Dynamic user-affinity based curation

Justin Lewis
Scott Davies

Follow this and additional works at: http://www.tdcommons.org/dpubs_series

Recommended Citation
Lewis, Justin and Davies, Scott, "Dynamic user-affinity based curation", Technical Disclosure Commons, (May 09, 2017)
http://www.tdcommons.org/dpubs_series/501

This work is licensed under a Creative Commons Attribution 4.0 License.
This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.
Dynamic user-affinity based curation

ABSTRACT

Online content providers recommend content based on the user’s viewing history, e.g. to sustain user interest and promote content that users are likely to be interested in. Current recommendation systems generally focus on one signal, e.g., content viewed by others who also viewed the present content (“co-watch signal”), to make recommendations. Co-watch signal-based recommendations tend to skew towards lowest-common-denominator content and are not capable of accurately recommending niche content. Niche content recommendations that are poorly targeted result in low click-through / view-through rates (CTR/VTR) and miss truly interested audiences. This disclosure describes recommendation systems that take contextual information into account. Context is established, with user permission, based on content that the user has interacted with in the past. Per techniques of this disclosure, content is recommended if the co-watch signal is strong and the context-matching score is high.

KEYWORDS

- Niche content
- Co-watch signal
- Recommender system
- Contextual information
- Shoulder content

BACKGROUND

Online content providers recommend content based on a visitor’s viewing history, when use of such history is permitted by the visitor. Current recommendation systems generally focus on a particular signal, e.g., another piece of content or topic, when determining
recommendations. Content can include audio content (e.g., music, podcasts, etc.), video content, etc. In making recommendations, contextual factors are not generally determined or accounted for. For example, a recommendation system on a video-hosting site can, with user permission, use a video that the user is currently watching, or has watched in the past, as a signal to determine recommendations. However, the relevance of niche content depends on multiple signals, e.g., multiple interests of the user. Often, niche content is not surfaced when recommendations are made based on broad entities (like genres and keyword metadata). Current recommendation systems are unable to make recommendations of niche content to a correct set of target users.

Due to sub-optimal targeting, recommendations of niche content suffer from low click-through or view-through rates (CTR/VTR). This is because viewers who are uninterested in the niche content are less likely to click through, since the recommended content is of low relevance to such viewers. The low CTR/VTR causes a further decline in the number of recommendations for the niche content, which can lead to a loss of an audience that is likely truly interested in the content. Sub-optimal recommendations consume valuable real estate that can otherwise be used to promote other, more relevant content.

Co-watch signal-based recommendations, that include content that is watched by other users who watched the present content, are also incapable of accurately determining niche-content recommendations. Rather, co-watch recommendation systems generally skew toward lowest-common denominator content.

Consider the below example that illustrates sub-optimal targeting that is a result of using conventional recommendation signals. The example presumes that the user has consented to use of their data for targeted content recommendations. A user is watching a music video by artist
A. There exists content discussing an upcoming collaboration between artist A and artist B. However, it is not known a priori that the user is interested in or aware of artist B. Thus, a recommendation of the content about the upcoming collaboration, when generated by conventional recommendation system, is likely to be deemed irrelevant by the user. Such content is niche content and benefits from being optimally directed at users who are identified as being interested in both artists, e.g., by user of user data, upon consent. If the content is recommended based on the music video of artist A alone, it is possible that the content is recommended too broadly and will have a low CTR. However, if such content is not recommended, users that are interested in both the artists A and B do not see the article and the system end up providing a less useful recommendation instead.

DESCRIPTION

The disclosure addresses the problem of accurate niche-content recommendations by taking into account contextual factors, upon user consent, to utilize such factors when making recommendations. Per techniques of this disclosure, context is established, e.g., based on content that the user has interacted with in the past. Co-watch session data is used to seed a candidate recommendation list. Recommendations are issued based on a good match of multiple signals that arise out of context and co-watch data that the user has consented for use for the purpose of providing content recommendations. For users that do not provide consent to utilize user data to evaluate context and provide recommendations, user data is not utilized, and the technique is not implemented, e.g., niche content is not shown, or is shown only when such content is identified for recommendation without the use of user data.

In the previous example, if the user, having watched a video involving just artist A, had previously watched videos involving artist B, then the content describing the collaboration
between artists A and B is recommended by techniques of this disclosure. A user who previously has not watched content relating to artist B or has otherwise not expressed interest in artist B is not shown the recommendation.

Although the above example illustrates just one match in contextual criteria to offer a recommendation, techniques of disclosure rely on matches in multiple contextual criteria to conclude that a particular recommendation be made. Such contextual criteria include, for example, an element describing other content the user has previously interacted with. Some proportion of the contextual requirements may need to match in order for content to be recommended. A score is determined based on matching context in order to rank recommendations, with the highest-ranking recommendations being shown to the user.

Fig. 1: A content, its related entities and its required content

As shown in Fig. 1, techniques disclosed herein associate an online content (102) with related entities (104) and required context (106). The related entities include, for example, online content that is identified based upon co-watch signals. The required context includes context descriptors that are matched, upon user permission, with user data for a user for content to be suggested to that user. Content that satisfies criteria specified in the related entities and the required context is considered as candidate for recommendation. Although Fig. 1 shows
illustratively two layers 102 and 106 of targeting criteria, it is possible to implement recommendations with any number of targeting criteria.

The required context comprises fields which are filled manually for a specific content to be a candidate for recommendation. Each field comprises online entities such as videos, channels, playlists, music stories, web-pages, shoulder content about artists, text, images, etc. that are associated with the candidate content. In order for content to be recommended, the corresponding required context field is matched, e.g., based on detecting that the user engaged with one or more of the entities in that field. Certain content, e.g., that has only one required context field, is recommended if the user has engaged with any of the entities in that field. Other content that has more than one required context fields. In such cases, the candidate content is recommended when certain conditions are fulfilled, e.g., only if the user has interacted with at least one entity from each of the required context fields. Content curators can specify one or more of the fields. Bitwise logical operators (e.g., OR, AND, etc.) can be applied within a field. Further, a user is deemed to have engaged in an interaction with an entity based on a single user visit to that entity. Alternately, interaction with an entity can be based on measurement of a user affinity score, e.g., a binary measure based on the user affinity score meeting a threshold.

Further, the techniques can be implemented such that both related entities and required contexts include a field that describes the genre of content. Genre ensures that topical content appears as recommendations over a broad range of different watched pages. The other context fields ensure that even as recommended content appears over a relatively broad range of watched pages, such content is suggested specifically to users that are likely to be interested in the recommendation. This manner of capturing a broad range of recommended content and
targeting recommendations towards specific users is especially important for current events. During current events, content recommendation is provided on a large surface area to ensure that the recommendation is seen by as many people as fast as possible. However, recommended content is targeted only at users who are likely to be interested in the content.

Once a piece of content is considered as a candidate for recommendation, it is ranked. Ranking of the content is done by combining the user-to-entity affinity scores for the related entities and the required contexts. This combined score is compared against user-to-content affinity scores for other content in order to produce a rank for the content. In different implementations, the score is normalized or not normalized. For example, not normalizing the score is useful to ensure that niche content recommendations outrank non-niche content recommendations.

Per techniques of this disclosure, niche content associated with particular contexts can be appropriately curated and recommended to users. By restricting the audience to whom the niche content is shown, the CTR and VTR for niche content are not degraded by unaware or uninterested people. This feature is generally valuable, and especially so, for example, for targeting shoulder content, e.g., interviews, news articles, collaborations that support and enhance a music video, etc. Furthermore, techniques disclosed herein enable niche content to be recommended successfully against broad entities such as a music genre, playlist, or keyword metadata. Techniques of this disclosure specifically target niche content at users who are interested in the content, while also broadly recommending content on watch pages of videos tangentially related to the content.

In situations in which certain implementations discussed herein may collect or use personal information about users (e.g., user data, information about a user’s social network,
user's location and time at the location, user's biometric information, user's activities, user’s online history and demographic information), users are provided with one or more opportunities to control whether information is collected, whether the personal information is stored, whether the personal information is used, and how the information is collected about the user, stored and used. That is, the systems and methods discussed herein collect, store and/or use user personal information specifically upon receiving explicit authorization from the relevant users to do so. For example, a user is provided with control over whether programs or features collect user information about that particular user or other users relevant to the program or feature. Each user for which personal information is to be collected is presented with one or more options to allow control over the information collection relevant to that user, to provide permission or authorization as to whether the information is collected and as to which portions of the information are to be collected. For example, users can be provided with one or more such control options over a communication network. In addition, certain data may be treated in one or more ways before it is stored or used so that personally identifiable information is removed. As one example, a user’s identity may be treated so that no personally identifiable information can be determined. As another example, a user’s geographic location may be generalized to a larger region so that the user's particular location cannot be determined.

CONCLUSION

Mechanisms are disclosed herein to curate content dynamically based not only on a content’s relationship with other content, but also the context required to make that content relevant. Per techniques of this disclosure, niche content is recommended specifically to users who are interested in the content, even in broad contexts where the content they are currently
interacting with does not exactly match with the recommended content. Recommendation of niche content to uninterested users is avoided. Techniques described herein result in high CTR and VTR, indicative of a good user experience.