Visualizing Program Genres' Temporal-Based Similarity in Linear TV Recommendations

Veronika Bogina University of Haifa Haifa, Israel

Tsvi Kuflik University of Haifa Haifa, Israel

ABSTRACT

There is an increasing evidence that data visualization is an important and useful tool for quick understanding and filtering of large amounts of data. In this paper, we contribute to this body of work with a study that compares chord and ranked list for presentation of a temporal TV program genre similarity in next-program recommendations. We consider genre similarity based on the similarity of temporal viewing patterns. We discover that chord presentation allows users to see the whole picture and improves their ability to choose items beyond the ranked list of top similar items. We believe that similarity visualization may be useful for the provision of both the recommendations and their explanations to the end users.

CCS CONCEPTS

• Human-centered computing \rightarrow Information visualization.

KEYWORDS

Visualization, similarity, recommender system

ACM Reference Format:

Veronika Bogina, Julia Sheidin, Tsvi Kuflik, and Shlomo Berkovsky. 2020. Visualizing Program Genres' Temporal-Based Similarity in Linear TV Recommendations. In *International Conference on Advanced Visual Interfaces (AVI '20), September 28-October 2, 2020, Salerno, Italy.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3399715.3399813

1 INTRODUCTION

We are living in the era of big data, where users can access online exceedingly large volumes of multifaceted information. Monitoring and understanding changes in online information consumption over time has become of interest to both professional analysts and the general public, in order to inform the investigation and analysis of trends in areas such as politics, economics, security, public opinion, and more [3, 11, 12]. Visualization plays a pivotal role here, since it can help users overcome perception- and attentionrelated limitations constraining information processing and help

AVI '20, September 28-October 2, 2020, Salerno, Italy

© 2020 Association for Computing Machinery.

ACM ISBN 978-1-4503-7535-1/20/09...\$15.00

https://doi.org/10.1145/3399715.3399813

Julia Sheidin University of Haifa Haifa, Israel

Shlomo Berkovsky Macquarie University Sydney, NSW, Australia

them focus on the most relevant or most interesting aspects of the data [1, 5].

Considering the linear TV domain, understanding viewing habits of users is crucial for providing relevant recommendations. However, the continuously growing quantity of available data poses a real challenge for anyone willing to make sense of it and further present it to the user in a clear and comprehensible manner.

Hence, we are considering the problem of personalized TV program recommendations. The recommendations are examined in the following use-case: a target user is watching TV and willing to switch to another program. They may possibly be willing to continue and watch programs from the same genre or experience diverse type of content. So, they will be shown with a list/table of currently broadcast program genres with the connections between them (through the lens of similarity). Afterwards, program recommendation will be provided according to the program genre selected by the user.

In this research we address the following question: *How can* similar and dissimilar TV program genres be visually presented to the end user for next program recommendations?

The outline of the paper is as follows. First, we introduce the problem we address and the main research questions. Afterwards, we elaborate on our approach and discuss the experiment conducted on a linear TV data set to visualize the degree of similarity between genres, together with a comparative study of two visualization techniques. Finally, we discuss the limitations and future implications of the proposed method and visualization technique.

2 RELATED WORK

Various approaches to incorporating temporal factors in similarity metrics have been studied in the past. Such metrics are frequently used in collaborative filtering to find item/user neighborhood and generate recommendations [6] and in sequential recommendations [10]. Temporal factors are explicitly or implicitly embedded into Pearson's correlation, Jaccard's similarity [14] or cosine similarity[2]. Furthermore, time weights and decays are used to prioritize recent data [4, 7, 9]. In contrast to past studies, we base genre similarity on the similarity of temporal patterns.

To explain recommendations, researchers often resort to textual and visual explanations, as well as their combination. Kouki et al. [8] suggested using three formats for music artist recommendation: Venn diagrams, static cluster dendrograms and text. They focused on the persuasiveness of explanations and showed that textual explanations were perceived more persuasive than the visual Vienn

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Veronika Bogina, Julia Sheidin, Tsvi Kuflik, and Shlomo Berkovsky



Figure 1: Similar temporal patterns



Figure 2: Dissimilar temporal patterns

diagram and the dendrogram. Tsai and Brusilovsky [13] stepped beyond relevance and focused on diversity visualization in people recommendations. They measured diversity using Shannon's entropy and used a scatter plot to visualize it, which was preferred by users over a ranked list.

3 METHODS

3.1 Genre Similarity

One of the challenges of linear TV is how to recommend to users next program to watch. One of the approaches for next item recommendation is to recommend an item with a genre similar to the one the user is currently consuming. Alternatively, a different genre can be recommended, if diversity or serendipity are prioritized. Hence, we decided to compare the temporal distribution of the program genre consumption throughout the day. Our basic assumption is that if the temporal patterns of two genres are similar, then these genres may also be similar from the *screen time availability perspective*, not necessarily for the content itself.

Consider the following example, where we split the day into 96 slots of 15 minutes each (starting from midnight). *Adventure* programs are popular during the afternoon and evening hours, while at

night these are not consumed (Figure 1). If the same trends are observed for the *Action* genre, then these genres have similar viewing patterns. On the contrary, if genre distributions are dissimilar, such as for the *Anthology* and *Adults* content (Figure 2), then they are not similar. Figure 1 and Figure 2 present similar and dissimilar genres, respectively, based on their time distributions. As mentioned, there are 96 slots of 15 minutes (numbers on on the X-axis), that start at midnight ("0" on the X-axis in these figures).

3.2 Similarity Visualization

To visualize similarity of various TV program genres, we propose to use a chord diagram¹ that visually represents connections between several entities. The diagram is based on a sphere, where each entity is a fragment of the circumference. Entities are connected to each other by arcs, the thickness of which is proportional to the degree of similarity between the entities (Figure 3).

In our case, the visualization presents the program genres that are currently broadcast on TV. The user can explore similar or dissimilar genres by selecting the genre she is currently watching (Figure 4). Upon selecting another program genre, the similarity score between the two genres is visualized by the arc thickness.

4 EXPERIMENT

We conducted an initial user evaluation, to examine the usability of our visualization. The goal of our experiment was to collect data about user perceptions and opinions regarding the studied visualizations in order to understand which one is more effective and which one is preferred by the users in the targeted case study. We initially describe the linear TV data set used for measuring the genre similarity and then we compare two genre similarity visualizations.

4.1 Data Set

To illustrate our assumptions, we use a real-world proprietary linear TV data set collected by FourthWall Media, containing household TV program views logged during the first five months of 2015. The data set contains viewing history of households in the US region of Little Rock Pine Bluff, which is represented by the household id, timestamp, program id, and program genre. There are 120 unique genres in the data, where one of them is *Idle*.

4.2 Genre Similarity

To quantify the genre similarity, we compare temporal distributions of the genres consumption and KL-divergence entropy² for all the pairs of genres. It measures the difference between two probability vectors P and Q and is calculated as follows:

$$D_{KL}(P||Q) = \sum_{x \in X} P(x) log(\frac{P(x)}{Q(x)})$$
(1)

where P(x) and Q(x) are the relative genres consumption at time x. Low values of KL-divergence indicate high similarity of genres and, vice versa, high values indicate dissimilarity.

At the pre-processing phase, we split each day into 15-minute slots. The duration of the slots was guided by the average duration

¹https://datavizcatalogue.com/methods/chord_diagram.html

²https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.entropy.html

Visualizing Program Genres' Temporal-Based Similarity in Linear TV Recommendations



Figure 4: Hovering over horror genre

of programs in the data set. Thus, a day is represented by 96 slots, starting from midnight. Then, for each slot we calculated how many times programs of each genre were viewed. These values were normalized to produce the daily distribution of each genre, which allowed to compute KL-divergence and find similar genres.

4.3 Visualization of Genre Similarity

As explained in Section 3.2, we used the chord diagram to visualize similarity and dissimilarity of the program genres³, where entropy was reversed to show similar genres with thicker arcs and less

AVI '20, September 28-October 2, 2020, Salerno, Italy

Table 1: An example of ranked list, showing 5 pairs of	f TV
genres and their similarity scores	

Shopping	Music	33.5
Action	Horror	32.3
Horror	Action	32.2
Music	Shopping	30.9
Biography	Comedy	28.2

similar — with the thinner ones. Figure 3 presents an example chord diagram visualization for all the broadcast genres. When hovering over a genre, as shown in Figure 4, the similarity of the selected genre with the other genres is visualized by the arcs. For example, *Horror* is similar to the *Action* and *Biography* genres, but different from *Children* and *Skateboarding*. Upon selecting a genre, the user is willing to switch to, the list of programs from this genre — broadcast at the given time — is shown.

Another common method for representing similar genres is a well-know ranked list, which is essentially a table showing pairs of TV genres and their similarity/dissimilarity scores (refer to Table 1). In our study, upon identifying the current genre, a list is shown with top 20 genres most similar to the selected genre. These are presented in a list sorted according to their similarity with the current genre. As such, the most similar genre appears at the top of the list and the least similar (out of the 20) — at the bottom.

4.4 Participants

While our main goal was to discover whether the chord visualization allows users to see the whole picture and improves their ability to choose items beyond the top 10 similar items, in the current paper we focus on an initial evaluation examining wherever the chord diagram facilitates a better understanding of the similarity of program genres. Hence, the users of our system is a general public and the participants were selected in a way representative of the entire population. We recruited participants from a range of ages (from fresh graduates to elderlies, average age was 41.23 (SD = 15)) and professions (not all technology savvy). 13 participants took part in the study, seven males and six females. They all had normal or corrected-to-normal eyesight, and no participants were color blind (self-reported). All the participants had an opportunity to withdraw from the study at any time.

4.5 Task and Procedure

We constructed three domain-oriented tasks that were highlyrelated to understanding of the genre similarity. The participants were asked to find (1) a specific genre and then the (2a) most similar and (2b) most dissimilar genres.

The participants were seated in front of a paper mock-up and were first briefed about the experiment. After the study introduction, a short questionnaire collecting personal information was administered. Afterwards, the participants were given two consecutive session blocks. Each block contained the ranked list and chord diagram visualizations and included a training session. The participants were asked to work as quickly and accurately as possible. The duration and the result for each task were recorded. After each

³https://github.com/sveron/TV-Genres-Visualization

AVI '20, September 28-October 2, 2020, Salerno, Italy

	Question	Ranked List	Chord Diagram
Q1	Was most effective?	11	2
Q2	Was most enjoyable?	7	6
Q3	Was easiest to learn?	11	2
Q4	Was the best to present similarity and dissimilarity between different program genres?	2	11
Q5	Do you prefer overall?	10	3

Figure 5: Overall preference between the two methods for five questions starting with: "Which method...". Each cell indicates how many participants preferred each method for that question (N = 13).

block, the participants' subjective assessment of each visualization (on a 7-Likert scale) was assessed. At the end of the study, the participants were asked to fill out a questionnaire directly comparing the two visualizations. The order of the visualizations was randomized using a Latin square design. Each session lasted for 10-15 minutes.

4.6 Results: Chord vs Ranked List Visualization

We present the study results in two parts: first effectiveness (measured by task completion time and accuracy) and then the reported subjective preferences.

We first compared the overall average completion times of tasks using the two visualizations of the genre similarity (ranked list: 1.62 and chord: 2.15) and observed a significant differences in the completion times. Our results reveal that ranked list was significantly faster than the chord (p=0.018). For accuracy, our results revealed that the ranked list was significantly more accurate than the chord diagram (error rate of 61.5% for chord and 8% for ranked list).

In order to analyze the participants' preferred visualization, we used a one-tailed t-test. The results reveal that the ranked list was significantly more effective than the chord (p=0.021), easier to use (p=0.008) and more clear (p=0.016). The results also show that chord was significantly more frustrating (p=0.009) and more complicated (p=0.0006) than the ranked list.

At the end of the experimental session, five comparative questions were presented to the participants, who were asked to choose their preferred visualization for each question. The results are presented in Figure 5. Most participants found the ranked list more effective (Q1), easier to learn (Q3), and preferable (Q5) over the chord diagram. This is unsurprising, since the list is considered as one of the most familiar visualizations [13]. For the most enjoyable visualisation (Q2), the participants showed no clear preference. However, most participants preferred the chord diagram for presenting similarity and dissimilarity of genres (Q4).

5 DISCUSSION AND LIMITATIONS

In this paper we proposed that temporal patterns of program consumption could be used for measuring similarity of TV program genres. Furthermore, recommendations of program genres can be visualized through such program genre similarity. To this end, we conducted a user study comparing two program genre similarity visualisations: chord diagram and ranked list. Our contribution is two-fold. First, we use temporal factors, such as program genre consumption times, to measure similarity of linear TV program genres. Second, we compare two techniques to visualize the genre similarity scores for users.

Although the participants generally preferred the ranked lists, we believe that the chord visualization may be useful for providing a more complete picture of the genres similarities and explaining similarity-based recommendations. It can show in a better way oneto-many relationships (k most similar genres) or turn out stronger for a complex multi-dimensional data. Afterwards, users can choose any type of similar/dissimilar programs, not only from the list of top n ranked items that may limit users' choice. When choosing from all available broadcast genres, users may feel that the system is more transparent by presenting all the data; therefore, they will build a stronger trust in such systems. Moreover, the user will be the one who will choose an appropriate genre for her current context, given all the available data, and thus may be more satisfied with the final outcome.

In a controlled experiment, each experimental design decision brings with it some trade-offs. First, we acknowledge that there are many other techniques that could be applied for the data and task at hand. Our results are therefore limited to the two visualization techniques used in the experiment. Second, the current design decisions pose limitations on the empirical evaluation. Ranked list had the similarity score ordered; hence, it was unsurprisingly easier to find the most similar and dissimilar genres. In addition, original chord diagrams do not use different colors when hovering over the selected genre. Thus, using the same color and adding transparency might have over-simplified the chord visualization. We believe that implementing these improvements could have made the chord diagram more effective and usable.

We believe that chord and ranked list can be combined together in a dashboard. Once a user understands connections between different genres, by clicking on the target genre, she can see the sorted list of related genres.

In the future, we intend to evaluate the same approach for music recommendations, with a finer granularity of the temporal sampling (down from the 15-minute intervals). It would be interesting to compare such a similarity of genres to the one created by music experts. We also intend to analyze KL-divergence values used as a similarity measure and consider specific ranges of values appropriate for serendipitous recommendations.

ACKNOWLEDGMENTS

The study was partially supported by the Israeli Science Foundation ISF 226/17 grant and by the CyCAT, which has received funding from the European Union's Horizon 2020 Research and Innovation Program under Grant Agreement No. 810105.

REFERENCES

- Wolfgang Aigner, Silvia Miksch, Heidrun Schumann, and Christian Tominski. 2011. Visualization of time-oriented data. Springer Science & Business Media.
- [2] Pedro G Campos, Fernando Díez, and Iván Cantador. 2014. Time-aware recommender systems: a comprehensive survey and analysis of existing evaluation protocols. User Modeling and User-Adapted Interaction 24, 1-2 (2014), 67–119.

Visualizing Program Genres' Temporal-Based Similarity in Linear TV Recommendations

AVI '20, September 28-October 2, 2020, Salerno, Italy

- [3] Siming Chen, Lijing Lin, and Xiaoru Yuan. 2017. Social media visual analytics. In Computer Graphics Forum, Vol. 36. Wiley Online Library, 563–587.
- [4] Yi Ding and Xue Li. 2005. Time weight collaborative filtering. In Proceedings of the 14th ACM international conference on Information and knowledge management. ACM, 485–492.
- [5] Jeffrey Heer, Michael Bostock, Vadim Ogievetsky, et al. 2010. A tour through the visualization zoo. Commun. Acm 53, 6 (2010), 59–67.
- [6] Jonathan L Herlocker, Joseph A Konstan, Loren G Terveen, and John T Riedl. 2004. Evaluating collaborative filtering recommender systems. ACM Transactions on Information Systems (TOIS) 22, 1 (2004), 5–53.
- [7] Christoph Hermann. 2010. Time-based recommendations for lecture materials. In *EdMedia+ Innovate Learning*. Association for the Advancement of Computing in Education (AACE), 1028–1033.
- [8] Pigi Kouki, James Schaffer, Jay Pujara, John O'Donovan, and Lise Getoor. 2019. Personalized explanations for hybrid recommender systems. In Proceedings of the 24th International Conference on Intelligent User Interfaces. ACM, 379–390.
- [9] Hao Liao, Ming-Kai Liu, Manuel Sebastian Mariani, Mingyang Zhou, and Xingtong Wu. 2019. Temporal similarity metrics for latent network reconstruction:

The role of time-lag decay. Information Sciences 489 (2019), 182-192.

- [10] Massimo Quadrana, Paolo Cremonesi, and Dietmar Jannach. 2018. Sequenceaware recommender systems. ACM Computing Surveys (CSUR) 51, 4 (2018), 66.
- [11] Steven F Roth and Joe Mattis. 1990. Data characterization for intelligent graphics presentation. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 193–200.
- [12] Julian B Rotter. 1966. Generalized expectancies for internal versus external control of reinforcement. *Psychological monographs: General and applied* 80, 1 (1966), 1.
- [13] Chun-Hua Tsai and Peter Brusilovsky. 2018. Beyond the ranked list: User-driven exploration and diversification of social recommendation. In 23rd International Conference on Intelligent User Interfaces. ACM, 239–250.
- [14] Dawei Yin, Liangjie Hong, and Brian D Davison. 2011. Exploiting session-like behaviors in tag prediction. In Proceedings of the 20th international conference companion on World wide web. ACM, 167–168.