Recommender Systems for MOOCs: A Systematic Literature Survey
(January 1, 2012 – July 12, 2019)
Asra Khalid, Karsten Lundqvist, and Anne Yates
Victoria University of Wellington

Abstract
In recent years, massive open online courses (MOOCs) have gained popularity with learners and providers, and thus MOOC providers have started to further enhance the use of MOOCs through recommender systems. This paper is a systematic literature review on the use of recommender systems for MOOCs, examining works published between January 1, 2012 and July 12, 2019 and, to the best of our knowledge, it is the first of its kind. We used Google Scholar, five academic databases (IEEE, ACM, Springer, ScienceDirect, and ERIC) and a reference chaining technique for this research. Through quantitative analysis, we identified the types and trends of research carried out in this field. The research falls into three major categories: (a) the need for recommender systems, (b) proposed recommender systems, and (c) implemented recommender systems. From the literature, we found that research has been conducted in seven areas of MOOCs: courses, threads, peers, learning elements, MOOC provider/teacher recommender, student performance recommender, and others. To date, the research has mostly focused on the implementation of recommender systems, particularly course recommender systems. Areas for future research and implementation include design of practical and scalable online recommender systems, design of a recommender system for MOOC provider and teacher, and usefulness of recommender systems.

Keywords: recommender system, massive open online course, MOOC, systematic review, implemented recommender system
Introduction

Access to higher education can be restrictive and expensive but it can also be improved by implementing enhanced and novel methods and solutions. Massive open online courses (MOOCs) are a potential solution that have been used for more than a decade. Their spread is enabling learners to satisfy learning needs in an open, participatory, and distributed way. The term MOOC was first introduced in 2008 when the course *Connectivism and Connective Knowledge* was offered by George Siemens and Stephen Downes (Downes, 2008). Siemens designed this course according to the principles of connectivism, and due to the vast number of participants, it was named a massive open online course (Adham & Lundqvist, 2015). In 2011, at Stanford University, a MOOC different from Siemens and Downes’ was designed. Learning objectives and plans were defined, and it followed a traditional teaching style (Sunar, Abdullah, White, & Davis, 2016). This is known as a content-based MOOC (xMOOC). Currently most MOOCs are not designed on the principles of connectivism, but instead are xMOOCs.

The number of MOOCs and the number of students registered in MOOCs are growing every year. By the end of 2018, more than 900 universities were offering MOOCs with 11,400 courses available, and around 101 million students had registered in them (Shah, 2018), providing learners with a wide variety of choices. With such a high number of courses available, learners now face the problem of selecting courses without being overwhelmed.

With the increase in e-commerce and online business, the number of users attracted to online Web services has increased. Both MOOC providers and online businesses advertise their courses and services while learners search for courses that match their interests and needs. In these situations, recommender systems play an important role, and have attracted the attention of researchers. Recommender systems are algorithms and techniques that recommend matching and relevant courses or services to the learner depending upon their interests, information about which comes from learner profiles and histories gathered by the systems. Recommender systems help MOOC providers grow and learners find more appropriate and customized services tailored to their personalities and interests. An example is provided below.

Mark has a free slot in the evening, and he wants to polish his professional skills by registering in a part-time course *Introduction to Java*. Mark has no idea about the course, and he does not want to waste his money on something that will not help his career. What will he do? Mark has different options: he can ask his friend who has completed this course, or he can observe details of the course, such as the content, length, pre-requisites, and instructors to reach a decision. In this case, Mark is searching for recommendations or inferring data to generate a recommendation for himself. What should we do if we face the same problem in our online learning life? We could use recommender systems, which help diminish information overload.

Recommender systems discover patterns in considerable datasets to learn the preferences of different users and predict items that correlate to their needs. Here *item* is a generic term that represents any course, learning element, book, service, application, or product. Recommender systems mostly use machine learning and data mining techniques to achieve their goals (Ricci, Rokach, Shapira, & Kantor, 2010). These systems are used intensively in e-commerce and by retailers to lift their sales and audience and now, increasingly, for learning purposes in MOOCs.
According to Manouselis, Drachsler, Verbert, and Duval (2013), recommender systems can be divided into two broad categories: collaborative filtering recommender systems and content-based recommender systems. There is a third type called the hybrid that contains characteristics of both collaborative filtering and content-based recommender systems.

Collaborative filtering recommender systems perform recommendations on the assumption that people who have had similar taste in the past will make similar choices in the future. This can be compared with real life scenarios in which, when we have to choose from multiple available options, we consider the recommendations of family and friends who have similar interests (Dakhel & Mahdavi, 2013).

Content based recommender systems consider the profile of users and items. Profiles of users can include age, gender, education, and residency area. Characteristics of items, for example in the case of movies, might include actor, genre, category, and type. These recommender systems analyze the items rated by a user and try to design a model that reflects the interests of that user. This model is employed to recommend new items to the user (Lops, de Gemmis, & Semeraro, 2011).

With the increased use of MOOCs, data produced by MOOCs is also expanding. This data contains information about the interests and behaviors of learners and the courses in which they are registered, and that data can be used by a recommender system to make recommendations (Ricci et al., 2010). Recommender systems in MOOCs can help the learner find related learning objects or elements. MOOC providers can also use these systems to inform MOOC design and creation.

The purpose of this systematic literature review was to fully scope and report on: (a) how recommender systems have been used in MOOCs between 2012 and 2019, (b) the trends over this period, and (c) the types of recommender systems yet to be explored. This research reviewed all related work between January 1, 2012 and July 12, 2019, in the English language only. We chose 2012 as the starting year because it was declared the Year of the MOOC by The New York Times (Pappano, 2012) and, from that year, publication of peer-reviewed research on recommender systems in MOOCs started.

**Method**

According to Fink (2005), a systematic literature review is an organized, comprehensive, and reproducible method of review. Using this definition as a framework, the purpose of this study was to

- report on work on recommender systems for MOOCs; and

- provide a comprehensive analysis that could be used to find opportunities for research and implementation in the field.

Our methodology consisted of two fundamental steps: data collection and data analysis. The analysis was further divided into quantitative and qualitative analyses.
Data Collection

Gathering data from the literature was performed with care to maximize accuracy. A set of rules describing the criteria for selection of research papers was established. These rules involved four significant points: (a) search terms, (b) research period, (c) sources, and (d) publication type. Search terms are used to find related published work from specific sources, while research period refers to the publication date, and publication type refers to the type of paper, such as journal article, conference paper, book chapter, or review article. The following sections explain these rules in more detail.

**Search terms.** This review involved two main concepts: massive open online courses and recommender systems. Therefore, we started with the following search terms: “massive open online courses” AND “MOOCs” AND “recommender system.” We added “RS,” a common abbreviation for recommender systems, but that resulted in many unrelated papers. Similarly, we used “MOOC” instead of “MOOCs,” which also resulted in many unrelated papers since MOOC is used as an abbreviation for other terms such as “multiple optical orthogonal code sequences” and “management of organizational change.” We also used “adaptive MOOCs” and “personalized MOOCs” along with “recommender system” and “massive open online courses.” With “personalized MOOCs,” we only found one related paper which was already in our database, whereas the term “adaptive MOOCs” resulted in seven papers, though they were also part of our database. Most of the unrelated papers in the latter case were about making MOOCs adaptive and not about recommending any resource or service to users.

Thus, we finalized the search terms: “massive open online courses” AND “MOOCs” AND “recommender system” because these were the most efficient for locating the literature we were seeking.

**Research period.** We reviewed papers published between January 1, 2012 and July 12, 2019.

**Sources.** To determine the sources of research, we followed the same methodology as Liyanagunawardena, Adams, and Williams (2014). We used Google Scholar, academic databases, and the reference chaining technique of Gao, Luo, and Zhang (2012). The initial searching was in Google Scholar, followed by selected academic databases. We chose five databases from the area of computer science and education: the Institute of Electrical and Electronics Engineers (IEEE) Xplore, the Association for Computing Machinery (ACM) journals/Transactions Springer Link, ScienceDirect, and the Education Resources Information Centre (ERIC). Reference chaining was performed at the end to find any further related work.

**Publication type.** Peer reviewed conference papers, journals, and book chapters were included in this literature review.

**Data Analysis**

We performed both quantitative and qualitative analysis on the data. In the quantitative analysis, we classified data based on publication year, publication type, and the geographical region of authors. In the qualitative analysis, we used open coding content analysis (Gao et al., 2012). In this technique, there were two phases; first, we read all papers to extract themes and, second, the themes were classified. Then the same process was repeated to refine the classification and synthesis.
Limitations

For this systematic literature review, we only considered:

- articles published between January 1, 2012 and July 12, 2019. (We note that there may have been conference papers presented before July 12, 2019 that were not published by the cutoff date for this study and that they were not included in our literature review.).

- five academic databases and Google Scholar.

- peer reviewed journal articles, conferences, and book sections.

- papers in which the recommender system for MOOCs is proposed, implemented, or discussed as a need, or in which different recommendation algorithms for MOOCs are compared.

- articles that were published in English. While searching Google Scholar and performing reference chaining, we found related articles in other languages, such as French. These other articles are not included.

The Google Scholar search returned more than 30,000 items (13 October 2019). These items included websites, blogs, videos, etc. However, we did not include these resources because they are subjective and usually not considered for peer review. We did, however, include existing literature reviews.

Results and Analysis

Descriptive/Quantitative Analysis

The initial Google Scholar search resulted in 424 research papers. After analyzing titles and abstracts, 124 papers were classified as relevant. After a detailed analysis of each of these papers, we considered only 89 to be related to the topic of recommender systems in MOOCs.

Table 1 contains the results of the searches from the academic databases. Springer Link showed 144 publications, of which 26 were related to our research. IEEE and Springer contained the highest number of related publications, but ERIC revealed no related research papers. Many of the unrelated papers were about recommender systems used in technology enhanced learning other than MOOCs.
Table 1

**Distribution of Papers Found in Academic Databases**

<table>
<thead>
<tr>
<th>Academic database</th>
<th>Number of related papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEEE</td>
<td>26</td>
</tr>
<tr>
<td>Springer</td>
<td>26</td>
</tr>
<tr>
<td>ACM</td>
<td>20</td>
</tr>
<tr>
<td>Science Direct</td>
<td>7</td>
</tr>
<tr>
<td>ERIC</td>
<td>0</td>
</tr>
</tbody>
</table>

After searching the databases, we performed reference chaining and found another 10 related papers. As a result, we had 116 papers, which included 88 conference papers, 26 journal articles, and 2 book chapters. Both book chapters were published in 2019. Figure 1 shows yearly distribution of literature in these categories.

![Figure 1. Yearly distribution of literature by type: journal article, conference paper, or book chapter. *2019 data includes research published only up to July 12, 2019.](image)
There were no publications on recommender systems in MOOCs in 2012, but subsequently, a gradual increase in the number of publications per year is visible. The highest number of publications was in 2017. (Note that 2019 covers only 6 months.)

Table 2 shows groups of authors who have had more than one publication in this research area. Of 319 authors, we found 68 had at least two papers in this area, and the maximum number of papers from a single author was five.

Table 2

Groups of Authors Having More Than One Publication

<table>
<thead>
<tr>
<th>Group of Authors</th>
<th>No. of Publications</th>
<th>Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ayse Saliha Sunar*</td>
<td>5</td>
<td>Sunar et al. (2016)</td>
</tr>
<tr>
<td>Nor Aniza Abdullah</td>
<td></td>
<td>Sunar, Abdullah, White, &amp; Davis (2015a, 2015b)</td>
</tr>
<tr>
<td>Su White</td>
<td></td>
<td>Sunar, Abdullah, White, &amp; Davis (2015c)</td>
</tr>
<tr>
<td>Hugh C. Davis</td>
<td></td>
<td>Yousef &amp; Sunar (2015)</td>
</tr>
<tr>
<td>Ahmed Mohamed Fahmy Yousef</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Covadonga Rodrigo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hugues Labarthe</td>
<td>4</td>
<td>Bouchet, Labarthe, Bachelet, &amp; Yacef (2017)</td>
</tr>
<tr>
<td>François Bouchet</td>
<td></td>
<td>Bouchet, Labarthe, Yacef, &amp; Bachelet (2017)</td>
</tr>
<tr>
<td>Rémi Bachelet</td>
<td></td>
<td>Labarthe, Bachelet, Bouchet, &amp; Yacef (2016)</td>
</tr>
<tr>
<td>Kalina Yacef</td>
<td></td>
<td>Labarthe, Bouchet, Bachelet, &amp; Yacef (2016)</td>
</tr>
<tr>
<td>Jian Zhao</td>
<td>4</td>
<td>Bhatt, Cooper, &amp; Zhao (2018)</td>
</tr>
<tr>
<td>Chidansh Bhatt</td>
<td></td>
<td>Cooper, Zhao, Bhatt, &amp; Shamma (2018a, 2018b)</td>
</tr>
<tr>
<td>Matthew Cooper</td>
<td></td>
<td>Zhao, Bhatt, Cooper, &amp; Shamma (2018)</td>
</tr>
<tr>
<td>David A. Shamma</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diyi Yang</td>
<td>3</td>
<td>Yang, Piergallini, Howley, &amp; Rosé (2014)</td>
</tr>
<tr>
<td>Jingbo Shang</td>
<td></td>
<td>Yang, Shang, &amp; Rosé (2014)</td>
</tr>
<tr>
<td>Carolyn Penstein Rosé*</td>
<td></td>
<td>Yang, Adamson, &amp; Rosé (2014)</td>
</tr>
<tr>
<td>Fei Mi</td>
<td>3</td>
<td>Mi &amp; Faltings (2016a, 2016b, 2017)</td>
</tr>
<tr>
<td>Boi Faltings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group of Authors</td>
<td>No. of Publications</td>
<td>Publications</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>---------------------</td>
<td>-------------------------------------------------------------------</td>
</tr>
<tr>
<td>Guanliang Chen</td>
<td>3</td>
<td>• Chen et al. (2016, 2018)</td>
</tr>
<tr>
<td>Dan Davis</td>
<td></td>
<td>• Chen, Davis, Krause, Hauff, &amp; Houben (2017)</td>
</tr>
<tr>
<td>Markus Krause</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efthimia Aivaloglou</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Claudia Hauff</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geert-Jan Houben</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hiba Hajri</td>
<td>3</td>
<td>• Hajri, Bourda, &amp; Popineau (2017, 2018, 2019)</td>
</tr>
<tr>
<td>Yolaine Bourda</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fabrice Popineau</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hao Zhang</td>
<td>3</td>
<td>• H. Zhang, Huang, Lv, Liu, &amp; Yang (2019)</td>
</tr>
<tr>
<td>Tao Huang</td>
<td></td>
<td>• H. Zhang, Huang, Lv, Liu, &amp; Zhou (2018)</td>
</tr>
<tr>
<td>Zhihan Lv</td>
<td></td>
<td>• H. Zhang, Yang, Huang, &amp; Zhan (2017)</td>
</tr>
<tr>
<td>Sanya Liu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heng Yang</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Olga C. Santos</td>
<td>2</td>
<td>• Santos &amp; Boticario (2015)</td>
</tr>
<tr>
<td>Jesus G. Boticario</td>
<td></td>
<td>• Santos, Boticario, &amp; Pérez-Marín (2014)</td>
</tr>
<tr>
<td>D.F.O. Onah</td>
<td>2</td>
<td>• Onah &amp; Sinclair (2015a, 2015b)</td>
</tr>
<tr>
<td>J.E. Sinclair</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fatiha Bousbahi</td>
<td>2</td>
<td>• Bousbahi &amp; Chorfi (2015)</td>
</tr>
<tr>
<td>Henda Chorfi</td>
<td></td>
<td>• Ouertani &amp; Alawadh (2017)</td>
</tr>
<tr>
<td>Panagiotis Adamopoulos</td>
<td>2</td>
<td>• Adamopoulos (2014a, 2014b)</td>
</tr>
<tr>
<td>Daniel Burgos</td>
<td>2</td>
<td>• Burgos &amp; Corbí (2014)</td>
</tr>
<tr>
<td>Alberto Corbí</td>
<td></td>
<td>• Corbí &amp; Burgos (2014)</td>
</tr>
<tr>
<td>Yifan Hou</td>
<td>2</td>
<td>• Hou et al. (2016)</td>
</tr>
<tr>
<td>Pan Zhou</td>
<td></td>
<td>• Hou, Zhou, Xu, &amp; Wu (2018)</td>
</tr>
<tr>
<td>Ting Wang</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li Yu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yuchong Hu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dapeng Wu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thanasis Daradoumis</td>
<td>2</td>
<td>• Bassi, Daradoumis, Xhafa, Caballé, &amp; Sula (2014)</td>
</tr>
<tr>
<td>Roxana Bassi</td>
<td></td>
<td>• Daradoumis, Bassi, Xhafa, &amp; Caballé (2013)</td>
</tr>
<tr>
<td>Fatos Xhafa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Santi Caballé</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Group of Authors</th>
<th>No. of Publications</th>
<th>Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marwa Harrathi</td>
<td>2</td>
<td>Harrathi, Touzani, &amp; Braham (2017, 2018)</td>
</tr>
<tr>
<td>Narjess Touzani</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rafik Braham</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sara Assami</td>
<td>2</td>
<td>Assami, Daoudi, &amp; Ajhoun (2018, 2019)</td>
</tr>
<tr>
<td>Najima Daoudi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rachida Ajhoun</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rodrigo Campos</td>
<td>2</td>
<td>Campos, dos Santos, &amp; Oliveira (2018a, 2018b)</td>
</tr>
<tr>
<td>Rodrigo Pereira dos Santos</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jonice Oliveira</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Naima Belarbi</td>
<td>2</td>
<td>Belarbi, Chafiq, Talbi, Namir, &amp; Benlahmar (2019a, 2019b)</td>
</tr>
<tr>
<td>Nadia Chafiq</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mohammed Talbi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdelwahed Namir</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elhabib Benlahmar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panagiotis Symeonidis</td>
<td>2</td>
<td>Symeonidis &amp; Malakoudis (2016)</td>
</tr>
<tr>
<td>Dimitrios Malakoudis</td>
<td></td>
<td>Symeonidis &amp; Malakoudis (2018)</td>
</tr>
<tr>
<td>Ivan Srba*</td>
<td></td>
<td>Macina, Srba, Williams, &amp; Bielikova (2017)</td>
</tr>
<tr>
<td>Joseph Jay Williams</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maria Bielikova</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peter Babinec</td>
<td></td>
<td></td>
</tr>
<tr>
<td>René F. Kizilcec</td>
<td>2</td>
<td>Alario-Hoyos et al. (2014)</td>
</tr>
<tr>
<td>Mar Pérez-Sanagustín*</td>
<td></td>
<td>Kizilcec, Pérez-Sanagustín, &amp; J. Maldonado (2017)</td>
</tr>
<tr>
<td>Jorge J. Maldonado</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carlos Alario-Hoyos</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Derick Leony</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iria Estévez-Ayres</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Israel Gutiérrez-Rojas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carlos Delgado Kloos</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Authors who have publications with more than one group of authors.

At this stage, we analyzed research links between authors and how they are grouped. Figure 2 shows the network of authors who have at least two papers in this area, and their links with other groups of authors.
Figure 2. Network diagram of authors who are linked with other groups of authors. Red nodes indicate authors who have publications with more than one group.

By observing the country of the first author, we determined that the majority of work (43%) is from Europe whereas 24% and 22% of research in this field was performed in Asia and the USA respectively. Ten percent of the research is from Africa, with 1% from Australia. In Asia, most of the research is from China. Figure 3 shows the distribution by country.
We classified the literature into four categories: need, design proposal, implementation, or other. These are defined as follows:

- **Need**: Papers that mainly focused on the importance of recommender systems in MOOCs.
- **Design proposal**: Papers in which the author has given an abstract proposal for a recommender system.
- **Implementation**: Research work in which authors designed an algorithm and implemented it on a dataset.
- **Other**: All other papers in which authors reviewed the current work, guidelines, or challenges.

Figure 4 shows trends in these categories between 2012 and 2019. Implementation was the main focus of research throughout this period, and from 2016 onwards, the number of published papers in this area rose rapidly. The reason for this rapid increase is that researchers not only implemented new techniques but also implemented their proposals from their 2014 and 2015 research work. Research on design proposals for recommender systems in MOOCs showed a gradual decrease after 2016. A similar pattern is evident in the need category.
In the implementation category, some authors also evaluated their work using metrics and baselines. Table 3 illustrates the number of implemented and evaluated recommender systems. Among all implemented systems, 42% were evaluated using different datasets and evaluation techniques. Most authors used datasets of edX and Coursera, but some also created their own datasets. For evaluation, most authors used receiver operating characteristic (ROC), recall and precision metrics, as well as accuracy metrics. The remaining 58% did not evaluate their proposed solutions and instead presented evaluation as future work.

Table 3

<table>
<thead>
<tr>
<th>Year</th>
<th>Implementation</th>
<th>Evaluated</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2013</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2014</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>2015</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>2016</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>2017</td>
<td>17</td>
<td>11</td>
</tr>
<tr>
<td>2018</td>
<td>18</td>
<td>11</td>
</tr>
<tr>
<td>2019*</td>
<td>8</td>
<td>7</td>
</tr>
</tbody>
</table>

Note: *2019 data includes research published only up to July 12, 2019.
Content/Qualitative Analysis

To carry out a comprehensive review of a topic, it is necessary to conduct an in-depth analysis through synthesis. In a systematic review, synthesis provides a bottom-line statement regarding any gaps and missing links through pooling and exploring the results (Fink, 2005). In this section, we highlight the main issues addressed and major contributions on recommender systems in MOOCs. We found that research could be broadly categorized into seven main themes.

- **Thread recommender**: Thread recommender involves thread/discussion, question recommendation, and question tag recommendations.

- **Learning element recommender**: Learning element recommender includes learning activities, suggestions on the study, video lectures, next page recommender, source, and learning path recommenders.

- **Course recommender**: Only involves course recommendation.

- **Student performance recommender**: Student performance recommender involves jobs, grades, student difficulty based, student dropout, work plan, and paid task recommenders.

- **Peer recommender**: Social interactions are a key factor in successful learning, and peer recommender involves systems that recommend related peers or fellow learners to interact with instead of recommending a learning resource or another class to follow. It uses demographics and progress made in a course for recommendations.

- **MOOC provider/teacher recommender**: This involves curriculum recommendations, news of MOOCs, and MOOC provider feedback.

- **Others**: This category involves improved and personalized MOOCs, adaptive content, and special user recommender systems.

We found that more than 60% of the literature is on course and learning element recommender systems for MOOCs. A possible reason for this is that universities or institutes that offer MOOCs do so to increase enrolment and throughput, and therefore, they recommend further courses to those already enrolled. Figure 5 shows the percentage distribution of research in different categories.
Figure 5. Distribution of work done on different types of recommender systems in MOOCs.

To analyze the type and trends of work found in the literature, we grouped the research work with respect to the area of MOOCs where the recommender system is applied. Table 4 shows a detailed categorization of different areas of MOOCs where recommender systems have been applied.

Table 4

*Distribution of Work in Recommender System Categories*

<table>
<thead>
<tr>
<th>Research concern</th>
<th>Related studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thread recommender</td>
<td>Cohen et al. (2013); Yang, Piergallini, et al. (2014); Sunar et al. (2015b); Jo, Tomar, Ferschke, Rosé, &amp; Gašević (2016); Mi &amp; Faltings (2016a, 2016b); Kardan, Narimani, &amp; Ataiefar (2017); Mi &amp; Faltings (2017); Lan, Spencer, Chen, Brinton, &amp; Chiang (2019).</td>
</tr>
<tr>
<td>Thread /discussion recommender</td>
<td></td>
</tr>
<tr>
<td>Question recommender</td>
<td>Yang, Adamson, et al. (2014); Yang, Shang, et al. (2014); Macina et al. (2017).</td>
</tr>
<tr>
<td>Learning element recommender</td>
<td></td>
</tr>
<tr>
<td>Research concern</td>
<td>Related studies</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td><strong>OER/learning element/activity recommender</strong></td>
<td>Piedra, Chicaiza, López, &amp; Caro (2014); Itmazi &amp; Hijazi (2015); Niu et al. (2015); Onah &amp; Sinclair (2015b); Paquette, Mariño, Rogozan, &amp; Léonard (2015); Kopeinik, Kowald, &amp; Lex (2016); Hajri et al. (2017); Harrathi et al. (2017); Hajri et al. (2018); Harrathi et al. (2018); Xiao, Wang, Jiang, &amp; Li (2018); Chanaa &amp; Faddouli (2019); Hajri et al. (2019); H. Zhang et al. (2019).</td>
</tr>
<tr>
<td><strong>Suggestion to study</strong></td>
<td>Corbi &amp; Burgos (2014); Niu et al. (2015).</td>
</tr>
<tr>
<td><strong>Video/lectures/clip recommender</strong></td>
<td>Agrawal, Venkatraman, Leonard, &amp; Paepcke (2015); Gómez-Berbís &amp; Lagares-Lemos (2016); Bhatt et al. (2018); Cooper et al. (2018a, 2018b); Mawas, Gilliot, Garlatti, Euler, &amp; Pascual (2018); Zhao et al. (2018); Belarbi et al. (2019a, 2019b).</td>
</tr>
<tr>
<td><strong>Next page recommender</strong></td>
<td>Pardos, Tang, Davis, &amp; Le (2017).</td>
</tr>
<tr>
<td><strong>Learning path recommender</strong></td>
<td>Popescu, Portelli, Anagnostopoulos, &amp; Ntarmos (2017).</td>
</tr>
<tr>
<td><strong>Course recommender</strong></td>
<td>Ahera &amp; Lobo (2013); Apaza, Cervantes, Quispe, &amp; Luna (2014); Bousbahi &amp; Chorfi (2015); Onah &amp; Sinclair (2015a); Fu, Liu, Zhang, &amp; Wang (2015); Yanhui, Dequan, Yongxin, &amp; Lin (2015); Fazeli, Rajabi, Lezcano, Drachsler, &amp; Sloep (2016); Gómez-Berbís &amp; Lagares-Lemos (2016); Hou et al. (2016); Piao &amp; Breslin (2016); Symeonidis &amp; Malakoudis (2016); Dai et al. (2017); Gope &amp; Jain (2017); He, Liu, &amp; Zhang (2017); Jing &amp; Tang (2017); Y. Li &amp; Li (2017); Ouertani &amp; Alawadh (2017); Shaptala, Kyselova, &amp; Kyselov (2017); EL Alami, Eddine, &amp; Mohamed (2017); Yuqin Wang, Liang, Ji, ShiweiWang, &amp; YiqiangChen (2017); Yuanyuan Wang, Maruyama, Yasui, Kawai, &amp; Akiyama (2017); H. Zhang et al. (2017); Assami et al. (2018); Campos et al. (2018a, 2018b); Chen et al. (2018); Hou et al. (2018); Iniesto &amp; Rodrigo (2018); Jain &amp; Anika (2018); Jun Xiao et al. (2018); X. Li, Wang, Wang, &amp; Tang (2018); Pang, Liao, Tan, Wu, &amp; Zhou (2018); Rabahallah, Mahdaoui, &amp; Azouaou (2018); Symeonidis &amp; Malakoudis (2018); H. Zhang et al. (2018); Agrebi, Sendi, &amp; Abed (2019); Aryal et al. (2019); Boratto, Fenu, &amp; Marras (2019); Chanaa &amp; Faddouli (2019); Margolis et al. (2019).</td>
</tr>
<tr>
<td><strong>Student performance recommender</strong></td>
<td>Symeonidis &amp; Malakoudis (2016).</td>
</tr>
<tr>
<td><strong>Jobs recommender</strong></td>
<td>Symeonidis &amp; Malakoudis (2016).</td>
</tr>
<tr>
<td>Research concern</td>
<td>Related studies</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Grades improvement recommender</td>
<td>Elbadrawy et al. (2016); Luacesa, Dieza, Alonso-Betanzosb, Troncosoc, &amp; Bahamondea (2017).</td>
</tr>
<tr>
<td>Student dropout based recommender</td>
<td>H. Zhang et al. (2019); M. Zhang, Zhu, Wang, &amp; Chen (2019).</td>
</tr>
<tr>
<td>Work plan recommender</td>
<td>Alario-Hoyos et al. (2014).</td>
</tr>
<tr>
<td>Paid task recommender</td>
<td>Chen et al. (2016); Chen et al. (2017); Chen et al. (2018).</td>
</tr>
<tr>
<td>Peers recommender</td>
<td></td>
</tr>
<tr>
<td>Peer recommender</td>
<td>Sunar et al. (2015a); Labarthe, Bouchet, et al. (2016); Bouchet, Labarthe, Yacef, et al. (2017); Prabhakar, Spanakis, &amp; Zaiane (2017); Potts et al. (2018).</td>
</tr>
<tr>
<td>MOOC provider/teacher recommender</td>
<td>Zhou et al. (2015); Medio et al. (2017).</td>
</tr>
<tr>
<td>News of MOOCs</td>
<td>Holotescu (2016).</td>
</tr>
<tr>
<td>Others</td>
<td></td>
</tr>
<tr>
<td>Improve and personalize MOOC</td>
<td>Daradoumis et al. (2013); Burgos &amp; Corbí (2014).</td>
</tr>
<tr>
<td>Special user</td>
<td>Iniesto &amp; Rodrigo (2016).</td>
</tr>
</tbody>
</table>

Table 5 shows research on the implementation or proposal of recommender systems in MOOCs. There are some papers in which authors have discussed the recommender systems in a generalized way, while in other papers they have provided guidelines or a literature review of existing work. We have classified these papers...
into four broad categories: preliminary study; literature review; challenges and effects of recommender systems in MOOCs; and design guidelines. A description of each category follows:

- **Preliminary study**: All research papers which discuss initial steps of the design of a recommender system in MOOCs. In these papers, the authors discuss steps and possible techniques for preprocessing of data.

- **Literature review**: We found two related literature reviews. However, these reviews discussed personalized MOOCs and not recommender systems.

- **Challenges and effects of recommender systems in MOOCs**: Papers in this category target the challenges of implementing a recommender system in MOOCs and the effects on MOOCs after introduction of a recommender system.

- **Design guidelines**: Papers in which authors have described guidelines to design a recommender system are in this category.

Table 5 shows the distribution of research work by year into these four categories.

Table 5

<table>
<thead>
<tr>
<th>Year</th>
<th>Preliminary study</th>
<th>Literature review</th>
<th>Challenges and effects of RS in MOOCs</th>
<th>Design guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>Kizilcec et al. (2017).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2019 (up to July)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 6 shows the trend in the types of recommender systems researched over time. Until 2017, a gradual increase in research was evident. Initially, researchers focused on thread and course recommender systems, which then extended to peer, learning element, and student performance recommenders. By 2016, MOOC provider recommender systems were added to the research stream, and this trend continued in 2017. In 2018, most of the research was into course recommender systems, while no work was found on thread recommenders. Up until July 2019, course and learning element recommenders were the focus of research.

Figure 7 presents the number of research publications based on the different types of recommender systems applied to MOOCs. Overall, course recommender and learning element recommender systems are the most popular areas of research in the application of recommender systems.

![Figure 6: Change over time in the types of recommender systems in MOOCs researched. *2019 data includes research published only up to July 12, 2019.](image-url)
Figure 7. Distribution of research based on type of recommender system in MOOCs (2012 to July 2019).

Figure 8 shows the percentage of published work falling into each of the four broad classifications of research described in the Descriptive/Quantitative Analysis section of this paper. These are: (a) need, (b) design proposal, (c) implementation, and (d) other. The overall focus of research is the implementation of recommender systems in MOOCs. The category of other includes no implementation, and around 70% of this research is about proposed systems only. A possible reason for this could be that the work included in this category is meant only for a specific group of people.

Figure 8. Proportion of literature categorized with respect to type of recommendation and type of research.
Discussion

To the best of our knowledge, this is the first systematic literature review on the use of recommender systems in MOOCs from 2012 to July 12, 2019. Published work falls into three major categories: the need for recommender systems, proposed recommender systems, and implemented recommender systems in MOOCs. We classified the types of recommender systems into seven themes: course recommender, learning elements recommender, peer recommender, thread recommender, student performance recommender, MOOC provider/teachers’ recommender, and others. In this section, we discuss the types and trends of research carried out within each of these themes. We also identify gaps in the current literature which may be areas for future research.

Course Recommender Systems

The implementation of course recommender systems was a key focus of much of the research. This could be due to the availability of data and the interests of MOOC providers, because course recommender systems can help in improving enrolment and the learning experience. From 2013 to 2016, most of the work on the implementation of the recommender system for courses used collaborative and content-based filtering (Onah & Sinclair, 2015a; Piao & Breslin, 2016). Some researchers discussed the need for course recommender systems in MOOCs (Campos et al., 2018a; Fu et al., 2015; Ouertani & Alawadh, 2017). Campos et al. (2018b) implemented a course recommender system using knowledge reuse in ecosystems, and Hou et al. (2016) considered the context of the learner while performing recommendations.

After 2016, along with collaborative and content-based filtering for course recommendation (Boratto et al., 2019; He et al., 2017; Hou et al., 2018; Rabahallah et al., 2018) researchers started to use neural networks, pattern mining, and deep learning for preprocessing of data and recommendations (Agrebi et al., 2019; Jain & Anika, 2018; Jing & Tang, 2017; H. Zhang et al., 2019). We also observed the introduction of association rule mining and hybrid algorithms (Xiao et al., 2018; Y. Li & Li, 2017; Pang et al., 2018) Gope and Jain (2017) used the learning style of the student to recommend courses. Their prototype was based on a learning system model and worked exclusively with edX courses. It scanned every course to identify learning objects and then made recommendations.

Learning Elements Recommender Systems

More than half the work we reviewed focused on the implementation of recommender systems and the work concerns content-based filtering or hybrid algorithms (Cooper et al., 2018a; Itmazi & Hijazi, 2015; Zhao et al., 2018). Researchers have designed recommender systems for different types of learning elements, such as video clips, next page, and additional resources helpful to the learner. Kopeinik (2016) compared existing algorithms that provide a recommendation of learning resources and tags to annotate these resources. Onah and Sinclair (2015b) recommended a suitable path to learners by considering scores on concept-based quizzes. A low score indicated that the learner needed more resources related to a concept, and so this system would recommend instructional material according to a learner’s profile.

While preprocessing the dataset, we observed an increase in the use of neural networks, pattern mining, and machine learning in more recent years (Cooper et al., 2018a; Hajri et al., 2019; Xiao et al., 2018; Pardos et al., 2017; H. Zhang et al., 2019). Cooper and his colleagues researched the recommendation of video...
lectures to learners by analyzing the content of videos. They designed a user-friendly interface that suggested related videos while learners were watching another video (Bhatt et al., 2018; Cooper et al., 2018a, 2018b; Zhao et al., 2018).

**Peer Recommender Systems**

Another group of authors completed detailed research in the field of peer recommender systems. They investigated the effect of a peer recommender on overall student performance (Labarthe, Bachelet, et al., 2016; Labarthe, Bouchet, et al., 2016). They compared the results of different peer recommender systems in terms of student engagement (Bouchet, Labarthe, Yacef, et al., 2017). These researchers also attempted to identify the reasons for peer communication usage in MOOCs through conducting surveys of peer recommender users and found that most students express their emotions about coursework through communicating with peers (Bouchet, Labarthe, Bachelet, et al., 2017). Reciprocal scores were used by some authors to find and recommend suitable peers for learners (Potts et al., 2018; Prabhakar et al., 2017).

**Thread Recommender Systems**

Most research on thread recommender systems focused on implementation. We found only two papers related to the algorithm proposals (Cohen et al., 2013; Sunar et al., 2015a). Implementation work was mostly performed by using matrix factorization, collaborative filtering, and content-based filtering for recommender systems (Garg & Tiwari, 2016; Yang, Piergallini, et al., 2014). Mi and Faltings (2016a, 2016b, 2017) used a context tree for online thread recommendations. Yang et al. (2014) designed a recommender system for threads in a forum that recommends questions for students to answer based on their expertise. They also managed learner workload by defining a threshold on the number of questions recommended to each learner. After proposing an initial algorithm, Yang, Shang, et al. (2014) then improved this algorithm by adding sub-modularity to make it computationally less expensive. Agrawal et al. (2015) designed a recommender system that recommends video clips from lectures based on questions asked in forums.

**Student Performance Recommender Systems**

To increase student performance and engagement, some researchers (Alario-Hoyos et al., 2014; Luacesa et al., 2017; M. Zhang et al., 2019) presented work focused on the design of recommender systems only. Chen et al. (2016) first proposed and then designed and implemented a system that recommends to learners course-related paid tasks from freelancing websites such as Upwork or Witzmart (Chen et al., 2018; Chen et al., 2017). The main idea was to make it possible for learners to earn money while using MOOCs.

**MOOCs Provider/Teacher’s Recommender Systems**

Only two studies paid attention to designing recommender systems for MOOC providers and teachers. Holotescu (2016) designed a chatbot for MOOCs that works with Facebook and provides news about MOOCs that can deliver the latest news about MOOCs to teachers and providers. Zhou et al. (2015) designed an Android application to improve a course, and this application takes feedback from students during the course and makes suggestions to the teacher based on this feedback.
Conclusion

The use of recommender systems in MOOCs presents exciting opportunities to increase the popularity of MOOCs and improve the learners’ experience. Research to date has mostly focused on the implementation of recommender systems in MOOCs, particularly course recommender systems which was the most prolific research line throughout the period.

From 2012 to 2016, researchers modified existing recommender systems that were designed for e-commerce, music, videos, or books, to make them appropriate for use in MOOCs; however, from 2017 onwards, researchers started to apply neural networks, deep learning and data mining techniques in data preprocessing to apply recommender systems in MOOCs. Researchers focused on learners and strived to exploit their learning habits.

Future Directions

Although a considerable number of recommender systems in MOOCs have been proposed and implemented, only a few authors have discussed the time and space complexity of their proposed and implemented algorithms (Ahera & Lobo, 2013; Hou et al., 2018; Mi & Faltings, 2016a, 2017; H. Zhang et al., 2018). MOOCs produce a large amount of data that can be used for recommender systems and researchers should focus on systems that scale well with the increase in data and have linear time and space complexity. In evaluating their solutions, authors have ignored the training and recommendation time that their recommender system is taking.

One reason for overlooking this aspect of their algorithms could be the batch/offline nature of proposed algorithms. Batch/offline algorithms use existing datasets for training and recommendations. For this purpose, algorithms require memory space and time, the amount of which depends upon the type of dataset. Online recommender systems consider only the current context of the user while computing recommendations. In MOOCs, the current context of the user is an important factor, and researchers should put more focus on this. We found only one such work, Mi and Faltings (2016b, 2017), that addressed online recommender systems.

There is also a lack of standardized datasets available for the evaluation of recommender systems in MOOCs. Researchers have mostly used publicly available datasets of Coursera, edX, and, in some cases, datasets from their own institutes to evaluate recommender systems (Aryal et al., 2019; Dai et al., 2017; Kardan et al., 2017; Mi & Faltings, 2016a; Shaptala et al., 2017; Yang, Adamson, et al., 2014; Yang, Piergallini, et al., 2014). Other authors have created datasets (Onah & Sinclair, 2015a; He et al., 2017; Iniesto & Rodrigo, 2019; Zhou et al., 2015). A lack of standardized datasets can be a significant limitation when benchmarking or comparing algorithms or techniques of different researchers. Furthermore, most researchers used datasets from computer science-related courses for testing their recommender systems (Aryal et al., 2019; Bhatt et al., 2018; M. Zhang et al., 2019; Zhou et al., 2015) which limits the research to one academic field.

Scant attention has been paid to designing recommender systems for MOOC providers and teachers. Such systems can help providers and teachers in planning course materials, delivery styles, and the content of the MOOC. Recommender systems could also help providers decide which courses should become MOOCs.
The effect of recommender systems on student engagement and completion rates is another useful topic to pursue.

We also observed that previous research has overlooked different types of MOOCs, such as cMOOCs, xMOOCs, and sMOOCs, and has not considered the characteristics of types of MOOCs while designing recommender systems. In future, recommender systems could cater to different types of MOOCs.
References


Chen, G., Davis, D., Krause, M., Aivaloglou, E., Hauff, C., & Houben, G.-J. (2016). Can learners be earners? Investigating a design to enable MOOC learners to apply their skills and earn money in


Khalid, Lundqvist, and Yates

Communications in computer and information science, vol 658. (pp. 269-280). doi: https://doi.org/10.1007/978-3-319-48024-4_21


283
Khalid, Lundqvist, and Yates

Interaction RoCHI'2016 (pp. 91-94). Retrieved from https://pdfs.semanticscholar.org/832e/8de6424644765f98094c0127281120fe66e5.pdf


