A Summary of the TribeFlow Model for Music Discovery Applications

Flavio Figueiredo FLAVIOVDF@BR.IBM.COM
Bruno Ribeiro RIBEIROB@PURDUE.EDU
Jussara M. Almeida JUSSARA@DCC.UFMG.BR
Christos Faloutsos CHRISTOS@CS.CMU.EDU

Abstract

How can we determine which artist a given user will listen to next? Is it possible to create interpretable representations of listening trajectories? In this paper, we present an overview of a recently proposed method, called TribeFlow, developed to tackle questions like these. In music streaming datasets, TribeFlow has been shown to be more accurate than other approaches. Also, TribeFlow is able to represent user behavior in an interpretable and probabilistic latent space.

1. Introduction

Which artist will Alice listen to next? Can listening habits be explained by common patterns? For a long time such questions have attracted the attention of researchers from different fields. In the fields of psychology and musicology (Rentfrow et al., 2011; Rentfrow & Gosling, 2003; Hargreaves, 2012), researchers exploit musical preferences to study social and individual identity (Rentfrow & Gosling, 2003), mood regulation (Saarikallio & Erkkilä, 2007), as well as the underlying factors of preferences (Rentfrow et al., 2011). More recently, computer scientists are tackling such questions as they become central to develop music recommender systems (Figueiredo et al., 2016; Chen et al., 2012; 2013).

With the rise of Online Music Streaming Services (OMSSs) over the last decade or so, large datasets of user behavior can be used to shed light on questions like the ones above. More specifically, digital traces of the listening habits music streaming users are readily available to researchers.

In this paper, we present an overview of TribeFlow. This method has been recently proposed to mine the online listening habits of users as trajectories (Figueiredo et al., 2016) (or trails (Singer et al., 2015)). TribeFlow has been shown to be more accurate and interpretable than state-of-the-art baselines in user trajectory mining (Figueiredo et al., 2016). We describe the method in the next section.

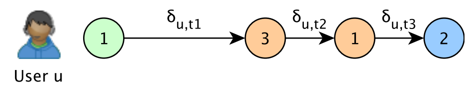

Figure 1. Example of a trajectory T_u. Nodes represent artists, edges represent time. Colors represent environments.

2. TribeFlow in a Nutshell

In order to apply TribeFlow, user listening habits are represented as trajectories. More formally, each trajectory defines the sequence of artists (ordered by time) that the user has listened to. Initially, let us define a user (or listener) as $u \in \mathcal{U}$ and an artist as $a \in \mathcal{A}$. $\mathcal{U}$ and $\mathcal{A}$ is the set of users and artists, respectively. In a song listening dataset, user plays are represented as triples (t, u, a), capturing that u, listened to a at time t. Finally, let $a_{u,t}$ be the artist listened by user u at timestamp t, and t_i simply denote t's position (i) when user time stamps are ordered (e.g., $t_1 < t_2$).

With the definitions above, the trajectory of a user u is defined as $T_u =< (a_{u,t_1}, \delta_{u,t_1}), (a_{u,t_2}, \delta_{u,t_2}), ... >$. T_u is a sequence with each entry $(a_{u,t_i}, \delta_{u,t_i})$ being a tuple representing the artist listened by u at time t (a_{u,t_i}), as well as the amount of time the user dedicates listening to a (δ_{u,t_i}). δ is also known as the inter-event time. For instance, if a user listens to *The Beatles* for one hour, then a_{u,t_i} is *The Beatles* and $\delta_{u,t_i} = 1h$. In Figure 1 we show an example a users trajectory. Each circle is an artist identified by numbers, the time is represented by the length of the arrows. Colors represent environments, which we now describe.

There exists a variety of latent factors that create user trajectories. For instance, the preferences of users is one such factor (e.g., users that like rock music). Geographical constraints are another factor. Some users will listen to bands that come from their home cities, or play where they live. Social factors, such as the preferences of friends, or even characteristics of the OMSS (e.g., recommendation

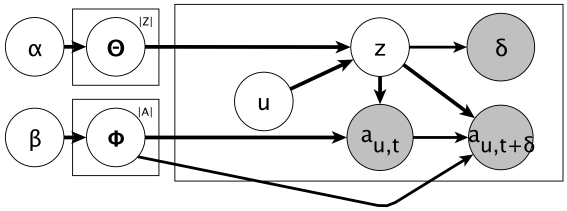

Figure 2. A simplified version of the graphical model for TribeFlow. For the full version see (Figueiredo et al., 2016)

engines) also impact trajectories. What is important to note is that, in the end, the result of listening habits is a trajectory as shown in Figure 1. We call the latent factors that lead to trajectories as environments. In Figure 1 environments are represented by colors. The goal of TribeFlow's inference algorithm is specifically to decompose user trajectories such environments. Trajectories are then aggregated in order to provide an interpretable probabilistic space.

In details, TribeFlow models the listening habits of users as random choices over random environments. That is, TribeFlow captures a generative process that represents users as initially choosing a random environment, say pop music. After choosing this environment, users will fixate their attention to an artist for a given time interval δ. After this period, users will choose a random environment again (which can be the same as before).

In Figure 2 we show a simplified version of the graphical model employed by TribeFlow. The difference between this model and the one originally proposed (Figueiredo et al., 2016) is that in our summary we focus on bursts of size one. That is, users choose environments after each play. See the original paper for a more in-depth discussion of the effect of burst sizes. Also, we removed the stick breaking prior. In this figure, α and β are Dirichlet hyperparameters. z is a latent environment.

TribeFlow works using as input a set of trajectories $T_c \in \mathcal{T}$. Let, $\mathcal{Z}$ be the set of latent environments. The size of this set $k = |\mathcal{Z}|$ is learned from the data using split and merge moves (Figueiredo et al., 2016). Θ and Φ are outputs of the model, initialized randomly and updated in each learning step. TribeFlow works by sampling from the posterior defined by the model in Figure 2:

$$P[z, a_{u,t_{i+1}}, a_{u,t_i}, s, \delta] = \frac{P[z|u]P[a_{u,t_i}|z]P[a_{u,t_{i+1}}|z]P[\delta|z]}{1 - P[a_{u,t_i}|z]}$$

The model is learned using a EM algorithm. The e-step is captured by a Gibbs sampling step over every play on the dataset. The m-step estimates $P[\delta|z]$. $P[z|u]$, and $P[z|a]$ are both sampled from a multinomial using the Dirichlet priors defined by Θ and Φ. The term $P[\delta|z]$ captures a distribution of inter-event times for a given environment. This term is heuristically captured using the survival probability of inter-event times assigned to z during the e-step.

	Last.FM-Groups	Last.FM-1k
FPMC	0.00043	0.00048
PRLME	0.10861	0.10354
TribeFlow	**0.18301**	**0.16735**

Table 1. Mean Reciprocal Rank Values for TribeFlow and state-of-the-art competitors: FPMC (Rendle et al., 2010) and PRLME (Feng et al., 2015)

3. Results

We now show the results of TribeFlow at work. We evaluate the model on two Last.FM datasets, namely Last.FM-1k and Last.FM-Groups. Last.FM-1k captures captures roughly 10 million plays from 1 thousand users. Last.FM-Groups captures over 80 million plays from 15,000 users.

Before describing some of our results, we note that TribeFlow was the only approach to execute in these datasets (without any filtering) in under a day (Figueiredo et al., 2016). Other methods did not finish in over 10 days. Thus, we present some results comparing TribeFlow with other methods based on sub-samples of 10k user plays from each dataset. 70% of each dataset is used for training and validation (the first 70% of plays). Evaluations are performed on the last 30% of plays.

Our first competitor is Factorizing Personalized Markov Chain (FPMC) (Rendle et al., 2010). FPMC was initially proposed to predict the next object a user will insert into an online shopping basket. If we consider a user has listening to $a_{u,t}$, FPMC can be used to rank the candidate artists, or destinations that a user will listen to next. Our second competitor is the best-performing Latent Markov Embedding (LME) (Chen et al., 2012; 2013) method in our datasets, called Personalized Ranking by Latent Markov Embedding (PRLME) (Feng et al., 2015).

We evaluate each method using the mean reciprocal rank over (MRR) (Figueiredo et al., 2016) every play on the evaluation set. Our results, depicted in Table 1, show that TribeFlow is at least 60% more accurate than competitors. PRLME, the second best results, achieved MRR values of roughly 0.10 on both datasets. In contrast, TribeFlow achieves 0.16 and 0.18 for Last.FM-Groups and Last.FM-1k respectively. Similar results are achieve on other subsets of the data (Figueiredo et al., 2016).

4. Conclusions

In this paper we presented a summary of TribeFlow for music discovery applications. TribeFlow is a novel approach, and it is more accurate than state-of-the art methods when predicting which artist a user will listen to next. As future work, we plan on experimenting with TribeFlow on other datasets and other domains.

Acknowledgments: Research supported by the project FAPEMIG-PRONEX-MASWeb, process number APQ-01400-14.Our work was also supported by NSF grant CNS-1065133, U.S. Army Research Laboratory under Cooperative Agreement W911NF-09-2-0053, and by the Defense Threat Reduction Agency contract No. HDTRA1-10-1-0120. The views and conclusions contained here are our own (the authors).

References

Chen, Shuo, Moore, Josh L, Turnbull, Douglas, and Joachims, Thorsten. Playlist prediction via metric embedding. In *Proc. KDD*, 2012.

Chen, Shuo, Xu, Jiexun, and Joachims, Thorsten. Multi-space probabilistic sequence modeling. In *Proc. KDD*, 2013.

Feng, Shanshan, Li, Xutao, Zeng, Yifeng, Cong, Gao, Chee, Yeow Meng, and Yuan, Quan. Personalized ranking metric embedding for next new POI recommendation. In *Proc. IJCAI*, 2015.

Figueiredo, Flavio, Ribeiro, Bruno, Almeida, Jussara M, and Faloutsos, Christos. TribeFlow: Mining & Predicting User Trajectories. In *Proc. WWW*, 2016.

Hargreaves, David J. Musical imagination: perception and production, beauty and creativity. *Psychology of music*, 40(5):539–557, 2012.

Rendle, Steffen, Freudenthaler, Christoph, and Schmidt-Thieme, Lars. Factorizing personalized Markov chains for next-basket recommendation. *Proc. WWW*, 2010.

Rentfrow, P. J. and Gosling, S. D. The Do Re Mi's of Everyday Life: The Structure and Personality Correlates of Music Preferences. *Journal of personality and social psychology*, 84(6), 2003.

Rentfrow, Peter J., Goldberg, Lewis R., and Levitin, Daniel J. The structure of musical preferences: a five-factor model. *Journal of personality and social psychology*, 100(6):1139–1157, June 2011.

Saarikallio, Suvi and Erkkilä, Jaakko. The role of music in adolescents' mood regulation. *Psychology of music*, 35 (1):88–109, 2007.

Singer, Philipp, Helic, Denis, Hotho, Andreas, and Strohmaier, Markus. HypTrails: a bayesian approach for comparing hypotheses about human trails on the web. In *Proc. WWW*, 2015.