February 2022

# Graph-based Multi-surface Recommender

Anonymous

# Graph-based Multi-surface Recommender


## ABSTRACT

Multi-surface applications are applications where multiple surfaces share common infrastructure and user management, each surface focusing on particular product goals. Casual users are a segment of users who have limited activities on a given surface. Data for casual users is usually sparse and noisy, making it hard to capture the casual user's true interest and to recommend items of interest to them. This disclosure describes a graph-based multi-surface recommender for casual user understanding and transfer, e.g., techniques of user understanding that transfer user interests, both latent and as disclosed by the users, across multiple surfaces. Leveraging centrally managed user accounts, a graph-theoretic understanding of the user is developed based on the diverse activities the user may have on different surfaces, each reflecting the user's interests or preferences. The user's interests are captured and modeled for personalization across surfaces.



## KEYWORDS

- Recommendation system
- Multi-surface application
- User understanding
- Casual user
- Multi-surface recommender
- Interest transferability
- User interest
- Interest profile
- Content labeling
- Cross-surface transfer
- Community detection

**BACKGROUND**

A large volume of information exists on the Internet and in private intranets. Substantial new content such as new articles, photos, videos, or other media are routinely added or updated. Recommendation systems play an important role in retrieving content relevant to a given user. In many cases, the lack of knowledge of a user's interests makes personalized recommendations challenging; this is especially true for new or casual users.

Multi-surface applications are applications where multiple surfaces share common infrastructure and user management, each surface focusing on particular product goals. For example, a video sharing service can offer different surfaces that enable users for specific activities such as sharing short videos, long-form videos, exploratory videos, shopping videos, etc. A user registered with a multi-surface application has access to all the surfaces, although users typically show different levels of activities across the surfaces.

In the context of recommendation systems, user understanding is a procedure that models users' interests to predict which candidate items a given user is likely to choose (personalized recommendation). A user's interests, a characteristic of the user, travel with the user from one surface to another, although with varying manifestations.

Casual users are a segment of users that have limited activity (e.g., low level of engagement) on a given surface. Data for casual users is usually sparse and noisy, making it hard to capture the user's true interest. Casual user recommendation, which is the recommendation of content relevant to casual users, is thus challenging in practice. In new apps or surfaces, a substantial number of consumers can be new or casual users. Although technically challenging, there is substantial business value in converting casual users to steady ones.

## DESCRIPTION

This disclosure describes a graph-based multi-surface recommender for casual user understanding and transfer, e.g., techniques of user understanding that transfer user interests (both latent and as disclosed by the users) across multiple surfaces. Leveraging centrally managed user accounts, a graph theoretic understanding of the user is developed based on the diverse activities the user may have on different surfaces, each reflecting the user's interests and/or preferences. The user's interests are captured and modeled for personalization across surfaces.

Given a source surface and a target surface, a user interest captured on the source surface may or may not be directly transferable as a recommender input to the target surface. For example, a source surface of short-form videos with a user interest of comic videos is not directly transferable to a target surface of documentary videos. Therefore, a transferability score is defined that measures the transferability of a given interest across a source-target pair. Because surfaces can effectively comprise differing products, e.g., short-form versus shopping videos, the respective recommenders on each surface address different business goals while leveraging a common user understanding for sharing personalization. Estimated interests on a source surface are used to predict a set of items the user is likely to prefer on a target surface where the user has limited activities and is a casual user.

The graph-based, multi-surface recommender comprises one or more of the following.

- Input: A multi-surface system, a registered user (and the user's historic engagement activity log), a target surface.
- Output: A set of items recommended to the user on the target surface.

- User interest capturing and modeling: A module that recommends items to users based on the relevance of the items with respect to the individual's interest. Personalization considers the similarity to a user's historic engagement as well as the diversity of recommendation.
- User interest transferability: A module that evaluates and measures the transferability of an interest of a user captured on a source surface to enhance the user's interest profile on a target surface. The user's interest profile refers to a group of interest topics and concepts that describe the user's preference with respect to the items to recommend. For a given user, an interest captured from one surface may not be transferable to another due to the differing experiences of the two surfaces. For example, a captured interest of comic videos on a short-form surface may not be immediately transferable to a surface with a focus on shopping.
- Cross-surface user interest transfer: A module that evaluates if a specified source surface qualifies to capture a given user's interest and to transfer the interest to the target surface.
- Within-surface user interest transfer: In contrast to cross-surface user interest transfer, which captures same-user interest according to same user engagement behaviors, the within-surface user interest transfer module presumes that users within the same group/segments have similar tastes and evaluates the transferability of the group's common interests.
- Blending strategy: A module that blends user interests captured from multiple surfaces to finalize the interest transfer for casual user recommendation.

The above modules are explained in greater detail below.

*System architecture*

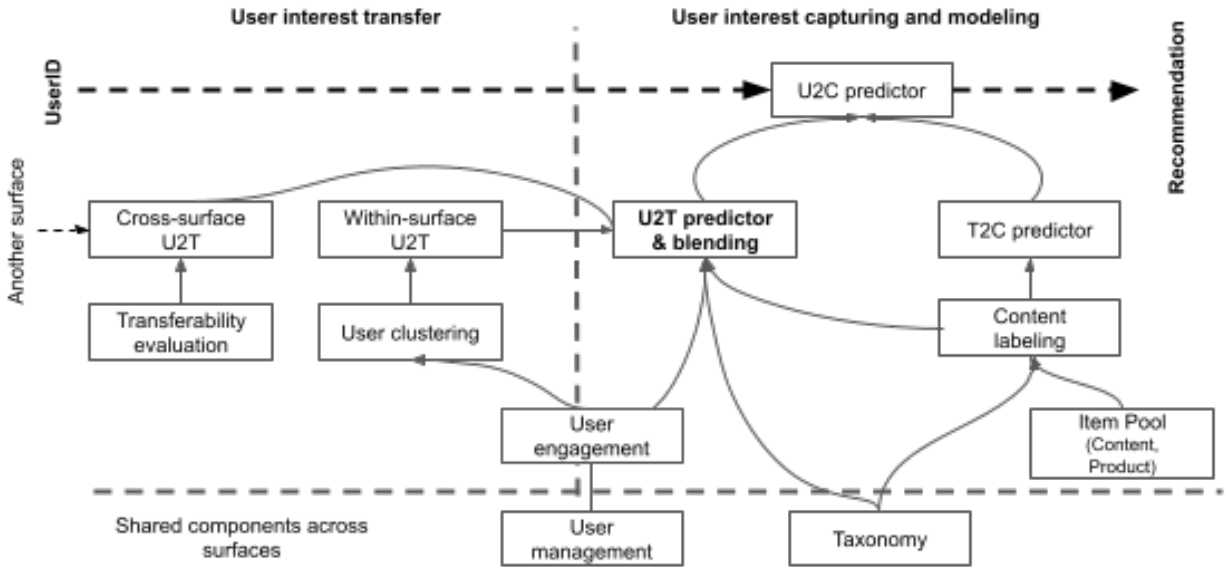


**Fig. 1: Architecture of the multi-surface recommender**

Fig. 1 illustrates an example architecture of a multi-surface recommender. As illustrated, there can be three parts to the architecture, as indicated by dashed lines. The bottom part includes components that are shared across surfaces; the right-hand part includes user-interest capturing and modeling for recommendation on a single surface; and the left-hand part includes mechanisms for transferring user interests either from other surfaces or through local, clustering-based transfer.

Components shared components across surfaces include:

- A user management module shared across multiple surfaces. Components to be shared across surfaces are defined such that channels exist to support user interest transfer.
- Taxonomy defines a set of topics and the likely relationship between topics.

User-interest capturing and modeling for recommendations includes:

- A graph-based recommendation sub-system to capture user interest and to support the recommendation.
- User-to-content (U2C) predictor
    - The goal of such a recommendation on a surface is to identify a subset of content for a given user, referred to as the user-to-content predictor (U2C).
    - The definition of content varies from one surface to another, e.g., a short-form video, a shopping video, a documentary-style video, etc.
    - The functionality is further decomposed into components, e.g., the user-to-topic (U2T) predictor and the topic-to-content (U2C) predictor.
- User-to-topic (U2T) predictor and blender
    - A component representing a user's interest or preference in a topic space specified in the given taxonomy. The prediction is based on the user's recent activities from the user engagement component and the taxonomy topics that describe the content/engagement, e.g., sports, landscape, funny, etc.
    - This component can work as a blender when it also receives U2T prediction from the components of the transferred interests such that a consolidated U2T prediction can be reached.
- Topic-to-content (T2C) predictor
    - For each topic from the taxonomy, an indexer is constructed to link the topic to a set of content in the content pool that has the topic along with a strength (weight). The topics and the corresponding weights are from the content labeling component.
- Content labeling

- For each content in the content pool, the content-labeling component detects a set of topics and the corresponding strength (confidence). Depending on the taxonomy, the topics can be of multiple categories ranging from concepts to emotions.

- User engagement
  - A log that keeps the recent activities of users in terms of the engaged contents from the content pool, e.g., news articles, videos, photos, etc. The type of engagement is also logged, such as views (watches), likes, shares, commenting, etc.
- Item pool
  - This is a set of items (content/products) to be considered. The content pool can be dynamic and substantial in size.

The user-interest transferring mechanism can include the following.

- Cross-surface U2T predictor
  - This component is associated with the U2T component from the user interest capturing and modeling for recommendation on another surface, such that some captured topics of a user can be transferred between the surfaces.
  - Transferability evaluation evaluates if a topic captured from another surface can be transferred to the current surface or not.
- Within-surface U2T predictor
  - This component identifies similar users from the current surface and transfers common interests to the current user.
  - User clustering is a component to identify a community of users on the local surface, assuming homophily exists, e.g., the users enjoy some similar interests.

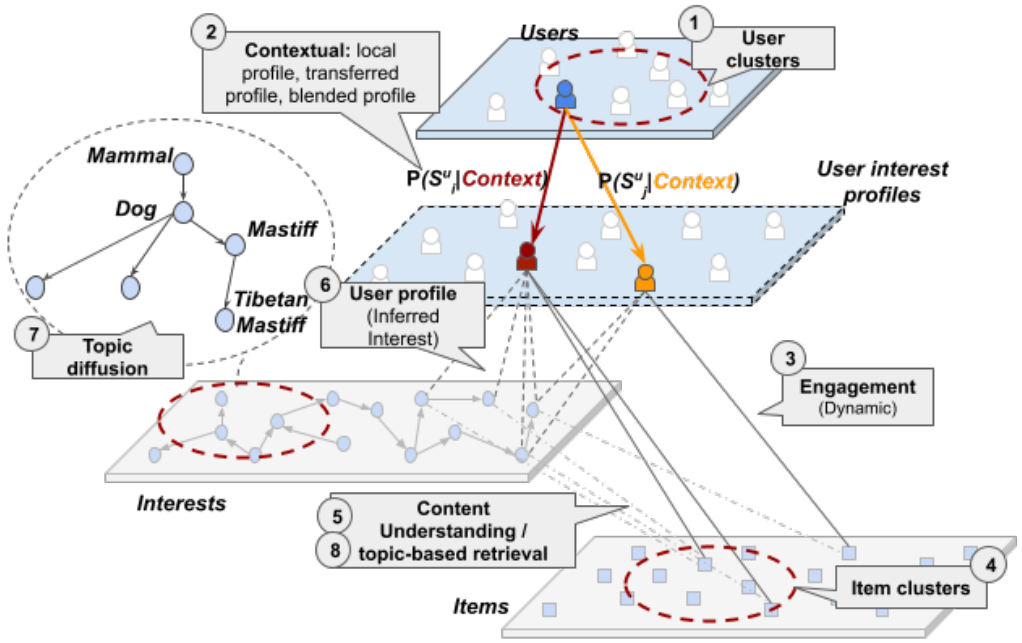


**Fig. 2: Graph-based, user interest capture and modeling**

Fig. 2 illustrates user interest capture and modeling, which, as explained earlier, is graph-based on each surface of the multi-surface recommender. The expressiveness and interpretability of graphs is leveraged to control and transfer captured user interests. In a stage entailing the capture or update of user interests based on topics, the following tasks are executed.

- Identify a user (or user group -- for within-surface transfer).
- Find the user's profile (local, transferred, or blended) that links to interest topics.
- Retrieve the recent engagement of the user.
- Optionally, retrieve similar items for diversity.
- Map the engaged items into interest topics.
- Optionally, diffuse on similar topics for diversity.
- Predict links from user profile to interest topics nodes.

In a stage that predicts a set of contents for the user, the following tasks are identified.

- Identify a user (or user group -- for within-surface transfer).

- Find the user's profile (local, transferred, or blended).
- Select topic(s) linked to the profile.
- Retrieve (new) items linked to the selected topics.

*Transferability*

A user interest captured on a source surface may or may not be transferable as a recommender input on the target surface. For example, for a source surface of short-form videos with a user interest of comic videos, the user interest may not be transferable to a target surface of documentary videos. A transferability score measures the transferability of a user interest captured from one surface to another.

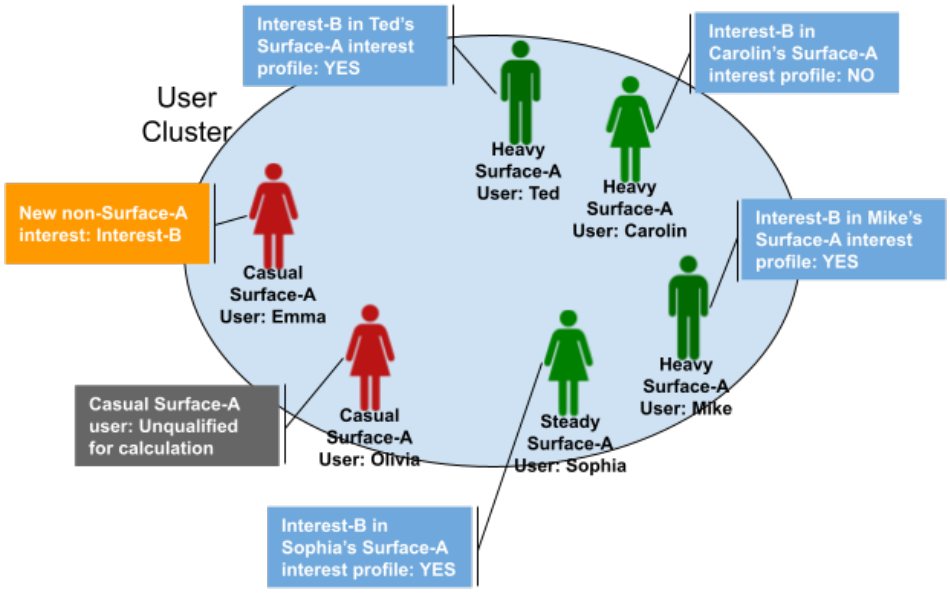


**Fig. 3: Calculating the transferability of a captured interest topic for a user. The green users are heavy or steady users of surface A. The red users are casual users of surface A. The transferability of interest B (in the green users' interest profiles) to the red users is to be evaluated.**

Fig. 3 illustrates an example mechanism of calculating the transferability of a captured interest topic. Within a user cluster, the green users are heavy or steady users of a certain surface A. Within the same user cluster, the red users are casual users of surface A. The transferability of

interest B (in the green users' profiles) to the red users is to be evaluated. The fraction of green users who have interest B in their interest profile is determined, in this example, to be 0.75. The transferability score for casual (red) user Emma is defined to be 0.75.

Since there is generally no direct data from a user's external interests to suggest that a user is likely to engage with a given surface, similar users' interests are leveraged to determine the transferability score. A general procedure to compute the transferability score is as follows. Users are segmented into clusters. Steady and heavy surface A users within a cluster are examined for each pair of surface A user and new interest. The user interest profiles of steady and heavy users are relatively stable compared to casual users. The transferability score is calculated based on the fraction of steady and heavy users in the same cluster that includes the target interest in their surface A interest profiles.

*Cross-surface transfer*

After transferable interests of a given user are identified from a source surface to a target surface, the transfer occurs as a new node being added to the plane of user interest profiles (as illustrated in Fig. 2). The context is parameterized such that a particular profile of a given user is correctly reached.

After the transfer, when recommending items for a user, the user profile is randomly walked through in accordance with the context for fetching related topics. A random walk refers to a movement on the graph where the edge probability is based on the strength of association of a profile with the given context. More complex cases that are inexpressible as graph random walks are discussed in the blending strategy below. Random walk is one example way to implement the cross-surface transfer; however, the described framework supports other ways for recommendation using the transferred profiles.

*Within-surface transfer*

Within-surface transfer is based on the rationale that although the casual user has sparse or noisy interest signals, the community's common profile is likely applicable because users of the same community share common interests (homophilic effect). User clusters being relatively stable, communities can be found using a known community detection algorithm, e.g., the Louvain algorithm. A virtual user profile node is created that links to the common interest topics of the group. Similarly, context is set such that such a profile can be fetched properly on the graph.

*Blending strategy*

Specifying the context enables choosing the local profile or the transferred profile (as illustrated in Fig. 2), since each is a node linked to the user. A more advanced blending strategy for the profiles of a given user can be incorporated, e.g., by selecting representative topics from both profiles, rather than simply choosing a profile. Regardless of the blending strategy, a new profile node is created in the graph (in the profile plane of Fig. 2) and is linked to the user along with the context when such a profile is used.

## CONCLUSION

This disclosure describes a graph-based multi-surface recommender for casual user understanding and transfer, e.g., techniques of user understanding that transfer user interests, both latent and as disclosed by the users, across multiple surfaces. Leveraging centrally managed user accounts, a graph-theoretic understanding of the user is developed based on the diverse activities the user may have on different surfaces, each reflecting the user's interests or preferences. The user's interests are captured and modeled for personalization across surfaces.