Usage of annotation tags in the problem of mining similar users

Oleksii Vedernikov
The University of Melbourne
Melbourne, Australia
oleksii.vedernikov@unimelb.edu.au

***Abstract*—This paper presents a novel method of measuring user similarity in *Location-Based Services (LBS)* via relationships between users and annotation tags of locations they attended. Collecting all check-in data together, matrix factorization methods are applied in order to find semantic similarity between tags. Next, an idea of *User Attendance Graph (UAG)* is proposed to represent user check-in history and describe importance of each tag together with transitions between them. Further, *Semantic Behavior Similarity (SBS)* algorithm is proposed to measure likeness between *UAG*. We evaluated this approach with a real dataset collected from Whrrl using *nDCG* measure. Results show efficiency of proposed method for finding *LBS* users with similar behavior, and it can be used in different applications, e.g. friend recommender systems.**

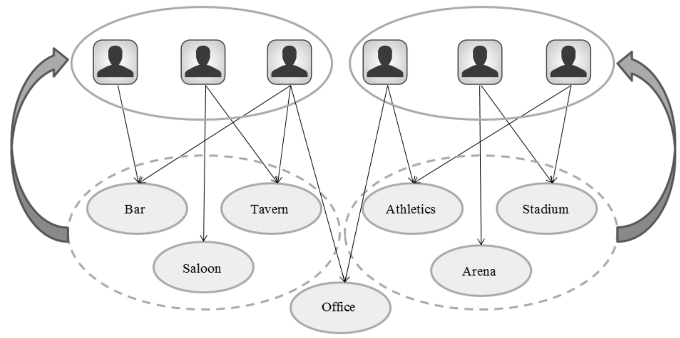


Fig. 1. Communities through user-tag interaction

I. Introduction

According to increasing usage of *Location-Based Services (LBS)*, problem of understanding user behavior arises consequently. Finding of people with similar actions can be used, for example, in friend recommender systems, community discovery[21], sociological researches of distinct human categories, different applications of geo-targeting advertisement. Contemporary approaches of handling this problem include investigation of social, temporal, spatial features. Since modern *LBS* include opportunity to label a place with an annotation tag, there is a possibility to extend possible solutions of this problem. Consider *Ivan* frequently attends *McDonalds* in Kyiv, and *Dai-Li* has a considerable amount of check-ins at *Burger King* in Hsinchu. Their locations are far away in space, annotation tags also differ, but somehow people tend treat them equally - as persons who eat at fast food restaurants. To solve this problem, a novel technique, named *Semantic Behavior Similarity (SBS)*, is proposed to investigate semantic nature of user-tag interaction and find users with similar behavior.

Having initial dataset $D = <U, L, C>$, where U is set of users u_i, $L = \{(l, S)\}$ is set of locations l_i with corresponding set of annotation tags S_i, $C = \{(u, l, t)\}$ is set of check-ins, each of them includes user u_i, location l_i and time stamp ts_i. $TagSet(u_i)$ is a set of all tags from the history of u_i. The first goal is to find a similarity between all pairs of users in a matrix $UserSim$ through similarity of visited locations, and then use it to estimate similarity. Figure 1 illustrates this idea.

Problem of finding similar users is related to search of similar tags. All information about tags can be obtained from users' check-in histories only, and vice versa. Hence, the first issuie is to reveal hidden significant relationships between users and tags of locations they attend, and one possible solution is to use *Singular Value Decomposition* of *User-Tag Matrix (UTM)*, which contains all the location visits information. *SVD* is widely used algorithm of matrix factorizations, and its main applications belong to areas of image processing[2] and text analysis[5]. Another common matrix decomposition method called *CUR* is compared. A hypothesis is that annotation tags describe semantic characteristics of places they are associated with, and the goal is to discover them.

In this case, frequency of tag for a given user and transitions between are the most important. Therefore, the next issue is to represent user check-in history in a structure, which should include information about both these feature. Single visit of place with very popular tag may be not significant, hence weights should be used to describe importance of a tag to a user. Time difference between two check-ins in user history should correspond to transition significance between tags. To handle these issues, an idea of *User Attendance Graphs (UAG)* is proposed, which have weight values for both nodes and edges to represent annotation tags and transitions between them respectively.

The issue of calculating similarity between *UAG* is complicated due to node relationships. As mentioned above, graph nodes (i.e. annotation tags) are similar to each other with some coefficients, which makes existing algorithms of graph comparison inappropriate. A new technique to measure similarity between *UAG* called *Semantic Behavior Similarity (SBS)* is proposed to consider distances between both nodes and edges of two *UAG*.

In order to evaluate proposed approach, *normalized Dis-*

counted cumulative gain (nDCG)[8] is used, making a survey about similarity of different *LBS* users. It is used for two major goals: estimating empirical parameters in this algorithm and verify efficiency of *SBS* algorithm. This work has made the following contributions.

- Semantic similarity between annotation tags from check-in histories of users is measured using matrix factorization method.
- A concept of *User Attendance Graph(UAG)* is introduced to describe not only the annotation tags of places visited by a user, but also significance of them as well as importance of transition between those tags.
- In order to compare two *UAG*, *Semantic Behavior Similarity(SBS)* algorithm is developed, which calculates similarity of user check-in histories using similarity between locations.

The remainder of this paper covers following topics. In Chapter II, existing works are examined. Chapter III describes full algorithm of finding communities from initial data with a representative example. Chapter IV is dedicated to the analysis of results from *Whrrl* dataset, and Chapter V summarizes all the points and provides future directions of this work.

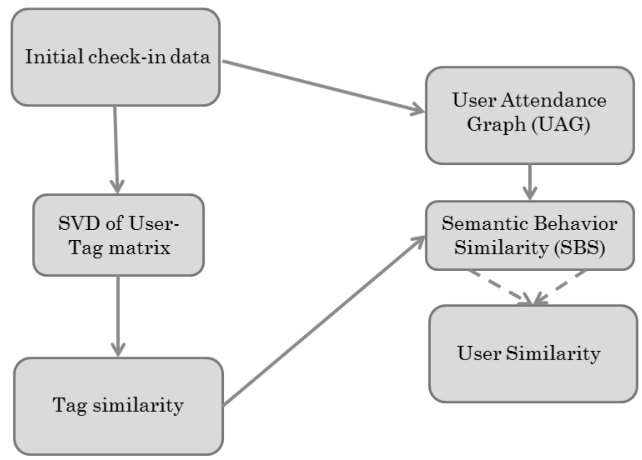
II. Related work

A considerable amount of works are related to mining *LBS* user similarity. Particularly, [19] proposes a *Maximal Semantic Trajectory Pattern Similarity (MSTP-Similarity)* in order to calculate similarity between GPS trajectories, while [16] uses *Maximal travel match* algorithm. Mining user commuities with a similar movement behavior from GPS data is introduced in [7]. However, GPS trajectories are different from check-in history and methods proposed there are inapplicable in this problem. In [3] authors investigate behavioral patterns of doing check-in. Since nobody used annotation tags to measure similarity between users, this idea could be considered as novel and unique.

Studies of user check-ins include [11], where usage of Foursquare check-ins is analyzed in the 24-hours domain for both weekends and weekdays. Additionally, authors investigate which Foursquare categories are used consecutively within certain time interval. [15] uses check-in data for constructing top-*k* popular routes for each user query, which consists of some desirable to visit locations, while [20] discovers check-in data for analysis user preferences and location properties in order to construct recommendation system. In work [17], technique to add annotation tags to places which have no tag is introduced. Mao Ye *et al.* investigate similarity between tags from the same *Whrrl* dataset using distance between temporal distribution [18]. Despite most works consider spatial or social characteristics of human movements, [10] is dedicated to investigation of semantic aspect of *Volunteered Geographic Information (VGI)*, espicially in *Open Street Map*. Some works investigate concept of friendship between users, for example between users of mobile phone[6] or *Online Social Networks* [13]. Link prediction between users of *LBS* is introduced in[14]. These works do not consider usage of annotation tags, which are used for semantic connection of different places. Hovewer, all these works do not research the user-tag relationships with a goal to estimate similarity between tags or users, so they can not handle stated problem.

III. Solution

The proposed solution to discover *SBS* has the following workflow:



The first step is to discover annotation tag similarity from the initial check-in data using *SVD*.

Second, all check-in histories should be mapped into *User Attendance Graphs (UAG)*, which describe importance of places and transitions between them for each user.

Finally, a technique to calculate *SBS* between two *UAG* is proposed, which uses results of two previous steps, and make user similarity matrix. It describes how these users are resembling in their check-in behavior.

A. Tag similarity

As mentioned above, ll the information about tags can be discovered from users' check-in histories only, and vice versa. Their relations are described by a *User-Tag Matrix (UTM)*, where each row represents one tag, and each column corresponds to a user. Each value of *UTM* means how many times user visited all the locations with given tag. Finding tag similarity in a given problem may be related to understanding which users attend some locations more frequent than other, and how it can influence relations between them. One possible approach of finding these hidden links between annotation tags is *Singular Value Decomposition*, what exposes most significant relationships between data items in a given matrix. An assumption is that annotation tags describe semantic characteristics of places they are associated with, and the goal is to investigate them. Proposed method to measure similarity between tags is inherited from *Latent Text Analysis* method [5], which investigates relationships between terms in a document collection. One important parameter in *SVD* is k, number of used dimensions of decomposed matrix, and this parameter should be estimated for each problem and dataset particularly. Check-in histories of users may vary significantly, so they are normalized using TF-IDF weighting, which is widely used in related problems. After that, given matrix is decomposed using *SVD* or *CUR* and cosine distance between tags is calculated, so $TagSim$ - matrix of similarities between tags - is obtained, where all the values belong to the interval $[0, 1]$. In our experiments, we use 3 distance functions: *cosine*, *Jaccard* and *Euclidean*.

B. UAG construction

While investigating tags, the order of users visited them is inessential. However, transitions between tags may contain some important information about user preferences, and temporal difference between check-ins is crucial in this case. Therefore, cosine or other distance between user columns is not complete, in contrast with document similarity in *LSA*. Regarding that a narrow interval between check-ins corresponds to a complete shift, and a wide one does not imply any meaning, a linear function to measure transition coefficient is used:

$$Tran = \begin{cases} 0 & \text{, for } \Delta_t > t_{max} \\ \frac{\Delta_t - t_{min}}{t_{max} - t_{min}} & \text{, for } t_{min} < \Delta_t \leq t_{max} \\ 1 & \text{, for } \Delta_t \leq t_{min} \end{cases}$$

where estimation of parameters t_{min} and t_{max} is described below. Rare check-ins imply unimportance of transitions for the user. All transition coefficients from one tag to another are summed up and divided to the maximum possible amount of junctions, making transition coefficient matrices $Trans(u_i)$ for all users.

Tag significance is indicated by TF-IDF weights from the previous step, so a vector $TagImp(u_i) = (\,\omega_j(u_i)\,)$ is used, where $\omega_j(u_i)$ is TF-IDF for tag a_j for user u_i. Therefore, *UAG* is a graph with weighted nodes and edges, some of them may be missing.

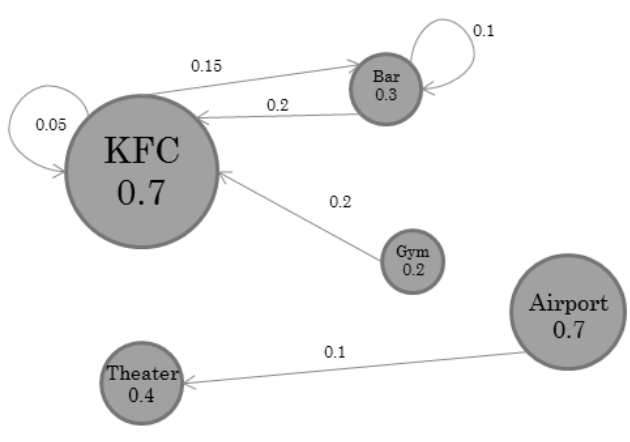


Fig. 2. *UAG* example

C. User similarity

The task to calculate similarity between *UAG* is difficult due to their structures. As mentioned above, graph nodes are dependent from each other. Consider example of one user, who makes check-ins between *Church* and *Mall*, and other, that attends *Temple* and places with tag *Shopping*. Generally, these tags are different, but empirically we understand their likeness. To handle this problem, a new technique to measure similarity between two *UAG* called *Semantic Behavior Similarity (SBS)* is proposed. For two given *UAG*, similarities of nodes and edges of *UAG* are combined, which represent tags and transitions between them.

$$SBS(u_i, u_j) = \mu \cdot SBSTag(u_i, u_j) + (1 - \mu) \cdot SBSTrans(u_i, u_j) \quad (1)$$

Weight μ will be estimated using *nDCG* [8]. Detailed approach is described below.

For given pair of users (u_i, u_j) tags from their check-in history are united to a set $TagSet(u_i, u_j) = TagSet(u_i) \cup TagSet(u_j)$. Then a matrix $TagSim(u_i, u_j)$ is extracted, which is a part from all-tag matrix $TagSim$, but includes only tags $TagSet(u_i, u_j)$ that are considered at this moment. In the same way two transition matrices are created for both users $TagTrans(u_i), TagTrans(u_j)$. Because in a general case $TagSet(u_i) \neq TagSet(u_j)$, they are adjusted by adding zeros to all values corresponding to $TagSet(u_i, u_j) \setminus TagSet(u_i)$ and $TagSet(u_i, u_j) \setminus TagSet(u_j)$ respectively .Program realization uses sparse matrices, hence zeros in matrices don't affect effectiveness of an *SBS* algorithm.

First, a likeness between nodes is estimated. Calculation of $SBSTag(u_i, u_j)$ is not a simple distance between $TagImp(u_i)$ and $TagImp(u_j)$, because likeness between these tags may vary. Therefore, their similarity should be considered, and adjustment operation described above should be used to make vectors $TagImp(u_i)$ and $TagImp(u_j)$, which have the same length $|TagSet(u_i, u_j)|$ and zeros on the places of remaining tags.

Consequently, $SBSTag(u_i, u_j)$ is computed by the normalized Euclidean distance between these vectors:

$$SBSTag(u_i, u_j) = \|(TagSim(u_i, u_j) \cdot TagImp(u_i)\,, TagSim(u_i, u_j) \cdot TagImp(u_j))\| \quad (2)$$

Second, edges difference is measured. Transition similarity is calculated in the same way, but instead of Euclidean distance for vectors Frobenius norm is used for adjacent matrices of transitions:

$$SBSTrans(u_i, u_j) = \|\,(\,TagSim(u_i, u_j) \cdot TagTrans(u_i) \cdot TagSim(u_i, u_j))\,, TagSim(u_i, u_j) \cdot TagTrans(u_j) \cdot) TagSim(u_i, u_j)\,)\| \quad (3)$$

D. Example

This approach is illustrated by the following example. There is a set of 8 users who visited places with 8 annotation tags, and goal is to estimate *SBS* similarity between them.

First, an example of *SVD* is used in problem of finding similar tags. Consider 8 users $u_1, u_2, \dots u_8$ and 8 annotation tags: *gym, swimming, stadium, sports, bar, night club, chillout, drinking*. Numbers in initial *User-Tag Matrix* are given in Table I indicate how many times user attended places with a given annotation tag.

tag/user	u1	u2	u3	u4	u5	u6	u7	u8
gym	8			6				1
swimming		4	6			1	1	
stadium	5	2	3	6				1
sports	5	4	6	3	1	1		
bar			1		7			11
night club		2				7	6	
chillout		1	2		5	5	5	6
drinking	1			1	8	7	4	7

TABLE I: Initial User-Tag Matrix

Both groups of users and tags we can divide intuitively to 2 groups of supporters of healthy lifestyle and party fans. However, it's only human assumption about their semantic difference, and *SBS* algorithm recognizes it automatically.

The cosine distance between rows is given in Table II. Note that 0 means that tags are similar, while 1 refers to totally distinct tags.

Tag	gym	swimming	stadium	sports	bar	night club	chillout	drinking
gym	0.0	1	0.1	0.35	0.87	0.1	0.9	0.83
swimming	1	0.0	0.62	0.27	0.95	0.74	0.72	0.88
stadium	0.1	0.62	0.0	0.15	0.82	0.98	0.82	0.82
sports	0.34	0.27	0.15	0.0	0.9	0.86	0.76	0.8
bar	0.87	0.95	0.82	0.90	0.0	1	0.19	0.2
night club	0.1	0.74	0.98	0.86	1	0.0	0.43	0.47
chillout	0.9	0.72	0.82	0.76	0.19	0.44	0.0	0.33
drinking	0.83	0.88	0.82	0.8	0.2	0.47	0.33	0.0

TABLE II: Initial Tag-Tag cosine distances

The table shows that tags *swimming* and *gym* are totally different, because there are no people that attend swimming pools and gyms in a given dataset. The same situation appears with tags *bar* and *night club*. However, a guess is that in real they may have some relationship, and *SVD* is useful instrument to discover these hidden relationships between users and locations. The more users visit a group of locations with specific semantic meaning (e.g. sport venues or relaxing night spots), the more likely these groups of tags will be tied between. After using *SVD* of initial matrix distances between tags will change, which is illustrated in Table III.

From the results, the distance between pairs of tags *(bar, night club)* and *(swimming, gym)* decreased, which emphasizes their dependence. Therefore, these two pairs of users tend to visit resembling places, and *SVD* reveals their connection. In addition, this example shows that given method is redundant to noise, and rare check-ins with values 1-2 in the table do not noticeably affect the result.

Tag	gym	swimming	stadium	sports	bar	night club	chillout	drinking
gym	0.00	0.00	0.01	0.00	0.84	0.49	0.70	1.00
swimming	0.00	0.00	0.01	0.00	0.87	0.52	0.73	1.00
stadium	0.01	0.01	0.00	0.01	0.70	0.38	0.57	0.86
sports	0.00	0.00	0.01	0.00	0.86	0.51	0.71	1.00
bar	0.84	0.87	0.70	0.86	0.00	0.07	0.01	0.01
night club	0.49	0.52	0.38	0.51	0.07	0.00	0.02	0.14
chillout	0.70	0.73	0.57	0.71	0.01	0.02	0.00	0.05
drinking	1.00	1.00	0.86	1.00	0.01	0.14	0.05	0.00

TABLE III: Final Tag-Tag cosine distances

Consider users u_2 and u_3. From Table I one can see that they attend 2 distinct sets of tags, and therefore should be different. However, compute weights of *UAG* nodes:

$TagImp(u_2) = \Big(0, 1.48, 0, 0.84, 0, 0.42, 0\Big)^\top$

$TagImp(u_3) = \Big(1.46, 0, 1.46, 0, 0.24, 0, 0.24\Big)^\top$

Note that value for 8th tag *drinking* is deleted for convenience due to its absence in check-in histories of both users u_2 and u_3. $TagSim(u_2, u_3)$ also contains tags 1-7 only. Then the formula mentioned above can be used to calculate $SBSTag(u_2, u_3)$ in III-D.

$$
\begin{aligned}
&SBSTag(u_2, u_3) = \\
&\left\| \left(\begin{pmatrix} 0 & 0.01 & 0 & 0.84 & 0.49 & 0.7 \\ 0 & 0.01 & 0 & 0.87 & 0.52 & 0.73 \\ 0.01 & & 0.01 & 0.7 & 0.38 & 0.57 \\ 00. & 0.01 & 0 & 0.86 & 0.51 & 0.71 \\ 0.87 & 0.7 & 0.86 & 0 & 0.07 & 0.01 \\ 0.52 & 0.38 & 0.51 & 0.07 & 0 & 0.02 \\ 1 & 0.86 & 1.02 & 0.01 & 0.14 & 0.05 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1.48 \\ 0 \\ 0.84 \\ 0 \\ 0.42 \\ 0 \end{pmatrix} - \right.\right. \\
&\left.\left. \begin{pmatrix} 0 & 0.01 & 0 & 0.84 & 0.49 & 0.7 \\ 0 & 0.01 & 0 & 0.87 & 0.52 & 0.73 \\ 0.01 & & 0.01 & 0.7 & 0.38 & 0.57 \\ 00. & 0.01 & 0 & 0.86 & 0.51 & 0.71 \\ 0.87 & 0.7 & 0.86 & 0 & 0.07 & 0.01 \\ 0.52 & 0.38 & 0.51 & 0.07 & 0 & 0.02 \\ 1.03 & 0.86 & 1.02 & 0.01 & 0.14 & 0.05 \end{pmatrix} \cdot \begin{pmatrix} 0.46 \\ 0 \\ 1.46 \\ 0 \\ 0.24 \\ 0 \\ 0.24 \end{pmatrix} \right) \right\| \\
&= 1.08
\end{aligned}
\tag{4}
$$

Note that these values represent distance between two vectors, and the less it is, the more similar users are. $SBSTrans(u_i, u_j)$ is calculated in a similar way, considering transition time between location and calculating Frobenius norm between matrices instead of Euclidean distance between vectors, after that normalize all the values and find resulting similarity scores.

IV. Evaluation

Because *SBS* method is based on tag similarity, it should be investigated first. For example, consider most similar and different tags for tag *Restaurant* in *Table IV*:

similar	different
Pho	Alternative
Hawaiian	Hungarian
Tours	Teams
Tempura	Video
Juices	Appliances
Wine	Rinks
Sum	Fitness
Wineries	Pets
Smoothies	Skating
Diner	Books

TABLE IV: Top-10 similar and dissimilar tags for *Restaurant*

Hence, most of similar tags are related to food *(Pho, Tempura, Juices, Wine, Wineries, Smoothies, Diner,)*, while dissimilar tags annotate locations which aren't connected with meal. These results state that $User - Tag$ matrix contains hidden semantic relationships between tags, and *SVD* can successfully investigate them.

Although results could be estimated only in empirical way and it's impossible to compare this algorithm with previous results due to its novelty, experiments show that proposed *SBS* method is accurate and able to find users with similar semantic behavior. For estimating empirical results, *nDCG* method is used, where a survey was conducted about 12 users from *Whrrl* dataset and asked how similar these users are according to $[0..3]$ scale, where 0 means users are different, and 3 corresponds to similar users. In the survey, both tags and time difference between them are included. After calculating average weights from user survey, experiments were made regarding different parameters from this algorithm: t_{max}, t_{min}, μ, k. In order to simplify calculation, $t_{max} = 6t_{min}$. First, t_{min} and k were estimated for different values of each other and k. Despite results don't vary significantly, after fixing $t_{min} = 120min, \mu = 0.9$ efficiency of algorithm for some values of k extended 90%, see in Fig.5.

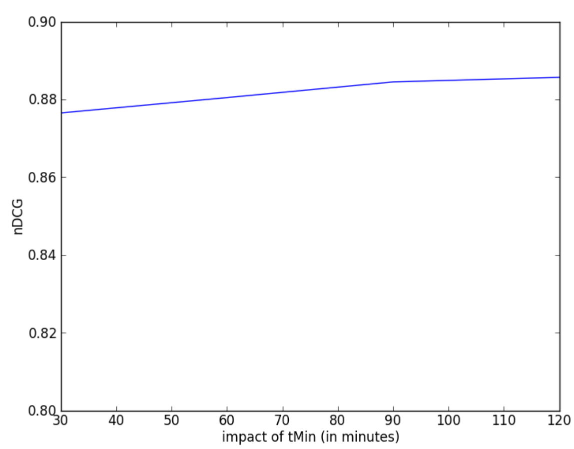


Fig. 3. Impact of t_{min}

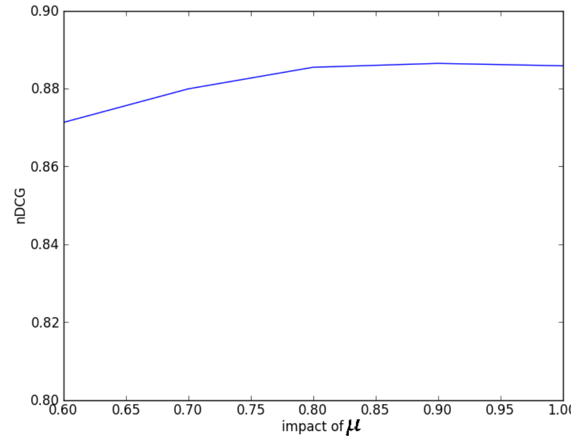


Fig. 4. Impact of μ

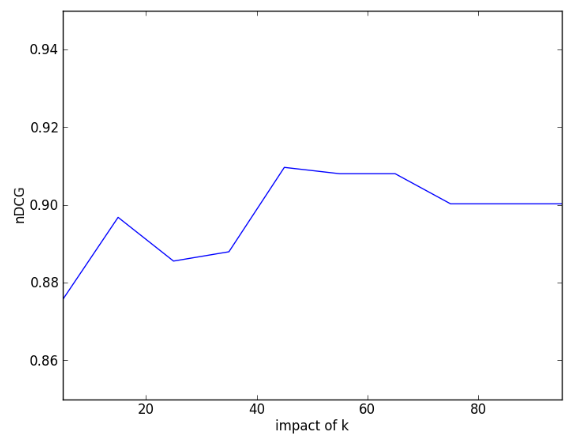


Fig. 5. Impact of k

In another experiment set, we compared 3 distance functions *euclidean, cosine, jaccard*, 3 time functions *linear, step, exponential* and 2 matrix decomposition methods *SVD, CUR* together with using only distance for measuring likeness between columns in *User-Tag* matrix. For results of those experiments, see Fig.6, Fig.7, Fig.8.

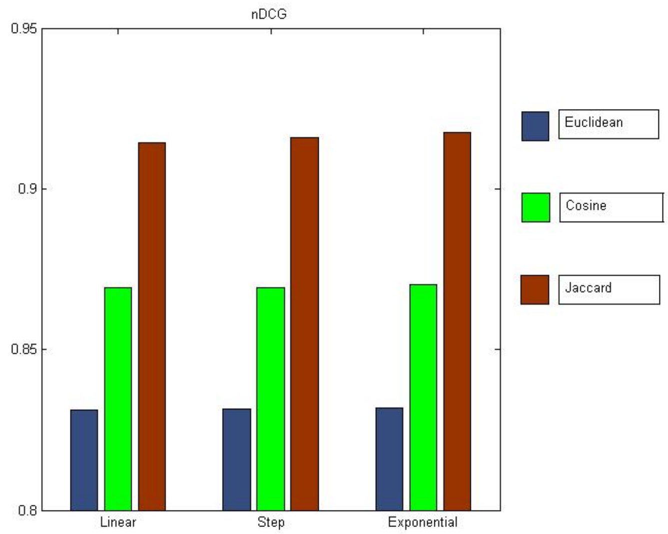


Fig. 6. Time functions vs. distances

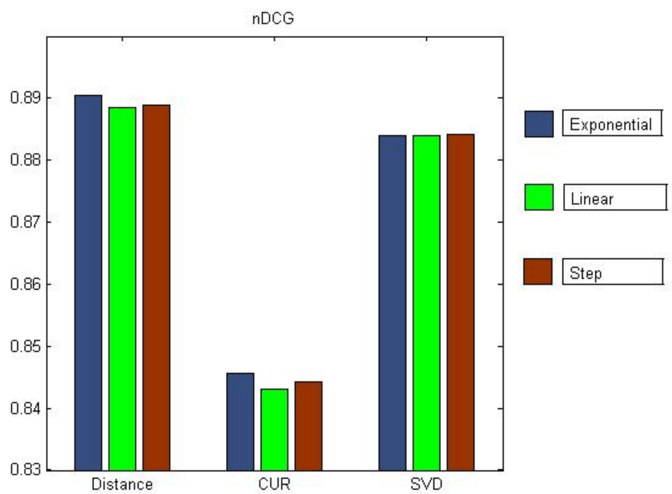


Fig. 7. Matrix factorizations vs. time functions

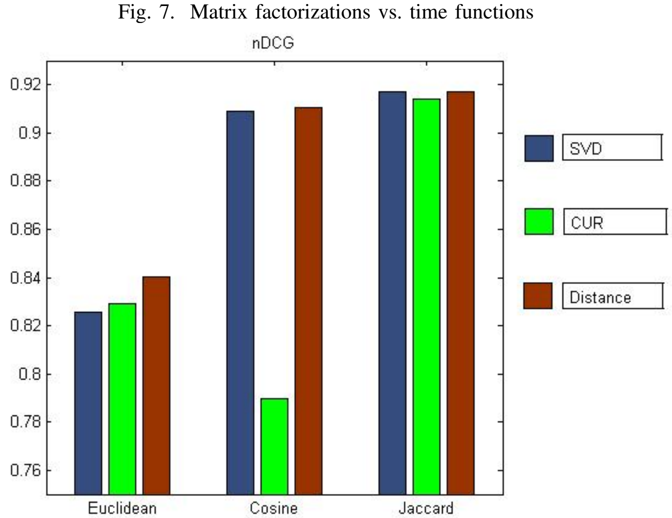


Fig. 8. Distances vs. matrix factorizations

As we see, SVD performs better than CUR, time functions do not affect final result significantly, and $Jaccard$ and $Cosine$ distances show better results than simple $Euclidean$ distance. Therefore, conducted experiments state that overall accuracy of given method reaches 90%, that verifies that proposed *SBS* algorithm can find semantic similarity of users as people do empirically, and therefore it is almost proper in a given problem.

Therefore, the experiments state that overall accuracy of this method reaches 90%, that verifies that proposed *SBS* algorithm can find semantic similarity of users as people do empirically, and therefore it is almost proper in a given problem.

V. Conclusion

In this paper, a novel technique to measure similarity of users in Location-Based social networks based on annotation tag use is presented. We estimated how these tags differ using matrix factorization methods, developed an idea of *User Attendance Graph* to represent user behavior in tag space, and created *Semantic Behavior Similarity* method between check-in histories of different user. Through a number of experiments on *Whrrl* dataset, effectiveness of proposed method is validated for measuring user similarity in Location-Based Services and reaches 90%. Therefore, it can be applied in current Recommender Systems or other services together with spatial and social approaches, which is the main future direction of this work.

References

[1] B. Berjani and T. Strufe, "A recommendation system for spots in location-based online social networks", 4th Workshop on Social Network Systems, 2011.

[2] Chen J., (2000), "Image compression with SVD", ECS 289K Scientific Computation

[3] H. Cramer, M. Rost, and L. Holmquist, "Performing a check-in: emerging practices, norms and 'conflicts' in location-sharing using foursquare", Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, 30.08-02.09, 2011, Stockholm, Sweden

[4] Justin Cranshaw, Eran Toch, Jason Hong, Aniket Kittur, and Norman Sadeh, "Bridging the gap between physical location and online social networks", Proceedings of the 12th ACM international conference on Ubiquitous computing, September 26-29, 2010, Copenhagen, Denmark

[5] S.Dumais, G. Furnas, T. Landauer, S. Deerwester, and R. Harshman, "Using latent semantic analysis to improve access to textual information." In Proceedings, CHI '88. 281-285.

[6] N. Eagle, A. Pentland, and D. Lazer, "Inferring friendship network structure by using mobile phone data", Proceedings of the Natural Academy of Sciences, Jan 2009

[7] C.-C. Hung , C.-W. Chang, and W.-C. Peng, "Mining Trajectory Profiles for Discovering User Communities" Proceedings of the 1st Workshop on Location-Based Social Networks (in conjunction with ACM GIS), Seattle, USA, Nov. 3-6, 2009

[8] K. Jrvelin and J. Keklinen, "Cumulated gain-based evaluation of IR techniques, ACM Transactions on Information Systems (TOIS)", v.20 n.4, p.422-446, October 2002

[9] Q. Li, Y. Zheng, X. Xie, Y. Chen, W. Liu, W.-Y. Ma, "Mining user similarity based on location history, Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems, November 05-07, 2008, Irvine, California

[10] C. Mlligann, K. Janowicz, M. Ye, and W.-C. Lee, "Analyzing the Spatial-Semantic Interaction of Points of Interest in Volunteered Geographic Information", Spatial Information Theory, 10th International Conference, COSIT 2011, pages 350-370, 2011.

[11] A. Noulas, S. Scellato, C. Mascolo, and M. Pontil, ”An empirical study of geographic user activity patterns in foursquare” In Proc. of the 5th Int’l AAAI Conference on Weblogs and Social Media (2011), pp. 570-573
[12] P. Symeonidis, A. Papadimitriou, Y. Manolopoulos, P. Senkul, and I. Toroslu, ”Geo-social recommendations based on incremental tensor reduction and local path traversal”, GIS-LBSN 2011: pages 89-96.
[13] S. Scellato, C. Mascolo, M. Musolesi, and V. Latora, ”Distance Matters: Geo-social Metrics for Online Social Networks”, In Proceedings of WOSN’ 10, June 2010.
[14] S. Scellato, A. Noulas, and C. Mascolo. ”Exploiting place features in link prediction on location-based social networks”, In ACM SIGKDD, 2011
[15] L.-Y. Wei, Y. Zheng, and W.-C. Peng, ”Constructing Popular Routes from Uncertain Trajectories”, KDD 2012
[16] X. Xiao, Y. Zheng, Q. Luo, and X. Xie, ”Finding similar users using category-based location history”, in Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, ser. GIS ’10. New York, NY, USA: ACM, 2010, pp. 442-445.
[17] M. Ye, D. Shou, W.-C. Lee, P. Yin, and K. Janowicz, ”On the Semantic Annotation of Places in Location-Based Social Networks”, 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 520-528. ACM, 2011.
[18] M. Ye, K. Janowicz, C. Mlligann, and W.-C. Lee, ”What you are is when you are: the temporal dimension of feature types in location-based social networks”, Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, November 01-04, 2011, Chicago, Illinois
[19] J. Ying, E. Lu, W.-C. Lee, T.-C. Weng, and V. Tseng, ”Mining User Similarity from Semantic Trajectories”. In LBSN , 2010.
[20] J. Ying, E. Lu, W. Kuo, and V. Tseng, ”Urban point-of-interest recommendation by mining user check-in behaviors”, UrbComp ’12 Proceedings of the ACM SIGKDD International Workshop on Urban Computing, pages 63-70, ACM New York, NY, USA 2012.
[21] Y. Zheng , X. Zhou, ”Computing with Spatial Trajectories”, Springer Publishing Company, Incorporated, 2011