopped and the learned parameters were used to evaluate the model on the test set.

Baselines. As baselines for our experiments, we used the original Rank-GeoFM approach, that models user-preferences as well as geographical influence into the model. Furthermore, we compare to the time-based method of Rank-GeoFM, that was also introduced in Li et al. [10].

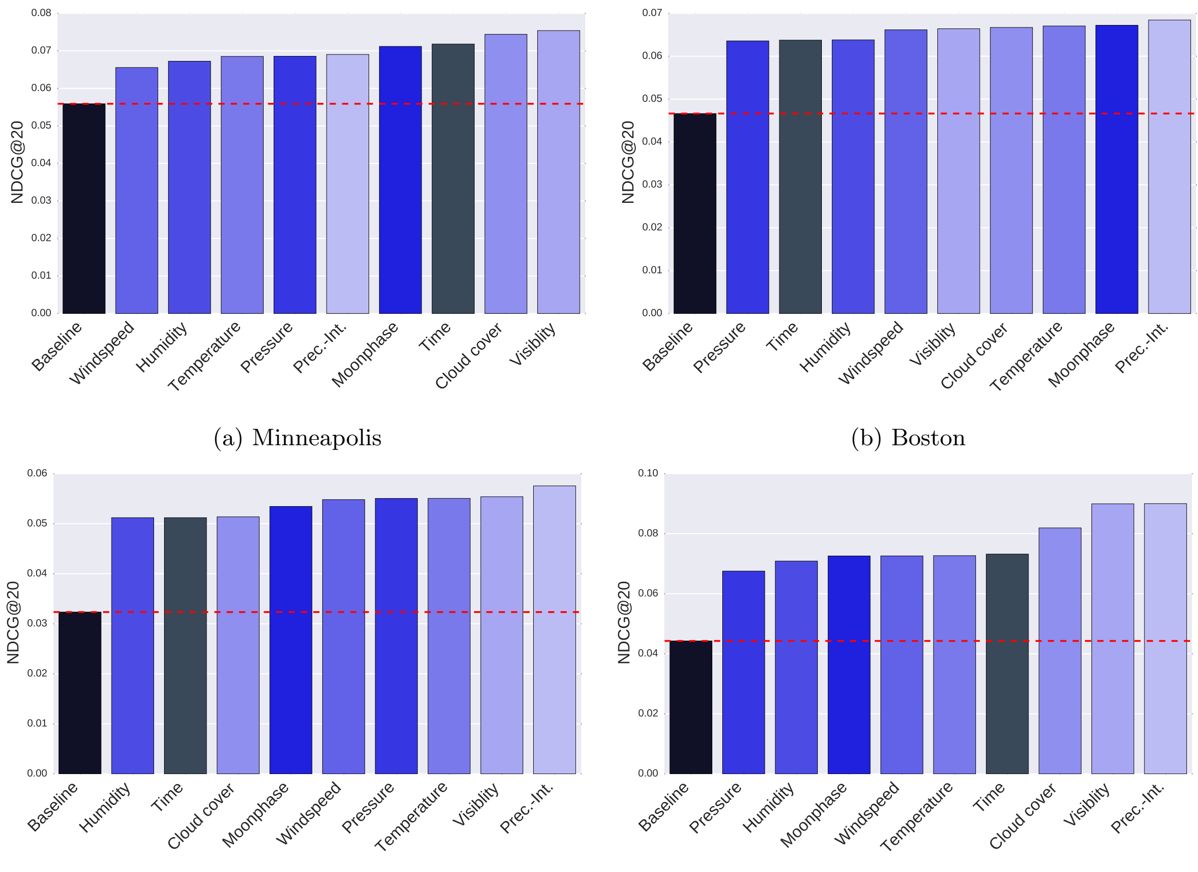
Metric. As evaluation metric NDCG@k (Normalized Discounted Cumulative Gain) with $k = 20$[4] was chosen, as we want to predict the top-k POIs for a user.

5. RESULTS

Figure 5 shows the results of our offline experiment. As shown, in all cases over all four cities, Rank-GeoFM enriched with our proposed weather features significantly outperform the original Rank-GeoFM algorithm, which answers our RQ2. For all pairwise-comparisons (recommenders with weather context vs. without) a standard t-test showed that the p-values were always smaller than $p < .001$. What is even more interesting to note is the performance of Rank-GeoFM that utilizes the time feature as contextual factor. As highlighted, in all cases, Rank-GeoFM with weather features, such as visibility and precipitation intensity outperforms the time-based variant, showing the indeed weather conditions may help to improve the recommendation quality.

We also highlight the fact that certain weather features perform better than others and this pattern seems to be city dependent. This can be clearly observed in Figure 5, where the results of Rank-GeoFM with each weather feature is shown. This answers RQ3, showing which features provide the highest gain in recommendation quality. For example, in Honolulu the best performing feature is precipitation intensity, while in Minneapolis visibility seems to work the best among all investigated weather features. Similar patterns can be observed for other features, such as temperature or cloud cover changing their relative importance across the four cities. These observations are in line with the results in Figure 1, showing a strong tendency of check-ins into POIs under certain weather conditions. However, what is also interesting to note is the good performance of the moonphase feature, which appeared to be uniformly distributed in general (cf. Figure 1). Hence, it appears, that at the level of locations, there is indeed a strong preference for check-ins in different phases of the moon. In a recent research, Kohyama et a. [9] found a relation between moonsphase, tidal

[4]Please note, that we also run simulations with k=5 or 10, with similar trends in the results as obtained with $k = 20$. However, due to limited space, they were not included into this paper.



(a) Minneapolis (b) Boston

(c) Miami (d) Honolulu

Figure 5: Recommender accuracy for the 8 different weather context features (sorted by importance) compared to Rank-GeoFM without weather context (denoted as “Baseline”). For further comparison the time-aware version of Rank-GeoFM is included, denoted as “Time”. The red dotted line denotes the basline.

variation, humidity and rainfall. Although further analysis should be performed to establish a link between our study and theirs, it might be indicative of an explanation regarding the effect of moonphase in our POI recommendation model.

Finally, the relative performance improvement over the original Rank-GeoFM seems to be also location dependent. Hence, while our approach work to a great extent better compared to the baseline for Miami and Honolulu, the differences are less pronounced for Minneapolis. One reason for this observation could be that the weather in this area is less often changing than in the other two cities or that people in Minneapolis are less weather sensitive. However, to further confirm these hypotheses, additional analyses are needed.

6. CONCLUSIONS

In this paper we presented our preliminary findings on how weather data may affect users' check-in behavior and how this information can be used in the context of a POI recommender system. As our preliminary analyses on the Foursquare check-in data showed, the weather factors have indeed a significant impact on the people's check-in behavior, showing different check-in profiles for different kinds of places (which answers RQ1). Further, we fed the proposed weather features into a state-of-the-art POI recommender and we were able to increase the recommender accuracy in comparison to the original method that does not use weather data (thus answering RQ2). Furthermore, our experiments revealed that the weather context is more useful than the context of time and, that the weather features used in this work are city-dependent. Finally, our study showed (see RQ3) that among the considered weather features, precipitation intensity and visibility are the most significant ones to improve the ranking in a weather-aware POI recommender system.

7. FUTURE WORK

Currently, our work only investigates one weather feature at a time. Investigating different hybridization or context-aware recommender system (CARS) methods and other context variables will be therefore a task to be conduct in our future work. Furthermore, it will help to investigate in more detail, how the algorithm performs on the whole Foursquare dataset, as more interesting patterns across cities may occur. Finally, we would like to extend our investigations also at user levels, since the current ones concentrate only on the weather profiles of the POIs.

8. OPEN SCIENCE

In order to make the results obtained in this work reproducible, we share code and data of this study. The proposed method Rank-GeoFM with weather context is implemented with the help of the MyMediaLite framework [7] and can be downloaded for free from out GitHub repository[5]. Furthermore, the data samples used in the experiments can be requested for free via email from CT.

Acknowledgements

This work is supported by the Know-Center. The Know-Center is funded within the Austrian COMET Program - managed by the Austrian Research Promotion Agency (FFG).

9. REFERENCES

[1] J. Bao, Y. Zheng, and M. F. Mokbel. Location-based and preference-aware recommendation using sparse geo-social networking data. In *Proceedings of the 20th International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '12, pages 199–208, New York, NY, USA, 2012. ACM.

[2] J. Bao, Y. Zheng, D. Wilkie, and M. Mokbel. Recommendations in location-based social networks: A survey. *Geoinformatica*, 19(3):525–565, July 2015.

[3] M. Braunhofer, M. Elahi, M. Ge, F. Ricci, and T. Schievenin. STS: design of weather-aware mobile recommender systems in tourism. In *Proceedings of the First International Workshop on Intelligent User Interfaces: Artificial Intelligence meets Human Computer Interaction (AI*HCI 2013) A workshop of the XIII International Conference of the Italian Association for Artificial Intelligence (AI*IA 2013), Turin, Italy, December 4, 2013.*, 2013.

[4] M. Braunhofer, M. Elahi, F. Ricci, and T. Schievenin. Context-aware points of interest suggestion with dynamic weather data management. In *Information and communication technologies in tourism 2014*, pages 87–100. Springer, 2014.

[5] C. Cheng, H. Yang, I. King, and M. R. Lyu. Fused matrix factorization with geographical and social influence in location-based social networks. In *Proc. of AAAI*, pages 17–23, 2012.

[6] G. Ference, M. Ye, and W.-C. Lee. Location recommendation for out-of-town users in location-based social networks. In *Proceedings of the 22Nd ACM International Conference on Information & Knowledge Management*, CIKM '13, pages 721–726, New York, NY, USA, 2013. ACM.

[7] Z. Gantner, S. Rendle, C. Freudenthaler, and L. Schmidt-Thieme. MyMediaLite: A free recommender system library. In *In Proc. of RecSys'11*, 2011.

[8] H. Gao, J. Tang, X. Hu, and H. Liu. Exploring temporal effects for location recommendation on location-based social networks. In *Proceedings of the 7th ACM Conference on Recommender Systems*, RecSys '13, pages 93–100, New York, NY, USA, 2013. ACM.

[9] T. Kohyama and J. M. Wallace. Rainfall variations induced by the lunar gravitational atmospheric tide and their implications for the relationship between tropical rainfall and humidity. *Geophysical Research Letters*, 43(2):918–923, 2016. 2015GL067342.

[10] X. Li, G. Cong, X.-L. Li, T.-A. N. Pham, and S. Krishnaswamy. Rank-geofm: A ranking based geographical factorization method for point of interest recommendation. In *Proc. of SIGIR'15*, pages 433–442, New York, NY, USA, 2015. ACM.

[11] D. Martin, A. Alzua, and C. Lamsfus. *A Contextual Geofencing Mobile Tourism Service*, pages 191–202. Springer Vienna, Vienna, 2011.

[12] K. Meehan, T. Lunney, K. Curran, and A. McCaughey. Context-aware intelligent recommendation system for tourism. In *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2013 IEEE International Conference on*, pages 328–331. IEEE, 2013.

[13] I. Nunes and L. Marinho. A personalized geographic-based diffusion model for location recommendations in lbsn. In *Proceedings of the 2014 9th Latin American Web Congress*, LA-WEB '14, pages 59–67, Washington, DC, USA, 2014. IEEE Computer Society.

[14] D. Yang, D. Zhang, and B. Qu. Participatory cultural mapping based on collective behavior in location based social networks. *ACM Transactions on Intelligent Systems and Technology*, 2015. in press.

[15] M. Ye, P. Yin, W.-C. Lee, and D. L. Lee. Exploiting geographical influence for collaborative point-of-interest recommendation. In *Proc. of SIGIR'11*, pages 325–334. ACM, 2011.

[16] H. Yin, Y. Sun, B. Cui, Z. Hu, and L. Chen. Lcars: a location-content-aware recommender system. In *Proc. of KDD'13*, pages 221–229. ACM, 2013.

[17] Q. Yuan, G. Cong, and A. Sun. Graph-based point-of-interest recommendation with geographical and temporal influences. In *Proc. of CIKM'14*, pages 659–668. ACM, 2014.

[5]https://github.com/aoberegg/WPOI