The Use of Recommender Systems in Demand Management in Intelligent Supply Chain Management

Qi Zhang

A research paper submitted to University of Dublin,

in partial fulfilment of the requirements for the degree of

Masters of Science Interactive Digital Media

2018

Declaration

I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at: http://www.tcd.ie/calendar

I have also completed the Online Tutorial on avoiding plagiarism 'Ready, Steady, Write', located at http://tcd-ie.libguides.com/plagiarism/ready-steady-write

I declare that the work described in this research paper is, except where otherwise stated, entirely my own work and has not been submitted as an exercise for a degree at this or any other university.

Signed: _____

Qi Zhang

10/05/2018

Permission to lend and/or copy

I agree that Trinity College Library may lend or copy this research paper upon request.

Signed: _____

Qi Zhang

10/05/2018

Acknowledgments

I would sincerely like to thank my supervisor, Fatemeh Golpayegani,
for all of her guidance and support throughout the process
of completing this paper.

Summary

With the rapid development of internet technology, electronic commerce (e-commerce) has become a new development direction of the application of internet technology. E-commerce has the characteristics of opening, global, offering low cost and high efficiency services. The traditional supply chain cannot match the speed of the development of e-commerce, as a result, the intelligent supply chain management shows up and plays an important role in e-commerce. It not only uses internet as a method of communication, but also deeply excavate the potential application of internet throughout the whole process of e-commerce, such as strategic planning, inventory management, demand management and so on, by using electronic data interchange technology (EDI), inventory management technology, and recommender systems. In intelligent supply chain management, demand management predicts the demand using collected information from users' behavior. For this reason, precise collection and analysis of information are quite necessary. This can be achieved by recommender systems, as they can collect data from users' historical behavior and feedback, and then identify and recommend product or content to users that they are looking for in the front end. It is based on user interest and offers personalized services to users. With the help of recommender systems, the intelligent supply chain in e-commerce process can achieve a great improvement in efficiency. So, it is an important choice for the e-commerce-based enterprises to improve the sales ability of themselves and the loyalty from the users.

This paper focuses on demand management in supply chain management, analyzing how recommender systems can affect the intelligent supply chain. Cases as Amazon, Alibaba are used to present explanations and make contrastive analysis. Based on theoretical exposition and case analysis, this paper discusses the concept of the recommender systems and their trend of development in the recent years, the problems appearing, and the potential solutions to the problems are put forward.

Table of Contents:

List of Figures

List of Tables

Chapter 1 - Introduction

The internet technology has been developing rapidly to web2.0 period in recent years. The concept of web2.0 was popularized by Tim O'Reilly and Dale Dougherty at the O'Reilly Media Web 2.0 Conference in late 2004 (Tim O'Reilly,2004). Compared with web1.0, web2.0 period focuses more on the interaction between computers and users, changed the way of communication, and enhanced the participation and spontaneity of all the subscribers. This is beneficial in different parts of our lives including business, traveling, entertainment, and social communication. One of the production of web2.0 is electronic commerce (e-commerce), which is a new development direction of the application of internet technology. With the similar characteristics with web2.0, e-commerce has the feature of opening, global, low cost and high efficiency, so it provides a vast space for traditional business activities, which is incomparable to the traditional channels. In this case, it has been growing rapidly worldwide.

However, developing at a high speed, the traditional mode of supply chain can hardly match its requirement of efficiency. Because none of the activities in e-commerce can be separated from internet platform, but for the traditional supply chain system, the use of internet is only at the level of information communication in the front end, and it takes a huge amount of manual cost in other links, such as warehousing, transportation and so on. To make e-commerce as a whole system, and deeply excavate the potential application of internet, it is necessary to bring in intelligent supply chain management in the process.

Supply chain is a logistics network composed of suppliers, manufacturers, warehouses, distribution centers and channel providers. There are 8 links in the supply chain management, which are customer relationship management, customer service management, demand management style, order fulfillment, manufacturing flow management, supplier relationship management, product development and commercialization, and returns management. Each link has the same target, which is to reduce cost, cut down time, so that to obtain more benefits, by making suitable plan for their supply chain. Among these links, the demand management is a key point that has a close connection with other links and makes use of various statistical

techniques to make predictions and forecasts based on historical data. It has a great influence on the two factors of revenue and inventory, and these two are the key factors to determine whether the supply chain is effective (Bisk). In another word, as an important part of e-commerce, the intelligent supply chain system can bring about improvement and optimization through improving the effectiveness of demand management. However, the historical data that demand management has to be based on is not easy to be collected and analyzed precisely and timely.

In recent years, the production and consumption of information have greatly promoted the development of information society, but information overload, information explosion and other phenomena make it very difficult for users to extract interesting and useful information in a timely and effective way. From the perspective of information providers, it is difficult too, because of the countless advertisements and other information that is not associated with them, it is difficult to transmit information to recipients quickly and directly. In view of this phenomenon, relying on modern information technology, there are two modes that can provide users with the information they need, which are push and pull media mode. Pull media is a passive way to deliver message to receiver, such as traditional way like newspaper, television or radio, while push media is used to push information to users actively, to reduce the time of blind search on the internet. Recommender systems are one of the effective ways of information push mode. On the basis of analyzing and predicting the user's needs, it recommends the potential interest that sometimes not quite easy to be obtained to the user.

In addition, from users' perspective, recommender systems can offer personalized recommendation service for them, based on gathering and exploring the information of user's behavior, interest and environment. Joe Pine argues in his book Mass Customization that, the old world of mass production, where "standardized products, homogeneous markets, and long product life and development cycles were the rule", should shifted to the new world, where "variety and customization supplant standardized products." That means, it is not adequate anymore to have only one product, instead, companies should develop at least multiple products that meet the needs of multiple customers (Joseph Pine B., 1993). Using recommender systems may help companies to achieve the purpose of meeting multiple needs of users so that

to firmly establish the relationship between users and enhance customer loyalty. Because of the above advantages of recommender systems, the prevalence and research attention rate of recommender systems have been keeping in a high degree, and such system has been used in a variety of aspects, including commercial or products in general, social tags, financial services. This paper will focus on recommender systems used in e-commerce, that how it can affect the process of demand management in intelligent supply chain management.

The second chapter in this paper reviews the literature and introduction of the theoretical background of recommender systems, then focus on the discussion of the functions of different types of recommender systems, the operating mechanism, and the problems have already appeared or will show up potentially. In the third chapter, there are cases of Amazon and Alibaba to present different type of recommender systems using in e-commerce. The directions of future studies are talking in the next parts, followed with the conclusion.

Chapter 2 - Literature review

2.1 E-commerce

This part reviews the definition of e-commerce, technologies used in it, and features and development of e-commerce.

2.1.1 Definition of e-commerce

E-commerce (electronic commerce), is the activity of buying or selling of products and services online or over the internet or Value-Added Network (VAN). E-commerce is the electronation and networking of all aspects of traditional business activities. IBM introduced this concept in 1996, and by 1997, the company also proposed the concept of Electronic Business. Based on different concept, e-commerce is divided into broad sense and narrow sense. The narrow sense e-commerce is defined as the main use of internet to engage in business or relative activities. In the broad sense e-commerce, which is so called electronic business or EB (E-Business), refers to the whole business activities. It means the application of computer, internet technology

and modern information communication technology, in accordance with certain standards, with the use of electronic tools (including broadcasting, television, electricity, telephone communication, etc.), to achieve the whole process of business transactions and administrative operations, including electronic commerce or electronic transactions. Whether in the broad sense or the narrow sense of the concept of electronic commerce, it covers two aspects: one is that any activities in e-commerce cannot be separated from internet platform; the second one is that the activities operating through the internet belong to business activities. In general, e-commerce refers to the narrow sense.

2.1.2 Technologies used in e-commerce

There are four main business models of e-commerce, which are B2B (business to business), B2C (business to customer), C2C (consumer to consumer), B2M (business to manager). In the process of all kinds of models of e-commerce, the information technologies used include: internet, extranet, e-mail, database, electronic directory, mobile phone and so on. There are three main aspects in e-commerce, which are information service, transaction and payment, including electronic commerce advertisement, electronic purchase and transaction, electronic transaction voucher exchange, electronic payment and settlement, and after-sales online service. In order to operate the three aspects, the management of e-commerce includes electronic funds transfer (EFT), supply chain management, mobile commerce, electronic trading market, internet marketing (or online advertising), online transaction processing (OLTP), electronic data exchange (EDI), inventory management system and automatic data collection system.

2.1.3 Features and development of e-commerce

With the help of the above technologies, e-commerce has become a new development direction of the application of internet technology. The internet itself has the characteristics of opening, global, low-cost and high-efficiency, which has also become the intrinsic features of e-commerce, and make e-commerce greatly surpass its value as a new form of trade. The electronic technology platform based on the internet provides a vast space for traditional business activities, and its outstanding superiority is incomparable by the traditional media

means. First of all, e-commerce makes traditional business online and digital, as a result, a huge range of real logistics is instead by electronic flow, which reduces a large number of manual labour, material resources and cost to a large extent, and the less limitation of time and space restrictions make it possible to make transactions at any time or any place, thus greatly improving the efficiency. Secondly, the characteristics of openness and globality of e-commerce have created more trade opportunities and even face-to-face communication opportunities for enterprises or individuals around the world. Thirdly, e-commerce enables enterprises to enter the global electronic market with similar costs, making it possible for small and medium-sized enterprises (SMEs) to have the similar information resources as large enterprises and improve the competitiveness of them. In addition, the reduction of intermediate links makes it possible for producers and consumers to make direct transactions, thus changing the way of the whole social economy to a certain extent. Moreover, the Internet is turning the whole earth into a global village, because of the speed of the transmission of information, and the strong interaction that greatly improves the efficiency and convenience for consumers to provide feedback to the seller. Because of the above advantages of e-commerce, it has been growing rapidly worldwide. In the last decades, it has a growth of 18% each year (Mackenzie Ian et al, 2013). E-commerce statistics confirm the explosive pace at which this industry has developed as worldwide B2C e-commerce sales amounted to more than 1.2 trillion US dollars in 2013[1]. According to popular e-commerce market data, US-founded Amazon is one of the leading e-commerce platforms worldwide. Asian competitors such as Rakuten or Alibaba are also constantly expanding their share within the B2C e-commerce market. Online auction website eBay is the most popular example for C2C e-commerce whilst also providing a platform for merchants to sell their goods.

2.2 Intelligent supply chain management

In this part, the definition and trend of development of supply chain management (SCM) are reviewed. After that, it talks about the superiority of intelligent supply chain management.

[1] Statistics and Market Data about E-commerce. https://www.statista.com/markets/413/e-commerce/

2.2.1 Definition of SCM

Supply chain management (SCM) is the management of the flow of goods and services[2].The original definition came in 1982 when Keith Oliver, a consultant at Booz Allen Hamilton (now Strategy), introduced the term "supply chain management" to the public domain in an interview for the Financial Times (Robert B. 1999). After developing for many years, it has been defined as the "design, planning, execution, control, and monitoring of supply chain activities with the objective of creating net value, building a competitive infrastructure, leveraging worldwide logistics, synchronizing supply with demand and measuring performance globally."[3] The definition shows that SCM plays an important role in most aspects of business activities, from the raw materials to finished goods (Duncan McFarlane& Yossi Sheffi, 2003). There are 8 links in supply chain management (Lambert, 2004), which are customer relationship management, customer service management, demand management style, order fulfillment, manufacturing flow management, supplier relationship management, product development and commercialization, returns management. One single enterprise can operate one link or several links at the same time, as long as the collaboration between the links can be managed well. Each link has the same target, which is to reduce cost and time by improving trust and collaboration, to obtain more benefits, to improving inventory visibility and the velocity of inventory movement.

2.2.2 Trend of development

In the web2.0 period, the new business environment has contributed to the changes in the supply chain systems. First of all, because of the globalization and high technologies, multinational companies and strategic alliances can expand their assembly lines throughout the world, thus to make the manufacturing more effective, lean, and agile. Secondly, the reduction of communication costs leads to the improvement of cooperation among the links of supply chain (Coase, 1998). Based on these changes, the term "SCM 2.0" has been proposed to

[2] Association of Employment and Learning Providers' Supply Chain Management Guide at aelp.org.uk published 2013, accessed 31 March 2015

[3] "supply chain management (SCM)". APICS Dictionary. Retrieved 2016-07-19.

describe the new type of management. It indicates the improvements of supply chain itself, as well as the evolution of processes, methods, and tools used in it.

However, the above developments are in the areas of manufacturing, or the collaboration using the newly developed communication technology, but more advanced changes are still needed. With the continuous development of e-commerce, there have been two kinds of challenges for the enterprises, the external challenge and the internal challenge. The external challenge means the competition among enterprises is becoming more and more intense; while the internal one is to describe the phenomenon that when determining the amount, type, or other sources in production process, lots of errors appear because it is becoming harder to predict the manufacturing plan. On top of that, enterprises need to be transformed and upgraded the SCM, to match the high speed of e-commerce developing. Martin Christopher, Emeritus Professor of Marketing and Logistics at Cranefield School of Management, further emphasized the importance of supply chain that "The competition in twenty-first Century is not the competition between enterprises, but the competition between supply chains" (Martin Christopher, 2016).

2.2.3 Superiority of intelligent supply chain management

With the help of internet technology, building a smart and efficient supply chain has become the key for enterprises to gain advantages in market competition, that is, the intelligent supply chain system. As saying in the book of "Encyclopedia of Information Science and Technology", the intelligent supply chain is smart enough to predict or anticipate customer demands, sense changes in supply and demand, and quickly respond to unplanned events. It can use information and quicker decision-making as a substitute for additional inventory (Mehdi Khosrow-Pour, 2017). Intelligent supply chain management is a typical production in the Industry 4.0 period, which is the name for the current trend of automation and data exchange in manufacturing technologies, including cyber-physical systems, cloud computing and cognitive computing (Kagermann, 2013).

Compared to the traditional supply chain, the intelligent supply chain management focuses more on marketing, technology and service, showing 5 significant characteristics.

- The first one is global optimization, through the automation and intelligence of all links, improve the effect of the entire supply chain.
- The second one is to emphasize the information sharing and cooperation between the customers and suppliers, to truly realize the demand planning through the demand perception;
- The third one is pay more attention to improving the accuracy and effectiveness of customer service, and promoting the iterative upgrading of products and services;
- The forth one is emphasizing more on the platform function, which involves multi-aspect like the product life cycle, market, supplier, process, information and so on;
- The fifth one is to attach importance to lean manufacturing based on the whole value chain, from lean production to lean logistics, lean procurement and lean distribution.

In another word, the intelligent supply chain management is no longer management of a single person or a single department but making plan and digitalizing with the entire supply chain.

Intelligent supply chain management has already been adopted by most companies involved in e-commerce. For instance, IBM's Watson artificial intelligence system has established a transparent, cooperative and highly predictive intelligent supply chain management system with the utilize of cognitive techniques, and serves many enterprises around the world. Amazon opens up a new ground of using "Prime Air" unmanned aerial vehicle (UAV) instead of manual delivery, and with the use of cloud technology, machine learning and big data analysis, it has achieved automatic forecast, procurement, replenishment, warehousing to address customer demand. Some leading companies in China like Jingdong, Alibaba, Kingdee, have also deployed the supply chain cloud, set up global logistic network, preparing for upgrading from traditional business modes.

2.3 Demand management

Demand management is a notion that can be used in many areas. It can be at macro-level, that government use it to control aggregate demand to avoid a recession (Sexton, Robert& Fortura, Peter, 2005). In natural resources level, it focuses on polices that control the assumption of

environmentally sensitive or harmful goods. In welfare economic level, it refers to optimal allocation resources to affect social welfare to achieve better overall social good[4]. This paper will discuss demand management as a business process, which is used mainly in manufacturing firms, to describe the activities of demand forecasting, planning, and order fulfillment.

2.3.1 Definition of demand management

Demand management is one important link out of 8 links of supply chain management, which is integrated into Sales and Operations Planning (S&OP) or Integrated Business Planning (IBP, 2017). According to Philip Kotler, demand forecast is not only a method of figure out or stimulate the demand, but more likely to influence demand, which is more effective in achieving the target of enterprises. Therefore, the essence of demand management is to promote the multifaceted capabilities of the enterprise throughout the whole supply chain, especially through the acquisition of production information from customers, to coordinate activities related to product flow, service flow, information flow and capital flow. The expectation is to meet users' personal needs through providing them with the minimum of cost and price. Traditional demand management methods include five main activities, which are demand elicitation, demand analysis and confirmation, demand document writing, demand validation, and demand alteration management.

2.3.2 Evaluation of demand management

The components of effective demand management involve more than just forecasting, and also encompassed planning demand, communicating demand, influencing demand, and managing and prioritizing demand (George Palmatier and Colleen Crum, 2003).

Demand forecasting, is the first step in the successful implementation of demand management. It is the basis for formulating demand plans. The higher the accuracy, the higher the reliability and feasibility of the demand plan. Its accuracy depends on whether the company can understand the market and the expectations of customers (Hyndman, 2005). The ability for a company to execute the marketing and sales strategies is also necessary.

[4] "Welfare Economics", Investopedia.

Planning demand, is supply and demand management during aggregate planning with predictable demand variation. It is at an aggregate level and determining the detail-level demand plan when there are few uncertainties (George Palmatier and Colleen Crum, 2003). The fundamental tradeoffs in planning demand involves capacity, inventory, and backlog. In this aspect, time flexibility from workforce or capacity, so that to reduce unnecessary waste is the main point. In another word, it needs to master and control the demand, coordinate the other links based on the demand, and enable them to communicate continuously.

Communicating demand, is the key point for company that has accurate demand planning. Because the process of demand management itself can be seen as a process of communication. There are many components involving in communication process, such as communicating input, validating the assumption, proposing a demand plan, and reaching consensus on the demand plan (Figure 2-1). Supply planning and master scheduling can be connected to each other by communicating. Another important function of communicating is feedback and performance monitoring, which can provide key information for demand managers, and alert for action of the sales, marketing, or supply organizations.

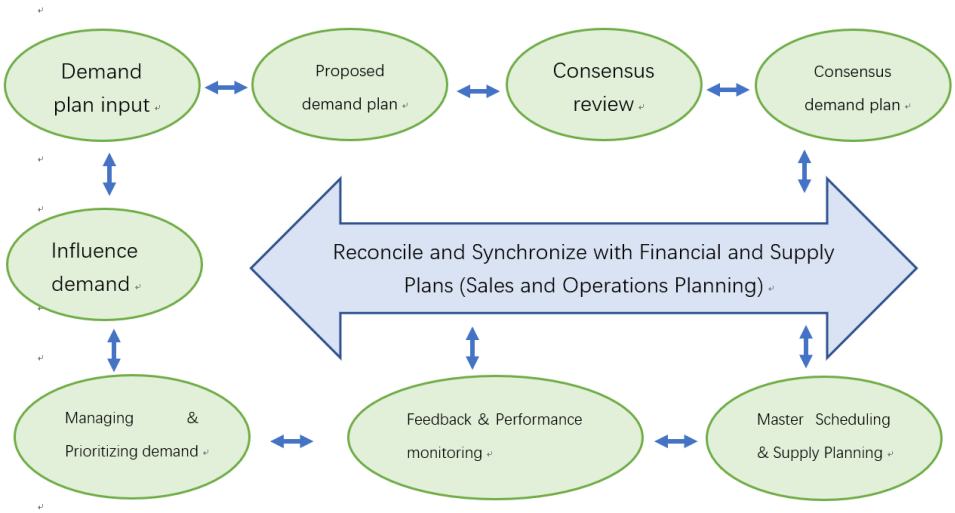

Figure 2-1: Demand management communication process

Influencing demand, has two aspects of external influence and internal influence. External influence includes macroeconomic situation, as well as competition situation and

consuming behaviors, and these are hard to control. While internal influence like sales and marketing organizations have influences on companies to meet the expectations of customers and markets. Both these aspects are in a dynamic condition in the business process, which means they are changing all the time, hence companies have to figure out solutions to adopt this condition.

Managing and prioritizing demand, many different aspects need to be taken into consideration in determining the priority of demands. The order of priority is usually made according to the characteristics of the project and its requirements, which generally including the factors such as cost, risk, manufacturing cycle, importance and so on. It is necessary to identify those key factors that need to be considered at the beginning of a project, as it is very unrealistic to consider all factors.

2.3.3 Importance of demand management

After 1990s, when WAL-MART appeared as the representative of the leading retailing company and rewrote the rules of production and sales on the supply chain, enterprises began to shift their focuses from supply to consumer demand. First of all, demand management process is no longer a simple process of event handling, it has become a dynamic management process. To fulfill the five demands mentioned above, it can connect other links in supply chain together and make the maximum effectiveness of it. As saying in a research report, “when demand and supply chain are integrated together to drive the execution and operation of supply chain, the business with the use of demand management will be successful” (Daniela de Castro Melo& Rosane Lúcia Chicarelli Alcântara, 2016). In addition, demand management is quite important in business activities, because it has a great influence on the two factors of revenue and inventory, and these two are the key factors to determine whether the supply chain is effective (Bisk). Moreover, it is an effective way for companies to meet the needs of diversification and individuation of consumers. Furthermore, when the actual situation or forecast changes, demand management can ensure that the process of sales and operation plan will be constantly revised to adapt to the changes.

2.4 Recommender systems

2.4.1 Definition and background of recommender systems

Recommender Systems (RSs) are software tools and techniques that providing suggestions to users for items to be of use (Francesco Ricci, 2011). The term “item” is a general term used to denote what the system recommends to users, and for each field there is a specific type “item” to be recommended. For example, in the field of news and propaganda, RS will push the information to users that relate to their regular browsing; in field of music, RS will recommend users music that has the similar style with the music they used to listen. RS also applies to many other fields, including restaurants, garments, financial services (Alexander Felfernig, Klaus Isak, Kalman Szabo, Peter Zachar, 2007), life insurance and online dating. The following discussion will focus on the use of RSs in e-commerce, which is according to users' characteristics of interest and previous purchase behavior, recommend interesting information and useful merchandise. It is a kind of advanced business intelligence platform, to help e-commerce websites provide personalized decision support and information service for their customers.

The emergence and popularization of the internet has brought a huge amount of information to users, which meets the users' needs of information. However, with the rapid development of the network, the amount of information has increased far more greatly, making it difficult for users to get useful information in a timely and effective way. The efficiency of the use of information and the quality of information have been reduced, which is the so-called problem of “information overload”, also named as information glut and data smog (Shenk, 1997). In a speech at the Web 2.0 Expo New York in 2008, Clay Shirky has indicated that in the modern age, information overload is a consequence of a deeper problem, which he calls filter failure (Asay, Matt, 2009). Information overload makes it inconvenient for both information receiver and supplier to communicate well. For the information receivers, they can hardly find or be convinced of some particular information, while for the suppliers, it is difficult to transfer information straightly to receivers among advertisements and spams.

In view of this phenomenon, from the perspective of information supplier, there are two

solutions based on modern technology, which are pull and push media mode, also known as pull and push marketing. Pull media is based on pull technology and is a passive way to deliver message to receiver, such as traditional means of news and propaganda, like newspaper, television, radio, or through internet like HTTP-Hypertext transfer protocol, audiences can decide which content they need and ignore the useless one. By contrast, push media is used to push information to users actively, to reduce the time of blind search on the internet. The advancement in mobile and internet technology has supplied more opportunities for this model[5], for example, advertisement and discount information sending through email.

Recommender systems are one kind of push media mode, the concept was first mentioned by Jussi Karlgren at Columbia University in a technical report in 1990 (Jussi Karlgren, 1990). Then in March 1995, Robert Armstrong and others at the University of Carnegie Mellon proposed a personalized navigation system "Web Watcher" at the American Artificial Intelligence Association, while Marko Balabanovic and others at the Stanford University introduced a personalized recommendation system "LIRA" at the same meeting. After that, RSs have been widely used in many fields, especially in e-commerce.

2.4.2 The operation mechanism of recommender systems

There are five core components (Maya Hristakeva, 2015) in every recommender system (Figure 2-2). The data collection and processing (step 1) need to be done very precisely, as the recommender models (step 2) has an upstream dependency of data. The step of recommender models is the most important one that provide recommendations to users and refers to data collection of user interest and description of items. Before the recommendations are shown to users, they need a post-processing (step 3) to filter out some common recommendations. After post-processing the recommendations, there is a set of online modules (step 4) that are responsible for serving them and tracking their use. After making the decisions on what to recommend, the user interface (step 5) will define a way to show them and to interactive with users.

[5] Definition of push media from Techopedia. https://www.techopedia.com/definition/21548/push-media

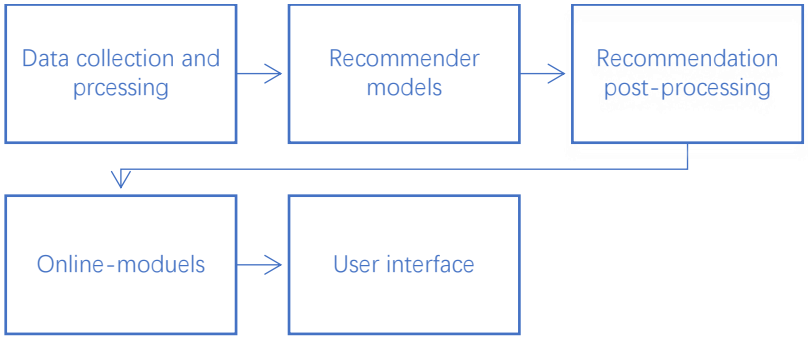

Figure 2-2: Components in a recommender system

Among the five steps, recommender models make full use of technologies like machine learning, data mining, search engines, natural-language processing and other related fields. It has 3 main links: user profile, product profile and recommendation algorithm. (Figure 2-3) The recommendation systems match the information of the interest and demand in the user profile with the feature information in the product profile, and uses the corresponding algorithm for calculation and filtering, then recommend to users the products they might be interested in (VenuGopal Reddy, 2016).

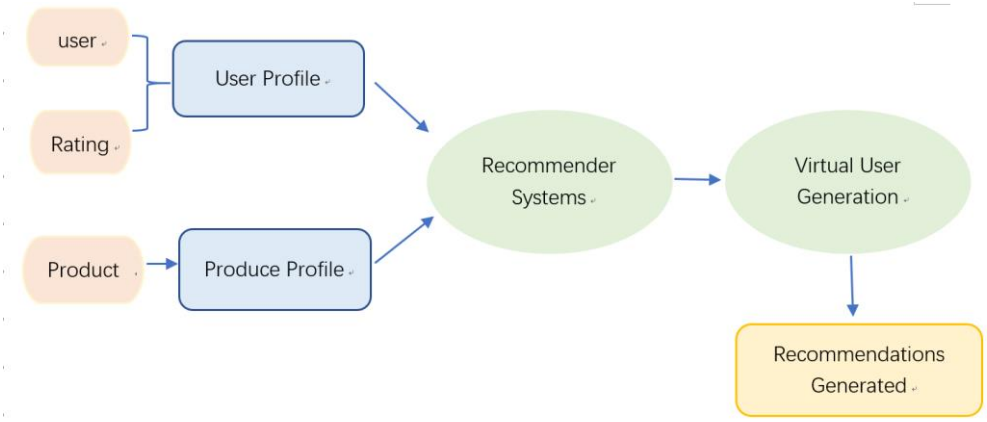

Figure 2-3: Data flow in recommender models

The recommender techniques of personalized recommender systems are based on the framework of personalized algorithms. The systems use the users interest that combined with the attributes, content, classification of the products, do research on social interaction between users, mining users' preferences and needs, and actively recommend the products that they

might be interested. To achieve this, they mainly based on **cloud computing** and a variety of **algorithm libraries**. Cloud computing can establish an accurate computing and analysis model, based on the preference trajectory of the users' activities throughout the entire network. The intelligent algorithm libraries can establish and optimize the algorithm model based on multi-dimensional data mining and statistical analysis. The complex utilization of multiple algorithms that based on content, user behavior and social relational networking can recommend products, services and contents to users that they like.

2.4.3 Types of recommender systems

The current recommender systems are divided into four main categories according to the different algorithms, they use collaborative filtering recommender, content-based filtering recommender, association rule-based recommender and hybrid recommender. (Figure 2-4)

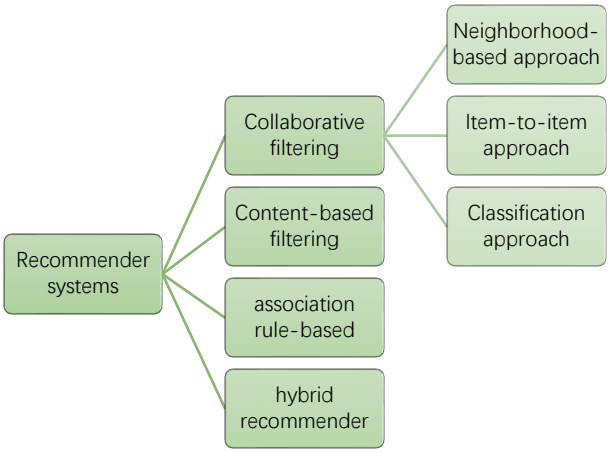

Figure 2-4: Types of Recommender systems

- **Collaborative filtering recommender**

Collaborative filtering (CF) is one of the earliest and most successful technologies being usd in recommender systems. It was introduced in the context of the first commercial recommender system, Tapestry (D. Goldberg, D. Nichols, B. Oki, and D. Terry, 1992), which was designed to recommend articles from newsgroups to a collection of readers.

The basic idea of CF is that if the user has some preferences in the past, he will have similar preferences in the future, and the similarities between users can provide advises to user who

has the same preferences. It has a narrow sense and a more general sense. In the narrower sense, it is a method of making automatic predictions (filtering) about the interests of a user by collecting information from users (collaborating). For example, if a person A has the same opinion on an issue as person B, compared with the other people, A might be more likely to have the same opinion as B in other issues, that is, they may like the same products in the future (F. Ortega, J.-L. Sánchez, J. Bobadilla and A. Gutiérrez, 2013). In the general and broad sense, collaboration among platforms including multiple agents, viewpoints, data sources are needed for filtering for information (Loren Terveen and Will Hill, 2012). Therefore, this method is based on a great amount of data (Figure 2-5).

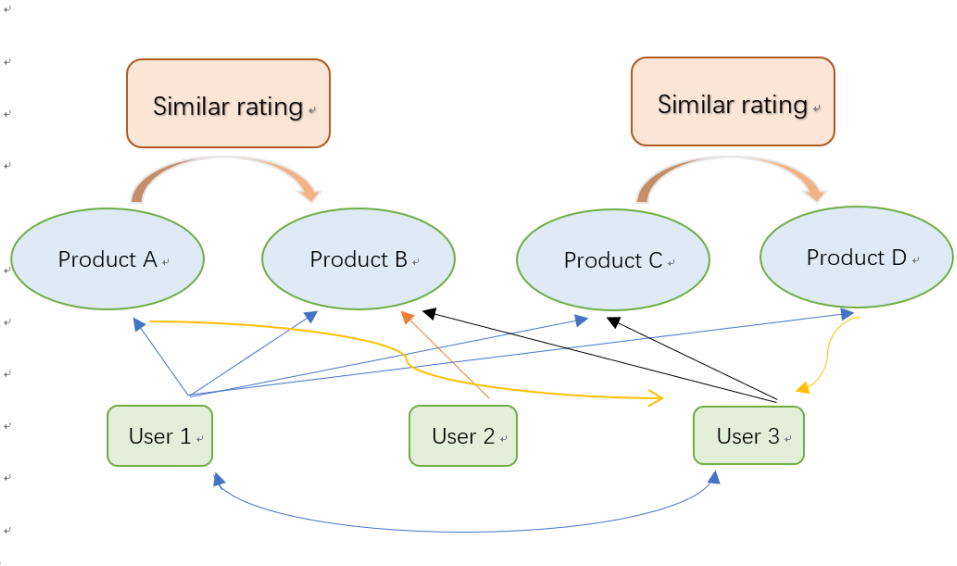

Figure 2-5: Collaborative filtering recommender

The **data collection** of user profile can be done in **explicit and implicit** forms. Explicit data collection includes data that users are asked to offer, such as to choose, to rate, to rank, or to fill a questionnaire. When entering personal information, users might have to stop the current browsing or reading behavior, which is the obvious shortcoming of explicit form. The users' reluctance to input personal data leads to a very sparse user rating data, and the recommender systems need a sufficient number of scoring data to produce more accurate results. The shortage directly leads to a decline in the quality of the system's recommendation. Implicit data collection mainly refers to analyzing the information that users have reviewed, searched for, or

interacted on social media. For instance, if the user A has recently purchased a product that the user B does not know and the system recommends this product to B, then the collaboration between A and B is implicit. During this process, the similarity between A and B is calculated by analyzing the behavior of users. Compared with explicit form, implicit of acquisition is much more flexible and its automation level is higher but is more difficult in the data acquisition. In general, the users' interest preference is diverse and dynamic, so in order to improve the quality of data collection, a mixture of explicit and implicit data collection forms is used in many practical applications.

Three **techniques** are mainly used in CF recommender to measure user similarity or item similarity.

- **First is neighborhood-based approach**, can also be called **user-based approach**, which uses both **algorithms** of Pearson Correlation (Allen, R.B, 1990) and the k-nearest neighbor (k-NN) (Sarwar, B.; Karypis, G.; Konstan, J.; Riedl, J, 2000). The k-NN approach is generally used recent years to calculate the distance between users by the user's historical preference information, which is one kind of machine learning algorithms. Then it uses the weighted evaluation value of the target users' nearest neighbor, to predict the preference of a target user to a particular commodity.
- **Second one is item-to-item approach**, can also be called item-based approach, which is an inversion of the neighborhood-based approach. Instead of measuring the similarities between users, the ratings are used to measure the correlation between items. Pearson correlation can be used in this approach. For example, the ratings of book A and B has correlation, so we can predict that person X is probably like book B based on his previous rating of book A.
- **The third technique is classification approach**. In this approach, to predict if person X likes book B, the learning method has to determine the class of this book. For example, Pazzani and Billsus (Pazzani & Billsus, 1998) use n boolean features for every user, where n indicates all the possible rating values. Once the user has given an item a rating i, a feature i will be assigned a value 1 (true) and vice versa.

There are both advantages and disadvantages of CF recommender. The biggest

advantage of CF recommender is that there is no special requirement for recommendation objects, complex and virtual products such as information quality and personal taste can also be handled, which is difficult for machine to automatically analyze the content; it can avoid incomplete and inaccurate content analysis by sharing the experience of others; it has the ability to recommend new information that is completely different from previous content, which cannot be expected by users in advance; It can effectively use similar users' feedback to reduce the amount of feedback from users. However, there are still a lot of problems to be solved. The most typical problems are cold start, sparsity and scalability. **Cold start** means in order to make accurate recommendations, it needs a large number of users' data (Rubens, Neil; Elahi, Mehdi; Sugiyama, Masashi; Kaplan, Dain, 2016). **Sparsity** refers to the problem of insufficient ratings. As there are variety of products sold on e-commerce platforms, the proportion between the most activate users and most popular items are very low, thus the rating of the most popular products is too low to be analyzed. **Scalability** means when facing millions of users and products, it often requires a lot of computing power to calculate.

- **Content-based filtering recommender**

Content-based filtering recommender (CBF, Figure 2-6) is the continuity and development of information retrieval and information filtering technology (G. Adomavicius and A. Tuzhailin, 2005). It provides recommendation based on the content information of the item and a profile of the users' preferences (Aggarwal, Charu C, 2016). There are two main components in CBF. First one is the **profile of products**, which is described by keywords. Second one is **profile of user interest** to indicate which type of items this user likes. In another word, CBF tries to recommend items that are similar to those that users liked in the past and provides variety of items to compare with the ones have been rating buy this user before, then recommend the best-match ones. For example, if a user likes detective novels, the system will be more likely to recommend "Sherlock Holmes" to him.

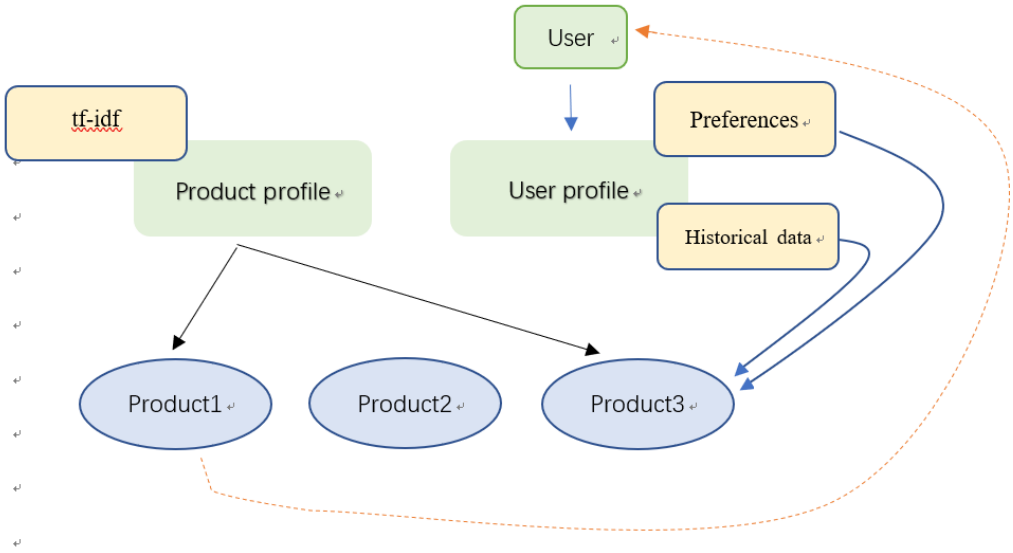

Figure 2-6: Content-based filtering recommender

The common used algorithm to abstract the feature of products includes frequency–inverse document frequency (tf-idf), which is a numerical statistic that intended to reflect how important a word is to a document in a collection or corpus (Rajaraman, A.; Ullman, J.D. ,2011).

For the user profile, the systems are mostly focusing on two kinds of information, one is user's preferences, while the other is the history data of user's interaction with the recommender systems. These kinds of data can be collected from users' previously liked items, ratings, search keyworks, and other behavioral data. Based on a weighted vector of item features, CBF creates a content-based profile of users and figure out the importance of each feature to the users, then computes from individually rated content vectors using a variety of techniques. Simple approaches use the average values of the rated item vector while other sophisticated methods use machine learning techniques such as Bayesian Classifiers, cluster analysis, decision trees, and artificial neural networks in order to estimate the probability that the user is going to like the item (Blanda, 2015). The users' history data is continually updating, and the preferences is changing all the time, therefore the user profile needs to be renovated according to these changes.

Different with CF, CBF does not need the active database of purchase history, the evaluation by users, or the data from others, so it does not have the problems of cold start and

sparsity that CF has. In addition, it can provide recommendations to users that have special interests, or of the products that are uncommon. Moreover, there are technologies used in these systems that are already mature, such as the technology of cluster analysis. But it also has some shortcomings. The content needs to be easily extracted into meaningful features, and features must have good structure. So, it has a requirement that the user profile should not be virtual but can be easily represented by some concrete details of features.

- **Association rule-based recommender**

The main idea of association rule is to recommend products based on their presence along with other products. the core mechanism is to find the frequent patterns, which are patterns occuring frequently in a database, by using frequent pattern mining to find regularities in data. When product A and B are purchased together, the presence of A in a transaction can be used to determine B that also being in the same transaction (Figure 2-7). An association rule consists of two Boolean propositions and states that if the left-hand side (the antecedent) is true, then the right-hand side (the consequent) is also true. But a probabilistic rule states that the right-hand side is true with probability p, given that the left-hand side is true (Ozgur Cakir, Murat Efe Aras, 2012). For example, when buying milk, many people will buy bread at the same time. Association analysis is the discovery of association rules showing attribute-value conditions that occur frequently together in a given set of data.

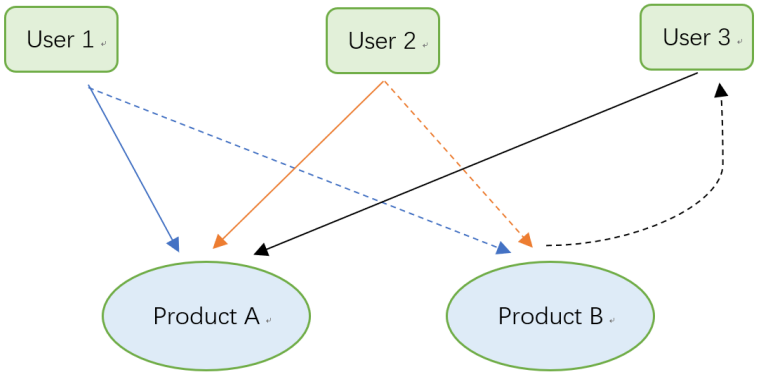

Figure 2-7: Association rule-based recommender

There are two sub-problems in association rule mining. One is to find frequent item sets. The second is to generate association rules from frequent item sets with the constraints of

minimal confidence. Most of the researches focus on the first sub-problem because the second sub-problem is quite straight forward [Kotsiantis, 2007].

One of the most common **algorithms** for finding frequent items and generating association rules is Apriori, which is first introduced in 1994 (Rakesh Agrawal). Main strategy of it is to reduce candidate frequent item sets. Because if an itemset is frequent, then all of its subsets must also be frequent, on contrast, if an itemset is infrequent, then all of its supersets must be infrequent too.

Association rules mining can discover the correlation of different commodities in the sales process and has been successfully applied in the retail industry for market basket or transaction data analysis, such as basket data analysis, cross-marketing, catalog design, sale campaign analysis, web log (click stream) analysis and so on.

This kind of system is more direct and simple, but the system will not automatically update, and it will be difficult in system management. In the first step of the algorithm, the discovery of association rules is the most critical and time-consuming step, which is the bottleneck of the algorithm, but it can be carried out offline. Secondly, the synonymy of commodity names is also a difficult point in association rules.

- **Hybrid recommender**

In addition to the above three recommender systems, there are several other methods such as feature combination (FC), feature augmentation (FA), demographic (DM), knowledge-based (KB). Because all kinds of recommender methods have their advantages and disadvantages, hybrid recommender is often adopted in practice. The one that is researched and applicated most are the combination of CBF and CF recommender. The simplest approach is to generate a recommended result, by using a content-based approach and a collaborative filtering approach separately, and then combine the results with a certain method.

Although there are many methods of recommender theoretically, they are not all effective in a specific problem. One of the most important principles of hybrid recommender is to avoid or remedy the weaknesses of their respective recommender technologies.

In the way of combination, researchers have put forward seven combinations (Robin Burke, 2007):

Weight. Combine the score of different recommender methods numerically.

Switch. Depending on the background of the problem and the actual situation or requirements, the system selects one method to use.

Mixed. Various recommendation technologies are used to provide multiple recommender results together.

Feature combination. The combination of features derived from different data sources is adopted by a single algorithm.

Cascade. First produce a rough recommender result, and the second recommendation techniques make more accurate recommendations based on the result.

Feature augmentation. One technology embeds feature information into the feature input of another recommendation technology.

Meta-level. The model generated by a recommender method is used as an input to another recommender method.

Netflix is a good example of the use of hybrid recommender systems. It compares the watching and searching habits of similar users (collaborative filtering) to make recommendations, and also recommend movies that share characteristics with films that a user has rated highly (content-based filtering).

2.4.4 Evaluation and comparison of the systems

To measure the effectiveness of a recommendation system, there are ten main indicators includes user satisfaction, accuracy, coverage rate, diversity, novelty, serendipity (recommended results are not similar to users' previous interest, but satisfy users), degree of trust, timeliness, robustness (logic of anti-cheating), business objectives (ability of making profits, but hard to evaluate).

There are three methods to measure the effectiveness of recommender systems, which are offline evaluations, user studies and online evaluations (Beel, J.; Genzmehr, M.; Gipp, B, 2013). Due to the high cost of user studies and online evaluations, most of the studies currently use off-line testing.

Offline evaluations: The user's behavior data is often obtained from the log file, then the dataset is divided into training data and test data, such as 80% training data and 20% test data (which can also be cross verifying). After that, the user's profile model is trained on the training data set, and the test it on the test set (Gareth James, 2013). The strong points of offline evaluations are as follows: it only need a data set without the actual recommendation system; off-line computing without human intervention can easily and quickly test a large number of different algorithms. The drawback is that it cannot get many practical indicators of recommendation systems, such as click rate and conversion rate.

User studies: off-line evaluations are most commonly measuring the accuracy rate, however, the accuracy is not equal to the satisfaction, so before using the algorithm, the user satisfaction needs to be investigated and test.

Online evaluations: The users are randomly divided into groups by certain rules, and different recommender algorithms are used for different groups of users. In this way, some performance indexes of different algorithms in actual use can be calculated fairly. But the disadvantage is that the cycle is relatively long, and long-term experiments are needed to get reliable results.

All the recommender systems have advantages and disadvantages, see Table 2-1 for the details.

Table 2-1 Advantages and disadvantages of main recommender systems

	Advantages	**Disadvantages**
Collaborative filtering	1. No special requirement for objects; 2. Complete and accurate; 3. New items can be recommended; 4. Don't need much feedback;	Cold start; Sparsity; Scalability;
Content-based	1. Purchase history, comments, evaluation are not necessary; 2. Others' data not necessary; 3. Can recommend special items;	Content should be concrete; Concrete and detailed user profile;

	4. Mature technologies;	
Association rule-based	Simple;	No auto update system; Association rules take time to find; Synonymy names of products;

In order to make a clear distinction of the above systems, the following Table 2-2 lists these characteristics separately to make comparison. In addition, the traditional method — information research— is added into the table, to see clearer the optimization of recommender systems.

Table 2-2 Comparison of different Recommender systems

	Information research	**Collaborative filtering**	**Content-based filtering**	**Association rule-based**
Automation	Low	High	High	Low
Sustainability	Low	High	High	Low
Individuation	Low	High	High	Low
Active & passive	Passive	Active	Active	Active
User involved	Yes	Yes	Yes	No
Special recommender	No	Yes	No	Yes
Main shortcomings	Low accuracy	sparsity	Limitation of content	Inaccurate rule
Examples	Google	WebSphere	WebMate	e-VZpro[6]

[6] Ayhan Demiriz, Enhancing Product Recommender Systems on Sparse Binary Data. Information Technology, Verizon Inc. 919 Hidden Ridge, Irving, TX 75038

2.4.5 The disadvantages and risks of recommender systems

- **Limited recommender effect**

The front end of the platform realizes thousands of recommenders for thousands people, but the server end needs to establish a complex process includes user behavior data collection of the whole network, storage and processing, data modeling and user modeling. It can only reach a small portion of the personalized recommendation effect if only simply collecting data. To raise the effect of personalized recommendation, it must cover the behavior track of users' entire network activities, even the offline activities. This requires the formation of an ecosystem with the core of the internet business platform, that is, a comprehensive link of e-commerce, physical stores, logistics and other links.

- **Problems of algorithms**

Sparsity problems are the manifestation of this problem. In any large-scale recommender system, for each user, there are always a large number of objects that are not evaluated or viewed by this user, and this kind of data is often larger than the amount of data that has been evaluated (Adomavicius G. and Tuzhilin A., 2005). The differences between users' choices are very large, resulting in sparse conditions, that is, the difference between the ratings of any two users is very large. Pazzani suggested a solution that add user's age, nationality, sex and other personal information as the basis for user similarity calculation, which is called the demographic filtering based on demographics (Pazzani, 1999). Dimension reduction algorithms for CF are also used for this problem. These algorithms try to solve the scarcity problem by reducing the user-item rating matrix. To do this, they employ different algorithms that use singular value decomposition of latent classes. (Saša Bošnjak, Mirjana Marić, Zita Bošnjak, 2008)

- **Multidimensional problem**

Most of the current research is based on a two-dimensional metric space of object-user, without considering contextual information. However, the user's evaluation and selection of objects is often determined by many environmental factors, such as the popularity of an object in a specific period of time, the user's bias towards choosing a special type of objects when the user is browsing in a certain website. Environmental factors cannot be obtained from the original

characteristics of users or objects, therefore, it is necessary to increase the feature dimensions that can be recommended according to specific problems. Take movie recommendation as an example: in addition to the character of the movie itself and people who is watching the movie, there are environmental factors including: is the person watching alone or watching with others; watching at the cinema or watching on the internet; the watching time, etc. Based on the idea of multidimensional features, Ansari proposed that this method could be further expanded by using Bayes model. (Ansari A, 2000)

- **Privacy Control**

Recommender algorithms are based on user's historical data, but complete and precise information is hard to be submitted to recommender system because of privacy protection considerations. On the other hand, the attackers who are familiar with the recommender algorithm cheat the recommender system by making use of the fabricated scores and object, so as to achieve the goal of being frequent recommendation. Therefore, a good privacy protection mechanism is one of the keys to obtain high quality data and loyal users. At the same time, the reputation of the users is also an important reference for the recommender algorithm.

Chapter 3 - Current development of recommender systems

At present, there are many types of applications that using different kinds of RSs, such as in movie, music and advertisement. See the main use of RSs in Table 3-1.

Table 3-1 The main use of RSs in different types

Types of application	**Examples**
E-commerce	Amazon, Alibaba
Movie	Netflix
Music	Pandora
Individualization read	Google Reader; Zite; Flipboard
Social media	Facebook; LinkedIn; Twitter
Location-based servise	FourSquare

Ad-Targeting	DoubleClick for Publishers (DFP)
Email	Function of Spam filtering

In 2006, Netflix organized a competition with a prize of one million dollars to the team that can take an offered dataset of over 100 million movie ratings and return recommendations that were 10% more accurate than the existing systems. Finally, the BellKor's Pragmatic Chaos team won the prize in 2009 (Steve Lohr, 2009). The new algorithms worked, as Netflix claims that about 75% of what people watch is from recommendation (Amatriain, Xavier, 2012).

3.1 Recommender systems used in e-commerce

Recently, the roles of RSs in the business process of companies are even more important, "The companies who rise to the top will be the ones who figure out the perfect balance of data, customer service and user experience", said Rachel Bogan, partner at the agency Work & Co (Hilary Milnes, 2016). Many large e-commerce companies in the world are using RSs to improve the business process, including Amazon, eBay, Alibaba, Jingdong, to name but a few. Amazon uses many forms of recommendations to engage the user (Table-4), increasing average value order and inviting them to acquire the latest items (Krawiec, Tom,2016). In addition, it develops its own homegrown math named "item-to-item collaborative filtering" (Linden, Greg et al., 2003) with a conversion rate of about 60%, claiming by itself. In Amazon's report, it had 29% sales increase to $12.83 billion during its second fiscal quarter, up from $9.9 billion during the same time in 2011 (Mangalindan, JP, 2012).

The personalized intelligence recommendation works in three channels. Firstly, when an ordinary user enters the home page of the e-commerce platform, the system can predict the consumer's desire to purchase according to the user's daily behavior preferences and purchasing habits. Before the click behavior has occurred, the system can automatically recommend so that to improve the conversion rate from bowering to buying. Secondly, when the user does not have access to the platform, the enterprise will cooperate with the platforms that the user daily browses, send links on the alliance platform with advertises of products the user is tend to buy, stimulate and guide the user to click to purchase. Thirdly, even when the user does not open

the computer, the recommendation can also be pushed to the user's mobile phone and computer at a specific time, according to the usual purchase frequency and circle of the user, through information and mail.

3.2 Recommender systems used in Amazon

Amazon has a heavy focus on data-driven marketing, generating 107.1 billion in net revenue in 2015 (Tom Krawiec, 2016). On most of its webpages, there are recommendations in different ways:

- **Personalized recommendation:** there is a user's profile page, chick which can link to a page full of different types of recommendations just for this user, and with a sentence "Recommended for you."
- **Bundling recommendation**: refers to "Frequently bought together" which aims to increase average order value. CF and CBF are used in this strategy by collecting data from user's shopping cart or the item that is currently searching for. Buying an addition commodity can increase the cross-selling benefit.
- **Feature recommendations:** Amazon knows which product a user is browsing right now, and will provide some very similar choices with different color, size, brand, etc. This is "You recently view items and Feature Recommendations". Because the user is already interested in this kind of product, so the possibility of buying is very high, while providing a variety of selections will also improve users' satisfaction with the platform.
- **Related recommendations:** has the similar target of "feature recommendations" but has another label with "Related to items you've viewed", and with some recommendations with similar feature but different brand, size and so on. In spite of being repeated with the previous one, Amazon can have the second remind for the user.
- **History recommender:** Amazon provides a "Your browsing history" to remind users that they were interested in these items.
- **User-based recommender:** refers to "Customers who bought this item also bought", which is also with a purpose of increasing average order value and cross-selling.
- **New version recommender:** "There is a new version of this item" which provides the

upgrading version of the previously purchasing item. This is association rule-based recommender that they have the same rules.

- **After-purchasing recommender:** "Recommended for you based on X" will show up after a user has bought X. The purpose is to encourage a second purchase.
- **Popular recommender:** there is a "Best-selling emerging" on the webpage to encourage users to buy the new or the most popular product. It is always a category that has not been purchased by the user before, so to broaden the selling range and increase cross-selling.

Table 3-2 Recommender systems used in Amazon

Recommender systems	**Recommendations**
Neighborhood-based CF recommender	User activities: 1. Customer watch this also watch…; 2. Customer buy this also buy…; 3. Commodities that related to the one the user bought before;
Item-to-item CF recommender	1. Commodities that related to the visiting one; 2. Commodity A has the similar rating with B by the same user;
Content-based filtering recommender	1. Tag for commodities: clothes, food 2. Keywords for commodities 3. Other commodities in the same shop
Association rule-based recommender	New versions of a bought item;

3.3 Recommender systems used in Alibaba

At the end of 2013, Taobao, the retailing website of Alibaba in China, had nearly 500 million registered users, more than 60 million fixed visitors a day, and more than 800 million online goods per day, with an average of 48 thousand sales per minute. With the expansion of its scale and the increase of the number of users, it has developed from a single C2C market to a

comprehensive retail business, including group buying, distribution, auction and other e-commerce models.

In addition, it has established a complete **credit rating system**, and expand the type of business to financial services including Alipay, Ant financial services and so on. The credit rating system plays an important role in the recommender systems, that people have an addition choice to choose whether or not to be recommended by the ratings. The Ant financial services include Ant micro loan, Ant points, etc. according to this, the platform can collect information from users at the same time when providing services.

Alibaba has a system named "All-in", is a client of Taobao website. It provides a search and class-lookup function of categories. It is embedded with WAP browser that make sure users can find and purchase anytime and anywhere. The "All-in" system has a special feature that is "**guess what you like**" and has three approaches. The first one is "**personalize in first level**", means when users open the website, it shows different content for different people. For example, female users will see dress, make-up, shoes while male users will see men's clothing, digital product, automotive supplies and so on. The second one is "**personalize in same type**", that even for the same type of products, users will see different recommendations. For example, for those who like snacks, it recommends spicy jerky to those who like spicy food and recommend chocolate to those who like to eat sweets. The third one is "**personalize in details**", means that for people who like the same products, it will recommend different brands or size to different users. For example, Ferrero and Dove will be recommended to different chocolate lover.

Based on the "All-in" system, Taobao uses multiple recommender systems in more than 60 ways, mainly in the following aspects:

- **Personalized recommendation according to "shopping cart" or "favorites":** when a user is checking the shopping cart or the folder of favorites, there is "Items you might be interested".
- **User-based recommender:** there are two types of user-base recommender. One refers to "Customers who bought this item also bought", which has a purpose of increasing average order value and cross-selling. The other one is "Customers are also interested at", to provides recommendations to a user according to the browsing history of both this user

and other users, it collects data of browse history from all web users, and calculate if a person X is checking product A, how many people (group Y, including person X) are also checking product A, then find the product B that checking by most people in group Y.

- **Shop-based recommendation:** because Taobao is a third-party platform, a same product can be found in many different shops, so there is a type of recommendation based on the credit systems and other elements, such as prices, locations, if delivering is included and so on. According to the credit systems, users can choose to see the ranking of rating, sales volume or the comprehensive rankings.
- **Related recommendation:** if a user is checking a product, the system will provide recommendations of products with similar features with a prompt of "related product", such as brand, function, color, size and so on.
- **After-purchasing recommendation:** according to the users' purchasing history, the platform keep providing recommendations of "guess what you like" with the same items as the ones bought before, but in different shops. When there is a shop has discount for these items, they will be recommended again and show up in the homepage of Taobao.
- **Time-based recommendations:** because the classification of Taobao is in a very detail way, the recommendations are impacted by the time and season to a large extent. For example, it will recommender dress for female users in summer on the homepage.

Table 3-3 Recommender systems used in Alibaba

Recommender systems	**Recommendations**
Neighborhood-based CF recommender	User activities: 1. Customer watch this also watch…; 2. Customer buy this also buy…; 3. Commodities that related to the one the user bought before;
Item-based CF recommender	1. Commodities that related to the visiting one; 2. Commodities that related to the bought one; 3. Commodities that are same/very similar with the bought

	one.
Content-based filtering recommender	1. Tag for commodities; 2. Keywords for commodities; 3. Other commodities in the same shop; 4. Same commodity in different shops;
Association rule-based recommender	Shopping cart; Favorites; Ratings; Sales volume; Locations, price, or delivering conditions; Time-based recommendations;

Multiple recommender systems are used together, and work well especially in the annual "double eleven". In 2015, the "double eleven" reduced the bounce rate (Farris, 2010) to a single digit number for the first time, while the number of guides and per capita guidance pages was 2-3 times that of the previous year, these improvements bring significant increases in profits.

3.4 Comparison of RSs between Amazon and Alibaba

Amazon and Alibaba are both the world's leading internet e-commerce platforms, but they have large differences in many ways, including the business mode, profit mode and so on. The Table 3-4 subdivide the differences.

Table 3-4 Comparison of Amazon and Alibaba

(Data from the two companies' Financial Statements of 2017)

	Amazon	**Alibaba**
Market Capitalization	458.6 billion dollars	345.6 billion dollars
Income	136 billion dollars	20.8 billion dollars
Net Margin	2.4 billion dollars	7.6 billion dollars
Business Mode	Platform + supply chain + storage (B2C)	Third party platform (C2C)

Service	Personalized service	Popular service
Products	Kindle and e-books; Commodities;	Commodities; Services (Ali-travel, film tickets, booking hotel, take away, etc.)
Profit Model	Reduce cost of supply chain and storage;	Advertising display; Value-added service (Alipay);
Marketing	in 12 countries, North America, Europe and China, Japan, India, Brazil, Australia; All have storage centers; International logistics sell China and India products to international market.	Southeast Asian; AliExpress , sell goods from China to international market.
Cloud Computing	AWS	Aliyun
Logistics	FBA	Cainiao logistics
Data of Products	Sparsity	Sparsity

From the above table, it can be seen that Amazon's income is about 6 times that of Ali, and the profit is only 1/3 of Ali, but in the global ecommerce market, Ali's market share still has a very large gap with Amazon's, because the main revenue of Alibaba is from China while Amazon has a worldwide market. According to this, the advertisement in Alibaba plays a major role, however, the value-added services are also nonnegligible. Since the products of the two websites have a common feature of sparsity, recommender systems are very necessary. The main differences in the recommender systems are in the following ways (Table 3-5).

Table 3-5 Comparison of Amazon and Alibaba in recommender systems

	Amazon	**Alibaba**
Communication	Rating	Rating (for seller, buyer and shop); Aliwangwang (instant messaging service);
Classification Products	Tags; Similar products classification;	Tags; Subdivision;
Recommender System	“Item-to item” and Multiple systems;	“All-in” and Multiple systems;

Recommender page	Yes;	No; But with "All-in" system, it will show the filter results automatically.
priority of recommendations	Relative degree; Ratings;	Advertisements; Relative degree; Ratings;
User-based CF	Customer watch/buy this also watch/buy;	Customer watch/buy this also watch/buy;
Item-based CF	Commodities that related/has similar rating to the visiting one;	Commodities that related to the visiting/bought one; Commodities that are same/very similar with the bought one.
Content-based filtering	1. Tags and key words; 2. Other commodities in the same shop	1. Tag and key words; 2. Other commodities in the same shop; 3. Same commodity in different shops;
Association rule-based	New versions of a bought item;	Shopping cart; Favorites; Ratings; Sales volume; Locations, price, or delivering conditions; Time-based recommendations;

The **communication** ways. First one is the rating systems are different. There is a feedback and a review module in Amazon, buyers can rate for the items or shops they have bought for the feedback, for review part, anyone can write something about a product whether or not they have bought it. Alibaba has a more complicated credit system for both sellers, buyers and shops. The two parties of the transaction can evaluate each other, which is one of the its recommender approaches: recommend according to the evaluate. The second difference in communication way is Alibaba has an instant message service called Aliwangwang, which enhances the

communicate effective between sellers and buyers, and help sellers know the needs of buyers in a detail level.

The way of **classifications** of products. Although both of them have tags for the products, Alibaba classify them in a more detail way. It's a better choice for people who has specific needs in style, brands or other component, while people who has a general preference will more likely to use the Amazon way.

The **recommender page** of the two platforms are different. There is a specific page in Amazon providing items recommended called "recommended for you", while in Alibaba there is not. Amazon uses an active way that people can choose what they need through this webpage. But for Alibaba, it is a passive system that provides recommendations that being filtered by "All-in" systems both in the homepage, and in the searching results, showing the items that based on users' preferences.

The **priority of recommendations** is different. Amazon ranks items according to the relative level and the ratings, while Alibaba has an extra one as well as the most important one that based on the advertisement fee. When a user is searching for a specific product, the first few items in the search results are always the ones that pay most advertisement fees. It is the shortage in Alibaba, however, it is its main profit model.

The **content-based filtering** recommender approaches are different. In the "guess what you like" module, the exact same or very similar products with the one bought before will always show up in Alibaba. But in Amazon, the system will recommender items in another way, such like a new version one or a relative one. For example, a user bought a mascara at Taobao, and it will keep asking if a second mascara is needed. But Amazon will ask if the user need eyelash curler, eyelash foundation, and eyelash growth liquid.

The **association rule-based approaches** are different. There are variety of rules in Alibaba as seen in Table-7, it has recommendations according to shopping cart, the folder of "favorites". In addition, because Alibaba has a different business model as a third-party platform, there are different conditions for shops, users can choose shops or items based on their preferences and needs. For example, if a user is looking for a quick-frozen food, the ones have nearer locations might be recommended as a priority. And among the shops with the same location, the user can

choose the one with a more sales volume or the one in the higher-rating shop.

3.5 Functions of RSs used in e-commerce

It's said in the book "The Everything Store: Jeff Bezos and the Age of Amazon" (Brad stone, 2014), if publishers do not meet the requirements of Amazon, Amazon will remove their books from the automated recommender system, so that the sales of publishers will generally fall by 40%. The role of recommender systems in e-commerce activities is enormous. In addition, because they provide recommendation as well as receiving feedback, they also have an important role in supply chain, especially in demand management. According to the previous literature review part, the key features of evaluating a demand management are as follows: (Table 3-6)

Table 3-6 Key features of evaluating demand management

	Key features of evaluating
Demand forecasting	1. Understand the market; 2. Understand the expectations of customs; 3. Good strategies of marketing;
Planning demand	Manage well at: 1. Capacity; 2. Inventory; 3. Backlog;
Communicating demand	1. Manage the process of: Communication input; Validating the assumption; Proposing a demand plan; Reaching consensus; 2. Provide feedback and performance monitoring;
Influencing demand	Figure out the dynamic conditions as follows: 1. External influence: Macroeconomic situation; Competition condition; Consuming behavior; 2. Internal influence: Sales and marketing organizations;

Managing and prioritizing demand	Know the characteristics of projects and their needs in the beginning of a project: 1. Cost; 2. Risk; 3. Manufacturing circle;

Recommender systems can effective on all of the above five elements.

Firstly, they can **provide precise forecast** for market and sales. No matter item-based recommender systems or user-based ones, they make profile of products and customers, which is help the business platform to understand customer and market, figure out the preferences of the market, so that to make marketing strategies.

Secondly, they can **improve the planning process** through enhancing cross-selling capacity. According to the purchase data of users, the systems can figure out an addition product and make recommendation to them. For example, when a user buys a pet food, he may be recommended pet care products and other products that he may need. Therefore, the company can make plans to reduce the inventory based on cross-selling.

Thirdly, recommender system can **develop the effectiveness of communication and enhance the loyalty of users** to the platform. In the competitive business world, it is an important strategy to reduce the unnecessary operation burden of the users, reduce the click times of the mouse, and establish the loyalty of the user (Reichheld and Sesser, 1990). A convenient and efficient product recommendation service may attract users repurchase behavior on the platform once again. And the more frequently users use the recommender system, the more data accumulated through feedback, the system will be more intelligent, and the recommended results will be more in line with the user expectations. Even if the competitor builds the same system and platform, the competitor is difficult to surpass in the short term because the user spends more time in the previous system.

In addition, recommender systems can **influence the dynamic process** of the intelligent supply chain of e-commerce. They can collect data of external influences such as consuming behavior so that to providing personalized service to users, so that to change the browsers into purchaser. Sometimes the users are not sure what exactly they need, then the system offers recommendations according to the users' interest and previous purchase data, thus the user's

potential demand may turn to the actual purchase behavior. They can also use the customer feedback as internal influences to make some changes in the marketing and sales organizations to adopt to the dynamic condition.

Moreover, recommender systems can **choose the priority of the marketing strategy.** The location-based recommender and time-based recommender provide the product preferences and circle to the platform. For example, electric carpet will be recommended in the cold area especially in winter, but seldom be recommended in tropical area. Another aspect is according to the data collection of recommender systems, a company will know which the most popular product is right now and make priority recommender for it to increase the benefits.

Chapter 4 - Conclusion and future work

4.1 Conclusion

With the help of recommender systems, enterprises involved in e-commerce can achieve a highly speed development. First of all, they can better understand the external environment, like consuming behavior, and better arrange the internal organizations, such as adjust the marketing plan and sales plan according to the forecast of market. It will benefit the demand management of the companies, as they can make a forecast in a detail way, raise the ability of cross-selling and reduce inventory, develop the effectiveness of communication with receiving feedback from users, and choose the priority of recommended products in the beginning of the project.

Secondly, based on the understanding of the users and market, recommender systems can provide personalized recommendations to users. There are different algorithms used in the recommender approaches, the four main ones are collaborative filtering recommender, content-based filtering recommender, Association rule-based recommender and hybrid recommender with combine multiple types of recommender approaches. There are also other types of recommender used in different conditions, such as location-based recommender used in restaurant recommender, time-based recommender used for seasonal products, and mobile recommender that provides more convenient and timely recommender for customers.

Thirdly, recommender systems will receive feedback data after providing services, so that to refresh the system through updating the data collected from users. The more data they collect, the more comprehensive they know the customer preference, so can provide more precise recommender services. It will form a virtuous circle, and the company can enhance the customer loyalty because their better understanding and services than other competitors. In this case, enterprises that make full use of the intelligent recommender systems will occupy an advantageous position in the fierce competition of e-commerce period.

4.2 Future of recommender systems

Over the years, recommender systems have brought great competitiveness to enterprises, but they also have many shortages at the same time. To have a better effect of recommendation, they have to form a comprehensive link of e-commerce, physical stores, logistics and other links to reduce the recommender limitations. The algorithms used in the systems are also need to be developed to solve the problems of cold start, sparsity, scalability, auto-updating and so on. Therefore, more researches are still need to be done in the future study.

Firstly, extract implicit negative ratings through the analysis of returned item. Most of the general recommender systems need users to provide appropriate feedback information based on their preference to the recommended object, which is called relevance feedback. Relevance feedback can be divided into explicit feedback and implicit feedback. At present, recommender systems mainly use explicit feedback, that is, users need to check their objects every time to score or other operations.

Second is the integrate community with recommendations. The information and communication technology have removed the geographical barriers, namely Balkans that means geographical space, such as locality restriction. However, the logic space Balkan is inevitably generated. The recommender systems are especially easy to lead to the local restriction and social differentiation when they are recommended according to individual interest, subject specialty, social status and opinion. A series of sociological problems arising from recommender and personalized technology are also important research directions. Other sociological problems reflected in Recommender systems include research on group

psychology and behavior, and the interaction between recommendation and mainstream interest.

The third one is the evolved of the technologies and algorithms. Until recently, recommender systems are generally based on single rather than combination of different data. New machine learning algorithms are emerging that solve this issue by building models based on carious product attributes, and user features (Scnafer, 2001). However, there is a big challenge that how to make full use with seasonal and temporary data. For example, based on a user's research history and behavior, a snow scraper might be a useful recommendation in winter, but it is irrelevant in the summer. Therefore, it needs to do more research on this emerging problem of temporal associations.

References

B. Joseph Pine II. (1993). *Mass customizing products and services*. Planning Review, Vol. 21 Issue: 4, pp.6-55

Colleen C.& George P. (2003). *Demand Management Best Practices: Process, Principles, and Collaboration*. J. Ross Publishing. ISBN-13: 978-1932159011

Hyndman, R.J.& Koehler, A.B. (2006). *Another look at measures of forecast accuracy*. International Journal of Forecasting. Monash University.

Duncan M.& Yossi S. (2003). *The Impact of Automatic Identification on Supply Chain Operations.* The International Journal of Logistics Management, Vol. 14 Issue: 1, pp.1-17

J.A. Konstan and J. Riedl. (2012) *Recommender systems: from algorithms to user experience. User Modeling and User Adapted.* Interaction, 2012, pp. 1–23.

Alexander F.& Klaus I., Kalman S., Peter Z. (2007). *The VITA Financial. Services Sales Support Environment*, pp. 1692-1699, Vancouver, Canada.

Levent V. (2007). *Consumer support systems*. Communications of the ACM. Volume 50 Issue 4. P49-54. 10.1145/1232743.1232747

Xinrui Z.& Hengshan Wang. (2005). *Study on Recommender Systems for Business-To-Business Electronic Commerce*. Communications of the IIMA: Vol.5: Iss. 4, Article 8.

Jie L.&Dianshuang W., Mingsong M., Wei W., Guangquan Z. (2015). *Recommender System Application Developments: A Survey.* Decision Support Systems archive. Vol.74: Iss. C. Pages 12-32

Business S. (2012): *2012 Supply Chain Survey Results and Insights ：IDC Manufacturing*

Insights. MI235719

D.B.A. Mehdi Khosrow-Pour. (2017). *Encyclopedia of Information Science and Technology*. Third Edition. Artificial Intelligence and Investing, Chapter 9, pp.85-93.

Brad S. (2013). *The Everything Store: Jeff Bezos and the Age of Amazon*. Little, Brown and Company. ISBN: 059307047X.

Ortega F.& Sánchez J-L., Bobadilla J and Gutiérrez A. (2013). *Improving collaborative filtering-based recommender systems results using Pareto dominance.* Information Science, pp. 50-61, 2013.

Francesco R.& Lior R., Bracha S. (2011). *Introduction to Recommender Systems Handbook*. Recommender Systems Handbook, Springer.

Heiner L.& Hans-Georg K., Peter F., Thomas F., Michael H. (2014). *Industry 4.0. In: Business & Information Systems Engineering 4 (6)*, pp. 239-242

Robert B.& Ernest L. (1999). *Introduction to Supply Chain Management.* New York: Prentice-Hall. p. 2. ISBN 0-13-621616-1.

Sobuj K.& Avizit B., Kanij F., Israfil S. (2016). *Study on Supply Chain Management of Industries in FMCG Sector in Bangladesh.* Global Journals Inc. (USA). Online ISSN: 2249-4596 & Print ISSN: 0975-5861

Lambert& Douglas M. (2008). *Supply Chain Management: Processes, Partnerships, Performance*, The Hartley Press. ISBN:978-0-9759949-3-1

Martin C. (2016). *Logistics and Supply Chain Management*. Financial Times/Pearson Education. ISBN: 978-1-08379-7

Sexton R.& Fortura P. (2005). *Exploring Economics*. Nelson College Indigenous. ISBN 0-17-

641482-7.

Demiriz A. (2004). *Enhancing Product Recommender Systems on Sparse Binary Data.* Information Technology. Kluwer Academic Publishers. ISSN:1384-5810. Online ISSN: 1573-756X

Adomavicius G& Tuzhilin A. (2005). *Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions.* IEEE Trans. on Knowledge and Data Engineering, 17(6):734−749.

Bisk. *How Demand Planning Can Improve the Supply Chain*. https://www.michiganstateuniversityonline.com/resources/supply-chain/how-demand-planning-improves-supply-chain/#.WvF4gPmnE2w

Beel J.& Genzmehr M., Gipp B. (2013). *A Comparative Analysis of Offline and Online Evaluations and Discussion of Research Paper Recommender System Evaluation*. http://docear.org/papers/a_comparative_analysis_of_offline_and_online_evaluations.pdf

Ulrich W. (2017). *Integrated Business Planning (IBP).* https://www.camelot-mc.com/en/integrated-business-planning-ibp/

Linden G & Brent S, Jeremy Y. (2003). *Amazon.com Recommendations. Item-to-Item Collaborative Filtering. Industry Report.* IEEE Computer Society. ISSN: 1089-7801 https://www.cs.umd.edu/~samir/498/Amazon-Recommendations.pdf

Definition of push media from Techopedia. https://www.techopedia.com/definition/21548/push-media

Tom K. (2016). *The Amazon Recommendations Secret to Selling More Online.* Rejoiner Website. http://rejoiner.com/resources/amazon-recommendations-secret-selling-online/

Maya H. (2015). *A Pratical Guide to Building Recommender Systems.* https://buildingrecommenders.wordpress.com/2015/11/10/the-components-of-a-recommender-system/

Sindhu V.& Saigurudatta P.V, Tharun K., Mounika P., Vikram K. (2016). *Developing A Recommender System on The Basis of Movie Lens Data Set.* https://www.slideshare.net/SaigurudattaPamulapa/recommender-system-62410707

Stephanie B. (2015). *Online Recommender Systems – How Does a Website Know What I Want?* American Mathematical Society. https://blogs.ams.org/mathgradblog/2015/05/25/online-recommender-systems-website-want/

Sanjeevan S.& Alireza S., Hossein R. (2014). Recommender Systems in E-Commerce. https://www.researchgate.net/publication/266970881_Recommender_Systems_in_E-Commerce