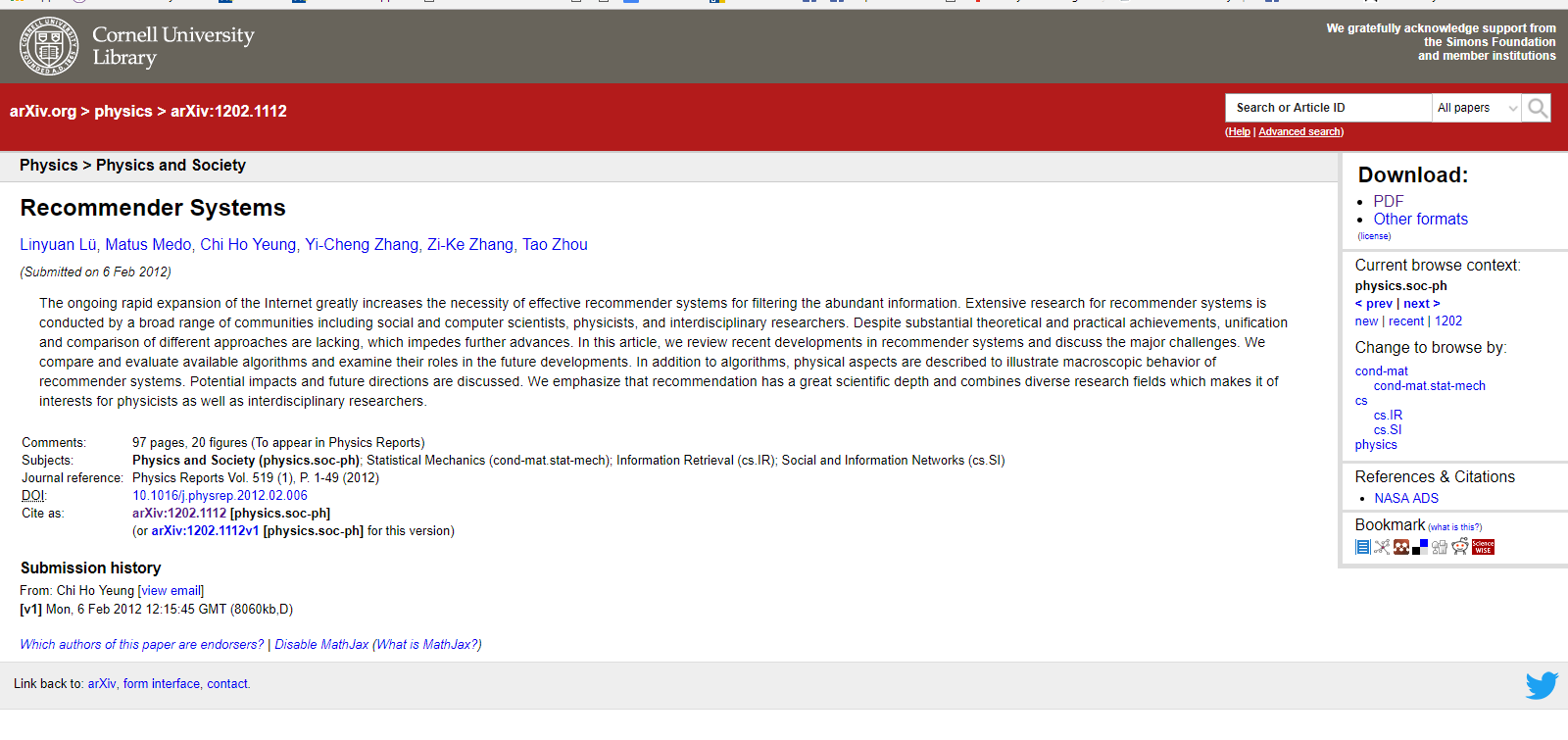
**Research on Recommender Systems**

**Source #1**



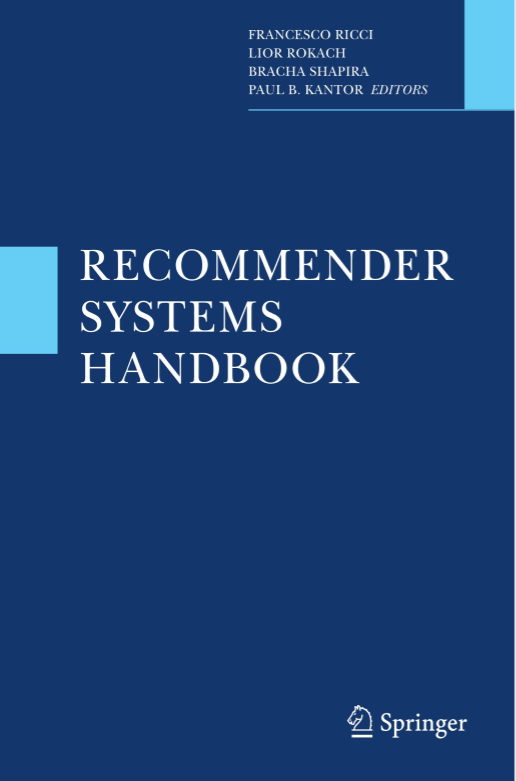
**Intro to Recommender Systems**

* Since most of our interactions are online/electronic, it has become easier for researchers to study socio-economical and techno-social systems at much greater level of detail.
* The task of recommender systems is to turn data on users and their preferences into predictions of users’ possible future likes and interests. The study of recommender systems is at crossroads of science and socio-economic life and its huge potential was first noticed by web entrepreneurs in the forefront of the information revolution.
* When computing recommendations, the very basic approach is to select the objects favored by other users that are similar to the target user -🡪 Known as user similarity
* Many issues arise with recommendations because there aren’t “first principles” to follow. E.g. what if user has little information?
* In consequence, similarly to physics, it is the experiment what decides which recommendation approach is good and which is not.
* Most e-commerce web sites now offer various forms of recommendation—ranging from simply showing the most popular items or suggesting other products by the same producer to complicated data mining techniques
* People soon realized that there is no unique best recommendation method. Rather, depending on the context and density of the available data, different methods adapting to particular applications are most likely to succeed. Hence there is no panacea, and the best one can do is to understand the underlying premises and recommender mechanisms, then one can tackle many diverse application problems from the real life examples. This is also reflected in this review where we do not try to highlight any ultimate approach to recommendation. Instead, we review the basic ideas, methods and tools with particular emphasis on physics-rooted approaches.
* While the availability of data is important for empirical evaluation of recommendation methods, the main driving force comes from practice: electronic systems give us too much choice to handle by ourselves.

**Real Applications of Recommender Systems**

* Thanks to the ever-decreasing costs of data storage and processing, recommender systems gradually spread to most areas of our lives. Sellers carefully watch our purchases to recommend us other goods and enhance their sales, social web sites analyze our contacts to help us connect with new friends and get hooked with the site, and online radio stations remember skipped songs to serve us better in the future (see more examples in Table 1). In general, whenever there is plenty of diverse products and customers are not alike, personalized recommendation may help to deliver the right content to the right person. This is particularly the case for those Internet-based companies that try to make use of the so-called long-tail [16] of goods which are rarely purchased yet due to their multitude they can yield considerable profits (sometimes they are referred to as “worst-sellers”). For ex3 ample on Amazon, between 20 to 40 percent of sales is due to products that do not belong to the shop’s 100 000 most saled products [17]. A recommender system may hence have significant impact on a company’s revenues: for example, 60% of DVDs rented by Netflix are selected based on personalized recommendations.
* As discussed in [18], recommender systems not only help decide which products to offer to an individual customer, they also increase cross-sell by suggesting additional products to the customers and improve consumer loyalty because consumers tend to return to the sites that best serve their needs
* Since no recommendation method serves best all customers, major sites are usually equipped with several distinct recommendation techniques ranging from simple popularitybased recommendations to sophisticated techniques many of which we shall encounter in the following sections. Further, new companies emerge (see, for example, string.com) which aim at collecting all sorts of user behavior (ranging from pages visited on the web and music listened on a personal player to “liking” or purchasing items) and using it to provide personalized recommendations of different goods or services.

**Source #2**



* Why service providers want to use RS include: increase number of items sold, sell more diverse items, increase user satisfaction, increase user fidelity and better understand what the user wants.
* Since the data and knowledge sources available for recommender systems can be very diverse, whether RS can be exploited depends on the technique.
* In general, there are recommendation techniques that are knowledge poor, i.e., they use very simple and basic data, such as user ratings/evaluations for items (Chapters 5, 4). Other techniques are much more knowledge dependent, e.g., using ontological descriptions of the users or the items (Chapter 3), or constraints (Chapter 6), or social relations and activities of the users (Chapter 19). In any case, as a general classiﬁcation, data used by RSs refers to three kinds of objects: items, users, and transactions, i.e., relations between users and items.
* Items = objects that are recommended. Item is positive if user found it useful and negative if the user made a wrong decision selecting it. We note that when a user is acquiring an item she will always incur in a cost, which includes the cognitive cost of searching for the item and the real monetary cost eventually paid for the item.

**Recommendation Techniques—Pg 38**

* In order to implement its core function, identifying the useful items for the user, a RS must predict that an item is worth recommending. To illustrate the prediction step of a RS, consider, for instance, a simple, nonpersonalized, recommendation algorithm that recommends just the most popular songs. The rationale for using this approach is that in absence of more precise information about the user’s preferences, a popular song, i.e., something that is liked (high utility) by many users, will also be probably liked by a generic user, at least more than another randomly selected song. Hence the utility of these popular songs is predicted to be reasonably high for this generic user.
* Itisalso importanttonotethatsometimestheuser utilityforan itemis observed to depend on other variables, which we generically call “contextual” [1]. For instance, the utility of an item for a user can be inﬂuenced by the domain knowledge of the user (e.g., expert vs. beginning users of a digital camera), or can depend on the time when the recommendation is requested. Or the user may be more interested in items (e.g., a restaurant) closer to his current location. Consequently, the recommendationsmustbeadaptedtothesespeciﬁcadditionaldetailsandasaresult it becomes harder and harder to correctly estimate what the right recommendations are.
* The six different types of Recommender Systems are:
  + **Content-based:** The system recommends items that are similar to the ones the user liked in the past. I.e. if they liked a comedy movie before then it recommends movies from the comedy genre.
  + **Collaborative filtering:** The simplest and original implementation of this approach [93] recommends to the active user the items that other users with similar tastes liked in the past. The similarity in taste of two users is calculated based on the similarity in the rating history of the users. This is the reason why [94] refers tocollaborativeﬁlteringas“people-to-peoplecorrelation.”Collaborativeﬁlteringis considered to be the most popular and widely implemented technique in RS.
  + **Demographic:** This type of system recommends items based on the demographic proﬁle of the user.The assumption is that different recommendations should be generated for different demographic niches. Many Web sites adopt simple and effective personalization solutions based on demographics. For example, users are dispatched to particular Web sites based on their language or country. Or suggestions may be customized according to the age of the user. While these approaches have been quite popular in the marketing literature, there has been relatively little proper RS research into demographic systems.
  + **Knowledge-based**: domain knowledge about how certain item features meet users needs and preferences and, ultimately, how the item is useful for the user. In these systems a similarity function estimates how much the user needs (problem description) match the recommendations (solutions of the problem).
  + **Community-based:** This type of system recommends items based on the preferences of the users friends.
  + **Hybridrecommendersystems**:These RSs are based on the combination of the above mentioned techniques. A hybrid system combining techniques A and B tries to use the advantages of A to ﬁx the disadvantages of B

Collaborative Filtering:

https://blog.dominodatalab.com/recommender-systems-collaborative-filtering/

Finding similarity:

* Shortest distance between 2 points you can use the Euclidean distance.
* When we compute similarity, we are going to calculate it as a measure of "anti-distance". The higher the distance between two objects, the more "farther apart" they are. On the other hand, the higher the similarity between two objects, the more "closer together" they are. Usually similarity metrics return a value between 0 and 1, where 0 signifies no similarity (entirely dissimilar) and 1 signifies total similarity (they are exactly the same).
* Cosine Similarity
  + Looking at two vectors, they are only similar if the angle between them is zero.
  + This type of similarity is a measure of orientation and not of similarity.
  + Just cuz similarity score is 1, doesn’t mean that they are the same.
* Model-based Collaborative filtering
  + Suppose you have an n X m matrix. Each element of the matrix(I,j) represents how user I rated item j.
* About the surprise library code:
  + Store the MovieLens 100k data in an user-interaction matrix.