Cross-Domain Recommendation: Challenges, Progress, and Prospects

Feng Zhu[1] , **Yan Wang**[2] , **Chaochao Chen**[1*] , **Jun Zhou**[1] , **Longfei Li**[1] , **Guanfeng Liu**[2]

[1] Ant Group, Hangzhou 310012, China

[2] Department of Computing, Macquarie University, Sydney, NSW 2109, Australia

{zhufeng.zhu, chaochao.ccc, jun.zhoujun, longyao.llf}@antgroup.com
{yan.wang, guanfeng.liu}@mq.edu.au

Abstract

To address the long-standing data sparsity problem in recommender systems (RSs), cross-domain recommendation (CDR) has been proposed to leverage the relatively richer information from a richer domain to improve the recommendation performance in a sparser domain. Although CDR has been extensively studied in recent years, there is a lack of a systematic review of the existing CDR approaches. To fill this gap, in this paper, we provide a comprehensive review of existing CDR approaches, including challenges, research progress, and prospects. Specifically, we first summarize existing CDR approaches into four types, including single-target CDR, single-target multi-domain recommendation (MDR), dual-target CDR, and multi-target CDR. We then present the definitions and challenges of these CDR approaches. Next, we propose a full-view categorization and new taxonomies on these approaches and report their research progress in detail. In the end, we share several promising prospects in CDR.

1 Introduction

In the past couple of decades, recommender systems (RSs) have become a popular technique in many web applications, e.g., *Youtube* (video sharing), *Amazon* (e-commerce), and *Facebook* (social networking), as they provide suggestions of items to users so that they can avoid facing the information overload problem [Ricci *et al.*, 2015]. In the existing RSs, collaborative filtering (CF) has been proven to be the most promising technique [He *et al.*, 2017]. The general idea of CF techniques is to recommend items to a user based on the observed preferences of other users whose historical preferences are similar to those of the target user.

Motivation of Cross-Domain Recommendations. In most real-world application scenarios, few users can provide ratings or reviews for many items [Ricci *et al.*, 2015] (i.e., the data sparsity problem), which reduces the recommendation accuracy of CF-based models. Almost all existing CF-based recommender systems suffer from this long-standing

---

*Contact Author

data sparsity problem to some extent, especially for new items or users (i.e., the cold-start problem). This problem may lead to over-fitting when training a CF-based model, which significantly reduces recommendation accuracy. To address the data sparsity problem, *cross-domain recommendation* (CDR) [Berkovsky *et al.*, 2007] has emerged to utilise the relatively richer information, e.g., user/item information [Chung *et al.*, 2007], thumbs-up [Shapira *et al.*, 2013], tags [Fernández-Tobías and Cantador, 2014], reviews [Tan *et al.*, 2014], and observed ratings [Zhu *et al.*, 2018], from the richer (source) domain to improve the recommendation accuracy in the sparser (target) domain. For example, the Douban's[1] RS can recommend books to users according to their movie reviews (i.e., CDR), since a common user in different domains is likely to have similar tastes.

Notion of Domain. In the literature, researchers have attempted to define the notion of domain. However, the concept of domain is still very confusing. For example, some of them treated items (e.g., movies and books) as different domains [Hu *et al.*, 2013], while others considered items in different sub-categories, e.g., such as textbooks and novels, as different domains [Cao *et al.*, 2010]. These definitions mainly focus on the difference of items but the difference of users is ignored. In this survey paper, we define domain from the following three perspectives, i.e., content-level relevance, user-level relevance, and item-level relevance.

- **Content-level relevance**. In the dual/multiple domains, there are the same content or metadata features, e.g., keywords and tags, from user preferences and item details. However, there are not common users/items in these domains, e.g., Amazon music[2] and Netflix[3].
- **User-level relevance**. In the dual/multiple domains, there are common users and different levels of items – such as *attribute-level* (i.e., items in the same type (e.g., book) with different attribute values, e.g., textbooks and novels) and *type-level* (i.e., items in different types, e.g., movies and books).
- **Item-level relevance**. In the dual/multiple domains, there are common items (e.g., movies) and different users, e.g.,

---

[1]Douban URL: https://www.douban.com
[2]Amazon music URL: https://www.amazon.com/music
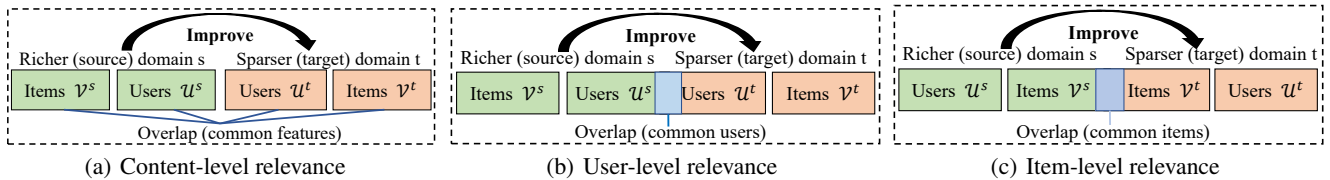[3]Netflix URL: https://www.netflix.com

Figure 1: Single-target CDR scenarios

the users in MovieLens[4] and Netflix systems. These users are totally different or it is difficult to distinguish the overlapped users among different recommender systems. In the literature [Zhu *et al.*, 2018; Zhu *et al.*, 2020], this type of cross-domain recommendation is also referred to as 'cross-system recommendation'.

Classification of Conventional CDR. Conventional CDR approaches can be generally classified into three groups: (1) content-based transfer approaches, (2) embedding-based transfer approaches, and (3) rating pattern-based transfer. *Content-based transfer* mainly handles the CDR problems with content-level relevance and tends to link different domains by identifying similar content information, such as item details [Winoto and Tang, 2008], user-generated reviews [Tan *et al.*, 2014], and social tags [Fernández-Tobías and Cantador, 2014]. In contrast, *embedding-based transfer* [Zhu *et al.*, 2020] mainly handles the CDR problems with user-level relevance and item-level relevance. This class involves first training different CF-based models (such as neural collaborative filtering [He *et al.*, 2017] and graphic models [Zhu *et al.*, 2020; Liu *et al.*, 2020b; Wang *et al.*, 2021]) to obtain user/item embeddings, and then transferring these embeddings through common or similar users/items across domains. Different from embedding-based transfer, *rating pattern-based transfer* tends to transfer an independent knowledge, i.e., rating pattern s, across domains. In contrast to the content-based transfer approaches, embedding-based and rating pattern-based transfer approaches typically employ machine learning techniques, such as multi-task learning [Singh and Gordon, 2008], transfer learning [Zhang *et al.*, 2016], clustering [Farseev *et al.*, 2017], and neural networks [Zhu *et al.*, 2018], to transfer knowledge across domains.

New Trends in CDR. The above conventional CDR approaches are *single-target* approaches that can only leverage the auxiliary information from a richer domain to help a sparser domain. However, each of the domains may be relatively richer in certain types of information (e.g., ratings, reviews, user profiles, item details, and tags); if such information can be leveraged well, it is possible to improve the recommendation performance in all domains simultaneously rather than in a single-target domain only. To this end, *dual-target* CDR [Zhu *et al.*, 2019; Li and Tuzhilin, 2020; Zhu *et al.*, 2020; Liu *et al.*, 2020b] and *multi-target* CDR [Cui *et al.*, 2020] have been proposed recently to improve the recommendation performance in dual/multiple domains.

The Motivation of This Survey. CDR is not a new research area, and in the literature, there are three survey papers

[Fernández-Tobías *et al.*, 2012; Cantador and Cremonesi, 2014; Khan *et al.*, 2017] and a systematical handbook [Ricci *et al.*, 2015] which have carefully introduced and analyzed this area. However, after these tutorials, there are many new challenges, e.g., feature mapping, embedding optimization, and negative transfer, and new directions, e.g., dual-target CDR and multi-target CDR. These new research trends motivate us to analyze the challenges in CDR and summarize the research progress.

Our Contributions. The main contributions of this survey paper are summarized as follows:

- We provide a detailed overview of the challenges in CDR and classify them from a developing perspective, which provides a whole view of the development in CDR area.
- We present a comprehensive review of current research progress in CDR. Specifically, we analyze the contributions of existing approaches and the similarities and differences between them.
- We outline some promising future research directions in CDR, which would shed light on the development of the research community.

2 Problems and Challenges

Cross-domain recommendation problem has been formulated in different recommendation scenarios, i.e., single-target CDR, single-target MDR, dual-target CDR, and multi-target CDR. The main differences among these scenarios are the scales of domains, overlaps, and improvement targets. In this section, we introduce these particular CDR scenarios and their corresponding challenges.

2.1 Single-Target CDR

Single-target CDR is a conventional recommendation scenario in CDR area and most of the existing CDR approaches focus on this scenario. We define this recommendation problem as follows.

Definition 1 *Single-Target Cross-Domain Recommendation: Given the source domain s (including a user set $\mathcal{U}^s$ and an item set $\mathcal{V}^s$) with richer data — such as explicit feedback (e.g., ratings and comments), implicit feedback (e.g., purchase and browsing histories), and side information (e.g., user profiles and item details) — and the target domain t (including a user set $\mathcal{U}^t$ and an item set $\mathcal{V}^t$) with sparser data, single-target CDR is to improve the recommendation accuracy in t by leveraging the auxiliary information in s.*

As introduced in Section 1, we define the notion of domain from three perspectives, i.e., content-level relevance,
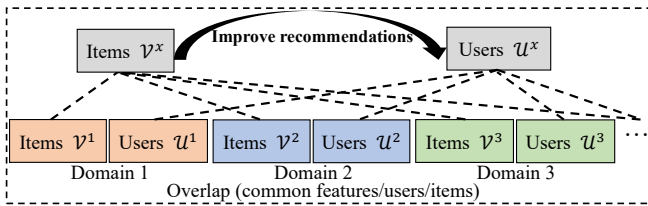
---

[4]MovieLens URL: https://movielens.org

Figure 2: Single-target MDR scenario
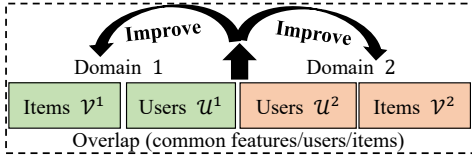

Figure 3: Dual-target CDR scenario

Figure 4: Multi-target CDR scenario

user-level relevance, and item-level relevance. Thus, single-target CDR scenario is divided into three corresponding sub-scenarios as well (see Figure 1). We summarize three main challenges for the single-target CDR scenarios and explain them respectively.

Building content-based relations (CH1). In Figure 1(a), to improve the recommendation accuracy in the target domain, one should first build content-based relations, then choose similar users/items according to their common features, and finally, transfer/share other features between similar users/items across domains. Therefore, how to build a suitable content-based relation in single-target CDR (content-level relevance) scenario is very important and challenging.

Generating accurate user/item embeddings or rating patterns (CH2). In Figures 1(b) and 1(c), to improve the recommendation accuracy in the target domain, one should first generate accurate user/item embeddings or rating patterns, and then transfer/share the embeddings of common users/items or rating patterns of common users across domains. Therefore, how to generate accurate embeddings or rating patterns is a fundamental and crucial challenge.

Learning accurate mapping relations (CH3). For the three scenarios of single-target CDR depicted in Figure 1, a naive transfer strategy is to directly replace the features/embeddings of users/items in the target domain with those of their similar users/items in the source domain [Zhao *et al.*, 2017]. This strategy is simple but not intelligent. An elegant way is to first learn accurate mapping relations between two domains, and then transfer the knowledge (e.g., user/item embeddings and rating patterns) learned from a source domain to a target domain according to the learned mapping relations. Following such an intuition, how to learn accurate mapping relations becomes a crucial challenge.

2.2 Single-Target MDR

Single-target multi-domain recommendation (MDR) is another direction in single-target CDR. It leverages the auxiliary information from multiple domains to recommend a set of items from multiple domains to a certain set of users (single-target) in the multiple domains. We define single-target multiple-domain recommendation as follows.
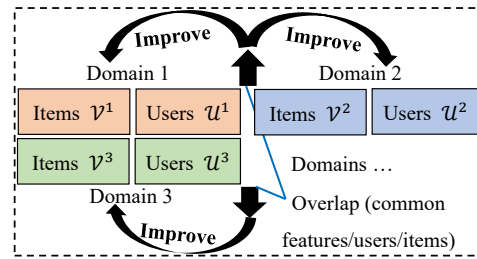
Definition 2 *Single-Target Multi-Domain Recommendation*: *Given the multiple domains* 1 *to* n, *including user sets* $\{\mathcal{U}^1, ..., \mathcal{U}^n\}$ *and item sets* $\{\mathcal{V}^1, ..., \mathcal{V}^n\}$, *single-target MDR is to recommend a set of items* $\mathcal{V}^x$ ($\mathcal{V}^x \in \mathcal{V}^1 \cup ... \cup \mathcal{V}^n$) *to a certain set of users* $\mathcal{U}^x$ ($\mathcal{U}^x \in \mathcal{U}^1 \cup ... \cup \mathcal{U}^n$) *and improve the corresponding recommendation accuracy.*

Single-target MDR faces the same challenges as single-target CDR.

2.3 Dual-Target CDR

Dual-target CDR is a new recommendation scenario in CDR area and it has attracted increasing attention in recent years. We define this recommendation problem as follows.

Definition 3 *Dual-Target Cross-Domain Recommendation*: *Given the two domains* 1 *and* 2, *including user sets* $\mathcal{U}^1$, $\mathcal{U}^2$ *and item sets* $\mathcal{V}^1$, $\mathcal{V}^2$ *respectively, dual-target CDR is to improve the recommendation accuracy in both domains* 1 *and* 2 *simultaneously by leveraging their observed information.*

Similar to the problem of single-target CDR, dual-target CDR scenario can be divided into three sub-scenarios according to the notion of domain. This means that dual-target CDR scenarios can also use common features (content-level relevance), common users (user-level relevance), and common items (item-level relevance), to link the two domains and share/transfer knowledge across domains based on these common entities. However, different from single-target CDR, dual-target CDR is to improve the recommendation accuracy in both target domains simultaneously (see Figure 3). To achieve dual-target CDR, there are two challenges.

Designing a feasible dual-target CDR framework (CH4). Unlike conventional single-target CDR, dual-target CDR should enhance the recommendation performance in the two domains, i.e., the source domain and the target domain. Therefore, how to design an effective framework for a dual-target CDR scenario is still very challenging because the auxiliary information from the target domain may negatively affect the performance in the source domain.

Optimizing the embedding of users and items (CH5). In a dual-target CDR scenario, to improve the recommendation accuracy in each domain, the researchers tend to share the common embeddings of common users/items for the two domains or enhance the embedding quality of users/items in each domain by leveraging the auxiliary information from another domain. Therefore, embedding optimization for dual-target CDR scenarios is particularly important.
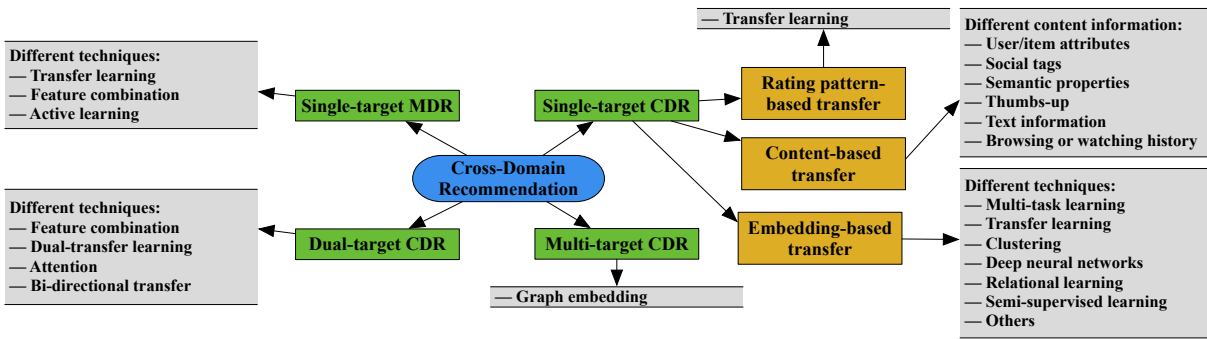
Figure 5: A categorization of CDR approaches

Category		Representative approaches	Technology adoption or basic idea
Content-based transfer	User/item attributes	[Berkovsky *et al.*, 2007]	Multi-source information
		[Leung *et al.*, 2007]	User-item & item-item relationships
	Social tags	[Fernández-Tobías and Cantador, 2014]	Rating-tag similarity
		[Wang and Lv, 2020]	Topic modeling
	Semantic properties	[Kumar *et al.*, 2014]	Semantic similarity
		[Zhang *et al.*, 2019]	Semantic correlation
	Thumbs-up	[Shapira *et al.*, 2013]	Preference similarity
	Text information	[Tang *et al.*, 2012]	Topic modeling
		[Tan *et al.*, 2014]	Topic modelling & transfer learning
	Browsing or watching history	[Elkahky *et al.*, 2015]	Multi-view learning
		[Kanagawa *et al.*, 2019]	Unsupervised domain adaptation

Table 1: The comparison of existing content-based transfer approaches in single-target CDR

2.4 Multi-Target CDR

Inspired by dual-target CDR, in a multi-target CDR scenario, the researchers aim to improve the recommendation accuracy in multiple domains simultaneously. The core idea of multi-target CDR is to leverage more auxiliary information from more domains to achieve a further improvement of recommendation performance. The problem of multi-target CDR can be defined as follows.

Definition 4 *Multi-Target Cross-Domain Recommendation: Given the multiple domains* 1 *to* n, *including user sets* $\{\mathcal{U}^1, ..., \mathcal{U}^n\}$ *and item sets* $\{\mathcal{V}^1, ..., \mathcal{V}^n\}$, *multi-target CDR is to improve the recommendation accuracy in all domains simultaneously by leveraging their observed information.*

To achieve multi-target CDR, in addition to the challenges in single-target CDR and dual-target CDR scenarios, there is a new challenge as follows.

Avoiding negative transfer (CH6). In a multi-target CDR scenario, the recommendation performance in some domains may decline as more domains, especially sparser domains, join in. This is the *negative transfer* problem that the transferred data/knowledge may negatively affect the recommendation performance in the target domain. In fact, in single-target CDR, single-target MDR, and dual-target CDR scenarios, the researchers may also face the negative transfer problem. However, this problem in multi-target CDR scenarios is more serious because the auxiliary information/knowledge in each domain should be transferred to other domains more than once. Therefore, avoiding negative transfer is an important prerequisite in multi-target CDR scenarios.

3 Research Progress

To correspond with the recommendation scenarios and challenges mentioned in Section 2, in this section, we summarize the existing CDR approaches according to their target scenarios, target challenges, data categories, and technical perspectives. We depict the general categorization of CDR approaches in Figure 5. Also, we summarize the popular datasets in CDRs.

3.1 Single-Target CDR

Most of existing single-target CDR approaches tend to leverage useful information from the source domain to the target domain. According to transfer strategies, these single-target CDR approaches are divided in three categories: content-based transfer, embedding-based transfer, and rating pattern-based transfer.

Content-Based Transfer. To target **CH1**, the existing content-based transfer approaches first create links based on the common contents, e.g., user/item attributes, social tags, semantic properties, thumbs-up, text information, metadata, and browsing or watching history. Then, they transfer user/item data or knowledge across domains. We clearly compare the difference of these approaches in Table 1.

Embedding-Based Transfer. To target **CH2** and **CH3**, the existing embedding-based transfer approaches employ some classical machine learning models, e.g., multi-task learning, transfer learning, clustering, deep neural networks, relational learning, and semi-supervised learning, to map or share embeddings, e.g., user/item latent factors, learned by CF-based models across domains. In addition to these frequently-used learning techniques, other embedding-based transfer ap-

Category		Representative approaches	Training data	Technology adoption or basic idea
Embedding -Based Transfer	Multi-task learning	[Singh and Gordon, 2008]	Ratings & item details	Multiple relations
		[Lu *et al.*, 2018]	Ratings	Multi-task learning
	Transfer learning	[Zhao *et al.*, 2017]	Ratings	Active transfer learning
		[Shang *et al.*, 2018] [Huang *et al.*, 2019] [Zhao *et al.*, 2020]	Ratings	Transfer learning
		[Hu *et al.*, 2019]	Ratings & text	Transfer learning & memory networks
	Clustering	[Rafailidis and Crestani, 2016]	Ratings	User clustering
		[Farseev *et al.*, 2017]	Ratings	Clustering
		[Wang *et al.*, 2019b]	Ratings	Clustering
	Deep neural networks	[Zhu *et al.*, 2018]	Ratings	Feature combination and mapping
		[He *et al.*, 2018a]	Ratings	Bayesian neural networks
		[Fu *et al.*, 2019]	Ratings & content	Stacked denoising autoencoders
		[Liu *et al.*, 2020a]	Ratings	Aesthetic preferences
	Relational learning	[Sopchoke *et al.*, 2018]	Ratings	Relational learning
	Semi-supervised learning	[Kang *et al.*, 2019]	Ratings	Semi-supervised learning
	Others	[Li *et al.*, 2011]	Ratings	Topic model & interest drift
		[Hu *et al.*, 2013]	Explicit and implicit feedback	Triadic relation (user-item-domain)
		[Liu *et al.*, 2018]	Ratings	Reinforcement learning
		[Ma *et al.*, 2019]	Item sequences	Sequential recommendations
		[Gao *et al.*, 2019]	Items' information	Data privacy

Table 2: The comparison of existing embedding-based transfer approaches in single-target CDR

proaches tend to employ different techniques or ideas, e.g., Bayesian latent factor models & interest drift, triadic relation (user-item-domain), reinforcement learning, sequential recommendations, and data privacy. We clearly compare the differences among these approaches in Table 2.

Rating Pattern-Based Transfer. To target **CH2**, the existing rating pattern-based transfer approaches tend to first learn an independent rating pattern of users from the source domain and then transfer the rating pattern for the target domain to improve the corresponding recommendation accuracy. The representative work of this type of approach includes [Gao *et al.*, 2013; He *et al.*, 2018b; Yuan *et al.*, 2019].

3.2 Single-Target MDR

Single-target multi-domain recommendation (MDR) is another direction in single-target CDR, but it achieves a different goal: it makes recommendations for different domains. Some of these multi-domain approaches can be applied in CDR scenarios, but they tend to make recommendations either for specific or common users who are selected from domains, or for the users in the target domain only.

Single-taregt MDR also faces the conventional challenges, e.g., **CH1** and **CH2**, in STCDR. To address these challenges, in [Zhang *et al.*, 2012], Zhang et al. proposed a multi-domain collaborative filtering (MCF) framework for solving the data sparsity problem in multiple domains. After this, the single-target MDR models proposed in [Cao *et al.*, 2010; Moreno *et al.*, 2012; Pan and Yang, 2013; Zhang *et al.*, 2016] employ different techniques, i.e., feature combination, transfer learning, and active learning to transfer the knowledge of similar/common users among multiple domains.

3.3 Dual-Target CDR

Dual-target CDR is still a novel but very attractive concept for improving the recommendation accuracies in both do-

mains simultaneously. Therefore, existing solutions are limited but the researchers are paying more and more attention to this direction. To target **CH4** and **CH5**, the existing dual-target CDR approaches mainly focus on either applying fixed or flexible combination strategies [Zhu *et al.*, 2019; Zhu *et al.*, 2020; Liu *et al.*, 2020b], or simply changing the existing single-target transfer learning to become dual-transfer learning [Li and Tuzhilin, 2020].

In [Zhu *et al.*, 2019], Zhu et al. first proposed the DTCDR, a dual-target CDR framework that uses multi-source information such as ratings, reviews, user profiles, item details, and tags to generate more representative embeddings of users and items. Then, based on multi-task learning, the DTCDR framework uses three different combination strategies to combine and share the embedding of common users across domains. Similarly, in [Liu *et al.*, 2020b], Liu et al. also use a fixed combination strategy, i.e., hyper-parameters and data sparsity degrees of common users, to combine the embedding of common users.

Additionally, a new dual-target CDR model (DDTCDR) was proposed in [Li and Tuzhilin, 2020], which considers the bi-directional latent relations between users and items and applies a latent orthogonal mapping to extract user preferences. Based on the orthogonal mapping, DDTCDR can transfer users' embeddings in a bidirectional way (i.e., Source → Target and Target → Source). Recently, Zhu et al. proposed another dual-target CDR framework in [Zhu *et al.*, 2020], which employs graph embedding to generate more informative embeddings of users and items, and employs element-wise attention to combine the embeddings of common users/items.

3.4 Multi-Target CDR

Although multi-target CDR is inspired by single-target multi-domain recommendation and dual-target CDR, it aims to achieve a bigger goal, i.e., providing a complete solution for

Datasets	Domains	Data types	Scale	Website
Arnetminer [Tang *et al.*, 2012]	Research domains (user-level relevance — attribute-level)	Paper & author & conference name ...	1 million	https://www.aminer.org/collaboration
Netflix + MovieLens [Zhu *et al.*, 2018]	Movie (item-level relevance)	Rating & tag	25 million & 100 million	https://www.kaggle.com/netflix-inc/netflix-prize-data https://grouplens.org/datasets/movielens/
Amazon [Fu *et al.*, 2019]	Book & music & movie ... (user-level relevance — type-level)	Rating & review & side information	100 million+	http://jmcauley.ucsd.edu/data/amazon/
Douban [Zhu *et al.*, 2019] [Zhu *et al.*, 2020]	Book & music & movie (user-level relevance — type-level)	Rating & review & side information	1 million+	https://github.com/FengZhu-Joey/GA-DTCDR/tree/main/Data

Table 3: Summary of datasets for CDR

data sparsity. In principle, if the multi-target CDR models can find enough related domains and utilize the auxiliary information from these multiple domains well, the long-standing data sparsity problem in recommender systems can be greatly alleviated and even solved. However, as introduced in Section 2.4, apart from the challenges in single-target CDR and dual-target CDR scenarios, a new challenge, i.e., negative transfer (**CH6**), is inevitable in real multi-target CDR scenarios.

Multi-target CDR is a challenging recommendation scenario, and thus, by now, there are few solutions [Cui *et al.*, 2020; Krishnan *et al.*, 2020] on achieving this goal. In [Cui *et al.*, 2020], the authors use a shared heterogeneous graph to generate more informative embeddings of users and items among multiple domains. Also, the MDCDR approach proposed in [Krishnan *et al.*, 2020] leverages the auxiliary information from a source domain to improve the recommendation accuracy of multiple domains. However, these approaches do not consider the negative transfer problem. Therefore, multi-target CDR is still a challenging task in CDR.

3.5 Summary of Datasets

In this section, we summarize several popularly used datasets for CDR tasks in Table 3. This will guide the researchers to obtain these CDR datasets conveniently. Anyone who wishes to use these datasets can refer to the corresponding citations and websites for more details.

4 Research Prospects

Although many efforts have been put to tackle the challenges of CDR, there remain some promising prospects, and we summarize three of them as follows.

Heterogeneous CDR. Most existing CDR approaches assume information across domains is homogeneous, which is not consistent with reality. For example, some researchers assume both domains have rating and content information [Winoto and Tang, 2008], while other studies assume the existence of rating data across both domains [Zhao *et al.*, 2017]. However, in practice, different domains are rich in different kinds of information. For instance, an e-commerce domain (e.g., Amazon) is rich in user-item interaction data while a social domain (e.g., Facebook) has plenty of user-user social data. Under such situations, new techniques should be proposed to identify the 'bridges' across domains so as to transfer information and improve the performance of CDR. How

to leverage these heterogeneous data across domains, to further improve the recommendation performance, becomes the first promising prospect in CDR.

Sequential CDR. Sequential recommendation has gained much attention since it can suggest items to users by modeling the sequential dependencies over the user-item interactions [Wang *et al.*, 2019a]. Naturally, CDR also faces the problem of sequentially modeling of users and items, the same as conventional recommender systems. Prior work on sequential recommendation mainly focuses on learning the high-order, long-term, and noisy user-item interactions in sequence. It becomes more challenging for sequential CDR since one not only needs to model sequential user-item interactions, but also transfer information across domains [Ma *et al.*, 2019]. Therefore, sequential CDR becomes the second promising research prospect.

Privacy-Preserving CDR. Most existing approaches in CDR assume that data across domains are available in plaintext, which ignores the data isolation problem in practice. Apparently, most recommender systems are built using users' sensitive data, e.g., check-in data, user profile, and browse history. And in CDR, these data are usually held by different domains, e.g., Amazon and eBay. In some cases, these data across domain cannot share with other directly since they contain sensitive information. Thus, it is urgent to build CDR meanwhile protect data privacy [Gao *et al.*, 2019]. Therefore, privacy-preserving CDR is the third pressing and promising research prospect.

5 Conclusion

Cross-domain recommendations (CDRs) have attracted increasing research attention, with the development of deep neural network and graph learning techniques. This paper has conducted a comprehensive survey on the following four scopes, i.e., single-target CDR, single-target MDR, dual-target CDR, and multi-target CDR. We first have presented the definitions and challenges of these scopes, then proposed a full-view categorization and new taxonomies on these scopes, and finally recommended several promising prospects in CDR. This survey summarizes the current representative research efforts and trends and we expect it can facilitate future research in the community.

References

[Berkovsky *et al.*, 2007] Shlomo Berkovsky, Tsvi Kuflik, and Francesco Ricci. Cross-domain mediation in collaborative filtering. In *International Conference on User Modeling*, pages 355–359. Springer, 2007.

[Cantador and Cremonesi, 2014] Iván Cantador and Paolo Cremonesi. Tutorial on cross-domain recommender systems. In *RecSys*, pages 401–402, 2014.

[Cao *et al.*, 2010] Bin Cao, Nathan Nan Liu, and Qiang Yang. Transfer learning for collective link prediction in multiple heterogenous domains. In *ICML*. Citeseer, 2010.

[Chung *et al.*, 2007] Ronald Chung, David Sundaram, and Ananth Srinivasan. Integrated personal recommender systems. In *ICEC*, pages 65–74, 2007.

[Cui *et al.*, 2020] Qiang Cui, Tao Wei, Yafeng Zhang, and Qing Zhang. Herograph: A heterogeneous graph framework for multi-target cross-domain recommendation. In *ORSUM@RecSys*, 2020.

[Elkahky *et al.*, 2015] Ali Mamdouh Elkahky, Yang Song, and Xiaodong He. A multi-view deep learning approach for cross domain user modeling in recommendation systems. In *WWW*, pages 278–288, 2015.

[Farseev *et al.*, 2017] Aleksandr Farseev, Ivan Samborskii, Andrey Filchenkov, and Tat-Seng Chua. Cross-domain recommendation via clustering on multi-layer graphs. In *SIGIR*, pages 195–204, 2017.

[Fernández-Tobías and Cantador, 2014] Ignacio Fernández-Tobías and Iván Cantador. Exploiting social tags in matrix factorization models for cross-domain collaborative filtering. In *RecSys*, pages 34–41, 2014.

[Fernández-Tobías *et al.*, 2012] Ignacio Fernández-Tobías, Iván Cantador, Marius Kaminskas, and Francesco Ricci. Cross-domain recommender systems: A survey of the state of the art. In *CERI*, page 24, 2012.

[Fu *et al.*, 2019] Wenjing Fu, Zhaohui Peng, Senzhang Wang, Yang Xu, and Jin Li. Deeply fusing reviews and contents for cold start users in cross-domain recommendation systems. In *AAAI*, volume 33, pages 94–101, 2019.

[Gao *et al.*, 2013] Sheng Gao, Hao Luo, Da Chen, Shantao Li, Patrick Gallinari, and Jun Guo. Cross-domain recommendation via cluster-level latent factor model. In *ECML-PKDD*, pages 161–176. Springer, 2013.

[Gao *et al.*, 2019] Chen Gao, Xiangning Chen, Fuli Feng, Kai Zhao, Xiangnan He, Yong Li, and Depeng Jin. Cross-domain recommendation without sharing user-relevant data. In *WWW*, pages 491–502, 2019.

[He *et al.*, 2017] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. Neural collaborative filtering. In *WWW*, pages 173–182, 2017.

[He *et al.*, 2018a] Jia He, Rui Liu, Fuzhen Zhuang, Fen Lin, Cheng Niu, and Qing He. A general cross-domain recommendation framework via bayesian neural network. In *ICDM*, pages 1001–1006. IEEE, 2018.

[He *et al.*, 2018b] Ming He, Jiuling Zhang, Peng Yang, and Kaisheng Yao. Robust transfer learning for cross-domain collaborative filtering using multiple rating patterns approximation. In *WSDM*, pages 225–233, 2018.

[Hu *et al.*, 2013] Liang Hu, Jian Cao, Guandong Xu, Longbing Cao, Zhiping Gu, and Can Zhu. Personalized recommendation via cross-domain triadic factorization. In *WWW*, pages 595–606, 2013.

[Hu *et al.*, 2019] Guangneng Hu, Yu Zhang, and Qiang Yang. Transfer meets hybrid: A synthetic approach for cross-domain collaborative filtering with text. In *WWW*, pages 2822–2829, 2019.

[Huang *et al.*, 2019] Ling Huang, Zhi-Lin Zhao, Chang-Dong Wang, Dong Huang, and Hong-Yang Chao. Lscd: Low-rank and sparse cross-domain recommendation. *Neurocomputing*, 366:86–96, 2019.

[Kanagawa *et al.*, 2019] Heishiro Kanagawa, Hayato Kobayashi, Nobuyuki Shimizu, Yukihiro Tagami, and Taiji Suzuki. Cross-domain recommendation via deep domain adaptation. In *ECIR*, pages 20–29, 2019.

[Kang *et al.*, 2019] SeongKu Kang, Junyoung Hwang, Dongha Lee, and Hwanjo Yu. Semi-supervised learning for cross-domain recommendation to cold-start users. In *CIKM*, pages 1563–1572, 2019.

[Khan *et al.*, 2017] Muhammad Murad Khan, Roliana Ibrahim, and Imran Ghani. Cross domain recommender systems: a systematic literature review. *ACM Computing Surveys (CSUR)*, 50(3):1–34, 2017.

[Krishnan *et al.*, 2020] Adit Krishnan, Mahashweta Das, Mangesh Bendre, Hao Yang, and Hari Sundaram. Transfer learning via contextual invariants for one-to-many cross-domain recommendation. In *SIGIR*, pages 1081–1090, 2020.

[Kumar *et al.*, 2014] Anil Kumar, Nitesh Kumar, Muzammil Hussain, Santanu Chaudhury, and Sumeet Agarwal. Semantic clustering-based cross-domain recommendation. In *CIDM*, pages 137–141. IEEE, 2014.

[Leung *et al.*, 2007] Cane Wing-ki Leung, Stephen Chi-fai Chan, and Fu-lai Chung. Applying cross-level association rule mining to cold-start recommendations. In *WI-IAT*, pages 133–136. IEEE, 2007.

[Li and Tuzhilin, 2020] Pan Li and Alexander Tuzhilin. Ddtcdr: Deep dual transfer cross domain recommendation. In *WSDM*, pages 331–339, 2020.

[Li *et al.*, 2011] Bin Li, Xingquan Zhu, Ruijiang Li, Chengqi Zhang, Xiangyang Xue, and Xindong Wu. Cross-domain collaborative filtering over time. In *IJCAI*, pages 2293–2298, 2011.

[Liu *et al.*, 2018] Bo Liu, Ying Wei, Yu Zhang, Zhixian Yan, and Qiang Yang. Transferable contextual bandit for cross-domain recommendation. In *AAAI*, pages 3619–3626, 2018.

[Liu *et al.*, 2020a] Jian Liu, Pengpeng Zhao, Fuzhen Zhuang, Yanchi Liu, Victor S Sheng, Jiajie Xu, Xiaofang

Zhou, and Hui Xiong. Exploiting aesthetic preference in deep cross networks for cross-domain recommendation. In *TheWebConf*, pages 2768–2774, 2020.

[Liu *et al.*, 2020b] Meng Liu, Jianjun Li, Guohui Li, and Peng Pan. Cross domain recommendation via bi-directional transfer graph collaborative filtering networks. In *CIKM*, pages 885–894, 2020.

[Lu *et al.*, 2018] Yichao Lu, Ruihai Dong, and Barry Smyth. Why i like it: multi-task learning for recommendation and explanation. In *RecSys*, pages 4–12, 2018.

[Ma *et al.*, 2019] Muyang Ma, Pengjie Ren, Yujie Lin, Zhumin Chen, Jun Ma, and Maarten de Rijke. π-net: A parallel information-sharing network for shared-account cross-domain sequential recommendations. In *SIGIR*, pages 685–694, 2019.

[Moreno *et al.*, 2012] Orly Moreno, Bracha Shapira, Lior Rokach, and Guy Shani. Talmud: transfer learning for multiple domains. In *CIKM*, pages 425–434, 2012.

[Pan and Yang, 2013] Weike Pan and Qiang Yang. Transfer learning in heterogeneous collaborative filtering domains. *Artificial intelligence*, 197:39–55, 2013.

[Rafailidis and Crestani, 2016] Dimitrios Rafailidis and Fabio Crestani. Top-n recommendation via joint cross-domain user clustering and similarity learning. In *ECML-PKDD*, pages 426–441. Springer, 2016.

[Ricci *et al.*, 2015] Francesco Ricci, Lior Rokach, and Bracha Shapira. *Recommender Systems Handbook*. Springer, 2nd edition, 2015.

[Shang *et al.*, 2018] Jin Shang, Mingxuan Sun, and Kevyn Collins-Thompson. Demographic inference via knowledge transfer in cross-domain recommender systems. In *ICDM*, pages 1218–1223. IEEE, 2018.

[Shapira *et al.*, 2013] Bracha Shapira, Lior Rokach, and Shirley Freilikhman. Facebook single and cross domain data for recommendation systems. *User Modeling and User-Adapted Interaction*, 23(2-3):211–247, 2013.

[Singh and Gordon, 2008] Ajit P Singh and Geoffrey J Gordon. Relational learning via collective matrix factorization. In *SIGKDD*, pages 650–658, 2008.

[Sopchoke *et al.*, 2018] Sirawit Sopchoke, Ken-ichi Fukui, and Masayuki Numao. Explainable cross-domain recommendations through relational learning. In *AAAI*, 2018.

[Tan *et al.*, 2014] Shulong Tan, Jiajun Bu, Xuzhen Qin, Chun Chen, and Deng Cai. Cross domain recommendation based on multi-type media fusion. *Neurocomputing*, 127:124–134, 2014.

[Tang *et al.*, 2012] Jie Tang, Sen Wu, Jimeng Sun, and Hang Su. Cross-domain collaboration recommendation. In *SIGKDD*, pages 1285–1293, 2012.

[Wang and Lv, 2020] Jiaqi Wang and Jing Lv. Tag-informed collaborative topic modeling for cross domain recommendations. *Knowledge-Based Systems*, 203:106–119, 2020.

[Wang *et al.*, 2019a] Shoujin Wang, Liang Hu, Yan Wang, Longbing Cao, Quan Z Sheng, and Mehmet Orgun. Sequential recommender systems: challenges, progress and prospects. In *IJCAI*, pages 6332–6338, 2019.

[Wang *et al.*, 2019b] Yaqing Wang, Chunyan Feng, Caili Guo, Yunfei Chu, and Jenq-Neng Hwang. Solving the sparsity problem in recommendations via cross-domain item embedding based on co-clustering. In *WSDM*, pages 717–725, 2019.

[Wang *et al.*, 2021] Shoujin Wang, Liang Hu, Yan Wang, Xiangnan He, Quan Z. Sheng, Mehmet Orgun, Longbing Cao, Francesco Ricci, and Philip S. Yu. Graph learning based recommender systems: a review. In *IJCAI*, pages 1–9, 2021.

[Winoto and Tang, 2008] Pinata Winoto and Tiffany Tang. If you like the devil wears prada the book, will you also enjoy the devil wears prada the movie? a study of cross-domain recommendations. *New Generation Computing*, 26(3):209–225, 2008.

[Yuan *et al.*, 2019] Feng Yuan, Lina Yao, and Boualem Benatallah. Darec: Deep domain adaptation for cross-domain recommendation via transferring rating patterns. In *IJCAI*, pages 4227–4233, 2019.

[Zhang *et al.*, 2012] Yu Zhang, Bin Cao, and Dit-Yan Yeung. Multi-domain collaborative filtering. *arXiv preprint arXiv:1203.3535*, 2012.

[Zhang *et al.*, 2016] Zihan Zhang, Xiaoming Jin, Lianghao Li, Guiguang Ding, and Qiang Yang. Multi-domain active learning for recommendation. In *AAAI*, pages 2358–2364, 2016.

[Zhang *et al.*, 2019] Qian Zhang, Peng Hao, Jie Lu, and Guangquan Zhang. Cross-domain recommendation with semantic correlation in tagging systems. In *IJCNN*, pages 1–8. IEEE, 2019.

[Zhao *et al.*, 2017] Lili Zhao, Sinno Jialin Pan, and Qiang Yang. A unified framework of active transfer learning for cross-system recommendation. *Artificial Intelligence*, 245:38–55, 2017.

[Zhao *et al.*, 2020] Cheng Zhao, Chenliang Li, Rong Xiao, Hongbo Deng, and Aixin Sun. Catn: Cross-domain recommendation for cold-start users via aspect transfer network. In *SIGIR*, pages 229–238, 2020.

[Zhu *et al.*, 2018] Feng Zhu, Yan Wang, Chaochao Chen, Guanfeng Liu, Mehmet A Orgun, and Jia Wu. A deep framework for cross-domain and cross-system recommendations. In *IJCAI*, pages 3711–3717, 2018.

[Zhu *et al.*, 2019] Feng Zhu, Chaochao Chen, Yan Wang, Guanfeng Liu, and Xiaolin Zheng. Dtcdr: A framework for dual-target cross-domain recommendation. In *CIKM*, pages 1533–1542. ACM, 2019.

[Zhu *et al.*, 2020] Feng Zhu, Yan Wang, Chaochao Chen, Guanfeng Liu, and Xiaolin Zheng. Graphical and attentional framework for dual-target cross-domain recommendation. In *IJCAI*, pages 3001–3008, 2020.