Metadata and Memory Cues in Collaborative Tagging: Music Listening and Tagging on Last.fm

Jared Lorince & Peter M. Todd

Abstract

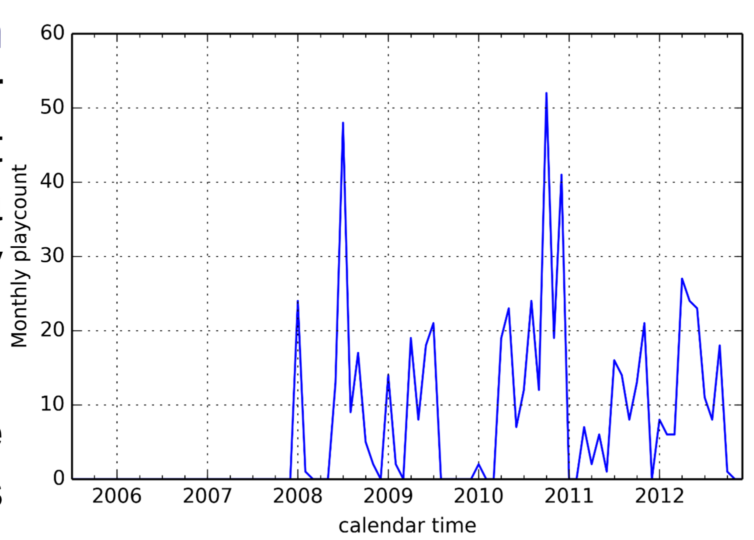
How do people create memory aids for themselves? In online collaborative tagging systems, users annotate resources with free-form textual metadata (tags). It is commonly proposed that users tag resources to facilitate re-finding them later, but this has not been directly tested. If users do tag items for later retrieval, we would expect different listening patterns for tagged versus untagged items. To explore this, we use a large-scale dataset from the social music site Last.fm to ask:

- **RQ1**: How can we measure and identify particular classes of listening patterns (e.g. regular, sporadic, or isolated listening) among the thousands of users and millions of songs in our data?
- **RQ2**: How are tags temporally distributed for items in these different classes? That is, when over the time course of listening to a song or artist is a user most likely to tag it?
- **RQ3**: Do we see evidence that application of tags predicts subsequent listening to an artist, supporting a possible memory-aid use?

Dataset

We collected complete listening histories from **~80,000 Last.fm users**, totaling **~1.4 billion individual listens**. The data for each user was represented as a set of time series – one per unique artist listened – consisting of the number of times a given user listened to a given artist each month within the time range of our data sample (July 2005 – December 2012).

Our data consisted of ~67 million such time series, trimmed to those in which a user listened to an artist at least 25 times total. This left us with **~8.6 million time series**, of which 1 **million** were included in the presented analyses, owing to computational constraints.

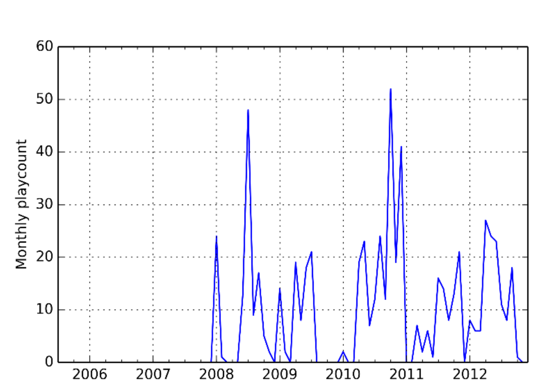


Example time series showing a user's monthly playcounts for a particular artist

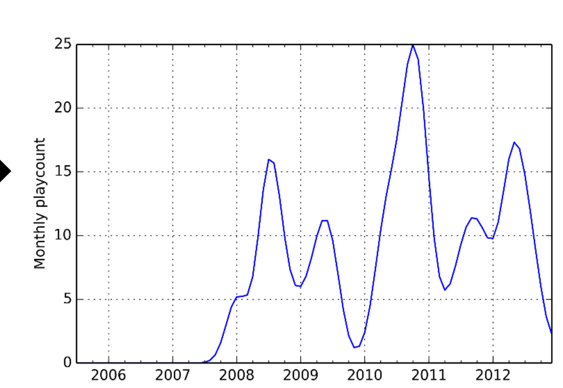
Clustering Process

The goal of clustering was to identify canonical, scale-invariant forms of listening series existing in the data. This involved a process of **smoothing**, **alignment**, and **normalization**, followed by the application of a standard **k-means clustering algorithm**.

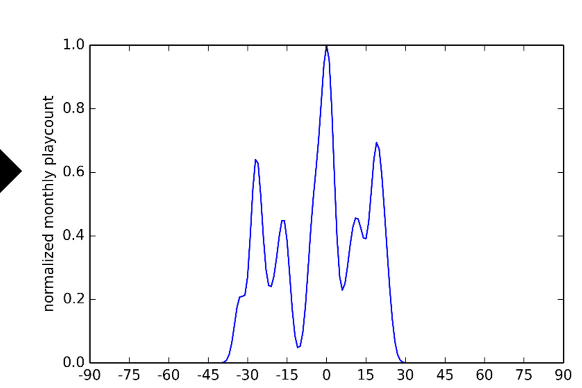
We are testing various transformation methods, all following the following framework:



Raw time series

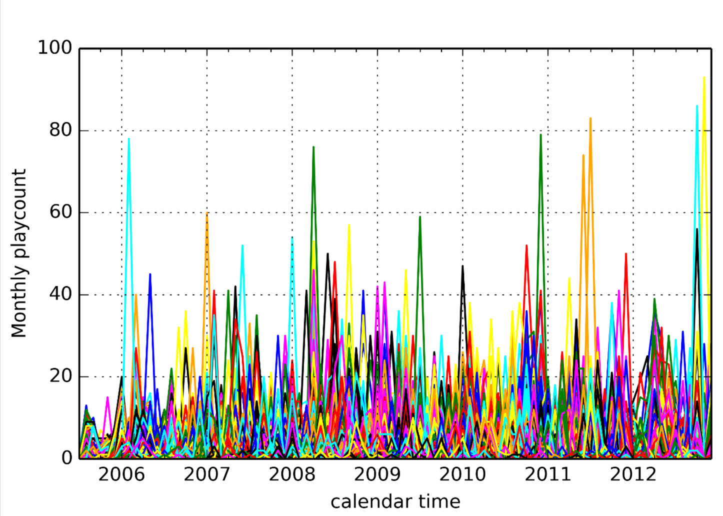


Smoothed time series

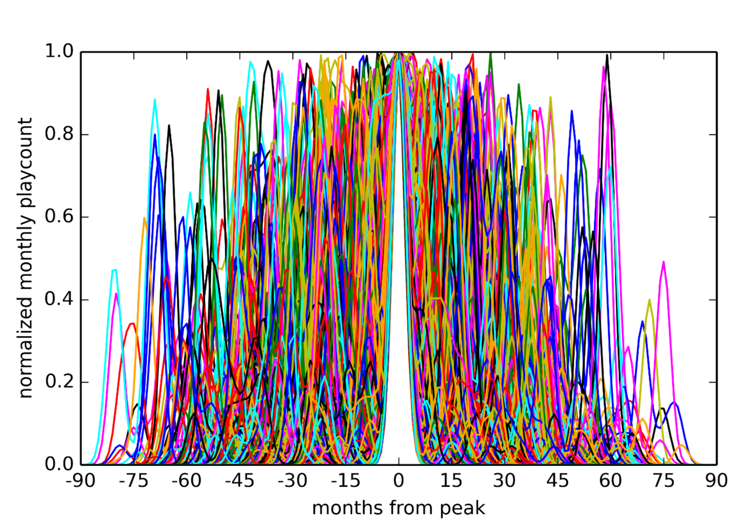


Aligned and normalized time series

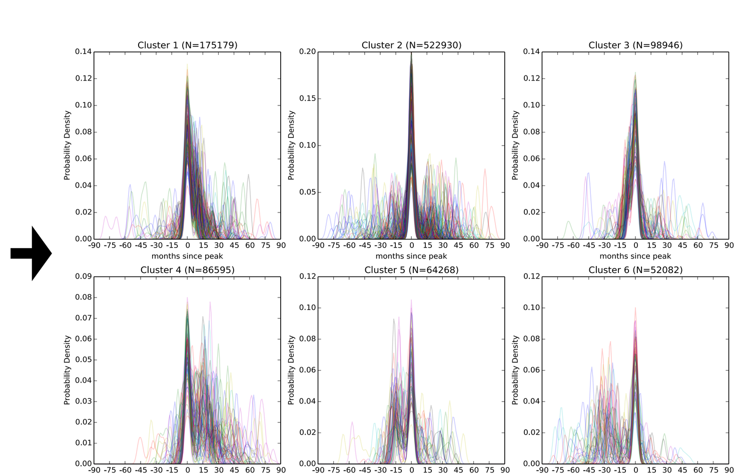
This process is repeated for the full set of 1 million time series to be clustered:



Raw time series



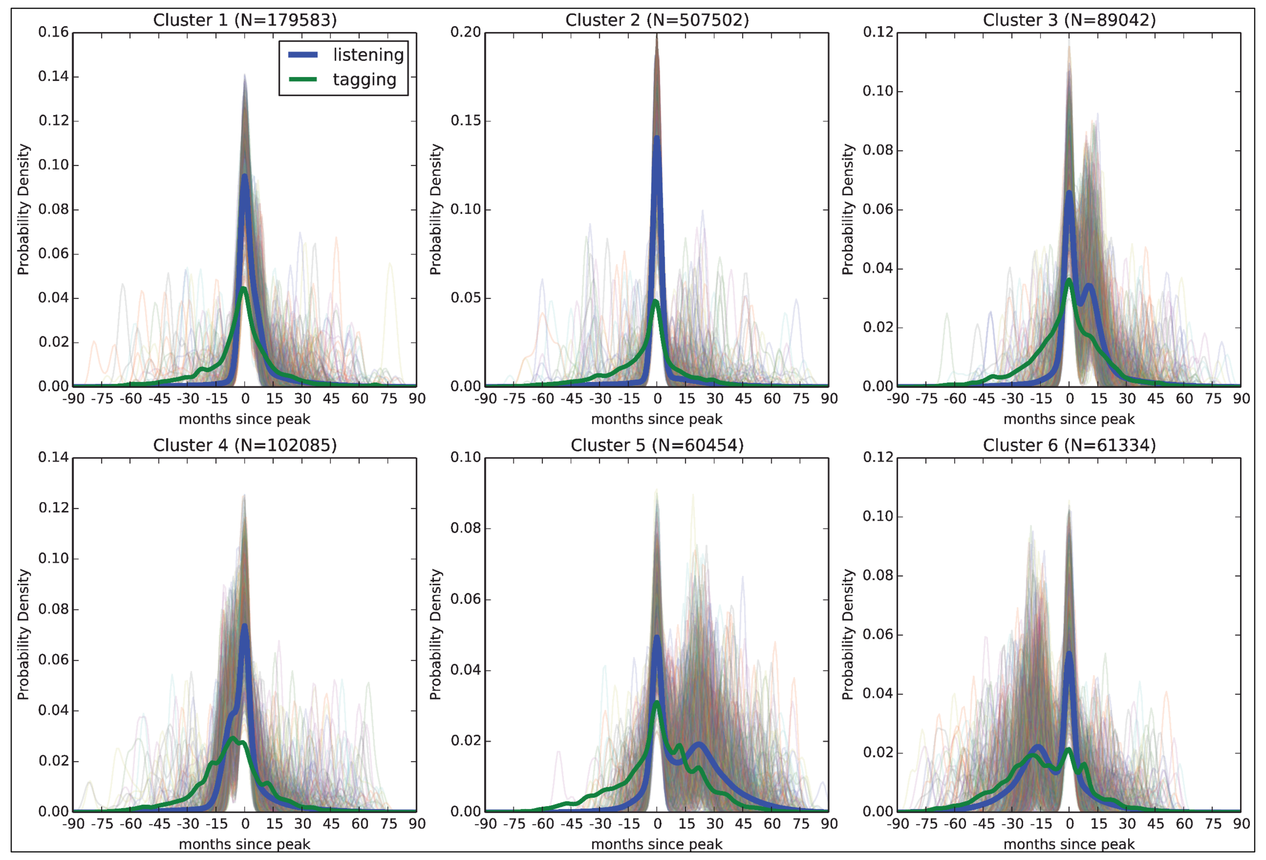
Transformed time series



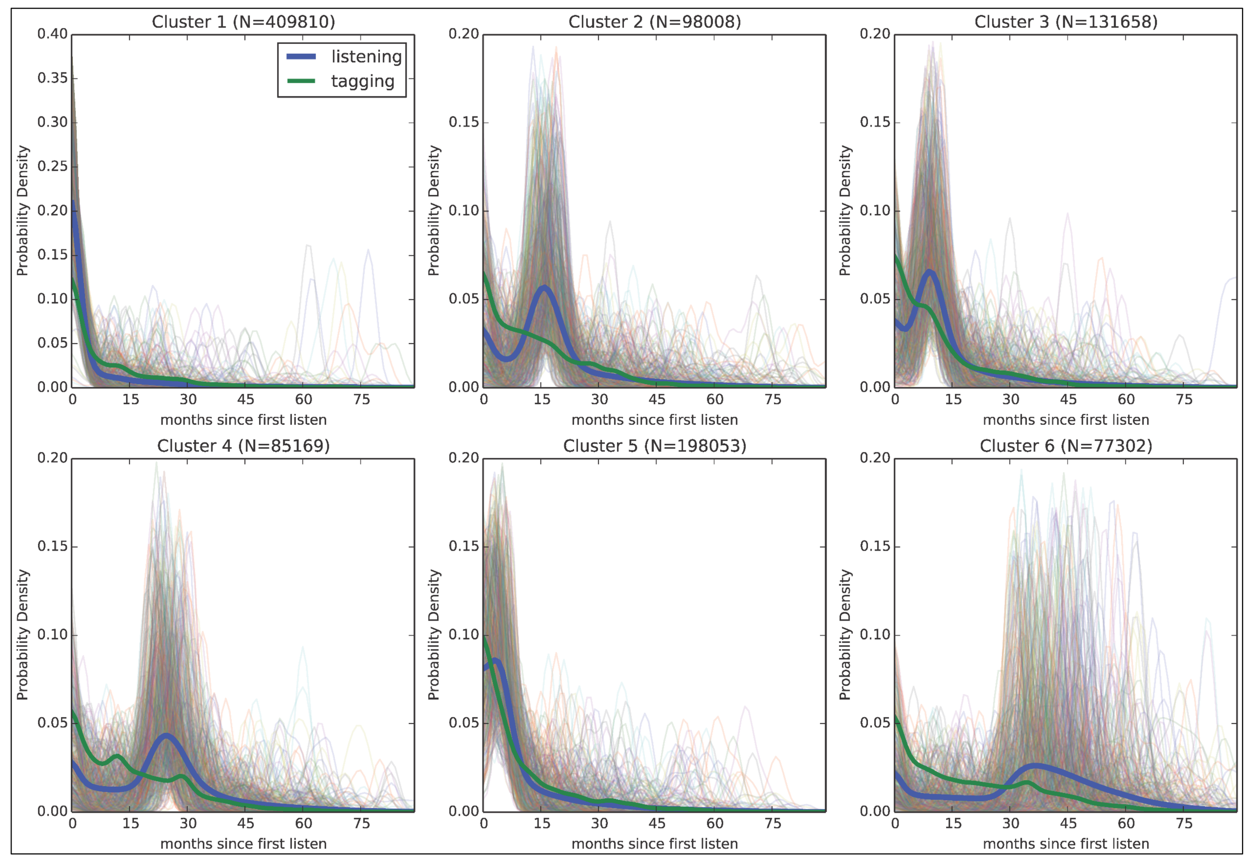
Clustered time series

Clustering Results

K-means categorizes the time series into maximally self-similar clusters, and defines a centroid for each. There is no ground truth dictating the "real" number of clusters in the data, and the algorithm requires a priori specification of the number of clusters, *k*. Preliminary analyses under varying *k*, however, suggest that the analysis is most informative using ~6 clusters. We present results for *k*=6 under two alignment methods.



"Peak" alignment: Each time series is aligned so that the month with the most listens is in the middle of the series



"Start" alignment: Each time series is aligned so that the first month in which a user listens to the artist is the first element of the series

In all plots, the blue line shows the **centroid** of a cluster, and the fine lines show a sample of the constituent time series in that cluster. These results reveal several distinctive and identifiable listening patterns, addressing **RQ1**. We find, for example, isolated bursts of listening (peak-aligned, cluster 2), slowly increasing listening (start-aligned, cluster 6), and several forms of bimodal listening distributions. Determining which normalization and alignment methods most faithfully represent the data remains for future work.

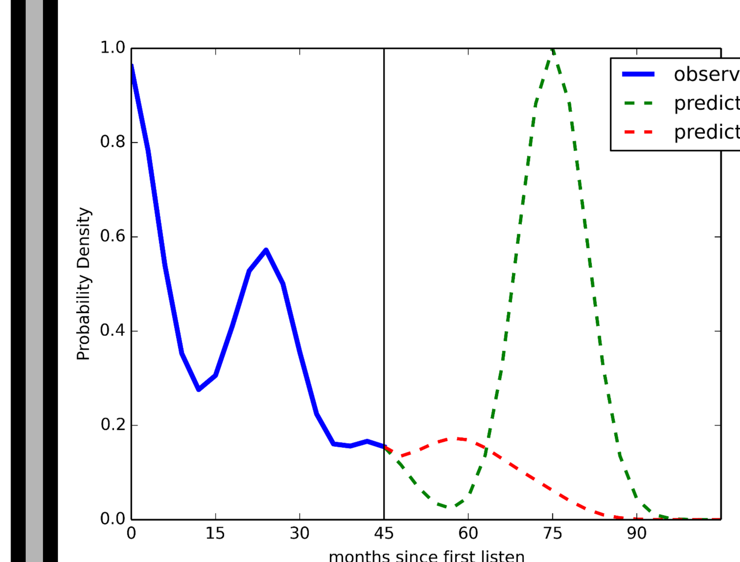
To address **RQ2**, we calculated the *aggregate* probability distribution of tagging across all time series in each cluster, temporally aligned to match the listening data and plotted in green. Thus each green line shows the distribution of tagging activity relative to the average listening within each cluster.

Unsurprisingly, the tagging distributions have roughly the same form as the listening centroids (i.e. users are more likely to tag an artist in the months they listen to that artist more). Two observations, however, are consistent with a "tags as memory cues" interpretation, addressing **RQ3**:

- **Tagging is most likely in the first month users listen to an artist**, even when their listening peaks later.
- Overall, **increases in tagging *precede* increases in listening** (evidenced by leftward shifts in the tagging distributions relative to the listening centroids)

Next Steps: A predictive model

Our next goal is to leverage our clustering analysis in the development of a **predictive model** of how the choice to tag affects future listening.



Hypothetical result of a predictive model. Given a listening time series, and assuming a tag is assigned at time *t* (in this example, 45 months after first listen, vertical line), what subsequent pattern of listening do we predict after tagging compared to the listening pattern if there were no tagging?

Conclusions

- Clustering methods identify well-defined classes of canonical listening patterns in our data.
- Comparison of aggregate listening and tagging patterns are *consistent* with tags being used as memory aids
- A model predicting subsequent listening given the application of a tag may provide stronger evidence for the "tags as memory cues" interpretation.

Selected References

- Anderson & Schooler (1991). Reflections of the environment in memory. *Psychological Science.*
- Glushko, et al. (2008). Categorization in the wild. *Trends in Cognitive Sciences.*
- Golder, & Huberman, (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science.*
- Gupta, et al. (2010). Survey on social tagging techniques. *SIGKDD Explorations Newsletter.*
- Yang & Leskovec (2011). Patterns of temporal variation in online media. *Proc. WSDM.*