Noname manuscript No.
(will be inserted by the editor)

Multistakeholder Recommendation: Survey and Research Directions*

Himan Abdollahpouri[1] · Gediminas Adomavicius[2] · Robin Burke[1] · Ido Guy[3] · Dietmar Jannach[4] · Toshihiro Kamishima[5] · Jan Krasnodebski[6] · Luiz Pizzato[7]

the date of receipt and acceptance should be inserted later

Abstract Recommender systems provide personalized information access to users of Internet services from social networks to e-commerce to media and entertainment. As is appropriate for research in a field with a focus on personalization, academic studies of recommender systems have largely concentrated on optimizing for user experience when designing, implementing and evaluating their algorithms and systems. However, this concentration on the user has meant that the field has lacked a systematic exploration of other aspects of recommender system outcomes. A user-centric approach limits the ability to incorporate system objectives such as fairness, balance, and profitability, and obscures concerns that might come from other stakeholders, such as the providers or sellers of items being recommended. Multistakeholder recommendation has emerged as a unifying framework for describing and understanding recommendation settings where the end user is not the sole focus. This article outlines the multistakeholder perspective on recommendation, highlighting example research areas, and discussing important issues, open questions, and prospective research directions.

1 Introduction

Recommender systems provide personalized information access, supporting e-commerce, social media, news, and other applications where the volume of content would otherwise be overwhelming. They have become indispensable

* This paper or a similar version is not currently under review by a journal or conference. This paper is void of plagiarism or self-plagiarism as defined by the Committee on Publication Ethics and Springer Guidelines.

[1]University of Colorado, Boulder · [2]University of Minnesota · [3]eBay, Inc. · [4]University of Klagenfurt · [5]National Institute of Advanced Industrial Science and Technology, Japan · [6]Expedia Group · [7]Accenture

features of the Internet age, found in systems of many kinds. One of the defining characteristics of recommender systems is personalization. In research contexts, recommender systems are typically evaluated on their ability to provide items that satisfy the needs and interests of the end user. Such focus is entirely appropriate. Users would not make use of recommender systems if they believed such systems were not providing items that matched their interests. Still, it is also clear that, in many recommendation domains, the user for whom recommendations are generated is not the only stakeholder in the recommendation outcome. Other users, the providers of products, and even the system's own objectives may need to be considered when these perspectives differ from those of end users.

In many practical settings, such as e-commerce, recommendation is viewed as a form of marketing and, as such, the economic considerations of the retailer will also enter into the recommendation function (Leavitt 2006; Pathak et al. 2010). A business may wish to highlight products that are more profitable or that are currently on sale, for example. More recently, system-level objectives such as fairness and balance have been considered by researchers, and these social-welfare-oriented goals may at times run counter to individual preferences. Sole focus on the end user hampers researchers' ability to incorporate such objectives into recommendation algorithms and system designs.

We believe that, far from being special "edge cases", these examples illustrate a more general point about recommendation, namely, that recommender systems often serve multiple goals and that the purely user-centered approach found in most academic research does not allow all such goals to enter into their design and evaluation. What is needed is an inclusive approach that expands outward from the user to include the perspectives and utilities of multiple stakeholders.

It is relevant to note a shift that occurred in microeconomics in the early part of the 21st century with the development of the theory of multisided platforms (Rochet and Tirole 2003; Evans et al. 2011). Prior to that time, the traditional business model focused on a firm's ability to produce products and deliver them to customers at a price that could ensure profitability. By contrast, multisided platforms create value by bringing buyers and sellers together, reducing search and transaction costs. Many online systems are exactly such multisided platforms (Evans and Schmalensee 2016).

As noted above, when it comes to the study of personalized information access in the form of recommender systems, academic research has, with few exceptions, examined only a single side of these interactions. The stage was set historically by the first recommender systems implementations, which either operated on objects with no associated price (newsgroup posts (Konstan et al. 1997)) or were external to any commerce associated with their recommendations (such as non-commercial music, movie, and restaurant recommenders (Shardanand and Maes 1995; Breese et al. 1998; Burke et al. 1997)). These systems brought users and products together, but they were not themselves party to any transactions. While academic research has largely concentrated on the user, commercial systems have regularly taken a broader view of recommendation objectives (Rodriguez et al. 2012a; Nguyen et al. 2017). There is, therefore, a gap between the complexity of real-world applications of recommender systems and those on which academic research has focused.

The integration of the perspectives of multiple parties into the design of recommender systems is the goal underlying the sub-field of *multistakeholder recommendation* (Abdollahpouri et al. 2017b; Burke et al. 2016; Nguyen et al. 2017). This article is intended to describe the current state of the art in multistakeholder recommendation research, to show some examples of current work in the area, and to outline research questions that should be addressed to support the demands of recommendation applications in environments where the perspectives of multiple parties are important.

2 Multistakeholder Recommendation

The concept of a stakeholder appears in business management literature as a way to discuss the complexities of corporate governance. According to Goodpaster (1991),

> the term 'stakeholder' appears to have been invented in the early '60s as a deliberate play on the word 'stockholder' to signify that there are other parties having a 'stake' in the decision-making of the modern, publicly-held corporation in addition to those holding equity positions.

In his classic work, *Strategic Management: A Stakeholder Perspective*, Freeman extends older definitions that emphasize a "stake" as a kind of investment, and instead defines stakeholders as "any groups or individuals that can affect, or are affected by, the firm's objectives" (Freeman 2010), pg. 25. We adopt this definition for our aims, focusing specifically on recommender systems.

Definition 1 A *recommendation stakeholder* is any group or individual that can affect, or is affected by, the delivery of recommendations to users.

As recommender systems are elements of an organization's operations, they will necessarily inherit the large and wide-ranging set of stakeholders considered in the management literature. However, only some of these stakeholders will be particularly salient in the generation of recommendations. In this work, we will consider three key groups of stakeholders who are particularly close to the recommendation interaction:

Consumers: The consumers are the end users who receive / consume recommendations. They are the individuals whose choice or search problems bring them to the platform, and who expect recommendations to satisfy those needs.

Providers: The (item) providers are those entities that supply or otherwise stand behind the recommended objects.

System: The final category is the organization itself, which has created a platform and associated recommender system in order to match consumers with providers. The platform may be a retailer, e-commerce site, broker, or other venue where consumers seek recommendations.

Of course, none of these are stakeholder groups are unitary entities — not even the system, which stands in for various internal groups within an organization who may have different, possibly competing demands on a recommender system. Differentiation among stakeholder groups may be necessary, depending on the application. This taxonomy does, however, highlight an important consequence of a multistakeholder perspective, namely the foregrounding of the multisided nature of recommendation, which has been slow to emerge in the research literature.

The multistakeholder perspective on recommendation can manifest itself in various aspects of recommender systems research and development. We may adopt a multistakeholder view of evaluation: asking the question of how different groups of stakeholders are impacted by the recommendations a system computes. A developer may employ the multistakeholder perspective in designing and optimizing an algorithm, incorporating the objectives of multiple parties and balancing among them. Finally, the creation of the recommender itself may be a multistakeholder endeavor, in which different stakeholders participate in developing a design.

2.1 Multistakeholder Evaluation

In research settings, recommender systems traditionally are evaluated from the users' perspective. Metrics such as precision, recall, NDCG, diversity, novelty, etc. all capture different aspects of recommendation quality as experienced by the end user. In on-line testing, click-through rate, dwell time, and other interaction metrics capture similar types of outcomes. These metrics are typically averaged over all users to produce a single outcome. This methodology in entirely reasonable and logical as, in the end, users are one of the most important stakeholders of any recommender system. We will take it as a given that any recommender system evaluation will include an outcome of this type, the *user summary* evaluation.

A multistakeholder perspective on evaluation, however, brings to light additional aspects of system performance that may be quite important. As mentioned above, multisided platforms such as eBay, Etsy, or AirBnB, have a key business requirement of attracting and satisfying the needs of providers as well as users. A system of this type will need to evaluate its performance from the provider perspective.

Even when only a single group of stakeholders is under consideration, a methodology that relies on a single point estimate of system performance under some metric may miss differences among stakeholder groups, even among users, who are typically the target of evaluation. Stakeholder theory recognizes that subgroups within the stakeholder categories may experience a range of different impacts from a firm's decisions (Freeman 2010). In recommendation, recent work has shown that, depending on the algorithm, male and female users may experience different outcomes in movie recommendations (Ekstrand et al. 2018a). Therefore, a multistakeholder evaluation is one that includes different stakeholder groups in its assessment, in addition to the user summary evaluation.

Definition 2 A *multistakeholder evaluation* is one in which the quality of recommendations is assessed across multiple groups of stakeholders, in addition to a point estimate over the full user population.

A multistakeholder evaluation may entail the use of different kinds of metrics and different evaluation methodologies than typically encountered in recommender systems research. For example, typical cross-validation methodologies that distribute user profiles between test and training sets may not yield reliable results when assessing outcome for other stakeholders, especially providers. We will take up the issue of provider-side metrics in Section 5.5.1.

2.2 Multistakeholder Algorithms

In implementing a given recommender system, a developer may choose to use the metrics associated with multistakeholder evaluation in algorithm design, implementation, and optimization. In general, this will entail balancing the objectives of different stakeholders, as it is unlikely that the optimal solution for one will be the best for all. Some solutions may combine all such objectives into a single optimization, a challenging approach given the methodological complexities noted above; others use a multi-stage approach incorporating different stakeholder concerns throughout a pipeline. Both approaches are discussed in the examples below.

Multistakeholder algorithms are particularly differentiated from typical recommendation approaches when the stakeholder concerns lie on different sides of the recommendation platform. For example, it is not difficult to change the loss function associated with a factorization algorithm to prefer balanced outcomes over multiple subgroups of users, rather than a simple overall mean. However, system or provider objectives may be orthogonal to user concerns and form a separate optimization problem that cannot be simply combined with the users'. We therefore define a multistakeholder recommendation algorithm with special attention to this subclass.

Definition 3 A *multistakeholder recommendation algorithm* takes into account the preferences of multiple parties when generating recommendations, especially when these parties are on different sides of the recommendation interaction.

2.3 Multistakeholder Design

Beyond implementing metrics and tuning algorithms, any fielded recommender system will also go through phases of design in which the system's role within a particular platform and its specific requirements are formulated. System designers may choose to engage in a design process such as participatory design (Kensing and Blomberg 1998) in which external stakeholders are incorporated into decision making. Although these techniques are well-established in the HCI community, they have not seen much discussion in recommender systems research. Anecdotal information suggests that commercial platforms with multiple stakeholders do engage in stakeholder consultation with item providers in particular (Semerci et al. 2019).

Definition 4 A *multistakeholder design process* is one in which different recommendation stakeholder groups are identified and consulted in the process of system design.

2.4 The Landscape of Multistakeholder Recommendation

The above list of the key recommendation stakeholders provides an outline for understanding and surveying the different types of multistakeholder recommendation. We can conceptualize a recommender system as designed to maximize some combined utility associated with the different stakeholders, and consider how different types of applications yield different stakegholder concerns.

2.4.1 Group recommendation

Group recommendation has a substantial history within recommender systems research as surveyed in (Masthoff 2011). In group recommendation, recommendation results are delivered to groups of users who are assumed to be experiencing them jointly. For example, a group of individuals traveling together will get a recommendation for a single trip, which must represent some balance between their individual interests (Garcia et al. 2011). Typically social choice mechanisms such as least misery, Borda count, or other preference aggregation methods are employed.

Group recommendation represents a case where each given recommendation is designed to satisfy a distinct and disjoint group of users. It can be effectively understood (and is often implemented) by considering each group to be a single super-user. Constraints and tradeoffs between individual users' preferences have ramifications only within their particular group, and do not need to be represented or reasoned about relative to the entire user population.

Group recommendation thus represents an atomized, local, type of multistakeholder recommendation, where each individual is considered as a solitary stakeholder and the impact of multiplicity is confined to the user groups receiving recommendations. For the purposes of evaluation, we may use standard evaluation metrics to see if individuals are experiencing utility loss due to their inclusion with a particular group. As such, group recommendation is quite different from other multistakeholder configurations in that it only arises when a recommendation application is such that recommended items must be experienced jointly. Other types of multistakeholder recommendation apply in the more typical case where recommendations are targeted to individual consumers and yet, multistakeholder issues still arise.

2.4.2 Consumer-side issues

If we concentrate only on the individuals consuming recommendations, multistakeholder issues arise when there are tradeoffs or disparities between groups of users in the provision of recommendations. For example, Ekstrand et al. (2018a) explored the performance of recommender system algorithms on users belonging to different demographic groups (gender, age) and observed that some algorithms perform better for certain groups than others. Other researchers have found that users with niche or unusual tastes can be poorly served by particular recommendation algorithms, lapses that are not detectable from point estimates of accuracy measures (Abdollahpouri and Burke 2019; Ghazanfar and Prügel-Bennett 2014).

In some settings, such differences in system performance for different users may be considered examples of unfair treatment. Section 4.3 examines algorithmic remedies for enhancing consumer-side fairness.

2.4.3 Provider-side issues

As noted above, multisided platforms need to satisfy both the consumers of recommendations and the providers of items that are being recommended. The health of such a platform depends both on a user community and a catalog of items of interest. Providers whose items are not recommended may experience poor engagement from users and lose interest in participating in a given platform. Platforms that facilitate peer-to-peer interactions, such as the online craft marketplace Etsy, may be particularly sensitive to the need to attract and retain sellers.

Depending on the application, providers may have particular audiences in mind as appropriate targets for their items. A well-known example is computational advertising in which advertisers seek particular target markets and ad platforms try to personalize ad presentation to match user interests (Yuan et al. 2012). In this application, market forces, expressed through auction mechanisms, serve to mediate between provider interests. In other cases, such as online dating, preferences may be expressed on both sides of the interaction but it is up to the recommender system to perform appropriate allocation. See Section 4.1 below.

2.4.4 System/platform issues

In many real-world contexts, the system may gain some utility when recommending items, and is therefore a stakeholder in its own right. Presumably, an organization creates and operates a recommender system in order to fulfill some business function and generally that is to enhance user experience, lower search costs, increase convenience, etc. This would suggest a consumer stakeholder point of view is sufficient.

However, there are cases in which additional system considerations are relevant, as noted above, and internal stakeholders have an impact on how a recommender system is designed. For example, in an e-commerce platform, the profit of each recommended item may be a factor in ordering and presenting recommendation results. This marketing function of recommender systems was apparent from the start in commercial applications, but rarely included as an element of research systems. Section 4.2 discusses value-awareness in recommender systems and potential trade-offs for user experience.

Alternatively, the system may seek to tailor outcomes specifically to achieve particular objectives that are separate from either provider or consumer concerns. For example, an educational site may view the recommendation of learning activities as a curricular decision and seek to have its recommendations fit a model of student growth and development. Its utility may, therefore, be more complex than a simple aggregation of those of the other stakeholders.

2.4.5 Other stakeholders

Complex online ecosystems may involve a number of stakeholders directly impacted by recommendation delivery beyond item providers and recommendation consumers. For example, an online food delivery service, such as UberEats, depends on delivery drivers to transport meals from restaurants to diners. Drivers are necessarily affected by recommendations delivered to users as the efficiency of routing and the distribution of order load will be a function of which restaurants receive such orders.

3 Related Research

Multistakeholder recommendation brings together research in a number of areas within the recommender systems community and beyond: (1) in economics, the areas of multisided platforms and fair division; (2) the growing interest in multiple objectives for recommender systems, including such concerns as fairness, diversity, and novelty; and, (3) the application of personalization to matching problems.

3.1 Economic foundations

The study of the multisided business model was crystallized in the work of Rochet and Tirole (2003) on what they termed "two-sided markets." Economists now recognize that such contexts are often multisided, rather than two-sided, and that "multisided-ness" is a property of particular business platforms, rather than a market as a whole (Evans and Schmalensee 2016).

Many of today's recommender systems are embedded in multisided platforms and hence require a multisided approach. The business models of multisided platforms are quite diverse, which means it is difficult to generalize about multistakeholder recommendation as well. A key element of the success of a multisided platform is the ability to attract and retain participants from all sides of the business, and therefore developers of such platforms must model and evaluate the utility of the system for all such stakeholders.

The theory of just division of resources has a long intellectual tradition going back to Aristotle's well-known dictum that "Equals should be treated equally." Economists have invested significant effort into understanding and operationalizing this concept and other related ideas. See (Moulin 2004) for a survey. In recommendation and personalization, we find ourselves on the other side of Aristotle's formulation: all users are assumed unequal and unequal treatment is the goal, but we expect this treatment to be consistent with diverse individual preferences. Some aspects of this problem have been studied under the subject of social choice theory (Arrow et al. 2010). However, there is not a straightforward adaptation of these classical economic ideas to recommendation applications as the preferences of users may interact only indirectly and in subtle ways. For example, if a music player recommends a hit song to user A, this will not in any way impact its desirability or availability to user B. On the other hand, if a job recommender system recommends an appealing job to user A, it may well have an impact on the utility of the same recommendation to user B who could potentially face an increased competitive environment if she seeks the same position.

3.2 Multi-objective recommendation

Multistakeholder recommendation is an extension of recent efforts to expand the considerations involved in recommender system design and evaluation beyond simple measurements of accuracy. There is a large body of recent work on incorporating diversity, novelty, long tail promotion and other considerations as additional objectives for recommendation generation and evaluation. See, for example, (Abdollahpouri et al. 2017a; Smyth and McClave 2001; Ziegler et al. 2005; Vargas and Castells 2011; Jannach and Adomavicius 2016; Abdollahpouri et al. 2019a). There is also a growing body of work on combining multiple objectives using constraint optimization techniques, including linear programming. See, for example, (Jambor and Wang 2010; Agarwal et al. 2011; Svore et al. 2011; Rodriguez et al. 2012b; Jiang and Liu 2012; Agarwal et al.

2012). These techniques provide a way to limit the expected loss on one metric (typically accuracy) while optimizing for another, such as diversity. The complexity of these approaches increases exponentially as more constraints are considered, making them a poor fit for the general multistakeholder case. Also, for the most part, multi-objective recommendation research concentrates on maximizing multiple objectives for a single stakeholder, the end user.

Another area of recommendation that explicitly takes a multi-objective perspective is the area of health and lifestyle recommendation. Multiple objectives arise in this area because users' short-term preferences and their long-term well-being may be in conflict (Lin et al. 2011; Ponce et al. 2015). In such systems, it is important not to recommend items that are too distant from the user's preferences – even if they would maximize health. The goal to be persuasive requires that the user's immediate context and preferences be honored.

Fairness is an example of a system consideration that lies outside the strict optimization of an individual user's personalized results. Therefore, recent research efforts on fairness in recommendation are also relevant to this work (Lee et al. 2014; Burke et al. 2017; Burke 2017; Kamishima et al. 2014; Kamishima and Akaho 2017; Yao and Huang 2017b; Mehrotra et al. 2018). The multi-stakeholder framework provides a natural "home" for such system-level considerations, which are otherwise difficult to integrate into recommendation. See Section 4.3 for a more in-depth discussion.

3.3 Personalization for matching

The concept of multiple stakeholders in recommender systems is suggested in a number of prior research works that combine personalization and matching. The earliest work on two-sided matching problems (Roth and Sotomayor 1992) assumes two sets of individuals, each of which has a preference ordering over possible matches with the other set. The task to make a stable assignment has been shown to have an $O(n^2)$ solution. This formulation has some similarities to reciprocal recommendation. However, it assumes that all assignments are made at the same time, and that all matchings are exclusive. These conditions are rarely met in recommendation contexts, although extensions to this problem formulation have been developed that relax some of these assumptions in online advertising contexts (Bateni et al. 2016).

Researchers on reciprocal recommendation have looked at bi-lateral considerations to ensure that a recommendation is acceptable to both parties in the transaction. A classical example is on-line dating in which both parties must be interested in order for a match to be successful (Pizzato et al. 2010b; Xia et al. 2015). Other reciprocal recommendation domains include job seeking (Rodriguez et al. 2012b), peer-to-peer "sharing economy" recommendation (such as AirBnB, Uber and others), on-line advertising (Iyer et al. 2005), and scientific collaboration (Lopes et al. 2010; Tang et al. 2012). See Section 4.1 for a detailed discussion.

The field of computational advertising has given considerable attention to balancing personalization with multistakeholder concerns. Auctions, a well-established technique for balancing the concerns of multiple agents in a competitive environment, are widely used both for search and display advertising (Yuan et al. 2012; Mehta et al. 2007). However, the real-time nature of these applications and the huge potential user base makes recommender-style personalization computationally intractable in most cases.

4 Examples

In the following sections, we introduce three applications of multistakeholder recommendation. These examples show some of the variety of contexts in which this concept can be applied.

Section 4.1 looks at what is perhaps the earliest application area for multistakeholder recommendation: reciprocal recommendation. In reciprocal recommendation, the recommender system matches users with other users, thus collapsing the distinction between consumers and providers: the consumers of recommendations are also the individuals who might be recommended to others. Section 4.2 examines the broad class of designs where the system's interest is enhancing profit or economic value related to recommendations that are produced.

The final example in Section 4.3 examines the problem of fairness in recommendation. Fairness is inherently a multistakeholder concept. If the only consideration in recommendation generation is matching individual user's known preferences, then the question of whether recommendations are fair does not arise. Fairness is therefore a quintessential system-level concern, not reducible to the problem of maximizing aggregate utility for either consumers or providers.

4.1 Example: People Recommendation

People recommendation is based on the notion of social matching (Terveen and McDonald 2005), as discussed above. The fact that the recommended entity is a person yields additional reciprocity and requires additional considerations of trust, privacy, reputation, and personal attraction. Recommendation in online dating is a paradigmatic example of people recommendation.

At the core of social media are individual relationships which serve as a fertile ground for recommendation. The underlying social network of a social media website – explicit through articulated connection or implicit via shared interests or goals – drives diffusion and engagement as well as key features such as news feeds and photo streams. The network's size is often considered a key metric of a social site's success. Recommendations of people on social media sites therefore play a key role in their success and have become ubiquitous (Guy 2018).

Many systems where people are in both sides of the recommendation process benefit from reciprocity even when reciprocity is not required in the system. In some of these domains such as Twitter, relationships have shown to be stronger, with a lower likelihood on one breaking a social link, when both users follow each other (Xu et al. 2013; Kwak et al. 2011; Kivran-Swaine et al. 2011).

4.1.1 Multistakeholder aspects

Where person-to-person recommendation is performed, any given individual can typically play both the role of consumer of recommendations and provider of an item to be recommended, namely their own profile. Thus the term *reciprocal recommendation*, which has often characterized this area of research.

In reciprocal recommendation, user profile data consisting of interactions with items may be sparser than in traditional recommendation contexts. In settings such as online dating and job recommendation, the task is often to find a suitable match quickly and exit the market. Users may only require a few interactions to achieve this goal, as opposed to consumption-oriented contexts where a user might rate dozens of books or hundreds of music tracks.

Because a successful recommendation in a reciprocal domain means that the user is likely to leave the system, conflicting incentives are created for platform owners. They want to have a profitable business by having repeated users, even though the best user experience would be for each user to instantly find a match (a successful date/partner) and never return.

Another consideration important in reciprocal recommendation is that, unlike other recommendation settings, the system's utility is not necessarily increasing function of the volume of recommendations. For instance, imagine if a highly qualified person is recommended to every single job position that they are fit to hold. This person is likely to be burdened by the amount of contact and might leave the website. A similar situation can occur for popular users in a dating website. These users are important as they represent the best of each of these services, but they can easily be overwhelmed by the interest of other users. Such a user should only be recommended to others when the recommender is highly confident that they will reciprocate.

Thus, people recommendation and reciprocal recommendation more generally will often have *value-awareness* aspects, where the system has recommendation objectives separate from those of consumers or providers.

4.1.2 Literature Review

In the area of online dating, RECON (Pizzato et al. 2010a) was the first recommender system to exploit the benefits of reciprocity. This system works by calculating a compatibility score between users and recommending people to people who have higher reciprocal compatibility scores. A number of studies followed this, including designs that focus on improving the cold-start problem of reciprocal recommenders (Akehurst et al. 2011; Yu et al. 2016).

Building on this work, Li and Li (2012) proposed MEET, which uses a bipartite graph that represent the mutual interest among a set of users to another set of users (men and women, in a heterosexual dating network). By creating subgraphs, it is able to perform graph inference and obtain a list of recommendations ranked based on mutual interests and filtered for users who exceed a certain availability budget.

Xia et al. (2016) proposed and compared a number of online dating recommenders including reciprocal content-based, memory-based and model-based collaborative filtering and found that memory-based methods outperform model-based models for female users who tend to have the largest sparsity in their interaction matrix.

Goswami et al. (2014) discusses reciprocal recommenders in more general terms as a two-sided market and proposes a two-layer architecture for recommendation ranking that looks at the preferences of both sides of the market. Alanazi and Bain (2016) developed a reciprocal recommender using hidden Markov models.

These approaches in online date have analogs in people recommendation in social media and social networks.

The most fundamental scenario of people recommendation on social media suggests familiar people for a long-term (permanent) connection, namely the recommendation of people to connect with on social network sites (SNSs), whose primary type of connection is symmetric (confirmed), such as Facebook and LinkedIn. This type of recommendation benefits both sides and reciprocity is its main characteristic. As a result, the person who receives the recommendation knows the other party (the recommended person) would have to confirm the connection and this party's anticipated reaction plays a key role in the decision making process leading to accepting or ignoring the recommendation. The permanent type of the connection also entails high weight: accepting such a recommendation may lead to a multi-year connection with another person on the SNS, which would involve receiving updates, news, photos, posts, and other types of information over a long period of time.

Widgets that proposed "people you may know" started to appear on leading SNSs at the end of the previous decade. Early work conducted on symmetric social networks within the enterprise, showed the benefit of aggregating multiple signals for recommendation (Chen et al. 2009) and indicated a dramatic effect on the number of connections on the site (Guy et al. 2009). It was also shown that providing evidence for a person's recommendation, such as the joint documents they have with the individual who receives the recommendation, helps making the latter feel more conformable accepting the recommendation and triggering the invitation to connect.

Two interesting followup studies were conducted by Guy et al. (2009), inspecting longer term effects. In the first, people recommendation was shown to increase engagement and retention rates on enterprise SNS when new users were introduced with people recommendations (Freyne et al. 2009). In this scenario, it was shown that the most effective recommendation ranking was by activity on the site rather than by the total weight of connection signals to the target user. Apparently, recommending active people has a special impact when trying to engage new users. The second followup study focused on the network effects of the provided recommendations (Daly et al. 2010). It was shown that different people recommendation algorithms render different network structures.

4.2 Example: Value-aware Recommendation

The literature in the field of recommender systems, as mentioned in Section 1, mainly focuses on the consumer perspective with the system being neutral regarding what items get recommended. The goal of most research efforts is therefore to design algorithms and systems that aim to provide value for the consumer in some form, e.g., by avoiding information overload or helping the consumer to discover new items. Even then, in many cases, researchers tend to abstract away from real-world consumer value metrics, such as consumer surplus or satisfaction, and focus on optimizing more general algorithmic metrics such as RMSE, NDCG, or precision and recall.

The underlying implicit assumption here is that recommending only assumedly relevant items to the user will also have a positive impact on the value for the provider or the platform. In fact, a number of studies support this hypothesis and show that providing personalized recommendations that are optimized to match the user's preferences lead to increased business value, e.g., in terms of increased sales or click-through rates (Garcin et al. 2014; Kirshenbaum et al. 2012; Jannach and Hegelich 2009); and, vice-versa, that unexpected or irrelevant recommendations can lead to a decreased quality perception and trust by consumers (Chau et al. 2013; Fitzsimons and Lehmann 2004). A recommender system can even be a competitive factor when other actors on the market do not have a recommendation service (Jannach and Adomavicius 2016).

A *value-aware recommender system* is distinguished from systems that assume platform value derives solely from maximizing user utility. It is a multistakeholder system in which the system's goal of realizing value from delivering recommendation is explicitly represented and optimized for.

4.2.1 Multistakeholder aspects

While factors like increased consumer engagement often considered to lead to indirect business value, e.g., in terms of an increased number of monthly re-subscriptions, recommendations can also be used to positively impact the business in a more direct way. Specifically, recommenders can be implemented as a tool that steers consumer demand, e.g., by promoting certain items. The particular goal in that context can be to drive demand in a direction that maximizes the platform's short-term or long-term profit, while also maintaining an acceptable level of consumer utility.

Simply recommending those items with the highest profit for the platform is probably in almost all cases not the optimal strategy, at least not in the long run, as consumers might distrust a recommendation service when its suggestions are not considered useful. Generally, we can hypothesize that in many domains there is a trade-off between suggesting items that are the most profitable for the platform and suggesting those that are considered the most relevant for the user.

However, an additional intuitive assumption in that context might also be that a recommender is still effective in steering consumer demand if the order of the suggestions is not strictly determined by the assumed relevance for the consumer, but also takes profitability considerations into account.

Figure 1 shows an example of simulated results from a value-aware recommender system. This system simulates differential profit across items and can be configured to promote more profitable items at the expense of fidelity to the user's interests. The curve shows the change in profit per user as we increase the threshold by which the system prefers profit to user utility, taking into account users who might reject inaccurate recommendations. Under the assumptions of this work, it is possible, to increase profit per user within a certain range, even while promoting items with higher profitability.

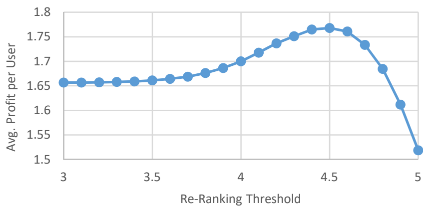

Fig. 1: Recommendation profit optimization, from (Jannach and Adomavicius 2017).

4.2.2 Literature Review

In the literature, a variety of methodological approaches of different complexities have been explored to incorporate profit information into recommenders and to balance relevance and profitability.

Considering not only the consumer preferences but also the profitability for the seller, as discussed in the example above, was in the focus in (Chen et al. 2008). In this work, the authors compared different recommendation strategies that combine general and customer-individual purchase probabilities of the items with profitability information on synthetic data. Their simulation results like those shown above indicate that higher overall profitability can be achieved without a loss of accuracy for personalized recommendations. Focusing too much on profitability, however, leads to an accuracy degradation.

Going beyond this comparably simple model, Wang and Wu (2009) framed the selection of products for customers as a constrained optimization problem. The constraints in the model ensure that the recommended products match the customer's preferences and their assumed budgets. This work therefore also considers the user's price sensitivities. Different optimization goals can be configured which either maximize the profit for the seller or lead to a win-win situation where seller profits and customer value are balanced. Alternative approaches that model the recommendation problem as a mathematical optimization task were later on put forward in (Das et al. 2009; Akoglu and Faloutsos 2010; Hammar et al. 2013; Azaria et al. 2013). The model proposed in (Das et al. 2009), for example, includes the concept of *trust*, assuming that a consumer will continue to make purchases as long as the system is able to predict their preferences to a certain extent. The proposed work unfortunately remained on a theoretical level and it is in particular unclear to what extent the assumed trust model is realistic.

Lu et al. (2014) take yet another set of factors into account in the revenue model of their optimization-based approach, including prices, saturation effects and competition effects. A specific aspect of their work is that they optimize the model over a finite time horizon, where the adoption probability at each time step can depend on different factors such as the previous purchases by the consumer in the same class, the number of times a certain item was already recommended, or the current price of the item and the individual consumer's willingness-to-pay (WTP) for it. Given the hardness of the resulting optimization problem, the authors propose greedy optimization strategies which they empirically evaluated on semi-synthetic data. Pricing based on the predicted WTP is also considered in (Kamishima and Akaho 2011).

A quite different optimization problem is formulated in (Hammar et al. 2013), where the goal is to generate a set of recommendations that maximizes the probability of a purchase. Therefore, instead of maximizing the revenue based on individual-item profitability considerations, the main short-term goal is to convert the visitor to a buyer. Challenging existing works that solely focus on purchase probabilities, Bodapati (2008) argues that one should also consider how consumer's would behave if no recommendations would be presented to them. If a certain product will be purchased by a consumer anyway with a certain probability, it might be better not to recommend it, given the limited number of recommendations that can be made.

Generally, all works discussed so far were mostly evaluated based on some form of simulations based either on synthetic or real-world data. The work of Azaria et al. (2013) is one of the few examples where the consumers' quality perception of and satisfaction with profit-optimized recommendation was assessed in a user study. The participants, who were recruited via Amazon Mechanical Turk, received different types of recommendations and were also queried about their willingness to pay. The results of a real-world field test on the effect of different recommendations, including a profit-oriented one based on the model of (Hosanagar et al. 2008), are reported Panniello et al. (2016). In their study, recommendations were provided to consumers through e-mail newsletters. Both these studies found that a profit-sensitive strategy led to an increased average revenue without a significant loss in terms of the participants' satisfaction.

Another real-world study reported in (Nguyen et al. 2017) looks at a multisided platform where the system earns a variable commission from sales to different providers. For example, a travel web site may earn a commission when a user uses its site to book a hotel room, and these commissions may vary by property. The choice of which hotel rooms to recommend therefore involves a distribution of utility among all three stakeholders. In this implementation, a learning-to-rerank algorithm was developed and testing showed that margins could be increased with acceptable loss in ranking accuracy.

The study by Panniello et al. is one of the few works that consider longer-term effects of profit-aware recommendations. Longer-term effects were studied in particular by Hosanagar et al. (2008) who concluded from their theoretical analyses that optimal recommendations balance profit margins and item relevance. Furthermore, they emphasized the importance of considering the current reputation of the provider when implementing the strategy.

The expected *customer lifetime value* (CLV) is a well-known instrument from the management and marketing literature. Limited work however exists that tries to connect CLV-related activities like promotions, cross-buying or retention-pricing in combination with recommender systems. Recency, frequency, and monetary (RFM) characteristics of consumers are usually used as a basis for CLV estimates. One possible approach, as discussed in (Liu and Shih 2005; Shih and Liu 2008), is to group consumers into segments if their CLV estimates are similar and to incorporate the consumer's segment assignment in the recommendation process. To what extent the consideration of these aspects has an impact on longer-term customer loyalty and the resulting CLV was however not yet been the focus of experimental evaluation.

4.3 Example: Fairness-aware Recommendation

The problem of bias and fairness in algorithmic systems generally and in machine learning systems in particular has been the subject of increasing research interest. Recommender systems may also be scrutinized relative to fairness considerations. However, because of their multi-sided nature, unfairness may arise in multiple ways. As noted in (Burke 2017), with respect to our stakeholder taxonomy, consumer-side fairness (C-fairness) and provider-side fairness (P-fairness) may be relevant depending on the specifics of a given application.

In C-fairness, the concern is that different users or groups of users may be receiving different types of recommended items or quality of recommendation, and that such differences rise to the level of discrimination on the part of the algorithm. For example, if a system recommending credit card offers were systematically showing those with higher interest rates to female users, regardless of other characteristics, its behavior might be considered discriminatory and even unlawful in some jurisdictions Steel and Angwin (2010).

Separate considerations arise when considering fairness towards providers (P-fairness). The concern in this case is that the benefits of participating in the system may be unevenly distributed to the individuals that are associated with items being recommended. For example, in a music recommendation system, fairness towards the artists whose music is being recommended may be important for a number of reasons (Mehrotra et al. 2018).

4.4 Multistakeholder aspects

Fairness concerns may exist on different sides of the recommendation platform and metrics related to fairness can therefore be employed for multistakeholder evaluation of recommendation outcomes. We discuss some of these measures below. However, that is not to say that fairness is particularly a consumer or provider concern. Self-interested actors in these spaces may prefer that the recommender systems be unfair in their favor rather than fair for all, and it is up to system designers and implementers to craft appropriate tradeoffs.

We can formalize C-fairness in different ways through underlying considerations of statistical independence. Effectively, if we have some sensitive feature S, along with fairness is sought, we can say that the system is fair if the results it produces are independent of this feature (Calders and Verwer 2010): for example, the requirement that a particular credit offer is recommended should be independent of gender, all other things being equal.

Table 1 lists three different independence criteria that might be employed. We use the notation $A \perp\!\!\!\perp B$ indicating the (unconditional) independence between variables A and B, and $A \perp\!\!\!\perp B \mid C$ indicating the conditional independence between A and B given C. The first criteria of unconditional independence requires that predictions be independent of the sensitive feature. From a viewpoint of information theory, this condition equivalent to specifying that mutual information between $\hat{Y}$ and S is zero. Conditioning the independence on X, as in the second definition, allows aspects of the user and item to be taken into consideration, but it is important to note that features in X that are correlated with S then may be used by proxy: the well-known "redlining effect" (Calders and Verwer 2010). Finally, if we assume that user judgements Y are not biased, we may want simply to ensure that prediction errors are not influenced by a sensitive feature due to inductive bias of a prediction algorithm. For example, a dimensionality-reduction algorithm may treat the preferences of a group of users as "noise" if they are a small minority in the overall population. Similar formulations can be made considering the sensitive features of providers.

As with value-aware recommendation, fairness considerations in recommendation will need to be balanced against overall system accuracy and the ability to present personalized results. The above independence measures are easy to achieve by predicting items at random. Each application context will be different to the extent that different aspects of fairness and different sensitive features are relevant.

Table 1: Independence criteria of consumer-side fairness. Y represents user feedback over some set of items, $\hat{Y}$ represents the predicted interest of user for items, S represents a sensitive feature, and X represents all other features used in prediction.

Type	Formal statement	Related concepts
Unconditional independence	$\hat{Y} \perp\!\!\!\perp S$	indirect fairness, statistical parity, disparate impact
Conditional independence	$\hat{Y} \perp\!\!\!\perp S \mid X$	direct fairness, disparate treatment
Error independence	$\hat{Y} \perp\!\!\!\perp S \mid Y$	equalized odds

4.5 Literature review

As fairness concerns in machine learning have grown, similar issues have been raised with regard to recommendation. The personalization and multisided aspect of recommendation mean that approaches from machine learning cannot be adapted directly, and recommendation-specific approaches to fairness have been developed.

One approach has been to control the model fitting process so that it prefers fairer outcomes. Regularization models proposed for classification fairness such as Kamishima et al. (2012) can be adopted for recommendation. Depending on how the independence criterion is formulated, different solutions can be found. (Kamishima et al. 2013) uses mean matching between protected and unprotected groups, but does not control for the difference in variance. Following research proposed distribution matching and mutual information as alternate ways to control the variance (Kamishima et al. 2018). In (Burke et al. 2018b), a regularization approach was applied to balance the weightings of different groups when generating recommendations. The error independence criterion was approached through regularization in (Yao and Huang 2017a). Fairness considerations can also be built into generative recommendation approaches as demonstrated in (Kamishima et al. 2016).

Many aspects of consumer-side fairness have seen little research. The causes of unfair outcomes in recommendation have only been lightly studied, although (Ekstrand et al. 2018a) and (Abdollahpouri et al. 2019b) provide some important clues, namely that the size of different user groups is not always a good indicator of outcomes. There has been little study of fairness in rivalrous contexts, where a recommendation to user A changes the utility of the same item for another user, except in the limited context of reciprocal recommendation noted above.

Provider fairness offers an incomplete symmetry with C-fairness. The biggest difference between consumers and item providers is that providers are more or less passive in the recommendation interaction. Items lie in wait for users seeking recommendations. The opportunity to appear on a recommendation list is therefore a quantity that is relatively fixed and fairness across providers is much more of a zero-sum game than fairness for consumers. Outcomes for providers are measured in different ways than for consumers as well, and while the definitions in Table 1 may still apply, there are a variety of ways to measure outcomes. See the discussion in Section 5.5.1 below as well as recent work in fairness in ranking (Zehlike et al. 2017b,a; Singh and Joachims 2018; Beutel et al. 2019). Some of the same techniques of regularization can be applied as in (Abdollahpouri et al. 2017a). Tensor factorization allows content information and preference information to be combined in a single optimization (Zhu et al. 2018), and there has been research applying fairness in a separate re-ranking step including (Liu et al. 2019; Karako and Manggala 2018b).

5 Methodological Issues

At this point in the development of multistakeholder recommendation research, there is a diversity of methodological approaches and little agreement on basic questions of evaluation. In part, this is a reflection of the diversity of problems that fall under the multistakeholder umbrella. Multistakeholder evaluation and algorithm development do not always use the same methodologies.

A key difficulty is the limited availability of real-world data with multistakeholder characteristics. The reason for this becomes clear if we consider the experiments discussed in Section 4.2. The data that makes value-aware experiments possible is highly business-critical, including such data as the margins associated with each provider and the commissions negotiated by the platform. Close collaboration is required to obtain such sensitive proprietary data. Some researchers have built such collaborations for multistakeholder research, but progress in the field requires replicable experiments that proprietary data does not support. Areas of multistakeholder research that involve public, rather than private, benefit may offer advantages in terms of the availability of data: see, for example, the data sets available from the crowd-funded educational charity site DonorsChoose.org[1].

5.1 Simulation

In the absence of real-world data with associated valuations, researchers have typically turned to simulations. Simulated or inferred provider data is useful for transforming publicly-available recommendation data sets in standard user, item, rating format into ones that can be used for multistakeholder experimentation. The experiments in (Sürer et al. 2018) provide an example of this methodology: each item in the MovieLens 1M data set was assigned to a random provider, and the distribution of utilities calculated. To capture

[1] https://data.donorschoose.org/explore-our-impact/

different market conditions, the experimenters use two different probability distributions: normal and power-law. There is no accepted standard for producing such simulations and what are reasonable assumptions regarding the distribution of provider utilities or the formulation of system utilities, except in special cases.

Researchers have also used objective aspects of data sets to infer proxy attributes for multistakeholder evaluation. In (Burke et al. 2016), the first organization listed in the production credits for each movie was treated as the provider – a significant over-simplification of what is a very complex system of revenue distribution in the movie industry. In other work, global metrics such as network centrality (Akoglu and Faloutsos 2010) have been used to represent system utility for the purposes of multistakeholder evaluation. (Burke et al. 2018a) demonstrated an alternate approach to generate synthetic attribute data based on behavioral characteristics that can be used to evaluate system-level fairness properties.

More sophisticated treatments of profitability and recommendation are to be found in the management and e-commerce literature, some using public data as seen in (Oestreicher-Singer and Sundararajan 2012; Chen et al. 2004; Adamopoulos and Tuzhilin 2015), but these techniques and associated data sets have not yet achieved wide usage in the recommender system community.

5.2 Models of utility

A multistakeholder framework inherently involves the comparison of outcomes across different groups of individuals that receive different kinds of benefits from the system. In economic terms, this entails utility calculation and comparison. As with data, different researchers have made different assumptions about what types of utilities accrue from a recommender system and how they are to be measured. A standard assumption is that the output of a recommendation algorithm can be treated as an approximation of user utility. Yet, research has confirmed that users prefer diverse recommendation lists (Pu et al. 2011), a factor in tension with accuracy-based estimates of user utility.

Most of the examples discussed above focus solely on the short-term perspective. More research is therefore required to understand the potential positive and negative long-term effects of profit-aware recommendation and other strategies that are not strictly user-focused. Future models could also consider the price sensitivity and willingness-to-pay of individual consumers in the recommendation process.

5.3 Off-line experiment design

A standard off-line experimental design in recommender systems is the creation of multiple folds of training and test data from a data set using random fixed-sized partitioning of user profiles. The benefit of this approach is that each partition contains a fixed proportion of each user's profile, guaranteeing a minimum profile size for recommendation generation. This makes sense when user outcomes are the highest priority, as it ensures that an evaluation data point can be produced for every user in every fold. All of the recommendations for a test fold are produced, in some sense, simultaneously, as a set of recommendation lists or rating predictions to which evaluation metrics can be applied.

This experimental framework makes a bit less sense in a multistakeholder context, and this is where the essential asymmetry of the stakeholders comes into play. Providers are, in a key sense, passive – they have to wait until users arrive at the system in order to have an opportunity to be recommended. The randomized cross-fold methodology measures what the system can do for each user, given a portion of their profile data, the potential benefit to be realized if the user visits and a recommendation list is generated. Evaluating the provider side under the same conditions, while a commonly-used methodology, lacks a similar justification.

A more realistic methodology from the provider's point of view is a temporal one, that takes into account the history of the system up to a certain time point and examines how provider utilities are realized in subsequent time periods. See (Campos et al. 2014) for a comprehensive discussion of time-aware recommender systems evaluation. However, time-aware methods have their own difficulties, forcing the system to cope with cold-start issues possibly outside of the scope of a given project's aims.

5.4 User studies

User studies are another instrument available to researchers that has not been extensively applied to multistakeholder recommender systems. As usual for such studies, the development of reliable experimental designs is challenging as the participants' decision situation typically remains artificial. Furthermore, as in the study by Azaria et al. discussed above (Azaria et al. 2013), familiarity biases might exist – in their study participants were willing to pay more for movies that they already knew – which have been observed for other types of user studies in the recommender systems domain (Jannach et al. 2015). Ultimately, more field tests – even though they are typically tied to a specific domain and specific business model – are needed that give us more insights into the effects of multistakeholder recommendations in the real world.

5.5 Evaluation Metrics

The building block of multistakeholder evaluation is the measurement of the utility each of the stakeholders gets within a recommendation platform. Common evaluation metrics such as RMSE, precision, NDCG, diversity, etc. are all different ways to evaluate the performance of a recommender system from the user's perspective. As noted above, these measures are implicitly a form of system utility measure as well: system designers optimize for such measures under the assumptions that (1) higher evaluation metrics correspond to higher user satisfaction and (2) higher user satisfaction contributes to higher system utility through customer retention, trust in the recommendation provided, etc. However, the formulation of multistakeholder recommendation makes it possible to characterize and evaluate system utility explicitly.

Typically, evaluation metrics are averaged over all users to generate a point score indicating the central tendency over all users. However, it is also the case that in a multistakeholder environment additional aspects of the utility distribution may be of interest. For example, in an e-commerce context, providers who receive low utility may leave the eco-system, suggesting that the variance of provider utility may be important as well as the mean. One suggested practice would be to report the first three moments of the distribution of utilities for each stakeholder – the mean, variance, and skewness – rather than just the mean value when reporting on multistakeholder evaluations, and to report results for significant user groups within their populations.

5.5.1 Provider metrics

When evaluating the utility of a recommender system for a particular provider, we may take several different stances. One views the recommender as a way to garner user attention. In this case, the relevance of an item to a user may be a secondary consideration. Another perspective views the recommender as a source of potential leads. In this view, recommending an item to uninterested users is of little benefit to the provider. In the first situation, simply counting (with or without a rank-based discount) the number of times a provider's products appears in recommendation lists would be sufficient. In the second situation, the metric should count only those recommendations that were considered "hits", those that appear positively rated in the corresponding test data.

Another provider consideration may be the reach of its recommended products across the user population. A metric could count the number of unique users to whom the provider's items are recommended. Multistakeholder applications may differ in their ability to target specific audiences for their items. In a targeted system, it would make sense to consider reach relative to the target population. For example, in an online dating application where the user can specify desired properties in a match, an evaluation metric might be the fraction out of the target audience receiving the recommendation.

Finally, where the consideration is the accuracy of system's predictions, we can create a provider-specific summary statistic of a measure like RMSE. (Ekstrand et al. 2018b) uses this method to examine differences in error when recommending books by male and female authors. Since the statistic by itself is not that useful for a single provider, a better metric would indicate the provider's position relative to other providers in the overall distribution

Table 2 shows example metrics for each of the provider cases. Note that these metrics can be normalized in different ways. For example, the count-oriented metrics may be normalized by the size of the provider catalog, and / or by the number of users, etc. For simplicity, we omit a complete enumeration of all such variants here. Note also that a provider might be interested in their rank or scoring relative to other providers. For example, an Exposure value of 600 might be more meaningful if the provider is also told that this value ranks 3rd across all providers or that it is 1.2 standard deviations above the mean value.

Type	Formula	Explanation
Exposure(p)	$\sum_{L_i \in \mathcal{L}} \sum_{j \in L_i} \mathbb{1}(j \in I_p)$	Count the number of recommendations given across all of p's items.
Hits(p)	$\sum_{L_i \in \mathcal{L}} \sum_{j \in L_i} \mathbb{1}(j \in I_p \wedge r_{ij} \in T)$	Count the number of hits in recommendation lists for all of p's items.
Reach(p)	$\sum_{L_i \in \mathcal{L}} \mathbb{1}(\lvert I_p \cap L_i \rvert > 0)$	Count how many users get at least one i_p item recommended.
TargetReach(p)	$\sum_{L_i \in \mathcal{L}} \mathbb{1}(\lvert I_p \cap L_i \rvert > 0 \wedge g_p(i))$	Count how many users in p's target set get at least one i_p item recommended.
PAccuracy(p,m)	$[\sum_{r_{ij} \in T_p} m(r_{ij}, \hat{r}_i j)]/\lvert T_p \rvert$	Average metric m score for predictions of p's items.

Table 2: Examples of provider metrics. Let p be a given provider, and $i_p \in I_p$ an item associated with p. Let $\mathcal{L} = L_0, L_1, ..., L_n$ be the recommendation lists calculated for n users. Let T be set of r_{ij} ratings in the test set over which $\mathcal{L}$ is calculated. Let T_p be provider p's subset of T: $T_p = \{r_{ij} : r_{ij} \in T \wedge i \in I_p\}$. Let $\mathbb{1}$ be the indicator function. Let $m(r_{ij}, \hat{r}_{ij})$ be an accuracy-oriented evaluation metric (such as RMSE) that evaluates a predicted rating $\hat{r}_{ij}$ relative to a known value $r_{ij} \in T$. Let $g_p(i)$ be a boolean function that returns true if user i is in the target market of provider p.

5.5.2 System metrics

System utility may in many cases be a simple aggregation of the utilities of other parties. For example, in a simple commission-oriented arrangement, the profit to the system might be some weighted aggregate of the *Hits* metric, taking item price and commission rate into account. However, other cases arise where the system has its own targeted utility framework.

An important such context is algorithmic fairness discussed in Section 4.3. In general, we should not expect that providers will care if the system is fair to others as long as it provides them with good outcomes. Any fairness considerations and related metrics will therefore be ones defined by system considerations. For example, we can define our provider metrics to map to providers with and without sensitive features, and $P(Y=1)$ as $\text{Reach(p)}/|\mathcal{L}|$. Then, a proportional impact metric such as $\Pr[\hat{Y}{=}1|Z{=}0]/\Pr[\hat{Y}{=}1|Z{=}1]$ can be defined to see how closely a particular set of recommendation results tracks the desired fairness outcome where this ratio approaches 1.

Other discussions of system utilities are relatively sparse in the multistakeholder recommendation literature. As noted in Section 4.2 above, considerations such as customer lifetime value are candidates for system-oriented metrics, as they are not simply reducible to the utilities of other stakeholders. Some of the applications discussed below, such as educational recommendation, also present some interesting challenging for defining and applying system metrics.

6 Research Directions

While there are a number of examples of notable research results in multistakeholder recommendation, a number of important unsolved challenges remain. In this section, we examine some important research directions in which we expect future progress.

6.1 Algorithms

As noted above, existing work has explored some algorithmic approaches to multistakeholder recommendation. Two approaches can be identified: the first situates the multistakeholder problem within the core recommendation generation function as a type of multi-objective optimization, the second applies multistakeholder considerations after an initial set of recommendations has been generated.

The multi-objective approach builds a loss function that incorporates multiple objectives and attempts to learn a recommendation function that is sensitive both to a standard accuracy-oriented objective, which can be understood as a consumer-side consideration, and to an objective that is oriented towards some other stakeholder, such as a fairness objective belonging to the system. Important research challenges remain in formulating and applying multiple objectives across the wide range of multistakeholder applications.

The second algorithm type is one that employs an existing recommendation algorithm to generate recommendations (again, this is understood as the user-oriented aspect of the system) and then other stakeholders' considerations are integrated through a re-ranking process. Such systems have the benefit of being modular, so that improvements can be made and analyzed for each part of the algorithm separately. Researchers have built on existing work in information retrieval such as maximum marginal relevance (Karako and Manggala 2018a) and xQuad (Liu and Burke 2018) as well as constraint satisfaction (Sürer et al. 2018) and probabilistic soft logic (Farnadi et al. 2018).

One discernible trend in algorithmic research in recommender systems has been the move from narrower objectives for recommendation algorithms to broader ones. Initially, point-wise accuracy metrics were developed, which evolved into pair-wise and list-wise metrics, and more recently to considering interactions extended in time. Multistakeholder recommendation raises the possibility of broadening the objective yet again towards optimizing over the entire set of recommendations delivered. The approach in (Sürer et al. 2018) is one step in this direction, as it formulates the re-ranking problem as approximating constraint satisfaction over all of the recommendation lists generated for the test data. While the interactive requirements of recommendation may seem to argue against the computation of a global optimum, many applications generate and cache recommendations and would be a good match for this kind of algorithm.

6.2 Applications

The pattern of multistakeholder recommendation can be observed in variety of different applications. The prior discussion has highlighted existing research in the reciprocal domains of job recommendation and online dating, in the value-aware environments in e-commerce and multisided platforms, and in the environments where fairness considerations apply. There are many additional areas of application for multistakeholder recommendation, some of which are listed here.

Education: Depending on the environment, recommendation of educational content may have a multistakeholder aspect. For example, there may be tension between the interests of students who want to pursue familiar content and those of the educational system that may be interested in producing students with a well-rounded range of experiences. When multiple educational providers are involved, there may be provider considerations as well (Burke and Abdollahpouri 2016; Zheng et al. 2019).

Philanthropy: Commerce-oriented multisided platforms are obvious examples where multistakeholder considerations are important. However, there are also multisided platforms that have philanthropic aims. The crowd-sourced microlending platform Kiva.org is such an example where fairness-aware recommendation has been applied (Lee et al. 2014; Burke et al. 2018b).

Tourism: Another example of recommender systems involving multiple stakeholders is tourism. For example when a travel recommender system recommends a destination or a travel package to a user, the stakeholders that are involved include the traveler, the airlines (or any other transportation provider), the host (destination) and also the system. The hotel recommendation system in Krasnodebski and Dines (2016) shows some of the multisided nature of interactions in travel. Peer-to-peer travel services like AirBnB may also have reciprocal aspects.

News recommendation: News recommendation can be viewed as strictly a matter of personalizing for user taste, but there are considerations – such as public service goals or regulatory requirements – that might require fairness, balance or other system objectives (Tintarev et al. 2018).

Social media: In social media platforms, users get a variety of different content merged as a ranked content stream, that can be understood as a set of recommendations. For example, on Facebook, users get friends' posts as one type of recommendation and ads as another type. Thus, we have multiple types of providers and a task of balancing the content of the feed so that multiple objectives are met.

As this set makes clear, the scope of multistakeholder recommendation is quite broad and incorporates systems of societal importance. It may be inevitable that, as recommender systems move further into applications with more significant stakes for individuals and society, it becomes more and more necessary that they serve a multiplicity of purposes.

6.3 Explanation / Transparency

Exploring how recommendation explanation could be done in a multistakeholder environment is also another direction for future research. Explanation is an important factor in recommendation interfaces, helping users understand how a given recommendation relates to their interests. It has been shown that explanations can enhance users' likelihood of adopting a given recommendation (Zhang and Chen 2018).

Multistakeholder recommendation poses some interesting challenges for recommendation explanation. First, there is the issue of complexity: a recommendation produced by a multistakeholder system will, by necessity, be one that incorporates multiple factors in its production, and therefore any explanation will be more complex than what would be needed if user preferences were the only consideration. In addition, there is the fact that the objectives of some of the other stakeholders may be in conflict with those of the user. In some contexts, one could imagine users finding it objectionable that their preferences are being downplayed in favor of others' interests. E-commerce sites that confront this problem often label items in recommendation lists as "promoted" or "sponsored" when they are being displayed because of provider considerations. It is more difficult to do this when a global optimization algorithm is being applied, as all results will potentially be influenced by the full set of stakeholders. Producing acceptable explanations in such contexts is an interesting challenge, but a good solution may be necessary to make multistakeholder recommendations broadly useful.

7 Conclusion

The emergence of multistakeholder recommendation as a research area is an important development in the evolution of the field, as researchers widen their view of those impacted by the results recommender systems produce. This is a natural progression from the initial academic research prototypes to today's fielded systems, key elements of online applications, with millions of users. It is not surprising that systems occupying key positions in complex commercial and social environments should have to answer to many masters.

While multistakeholder issues have surfaced regularly in the history of recommender systems research and have been a constant constraint in fielded applications, the recognition of common threads and research questions has been a more recent occurrence. This article has presented a synthesis of the landscape of this research past and present, demonstrated some important current applications, and raised important questions for future investigation.

References

Himan Abdollahpouri and Robin Burke. Multi-stakeholder recommendation and its connection to multi-sided fairness. In *Workshop on Recommendation in Multi-stakeholder Environments (RMSE'19), in conjunction with the 13th ACM Conference on Recommender Systems, RecSys'19, 2019*, 2019.

Himan Abdollahpouri, Robin Burke, and Bamshad Mobasher. Controlling popularity bias in learning to rank recommendation. In *Proceedings of the 11th ACM conference on Recommender systems*, 2017a.

Himan Abdollahpouri, Robin Burke, and Bamshad Mobasher. Recommender Systems as Multistakeholder Environments. In *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization (UMAP'17)*, page To appear, Bratislava, Slovakia, 2017b.

Himan Abdollahpouri, Robin Burke, and Bamshad Mobasher. Managing popularity bias in recommender systems with personalized re-ranking. In *The Thirty-Second International Flairs Conference*, 2019a.

Himan Abdollahpouri, Masoud Mansoury, Robin Burke, and Bamshad Mobasher. The unfairness of popularity bias in recommendation. In *RecSys workshop on Recommendation in Multi-Stakeholder Environments*, 2019b.

Panagiotis Adamopoulos and Alexander Tuzhilin. The business value of recommendations: A privacy-preserving econometric analysis. In *Proc. ICIS '15*, 2015.

Deepak Agarwal, Bee-Chung Chen, Pradheep Elango, and Xuanhui Wang. Click Shaping to Optimize Multiple Objectives. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '11, pages 132–140, New York, NY, USA, 2011. ACM.

Deepak Agarwal, Bee-Chung Chen, Pradheep Elango, and Xuanhui Wang. Personalized Click Shaping Through Lagrangian Duality for Online Recommendation. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '12, pages 485–494, New York, NY, USA, 2012. ACM.

Joshua Akehurst, Irena Koprinska, Kalina Yacef, Luiz Pizzato, Judy Kay, and Tomasz Rej. CCR - a content-collaborative reciprocal recommender for online dating. In *Proceedings of the 22nd International International Joint Conference on Artificial Intelligence (IJCAI 2011)*, Barcelona, Spain., July 2011.

L. Akoglu and C. Faloutsos. Valuepick: Towards a value-oriented dual-goal recommender system. In *ICDM '10 Workshops*, pages 1151–1158, 2010.

Ammar Alanazi and Michael Bain. A scalable people-to-people hybrid reciprocal recommender using hidden markov models. *2nd Int. Work. Mach. Learn. Methods Recomm. Syst*, 2016.

Kenneth J Arrow, Amartya Sen, and Kotaro Suzumura. *Handbook of social choice and welfare*, volume 2. Elsevier, 2010.

Amos Azaria, Avinatan Hassidim, Sarit Kraus, Adi Eshkol, Ofer Weintraub, and Irit Netanely. Movie recommender system for profit maximization. In *Proc. RecSys '13*, pages 121–128, 2013.

Mohammad Hossein Bateni, Yiwei Chen, Dragos F Ciocan, and Vahab Mirrokni. Fair resource allocation in a volatile marketplace. 2016. URL `AvailableatSSRN:https://ssrn.com/abstract=2789380`.

Alex Beutel, Jilin Chen, Tulsee Doshi, Hai Qian, Li Wei, Yi Wu, Lukasz Heldt, Zhe Zhao, Lichan Hong, Ed H Chi, et al. Fairness in recommendation ranking through pairwise comparisons. *arXiv preprint arXiv:1903.00780*, 2019.

Anand V. Bodapati. Recommendation systems with purchase data. *J. Mark. Research*, 45 (1):77–93, 2008.

J.S. Breese, D. Heckerman, C. Kadie, et al. Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the 14th conference on Uncertainty in Artificial Intelligence*, pages 43–52, 1998.

R.D. Burke, K.J. Hammond, and BC Yound. The FindMe approach to assisted browsing. *IEEE Expert*, 12(4):32–40, 1997.

Robin Burke. Multisided Fairness for Recommendation. In *Workshop on Fairness, Accountability and Transparency in Machine Learning (FATML)*, page 5 pages, Halifax, Nova Scotia, 2017.

Robin Burke and Himan Abdollahpouri. Educational Recommendation with Multiple Stakeholders. In *Third International Workshop on Educational Recommender Systems*, Omaha, 2016.

Robin Burke, Himan Abdollahpouri, Bamshad Mobasher, and Trinadh Gupta. Towards multi-stakeholder utility evaluation of recommender systems. In *Proceedings of the International Workshop on Surprise, Opposition, and Obstruction in Adaptive and Personalized Systems (SOAP 2016)*. ACM, 2016.

Robin Burke, Nasim Sonboli, Masoud Mansoury, and Aldo Ordoñez-Gauger. Balanced neighborhoods for fairness-aware collaborative recommendation. In *Workshop on Responsible Recommendation (FATRec)*, 2017.

Robin Burke, Jackson Kontny, and Nasim Sonboli. Synthetic attribute data for evaluating consumer-side fairness. *arXiv preprint arXiv:1809.04199*, 2018a.

Robin Burke, Nasim Sonboli, and Aldo Ordonez-Gauger. Balanced neighborhoods for multi-sided fairness in recommendation. In *Conference on Fairness, Accountability and Transparency*, pages 202–214, 2018b.

T. Calders and S. Verwer. Three naive Bayes approaches for discrimination-free classification. *Data Mining and Knowledge Discovery*, 21:277–292, 2010.

Pedro G. Campos, Fernando Díez, and Iván Cantador. Time-aware recommender systems: a comprehensive survey and analysis of existing evaluation protocols. *User Modeling and User-Adapted Interaction*, 24(1):67–119, 2014. ISSN 1573-1391.

Patrick Y.K. Chau, Shuk Ying Ho, Kevin K.W. Ho, and Yihong Yao. Examining the effects of malfunctioning personalized services on online users' distrust and behaviors. *Decis. Support Syst.*, 56(C):180–191, 2013.

Jilin Chen, Werner Geyer, Casey Dugan, Michael Muller, and Ido Guy. Make new friends, but keep the old: Recommending people on social networking sites. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 201–210, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-246-7. doi: 10.1145/1518701.1518735. URL `http://doi.acm.org/10.1145/1518701.1518735`.

Long-Sheng Chen, Fei-Hao Hsu, Mu-Chen Chen, and Yuan-Chia Hsu. Developing recommender systems with the consideration of product profitability for sellers. *Information Sciences*, 178(4):1032 – 1048, 2008.

Pei-Yu Sharon Chen, Shin-yi Wu, and Jungsun Yoon. The impact of online recommendations and consumer feedback on sales. In *Proc. ICIS '04*, pages 711–724, 2004.

Elizabeth M. Daly, Werner Geyer, and David R. Millen. The network effects of recommending social connections. In *Proceedings of the Fourth ACM Conference on Recommender Systems*, RecSys '10, pages 301–304, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-906-0. doi: 10.1145/1864708.1864772. URL `http://doi.acm.org/10.1145/1864708.1864772`.

Aparna Das, Claire Mathieu, and Daniel Ricketts. Maximizing profit using recommender systems. *CoRR*, abs/0908.3633, 2009.

Michael D Ekstrand, Mucun Tian, Ion Madrazo Azpiazu, Jennifer D Ekstrand, Oghenemaro Anuyah, David McNeill, and Maria Soledad Pera. All the cool kids, how do they fit in?: Popularity and demographic biases in recommender evaluation and effectiveness. In *Conference on Fairness, Accountability and Transparency*, pages 172–186, 2018a.

Michael D Ekstrand, Mucun Tian, Mohammed R Imran Kazi, Hoda Mehrpouyan, and Daniel Kluver. Exploring author gender in book rating and recommendation. In *Proceedings of the 12th ACM Conference on Recommender Systems*, pages 242–250. ACM, 2018b.

David Evans, Richard Schmalensee, Michael Noel, Howard Chang, and Daniel Garcia-Swartz. *Platform economics: Essays on multi-sided businesses*. Competition Policy International, 2011. URL http://ssrn.com/abstract=1974020.

David S. Evans and Richard Schmalensee. *Matchmakers: The New Economics of Multisided Platforms*. Harvard Business Review Press, 2016.

Golnoosh Farnadi, Pigi Kouki, Spencer K Thompson, Sriram Srinivasan, and Lise Getoor. A fairness-aware hybrid recommender system. 2018. Presented at the 2nd FATRec Workshop on Responsible Recommendation held at RecSys 2018, Vancouver, CA.

Gavan J. Fitzsimons and Donald R Lehmann. Reactance to recommendations: When unsolicited advice yields contrary responses. *Marketing Science*, 23(1):82–94, 2004.

R Edward Freeman. *Strategic management: A stakeholder approach*. Cambridge university press, 2010.

Jill Freyne, Michal Jacovi, Ido Guy, and Werner Geyer. Increasing engagement through early recommender intervention. In *Proceedings of the Third ACM Conference on Recommender Systems*, RecSys '09, pages 85–92, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-435-5. doi: 10.1145/1639714.1639730. URL http://doi.acm.org/10.1145/1639714.1639730.

Inma Garcia, Laura Sebastia, and Eva Onaindia. On the design of individual and group recommender systems for tourism. *Expert systems with applications*, 38(6):7683–7692, 2011.

Florent Garcin, Boi Faltings, Olivier Donatsch, Ayar Alazzawi, Christophe Bruttin, and Amr Huber. Offline and online evaluation of news recommender systems at swissinfo.ch. In *Proc. RecSys '14*, pages 169–176, 2014.

Mustansar Ali Ghazanfar and Adam Prügel-Bennett. Leveraging clustering approaches to solve the gray-sheep users problem in recommender systems. *Expert Systems with Applications*, 41(7):3261–3275, 2014.

Kenneth E Goodpaster. Business ethics and stakeholder analysis. *Business ethics quarterly*, pages 53–73, 1991.

A. Goswami, F. Hedayati, and P. Mohapatra. Recommendation systems for markets with two sided preferences. In *2014 13th International Conference on Machine Learning and Applications*, pages 282–287, Dec 2014. doi: 10.1109/ICMLA.2014.51.

Ido Guy. People recommendation on social media. In Peter Brusilovsky and Daqing He, editors, *Social Information Access: Systems and Technologies*, pages 570–623. Springer International Publishing, Cham, 2018. ISBN 978-3-319-90092-6. doi: 10.1007/978-3-319-90092-6_15.

Ido Guy, Inbal Ronen, and Eric Wilcox. Do you know?: Recommending people to invite into your social network. In *Proceedings of the 14th International Conference on Intelligent User Interfaces*, IUI '09, pages 77–86, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-168-2. doi: 10.1145/1502650.1502664. URL http://doi.acm.org/10.1145/1502650.1502664.

Mikael Hammar, Robin Karlsson, and Bengt J. Nilsson. Using maximum coverage to optimize recommendation systems in e-commerce. In *Proc. RecSys '13*, pages 265–272, 2013.

Kartik Hosanagar, Ramayya Krishnan, and Liye Ma. Recommended for you: The impact of profit incentives on the relevance of online recommendations. In *Proc. ICIS '08*, 2008.

Ganesh Iyer, David Soberman, and J Miguel Villas-Boas. The targeting of advertising. *Marketing Science*, 24(3):461–476, 2005.

Tamas Jambor and Jun Wang. Optimizing Multiple Objectives in Collaborative Filtering. In *Proceedings of the Fourth ACM Conference on Recommender Systems*, RecSys '10, pages 55–62, New York, NY, USA, 2010. ACM.

Dietmar Jannach and Gediminas Adomavicius. Recommendations with a purpose. In *Proceedings of the 10th ACM Conference on Recommender Systems*, pages 7–10. ACM,

2016.
Dietmar Jannach and Gediminas Adomavicius. Price and profit awareness in recommender systems. In *Proceedings of the ACM RecSys 2017 Workshop on Value-Aware and Multi-Stakeholder Recommendation*, Como, Italy, 2017.
Dietmar Jannach and Kolja Hegelich. A case study on the effectiveness of recommendations in the mobile internet. In *Proc. RecSys '09*, pages 205–208, 2009. doi: 10.1145/1639714.1639749. URL http://doi.acm.org/10.1145/1639714.1639749.
Dietmar Jannach, Lukas Lerche, and Michael Jugovac. Item familiarity as a possible confounding factor in user-centric recommender systems evaluation. *i-com Journal of Interactive Media*, 14(1):29–39, 2015. doi: 10.1515/icom-2015-0018.
Yuanchun Jiang and Yezheng Liu. Optimization of online promotion: a profit-maximizing model integrating price discount and product recommendation. *International Journal of Information Technology & Decision Making*, 11(05):961–982, 2012.
T. Kamishima and S. Akaho. Personalized pricing recommender system — multi-stage epsilon-greedy approach. In *The 2nd Int'l Workshop on Information Heterogeneity and Fusion in Recommender Systems*, 2011.
T. Kamishima, S. Akaho, H. Asoh, and J. Sakuma. Efficiency improvement of neutrality-enhanced recommendation. In *The 3rd Workshop on Human Decision Making in Recommender Systems*, 2013.
T. Kamishima, S. Akaho, H. Asoh, and I. Sato. Model-based approaches for independence-enhanced recommendation. In *Proc. of the IEEE 16th Int'l Conf. on Data Mining Workshops*, pages 860–867, 2016.
T. Kamishima, S. Akaho, H. Asoh, and J. Sakuma. Recommendation independence. In *Conf. on Fairness, Accountability and Transparency*, volume 81 of *PMLR*, pages 187–201, 2018.
Toshihiro Kamishima and Shotaro Akaho. Considerations on recommendation independence for a find-good-items task. In *Workshop on Responsible Recommendation (FATRec)*, Como, Italy, 2017.
Toshihiro Kamishima, Shotaro Akaho, Hideki Asoh, and Jun Sakuma. Fairness-aware classifier with prejudice remover regularizer. In *Proc. of the ECML PKDD 2012, Part II*, pages 35–50, 2012.
Toshihiro Kamishima, Shotaro Akaho, Hideki Asoh, and Jun Sakuma. Correcting popularity bias by enhancing recommendation neutrality. In *Poster Proceedings of the 8th ACM Conference on Recommender Systems, RecSys 2014, Foster City, Silicon Valley, CA, USA, October 6-10, 2014*, 2014. URL http://ceur-ws.org/Vol-1247/recsys14_poster10.pdf.
Chen Karako and Putra Manggala. Using image fairness representations in diversity-based re-ranking for recommendations. 2018a. Presented at the 2nd FATRec Workshop on Responsible Recommendation held at RecSys 2018, Vancouver, CA.
Chen Karako and Putra Manggala. Using image fairness representations in diversity-based re-ranking for recommendations. In *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization*, pages 23–28. ACM, 2018b.
Finn Kensing and Jeanette Blomberg. Participatory design: Issues and concerns. *Computer Supported Cooperative Work (CSCW)*, 7(3-4):167–185, 1998.
Evan Kirshenbaum, George Forman, and Michael Dugan. A live comparison of methods for personalized article recommendation at Forbes.com. In *Proc. ECML/PKDD '12*, pages 51–66, 2012.
Funda Kivran-Swaine, Priya Govindan, and Mor Naaman. The impact of network structure on breaking ties in online social networks: Unfollowing on twitter. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1101–1104, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0228-9. doi: 10.1145/1978942.1979105. URL http://doi.acm.org/10.1145/1978942.1979105.
J.A. Konstan, B.N. Miller, D. Maltz, J.L. Herlocker, L.R. Gordon, and J. Riedl. GroupLens: Applying Collaborative Filtering to Usenet News. *Communications of the ACM*, 40(3): 77–7, 1997.
Jan Krasnodebski and John Dines. Considering supplier relations and monetization in designing recommendation systems. In *Proceedings of the 10th ACM Conference on Recommender Systems*, pages 381–382. ACM, 2016.

Haewoon Kwak, Hyunwoo Chun, and Sue Moon. Fragile online relationship: A first look at unfollow dynamics in twitter. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1091–1100, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0228-9. doi: 10.1145/1978942.1979104. URL http://doi.acm.org/10.1145/1978942.1979104.

Neal Leavitt. Recommendation technology: Will it boost e-commerce? *Computer*, 39(5): 13–16, 2006.

Eric L Lee, Jing-Kai Lou, Wei-Ming Chen, Yen-Chi Chen, Shou-De Lin, Yen-Sheng Chiang, and Kuan-Ta Chen. Fairness-aware loan recommendation for microfinance services. In *Proceedings of the 2014 International Conference on Social Computing*, page 3. ACM, 2014.

Lei Li and Tao Li. Meet: A generalized framework for reciprocal recommender systems. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, CIKM '12, pages 35–44, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1156-4. doi: 10.1145/2396761.2396770. URL http://doi.acm.org/10.1145/2396761.2396770.

Yuzhong Lin, Joran Jessurun, Bauke De Vries, and Harry Timmermans. Motivate: Towards context-aware recommendation mobile system for healthy living. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2011 5th International Conference on*, pages 250–253. IEEE, 2011.

Duen-Ren Liu and Ya-Yueh Shih. Hybrid approaches to product recommendation based on customer lifetime value and purchase preferences. *Journal of Systems and Software*, 77 (2):181 – 191, 2005.

Weiwen Liu and Robin Burke. Personalizing fairness-aware re-ranking. *arXiv preprint arXiv:1809.02921*, 2018. Presented at the 2nd FATRec Workshop held at RecSys 2018, Vancouver, CA.

Weiwen Liu, Jun Guo, Nasim Sonboli, Robin Burke, and Shengyu Zhang. Personalized fairness-aware re-ranking for microlending. In *Proceedings of the 13th ACM Conference on Recommender Systems*, pages 467–471. ACM, 2019.

Giseli Rabello Lopes, Mirella M Moro, Leandro Krug Wives, and José Palazzo Moreira De Oliveira. Collaboration recommendation on academic social networks. In *International Conference on Conceptual Modeling*, pages 190–199. Springer, 2010.

Wei Lu, Shanshan Chen, Keqian Li, and Laks V. S. Lakshmanan. Show me the money: Dynamic recommendations for revenue maximization. *Proc. VLDB Endow.*, 7(14):1785–1796, 2014.

Judith Masthoff. Group recommender systems: Combining individual models. In *Recommender systems handbook*, pages 677–702. Springer, 2011.

Rishabh Mehrotra, James McInerney, Hugues Bouchard, Mounia Lalmas, and Fernando Diaz. Towards a fair marketplace: Counterfactual evaluation of the trade-off between relevance, fairness & satisfaction in recommendation systems. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 2243–2251. ACM, 2018.

Aranyak Mehta, Amin Saberi, Umesh Vazirani, and Vijay Vazirani. Adwords and generalized online matching. *Journal of the ACM*, 54(5):22, 2007.

Hervé Moulin. *Fair division and collective welfare.* MIT press, 2004.

Phong Nguyen, John Dines, and Jan Krasnodebski. A multi-objective learning to re-rank approach to optimize online marketplaces for multiple stakeholders. *arXiv preprint arXiv:1708.00651*, 2017.

Gal Oestreicher-Singer and Arun Sundararajan. The visible hand? demand effects of recommendation networks in electronic markets. *Manage. Sci.*, 58(11):1963–1981, 2012.

Umberto Panniello, Shawndra Hill, and Michele Gorgoglione. The impact of profit incentives on the relevance of online recommendations. *Electronic Commerce Research and Applications*, 20:87 – 104, 2016.

Bhavik Pathak, Robert Garfinkel, Ram D Gopal, Rajkumar Venkatesan, and Fang Yin. Empirical analysis of the impact of recommender systems on sales. *Journal of Management Information Systems*, 27(2):159–188, 2010.

Luiz Pizzato, Tomek Rej, Thomas Chung, Irena Koprinska, and Judy Kay. RECON: a reciprocal recommender for online dating. In *RecSys '10: Proceedings of the fourth*

ACM conference on Recommender systems, pages 207–214, New York, NY, USA, 2010a. ACM. ISBN 978-1-60558-906-0.

Luiz Pizzato, Tomek Rej, Thomas Chung, Irena Koprinska, and Judy Kay. Recon: a reciprocal recommender for online dating. In *Proceedings of the fourth ACM conference on Recommender systems*, pages 207–214. ACM, 2010b.

Victor Ponce, Jean-Pierre Deschamps, Louis-Philippe Giroux, Farzad Salehi, and Bessam Abdulrazak. Quefaire: context-aware in-person social activity recommendation system for active aging. In *Inclusive Smart Cities and e-Health*, pages 64–75. Springer, 2015.

Pearl Pu, Li Chen, and Rong Hu. A user-centric evaluation framework for recommender systems. In *Proceedings of the fifth ACM conference on Recommender systems*, pages 157–164. ACM, 2011.

Jean-Charles Rochet and Jean Tirole. Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4):990–1029, 2003.

Mario Rodriguez, Christian Posse, and Ethan Zhang. Multiple objective optimization in recommender systems. In *Proceedings of the sixth ACM conference on Recommender systems*, pages 11–18. ACM, 2012a.

Mario Rodriguez, Christian Posse, and Ethan Zhang. Multiple objective optimization in recommender systems. In *Proceedings of the sixth ACM conference on Recommender systems*, pages 11–18. ACM, 2012b.

Alvin E Roth and Marilda Sotomayor. Two-sided matching. *Handbook of game theory with economic applications*, 1:485–541, 1992.

Oguz Semerci, Alois Gruson, Catherinee Edwards, Ben Lacker, Clay Gibson, and Vladan Radosavljevic. Homepage personalization at spotify. In *Proceedings of the 13th ACM Conference on Recommender Systems*, pages 527–527. ACM, 2019.

U. Shardanand and P. Maes. Social information filtering: algorithms for automating "word of mouth". In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 210–217. ACM Press/Addison-Wesley Publishing Co., 1995.

Ya-Yueh Shih and Duen-Ren Liu. Product recommendation approaches: Collaborative filtering via customer lifetime value and customer demands. *Expert. Syst. App.*, 35(1-2): 350–360, 2008.

Ashudeep Singh and Thorsten Joachims. Fairness of exposure in rankings. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2219–2228. ACM, 2018.

Barry Smyth and Paul McClave. Similarity vs. diversity. In *Case-Based Reasoning Research and Development*, pages 347–361. Springer, 2001.

E. Steel and J. Angwin. On the web's cutting edge, anonymity in name only. The Wall Street Journal, 2010. ⟨http://on.wsj.com/aimKCw⟩.

Özge Sürer, Robin Burke, and Edward C Malthouse. Multistakeholder recommendation with provider constraints. In *Proceedings of the 12th ACM Conference on Recommender Systems*, pages 54–62. ACM, 2018.

Krysta M. Svore, Maksims N. Volkovs, and Christopher J.C. Burges. Learning to Rank with Multiple Objective Functions. In *Proceedings of the 20th International Conference on World Wide Web*, WWW '11, pages 367–376, New York, NY, USA, 2011. ACM.

Jie Tang, Sen Wu, Jimeng Sun, and Hang Su. Cross-domain collaboration recommendation. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1285–1293. ACM, 2012.

Loren Terveen and David W. McDonald. Social matching: A framework and research agenda. *ACM Trans. Comput.-Hum. Interact.*, 12(3):401–434, September 2005. ISSN 1073-0516. doi: 10.1145/1096737.1096740. URL http://doi.acm.org/10.1145/1096737.1096740.

Nava Tintarev, Emily Sullivan, Dror Guldin, Sihang Qiu, and Daan Odjik. Same, same, but different: Algorithmic diversification of viewpoints in news. In *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization*, pages 7–13. ACM, 2018.

Saúl Vargas and Pablo Castells. Rank and relevance in novelty and diversity metrics for recommender systems. In *Proceedings of the fifth ACM conference on Recommender systems*, pages 109–116. ACM, 2011.

Hsiao-Fan Wang and Cheng-Ting Wu. A mathematical model for product selection strategies in a recommender system. *Expert Systems with Applications*, 36(3, Part 2):7299 – 7308, 2009.

Peng Xia, Benyuan Liu, Yizhou Sun, and Cindy Chen. Reciprocal recommendation system for online dating. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*, pages 234–241. ACM, 2015.

Peng Xia, Shuangfei Zhai, Benyuan Liu, Yizhou Sun, and Cindy Chen. Design of reciprocal recommendation systems for online dating. *Social Network Analysis and Mining*, 6(1): 32, Jun 2016. ISSN 1869-5469. doi: 10.1007/s13278-016-0340-2. URL `https://doi.org/10.1007/s13278-016-0340-2`.

Bo Xu, Yun Huang, Haewoon Kwak, and Noshir Contractor. Structures of broken ties: Exploring unfollow behavior on twitter. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, CSCW '13, pages 871–876, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-1331-5. doi: 10.1145/2441776.2441875. URL `http://doi.acm.org/10.1145/2441776.2441875`.

S. Yao and B. Huang. Beyond parity: Fairness objectives for collaborative filtering. In *Advances in Neural Information Processing Systems 30*, 2017a.

Sirui Yao and Bert Huang. New fairness metrics for recommendation that embrace differences. In *Workshop on Fairness, Accountability and Transparency in Machine Learning (FATML)*, page 5 pages, Halifax, Nova Scotia, 2017b.

M. Yu, X. Zhang, and D. Kreager. New to online dating? learning from experienced users for a successful match. In *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 467–470, Aug 2016. doi: 10.1109/ASONAM.2016.7752276.

Shuai Yuan, Ahmad Zainal Abidin, Marc Sloan, and Jun Wang. Internet advertising: An interplay among advertisers, online publishers, ad exchanges and web users. *arXiv preprint arXiv:1206.1754*, 2012.

M. Zehlike, F. Bonchi, C. Castillo, S. Hajian, M. Megahed, and R. Baeza-Yates. Fa∗ir: A fair top-k ranking algorithm. In *Proc. of the 25th ACM Conf. on Information and Knowledge Management*, 2017a.

Meike Zehlike, Francesco Bonchi, Carlos Castillo, Sara Hajian, Mohamed Megahed, and Ricardo Baeza-Yates. Fa*ir: A fair top-k ranking algorithm. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pages 1569–1578. ACM, 2017b.

Yongfeng Zhang and Xu Chen. Explainable recommendation: A survey and new perspectives. *arXiv preprint arXiv:1804.11192*, 2018.

Yong Zheng, Nastaran Ghane, and Milad Sabouri. Personalized educational learning with multi-stakeholder optimizations. In *Adjunct Proceedings of the ACM conference on User Modelling, Adaptation and Personalization. ACM*, 2019.

Ziwei Zhu, Xia Hu, and James Caverlee. Fairness-aware tensor-based recommendation. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 1153–1162. ACM, 2018.

Cai-Nicolas Ziegler, Sean M McNee, Joseph A Konstan, and Georg Lausen. Improving recommendation lists through topic diversification. In *Proceedings of the 14th international conference on World Wide Web*, pages 22–32. ACM, 2005.

Himan Abdollahpouri is a Ph.D. candidate in Information Science at the University of Colorado Boulder. He was a pioneer in research on multistakeholder recommendation and has also investigated popularity bias in recommendation and its impact on different recommendation stakeholders. His work has appeared in leading conferences including RecSys, UMAP, and WWW.

Gediminas Adomavicius is a professor and department chair of information and decision sciences at the University of Minnesota. His general research interests revolve around computational techniques for aiding decision-making in information-intensive environments and include personalization technologies and recommender systems, machine learning and data mining, and electronic market mechanisms. His work has been published in a number of leading academic journals and conferences in information systems and computer science and has received several grants from major funding institutions, including the U.S. National Science Foundation CAREER award for his research on recommender systems.

Robin Burke is a professor and chair of the Information Science department at the University of Colorado, Boulder. His research group explores fairness, accountability and transparency in recommendation through the integration of objectives from diverse stakeholders. Professor Burke is the author of more than a hundred peer-reviewed articles in various areas of artificial intelligence including recommender systems, machine learning and information retrieval. His work has received support from the National Science Foundation and the National Endowment for the Humanities, among others.

Ido Guy is a Director of Applied Research and Chief Scientist at eBay Israel, leading teams practicing artificial intelligence to solve e-commerce problems, from generation of product titles and descriptions to better understanding of e-commerce search queries. Prior to eBay, Ido has been with Yahoo Labs and IBM Research. Ido's work has focused on social media mining and analysis, web search and question answering, and recommender systems. In the latter field, he has specialised in social recommender systems and, in particular, people recommendation. Ido has published over 70 research papers in leading conferences and journals in the areas of information retrieval, recommender systems, computer-human interaction, knowledge management, web data mining, and others.

Dietmar Jannach is a Professor of Computer Science at AAU Klagenfurt, Austria, and head of the department's information systems research group. Dr. Jannach has worked on different areas of artificial intelligence, including recommender systems, model-based diagnosis, and knowledge-based systems. He is the leading author of a textbook on recommender systems and has authored more than hundred research papers, focusing on the application of artificial intelligence technology to practical problems.

Toshihiro Kamishima is a senior researcher of the National Institute of Advanced Industrial Science and Technology (AIST), Japan. His research interests are recommender systems, data mining, and machine learning. In particular, he currently tackles with the problem of fairness in recommendation and machine learning.

Jan Krasnodebski is Senior Director of Data Science, Marketplace and Yield, for Expedia Group. He leads the data science team responsible for the customer lodging booking journey, from type-ahead suggestions, geography, destination recommendation to real-time personalized recommendation of specific lodging options. His team also develops hotel revenue management and air pricing algorithms. A particular focus of his work, which he has published on, is multi-stakeholder recommendations that combine the interests of consumers and suppliers along with the profitability of the platform where they interact.

Luiz Pizzato is an expert data scientist with nearly 20 years of experience working with Machine Learning and Artificial Intelligence. He has a PhD and an MSc in Computer Science both with a focus on Natural Language Processing. During his career Luiz has built commercial AI systems for matching people in online dating platforms as well as for finding the best candidates for job positions. Luiz also worked on complex NLP problems for some of the large financial institutions in the world and worked on AI solutions to improve the banking experience for businesses in Australia. Currently, Luiz leads Accenture's Liquid Studio Sydney team, which is a highly specialised group of engineers focusing on Data Science and Robotics. Luiz's work involves cutting edge and emerging technologies in AI including eXplainable AI, Homomorphically Encrypted Machine Learning and other innovations in the robotics space.