

Are Real-World Place Recommender Algorithms Useful in Virtual World Environments?

Leandro B. Marinho
UFCG
Campina Grande, Brazil
lbmarinho@dsc.ufcg.edu.br

Christoph Trattner
NTNU & Know-Center
Trondheim, Norway & Graz,
Austria
ctrattner@know-center.at

Denis Parra
Pontificia Universidad Católica
de Chile
Santiago, Chile
dparra@ing.puc.cl

ABSTRACT

Large scale virtual worlds such as massive multiplayer online games or 3D worlds gained tremendous popularity over the past few years. With the large and ever increasing amount of content available, virtual world users face the information overload problem. To tackle this issue, game-designers usually deploy recommendation services with the aim of making the virtual world a more joyful environment to be connected at. In this context, we present in this paper the results of a project that aims at understanding the mobility patterns of virtual world users in order to derive place recommenders for helping them to explore content more efficiently. Our study focus on the virtual world SecondLife, one of the largest and most prominent in recent years. Since SecondLife is comparable to real-world Location-based Social Networks (LBSNs), i.e., users can both check-in and share visited virtual places, a natural approach is to assume that place recommenders that are known to work well on real-world SBSNs will also work well on SecondLife. We have put this assumption to the test and found out that (i) while collaborative filtering algorithms have compatible performances in both environments, (ii) existing place recommenders based on geographic metadata are not useful in SecondLife.

Categories and Subject Descriptors

H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval—*Information filtering*

Keywords

virtual environments; place recommender systems; location-based social networks

Permission to make digital or hard copies of all or part 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 components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
RecSys '15, September 16 – 20, 2015, Vienna, Austria.
ACM 978-1-4503-3692-5/15/09.
<http://dx.doi.org/10.1145/2792838.2799674>.

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 others in exploring novel and interesting places they might not have been before. Foursquare¹, for example, is a popular LBSN with millions of subscribers doing millions of check-ins everyday all over the world. 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 behavior with the aim of assisting users, by means of personalized recommendation services, in exploring their surroundings more efficiently [19, 20, 2].

Problem Statement. Virtual world environments, on the other hand, are interactive 3D worlds where millions of users all around the globe spend a lot of time every day to create and explore new content. Rules are typically drawn from reality, such as gravity, topography, locomotion, real-time actions, and communication. These systems gained tremendous popularity over the past few years and are poised for continued growth. Facebook, for example, recently bought Oculus VR that produces virtual reality headsets for allowing game players to immerse themselves in virtual worlds, which indicates that there are big plans for virtual worlds on the table². With the massive amount of content available all the time, virtual world users, especially newbies and the inexperienced ones, typically suffer from the information overload problem which undermines their ability to find relevant and exciting content. To tackle this issue content suggestion services are typically provided to help people in exploring content more efficiently [17].

While there are many works on recommender systems for virtual worlds that focus on economical [7, 16] and social aspects of these systems [12, 11, 13], only a few of them try to understand mobility patterns in virtual worlds (see e.g., [15, 1]). These latter works indicate a strong correlation between virtual and real world movement patterns and although they suggest that existing real world LBSN recommender approaches would perform similarly in virtual worlds, to the best of our knowledge there are no studies yet showing explicitly if this assumption really holds. The findings with this respect have important implications for recommender systems designers, because if this assumption holds, a whole suite of place recommendation algorithms can be transferred from the LBSN domain to virtual worlds. If not, new algorithms need to be devised for characterizing the specificities of virtual world users mobility patterns.

¹<http://www.foursquare.com>

²<http://www.theguardian.com/technology/2014/jul/22/facebook-oculus-rift-acquisition-virtual-reality>

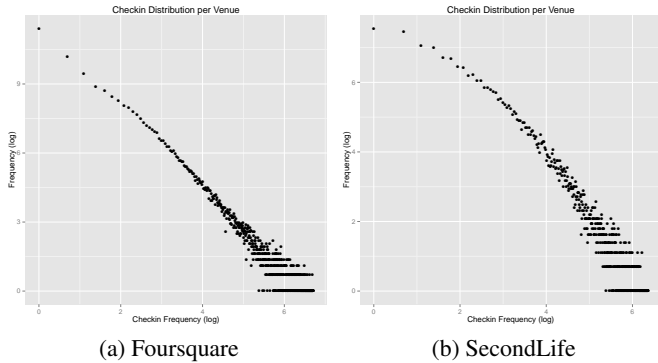


Figure 1: Distribution of Check-Ins per Venue in Foursquare and SecondLife (log-log scale).

In this paper we compared the check-in behavior of users in the virtual world of SecondLife³ (with over 7 million active users one of the largest virtual environments world wide) and the real-world LBSN Foursquare and explored in detail the extent to which place recommender systems that are known to work well on real-world LBSNs are applicable to virtual worlds.

Research Question. The research question that this research was driven by is the following: Knowing that Foursquare and the virtual world of SecondLife are comparable environments, to what extent are current real-world LBSN place recommender algorithms applicable to SecondLife?

Results. Based on a number of experiments, we find that while both worlds are indeed comparable in some aspects, such as the check-in distributions of Foursquare and SecondLife users, recommender algorithms purely based on a distance function however are not suitable for the virtual world of SecondLife, i.e., they do not have comparable performances. We also evaluated state-of-the-art collaborative-filtering algorithms based on user-based k-nearest neighbors (KNN) and matrix factorization and found out that these algorithms have comparable performances in Foursquare and SecondLife.

Contributions. The main contributions of this work are: (i) the study of current real-world LBSN recommender algorithms in the virtual world context and (ii) the insight that, although comparable in some aspects as suggested by related works, some of the assumptions upon which location recommenders of real-world LBSN recommenders operate, does not seem to hold in SecondLife. Another contribution of this work is the introduction and provision of a novel data set that allows the study of human mobility patterns in virtual worlds.

2. EXPERIMENTAL SETUP

In this section we describe in detail our experimental setup, i.e., the data sets and recommender approaches, as well as the evaluation protocol used to investigate our research question.

2.1 Datasets

Our study relies on data from Foursquare containing the check-in history of 11,326 users collected from January 2011 to December 2011. This data set was collected and used by the authors of [6] and is publicly available under request at [5]. The SecondLife data⁴

³<http://www.secondlife.com>

⁴The data set can be obtained for free per request.

Data	#users	#venues	#checkins	Sparsity
Foursquare	11,326	176,452	537,877	99.9%
SecondLife	4,771	17,167	120,078	99.8%

Table 1: Basic statistics of the datasets.

was collected from October 2014 to November 2014 containing the check-in history of 34,277 users⁵. Figure 1 depicts the distribution of check-ins per venue in Foursquare (a) and SecondLife (b) in log-log scale. Please note that they are very similar to each other, i.e., some venues are very popular and receive most of the check-ins, whereas the rest lies in the long tail of the distribution. The distributions of check-ins per user in both worlds follow a similar pattern, which we omitted due to space constraints.

2.2 Place Recommender Approaches

A key feature of LBSNs is the availability of geographic meta-data about checked-in places. In Foursquare and SecondLife, for example, the lat-long coordinates of the checked-in places are available. Several place recommender approaches have appeared that exploits geographic metadata of users and venues in order to improve the recommendation quality, under the assumption that users prefer to check-in venues that are nearby the venues they have visited in the past [9, 3, 2, 19].

To test whether this assumption holds in SecondLife, we have chosen two state-of-the-art place recommenders, purely based on geographic distance (aka location-aware), proposed in [19] and [2]. Given that such recommenders work well in Foursquare, as demonstrated in the aforementioned works, it should also work in SecondLife, assuming that the check-in behavior of users in both systems are comparable.

The approaches proposed at [19, 2] assign higher weights to candidate places that are nearby the places that the target user has already checked-in. Ye et al. [19] assume that the pairwise distance distribution of check-ins per user follows a power-law and propose to learn the parameters of the distribution by a simple linear regression on a log-log transformation of the distribution. Thus, the probability of any pair of places x and y being checked-in is computed as follows:

$$P(\text{dist}(x, y)) = a * \text{dist}(x, y)^b$$

$$\log P(\text{dist}(x, y)) = w_0 + w_1 * \log \text{dist}(x, y)$$

where $a = 2^{w_0}$, $b = w_1$ and dist returns the distance between x and y (Great-Circle on Foursquare and Euclidian in SecondLife). Now, the probability that a given user u will check-in venue l is modelled as the conditional probability $P(l|L_u)$, where L_u is the set of all venues checked-in by user u , computed as follows:

$$P(l|L_u) = \prod_{l' \in L_u} P(\text{dist}(l, l')) \quad (1)$$

A ranked list of venues is then generated using Equation 1 above and the top-N are recommended. We will refer to this method as *LinearRegression*.

A similar approach was proposed by [2] where Gaussian kernels were used for modeling the check-in behavior of users. The probability that a given user u will check-in place l is then modeled

⁵Similar to the real world, SecondLife offers since 2009 a location-based service which allows users to check-in places and share this information with friends in a Facebook-like social network called mySecondLife (see [12, 11] for more information).

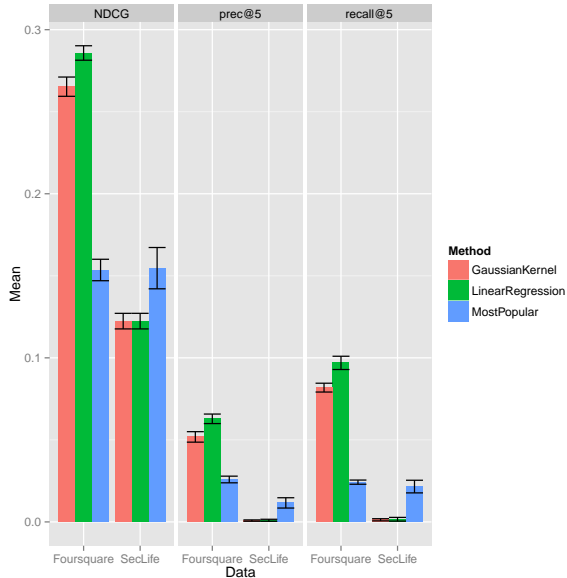


Figure 2: Results of the compared algorithms in terms of precision, recall and NDCG.

as:

$$P(l|L_u) = \frac{1}{|L_u|} \sum_{l' \in L_u} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \text{dist}(l, l')^2} \quad (2)$$

We will refer to this method as *GaussianKernel*.

The algorithms above achieve better performance when combined with collaborative filtering-based algorithms such as the classic user-based KNN or matrix factorization [3, 19]. These algorithms operate under the assumption that users who checked-in the same places in the past tend to check-in the same places in the future. Therefore, we also considered in our study two collaborative filtering algorithms that are solely based on the check-in history of users (i.e. do not use geographic information, but only the user-place interactions in terms of check-ins) observed in the training data, namely, the user-based KNN and a state-of-the-art based on matrix factorization known as BPRMF (Bayesian Personalized Ranking Matrix Factorization) [10].

To run the collaborative-filtering algorithms we have used the recommender systems library MyMediaLite [4] adopting the default settings for the collaborative-filtering recommenders. The location-aware algorithms were implemented on top of MyMediaLite in order to reuse the evaluation functions available.

2.3 Evaluation Protocol

For computing the recommendations we considered only the users with 10 or more check-ins in distinct venues. We split the data randomly keeping 90% of users' check-ins for training and the rest for testing. We repeated this process 5 times in order to measure some variability in the results. Table 1 summarizes the training data after this pre-processing. To evaluate the recommenders we have used prec@5, recall@5 and NDCG (Normalized Discounted Cumulative Gain), which are typical evaluation metrics in the recommender systems field [8].

3. RESULTS

Figure 2 shows the results of the location-aware recommenders in comparison to the most popular recommender on both data sets.

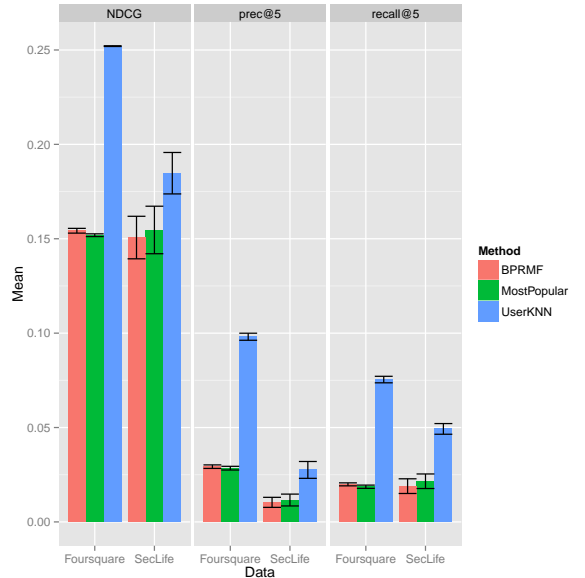


Figure 3: Comparison of pure CF-based algorithms for place recommendations in SecondLife and Foursquare.

The most popular simply recommend the most frequently checked-in venues for all users and is usually used as a baseline for testing new recommender algorithms. Please note that while in Foursquare the location-aware recommender is clearly better than the Most-Popular, in SecondLife the location-aware has worse or equal performance than the most popular approach.

In general, location algorithms based on geographic distances alone, like the ones we are using here, do not produce highly accurate recommendations. Instead, they are used as additional contextual signals on an ensemble of recommenders that usually include collaborative-filtering algorithms.

The fact that such location-aware algorithm is better than the most popular recommender in Foursquare provides a clear indication about the importance of this kind of signal for predicting users' check-in behavior, in contrast to SecondLife where this does not seem to be the case.

In order to better understand the aforementioned results, we plotted the pairwise distance distribution between the users' check-ins in Figure 4. In the case of the Foursquare distribution (Figure 4.a), there is a clear pattern indicating that users prefer to check-in venues within a small distance from his/her past check-ins. Please note that such pattern does not appear in the SecondLife distribution (Figure 4.b). In fact, in SecondLife users seem to prefer visiting places that are farther apart from each other. A possible explanation to this lies in the fact that SecondLife users can easily teleport anywhere in the virtual world map, thus diminishing their distance dependence when deciding to visit a new place.

Figure 3 shows the results of the collaborative-filtering algorithms. Note that in this case the recommender performances are comparable. The best result comes from UserKNN, whereas BPRMF and the MostPopular have comparable performance. In case of Foursquare, the performance can be improved by combining the location-aware and collaborative filtering recommenders. As for SecondLife, this would probably not work, which suggests that new approaches for modeling the mobility patterns of SecondLife users are needed, putting, for instance, a higher preference to places which are farther away from each other.

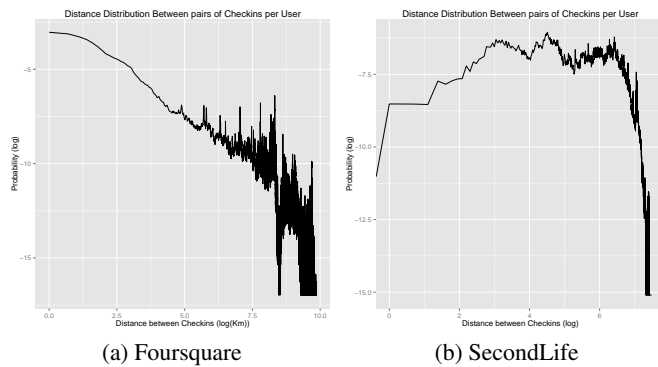


Figure 4: Distribution of pairwise check-in distances per user.

4. CONCLUSIONS & FUTURE WORK

Being rich data environments, many research works have appeared recently with the aim of using virtual worlds as proxies for studying and predicting real world social and behavioral characteristics of humans. But for this to work, it is required that human behaviors in virtual spaces and their counterparts in the offline world are similar. This requirement is known as the Mapping Principle and has to be established on a case-by-case basis, i.e., not all virtual worlds and virtual behaviors map to the offline world [18]. In this paper, we also contribute to this strand of research by (i) comparing the check-in behavior of SecondLife and Foursquare users and (ii) evaluating the extent to which real-world LBSN recommender systems are useful in virtual worlds such as SecondLife. Our main findings are summarized below:

- The check-in distribution, per venue and user, of SecondLife and Foursquare are very similar.
- While Foursquare users tend to check-in places that are nearby the places they visited in the past, SecondLife users present the opposite behavior, tending to check-in places far apart from one another.
- Most popular and collaborative-filtering algorithms, based on KNN and matrix factorization, have comparable performances considering both environments.

For future work we plan to investigate alternative approaches for modeling the distance dependence of users in virtual worlds, as well as other factors that influence virtual world users' decision on checking-in a new place. In this context we plan to investigate recommender models for other interesting entities such as events [14] and along other virtual environments.

Acknowledgments. This work was carried out during the tenure of an ERCIM “Alain Bensoussan” fellowship program. This work is also supported by the National Institute of Science and Technology for Software Engineering (INES), funded by CNPq and FACEPE, grants 573964/2008-4 and APQ-1037-1.03/08.

5. REFERENCES

- [1] M. Ahmad, C. Shen, J. Srivastava, and N. Contractor. *Predicting Real World Behaviors from Virtual World Data*. Springer Proceedings in Complexity. Springer International Publishing, 2014.
- [2] S. Author. Omitted for double-blind review. 2013.
- [3] 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. AAAI'12*, 2012.
- [4] Z. Gantner, S. Rendle, C. Freudenthaler, and L. Schmidt-Thieme. MyMediaLite: A free recommender system library. In *Proc. RecSys'11*, 2011.
- [5] H. Gao and H. Liu. Location-based social network data repository, 2014.
- [6] H. Gao, J. Tang, and H. Liu. Exploring social-historical ties on location-based social networks. In *Proc. ICWSM'12*, 2012.
- [7] G. Guo and M. Elgendi. A new recommender system for 3d e-commerce: An eeg based approach. *Journal of Advanced Management Science*, 1(1), 2013.
- [8] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl. Evaluating collaborative filtering recommender systems. *ACM TOIS*, 22(1):5–53, 2004.
- [9] B. Jie, Y. Zheng, and M. F. Mokbel. Location-based and preference-aware recommendation using sparse geo-social networking data. In I. F. Cruz, C. Knoblock, P. Kroeger, E. Tanin, and P. Widmayer, editors, *SIGSPATIAL/GIS*, pages 199–208. ACM, 2012.
- [10] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme. BPR: Bayesian personalized ranking from implicit feedback. In *Proc. UAI '09*, pages 452–461, Arlington, Virginia, United States, 2009. AUAI Press.
- [11] M. Steurer and C. Trattner. Acquaintance or partner?: predicting partnership in online and location-based social networks. In *Proc. ASONAM'13*, pages 372–379. ACM, 2013.
- [12] M. Steurer and C. Trattner. Predicting interactions in online social networks: an experiment in second life. In *Proc. MSM'13*, page 5. ACM, 2013.
- [13] M. Steurer and C. Trattner. Who will interact with whom? a case-study in second life using online social network and location-based social network features to predict interactions between users. In *Ubiquitous Social Media Analysis*, pages 108–127. Springer, 2013.
- [14] M. Steurer, C. Trattner, and F. Kappe. Success factors of events in virtual worlds a case study in second life. In *Proc. NetGames'12*, pages 1–2. IEEE, 2012.
- [15] M. Szell, R. Sinatra, G. Petri, S. Thurner, and V. Latora. Understanding mobility in a social petri dish. *Scientific reports*, 2, 2012.
- [16] C. Trattner, D. Parra, L. Eberhard, and X. Wen. Who will trade with whom?: Predicting buyer-seller interactions in online trading platforms through social networks. In *Proc. WWW'14*, pages 387–388. ACM, 2014.
- [17] C. V. D. Weth, V. Hegde, and M. Hauswirth. Virtual location-based services: Merging the physical and virtual world. In *Proc. ICWS'14*, pages 113–120. IEEE, 2014.
- [18] D. Williams. The mapping principle, and a research framework for virtual worlds. *Communication Theory*, 20(4):451–470, 2010.
- [19] M. Ye, P. Yin, W.-C. Lee, and D. L. Lee. Exploiting geographical influence for collaborative point-of-interest recommendation. In *Proc. SIGIR'11*, pages 325–334. ACM, 2011.
- [20] H. Yin, Y. Sun, B. Cui, Z. Hu, and Chen. Lcars: a location-content-aware recommender system. In *Proc. KDD'13*, pages 221–229. ACM, 2013.