On Privacy in Spatio-Temporal Data: User Identification Using Microblog Data

***Abstract*—Recommendation systems have become extremely common in recent years, and are utilized in a variety of areas to predict the "rating" or "preference" that a user would give to a point of interest (PoI), such as a restaurant, a hotel, or a bar. Such systems typically produce a list of recommendations by considering previous ratings of the user, as well as ratings of other users. Not every person rates every point of interest they visit. In this work, we want to explore the use of spatio-temporal data to improve recommendation systems: We postulate that spatio-temporal user data may indicate the liking or disliking of a point of interest. Clearly, if a user frequently visits the same PoI, stays at the PoI for long times, and is willing to travel a long distance to visit a PoI, that might indicate that user likes that PoI. Thus, we propose to extract user-PoI relation features from spatio-temporal trajectory data only. Using these features, we use out-of-the-box data mining and machine learning solutions, to estimate the popularity of a PoI. Our experimental evaluation shows, that the features extracted from spatio-temporal data able to accurately predict the popularity of a PoI, using ground-truth data from Yelp as a baseline.**

I. Introduction

Dummy reference [1]

II. Related Work

III. Problem Definition

IV. Spatio-Temporal User-Site Feature Extraction

V. User-Site Rating Prediction

VI. Experimental Evaluation

VII. Conclusions

References

[1] J. Liu, F. Zhang, X. Song, Y.-I. Song, C.-Y. Lin, and H.-W. Hon. What's in a name?: an unsupervised approach to link users across communities. In *Proceedings of the sixth ACM international conference on Web search and data mining*, pages 495–504. ACM, 2013.