Unfolding Temporal Dynamics: Predicting Social Media Popularity Using Multi-Scale Temporal Decomposition

Bo Wu[1,2], **Tao Mei**[3], **Wen-Huang Cheng**[4], **Yongdong Zhang**[1]
[1]Key Laboratory of Intelligent Information Processing,
Institute of Computing Technology, Chinese Academy of Sciences, China
[2]University of Chinese Academy of Sciences, China
[3]Microsoft Research, China
[4]Research Center for Information Technology Innovation, Academia Sinica, Taiwan
{wubo, zhyd}@ict.ac.cn tmei@microsoft.com whcheng@citi.sinica.edu.tw

Abstract

Time information plays a crucial role on social media popularity. Existing research on popularity prediction, effective though, ignores temporal information which is highly related to user-item associations and thus often results in limited success. An essential way is to consider all these factors (user, item, and time), which capture the dynamic nature of photo popularity. In this paper, we present a novel approach to factorize the popularity into user-item context and time-sensitive context for exploring the mechanism of dynamic popularity. The user-item context provides a holistic view of popularity, while the time-sensitive context captures the temporal dynamics nature of popularity. Accordingly, we develop two kinds of time-sensitive features, including user activeness variability and photo prevalence variability. To predict photo popularity, we propose a novel framework named Multi-scale Temporal Decomposition (MTD), which decomposes the popularity matrix in latent spaces based on contextual associations. Specifically, the proposed MTD models time-sensitive context on different time scales, which is beneficial to automatically learn temporal patterns. Based on the experiments conducted on a real-world dataset with 1.29M photos from Flickr, our proposed MTD can achieve the prediction accuracy of 79.8% and outperform the best three state-of-the-art methods with a relative improvement of 9.6% on average.

Introduction

In recent years, popularity prediction on social media has attracted extensive attention because of its widespread applications, such as online marketing, trend detection and resource allocation (Tatar et al. 2014). Generally, given historical user-item pairs, popularity prediction is defined as the problem of estimating the rating scores, view counts or click-through of a newly post in social media (Pinto, Almeida, and Goncalves 2013; McParlane, Moshfeghi, and Jose 2014). Existing research on popularity prediction predominantly focuses on exploring the correlation between popularity and user-item factors, such as item content (Hong, Dan, and Davison 2011; Cappallo, Mensink, and Snoek 2015), user cues (Khosla, Sarma, and Hamid 2014), social relation (Nwana, Avestimehr, and Chen 2013), and user-item interaction (Niu et al. 2012). In fact, time also exerts crucial impact on the popularity but is often overlooked. According to the 2015 survey (Patel 2015), it makes a big difference on when content should be shared in order to receive the most social traffic. For example, best times to pin on Pinterest are Saturday, 2am-4am and 8pm-11pm; Facebook has 86% of its posts published during the workdays with engagement peaking on Thursday and Friday. These phenomena show the importance using such time-sensitive context for popularity prediction. In other words, considering user-item factors only is insufficient for capturing popularity dynamics over time. As a result, our problem can thus be redefined as follows: how to predict the popularity s of a post v from a user u with sharing time t.

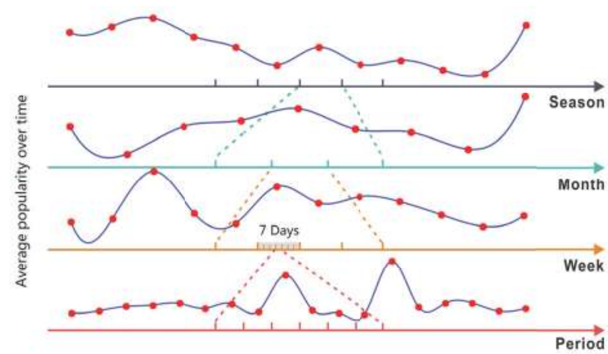

Figure 1: Popularity is observed to exhibit varying trends of changes at different time scales.

In this paper, we propose to investigate the popularity prediction by factoring popularity into two contextual associations, i.e., user-item context and time-sensitive context. The user-item context is linked to popularity with user-specific and item-specific contextual information, which can be derived from user-item sharing behaviors on social media (Niu et al. 2012; He et al. 2014). The time-sensitive context is affected by 'change over time' information (associated with sharing time of photos) including user activeness variability and photo prevalence variability. However, modeling the impact of temporal information for popularity is a non-trivial task. To conclude this, popularity naturally changes over time and the trends of changes at different time granularities are often observed to be distinct, e.g., Figure 1.

In responding to the above challenges, we propose a novel framework based on matrix factorization called Multi-scale Temporal Decomposition (MTD), whereby the time-sensitive popularity can be estimated through a joint latent space by the user-item sub-matrix and a series of time-sensitive sub-matrices at multiple time scales. Accordingly, to obtain time-sensitive features at multiple time scales, we define four temporal granularities from period, week, month, to season. Finally, we evaluated our method on a dataset containing 1.29 million photos with 55,000 users collected from Flickr and demonstrate the outperformance of our approach over the state-of-the-art methods.

Our contributions in this paper include:

- To our best knowledge, our study is the first work to explore social media popularity by two complementary perspectives: user-item context and time-sensitive context.
- We designated the temporal context of popularity variations with two 'change over time' features in our method: user activeness variability and photo prevalence variability. Also, we demonstrate the effectiveness of the proposed temporal features by statistical analysis.
- We developed a new framework named Multi-scale Temporal Decomposition (MTD) for popularity prediction, enabling popularity analysis in multiple time granularities.

The rest of this paper is organized as follows. We first review related work in Section 2. Section 3 presents two contextual perspectives for popularity: user-item context and time-sensitive context. Then we propose MTD method in Section 4. Section 5 provides evaluations, followed by conclusions and future work in Section 6.

Related Work

Popularity prediction problem has received a wide range of attentions on varied web contents. Most of these works exploited popularity by investigating into user, item or user-item behaviors. Hong *et al.* and Cappallo *et al.* predicted the popularity by utilizing properties of large-scale user generated content (Hong, Dan, and Davison 2011; Cappallo, Mensink, and Snoek 2015). Khosla *et al.* and Nwana *et al.* provided evidences that user cues or social contexts have contributions to popularity in social media (Khosla, Sarma, and Hamid 2014; Nwana, Avestimehr, and Chen 2013). Shamma *et al.* and Niu *et al.* presented that modeling with user-item interactions can explore implicit patterns of popularity (Shamma et al. 2011; Niu et al. 2012). Existing researches of popularity prediction pay attention on user and item aspects and lack time-sensitive factors for capturing the dynamic evolution of popularity.

Few existing works on dynamic popularity can be grouped into two main paradigms, each with known strengths and limitations. One paradigm studies the dynamic popularity based on statistical quantities (Ratkiewicz et al. 2010; Yang and Leskovec 2011; Figueiredo, Benevenuto, and Almeida 2011). These works found that a popularity change is related to time series features so as to measure dynamic trends of general popularity. However, as they do not provide a way to extract item-specific parameters (Shen et al. 2014), these models are lacking of predictive power for the popularity dynamics of specific user-item pairs. The other paradigm extracts dynamical features or function terms by using temporal information to estimate popularity on social media (Lerman and Hogg 2010; Ahmed et al. 2013; Shen et al. 2014; McParlane, Moshfeghi, and Jose 2014). Although these models succeed in using temporal features or variables, these approaches treated temporal information as time-series data or individual variables without scale linking. Therefore, how to utilize time-sensitive factors for predicting photo popularity is still an open research issue.

Popularity with Contextual Information

In this section, we first define photo popularity and demonstrate its contextual environment (including user-item context and time-sensitive context) on popularity with different contextual factors. Then we investigate the temporal dynamics of popularity with the time-sensitive context based on two 'change over time' features: user activeness variability and photo prevalence variability. Finally, we reveal the impacts of these features on photo popularity.

Photo Popularity

Popularity on social media is resulted from various selection behaviors, such as "view", "click" and "comment". As for Flickr, the largest photo-sharing site, users are allowed to view details of a photos content and its related information by clicking the thumbnails from public newsfeeds or image search engines. Thus the popularity of a photo is relied by a sharing behavior p (can be defined by a triple array$\langle u, v, t\rangle$ that denotes a user u shares a new photo v at time t).

To validate our proposed approach on real-world data, we collected 1.29 million photos from 55,000 users (Please see the experiments). To alleviate the large variation (i.e. the number of views of different photos varies largely from zero to millions), we apply the log-normalization approach (Khosla, Sarma, and Hamid 2014) on the popularity formulation. As a result, the log-normalized popularity of a photo can be defined as

$$s = \log_2 \frac{r}{d} + 1 \tag{1}$$

where r is the original view count of each photo, and d is the number of days since the photo was shared.

Contextual Information of Popularity

Popularity is related to both of online interactions (who are interested in what) and offline patterns which contain rich

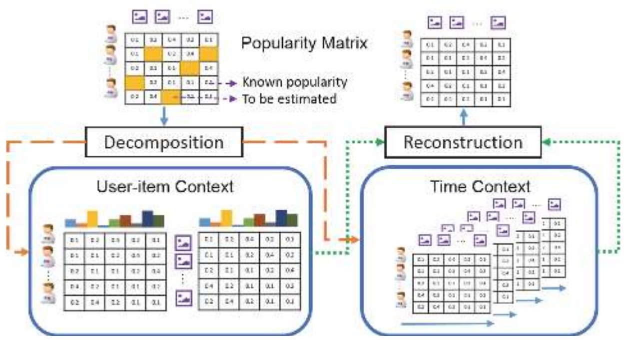

Figure 2: An overview of our proposed Multi-scale Temporal Decomposition (MTD).

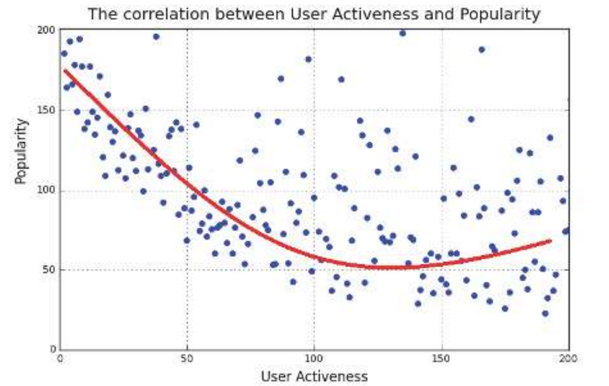

Figure 3: Popularity versus the user activeness: The example scatter graph derived from our dataset shows a negative correlation.

meanings of activities in real-life, such as work, leisure and travel. Therefore, recognizing impacts of these contextual information is important for understanding popularity, and it can be factorized into user-item context and time-sensitive context, as shown in Figure 2. User-item context is the environment with users, photos, and links of user-item interactions. Time-sensitive context is described as the temporal environment which affects popularity patterns with sharing time on different time scales.

User-Item Context User-item interaction is the fundamental momentum of popularity generation on social networks (Niu et al. 2012; Jiang et al. 2014).

Definition 1. **(User-item Context)** Given a user u and an item v, F_u and F_v are two feature representations of similar users and photos, respectively. User-item context is a distribution embedding F_{uv} with factors of F_u and F_v which is connected by user-item sharing behaviors.

Personal information reflects the user-specific reputation on social media, and we select user features as follows: mean views, photo counts, the number of contacts, average group members of each user, and having Pro Flickr account or not. To quantity a photo, we extract both visual and perceptive features. Visual information refers to the visual quantity of photos and we extract color patch features (Khan et al. 2013), Local Binary Pattern (LBP) features (Dalal and Triggs 2005) and locality-constrained gradient features (Dalal and Triggs 2005; Khosla, Sarma, and Hamid 2014). Perceptive features describe the user's response to a photo, containing the number of image tags, length of image title and length of image descriptions. Meanwhile, classification features by convolutional neural networks (CNNs) with the DeCAF method (Donahue et al. 2013) are adopted. As a result, we concatenate the extracted features above to a composite one, resulting in a 20,294 dimensional feature.

Time-sensitive Context To measure time-sensitive factors of popularity, we developed two temporal features (user activeness variability and photo prevalence variability) for modeling its temporal dynamics.

Definition 2. **(Time-sensitive Context)** At time t, without loss of generality, the temporal context is constituted by the user activeness variability vector A_u and the photo prevalence variability vector A_v, as detailed below. Time-sensitive context is thus a distribution embedding of A_{uv} consisting of the temporal factors A_u and A_v.

User Activeness Variability. The user activeness can be defined by the number of online behaviors (i.e. sharing records, comment interactions, view behaviors), denoted as c_t, at a time interval t_{span}. We showed the correlation between popularity and the user activeness at the same time scale of week in Figure 3. It demonstrates that user activeness is a time-aware and correlative factor to reflect the changing patterns of popularity dynamics. Therefore the user activeness variability in social media is a signal to indicate popularity variations related with the human activities of users in the real world. For example, 5pm is the time of working day for highest retweets and this is observed because people look for something to keep them occupied during commuting time after work (Patel 2015). To conclude that, user activeness variability should be considered in popularity prediction and can be defined as

$$f_{ua}{=}\delta(\frac{a_t - \bar{a}}{\bar{a}}) \tag{2}$$

where a_t is the active frequency of a user $\frac{c_t}{t_{span}}$ in a time interval, and $\bar{a}$ is the mean of a user's activeness based on the Direct delta function $\delta(\cdot)$.

Photo Prevalence Variability. Similar photos tend to have popularity consistency on social media (Khosla, Sarma, and Hamid 2014; Cappallo, Mensink, and Snoek 2015). Some of the recent works attempted to infer photo popularity by this property. However the popularity consistency of similar photos is a dynamic factor, because the users' preferences on varied categories are changing over time. Then, predicting photo popularity based on the photos with similar contents or visual information only will ignore the temporal variations of popularity for prediction. To measure this influence, we describe the prevalence variability of a photo to be defined as a ratio of its popularity with respect to the mean value of the popularity of the other photos with

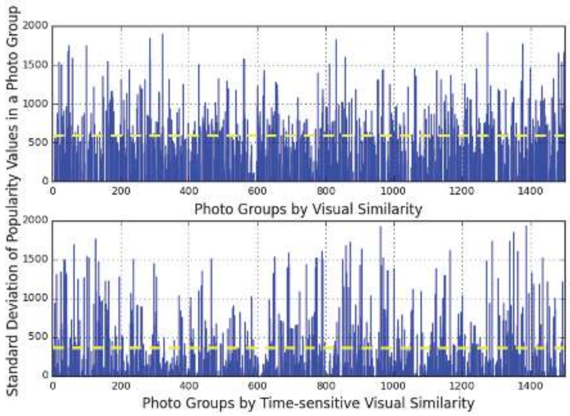

Figure 4: Histograms of standard deviations of the mean popularity (average views) for photo groups clustered by the visual similarity (top sub-figure) and by the proposed time-sensitive visual similarity (bottom sub-figure).

similar contents and in the corresponding time:

$$f_{va} = S(s_v \frac{K}{\sum_{t=1}^{K} \bar{s}_t}) \tag{3}$$

where s_v is the popularity of a photo v and $\bar{s}_t$ is the average popularity of similar photos measured at a time scale t. $S(\cdot)$ is Sigmoid function.

To enable the computation of the photo prevalence variability at different time scales, we grouped photos based on not only the visual content but also if two photos were shared in the same time interval at a time scale. By experimenting with the UsD dataset (see the experiments), Figure 4 illustrates two histograms of standard deviations of the mean popularity for photo groups clustered by the visual similarity only and by the combination of visual similarity and the condition of same time intervals, respectively. We can find that the average of standard deviations over the group by the time-sensitive visual similarity is relatively lower, implying a higher consistency in the popularity values of clustered photos in each group.

Popularity Prediction by Context Decomposition

Popularity Matrix

Given M users $\{u_i\}_{i=1}^{M}$ and N shared photos $\{v_j\}_{j=1}^{N}$, the popularity matrix can be expressed as $\mathbf{R}$, where each element R_{ij} is the log-normalized popularity score s of photo v_j shared by user u_i. The prediction task is to use the observed parts of the popularity matrix $\mathbf{R}$ and historical information to estimate real values of unknown popularities. Before the prediction phase, we first factorize the popularity matrix using two contextual information: user-item context and time-sensitive context.

Decomposition Based on User-item Context Only

Inspired by the previous work (Salakhutdinov and Mnih 2007), we model the popularity matrix with the user-item context by matrix factorization. Assuming that there exists a latent space to describe the popularity distribution between users and items with latent factors of dimensionality l (e.g., inner momentum of popularity to cause user-item sharing behaviors and they may correspond to certain semantic or interesting information), a popularity matrix can be decomposed as the inner product of sub-matrices of users and items with common latent factors. We construct a sub-matrix $\mathbf{U} \in \mathbb{R}^{M \times l}$ as the user-latent matrix, and a sub-matrix $\mathbf{V} \in \mathbb{R}^{l \times N}$ as the latent-item matrix. Having some elements of the popularity matrix $\mathbf{R}$ with known popularity scores, our goal is to estimate the unknown popularities based on $\mathbf{U}$ and $\mathbf{V}$ by minimizing the following objective function:

$$\arg\min_{\mathbf{U},\mathbf{V}} \|\mathbf{R} - \mathbf{UV}\|_F^2 + \lambda_{\mathbf{U}} \|\mathbf{U}\|_F^2 + \lambda_{\mathbf{V}} \|\mathbf{V}\|_F^2 \tag{4}$$

where $\lambda_{\mathbf{U}}$ and $\lambda_{\mathbf{V}}$ are control parameters, and F represents Fobenius Norm. Furthermore, $\mathbf{U}$ and $\mathbf{V}$ can also be estimated by the user similarity matrix $\mathbf{S} \in \mathbb{R}^{M \times M}$ and the content similarity matrix $\mathbf{C} \in \mathbb{R}^{N \times N}$, as follows:

$$\begin{aligned} &\arg\min_{\mathbf{U},\mathbf{V}} \|\mathbf{R} - \mathbf{UV}\|_F^2 + \lambda_U \|\mathbf{U}\|_F^2 + \lambda_V \|\mathbf{V}\|_F^2 \\ &+\lambda_S \left\|\mathbf{S} - \mathbf{U^T U}\right\|_F^2 + \lambda_C \left\|\mathbf{C} - \mathbf{VV^T}\right\|_F^2 \end{aligned} \tag{5}$$

where $\lambda_{\mathbf{S}}$ and $\lambda_{\mathbf{C}}$ are control parameters.

Multi-scale Temporal Decomposition

We further incorporate the temporal decomposition for time-sensitive context into the above formulation, i.e. Equation (5). As aforementioned, time-sensitive context is affected by time-sensitive factors, A_u (user activeness variability) and A_v (photo prevalence variability). Accordingly, a set of pairs of time-sensitive matrices including the user-time matrix $\{\mathbf{W}_{T_d}\}_{d=1}^{D}$ and the time-item matrix $\{\mathbf{H}_{T_d}\}_{d=1}^{D}$ are the feature distribution matrices ($W_{T_d} = rows\_concat\{A_u \ over \ all \ users\}$, $H_{T_d} = cols\_concat\{A_v \ over \ all \ photos\}$) for modeling the time-sensitive context in various time scales, which consist of A_u and A_v over all users and photos, respectively. $\{1, \cdots, D\}$ is the index set of our predefined time scales {period, week, month, season}. With regard to periods of day, we divide 24 hours of a day into six periods, i.e. "morning (8:00am-12:00am)", "lunch time (12:00am-14:00pm)", "afternoon (14:00am-17:00pm)", "dinner time (17:00am-20:00pm)", "evening (20:00am-24:00pm)" and "sleeping (0:00am-8:00am)". Thus, given the popularity matrix $\mathbf{R}$, our purpose is to estimate nonnegative matrices $\mathbf{U}$ and $\mathbf{V}$ by minimizing the errors between the popularity matrix and the joint sum of two groups of the sub-matrices ($\mathbf{S}$, $\mathbf{C}$, $\{\mathbf{W}_{T_d}\}_{d=1}^{D}$ and $\{\mathbf{H}_{T_d}\}_{d=1}^{D}$) for popularity prediction. The non-negative constraint is also applied in our model, and the cost function of our model is expressed as follows:

$$\begin{aligned} &\arg\min_{\mathbf{U},\mathbf{V}} \left\|\mathbf{R} - (\mathbf{UV} + \sum_{d=1}^{D} \lambda_{T_d} \mathbf{W}_{T_d} \mathbf{H}_{T_d})\right\|_F^2 + \lambda_U \|\mathbf{U}\|_F^2 \\ &+\lambda_V \|\mathbf{V}\|_F^2 + \lambda_S \left\|\mathbf{S} - \mathbf{U^T U}\right\|_F^2 + \lambda_C \left\|\mathbf{C} - \mathbf{VV^T}\right\|_F^2 \end{aligned} \tag{6}$$

where $\{\lambda_{T_d}\}_{d=1}^{D}$ are control parameters for multi-scale temporal information.

Optimization

To solve the optimization objective, we minimize the loss between prediction values of popularity and the ground truth. For solving the Equation (6), we apply the Multiple Update Rule (D.Lee and Sebastian.Seung 1999; Lee and Seung 2001) to our model. The popularity matrix $\mathbf{R}$ is estimated by the sum of $\mathbf{UV}$ and $\{\mathbf{W}_{T_d}\mathbf{H}_{T_d}\}_{d=1}^{D}$. The solution can be formulated as Algorithm 1. To solve the objective function of our model, we denote gradient functions for $\mathbf{U}$ and $\mathbf{V}$ by $\frac{\partial J}{\partial U}$ and $\frac{\partial J}{\partial V}$. Once the convergence condition is satisfied or the number of iterations is larger than a threshold, the unknown popularity values will be harvested.

Algorithm 1 The Gradient Algorithm of Multi-scale Temporal Decomposition

Require: constraint matrices: $\mathbf{R}$, $\mathbf{C}$, $\mathbf{S}$, $\{\mathbf{W}_{T_d}\}_{d=1}^{D}$ and $\{\mathbf{H}_{T_d}\}_{d=1}^{D}$; maximal interaction: n;
1: Initial $\mathbf{U}$ and $\mathbf{V}$ randomly;
2: **repeat**
3: Compute gradient directions:
$\frac{\partial J}{\partial U} = 2\left(-\mathbf{R}\mathbf{V}^{\mathrm{T}} + (\mathbf{UV})\mathbf{V}^{\mathrm{T}} - 2\lambda_{\mathbf{S}}\mathbf{SU} + 2\lambda_{\mathbf{S}}\mathbf{U}\mathbf{U}^{\mathrm{T}}\mathbf{U} + \lambda_{\mathbf{U}}\mathbf{U}\right)$
$\frac{\partial J}{\partial V} = 2\left(-\mathbf{U}^{\mathrm{T}}\mathbf{R} + \mathbf{U}^{\mathrm{T}}(\mathbf{UV}) - 2\lambda_{\mathbf{C}}\mathbf{VC} + 2\lambda_{\mathbf{C}}\mathbf{V}\mathbf{V}^{\mathrm{T}}\mathbf{V} + \lambda_{\mathbf{V}}\mathbf{V}\right)$
4: Compute results of objective function:
$\arg\min_{\mathbf{U},\mathbf{V}} \left\|\mathbf{R} - (\mathbf{UV} + \sum_{d=1}^{D}\lambda_{T_d}\mathbf{W}_{T_d}\mathbf{H}_{T_d})\right\|_F^2 + \lambda_U\|\mathbf{U}\|_F^2$
$+\lambda_V\|\mathbf{V}\|_F^2 + \lambda_S\left\|\mathbf{S} - \mathbf{U}^{\mathbf{T}}\mathbf{U}\right\|_F^2 + \lambda_C\left\|\mathbf{C} - \mathbf{V}\mathbf{V}^{\mathbf{T}}\right\|_F^2$
5: Update $\mathbf{U}$ and $\mathbf{V}$ followed by Multiple Update Rule
6: **until** procedure convergence or the number of interactions is over n

Experiments

Experimental Settings

User-specific Dataset (UsD): We collected 750K photos from the personal albums of 400 different users from Flickr, and organized the collection into training data and testing data by users. This setting is built for user-specific applications to predict the popularity of images in other users' collections. When users view a photo album of another person, they may want to find several popular photos to share.

User-mix Dataset (UmD): We downloaded the Visual Sentiment Ontology dataset (Borth et al. 2013) consisting of approximately 540K images from about 75K users. Each user has five published photos at least. This setting often occurs on search engines or photostreams where people see multiple images from other people. We put all the images from various users together and performed popularity prediction on the full data.

Performance Metric: Assuming that the label set of testing data is P, we can obtain a predicted result set P' by running each approach with 10 fold cross validation on the training data. Additionally, we calculate the correlation between P and P' to measure performances of different methods by Spearman Correlation:

$$r = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{P_i - \bar{P}}{\sigma_P}\right)\left(\frac{P'_i - \bar{P}'}{\sigma_{P'}}\right) \tag{7}$$

where $\bar{P}$ and σ_P are the mean and the variance of P.

Performance Comparison with Different Methods

To compare performances with state-of-the-art models, we implemented the following approaches. For showing results on both of "big data" and "small data", we also selected photos randomly from UsD and UmD to organize two smaller subsets in size of 40K and 30K photos respectively.

Average Views: Since similar photo contents may obtain similar influences, the popularity of testing photos were estimated by the average views of its top five similar photos from training data. The popularity s_j can be formulated by $s_j = \frac{1}{k}\sum_{n=0}^{k} s_n$,where k is the number of similar photos, s_n is the popularity of each similar photo with v_j. If the number of similar photos is over five, we rank them by the posting time and select more recent photos.

Bipartite Graph (Niu et al. 2012): Bipartite Graph model is widely used in popularity prediction and ranking (Niu et al. 2012; He et al. 2014). Let $G = (< U \cup V >, E)$ be a bipartite graph, where the set U and set V represent users and items respectively, and the edge E are posting behaviors. We use a regularization term $R(f)$ as

$$R(f) = \frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{m}\omega_{ij}\left(\frac{f(u_i)}{\sqrt{d_i^u}} - \frac{f(v_j)}{\sqrt{d_j^v}}\right) \tag{8}$$

where w_{ij} is defined by the posting behaviors between users and items. d_i^u and d_j^v are the weighted degrees of photo v_j and user u_i for normalization, respectively.

SVR (Khosla, Sarma, and Hamid 2014): Khosla *et al.* used Support Vector Regression (SVR) in understanding the importance of photo content. It can combine different factors by feature vectors. As the same with them, we use SVR with a linear kernel.

BasicMF (Kong, Ding, and Huang 2011): By basic Matrix Factorization (BasicMF), we consider the user-item interaction and regular terms for $\mathbf{U}$ and $\mathbf{V}$ only. The optimization objective is

$$\arg\min_{\mathbf{U},\mathbf{V}} \|\mathbf{R} - \mathbf{UV}\|_F^2 + \lambda_U\|\mathbf{U}\|^2 + \lambda_V\|\mathbf{V}\|^2 \tag{9}$$

PreferenceMF (Cui et al. 2011): This approach is applied to social recommendation with users and items, which predicts popularity by preference rating estimation with the user context constraint and the item context constraint. The objective function is

$$\begin{aligned}&\arg\min_{\mathbf{U},\mathbf{V}} \|\mathbf{R} - \mathbf{UV}\|_F^2 + \lambda_U\|\mathbf{U}\|_F^2 + \lambda_V\|\mathbf{V}\|_F^2\\&+\lambda_C\|\mathbf{C} - \mathbf{VV}\|_F^2 + \lambda_S\|\mathbf{S} - \mathbf{UU}\|_F^2\end{aligned} \tag{10}$$

Multi-scale Temporal Decomposition (MTD): This is our proposed approach, which combines the user-item context constraint and temporal context constraint. The objective function for optimizing the estimation loss is defined in

Table 1: Performances (Spearman Correlation) on user-specific and user-mix datasets

Method	40K UsD	750K UsD	30K UmD	540K UmD
Average Views	0.0574	0.0854	0.1342	0.3202
Bipartite Graph	0.2286	0.2686	0.2345	0.6895
SVR	0.1958	0.2522	0.2316	0.7432
BasicMF	0.1043	0.2425	0.2783	0.4822
PreferenceMF	0.1835	0.2816	0.3172	0.7532
MTD	**0.2462**	**0.3491**	**0.3425**	**0.7977**

Table 2: Performances (Spearman Correlation) of MTD with different time scales on user-specific and user-mix Datasets

Method	750K UsD	540K UmD
Season of Year	0.3025	0.7468
Month of Year	0.3281	0.7542
Week of Year	**0.3417**	**0.7950**
Period of Day	0.3341	0.7907

Equation (6). By our multi-scale model, we can decompose contextual factors into different latent sub-spaces simultaneously. Note that, for BasicMF, PreferenceMF, and MTD, we adjusted parameters by variances of the corresponding matrices as like in the previous work (Jiang et al. 2014).

The results of our approach and the five comparing methods are listed in Table 1. We can observe that:

- The MTD achieves the best performance. This indicates that the user-item context and time-sensitive context can complement each other in the performance.
- The more entries are used for training, the higher correlation coefficient the approaches can achieve. This is consistent with the intuition that the prediction performance depends heavily on the percentage of training data, especially in sparse dataset where the model can be insufficiently trained.
- The comparisons between the BasicMF, PreferenceMF and MTD show that contextual information is effective in popularity prediction.
- The baseline method, Average View, does not fit popularity prediction well on new photos. Our observation is that it ignores other modalities except for content factors.

Relative importance of the adopted time scales. We run the MTD with each time scale individually, and the results are given in Table 2. We can find that 'week of year' is the best time scale to model temporal context, while the period of day is also relatively effective on different data settings.

Parameter analysis. We found that $l = 500$, $\lambda_U = 10^{-1}$, $\lambda_V = 10^{-1}$, $\lambda_S = 10^{-2}$, $\lambda_C = 10^{-1}$ and $\{\lambda_{T_d}\}_{d=1}^{D} = \{10^{-4}, 10^{-3}, 10^{-2}, 10^{-3}\}$ offered the best performance on UmD in our experiments. Note that $\{\lambda_{T_d}\}_{d=1}^{D}$ are tuning parameters ($0 < \lambda_{T_d} < 1$) for controlling the relative importance of the time scales for capturing popularity dynamics. To study hew these parameters affect the performance of popularity prediction by MTD, we set $\lambda_U = 10^{-1}$, $\lambda_V = 10^{-1}$, $\lambda_S = 10^{-2}$, $\lambda_C = 10^{-1}$, and vary the value of each time-sensitive parameter λ_{T_d} separately. Figure 5 shows the results. The larger a λ_{T_d} value is, the more weight will be given to the corresponding time scale. We can see the performance curves of the four time scales are intertwined and it indicates that considering all scales is indeed an effective solution for popularity prediction.

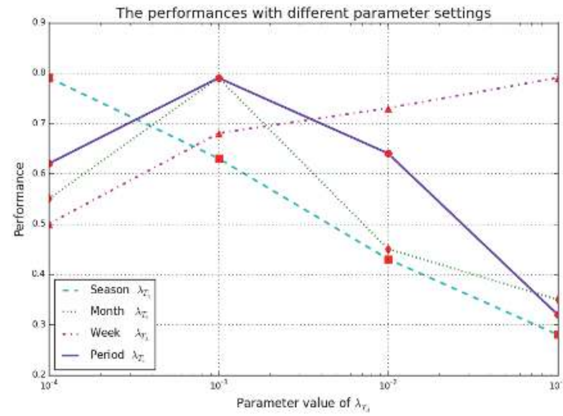

Figure 5: The performance curves of different time scales change with the corresponding λ_{T_d} parameters.

Conclusions and Future Work

In this paper, we have presented a novel approach named Multi-scale Temporal Decomposition (MTD) with context factorization for photo popularity prediction in social media. Specially, we modeled temporal patterns at multiple time scales to capture dynamics of photo popularity. The experimental results showed that our approach outperformed the state-of-the-art methods, with a relative improvement of averagely 9.6% over the best three methods, which demonstrates the effectiveness of multi-scale decomposition for popularity prediction. Besides, we found that the two time scales (week of year and period of day) have the best predictive power than other scales.

There are several research topics for future investigation. One interesting topic is to optimize and suggest best posting behaviors for receiving the highest popularity in social media, such as tagging vivid keywords. Another open question would be investigating the effect of image sentiment on photo popularity and the correlation between user emotion and viewing behaviors. Furthermore, the detection of locality popularity changes or decays with different topics or events is also essential for exploring user behaviors with contextual information in social media.

Acknowledgements

This work was supported in part by National High Technology Research and Development Program of China (2014AA015202), and National Nature Science Foundation of China (61428207, 61571424, 61525206).

References

Ahmed, M.; Spagna, S.; Huici, F.; and Niccolini, S. 2013. A peek into the future: Predicting the evolution of popularity in user generated content. In *WSDM'13*, 607–616.

Borth, D.; Ji, R.; Chen, T.; Breuel, T.; and Chang, S.-F. 2013. Large-scale visual sentiment ontology and detectors using adjective noun pairs. In *MM'13*.

Cappallo, S.; Mensink, T.; and Snoek, C. G. M. 2015. Latent factors of visual popularity prediction. In *ICMR'15*.

Cui, P.; Wang, F.; Liu, S.; Ou, M.; Yang, S.; and Sun, L. 2011. Who should share what? item-level social influence prediction for users and posts ranking. In *SIGIR'11*, 185–194.

Dalal, N., and Triggs, B. 2005. Histograms of oriented gradients for human detection. In *CVPR'05*, 886–893.

D.Lee, D., and Sebastian.Seung, H. 1999. Learning the parts of objects by non-negative matrix factorization. *Nature*.

Donahue, J.; Jia, Y.; Vinyals, O.; Horman, J.; Zhang, N.; Tzeng, E.; and Darrell, T. 2013. Decaf: A deep convolutional activation feature for generic visual recognition. In *arXiv*, 1310–1531.

Figueiredo, F.; Benevenuto, F.; and Almeida, J. M. 2011. The tube over time: Characterizing popularity growth of youtube videos. In *WSDM'11*, 745–754.

He, X.; Gao, M.; Kan, M.-Y.; Liu, Y.; and Sugiyama, K. 2014. Predicting the popularity of web 2.0 items based on user comments? In *SIGIR'14*, 233–242.

Hong, L.; Dan, O.; and Davison, B. D. 2011. Predicting popular messages in twitter. In *WWW'11*, 57–58.

Jiang, M.; Cui, P.; Wang, F.; Zhu, W.; and Yang, S. 2014. A survey on predicting the popularity of web content. *TKDE*.

Khan, R.; de Weijer, J. V.; Khan, F. S.; Muselet, D.; Ducottet, C.; and Barat, C. 2013. Discriminative color descriptors. In *CVPR'13*.

Khosla, A.; Sarma, A. D.; and Hamid, R. 2014. What makes an imag e popular? In *WWW'14*.

Kong, D.; Ding, C.; and Huang, H. 2011. Robust nonnegative matrix factorization using l21-norm. In *CIKM'11*, 673–682.

Lee, D. D., and Seung, H. S. 2001. Algorithms for non-negative matrix factorization. In *NIPS'01*.

Lerman, K., and Hogg, T. 2010. Using a model of social dynamics to predict popularity of news. In *WWW'10*.

McParlane, P. J.; Moshfeghi, Y.; and Jose, J. M. 2014. Nobody comes here anymore, it's too crowded;predicting image popularity on flickr. In *ICMR'14*.

Niu, X.; Li, L.; Mei, T.; and Xu, K. 2012. Predicting image popularity in an incomplete social media community by a weighted bi-partite graph. In *ICME'12*.

Nwana, A. O.; Avestimehr, S.; and Chen, T. 2013. A latent social approach to youtube popularity prediction. In *GLOBECOM'13*, 3138–3144.

Patel, N. 2015. What are the best times to post on social media. In *http://www.quicksprout.com/*.

Pinto, H.; Almeida, J. M.; and Goncalves, M. A. 2013. Using early view patterns to predict the popularity of youtube videos. In *WSDM'13*, 365–374.

Ratkiewicz, J.; Fortunato, S.; Flammini, A.; Menczer, F.; and Vespignani, A. 2010. Characterizing and modeling the dynamics of online popularity. *Physical Review Letters*.

Salakhutdinov, R., and Mnih, A. 2007. Probabilistic matrix factorization. In *NIPS'07*.

Shamma, D. A.; Yew, J.; Kennedy, L.; and Churchill, E. F. 2011. Viral actions: Predicting video view counts using synchronous sharing behaviors. In *ICWSM'11*.

Shen, H.; Wang, D.; Song, C.; and Barabasi, A.-L. 2014. Modeling and predicting popularity dynamics via reinforced poisson processes. In *AAAI'14*.

Tatar, A.; de Amorim, M. D.; Fdida, S.; and Antoniadis, P. 2014. A survey on predicting the popularity of web content. *JISA*.

Yang, J., and Leskovec, J. 2011. Patterns of temporal variation in online media. In *WSDM'11*, 177–186.