Reviewing: 1

Comments to the Author

This paper focuses on the problem of shop recommendation, a subtype of POI recommendation. The problem is very important and useful in the mobile internet era with the development of smart phones. Generally, this paper is technically sound. But, it lacks of technical novelty and contribution. It combines several advanced recommendation techniques, including Session-based temporal graph (STG), BRP and Tuck decomposition approach. Considering this journal is focused on the practice and experience, I think the technical contribution aspect of this paper is fine. The paper is well-written and is easy to follow.

However, there are some issues needed to address in the revision.

1. In Equation (8), the term “bias\_{ui}” is precomputed by the STG method. Thus, the term is a constant. As a constant, it cannot be regularized.

Answer: The constant term “bias\_{ui}”should not be regularized, we revise the objective function and remove the regularization term “bias\_{ui}”. See equation (8).

1. In Algorithm 1, the authors chooses to randomly initialize the latent factors P and Q. According to my experience, the BPR model with the random initialization performs poorly. It can achieve better performance with the initialized parameters following Gaussian distribution.

Answer: Yes we achieve better performance by initializing P and Q by Gaussian distribution. And we find that the result will be even better if P and Q are initialized by N~(0,1)/10. See Algorithm“MF-STG-BPR”.

(3) I do not think that the repetitive shop recommendation is necessary in the reality. The problem of new shop recommendation is more meaningful. As a recommender system, its function is to help users explore and find new items that are hard to find. If the user has known the item, it is a waste of resources to still recommend this item. Please refer to the following literature:

Hongzhi Yin, Bin Cui, JingLi, Junjie Yao, Chen Chen. "Challenging the Long Tail Recommendation". VLDB 12.

Answer: Actually all of our proposed methods do aim to recommend new shops. At the same time we fully consider the possible repetitive recommendation problem and deal with it by designing rating update rules. We revise the abstract and introduction part. The paper "Challenging the Long Tail Recommendation" is referred in Section 1, paragraph 1.

(4) The proposed Tensor-STG achieves the best performance. It is hard to judge the reasons. Is it due to the integration of STG or the tensor factorization method? It is suggested to compare with the following tensor factorization-based recommendation method.

Temporal Collaborative Filtering with Bayesian Probabilistic Tensor Factorization. SDM 2010.

Answer: In the last version, we use HR and ARHR as the metrics. In the revision, we use leave-one-out HR and ARHR, so we redo all the experiment. The new results are shown in Table II “Baseline Algorithm Comparison”. We notice that MF-STG-BPR performs the best while Tensor-STG doesn’t perform as well as MF-STG and MF-STG-BPR. We investigate it and find that the advantage of Tensor-STG is to predict customers’ future ratings based on their past ratings. Since the leave-one-out HR and ARHR evaluation method randomly take one non-zero rating from the dataset as the validation set, the training set and the validation set are not strictly time evolved, thus decreases prediction accuracy. We provide another set of recall and precision metrics to support this, see Figure 5 in section 4.2.3. Tensor-STG performs the best. We find that Tensor-STG can make more accurate prediction given time evolving dataset, and at the same time, it is more sensitive to the dataset.

The “BPTF”(Bayesian Probabilistic Tensor Factorization) approach is introduced in section 4.2.2 and tested in section 4.2.3.

(5) As shop recommendations belong to POI (point of interest) recommendation, many important related work is missing. Please refer to them in the revision.

1) Hongzhi Yin, Bin Cui, Yizhou Sun, Zhiting Hu, Ling Chen. "LCARS: A Spatial Item Recommender System". ACM Transaction on Information Systems. 2014

2) Hongzhi Yin, Xiaofang Zhou, Yingxia Shao, Hao Wang, Shazia Sadiq. "Joint Modeling of User Check-in Behaviors for Point-of-Interest Recommendation ". The 24th ACM International Conference on Information and Knowledge Management (CIKM'15), Melbourne, Australia, October, 2015.

3) Weiqing Wang, Hongzhi Yin\*, Ling Chen, Yizhou Sun, Shazia Sadiq, Xiaofang Zhou. "Geo-SAGE: A Geographical Sparse Additive Generative Model for Spatial Item Recommendation" . Proc. of 2015 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'15), Sydney, Australia, August, 2015.

4) Hongzhi Yin, Bin Cui, Zi Huang, Weiqing Wang, Xian Wu, Xiaofang Zhou. "Joint Modeling of Users’ Interests and Mobility Patterns for Point-of-Interest Recommendation ". The 2015 ACM Multimedia Conference(ACM-MM'15), Brisbane, Australia, October, 2015

5) Hongzhi Yin, Yizhou Sun, Bin Cui, Zhiting Hu, Ling Chen. "LCARS: A Location-Content-Aware Recommender System". Proc. of 2013 ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD’13), Chicago, IL, Aug. 2013

6) Hongzhi Yin, Bin Cui, Xiaofang Zhou, Weiqing Wang, Zi Huang, Shazia Sadiq. "Joint Modeling of User Check-in Behaviors for Real-time Point-of-Interest Recommendation". ACM Transaction on Information Systems. 2016. (TOIS'16)

Answer: All of these six papers are referred in Section 2.

Reviewing: 2

Comments to the Author

This article presents a new approach for shop recommendations where the aim is to establish predictive models in order to solve temporal dynamics with re-evaluation of customer’s rating by using update rules for solving repetitive recommendations. The authors pointed to three main contributions. The first one is related to matrix factorisation. The second is Bayesian Personalized Ranking from explicit feedback by Matrix Factorisation – Session-based Temporal Graph (MF-STG). Finally, the third one is the design rating update rules and mine customer RFID trajectories to solve repetitive recommendation.

General view. The article should be rechecked for the English and also for the typos. Example: page 3, section 3, first paragraph, second line, “The we expound” should be “Then, we expound”, in the same paragraph “Furthermore, we established context-ware”, it should be “context-aware” ect. Section 3, entitled approaches, I suggest the authors to rewrite it using short sentences rather than long ones and maybe it will be interesting to add a schema in order to make it easy for the reader to understand the whole architecture of their approach and also to understand how the algorithms developed interact with each other.

In section 3.5, the authors describe how they divided the customer’s RFID trajectory data and how they used it. However, the description provided is not enough. There is still some ambiguity in understanding how exactly they make use of the RFID trajectories. There is also a typo “Apriori” instead of “A priori” in the same section. I suggest the authors to provide more description about the methodology and also the nature of the data itself.

Answer:

1. Section 3, first paragraph, “The we expound🡪 “Then we expound”
2. Section 3, “Furthermore, we established context-ware”🡪 “context-aware”
3. We reorganize the entitled approaches in Section 3, we add content of subsection 3.1 and use a figure to describe the architecture. Subsection 3.1 explains the architecture of shop recommendation.
4. We use capital letters “APRIORI”to present the algorithm. We give an example of mining a customer’s maximum frequent visited areas in section 3.6. See figure 4. In section 4.1, we give samples of data.

Reviewing: 3

Comments to the Author

Authors propose three approaches for shop recommendation, including MF-STG, MF-STG-BPR, and Tensor-STG. In addition, they exploit customer historical trajectory information to revamp rating value. Experimental results show the prospect of their approaches. The contributions of this paper are satisfactory. However, there are also some problems.

Major problems:

The motivations of proposed approaches are vague. I suppose there are some relevance between them. In MF-STG, they fuse STG into Matrix Factorization Model, and the output is the rating prediction r’\_{ui}, which can be seen as the value of preference. But in the second approach MF-STG-BPR, the outputs are latent vectors P\_{u}, Q\_{i}, and Q\_{j}, which are also used to predict rating value. From the formula above Equation (15), we can see the \hat{x}\_{uij} just the difference value between predict ratings. So I’m confused with the relevance between MF-STG and MF-STG-BPR. The same thing is with Tensor-STG. I strongly suggest that authors should give the clear motivations and interpretations of the relevance between these approaches.

Answer: We revise the title of Algorithm 1 as “The Learning Algorithm of MF-STG-BPR”, the outputs are parameters P\_{u}, Q\_{i}, and Q\_{j} , the rating prediction r’\_{uv} that represents user u rates on item v is computed by : r’\_{uv} = P\_{u}\* Q\_{v} + tau \* bias\_{uv}. See the last paragraph in section 3.4.

For Tensor-STG, we add Figure 4 “Tuck Decomposition of a Three-way Array”and the elementwise calculation of Tucker Decomposition. We reorganized section 3.5.3 to explain how to use STG result and Tuck Decomposition to make prediction. We add a new section 3.1 to introduce the architecture of shop recommendation. In the first paragraph of section 3.1, we explain the common ground and the differences among these approaches. In the last paragraph of section 3.5.3, we further explain the difference between Tensor-STG and other two approaches.

Minor problems:

1. Please give the full name at the first time using abbreviation especially in Abstract and Introduction.

Answer: We rewrite the abstract and we revise paragraph 2 ,3 and 4 in section 1. We give the full name or make explanation at the first time using abbreviation.

1. The contributions presented in Introduction should be improved. Authors should not only extract the contributions but also explain why it is the contribution.

Answer: We revise paragraph 2 ,3 and 4 in section 1. We add paragraph 5 to explain why it is the contribution. See the follows,

“The graph computing for temporal dynamics is intuitive and easy to interpretation, in our approaches, all customers’ long- and short-term preferences are estimated in the form of matrix which is calculated by STG. The flexible matrix factorization and tensor decomposition method nicely fuse this estimated preference matrix and achieve good performance. The BPR method that utilizes the estimated preference matrix further improve the ranking result. The indoor RFID trajectories information is special and helpful when making indoor shop recommendation. The experimental results also show that mining RFID trajectories do help to improve recommendation accuracy.”

1. The difference between the proposed approaches and related works should be described.

Answer: We explain the differences in the last paragraph of section 2. As the follows,

“Different with LBSN based POI recommendation that puts emphasis on exploiting users' shared experiences and tips for recommendation, Our proposed approaches for indoor shop recommendation focus on predicting customers' evolving preferences. POI recommendations utilize social information, geographical information and user "check-in" information to improve prediction accuracy, we use indoor RFID trajectories and design rating update rules to improve prediction accuracy.”

4. The references are old and not sufficient. The latest reference was published in 2014. Most of the references were published before 2010. There are some recent and important related works on recommender system should be cited, including shop recommendation [1]-[7], POI recommendation [8]-[11].

1) W. Zhao, Y. Guo, Y. He, et al., “We Know What You Want to Buy: A Demographic-based System for Product Recommendation on Microblogs,” ACM KDD’14, 2014.

2) S. Ahn, A. Korattikara, N. Liu, et al., “Large-Scale Distributed Bayesian Matrix Factorization using Stochastic Gradient MCMC,” ACM KDD’15, 2015.

3) L. Hu, A. Sun, Y. Liu, “Your Neighbors Affect Your Ratings: On Geographical Neighborhood Influence to Rating Prediction,” ACM SIGIR’14, 2014.

4) N. Kawamae, “Real-time Recommendations from Connoisseurs,” ACM KDD’15, 2015.

5) H. Wang, N. Wang, D. Yeung, “Collaborative Deep Learning for Recommender Systems, ACM KDD’15, 2015.

6) G. Zhao, X. Qian, X. Xie, “User-Service Rating Prediction by Exploring Social Users’ Rating Behaviors,” IEEE Trans. Multimedia, 2016.

7) G. Zhao, X. Qian, and C. Kang, “Service Rating Prediction by Exploring Social Mobile Users’ Geographic Locations,” IEEE Trans. Big Data, 2016.

8) J. Bao, Y. Zheng, et al. "Recommendations in location-based social networks: a survey." Geoinformatica, 19.3(2015):525-565.

9) S. Jiang, X. Qian, T. Mei, Y. Fu, “Personalized Travel Sequence Recommendation on Multi-Source Big Social Media,” IEEE Trans. Big Data, 2016.

10) P. Lou, G. Zhao, X. Qian, et al., "Schedule a Rich Sentimental Travel via Sentimental POI Mining and Recommendation," in Proc. BigMM, 2016.

11) S. Jiang, X. Qian, J. Shen, Y. Fu, and T. Mei, “Author Topic Model-based Collaborative Filtering for Personalized POI Recommendations”, IEEE Trans. Multimedia, 2015.

Answer: All of the eleven paper have been cited in section 2.

5.Complexity analysis of these approaches should be given.

Answer: We supplement section 3.7 as complexity analysis.

6.Below Equation (16), since P, Q, R are matrices, then what’s the meaning of R\_{I×P}, R\_{J×Q}, and R\_{K×R}?

Answer: We explain P,Q and R. In this subsector, they are integers that represent the indices of a tensor. See line 6 to line 8 in section 3.5.2.

“Here, I,J and K are integers that represent the indices of tensor \chi. For i = 1,2...,I,j =

1,2...,J,k = 1,2...,K, the element (i,j,k) of tensor \chi is denoted by x\_{ijk}.”

We also add figure 4 to illustrate Tuck decomposition.

7. Above Equation (17), authors said “if a customer consumed in a shop and his consumption is lower than his historical average level, his potential interest in this shop should be adjusted higher”. Why his interest should be adjusted higher while his consumption is low?

Answer: The rating adjust rule is only applied on shops that the customer has consumed. We make a more clear explanation on it. The intuition of designing rating update rules is to deal with the possible repetitive problem.

8. For the metrics of this paper, are HR and ARHR authoritative? Or just proposed in this paper. If related works used HR and ARHR, authors should cited them.

Answer: The metrics of HR and ARHR were proposed in paper “Item-Based Top-N Recommendation Algorithms, MUKUND DESHPANDE and GEORGE KARYPIS , ACM Transactions on Information Systems, Vol. 22, No. 1, January 2004. ” In paragraph 2, section 6.1 “Experimental Design and Metrics”, the authors give the explanation of HR and ARHR. See the follows copies.

“*The quality was measured by looking at the number of hits and their position within the top-N items that were recommended by a particular scheme. The number of hits is the number of items in the test set that were also present in the top-N recommended items returned for each user.*”

And the authors supplement that,

“*In order to ensure that our results are not sensitive to the particular training-test partitioning of each dataset, for each of the experiments we performed ten different runs, each time using a different random partitioning into training and test sets.*”

The authors adopt the following methods:

“*To evaluate the quality of the top-N recommendations, we split each of the datasets into a training and test set by randomly selecting one of the nonzero entries of each row to be part of the test set, and used the remaining entries for training.*”

In the last version of this paper, we separate out training set and validation set by time sequence strictly and then test HR and ARHR. At first the intuition is to test the recommendation accuracy if we use customer’s past behaviors to predict their future behavior. But we notice the possible problem that the particular training and test partitioning of dataset may cause sensitive result. So in the revision, we use **5-time leave-one-out cross validation and we redo all the experiment**, see section 4.2.3. The metrics are still HR and ARHR. Since leave-one-out HR and ARHR are common metrics used in many related papers such as,

“SLIM : Sparse Linear Methods for Top-N Recommender Systems, Xia Ning and George Karypis, ICDM 2011”.

We also cite this paper in section 4.2.2.

9. In performance comparison, authors should show us the effect of their proposed factors. For example, the performance comparison between the method of considering Long- and Short-term Preferences and just Long-term Preference, just Short-term Preferences, and no preference. The effect of Rating Update Rules Using Context-aware is also should be discussed.

Answer: In section 4.2, we compare the performance of considering Long- and Short-term preferences , just Long-term Preference , just Short-term Preferences and no preference. We also compare the performance of considering RFID trajectories only, rating update rule only and by both of them. The results are shown in figure 8 and figure 9.

10. I strongly suggest that authors should improve their Figures, especially for Figure 3, 4, and 5. Figure 4 and 5 are upturned.

Answer: All the figures in section 4.2 have been replaced.

Reviewing: 4

Comments to the Author

In this paper, authors present a set of three methods for the top-N recommendation problem. The presented methods are extension and amalgamation of existing methods (viz. MF, STG and BPR). The basic idea in the proposed method is to utilize the STG method to precompute the user-item bias matrix and then using it with MF, BPR and Tensor decomposition to provide top-N item (shop) recommendations. The proposed approach is interesting and the presented results show that the proposed method outperforms a set of baselines in terms of hit rate and ARHR.

However, there are many shortcomings to the proposed approach:

\* The proposed approach is not significantly novel. It is a simple extension (and amalgamation of existing approaches).

Answer: We understand your concern about novelty. As you are the expert in machine learning area, you have authoritative perspective on novelty. Our proposed approaches are the extension and amalgamation of existing approaches, from machine learning point of view, these methods may not be novel, but

(1) Our methods are different with the existing approaches. The MF-STG method nicely absorb the advantage of MF for its flexibility and the advantage of STG for its good performance on processing temporal dynamics. The BPR based method utilizes the bias term and provide further accurate prediction on ranking. The Tensor-STG method is different with the existing methods that are good at restoring the tensor while hard to perform ranking task, we utilize the STG result matrix to predict customers’ preference in a new time window in tensor structure.

(2) The proposed methods have novelty for application. We give the architecture for shop recommendation. The proposed methods outperform the state-of-the-art algorithms. Furthermore, the application of mining RFID trajectories to improve recommendation accuracy is first proposed.

(3) The proposed methods have the value in practical application and easy to deploy. Nowadays shop recommendation is useful in the mobile internet era, the RFID technology has been maturely deployed in large shopping center. The proposed methods provide effective ways for practice.

\* The writing and overall flow of the manuscript is confusing and hard to follow. It needs a major overhaul to make the manuscript easy to understand.

Answer: We have carefully revised the whole paper to make better presentation.

\* There is no clear motivation for the need of proposed methods. The results section just presents the results and does not discuss them.

Answer: We add according contents in section 1 and section 2. See the follows,

“The graph computing for temporal dynamics is intuitive and easy to interpretation, in our approaches, all customers’ long- and short-term preferences are estimated in the form of matrix which is calculated by STG. The flexible matrix factorization and tensor decomposition method nicely fuse this estimated preference matrix and achieve good performance. The BPR method that utilizes the estimated preference matrix further improve the ranking result. The Tensor-STG aims to make rating prediction in a new time window on three dimensional tensor structure, which has better explanation to the result. The indoor RFID trajectories information is special and helpful when making indoor shop recommendation. The experimental results also show that mining RFID trajectories do help to improve recommendation accuracy.”

“Different with LBSN based POI recommendation that puts emphasis on exploiting users’shared experiences and tips for recommendation, Our proposed approaches for indoor shop recommendation focus on predicting customers’evolving preferences. POI recommendations utilize social information, geographical information and user ”check-in” information to improve prediction accuracy, we use indoor RFID trajectories and design rating update rules to improve prediction accuracy. The fusion of matrix factorization and STG nicely absorb the advantage of MF for its flexibility and the advantage of STG for its good performance on processing temporal dynamics. The result of STG as a bias matrix for MF can overcome the disadvantage of STG that the determination of time window size may affect recommendation accuracy and computational complexity. The MF-STG method can provide more stable and flexible solution without paying attention to time window size selection in STG. The adoption of BPR method helps to get more accurate ranking result. The Tensor-STG method is different with the existing methods that are good at restoring the tensor while hard to perform ranking task, we utilize the STG result matrix to predict customers preference in a new time window in tensor structure.”

The results section, the discussion is added in paragraph 2-7, section 4.2.3.

Following are the specific issues:

\* Opening paragraph in section 3 is hard to follow. Please consider changing the flow.

Answer: We rewrite the opening paragraph in section 3. We add a new subsection 3.1 “The Architecture of Proposed Shop Recommendation”to introduce the system architecture, the data flow, the relevance and differences of proposed algorithms and the outputs.

\* In Section 3.2.1, it is not clear what "sessions are inflexible" in STG means.

Answer: We make the following explanation:

“When creating session nodes, the determination of time window size has obvious influence on recommendation accuracy and computational complexity. The smaller size of time window achieves better accuracy while increasing complexity. There is no flexible way to determine the size sessions.”

\* In Section 3.2.1, it is not clear what PATH means. Does it always start with a user node and end in an item node? Better notation instead of just PATH would make it clear.

Answer: We explain it as the follows,

“Define set PATH is the paths that source node v\_0 to an unknown item node v\_n.”

“we can find that the source point could be a user node or a session node, the end point is an item node.”

\* Looks like majority of section 3.2.1 is reproduced from the original STG paper. Section 3.3 also has the same issue, it is reproduced from the BPR paper.

Answer: In section 3.3.1 (corresponds to section 3.2.1 in the original paper), we retain the least necessary part to present STG construction, which is in dispensable for the rest of presentation. And for section 3.4(corresponds to section 3.3 in the original paper), we give necessary introduction and deduction of BPR, the latter part of 3.4 is original, including the definition of x^\_{uik}, the gradient, the prediction of rating and Algorithm 1 “The Learning Algorithm of MF-STG-BPR”.

\* Equation 6, what does pre \in mean? (in summation).

Answer: We explain it as the follows,

“Define pre\_{u\!i} as user u's estimated preference on unrated item i, the final calculation of pre\_{u\!i} is defined by Equation (8).”

\* Sec 3.3, \hat{x\_{uij}} is never defined.

Answer: We add the definition of define \hat{x}\_{uij}, see section 3.4, “In order to get user's personalized ranking result, define \hat{x}\_{uij} as user individual probability that user u prefers item i to item j”.

\* Sec 3.3 page 7, the vector notation is not consistent. Upper case and lower case both are used. Q\_i and Q\_j are used for both vectors and the set of items.

Answer: The upper case Q\_i and Q\_j have been replaced by lower case q\_i and q\_j in section 3.4.

\* Sec 3.4.3: It is not clear what the different dimensions of the tensor represent?

Answer: We explain it in section 3.5.3“Construct a Customer-Shop-Time three dimensional tensor \chi which is composed of slices of \chi\_{::t\_1}, \chi\_{::t\_2},......,\chi\_{::t\_{prediction}}.”

\* Sec 3.5: It is not clear what "context-aware" mean.

Answer: We explain “context-aware”in introduction and in section 3.6.

“The context-aware information includes two parts, customers’ past RFID(radio frequency identificationdevices) trajectory data and customers’ past consumption.“

\* Sec 4.1: Data attributes are missing. #rows, #cols, density?

Answer: In section 4.1, we give details of test data.

“Our test dataset is a small part of the whole data. The test dataset contains customer RFID trajectories records and customer historical consumption records spanning 16 consecutive months. The total number of customers is 300, the total number of shops is 280. There are 2019 consumption records and the total amount of customer consumption is 1,826,028 RMB. If the consumption records are aggregated into a customer-shop matrix, the number of converted ratings is 1309, the sparsity is 1.56%”.

\* Sec 4.2.3: What is the hyper parameter search space for each of the baseline algorithms? Given that many of the algorithms are very sensitive to the values of the hyper parameters, if a proper parameter search method is not employed, then the results might be the best one.

Answer: We make explanation in section 4.3.2, see as follows,

“We use a grid search to find the most suitable parameter. The search space is from 0 to 1 and the increment is 0.1. For UserKNN and ItemKNN, the number of k is searched from 10 to 100 and the increment is 10. For the number of latent factors in SVD++ and FISMrmse is searched from 10 to 100 and the increment is 10.”

\* Figure 3: A multi line chart might be better to present the same information.

Answer: Multi line charts Figure 6 and figure 7 have been replaced in section 4.2.3.

\* Figure 4 and 5 are mirror images and cannot be followed.

Answer: All the figures in section 4.2 have been replaced.

Possible typos:

\* Page 3:

\*\* Line 18: The -> Then?

\*\* Line 19: Expound?

\*\* Line 44: 't' is part of the expression A(ui)

\* Page 4:

\*\* Line 7: shot -> short

\* Page 5:

\*\* Line 24: constrain -> constraint

\* Section 3.3 - all equations are not numbered.

\* Page 7:

\*\* Line 25: "Latent Feature Number" -> "Number of Latent Features"

Answer: We fixed the above seven typo problems, all equations are numbered in this paper except some expressions.

\* Equation 16: what do those operators mean? x\_1, x\_2, x\_3

Answer: We revise the explanation of tucker decomposition. We use Figure 3 to illustrate tucker decomposition and give equation (31) to computer the elementwise of tensor \chi. See 3.5.2.

\* Page 9:

\*\* Line 42: 'p' -> 'p\_i'

Answer: It has been fixed.

Overall the presented approach is not significantly novel in extending the state-of-the-art in top-N recommender algorithms and the writing needs a significant improvement.

Answer: Replied in the first and second point.