Journal of Korea Multimedia Society Vol. 22, No. 4, April 2019(pp. 515-526)
https://doi.org/10.9717/kmms.2019.22.4.515

Context-Aware Ad Contents Scheduling over DOOH Networks based on Factorization Machine

Van Hoang Nguyen†, Thanh Binh Nguyen††, Sun-Tae Chung†††

ABSTRACT

DOOH(Digital Out Of Home) advertising targets for reaching consumers through outdoor digital display medias. Traditionally, scheduling of advertisement contents over DOOH medias is usually done by operator's strategy, but an efficient ad scheduling strategy is not easy to find under various advertising contexts. In this paper, we present a context-aware factorization machine-based recommendation model for the scheduling under various advertising contexts, and provide analysis for understanding of the contexts' effects on advertising based on the recommendation model. Through simulation results on the dataset adapted from a real dataset of RecSys challenge 2015, it is shown that the proposed model and analysis based on the model will be effective for better scheduling of ad contents under advertising contexts over DOOH networks.

Key words: DOOH, Recommendation, Ad Contents Scheduling, Advertising Contexts, Factorization Machine,

1. INTRODUCTION

Digital Out of Home or DOOH refers to digital media advertising which reaches the consumers while they are outside their homes[1, 2]. This excludes TV and radio advertising, but includes digital signage advertising to people who are in transit, waiting, or at commercial locations over DOOH inventories. DOOH inventories mean DOOH medias such as billboards, kiosks, LCD screens, video walls and whatever are available for displaying digital ad contents. DOOH advertising has been expanded rapidly. The global DOOH market is expected to grow at a compound annual growth rate of 12.6 percent from 2017 to 2023, according to a report by Allied Market Research. It also predicts the market will reach more than \$8 billion by 2023[3].

The challenge of DOOH has been how to reach the right people with the right advertisement contents (ad contents) at the right time. Conventional approach is that an operator schedules distribution of ad contents over selected DOOH medias via content management systems based on numerous user information collected by advertisers or others. Current trends in DOOH advertising move toward a programmatic way, where matching between advertisers's ad contents and DOOH inventories are automatically achieved through a program[1]. This programmatic advertising in DOOH industry has been evolved from online advertising, where real-time bidding (RTB) has been already set up as a major programmatic advertising method on commercial advertising systems, like Google Double

※ Corresponding Author : Sun- Tae Chung, Address: (06978) 369 Sangdo-ro, Dongjak-gu, Seoul, Korea, TEL : +82-2-820-0638, FAX : +82-, E-mail : cst@ssu.ac.kr
Receipt date : Feb. 11, 2019, Revision date : Mar. 22, 2019
Approval date : Apr. 1, 2019

† Dept. of Information and Telecommunication Eng., Graduate School, Soongsil University
(E-mail : hoangdktd@soongsil.ac.kr)
†† CublickDigital, Inc., Seoul
(E-mail : binh.nguyen@sectic.com)
††† Dept. of Intelligent Systems, Graduate School, Soongsil University

Click or Facebook Ads[4]. Matching between ad contents and inventories is a kind of recommendation system like Amazon's product recommendation[5]. Similarly, one can approach programmatic scheduling of ad contents over a group of DOOH medias as a recommendation system. Correspondingly to the rating score in a recommendation system, GRP (Gross Rating Point) is normally adopted as a metric to measure the effectiveness of advertising in DOOH. GRP is defined as an audience size reached by an ad content on a DOOH media[6]. DOOH medias such as kiosks display ad contents on time-slot base. For a given time-slot, candidate ad contents with the highestly predicted GRP could be recommended as the best ad content candidate for the DOOH medias. In more real operation situations, some of candidate recommended ad contents for each DOOH media would be filtered according to marketing strategies, and the ad contents with the best GRP among the remaining candidates will be scheduled to be displayed for the DOOH medias. Then, the issue is how to predict GRPs for ad contents against DOOH medias.

Typical recommendation systems like a product recommendation take care of users and items as major input variables(fields), and predict rating scores of items for users. In DOOH media scheduling, one can consider inventories, ad contents, and GRP, correspondingly to users, items, and rating scores. GRP may not be dependent only on inventories and ad contents, but GRP may be affected also by other advertising contextual information such as location of the DOOH media, content categories, weather conditions, event situation, displaying time and so on.

In this paper, we formulate the scheduling of ad contents under advertising contexts over a group of DOOH medias as a context-aware recommendation problem, where effective prediction of GRPs under contextual information is sought. For that, we present a context-aware factorization machine-based recommendation model for more precise GRP prediction. Through the training and testing processes based on the presented model, we provide a simulation analysis for more precise understanding of each context's effect on advertising by calculating GRP prediction error. Through simulation analyses on the dataset made from adaptation of the real dataset given in an online recommendation competition, RecSys Challenge 2015[7] for our purpose, it is shown that the proposed context-aware factorization machine model and analysis based on the model improves GRP prediction by incorporating contexts and enables better understanding of the effect of each context on advertising.

Even though many of context-aware factorization machine-based approaches have been reported in online recommendation systems, such a context-aware factorization machine-based modeling and analysis have not been reported yet for DOOH ad contents scheduling, to the best of our knowledge.

The rest of the paper is organized as follows. In Section 2, technical backgrounds and related works are described, and in Section 3, a context-aware factorization machine model formulation is provided. Evaluations through simulation analysis are explained in Section 4, and Conclusion and future work are given in Section 5.

2. TECHNICAL BACKGROUNDS AND RELATED WORKS

2.1 Technical Backgrounds

2.1.1 Mathematical Formulation of Collaborative Recommendation System

Recommendation systems provides suggestion for items for a user[8]. Basically, suggestions are based on rating scores and recommendation systems predict rating scores about items (ad contents) for each user (inventory) from the past transaction data.

Then, the collaborative approach to a recommendation system is formulated as a regression task as follows.

For given a set of users $U=\{u_{1,}\cdots,u_m\}$ and a set of candidate items $I=\{i_{1,}\cdots, i_n\}$, target function $y: U\times I \rightarrow R$ (1)
has to be predicted.

The target function represents a rating score, e.g. $y(u,i)$ is a rating score (eg., recommendation score from 0 to 5) of item i for user u. We denote the subset with known rating scores of $U\times I$ by S ($\subset U\times I$) ; that is, for all $(u,i)\in S$, the rating score $y_{(u,i)}$ is known in advance, which is supposed to be available by measurements or some other rating mechanisms.

Then, more clearly, the task of recommendation system can be stated as estimating the target function so as to predict a rating score for any item i for any user u. The estimation of the target function

$$\hat{y}(u,i) \text{ for } all \ (u,i)\in U\times I \tag{2}$$

is obtained by minimizing an estimation error on the measurement dataset, S ($\subset U\times I$).

The performance of a recommendation system is usually evaluated by calculating RMSE(Root Mean Square Error), which is defined as

$$\mathrm{RMSE} := \sqrt{\frac{\sum_{k=1}^{m}\sum_{l=1}^{n}(y(u_k,i_l)-\hat{y}(u_k,i_l))^2}{mn}} \tag{3}$$

where $y(u,i)$ for $(u,i)\not\in S$ is considered as 0.

Since among the domain set of the target function (1), the subset S , a measurement dataset with known rating scores are usually sparse, estimating the target function (2) is a sparse regression problem [9].

2.2.2 Factorization Machine (FM) [9]

Factorization machine formulates the target function in (2) as

$$\hat{y}(x) = w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n}\sum_{j=i+1}^{n}\widehat{w_{i,j}} x_i x_j \tag{4}$$

$$\widehat{w_{i,j}} := <\boldsymbol{v_i}\ ,\ \boldsymbol{v_j}\geqq \sum_{f=1}^{k} v_{i,f}\cdot v_{j,f} \tag{5}$$

where $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ is a feature vector and x_i $(i=1,\cdots,n)$ is an i-th input variable. And, Θ that have to be estimated are:

$$w_0\in R\ ,\ \boldsymbol{w}\in R^n\ ,\ \boldsymbol{V}\in R^{n\times k} \tag{6}$$

where $\boldsymbol{w} := (w_{1,}\cdots,w_n)$ and a i-th row of $\boldsymbol{V}$ is $\boldsymbol{v_i}$ which describes the i-th latent vector with k factors (components). $k\in N_0^+$ is a hyperparameter that defines the dimensionality of the factorization.

A 2-way FM (degree $d=2$ as (4)) captures all single and pairwise interactions of variables with the target.

- w_0 is the global bias.
- w_i models the interaction of the i-th input variable with the target
- $\widehat{w_{i,j}}=<\boldsymbol{v_i}\ ,\ \boldsymbol{v_j}>$ models the factorized interaction of a pair of variables (i-th and j-th variable) with the target.

Instead of using an own independent model parameter $w_{i,j}\in R$ for each interaction, the FM models the interaction by the inner product of two latent vectors associated with input variables. This is the key point which allows high-quality parameter estimation of higher-order interactions $(d\geqq 2)$ under sparsity.

For the task of regression, the most widely used loss function is square loss. To prevent overfitting it is common to add a regularization term - usually L_2. In total, the following regularized least square criterion for optimization is normally adopted:

$$RLS-Opt = \sum_{(\boldsymbol{x},y)\in S} l(\hat{y}(\boldsymbol{x}|\Theta),\ y) + \sum_{\theta\in\Theta}\lambda_{(\theta)}\theta^2 \tag{7}$$

where $l(\hat{y}(\boldsymbol{x}|\Theta),\ y) := (\hat{y}(\boldsymbol{x})-y)^2$

2.2.3 Field-aware Factorization Machine (FFM) [10]

A feature vector $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ can be grouped into fields. Field-aware Factorization Machine

(FFM) is a variant of FM that utilizes field information more. Thus, a target function is formulated as

$$\hat{y}(\boldsymbol{x}) = w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n} \sum_{j=i+1}^{n} < \boldsymbol{v}_{i,f_j}, \ \boldsymbol{v}_{j,f_i} > x_i x_j \quad (8)$$

where f_i and f_j represents respectively the fields that x_i and x_j belongs to.

2.2.4 GRP (Gross Rating Point)

In online situation (web, mobile app), preferable rating scores are recommendation score, or the number of clicks, or others. However, in the DOOH world, which is offline and outdoor, rating is different from online and should be measured by different metrics.

An impression in the context of online advertising is when an ad is fetched from its source, and is countable. Each time an ad is fetched and displayed, it is counted as one impression. Likewise, in DOOH, when an ad is displayed, then it is counted as one impression. Audience impression measures the total number of times people passing a digital out-of-home display are likely to notice the ad contents. When 10 people watch an ad content, then the ad content has 10 audience impression.

GRP, or gross rating point, is a conventional standard metric in DOOH advertising used to measure the size of an advertising campaign by a specific medium or schedule. The purpose of the GRP metric is to measure audience impressions in relation to the number of people in the target for an advertising campaign[6]. GRP values are commonly used by media buyers to compare the advertising strength of an advertising system. GRP is calculated as follows.

*GRPs (%) := 100 * Audience Impressions ÷ Defined population*

2.2 Related Works

A recommender system or a recommendation system (RS) is a subclass of information filtering system that seeks to predict the "rating" or "preference" a user would give to an item [8]. Recommender systems are utilized in a variety of areas including movies, music, news, books, research articles, search queries, social tags, and products in general. There are also recommender systems for experts, collaborators, jokes, restaurants, garments, financial services, life insurance, romantic partners (online dating), and Twitter pages.

Usually, recommendation systems are classified as three approaches; content-based, collaborative, and hybrid. Content-based approach recommends items based on a comparison between the content of the items and a user profile. A user profile has information about a user and their tastes toward items. The content-based recommendation has weakness in that it will only recommend items related to the categories which a user viewed before. This weakness can be remedied by using another variant of recommendation algorithm known as collaborative approach. Collaborative approach is a method of making automatic predictions (filtering) about the interests of a user by collecting preferences or taste information from many users (collaborating). A hybrid recommender system is one that combines multiple techniques together to achieve some synergy between them. Many recommendation systems adopt collaborative approaches since they support mathematically tractable models and produce robust systems, and have shown many successful stories[8].

For a typical collaborative recommendation system as (1), the target function can be represented as a rating matrix R where i-th row of R is modeled as user i and j-th column is modeled as item j and the ij-th component, $R_{i,j}$ of the rating score matrix is represented by $y(u_i, i_j)$. Thus, the matrix factorization method[11], which characterizes both items and users by vectors of factors inferred from item rating patterns. High correspondence between item and user factors leads to a recommendation. These methods have become popular in recent years by combining good scalability with predictive accuracy. In addition, they offer much flexibility for modeling various real-life situation[11].

Many recommender systems simply ignore other contextual information existing alongside user's rating in providing item recommendation. However, by pervasive availability of contextual information such as time, location, social information, and type of the device that user is using, it is becoming more important than ever for a successful recommender system to provide a context-aware recommendation. The most important disadvantage of taking context into recommendation model is to be able to deal with a larger dataset that contains much more missing values in comparison to user-item rating matrix. For the case with contexts, some collaborative methods like matrix factorization methods[11], which have been successfully applied to a recommendation system without context consideration, could not be simply extended into a tensor of higher order[12].

Such a context-aware recommendation is more appropriately handled by factorization machine [9,13,14]. Factorization machine formulates the predictor of n variables as a target function of the second degree polynomial of n input variables. A factorization machine is a general-purpose supervised learning algorithm that one can use for both classification and regression tasks. It is an extension of a linear model that is designed to capture interactions between variables within high dimensional sparse datasets economically. For sparse datasets, a second order FM model suffices, since there is not enough information to predict more complex interactions. FM solves the problem of considering all interaction of a pair of input variables with the target. by formulating the pairwise interaction coefficient as an inner product of latent feature vectors associated with each input variable, which reduces the polynomial computation time to linear complexity. Factorization machines are easily applicable to a wide variety of context by specifying only the input data[9] and are successfully applied to context-aware recommendation problems[13]. [13] develops a fast iterative optimization method for learning factorization machine, called 'alternating least square optimization' that analytically finds the least-square solution for one parameter given the other ones.

In [10], field-aware FM (named as FFM) is proposed, where the interaction of a pair of input variables with the target in FM was refined to consider the field which each input variable belongs to. FFM has turned out to performs better than FM for many recommendation problems[15].

Recommendation on online advertising targeting individual user has proven successful. However, direct application of it over DOOH networks has its limitation due to privacy concerns. There has been intensive research both in academia and industry in the field of DOOH attempting to improve ad distribution without targeting individual user.

Our approach in this paper proposed to model context-aware scheduling of ad contents for a group of DOOH media and to analyze the effects of each context on GRP is based on factorization machine model. To the best of our knowledge, formulating scheduling of ad contents over a group of DOOH media as context-aware factorization machine-based recommendation and analysis of effects of contexts on DOOH advertising based on the formulation has not been reported yet,

3. FACTORIZATION MACHINE-BASED CONTEXT-AWARE AD CONTENTS SCHEDULING

3.1 The target DOOH Advertisement System

Fig. 1 shows an overview of our target DOOH advertisement system. Working environments of the target DOOH advertisement system is composed of many components; inventories, advertiser, ad contents, advertisement management system including context-aware factorization machine-based recommendation system module.

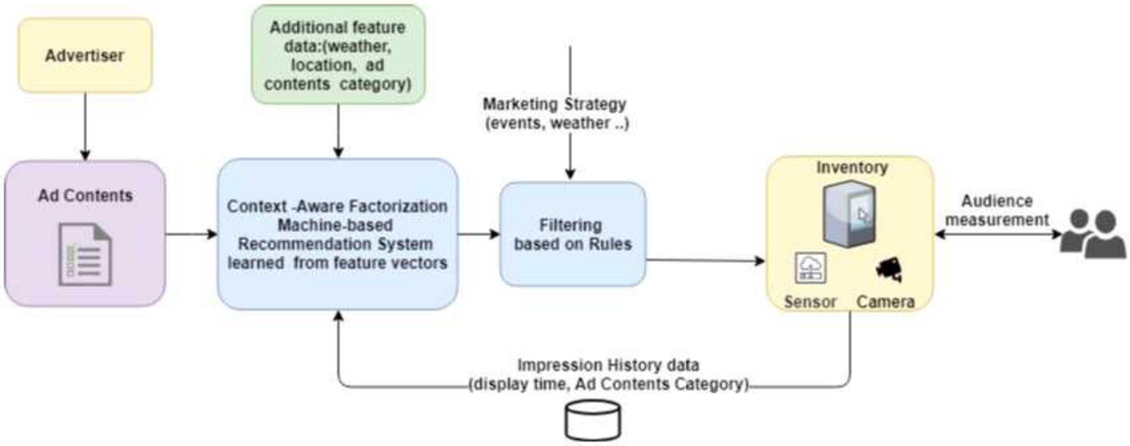

Fig. 1. Target DOOH Advertisement System.

The recommendation system predicts GRPs for candidate ad contents based on given data for each DOOH media, and ranks them. After marketing strategy and consideration of urgent events are applied and some of candidate ad contents are filtered out. Then, the highest ranking ad content among the remaining candidate ad contents at the scheduling is chosen and scheduled for display on the inventory. GRP is affected by contexts. For more precise GRP prediction, one may need to include proper contexts in a prediction model.

3.2 Contexts of DOOH Advertising

Advertising of ad contents, evaluated by how many people watch them, is affected by advertising contexts such as weather, location, content category, and so on. By defining relevant and meaningful context model for the ad content recommendation system, one can express audience target more compactly, and this is expected to improve the prediction accuracy of the ad content recommendation.

The target DOOH advertisement system manages system-related information such as inventory id, inventory type, inventory location, ad content id, ad content category, and it further collects impression history log from inventories, which contains a list of record with inventory id, ad content id, timestamp, and audience impression measurement. Timestamp contains an information about when and on which day the ad content starts to display, and audience impression measurement has the information about how many people watch the ad contents, which is achieved by analyzing images or data captured by cameras or sensors. Audience measurement can have more details like age and gender depending on the audience measurement processing. In addition to impression history log, the system collects more contextual information such as weather, event types (sports game event, entertainment or art performance events, and other city events), and event times from outside.

Some contextual information like inventory location, weather, time will affect GRP; for example, weather affects people's going out and watching outdoor DOOH media. Thus, it is import to include contexts in the prediction model.

3.3 DOOH Ad Content Recommendation with Contexts

By extending a typical formulation of recommendation system such as (1), GRP prediction context over DOOH medias can be basically formulated as a regression task as follows.

For given a set of DOOH medias $K=\{k_{1,} \cdots, k_p\}$, a set of candidate ad contents $A=\{a_{1,} \cdots, a_q\}$, and a set of contexts $C_3, C_4, \cdots, C_k$, target function $y: K \times A \times C_3 \times \cdots \times C_k \rightarrow R$ has to be predictd.

The target function represents the GRP, e.g. $y(k, a, c_3, \cdots, c_k)$ is the GRP of ad content c for DOOH media k under contexts $c_3, c_4, \cdots, c_k$.

Any data $(k, a, c_3, \cdots, c_4)$ $(\in K \times A \times C_3 \times \cdots \times C_k)$ is called context-aware DOOH data.

One can consider ad content category, displaying time in a day, displaying day in a week, weather, location, and so on,

3.4 DOOH Ad Contents Recommendation based on Factorization Machine

The context-aware DOOH data $(k, a, c_3, c_4, \cdots, c_k)$ can be transformed into a feature vector $\boldsymbol{x} = (x_1,\ x_2,\ \cdots,\ x_n) (\in X \subset R^n)$ by using binary or real valued mappings $z: C \rightarrow R^{n_z}$ as follows.

$Z: K \times A \times C_1 \times \cdots \times C_K \rightarrow X$ defined by

$$Z(k, a, c_1, \cdots, c_k) = (z_1(k),\ z_2(a),\ z_3(c_3), \cdots, z_k(c_k)) \quad (9)$$

Then, target function $\hat{y}(k, a, c_3, \cdots, c_k)$ of DOOH ad contents recommendation can be formulated as a factorization machine as follows.

$$\hat{y}(k, a, c_3, \cdots, c_k) = \hat{y}(Z(k, a, c_3, \cdots, c_k)) = \hat{y}(\boldsymbol{x}) = w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n} \sum_{j=i+1}^{n} \widehat{w_{i,j}} x_i x_j \quad (10)$$

As explained in Section 2.1, for general factorization machine (FM), $\widehat{w_{i,j}} := <\boldsymbol{v_i}\ ,\ \boldsymbol{v_j}>$ as in (5) and for field-aware factorization (FFM), $\widehat{w_{i,j}} := <\boldsymbol{v_{i,f_j}}\ ,\ \boldsymbol{v_{j,f_i}}>$ as in (8).

Fig. 2 shows an example of feature vectors for DOOH advertising made by utilizing the mappings z_i.

For example, about weather data transformation, one can take the function z as $z(sunny) = (1,0,0,0)$, $z(rainy) = (0,1,0,0)$. And about displaying time in a day, one can take a function which quantizes the display time in a minute unit and normalizes the quantized integer number to the range of 0 and 1 with two decimal places of accuracy. For example, 10:32:40 (10 o'clock 32 minutes 40 seconds) is quantized into 10×60+33 and normalized $\frac{10 \times 60 + 33}{24 \times 60} \simeq 0.44$. Location data can be real valued using such as zip code of the location.

3.5 Training Factorization machine based DOOH Ad Contents Recommendation

Stochastic gradient descent (SGD) algorithms are popularly adopted for optimizing loss function (6) in training of factorization machine model of DOOH ad contents recommendation (9) since SGD algorithms are simple, work well with different loss functions, and have low computational and storage complexity. Algorithm 1 shows how FM can be optimized with SGD.

The Algorithm 1 iterates over cases $(\boldsymbol{x}, y) \in S$ and performs updates on the model parameters.

$$\theta \leftarrow \theta - \eta\left(\frac{\partial}{\partial \theta} l(\hat{y}(\boldsymbol{x}),\ y) + 2\lambda_\theta \theta\right) \quad (11)$$

where $\eta \in R^+$ is the learning rate or step size for gradient descent.

SGD-based FFM learning algorithm utilized in this paper is similar to Algorithm1, which is SGD-

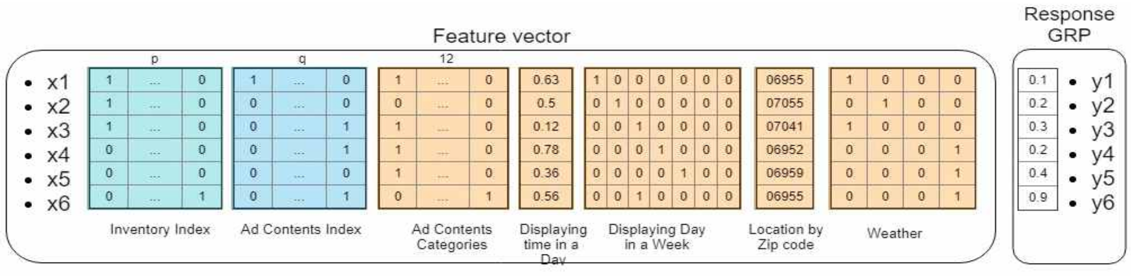

Fig. 2. An example of feature vectors for DOOH advertising.

Algorithm 1 : Stochastic Gradient Descent (SGD)-based FM Training [14]

Input: Training data S , regularization parameters λ, learning rate η, initialization σ
Output: Model parameter $\Theta = (w_0, \boldsymbol{w}, \boldsymbol{V})$
$w_0 \leftarrow 0$; $\boldsymbol{w}=(0,\cdots,0)$; $\boldsymbol{V} \sim N(0, \sigma)$;
repeat
for $(\boldsymbol{x},y)\in S(\mathrm{x}, \mathrm{y})$ do
$w_0 \leftarrow w_0 - \eta(\frac{\partial}{\partial w_0} l(\hat{y}(\boldsymbol{x}|\Theta), y) + 2\lambda^0 w_0)$;
for $i\in\{1,\cdots,\mathrm{n}\} \wedge (\mathrm{x_i} \neq 0)$ do
$w_i \leftarrow w_i - \eta(\frac{\partial}{\partial w_i} l(\hat{y}(\boldsymbol{x}|\Theta), y) + 2\lambda^w_{\pi(i)} w_i)$;
for $f\in\{1,\cdots,\mathrm{k}\}$ do
$v_{i,f} \leftarrow v_{i,f} - \eta(\frac{\partial}{\partial v_{i,f}} l(\hat{y}(\boldsymbol{x}|\Theta), y) + 2\lambda^v_{f,\pi(i)} v_{i,f})$;
end
end
end
until stopping criterion is met;

based FM learning algorithm except that the updating the latent vector changes from

$$v_{i,f} \leftarrow v_{i,f} - \eta(\frac{\partial}{\partial v_{i,f}} l(\hat{y}(\boldsymbol{x}|\Theta), y) + 2\lambda^v_{f,\pi(i)} v_{i,f}) \text{ to}$$

$$v_{i,f_j,k} \leftarrow v_{i,f_j,k} - \eta(\frac{\partial}{\partial v_{i,f_j,k}} l(\hat{y}(\boldsymbol{x}|\Theta), y) + 2\lambda^{v,f}_{k,\pi(i)} v_{i,f_j,k})$$

where $v_{i,f_j,k}$ is the $k-th$ component of $\boldsymbol{v_{i,f_j}}$.

For general FM model, [13] proposes a fast optimization algorithm, 'Alternating Least-Squares (ALS)' . For the details, the reader needs to refer to [13].

4. EVALUATION

4.1 Evaluation Environments

Since we neither have any real data set for DOOH ad contents scheduling nor find any dataset for DOOH from Internet, we decide to utilize an existing real online recommendation dataset, 'youchoose-data' from RecSys Challenge 2015[7]. Given a sequence of click events performed by some users during a typical session in an e-commerce website, the goal of RecSys Challenge 2015 is to predict whether the user is going to buy something or not, and if he or she is buying, what would be the items he or she is going to buy.

RecSys Challenge 2015 dataset comprises two kinds of datasets; 1) Click event (which contains information about session ID, item ID, timestamp, category), 2) Buys event (session id, item id, timestamp, price, quantity). Session ID is the id of the session. In one session there are one or many clicks or buying events. Timestamp is the time when the click event or buys event occurred. Category means the category of the item (food & drink, clothing & fashion, travel, hospitality, leisure & entertainment, health & beauty, pharmaceuticals, automotive, energy, insurance, electronics). Quantity represents how many of this item were bought.

We adapt these dataset and build a dataset for our evaluation. Our dataset consists of the following fields; DOOH inventory ID, ad contents ID, ad contents category, displaying time in a day, and displaying day in a week, which are analogously taken from session ID, item ID, item category, timestamp of click events, respectively. For GRP, we obtain it by calculating quantity over an item. Since there are no fields corresponding to weather and location information in RecSys 2015 dataset, we ignore them in this evaluation. The built dataset has 48,505 data records of 1000 inventories and 6029 ad contents with other context data. After we convert all data records into feature vectors to suit the factorization model (10), we split the dataset to two parts for training and testing, respectively. 20% samples from the dataset were randomly selected as the test set.

For evaluation, we test FFM model with SGD optimization by using libFFM [16] as well as FM model with ALS optimization [13] by using libFM [14].

4.2 Simulation Analysis Results

4.2.1 The Effects of Contexts on Advertising (GRP)

In order to analyze the effects of contexts on advertising (GRP), we first trained 5 FFM models of ad contents recommendation; ‘Context’, ‘Hour’, ‘Weekday’, ‘Category’, and ‘MF’.

‘Context’ is a full context-aware FFM model considering all contexts(displaying time in a day, displaying days in a week, ad contents category in addition to the basic input variables, inventory ID and contents ID), ‘Hour’, ‘Weekday’, ‘Category’ are FFM models with one context considered (consider displaying time in a day only, displaying day in a week, and ad contents category, respectively) in addition to the basic input variables. ‘MF’ is a FFM model without considering any contexts.

Then, we simulated the test set with respect to those trained FM models, respectively and compare RMSEs of GRP prediction over the trained FFM models. Fig. 3 shows the simulation results. In FM tasks, one epoch represents one full training cycle on the training set.

As shown in Fig. 3, the full context-aware recommendation FFM model (‘Context’) predicts better than the recommendation FFM model without considering contexts. Against the evaluation dataset, ad content category turns out to be a much more effective context compared to temporal contexts (displaying time in a day, displaying day in a week). These temporal contexts and GRP of the evaluation dataset are taken from timestamps and quantity of online RecSys 2015 dataset. How many items are bought on an online web market place may not have much interaction with click times, at least on the RecSys 2015 dataset. That is why the simulation result of Fig. 3. shows much lower effects of temporal contexts. In Fig. 4 of the simulation results based on context-aware FM models against the same training and testing data set as those of Fig. 3, it was observed that the context-aware FM models with ALS optimization perform worse than FFM models.

Another simulation analysis to see the effectiveness of a pairwise combination of contexts on the prediction of FFM models against the evaluation dataset was also conducted. In Fig. 5 of the simulation results, one can understand that ‘Category’ context combined together with ‘Weekday’ context shows the best interaction with prediction among all context combinations.

From simulation analysis results as in Fig. 3 and Fig. 5, one can understand that the context-aware FFM model is effective to analyze the effects of advertising contexts for scheduling of ad contents under contexts.

4.2.2 The Effects of Hyperparameters of FM(FFM) Modeling

We also conducted simulation to investigate the effects of hyperparameters of FM modeling such

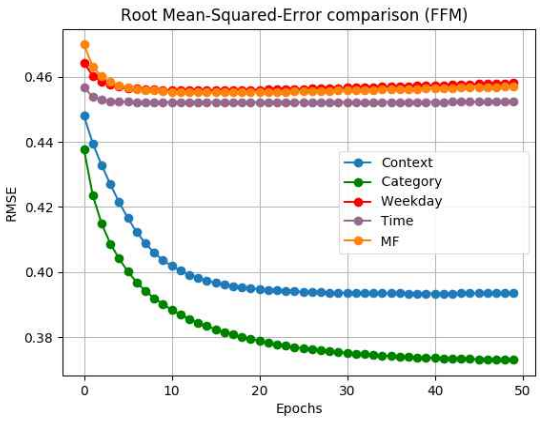

Fig. 3. Comparisons of the effects of contexts based on FFM models.

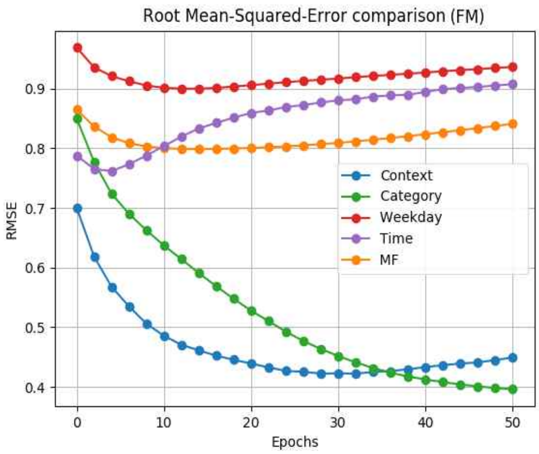

Fig. 4. The simulation results based on context-aware FM models.

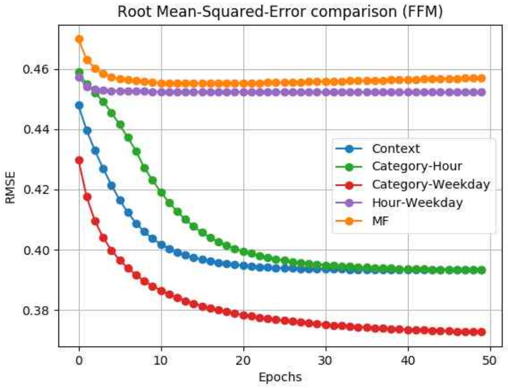

Fig. 5. Comparisons of the effects of pairwise combination of contexts based on FFM models.

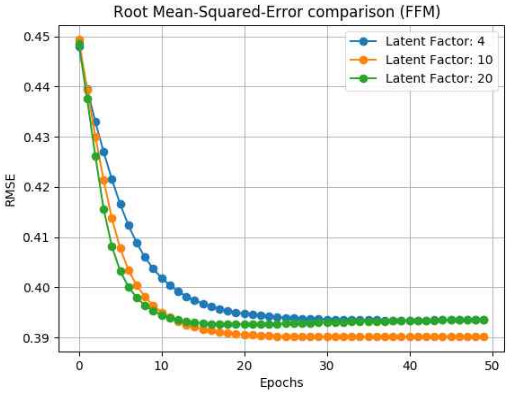

Fig. 6. Effects of the latent facor k.

as latent factor k in (6), regularization parameter λ in (7), and learning rate η in (11). In FM(FFM) modeling, the latent factor k, which is the dimension of the latent vector $\boldsymbol{v_i}$ $(i=1,\cdots,n)$ is one of important hyperparameters, which affects the prediction performance, and can be decided from simulation. Simulation result in Fig. 6 shows that larger latent factor improves the performance before threshold, but the more larger latent factor after threshold causes overfitting.

Through evaluations, it is observed that regularization parameter values do not make any big difference with respect to the performance as long as they are some small positive values, which is shown in Fig. 7(a) but learning rate parameter η in training should be carefully determined, which is shown in Fig. 7(b).

5. CONCLUSION AND FUTURE WORK

In this paper, we presented a context-aware factorization machine-based recommendation model for scheduling of ad contents under advertising contexts over a group of DOOH medias. Then, it was shown through simulations based on the model over a realistic dataset that one could analyze effectively effects of contexts such as categories of ad contents, weather, displaying time, location and other contexts on the ad contents advertising over DOOH medias.

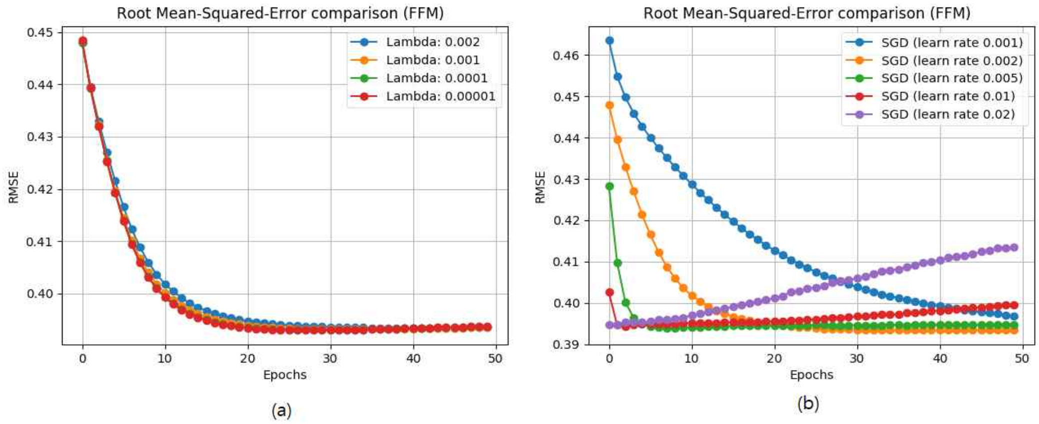

Fig. 7. Effects of the regularization parameter λ and learning rate η ; (a) regulation parameter λ , (b) learning rate η.

The analysis was achieved as follows. Firstly, one trains the context-aware factorization machine model for DOOH ad contents scheduling. Then, one predicts GRPs(Gross Rating Points) against test sets, calculates prediction error, and analyzes the factors affecting the prediction errors. Thus, theses simulation analyses are considered as useful guidances for efficient scheduling of ad contents under advertising contexts over a group of DOOH medias. The dataset utilized for the evaluation in this paper, is made from adapting a real dataset given by RecSys Challenge 2015 for our DOOH advertisement environments.

For measuring real effectiveness of the proposed model, a context-aware factorization machine-based recommendation for scheduling of ad contents needs to be implemented on a real DOOH advertising system like Fig. 1. Then, over the real advertising context data including weather, events, and more accumulated through operations over a group of DOOH medias, we can test the effectiveness, which is worthwhile to report for the contribution to development of well-performing programmatic DOOH systems.

REFERENCE

[1] What Is Digital Out-of-Home (DOOH) Advertising and How Does It Work?, https://clearcode.cc/blog/what-is-digital-out-of-home-dooh (accessed Feb., 08, 2019).

[2] F. Jiang, K. Thilakarathna, M. Hassan, Y. Ji, and A. Seneviratne, "Efficient Content Distribution in DOOH Advertising Networks Exploiting Urban Geo-Social Connectivity," *Proceedings of International World Wide Web Conference Committee*, pp. 1363-1370, 2017.

[3] DOOH Market to Grow by 12.6% from 2017 to 2023, https://www.digitalsignagetoday. com/news/dooh-market-to-grow-by-1260-percent-from-2017-to-2023 (accessed Feb., 08, 2019).

[4] J. Wang, W. Zhang, and S. Yuan, *Display Advertising with Real-Time Bidding (RTB) and Behavioural Targeting*, Now Publishers Inc., U.S.A. , 2017.

[5] G. Linden, B. Smith, and J. York, "Amazon.com Recommendations Item-to-Item Collaborative Filtering," *Journal of IEEE Internet Computing*, Vol. 7, No. 1, pp. 76-80, 2003.

[6] Gross Rating Point, https://en.wikipedia.org/wiki/Gross_rating_point (accessed Feb., 08, 2019).

[7] Recsys Challenge 2015, https://www.kaggle.com/chadgostopp/recsys-challenge-2015 (accessed Feb., 08, 2019).

[8] J.I. Kim, "An Implementation of Automatic Upper-Lower Clothes Matching System Using Machine Learning," *Journal of Korea Multimedia Society*, Vol. 13, No. 3, pp. 467-474, 2010.

[9] S. Rendle Steffen Rendle, "Factorization Machines," *Proceeding of the 10th IEEE International Conference on Data Mining, IEEE Computer Society*, pp. 995-1000, 2010.

[10] Y. Juan, Y. Zhuang, and W. Chin, "Field-aware Factorization Machines for CTR Prediction," *Proceedings of the 10th Association for Computing Machinery Conference on Recommender Systems,* pp. 43-50, 2016.

[11] Y. Koren, R. Bell, and C. Volinsky, "Matrix Factorization Techniques for Recommender Systems," *IEEE Computer*, Vol. 42, Issue 8, pp. 30-37, 2009.

[12] Y. Fang and Y. Guo, "A Context-aware Matrix Factorization Recommender Algorithm," *Proceedings of IEEE 4th International Conference on Software Engineering and Service Science*, pp. 914-918, 2013.

[13] S. Rendle, Z. Gantner, C. Freudenthaler, and L.S. Thieme, "Fast Context-aware Recommendations with Factorization Machines," *Proceeding of the 34th International ACM SIGIR Conference on Research and Develop-*

ment in Information Retrieval, pp. 635-644, 2011.

[14] S. Rendle, "Factorization Machines with LibFM," *Association for Computing Machinery Transactions on Intelligent System and Technology,* Vol. 3, No. 3, Article No. 57, pp. 1-22, 2012.

[15] Y. Juan, D. Lefortier, and O. Chapelle, "Field-aware Factorization Machines in a Real-world Online Advertising System," *Proceedings of the 26th International Conference on World Wide Web Companion*, pp. 680-688, 2017.

[16] LIBFFM: A Library for Field-aware Factorization Machines, https://www.csie.ntu. edu. tw/~cjlin/libffm (accessed Feb., 08, 2019).

Van Hoang Nguyen

He received B. Eng. degree in Control and Automation from Hanoi University of Science and Technology, Ha Noi, Viet Nam, in 2013. He is studying for a Master degree in Information and Telecommunication Engineering at the University of Soongsil, Seoul, Korea. His research covers web development, digital signage systems, big data, and machine learning.

Thanh Binh Nguyen

He received B. Eng. Degree in computer science from the University of Science, Ho Chi Minh, Vietnam, in 2005, and M. Sc. degree and Ph.D. degree in Information and Telecommunication Engineering from the University of SoongSil, Seoul, South Korea, in 2010, and 2017, respectively. He is currently a principal software R&D researcher at CublickDigital, Inc., Seoul, South Korea. South Korea. His research interests cover the design and analysis of various smart embedded software system, IoT and also intelligent imaging systems, video analytic algorithms which are applied to visual surveillance, recognition systems, and etc.

Sun-Tae Chung

He received B.E. degree from Seoul National University, and M.S. degree and Ph.D. degree in Electrical Eng. and Computer Science from the University of Michigan, Ann Arbor, USA, in 1986 and 1990, respectively. Since 1991, he had been with the School of Electronic Eng. at the Soongsil university, Seoul, Korea where he is now a full professor. Now, he has been with the Dept. of Smart Systems Software, at the Soongsil Univ. since 2015. His research interests include: computer vision, visual surveillance, digital signage systems, and digital marketing.