A FILTERING BUBBLE TECHNIQUE FOR A RECOMMENDER SYSTEM USING COT MODEL

Dharani. C[1], Mahalashmy. M[2] and Rupa Kesavan[3], Kapila Vani. R. K[4]

[1, 2] *Department of Computer Science and Engineering, Prince Shri Venkateshwara Padmavathy Engineering College, Chennai, India.*
[3]*Assistant Professor, Department of Computer Science and Engineering, Prince Shri Venkateshwara Padmavathy Engineering College, Chennai, India*
[4]*Assistant Professor, Department of Computer Science and Engineering, Prince Dr.K.Vasudevan College of Engineering and Technology, Chennai, India.*

ABSTRACT

With rapid growth of information on the internet, recommender systems become fundamental for helping users alleviate the problem of information overload. Since contextual information can be used as a significant factor in modeling user behavior, various context-aware recommendation methods are proposed. However, the state-of-the-art context modeling methods treat contexts as other dimensions similar to the dimensions of users and items, and cannot capture the special semantic operation of contexts. On the other hand, some works on multi-domain relation prediction can be used for the context-aware recommendation, but they have problems in generating recommendation under a large amount of contextual information. In this work, we propose Contextual Operating Tensor (COT) model, which represents the common semantic effects of contexts as a contextual operating tensor and represents a context as a latent vector. Then, to model the semantic operation of a context combination, we generate contextual operating matrix from the contextual operating tensor and latent vectors of Contexts. Thus latent vectors of users and items can be operated by the contextual operating matrices. Experimental results show that the proposed COT model yields significant improvements over the competitive compared methods on three typical dataset, i.e., Food, Accommodation and Movielens-1 Mdatasets.

Keywords: *Recommender systems, context aware recommendations, contextual operation, context representation.*

---

1. INTRODUCTION

With rapid growth of available information on the internet, users are getting in trouble with the problem of information overload. Recommender systems have become important for helping user to select interesting information in many Web applications such as social networks, e-commerce, online reading, and review websites and so on. Nowadays, with enhanced ability of systems in collecting information, a great amount of contextual information in recommender systems has been collected. The contextual information in real-world application includes location,time,weather,companion and so on.These kinds of contexts have significant effect on the user behavior.For instance, a man may like to watch cartoon with his children while may like to watch romantic movies with his wife.He may prefer to read novels during weekend while may tend to read professional books during weekdays.

Due to the fundamental effect of contextual information in recommender systems, many different kinds of context modeling methods have been developed.These methods incorporate transfer matrix to map latent vectors of entities from one domain to another domain .To deal with context-aware recommendation, using transfer matrix,latent vector scan be mapped from one context to another context.However,since there are multiple contexts in the real world(e.g.,location,time, companion),using a transfer matrix for each specific context combination,these methods have problem in confronting with large amount of contextual information.

To overcome the shortages mentioned above,we propose a novel context modeling method,which uses contextua operating tensor to capture the operation of contexts.For instance,in the phrase "excellent product",the noun"product"has a latent vector,while the adjective"excellent"has the semantic operating matrix which operate the vector of the noun"product".Thus, the phrase"excellent product"has a new latent vector which represents a positive attitude to the "product".In this work,we assume that context combinations have similar properties of adjectives and can operate the latent characteristics of entities.For instance,when a man is with children,the context of this companion operates his latent interests,and then he may like to watch cartoons.Here,for the context combination in a user-item rating behavior,we use contextual operating matrix to represent the semantic operation.The main contributions of this work are listed as follows:

- We model the contextual information as the semantic operationon entities,which presents a novel perspective on modeling of the contextual information.
- We use contextual operating tensor to capture the common semantic effects of contexts,and latent vectors to capture the specific properties of contexts. Then the contextual operating matrix can be generated from them.
- Experiments conducted on three real datasets show thatCOT are effective and evidently out performs the state-of-the-art context-aware models.

2. RELATED WORKS

2.1. Context-aware Recommender Systems

Contextual information has been proved to be useful for recommender systems and these methods can be categorized into pre-filtering,post-filtering and context modeling.Employing pre-filtering or post-filtering strategies,conventional methods utilize contextual information to drive data selection or adjust the resulting set.These ad-hoc methods may work in practice,but they requiret he supervision and the fine-tuning in all steps of recommendation.

The context model in approaches, which use the contextual information directly in the model,have become popular.Recent works on context modeling have focused on integrating contextual information with user item rating matrix and building models based on factorization models.Multi-verse recommendation represents the user-item rating matrix with contextual information as a user-item-context rating ensor. And FM is easily applicable to a wide variety of context by specifying only the input data. However,these methods treat contexts as one or several dimensions besides the dimensions of users and items,and cannot capture the special semantic operation of contexts.

Research on multi-domain relation prediction can also be used for he context-aware recommendation. Collective Matrix Factorization (CMF) factorizes the rating matrixin each domain,and the latent vectors of some entities are shared among different domains.CMFcan also be applied in social rating networks and attribute-aware relation prediction. For context-aware recommendation, with a transfer matrix for each context combination,these methods cannot be implemented for real applications with large amount of contextual information.

3. PROPOSED COT MODEL

In this section,we introduce our proposed contextual operatingtensor model. We introduce the notations at first,then present the proposed model thoroughly,and finally describe the process of parameter inference.

3.1. Notations

In typical recommender systems, there are users and items denoted by $U=\{u1,u2,...\}$ and $V=\{v1,v2,...\}$. The latent vectors of user I and item j can be denoted by $u_i \in R^d$ and $v_j \in R^d$. There are multiple contexts $C1,...,Cn$, such as location, time, companion and so on. A context value c iss a variable of the context C,and a specific combination $\{c1,k,...,cn,k\}$is named context ombination k. The latent vector of the context value $c_{m,k}$ is denoted as $h_{m,k} \in R^{dc}$, then the context combination k can be represented as a latent matrix $H_k = [h1,k,...,hn,k] \in R^{dc \times n}$. In this work,the rating that user I provides to item j under the context combination can be written as $r_{i,j,k}$.Note that,in some cases,I and j can be the same entity.For example,in social net work, both i and j are users.

3.2. Contextual Operating Matrix

In typical matrix factorization methods, latent vectors of users and items are constant with varying contexts. But in real world applications, user interests and item properties are changed with different context combinations. Here, we use context-specific latent vectors for users and items under different context combinations. Then, the matrix factorization equation can be rewritten as:

$$\hat{r}_{i,j,k} = \omega 0 + \omega_i + \omega_j + \sum_{m=1}^{\cdots} \omega_{m,k} + u^{i}\mathbf{k}$$

where ω0 denotes the global bias, ωi and ωj denote the bi- ases of user i and item j, ωm,k is the bias of the context value cm,k , ui,k and vj,k are context-specific latent vectors of user i and item j under the context combination k. Similar to a phrase of noun and adjective, where the noun

has semantic information and the adjective has semantic operation on the noun, in recommender systems, entities have rich semantic information and contexts act like adjectiveswhich have the operation on entities. For example, companion with children can make users would like to watch cartoons, and romantic films become popular during Valentine's Day. We use two contextual operating matrices for a specific context combination. These matrices describe how a context combination affects the properties of entities and how context-specific latent vectors of users and items can be generated from their original ones. Here, we can calculate context-specific latent vectors as: $u_{i,k} = M_{U,k} u_i$, $v_{j,k} = M_{V,k} v_j$

3.3. Contextual Operating Tensor

As discussed above,with two contextual operating matrices for each context combination,the number of parameters will grow rapidly as the number of context combinations grows.

Besides,different contextssharesimilarcommonsemanticeffects,forexample,both weekendandbeingathomecanmakeyouwouldliketoreadnovels.Contextoperationcanberepresentedasaunityofcommonsemanticeffectsandspecificpropertiesofcontexts.Therefore,itwillbewonderfulandplausibleifwecangeneratecontextualoperating matrices from several basic matrices which represent some common semantic effects of contexts. In this way, we can not only reduce the number of parameters tobe estimated,but also model the underlying characteristics of contexts.Here,incorporating contextual operating tensors,

we have:

$M_{U,k} = a^T T^{[1:d]}$,

$M_{V,k} = a^T T^{[1:d]}$,*where T [1:d]*
k V

[1:d]

c and TV are both d ×d× d tensor
$M_{U,j}$ and $M_{V,j}$are both $d \times d$ matrices,denoting the contextual operating matrices for users and items under the context combination k.

Table1:Performancecomparisononthreedatasetsandtwokindsofsplitting,measuredbyRMSEandMAE(d=8,d_c=4).

	FoodDataset				AdomDataset				Movielens-1M			
	AllUsers		ColdStart		AllUsers		ColdStart		AllUsers		ColdStart	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
SVD++	1.155	0.948	1.278	1.086	2.782	2.093	3.421	2.436	0.942	0.721	1.248	0.956
Multiverse	1.063	0.841	1.121	0.921	1.833	1.383	2.168	1.556	0.883	0.669	1.025	0.771
FM	1.055	0.845	1.115	0.918	1.852	1.446	2.125	1.563	0.878	0.672	1.001	0.766
HeteroMF	1.072	0.862	1.136	0.932	2.084	1.552	2.384	1.782	0.902	0.686	1.072	0.792
COT	**1.002**	**0.792**	**1.098**	**0.898**	**1.726**	**1.367**	**2.056**	**1.518**	**0.841**	**0.645**	**0.987**	**0.759**

- **MultiverseRecommendation:** State-of-the-artmodel andhas beenshown tooutperformothercontext-awarerecommendationmodelsontheFooddatasetandtheAdomdataset.
- **FM**isapplicabletothecontextualinformationbyspecifyingtheinputdata.WealsouseLibFMasitsimplementation.
- HeteroMFusestransfermatrixtomodelthecontextualinformation.EachspecificcontextcombinationhasatransfermatrixinHeteroMF.

A. *EvaluationMetrics*

To measure the performance of rating prediction,we use the most popular metrics,Root MeanS quareError(RMSE)andMean Average Precision(MAE):

$$\text{——}), \quad \overline{(r_{i,j,k}-\hat{r}_{i,j,k})^2}$$

B. *PerformanceComparison*

Table1 illustrates all the experiment results measured by RMSE and MAEonthreedatasetsandtwokindsofsplit-ting. ThistableshowsthatCOTachievesthebestresultsconsistently,whichdemonstratestheeffectivenessofusingcontextualoperatingtensortomodelthecontextualinformation.Onallusersplitting,comparingwiththebestperformanceofothermodels,COTimprovestheRMSEvaluesby5.2%,6.0% and 4.2% on the Food, AdomandMovelen-1M dataset respectively.Oncoldstartsplitting,theimprovementbecome1.5%,3.2%and1.0%.COTimprovesgreatlyontheAdomdataset,whichprovesCOTtobeparticularlyhelpfulforthedatasetwithrichcontextualinformation.Throughalltheexperiments,context-awaremodelsoutperformthecontext-unawaremodelSVD++,whichdemonstratestheimportanceoftheutilizationofcontextualinformationinrecommendersystems.MultiverseRecommendationand FMachieve close performance on all thedatasets.OnFoodandMovielens-1Mdatasets,HeteroMFperformsclosetoMultiverseRecommendationandFM,but fails on the Adom dataset. This may be because HeteroMF needs to estimate too manytransfer matrices with richcon-textualinformationintheAdomdataset. Moreover,wecomputeandillustratetheRMSE

$$MAE=\frac{(i,j,k)\in\Omega_{test}}{n_{test}}$$

Improvement ofthecontext-aware models comparing with SVD. We can observe that, on all three datasets, the RMSE improvements are larger on cold start.

C. *Experiment Methodology*

Toexaminetheperformanceonallusersandcoldstartusers,weadopttwodifferentwaystosplitthedatasets.

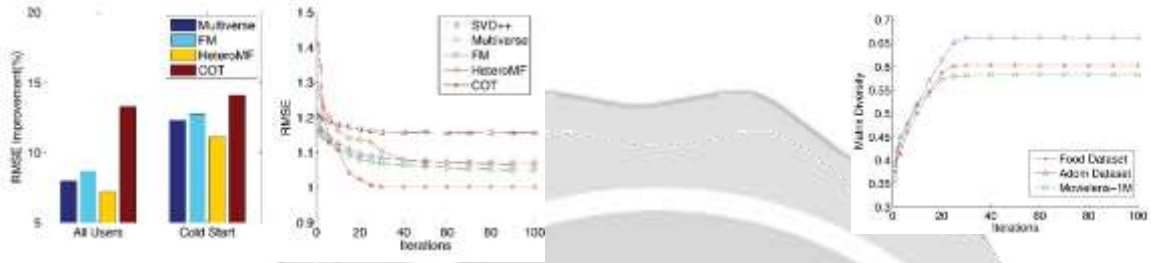
- **All Users**:
- Werandomlysampleabout10%oftheratingsfromthe original dataset tocreate the test set, andtheremaining90%ratingsaretreatedasthetrainset.

ColdStart:Werandomlysamplesomeusersfromtheoriginaldataset then selectless than threeof their ratingsasthe trainset, andleave allremaining ratings asthe testset.Thenumbersofratingsofeachuserintestsetarerandomlydecided.Also,thetestsetcoversabout10%oftheoriginaldataset,andthetrainsetcoversabout90%.splittingthanthoseonallusersplitting.Thisshowsthatthecontextualinformationismoreimportantinthecoldstartsituationandcanbeusedtocompensateforthelackofhistoryinformation

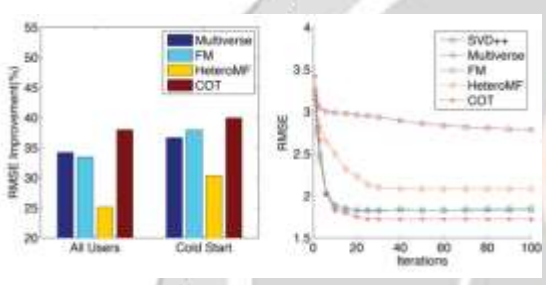
The convergence curve are illustrated in the right part of Figure2. The convergence curves how that the RMSE of COT becomesstableafterabout30iterations.TheseindicatethatCOThasasatisfactoryconvergencerateandcanbetrainedrapidlyandeffectivel yinpracticalapplications.

D. *AnalysisofContexts*

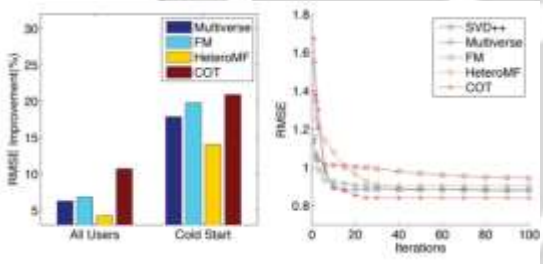
AswediscussedinthesectionoftheCOTmodel,eachsliceofcontextualoperatingtensorrepresentsonekindof commonoperation.Withlargerdifferenceamongalltheslices,thecontextualoperatingtensorismorepowerfulinmodelingthecontexto peration.Similartothecontentdiversitymeasuringthedifferenceamongcontentofmovies we usea metric *matrix diversity* measuring the



(a) FoodDataset



(b) AdomDataset



(c) Movielens-1M

Figure2:Performancecomparisononthreedatasets.TheleftpartillustratestheRMSEimprovementofthecontext-awaremodelscomparedtoSVD++.Therightpartillustratestheconvergencecurvesofthemodels.

differenceamongallslices.Matrixdiversityisdefinedasaverage RMSEofallslicepairsoftensor.Figure3illus-trateshowmatrixdiversityofthecontextualoperatingten-sorschangeswithincreasingnumberofiterations.Thefig-ureshowsthatvaluesofmatrixdiversityincreasewhenthenumber of iterationsgrows from 1to about 30.After theit-eration30,thevaluesofmatrixdiversitybecomestable.Be-sides,wecanseethatCOTachievesconvergenceinFigure2atthesametimeasthevaluesofmatrixdiversityconvergedinFigure3onthreedat asets.Theseevidencesindicatethatinthetrainingprocess,whenmatrixdiversityachievessta-bleresults,COTachievesthebestperformance.Moreover,thefinalvalueofmatrixdiversityontheAdomdatasetislargerthanthoseontheot her twodatasets.Thismaybebe-causerichercontextualinformationontheAdomdatasethasmorepowerfuloperatingabilityinchangingthepropertiesofusersanditems.

Theweightsofdifferentcontextscapturedby$W$onthreedatasetsareillustratedinFigure4.Thefigureshowsthatthecontext*hunger*ismor eimportantthan*virtuality*ontheFooddataset,and*day*ismoreimportantthan*hour*onMovielens-1M.These observations follow ourintuition. For the Adom

Figure1:Matrixdiversityoftensor$T_U$and$T_V$withincreas-ingnumberofiterationsonthreedatasets.

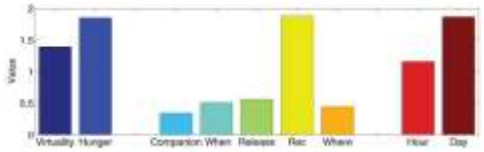


Figure2:Weightsofdifferentcontextsonthreedatasets.

dataset,thecontext*rec*hasmuchhigherweightandbecomesthe dominantcontext. Thismaybe becausethe context*rec*,whichindicateshowwilltheuserrecommendthemovie,hashighrelevancewiththefinalrating.

E. ImpactofParameters

AsshowninFigure5,westudythemodelperformancewithvaryingdimensionalitiesofentityvector$d$andcontextvector$d_c$ontheFooddataset.With increasing$d$andd_c,thevalueofRMSEdecreasesatfirst,thenstaysnearlystableafter$d$=5andd_c=3.Theparameter$d_c$canbeselectedina largerange,whichmeansthattheperformanceofCOTdoesn'trely on parameter selectionverymuch.Since the performancesof COT withdifferentparametervaluesselectedfromtheserangesarevery similar,inprevioussub-section,weonlyillustratethe results with d=8andd_c=4on three datasets for simplicity.

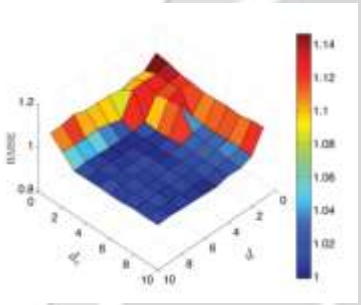


Figure3:*PerformanceofCOTontheFooddatasetwithvaryingdimensionalitiesdand*d_c.

Algorithm Optimization:

Steps:

1.Cascade multiple algorithms to cover different users.

2.Combine algorithm results in different ways, e.g. weighted scores, rank combine.

3.Run A/B and Multivariate tests with no redeploy

4.Select algorithm strategies via API tags:

- to handle user cohorts: mobile users, desktop, tablet
- to provide multiple content recommendations per page, sitewide, insection.

5.Change all configuration in real time with no redeployment.

6.Dynamicoptimisation with multi-armed bandits.

Markov modelsare a form of probabilistic model that can be used to predict elements of a sequence, usually a temporal sequence (e.g., the weather, the stock market, network traffic, web clicks within a site, or purchase patterns). Markov models have a restricted form, mathematically speaking, and this restriction can sometimes make it possible to learn a Markov model from past data (training data). This model can be used to predict probabilities of future events, such as the next most likely thing that the customer will do or buy.

Passive filtering andActive Filtering:

Once you've gathered your data, there are two basic ways of filtering through it to make predictions. The most basic method is passive filtering, which simply uses data aggregates to make method is to predictions (such as the average rating for an item). The more advanceduse active filtering, which uses patterns in user history to make predictions. An example of this would be finding similar users to the current user and using their history to predict a rating.

User-centric vs. Item-centric Filtering:

All recommender systems must decide whether or not it will attempt to see patterns between users or between items. A user-centric system will find similarities between users then use the similar users' preferences to predict ratings. An alternative to this is an item-centric system which will attempt to find the relationships between items, and make predictions then only based on a user's preferences and these relationships.

It is not necessary that a recommender system focus only on users or items, but most typically only find similarities between users or similarities between items and not both.

Memory-Based vs. Model-Based Algorithms:

One big distinction between algorithms is that of memory-based algorithms and model-based algorithms. The basic difference is that memory-based algorithms uses all the data all the time to make predictions, whereas model-based algorithms use the data to learn/train a model which can later be used to make predictions. This means that the memory-based algorithms generally should have all data in memory, whereas model-based can make fast predictions using less data than the original (once you build the model).

The individual advantages and disadvantages of each method will be discussed in their own sections.

Clustering• Clustering algorithms find group of items that are similar • Basically divides a dataset so that records with similar content are in the same group and group are as different as possible from each other • K-Nearest Neighbor – a classification method that clasifies based on calculating the distances between point and other pointsin the training dataset • Example Car Sales

Regression • Deals with prediction ofvalue rather than class • Given x1, x2, x3….. Predict Y • Use Linear regression and predict variables a0, a1, a2… in Y=a0+a1x1+a2x2….. • Use Line fitting, Curve fitting methods • Example find a relationship between smoking patients and cancer related illness

Association Rules

• These algorithms create rules that describe how often events have occurred together • Example when a customer buys a hammer then 90% of the time they buy nails • Spam classification based on conditional probability • Support is a measure of what fraction of the population satisfies both the antecedent and the consequent of the rule • Confidence is the measure of how often the consequent is true when the antecedent is true • Outlier Analysis • Most Data mining methods discard outliers as noise or exceptions • However in some applications such as fraud detection, these rare events can be more interesting.

EXPERIMENTAL PLATFORM

Our experiments were implemented using Java and JSP servlet.We ran our experiments on Netbeans7 and glassfish webserver with Intel Dual Core Processor with a speed of 3.12GB and 2GB RAM.

Hybrid Recommenders

It is natural to consider combining several different recommenderalgorithms into a hybrid recommender system .In some applications, hybrids of various types have been found to outperform individual algorithms . Hybrids can be particularly beneficial when the algorithms involved cover different use cases or different aspects of the data set. For example, item–itemcollaborative filtering suffers when

no one has rated an item yet, but content-based approaches do not. A hybrid recommender could use description text similarity to match the new item with existing items based onmetadata, allowing it to be recommended anyway, and increase theinfluence of collaborative be defined by the

content of the items they like as well as the item themselves.

Grouping them into seven classes:

• Weighted recommenders take the scores produced by several recommenders and combine them to generate a recommendation list (or prediction) for the user.

• Switching recommenders switch between different algorithms and use the algorithm expected to have the best result in a particular context.

• Mixed recommenders present the results of several recommenders together. This is similar to weighting, but the results are not necessarily combined into a single list.

Collaborative Filtering Methods

• Feature-combining recommenders use multiple recommendation data sources as inputs to a single meta-recommender algorithm.

• Cascading recommenders chain the output of one algorithm

u,i (2.21)into the input of another.

• Feature-augmenting recommenders use the output of one

algorithm as one of the input features for another.

• Meta-level recommenders train a model using one algorithm and use that model as input to another algorithm.

Many weighting algorithms use a linear combination of the predictions of many recommenders:

pu,i = α1p

(1)

u,i + ⋯ + αnp

(n)

u,i (2.20)

Feature-weighted linear stacking replaces each coefficient αj with a ratings or genre, to alter the blending ratio of the various algorithms' predictions on an item-by-item basis:

pu,i = g1(i)p

(1)

u,i + ⋯ + gn(i)p

(n)

CONCLUSION

In this paper, a novel context-aware recommendation method,i.e.COT,is proposed. InCOT,the context iscon-verged to be a vector and the common semantic effects ofcontexts arecapturedbyacontextualoperatingtensor.Be-sides,thesemanticoperationofacontextcombinationonentities,i.e., users and items,a contextual operating matrix,which can be generated from the contex-tual operating tensor and latent vectors of contexts. The experimentalresultsonthreerealdatasetsshowthatCOTout-performsthestate-of-the-artcontext aware models and can well reveal the relationshipbetweencontextsandentities.

REFERENCES

Pedro G. Campos Departamento de Sistemas de Información Universidad del Bío-Bío 4081112 Concepción, Chile pgcampos@ubiobio.cl.A criterion based on Fisher's exact test for item splitting in Context-Aware Recommender Systems.2014 33rd International Conference of the Chilean Computer Science Society

Liang He Faqing Wu Department of Computer Science & Technology East China Normal University Shanghai, China..A Time-context-based Collaborative Filtering Algorithm

Qiang Liu, Shu Wu, Liang Wang Center for Research on Intelligent Perception and Computing, National Laboratory of Pattern Recognition Institute of Automation, Chinese Academy of Sciences, China {qiang.liu, shu.wu, wangliang} nlpr.ia.ac.cn.COT:

Contextual Operating Tensor
for Context-Aware Recommender Systems.Proceedings of the Twenty-Ninth AAAI Conference on Artificial