Software and Systems Modeling
https://doi.org/10.1007/s10270-020-00790-w

SPECIAL SECTION PAPER

Profiling users via their reviews: an extended systematic mapping study

Xin Dong[1] · Tong Li[1] · Rui Song[1] · Zhiming Ding[1]

Received: 30 September 2019 / Revised: 10 March 2020 / Accepted: 12 March 2020

Abstract

With the extensive development of big data and social networks, the user profile field has received much attention. User profiling is essential for understanding the characteristics of various users, contributing to better understanding of their requirements in specific scenarios. User-generated contents which directly reflect people's thoughts and intention are a valuable source for profiling users, among which user reviews by nature are invaluable sources for acquiring user requirements and have drawn increasing attention from both academia and industry. However, review-based user profiling (RBUP), as an emerging research direction, has not been systematically reviewed, hindering researchers from further investigation. In this work, we carry out a systematic mapping study on review-based user profiling, with an emphasis on investigating the generic analysis process of RBUP and identifying potential research directions. Specifically, 51 out of 2478 papers were carefully selected for investigation under a standardized and systematic procedure. By carrying out in-depth analysis over such papers, we have identified a generic process that should be followed to perform review-based user profiling. In addition, we perform multi-dimensional analysis on each step of the process in order to review current research progress and identify challenges and potential research directions. The results show that although traditional methods have been continuously improved, they are not sufficient to unleash the full potential of large-scale user reviews, especially the use of heterogeneous data for multi-dimensional user profiling.

Keywords User-generated reviews · User profiling · Systematic mapping study · Software requirements

1 Introduction

Software requirements analysis is a crucial stage of the software development life cycle [42]. In order to precisely elicit user requirements, in particular understanding their requirements, it is essential to establish user profiles. Typically, user profiles are elicited and tagged in terms of statistic attributes, e.g., the number of products a person has bought [1]. Moreover, rate-based user profiling [1] has been investigated for years, which judges whether users like the product according to their ratings. However, rate-based user profiling cannot mine fine-grained information such as about user's opinions and rationale.

Recently, user reviews are becoming an ideal data source for profiling users [21]. In particular, user reviews are an important way for users to express their requirements and preferences on specific products or services, by mining which we can obtain more precise user profiles. Given such a promising research direction, currently there is a lack of systematic survey and comparison of new technologies for review-based user profiling (RBUP), which may result in reinventing wheels and hinder the development of this field. For example, in the literature [21], the current status and research trends of RBUP are not clearly stated, and it focuses on the application field of the user profiling, not the development of RBUP itself. In addition, the techniques of user profiles are divided into three categories in the literature [8], content-based method, hybrid method, and collaborative

Communicated by Jelena Zdravkovic and Iris Reinhartz-Berger.

✉ Tong Li
litong@bjut.edu.cn

Xin Dong
dongxin19@foxmail.com

Rui Song
songrui@emails.bjut.edu.cn

Zhiming Ding
zmding@bjut.edu.cn

1 Beijing University of Technology, Beijing, People's Republic of China

method. The modeling methods of user profiles have been classified, but the historical development of user profiles has not been investigated from the perspective of systematic mapping study. Therefore, it is important to provide an overview of the review-based user profile area, which will help reflect the current state of research and the stage of development in this field.

In this work, we conduct a systematic mapping study (SMS) to investigate publications that fall into the scope of RBUP [14,17]. This SMS can be beneficial for both researchers and practitioners. For researchers interested in RBUP, the map helps them to build upon existing work and understand the research trends, shedding light on promising directions. For practitioners, a generic processing framework for RBUP identified in our study would help them to better implement RBUP and select appropriate techniques, promoting this field from barbaric growth to unified and stable development.

As per available SMS guidelines [14,16], we focus our investigation on a set of particular research questions. First, we aim to present the current state of user profiling based on the systematic investigation of the selected studies. Secondly, we focus on identifying the generic research process of RBUP and the research topics in different stages of the process. Finally, we discuss and explore the challenges regarding this research area and point out the direction for subsequent research in this field. Overall, we focus on the trend of advanced technologies and hot spots in recent years, as well as discussing the potential reasons for the development of this field, and map the results from multiple aspects, including time evolution, publishing venues, application areas, and co-author networks.

This work is an extended version of our previous conference paper [10]. In particular, we have made significant improvements in the following aspects.

- We have carefully revised the search query of our study to include a broader scope of relevant studies. Thus, we carried out another three-round paper selection, resulting in 34 new papers selected from 1106 papers. All such new papers were carefully reviewed and tagged through the defined analysis process in the paper. Previously, we have analyzed 21 relevant, and after removing the repetition of the two sets, now we deal with 51 papers in total, significantly improving the comprehensiveness of our study.
- We have carried out a more useful analysis over the study results, including co-author network (Sect. 5.1.4) and paper venue (Sect. 5.1.2), to gain insight on review-based user profiling. In addition, we provide more detailed statistics on the research activities of each relevant paper, clearly justifying the generic analysis process identified in this paper.
- As for all the tags we used for analyzing papers, including review elements classification (Sect. 5.3.2), user profiling techniques classification (Sect. 5.3.3), etc., we have carefully defined each of them in order to improve validity and repeatability of our study.

The rest of the paper is structured as follows. Section 2 refers to the related work and introduces the background for this paper. Section 3 describes the research methodology and the systematic mapping process. Section 4 presents the statistical results of the obtained data. The analysis of the activities in the generic process is detailed in Sect. 5. Section 6 discusses the threats to validity, followed by the conclusions in Sect. 7.

2 Related work

As the user profiling research field matures, the number of available reports and results tends to increase dramatically; thus, providing an overview of RBUP becomes important. Many researchers have generated a literature review to analyze the research status and future direction of the field. In this paper, we investigate the difference between systematic mapping study (SMS) and systematic literature review (SLR). Then, we introduce the background of user profiling and compare the difference between our work and other SMS work.

2.1 SMS & SLR

The systematic literature review is one secondary study method that has gotten much attention lately in software engineering (SE) [4] and is inspired by medical research. The SLR is used to review existing significant reports, study them in-depth, and describe their methods and results. The SLR incorporates statistical synthesis analysis in the traditional literature review process [7,34]. The advantages of SLR are as follows: (1) At the knowledge level, the SLR develops rapidly in the mechanism of knowledge innovation, and the statistics of the literature establishes the relationship between knowledge. (2) At the technical level, the SLR not only systematically collects documents within a specific range but also expands the scope of research. However, the disadvantage of the SLR is that it requires much work.

Systematic mapping study is a set of research methods that use standard techniques to extract and integrate selected documents and obtain new problems and new theories [18]. It can be seen as a research advancement in primary research based on empirical evidence [11]. SMS has a particular research process: (1) Solving specific problems according to research purposes. (2) Developing inclusion and exclusion criteria to follow the screening stage. (3) SMS adopts standardized technology and scientific methods for research and analysis. (4) Under the control of specific methodologies, the literature is qualitatively evaluated, and on this basis, new research results and new theoretical perspectives are generated.

2.2 User profiling & review-based user profiling

A user profile is a set of information representing a user via user-related rules, settings, needs, interests, behaviors, and preferences [8,15]. This collection of personal information can either be represented as static data or dynamic data. The static attribute of the user refers to the relatively stable data of the user, such as age, gender, and region. The user's dynamic properties vary with the user's behavior, and the data in different domains are different. However, regardless of the type of data, the accuracy of the user profile depends on how the user information is collected and organized [32]. In other words, it depends on the user analysis process that gathers, organizes, and interprets information to create user summaries and descriptions.

There are two main challenges in the user profiling process. The first is to generate an initial user profile for new users. The second is to continually update the profile information to adapt to the changing preferences of the user [15,25]. To solve the first challenge, basic portraits of users can be generated from popular tags or basic information. For the second challenge, the large amount of review information in social platforms can reflect the changing preferences of users. By mining the implicit user demand information in the reviews, we can describe the user profiles through the in-depth semantic information. This semantic information contains text patterns, review topics, emotional preferences, and other factors in user reviews. Through this review element, the researchers explore features that may reflect user preferences from various aspects of the review to form user tags. At the technical level, similar user groups can be clustered and classified using user profile techniques. Therefore, when further defining user interests, it is necessary to prevent the short-term interest of the user that may lead to the problem of lack of diversity.

2.3 SMSs applied to RBUP

Over the past decade, a large number of SMS and literature reviews on user profiles have summarized the status of RBUP and provided reports about this field [27].

Chen and Wang [21] summarized the present situation from review-based user profiling and review-based product profiling. They proposed possible research directions which included combining different types of review elements and producing review-based explanations. In [22], the techniques of various user profiling are combined and classified to form a classification tree. The article [22] emphasizes that for new customers who have already registered and have not purchased, as well as old customers who have a purchase record, the same recommendation method cannot be used to personalize. However, these studies only describe some related technologies and do not make a comprehensive and systematic analysis of the user's profile. The work in the literature [8,15] provides theoretical and practical proof for our research. The literature [8] investigates different methods, techniques, and algorithms of the user profiling process. The author aims to give an overview of the user profiling and its related concepts and discuss the pros and cons of the current methods for the future service personalization. In the article [8], user profiling methods are divided into three categories: content-based method, collaborative method, and hybrid method. At the same time, the article also describes the comparison of user profiling methods at the algorithm level (classification and clustering algorithms). The paper [15] analyzes the development trend of user profiles and explains how user profiles have evolved from the recommendation system, then discusses user profile techniques, and finally gives examples of user profiling applications in various fields.

In addition, since the field of user profiles about reviews is an emerging research direction, there are not many related research reviews. Therefore, we refer to SMS in related fields to help us conduct systematic research. We refer to Sabine Wolny et al.'s SMS about systems modeling language (SysML). They screened 579 documents according to the general method of SMS and has done a detailed study on its development stage and contribution [38]. We also refer to Jennifer Horkoff's research on goal-oriented requirements engineering (GORE). The article elaborates on the intercoder agreement and analyzes the multifaceting mapping properties of GORE [12].

By reading the SMS in related fields, we refer to the specific definition of SMS in the field of software engineering by Petersen [14] to explain the systemic mapping process used in this article. Specifically, we focus on maps of existing work rather than detailed surveys to assess the quality of publications. The process of discovery and inclusion of papers is clearly defined in our work, which makes our research questions more evident [33]. At the same time, we also referred to the indispensable aspects of the systematic literature review, which provided a comprehensive understanding of the development of the field. Kitchenham et al. [18] provide guidelines for SLRs in software engineering. When applicable, we apply these guidelines to our SMS, including specifying a hypothesis, defining populations, defining a process, providing raw data, and making extensive use of graphics [17].

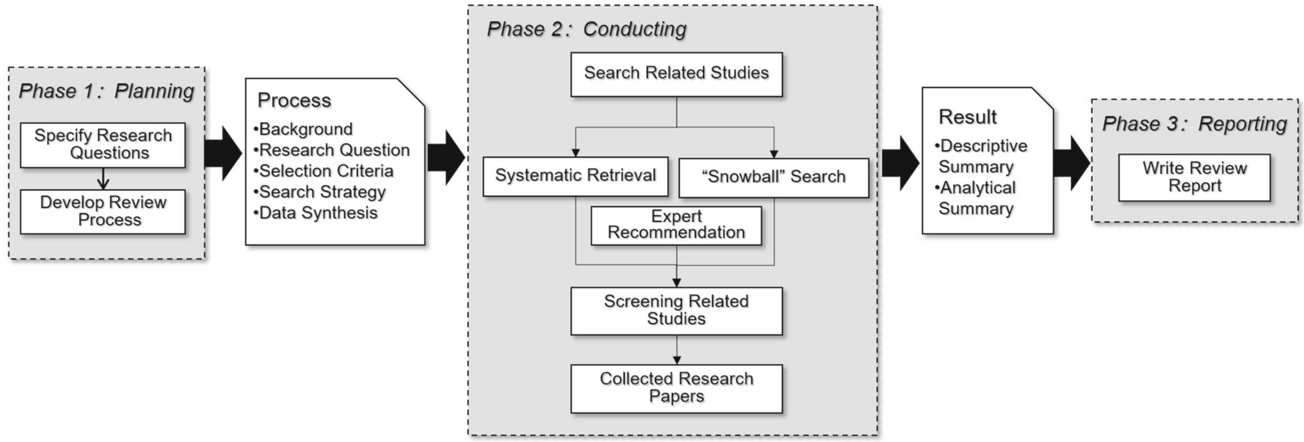


Fig. 1 Systematic mapping process

3 Research methodology

This study followed the guidelines for SMS developed by Kitchenham et al. [17,42], and the specific SMS process adopted in this paper is presented in Fig. 1. In particular, we schedule three main stages: planning, conducting, and reporting. First, we identified meaningful research questions based on the research topic and determined the scope of subsequent research (Sect. 3.1). Second, in the conducting phase, we systematically researched relevant publications in the database and screened papers according to the criteria (Sects. 3.2 and 3.3). Then, we performed a mapping study of the selected papers and responded to the research questions (Sect. 4).

3.1 Research questions

To accomplish the goal of this study (Sect. 1), we formulate four research questions as listed below:

– *RQ-1 What is the overview of selected studies?*
  – What is the evolution trend of RBUP research?
  – In which venues are most of the work in this area published?
  – What are the application areas of RBUP?
  – What is the co-authoring pattern in this field?

  By answering this research question, we can introduce the overall status of the field clearly and help the interested readers to be familiar with RBUP field in more detail. Secondly, through the statistics of bibliometrics, we can explore the development of maturity and trend of the development of this field.
– *RQ-2 What is the generic research process in this field?*
  This question is for the existing research process of RBUP. Specifically, the purpose is to collect the processing flow commonly used in the RBUP field and form the conversion relationship between modules. There is currently no uniform report on research methods and processes in this field, so the results of this question can act as a reference list for authors while designing and conducting RBUP studies.
– *RQ-3 Based on a unified generic process in RQ2, what topics and valuable research aspects are covered by each activity?*
  Answering this question will extend the general process obtained in RQ2. The results provide the topics and research points included in each process stage. By understanding the research methods and hot spots in various activities, researchers are provided with comprehensive research directions.
– *RQ-4 What are the opportunities and challenges in this field?*
  RQ4 is related to opportunities and challenges in this field. We hope that the work in this article can provide new research ideas for future development and provide valuable analysis reports on the status of research. Most importantly, this will allow researchers to better understand and promote the maturity of the field.

3.2 Search process

After identifying the research questions, the next step is to formulate the search query, which is used to retrieve papers about RBUP. In order to systematically and comprehensively search for papers in the field, we use systematic retrieval, "Snowball" search, and expert recommendation methods. The systematic literature search process for SMS is shown in Fig. 2.

Since the electronic library contains a large number of conference, journal, patent, and workshop publications, we first automatically retrieve the related literature from the database through keywords. After automatically retrieving

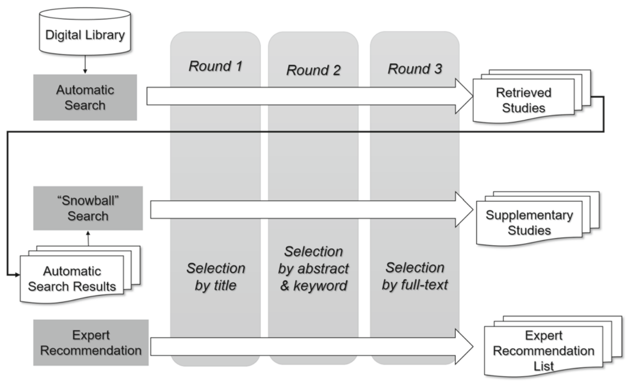


Fig. 2 Paper screening process

papers, researchers need to filter the collected data because of the uneven quality of the literature. The literature retained after three rounds of screening by researchers is the final result of SMS. Secondly, in order to expand the potentially related research, we use the "snowball" method to search for relevant literature manually [37]. In addition, we supplement the existing literature from previous work [10] to the latest research statistics of this article. In summary, we use three search strategies to collect the final research literature.

3.2.1 Determine data sources

We evaluated various potential sources, including Google Scholar and authoritative digital library, and we decided to perform our search through the authoritative digital library as it covers major publishers and is less inclusive than Google Scholar, which may include many non-peer-reviewed papers such as technical reports. Moreover, to ensure the comprehensiveness of the research, we use multiple digital libraries, instead of one single library. In particular, we search in the following digital libraries, which have been widely used in SMS [14,17,33].

– IEEE Xplore[1] An academic literature database provides indexing, abstracting, and full-text downloading services for documents related to computer science and electrical engineering.
– Web of Science[2] It is an online subscription-based scientific citation indexing service that provides a comprehensive citation search.
– Scopus[3] It is a literature database which includes more than 5,000 publishers in the world of science, technology, medicine, and social sciences.
– ACM Digital Library[4] It is a full-text scholarly online databases, where users can find full-text articles and bibliographic records.

3.2.2 Define the search query

The keywords in our research topic are *review-based user profiling*. Therefore, we derived our search query from our research questions and evaluated whether it is feasible. We identified two essential keywords of our study, i.e., "user profile" and "review." Moreover, we use word variants (singular/plural) and Boolean operators for optimization. In addition, we have formed multiple keyword strings based on a combination of different keywords and tested them in multiple databases. The reason for the test is to make the retrieved literature as relevant as possible to the keywords. The search deadline for this extended version is August 1, 2019. Based on these considerations, an optimized search query was obtained:

("user profiling" OR "user profile") AND ("review" OR "reviews" OR "review-based")

In order to ensure that the search scope and screening standards of the database are consistent, we set the following constraints: *a) The time is setting 1999–2019.* We are concerned about the development of RBUP for 20 years in order to explore the evolution trace of this research field. *b) The language is English. c) Search in metadata.* The word "review" has multiple meanings; thus, if we search in full paper, it will produce much irrelevant literature. However, titles, abstracts, and keywords often represent the main content of the literature. Retrieval from metadata not only guarantees the quality of retrieval but also limits the number of related documents.

3.2.3 "Snowball" search

In order to ensure the integrity of the research literature, we carried out a "snowball" search [12]. The "Snowball" method is to track the references of the paper on the results of Sect. 3.2.2. In the same screening process, the most relevant literature was selected and included in our study. In this work, we only roll once as further iterations rarely produce new papers. The specific screening rules are described in detail in Sect. 3.3.

3.3 Screening process

After obtaining the retrieved publications, we need to select the target literature that meets the screening criteria. First, a pilot study [19] was adopted to enhance the feasibility of screening and reduce the bias of human judgment. Second, through a pilot study, researchers can improve screening

[1] http://ieeexplore.ieee.org.
[2] https://www.webofknowledge.com.
[3] http://www.scopus.com.
[4] http://dl.acm.org.

Table 1 The inclusion and exclusion criteria

Inclusion criteria (IC)	Exclusion criteria (EC)
IC1: More than four pages of journals, conferences, book chapters, or workshop articles	EC1: Non-peer-reviewed paper (e.g., technical report), less than four pages (e.g., extended abstract, poster, short paper), non-reviewed articles
IC2: The data collected and used must contain user-generated reviews	EC2 The data collected and used contain only rate, tag, twitter, MicroBlog, or demographic attribute information
IC3: At least one module in the article describes and mines user preferences, including tagging modules for users	EC3: Text processing and analysis were only performed on the reviews, and no further statistics on user preferences were made
IC4: The main goal and contribution is the user profiling model construction, including the research of mining method, the study of user profiling method, and the construction of user profiling model	EC4: The analysis of the user's profiles is not taken as the primary analysis object of the article. In particular, the article mentioned the term "user profile" but without discussing its research methods and theories

methods and estimate the time and resources required to complete larger versions (Sect. 3.3.1). Then, in the large-scale screening, the target literature is selected by three rounds of iteration (Sect. 3.3.2).

3.3.1 Pilot study

In our experiments, we used a pilot study method [19] to ensure the reliability of the publication, which is important for conducting reasonable and full-scale research project [36]. Specifically, we selected ten studies from different databases. These studies are from our previous work [10], which contains seven studies that can be included in the results list, and three studies that should be excluded. Two researchers completed the screening phase. According to the results selected by the two researchers, one article was misclassified. In particular, they classified six studies correctly. Then, we looked for another expert who was not involved in the pilot study process to re-evaluate ten studies. The result is the same as the initial classification of the literature. Therefore, the researchers discussed the misclassified literature and refined the screening criteria to make it clearer.

3.3.2 Paper screening process

In this section, the screening criteria mentioned in Sect. 3.3.1 are given, and the details of multiple rounds of screening in the process of investigation are shown.

The inclusion and exclusion criteria in the screening process are defined in Table 1. Our research focuses on publications in journals or conferences (IC1). Less than four pages of essays and some low-quality reports were removed from the survey (IC1). The data source of the research article must contain user-generated reviews (IC2). The research literature must use text information to analyze user behavior and mining user preferences in the process of modeling (IC3). IC4 means that the main goal and contribution of this paper is the construction of the user profiling model, including the research of mining methods and user profile evaluation methods. The relationship between the inclusion criteria is *(IC1 AND IC2 AND(IC3 OR IC4))*.

We ignore publications in the form of the editorial poster because their quality is uneven and affects the quality of research (EC1). At the same time, we ignore papers that are less than four pages because their content is relatively simple, and the research value is weak (EC1). For the content, papers that are not related to the process about the user's profile should be deleted (EC2). More importantly, if the paper only processes the text of the review without the analysis of the user tag, we should also delete it (EC3). We removed papers that only mention user profiles but do not focus on them (EC4). The relationship between the exclusion criteria is *(EC1 OR EC2 OR EC3 OR EC4)*.

Using the criteria for inclusion and exclusion presented above, we conducted three rounds of paper screening (Round1, Round2, and Round3). Based on the principle of objectivity and impartiality, three researchers were assigned to conduct three rounds of paper screening. Two researchers reviewed each paper in each round. In case two researchers come out with conflicting decisions, an additional researcher will be involved to resolve the conflict through a joint meeting. Each paper is marked as relevant, uncertain, or irrelevant. Any articles classified as irrelevant are directly excluded, and uncertain articles are left to the next round. Figure 2 shows the process in the specific literature screening stage.

- *Round1* Browse the paper title to eliminate irrelevant papers. If the title of the paper is not relevant to the topic of our research, we removed it directly according to EC3. Any paper that any researcher believes should be included or uncertain was passed to Round2.

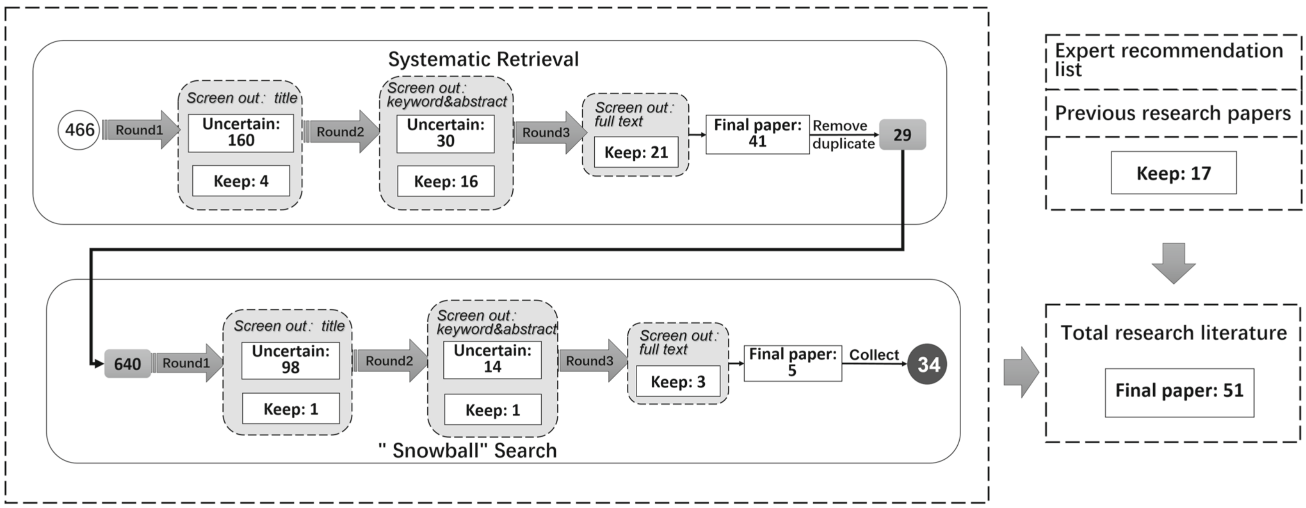


Fig. 3 Paper screening results

- *Round2* Read the abstract and keywords of the paper by Round1. Through the abstract and the title, we can roughly grasp the main topic of the paper and classify it according to the criteria. In this phase, we also search keywords across the paper to assist the selection. By searching for "review" and "profile" in the full text, we can efficiently evaluate whether the topic of this paper is related.
- *Round3* Remaining papers in this stage were carefully read and evaluated. In particular, we focus on whether the study includes profiling modeling methods and evaluation algorithms, or mines reviews to classify users' preferences.

4 Search and selection results

In the current paper, the literature for statistical analysis is obtained from three kinds of channels. One part is the literature of automatic retrieval mentioned in this paper. The second part is the literature selected from "snowball," and the third part is the expert recommendation. The results of the final literature are shown in Fig. 3.

We have carefully revised the search query of our study to include a broader scope of relevant studies. Thus, we carried out another three-round paper selection, resulting in 34 new papers selected from 1106 papers. Previously, we have analyzed 21 relevant papers. Since there are four duplicate papers between them, now we deal with 51 papers in total, significantly improving the comprehensiveness of our study. The titles and tagged information of the total research literature are shown in "Appendix A." We provide the tagging results of 51 papers by the link in Dropbox.[5]

[5] https://www.dropbox.com/sh/v63qqelddxg1lim/AABjLszZekUOtDiKV8q8aOROa?dl=0.

5 Analysis of research results

Based on the research results (Sect. 4), this section analyzes the research results and answers the corresponding research questions. In Sect. 5.1, we use the methods of bibliometrics to analyze the paper's attributes information. In Sect. 5.2, we developed a common process of user profiling based on reviews from four aspects. In Sect. 5.3, the research status and topics of each activity in the generic process are introduced in detail. Finally, we make constructive suggestions on challenges and potential opportunities in this area (Sect. 5.4).

5.1 Bibliometrics of user profiling publications RQ-1

In this section, the overall status of the selected studies is reported. We focus on the time distribution of the selected papers, as well as the publishing venue and the application domain. In addition, we use co-author networks to explore the relevance of authors. These dimensional indicators demonstrate the maturity of the field from different aspects and reflect the overall situation of research in this area.

5.1.1 The evolution trend over time

Because user profiling is a relatively new study in the big data field, it is interesting to observe the convergence of these papers. In particular, this is useful to identify when such areas have begun to converge. The time distribution of 51 papers is shown in Fig. 4. Since 2010, many researchers have focused on the potential value of user reviews. Most of them develop user preferences through data mining techniques. A plausible reason for this phenomenon is that the emergence of text analysis technology and unstructured text reviews can

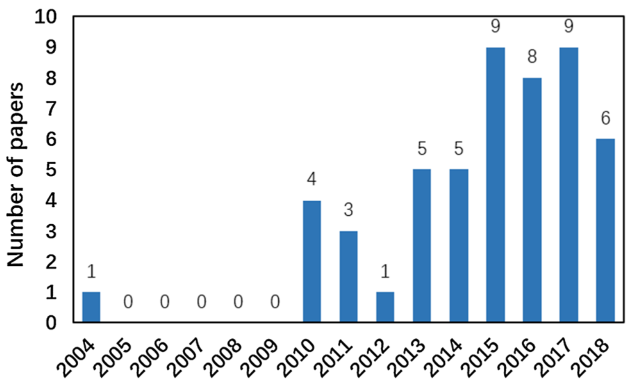


Fig. 4 Time distribution

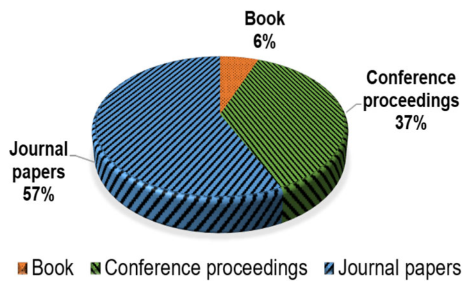


Fig. 5 Paper venue distribution

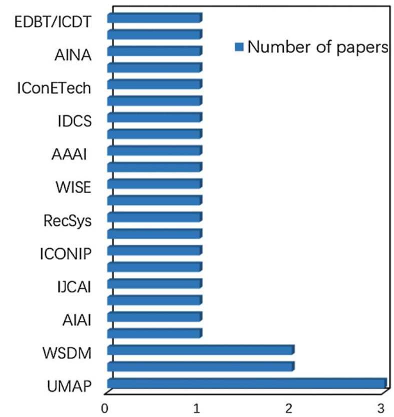


Fig. 6 Publication conference venue

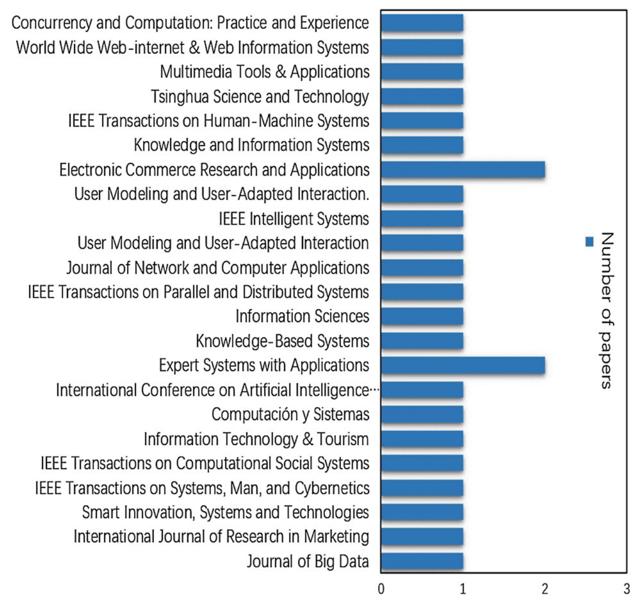


Fig. 7 Publication journal venue

be automatically extracted into structured forms [21]. At the same time, new technologies such as natural language processing (NLP) have gradually promoted the development of user profiles in software requirements area. The reduction in the number of papers in 2019 may be because our search time expired in September 2019.

5.1.2 Paper venues

Generally, in the early stages of a field, the number of conference papers is large. With the development of this field, the number of journal papers will gradually increase. As shown in Fig. 5, there are 19 studies (51%) published in the conference, and the rest are published in journals and books. As can be seen from the figure, this study has gradually transitioned from the early development stage to the steady development stage. Although it has been partially verified, a sophisticated research system has not yet been formed. In addition, we also want to know what important research venues are published in these papers? The results show that the amount of publication on the ACM user modeling, adaptation, and personalization (UMAP) conference is large (Fig. 6). The uniform distribution of journals in the surveyed literature proves that a more extensive community is not formed in this area. Overall, it is in the early stages of development in this field (Fig. 7).

5.1.3 Application domain

The distribution of the applications in the selected papers is displayed in Fig. 8. These specific application areas represent typical scenarios in this field. In different scenarios, we can abstract the multimodal user profiling model based on the user's social attributes, living habits, and consumer behavior.

The data mentioned in our research literature are concentrated on several main modules, such as products, restaurants, and tourism. Product reviews have been a hot topic because of the wide range of data sources [30]. The hotel is also an area that is often involved. According to the user's review, the aspect information can be analyzed, so that the preferred accommodation can be recommended according to the user's historical records [9]. Research on health and medical care has appeared more in the past two years, probably because people have paid more attention to medical health with the

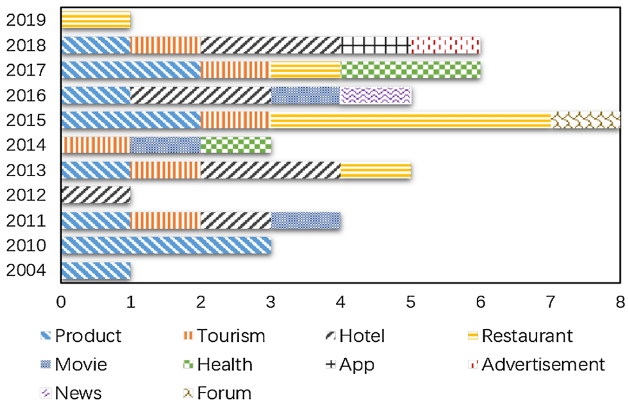


Fig. 8 Number of publications per year regarding application domain

deterioration of the ecological environment. However, the review data generated in other fields are not well studied, which limit the diversity of user review research. Therefore, this situation is not optimistic. The data need to be further expanded and aggregated to study the semantic information expressed by reviews in multiple domains.

5.1.4 Author connections

We want to know if there is a group effect in the RBUP field. Is there a connection between the authors of these papers? In Fig. 9, the circle represents the author entity and the edges between the entities represent the co-author relationship for each article. If the same author appears in multiple articles, it is represented as multiple connections. The results show that there are 36 clusters formed by 137 authors in 51 studies. According to Fig 9, we found that there is not much cooperation between authors and no large research community has been formed. In other words, there is a lack of communication and cooperation among researchers in this research field. However, the author of the cooperation between the authors still exists. For example, the authors of paper [S18] are Nostrawaliza Abdullah, Yue Xu, and Shlomo Geva, and the authors of paper [S07] are Noraswaliza Abdullah, Yue Xu, Shlomo Geva, and Jinghong Chen. Several identical authors appeared in these two articles, and we observed this in our results. This phenomenon indicates that, although there are differences in each research group, they also include interactions. In future research, communication and exchanges between authors in different regions should be strengthened, and the scope of research and the breadth of cooperation should be gradually expanded.

In addition, we counted their countries in order to get the author's distribution around the world. We would like to know if there are some concentrated countries or regions that have made significant contributions to the development of this field. We found that most of the authors were from

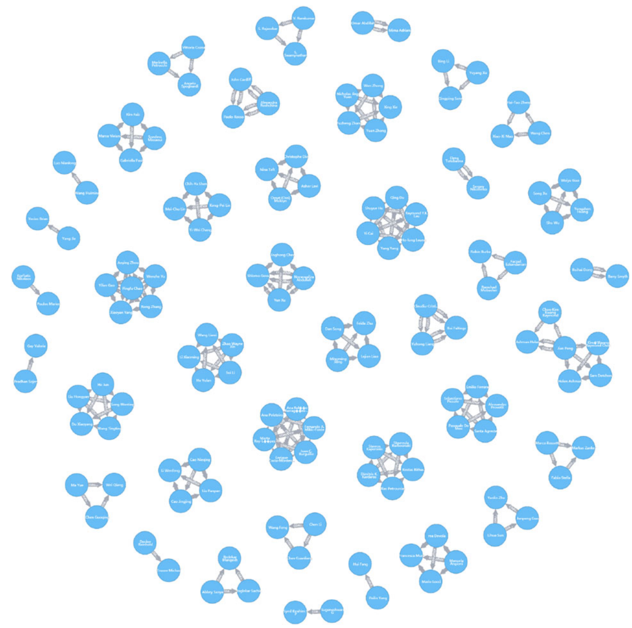


Fig. 9 Connections between authors

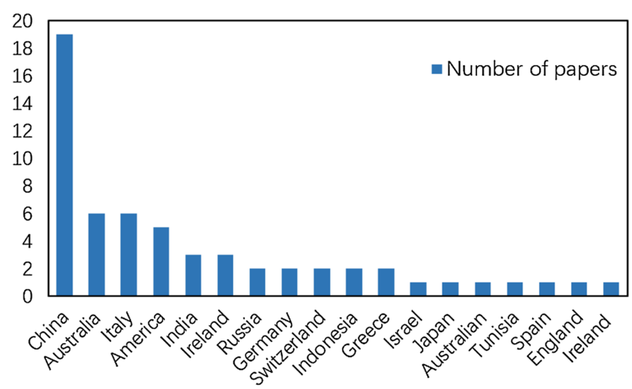


Fig. 10 Number of authors per country

China, followed by Australia and Italy (Fig. 10). According to the authors' associations of each article, Chinese researchers have more cooperation with each other, and they attach more importance to the field of user profiles based on reviews. Especially with the development of technology on the Internet, modeling people, mining user preferences, and recommending have become a crucial part of software requirements analysis.

5.2 Exploring the generic process of user profiling research RQ-2

For each field of research, the basic framework for analysis and processing is essential. The analysis of the generic process helps to segment the research process and more clearly define the direction available for the research area.

This section summarizes the general process of processing user profiles in the filtered literature.

Here, we have further detailed the method of obtaining the four stages of the generic process. The purpose is to clearly explain the labeling methods of generic processes. Therefore, the researchers labeled the obtained literature with different modules. The standard for labeling is the chapter name or the architecture diagram shown in the article. Indeed, in different articles, the author expressed the same concept or meaning as different expressions. Based on the expressions in the article, we synthesize the stages and activities in the research literature with the terms of synonyms and similar semantics. It can be seen from Fig. 11 that most of the researches preprocessed the review data after determining the research topic. However, some articles use raw text directly as input. The features of reviews obtained through natural language processing technology can represent the user's preferences, which can be used as input for the user profiling model. The user profile building method can classify similar people. The label feature of similar people can be applied to a variety of scenarios to measure the quality of the user profile model. Most articles group the general activities of user profiles into the following four steps: review data collection, review information mining, user profile modeling, and application. It should be noted that not all stages appearing in the researched articles are parts of the derived generic process, but only those that occur more than (41 times/80%) are included in the generic processes. These stages are important research steps in the field of user profiles, and they are closely related between each other.

By exploring this issue, we have found a unified research framework in this field. This research framework enables researchers to define the generic processes and essential activities of existing research.

5.3 The theme and valuable research aspects of each activity in generic process RQ-3

As shown in Fig. 12, we summarized a generic process based on 51 papers, including review data collection, review information mining, user profile modeling, and model application evaluation. The essential activities that exist in the research process include a large number of research directions and topics. In this section, we carry out a detailed expansion and analysis for each stage and discuss the methods they use, e.g., methods for mining review data, mature or emerging user profiling models used at the present stage, and directions widely used in subsequent research.

5.3.1 Review data collection

The collection and accumulation of review data is the basis of user profiling. The collection method is mainly through crawling or using existing classic data sets (Fig. 13). It takes more effort to get the data set by climbing, but it also has its advantages. As user behavior changes over time, we need to extend new user data to update user characteristics continually. The advantage of using crawlers is that they can be changed according to the requirement of researchers and projects. The crawling method is very flexible, so it provides researchers with rich data by collecting ratings, operation logs, or attribute information.

However, due to the explosive growth of review data in recent years, some publicly recognized data sets have gradually formed from the original data crawling stage (Fig. 13). A large number of researchers mine user behavior information based on the provided data sets to produce better recommendations. From the perspective of data evolution, the field has gradually gained attention, and more and more researchers are committed to the study of user profiles.

5.3.2 Review information mining

Solving data problems is a prerequisite for any research problem [25]. Data-driven scientific research has become a trend in the Internet age. In particular, various analysis methods have been generated based on user-generated content. However, more important is to improve the in-depth mining and understanding of content data.

Although the raw review data are unstructured that cannot be easily understood by the system, the advanced technology in the field of topic modeling and opinion mining (also called sentiment analysis) makes it possible to interpret reviews and extract useful elements from them. Figure 14 is the result of investigation and statistics. We get the definition of each label by referring to the definitions in other literature and combining the definitions in our research results. Specifically, researchers summarize the classification of review elements used in the surveyed literature. These classifications are then mapped onto existing review elements classifications. It should be noted that an article may have multiple tags. Figure 14 shows the review element classification and the corresponding number of papers.

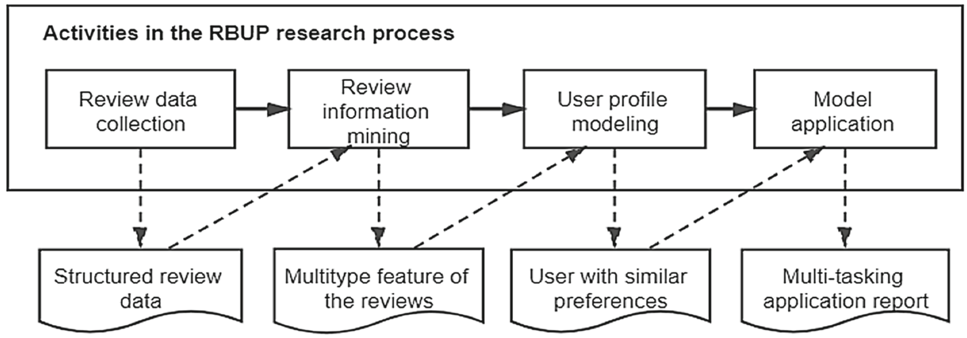
Review preference (RP) Review preference refers to the stable features of the user that can be expressed in the review. Review preferences represent users' attitudes toward the product, or the emotions they experience (sadness, anger, joy, excitement). The sentiment classifier can mark reviews with specific sentiments, reflecting the likelihood that a user will prefer a product. Review preferences can be divided into overall opinions, feature/aspect sentiment polarity, and sentiment strength. The specific explanation is given in Table 2.

Title	Architecture	Review collection	Review information mining	User profile modelling	Implementation	Application	Demonstration	Validation
S1		X	X	X		X		
S2		X		X		X		
S3	X	X		X		X		
S4		X	X	X		X		
S5		X	X	X		X	X	
S6		X		X		X		
S7	X	X		X		X		
S8		X	X	X		X		
S9		X	X		X	X	X	
S10		X	X	X				
S11		X	X	X				
S12		X	X	X		X		
S13		X	X	X				
S14	X	X	X			X		
S15		X		X		X		
S16		X	X	X		X		
S17		X	X	X		X		
S18		X	X	X		X		
S19		X	X	X		X		
S20		X	X	X		X		
S21		X	X	X		X		
S22		X	X	X				
S23		X	X	X		X		
S24	X	X		X		X		
S25		X	X	X		X		
S26		X	X	X		X		
S27		X	X	X		X		
S28		X	X	X		X		
S29		X	X	X		X		
S30		X	X			X		X
S31		X	X			X		
S32		X	X	X		X		
S33		X	X	X		X		
S34	X	X	X	X		X		
S35		X	X	X		X		
S36		X	X	X		X		
S37	X	X	X	X		X		
S38		X	X			X		
S39	X	X	X	X		X		
S40	X	X	X	X		X		
S41	X	X	X	X		X		
S42		X	X	X		X		
S43	X	X	X	X		X		
S44		X	X	X		X		
S45		X	X	X	X	X		
S46		X	X	X		X		
S47	X	X	X	X		X		
S48		X				X		
S49	X		X	X		X		
S50	X		X	X		X		
S51		X				X		
Total	**13**	**49**	**43**	**44**	**2**	**47**	**2**	**1**

Fig. 11 Data statistics of generic process

Review topic (RT) Review topics are key opinions that reflect user reviews. The topic extraction method can mine the user's reviews with some keywords and then cluster similar topics according to the keywords. The research of the review topic can be divided into two categories: frequency-based approach and topic model algorithm.

Fig. 12 Generic process framework



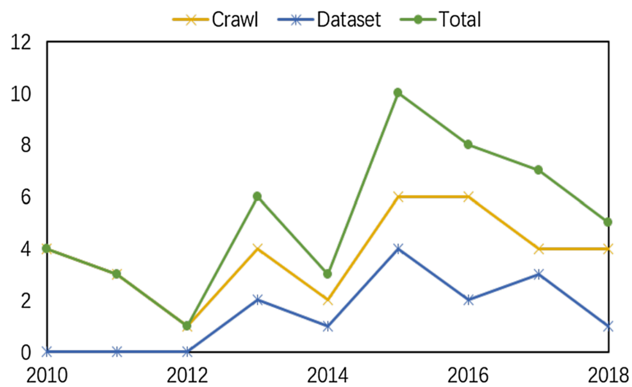


Fig. 13 Data source regarding user profiling per year

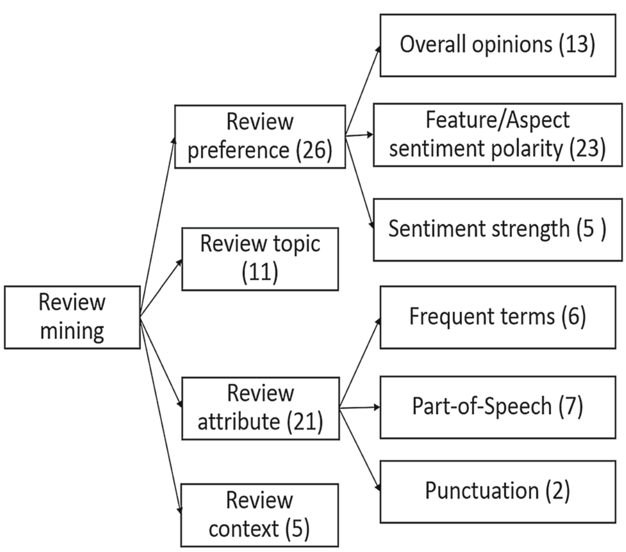


Fig. 14 Review elements classification (Number of papers)

– *Frequency-based approach* Word frequency is an important lexical feature in the review. The main research method is first to extract frequently occurring nouns based on a set of seed words and then group the nouns by topic or according to a predefined dictionary [20].
– *Topic model algorithm* Topic extraction is an important research method in text processing, such as latent Dirichlet allocation (LDA) and probabilistic latent semantic analysis (pLSA). LDA uses the Gibbs sampling method and the Dirichlet distribution. pLSA extracts implicit features based on the EM (maximum expectation) algorithm. The above methods all use the review topic to predict the user's label, thereby improving the quality of the recommendation system.

Review attribute (RA) The review attribute focuses on the structure of the text and the lexical mining of the word. Its purpose is to explore the features of texts with different preferences from the sentence itself. The review attribute contains a variety of information, such as frequent terms, part of speech, and punctuation.

– *Frequent terms* The word frequency can be used to determine the representativeness of each term in the review. The word-based keyword extraction method, term frequency-inverse document frequency (TF-IDF), can be used to evaluate the word importance of a file in a corpus.
– *Part of speech (POS)* POS is the way to discover and extract user demands from the lexical level. In general, we use adjectives to express positive (good) or negative (bad) opinions, using nouns as the subject of reviews.
– *Punctuation* The punctuation marks in the reviews show the degree of preference in the user experience process. For example, symbols such as "!", "!!", "?" in the text have an emphasis on the emotion.

Review context (RC) The review context refers to the user's opinion on a specific aspect. This kind of contextual opinion can reflect the background conditions of an item or a specific feature, which can be discovered from reviews through keyword matching, rule-based reasoning, or a LDA-based classifier. For example, "When I take pictures with this camera at night, its image quality is not good." "At night" is the context, "image quality" is the characteristic, and "bad" is the negative opinion. The context of the review highlights the background and conditions under which the user's demands occur.

Table 2 Review preference

RP	Description	Example
Overall opinions	The overall preferences of the reviews are expressed in terms of probability values or are divided into three categories: negative, positive, and neutral	Ratings or emojis (such as smiley and sad faces) can be used as a symbol of emotion. They can be combined with review words to infer the reviewer's overall score
Feature or Aspect sentiment polarity	The user's emotions are more than one type. People will suggest certain features or aspects of the item based on their own experience	When evaluating a restaurant, users usually pay attention to the atmosphere, service, taste, etc., of the restaurant, and for different attention points, the user may give the opposite conclusion, e.g., "The taste is good, but the restaurant environment is poor"
Sentiment strength	Different degrees of preference for the same emotion appear in the comments. These sentimental words reflect different levels of the same emotion, helping to describe the user's real emotions better	Positive emotions will include good, great, excellent, and other vocabularies. Negative emotions will include words such as bad and terrible

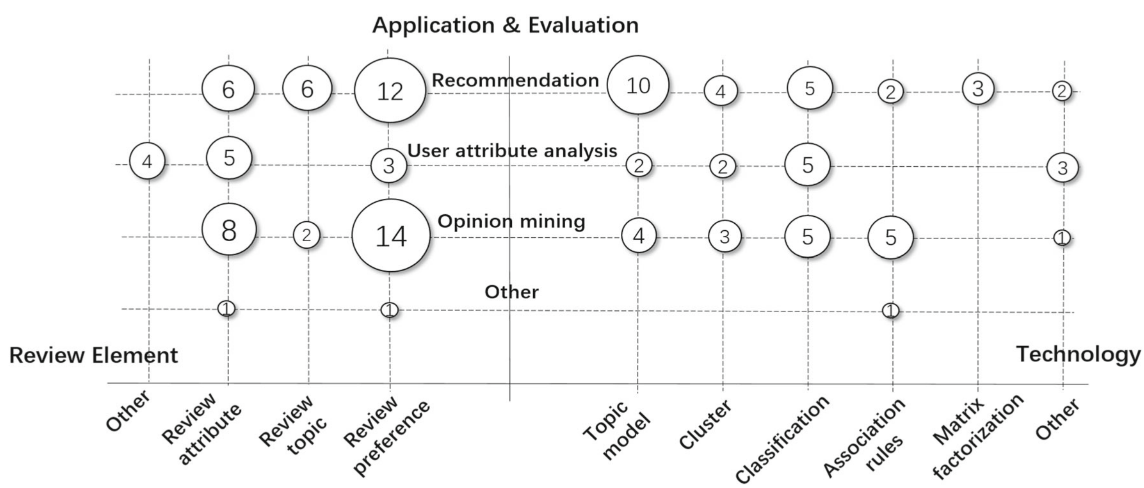


Fig. 15 Systematic map of RBUP papers (Number of papers)

Besides, we analyzed the relationship between the review element and application scenarios. As shown in the left side of Fig. 15, review preference is a frequently used method in the recommendation. We found that many papers used aspect-level information of reviews, which means that review mining can effectively help to describe user characteristics. At the technical level, techniques such as topic models and keyword extraction methods provide tools for extracting valid information from reviews. The attribute information of the review is often used as side information for auxiliary research. The combination of multiple review elements can reflect user preferences and improve the accuracy of user profiles.

5.3.3 User profile modeling

User profile modeling is an important part of the generic process. Figure 16 is the result of our investigation and statistics. Through data statistics, we synthesized techniques commonly used in the field of user profiles. The classification was obtained through the mapping operation. Therefore, it will be consistent with the existing classification section. However, the classification results obtained through our statistics are suitable for the RBUP domain. We summarize the types of techniques commonly used in the field of user profiles, including topic models, clustering, classification, and association rule models (Fig. 16). We have given the concepts and content of each techniques to improve the validity and repeatability of our research. Table 3 shows the specific modeling techniques in the literature investigated in the current work. Moreover, we mapped the distribution of user profiling techniques in different application scenarios. As shown in the right side of Fig. 15, it can be seen that the topic model is the most widely used in the user profiling field.

Topic model The topic model is a statistical model that clusters the latent semantic structure of a corpus in an unsupervised learning manner [26]. The topic is a probability distribution with all the features in the text as a support set, indicating how often the character appears in the subject, that is, features with high relevance to the subject have a greater probability of appearing. The topic model plays a crucial role

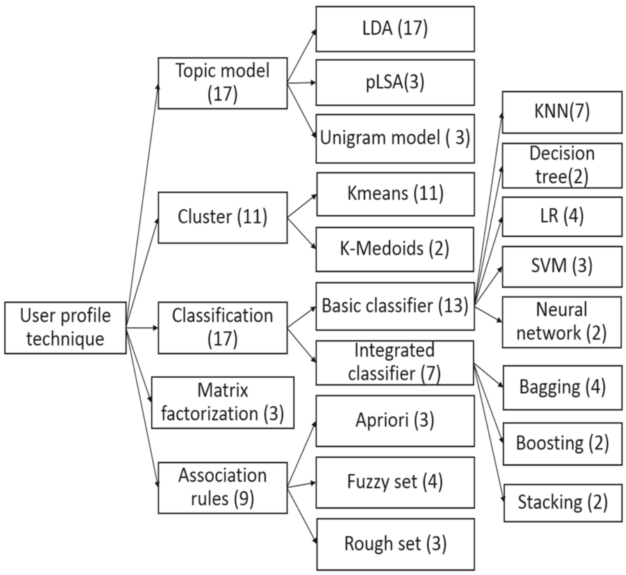


Fig. 16 Technique classification of user profiles (Number of papers)

in the field of text representation, dimensionality reduction, and text recommendation systems.

Clustering Cluster analysis is a statistical analysis technique that divides subjects into relatively homogeneous clusters. It can be used as a tool to obtain data distribution. Through the results of clustering, we can observe the characteristics of each cluster and focus on analyzing specific clusters.

Classification Classification is a supervised learning algorithm. According to the basic classification methods of labels, it mainly includes decision tree, naive Bayes, neural network, K-nearest neighbor, support vector machine, and so on. There are also integrated learning algorithms for combining basic classifiers. The representative algorithms for integrated learning are the random forest, AdaBoost, Xgboost, etc.

Association rule Association rule mining is a widely applied method for discovering frequent patterns, correlations, and causal structures in transaction and relational databases, as well as in other information repositories. Therefore, for a given set of items, the association rule mining technique allows defining rules to predict the occurrence of one item (or multiple items) while giving other items in the same item set.

5.3.4 Model application

Multitasking applications based on the user profile model are described in this section. Based on the generic processes, we found that most of the research conducted subsequent evaluation and analysis after constructing user profiles. According to the research literature in this paper, the application scenarios include user attribute analysis and recommendation systems (Fig. 17).

User attribute analysis User attribute analysis is a method of mining demographic information from the user-generated text. For example, most of the papers predict the user's age, gender, or group characteristics through user reviews and behavioral data. It mainly contains the following two tasks.

– *User demographic attribute prediction.*

  User demographic attribute prediction refers to predicting user's age, gender, occupation, and other related attribute information from the information generated by the user. Natural language processing techniques are often used in this field, including feature classifiers, extended topic models, and deep neural networks (convolution and recursive architecture). In [43], the author infers users' demographics and proposes a simple general location to profile (L2P) framework by location check-ins.
– *Author attribute recognition.*

  Author attribute recognition focuses on the notion of authorship attribution, which is to collect data from public social media sites to analyze the putative authors to see if they are the same person. The analysis comprises a binary n-gram method, which was previously shown to be effective at accurately identifying authors on a training set from the same authors, while [28] shows how authors on different social media turn out to be the same author.

Recommendation system The recommendation system is another broad application of RBUP. The recommendation task is to match the appropriate items based on the user's historical or real-time behavior. User profiles are not the purpose of the recommendation system, but the critical phases.

The problem of data sparse and cold start is often generated in the recommendation system. The user–item matrix is sparse because the user reviews or rated items account for a lower proportion of all items. Many articles [13,39] have proposed corresponding solutions to solve these problems. The technique of matrix decomposition is used in the recommendation system [6], which reduces the dimension of the matrix. In knowledge-based recommendations, ontology [3] and knowledge graphs [5] solve the cold start problems with auxiliary information. Relational extension and reasoning techniques in knowledge graphs can improve the ability of user behavior representation learning in big data environments. In addition, the computational framework and fast algorithms in the real-time recommendation domain are also important application scenarios.

Table 3 User profiling strategy

Technology category	Specific profiling studies	Paper
Topic Model	A Bayesian model that links a traditional collaborative filtering (CF) technique with a topic mode.	[S01]
	A topic criteria (TC) model, which exploits the topic model method to extract latent features from textual reviews and discuss its application for several application scenarios in tourism	[S18]
	Topic profile collaborative filtering method was used to verify the review texts, and topic profiles indeed correlate with ratings	[S16]
	The coupled topic model (CoTM) was proposed, which could capture both the content-based preferences and collaborative preferences	[S20]
	A partially labeled topic model (PLDA) was used to compare the effects of feature-rich classifiers, extensions of topic models, and deep neural networks (both convolutional and recurrent architectures)	[S21]
	LDA is used to find hidden aspects from the review text while using regression models to detect the relationship between the user	[S05][S08][S10][S13]
	The unigram model describes user preferences by defining two language models: a positive language model and a negative language model	[S04]
Clustering	K-Medoids clustering approach is used to automatically segment users who share the same level of diversity in their user profiles	[S02]
	Factorized latent Dirichlet allocation (FLDA) technology is used to integrate the cluster semantically related aspects	[S14]
	The KNN and K-means algorithms based on the uncertainty theory can be used to explore the accuracy of the recommender systems based on the opinions of various users	[S17]
	An approach combines numerous elements, including unsupervised clustering to build a vocabulary for hotel aspects, semantic analysis to understand sentiment toward hotel features, and the profiling of intent and nationality groups	[S27][S30]
Classification	tensor factorization is used to draw out low-dimensional representations of users' intrinsic check-in preferences. Meanwhile, the extracted features are used to train predictive models for inferring various demographic attributes	[S26]
	Comparison of the classification model (SVM and LambdaMART) and the regression model (linear regression, Poisson regression, and boosted tree regression)	[S03][S05][S12][S22]
	Neural network classifiers have gradually become one of the main classification models, especially LSTM-based recurrent networks and convolutional neural network (CNN) models	[S21]
	Deep semantic hybrid recommendation method (DSHRM) utilizes variational autoencoder (VAE) model to extract user profiles and item representations and make sure both of them are in a consistent latent semantic space	[S24]
	The K-nearest neighbor method (KNN) is used to calculate the user similarity, and then the reviews with the same characteristics are clustered	[S15]
	The semi-supervised sentiment classification model is proposed. By calculating the review similarity, the label propagation algorithm is used to predict the sentimental polarity of the unlabeled reviews	[S25]
Association	TF-IDF and fuzzy sets qualitative comparative analysis are used to model user interests	[S11][S29]
	The association between product attribute values can be represented by rules	[S28]
	A priori algorithm is used to mine implicit data association from Trip advisor hotel reviews	[S31]
	Adaptive collaborative filtering (ACF) for users and product profiles is presented. Use rough set association rules to mine the preferences of goods and users	[S06]
Matrix factorization	Matrix factorization (MF) techniques are used to reduce dimensionality and solve data sparsity in recommendation systems, including eigendecomposition and singular value decomposition (SVD)	[S24]

5.4 Challenge and promising research directions RQ-4

User profiles use data mining techniques to learn user preferences and build related models. With the development of natural language processing technology and big data platforms in social networks, user reviews have played a major role in the field of user profiles. However, according to the above analysis, it can be concluded that the field is still in the development stage. Therefore, possible future opportunities and challenges are discussed in this section.

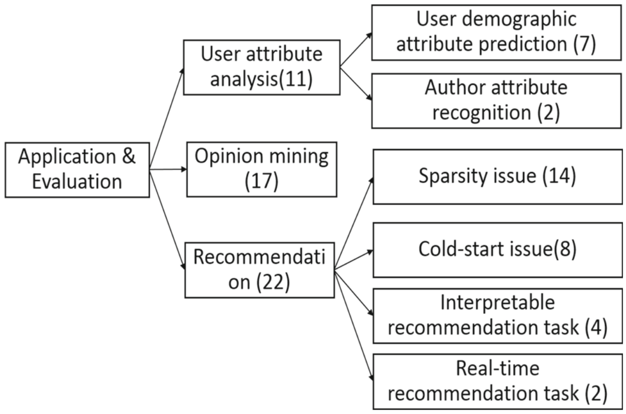


Fig. 17 Application scenario distribution (Number of papers)

5.4.1 Knowledge fusion among multiple data sources

In the research of user profile, it is inevitable to mention the data issue. Both RQ1 and RQ2 involve the topic of data analysis. The survey results in Sect. 5.1.3 show that with the widespread application in social platforms, review information has been extended to various fields. However, most of the research only focuses on the information in this field, and cross-field data integration is still relatively rare. In Sect. 5.3.2 and Fig. 15, most of the current data types used in the analysis focus on review preferences. In addition, when further studying the user profile model, we can consider the integration of user information on different social platforms. This allows integration of cross-domain knowledge to achieve the integrity of information in a single domain. At the same time, the diversity of data patterns brings more possibilities for user profiling to extract the specific user preferences. The difference in application context also deepens the complexity of data analysis. For example, in the field of transportation, multi-source data are used to depict profiles of pedestrians and traffic behavior. By analyzing the similarities and differences of travelers' traffic choice behavior under different traffic conditions, it can provide quantitative support for traffic planning decisions [31]. In addition, the integration of multi-source information also includes geographic location information, time information, and image information.

In order to establish a more comprehensive user profile, we need to build a multi-dimensional data model to solve the sparsity problem. The knowledge graph offered a new direction for solving this problem. We can use deep semantic learning-based entity recognition, relationship extraction, and attribute extraction techniques to extract multi-source user information. In particular, the knowledge reasoning method and entity extension techniques based on the deep neural language model can realize the dynamic evolution and update of the knowledge graph. At the same time, the knowledge fusion technology based on cross-modal shared subspace learning can realize the information fusion. Collaborative learning and multi-channel models based on deep neural networks may be a technique worth trying.

5.4.2 In-depth review mining with advanced NLP and ML/DL techniques

In Sect. 5.3.3, user profile modeling technique has been discussed in detail, in which the topic model is a relatively common technique, especially in the field of recommendation. From Fig. 15, we can see that most of the techniques of mining user profiles are focused on the related aspects of machine learning, while some of deep learning techniques have not been well used in RBUP. However, both machine learning (ML) and the deep learning (DL) technologies have shown significant results in many scenarios, such as text analysis task. In particular, there is little work to transfer these techniques to the field of user profiles. The development of user profiles has been driven by these new techniques, enhancing the reliability and interpretability of results from the data level. Next, we explain the development direction and challenges of user profiling from these two aspects.

Review mining based on machine learning The accuracy of user profiling is highly dependent on the quality of the data source. With the development of NLP technology, reviews can be used in combination with graphics algorithms, which provides in-depth ideas for characterizing user behavior.

In addition, in order to obtain precise semantic analysis, technologies like Word2vec [41] and Doc2vec [23] continue to emerge. These new technologies are innovative in dealing with dimensionality disasters, words similarity, and model generalization capabilities. Based on word vector technology, text similarity can be evaluated and reviews with the same semantics can be aggregated. From the perspective of the syntax tree, structural information can assess the quality of reviews. For text classification and clustering algorithms, specific algorithmic improvements such as support vector machine (SVM), K-nearest neighbor algorithm (KNN), and naive Bayes are also worth studying.

Review mining based on deep learning Traditional user analysis uses non-end-to-end approaches in feature extraction and clustering. Multiple steps are carried out independently in turn. Therefore, the quality of each task affects the results of the full training stage. However, deep learning uses the end-to-end approach to avoid a lot of manual tagging problems. For example, various types of data can be uniformly embedded into the same vector space, and then the representation of heterogeneous data can be solved by building a deep prediction model [23].

During the research process, the role of users is continuously changing with time. For example, the pregnancy label is time-sensitive and differs from the user's previous preferences. There are many methods for processing temporal data. For example, the recurrent neural network (RNN) model can memorize temporal information based on the user's check-in records, especially in the user check-in recommendation [40]. Therefore, dynamic user profiles can complement the shortcomings of traditional static user profiles. How to combine the long-term preferences of users with short-term needs based on context factors is also a research direction.

With the maturity of deep learning technology, deep neural network technique [24] can automatically extract the feature representation of the original data, among which the widely used methods are DeepWalk [29], SDNE [35], etc. Automatic learning of high-level feature interaction mode through the neural network can make up for the limitation of artificial feature engineering. However, the deep neural network and knowledge graph [5] are in the early stage of user profiling technology. The current user profiling technology is not sufficiently scalable in behavior analysis and multi-source information fusion. User profiling techniques for massive user behavior data require to be driven by more sophisticated algorithmic research.

5.4.3 User privacy protection and sharing of profiling data from different platforms

Currently, multiple types of user data are distributed across different platforms. For example, the search engine has the user's search and Web browsing records, and e-commerce Web sites have the user's product browsing, shopping, and purchasing information. User data from these different platforms are valuable to user profiles and provide complementary information to each other, helping to build a comprehensive user representation. However, sharing user information directly between platforms may leak and compromise user privacy. How to use the user information of different platforms to implement collaborative user profiles without transmitting and sharing user data is a direction worth studying.

6 Threats to validity

In this section, in order to indicate the possible threats in this work and to evaluate the validity of the SMS, we follow the classification of validity in paper [2] to summarize the threats in the research process. For each type of validity, the classification definition of the validity threats given by the secondary research method (SMS, SLR) [2] is used in this paper. The specific content is divided into the following three threats: study selection validity threats, data validity threats, and research validity threats.

6.1 Study selection validity threats

Study selection validity can be identified in the search process and screening phase of secondary studies planning. Study selection issues threat the validity of searching and including primary studies in the examined set [2]. This involves threats like the selection of digital libraries, search string construction, study selection bias, etc. For this article, we first used four public digital libraries. Although we try to supplement the important literature in this field, we cannot guarantee that our survey will include all the studies. In response to this threat, the method of snowballing was added to the stage of literature selection. In addition, we added the manual search to include the literature recommended by experts in our research. The second threat is the inclusion/exclusion of the gray literature. Based on the goal of the study, including or excluding the gray literature can pose a threat. For this threat, we set criteria for literature screening phase before data processing. In particular, we revisited these criteria after the pilot study. In this current study, the gray literature should be considered in multi-vocal literature reviews (MLRs), in which practitioners' view should be examined.

6.2 Data validity threats

This category includes threats that can be identified in the data extraction and analysis phases of secondary studies and threaten the validity of the extracted dataset and its analysis [2]. Examples of threats in this category are data collection bias, researcher bias, etc. In terms of data validity threats, data extraction may exist inaccuracies, because during the screening process, researchers from different backgrounds in the field are involved. In the subsequent data analysis stage, the article may not be carefully performed or might not follow strict guidelines due to the subjective factors of the researchers. For example, the same concept might be inconsistently classified into two primary studies. This leads to inaccuracies in the data statistics. In order to mitigate such threats, we have first conducted a pilot study to ensure that different researchers can reach a consensus regarding the tagging criteria. In addition, during the data tagging stage, when there are conflicting tagging results, we will introduce additional researchers to discuss and achieve a consensus in the end.

6.3 Research validity threats

Research validity threats can be identified in all SMS processes and concerned with the overall research design, such as repeatability, and coverage of research questions. First of all, for the repeatability problem of this work, researchers screened papers against strict include and exclude criteria to ensure the repeatability of screening process. In addition, the developed protocol of this article was sorted out in the process of SMS, so that researchers can make subsequent additions and extensions. There is also a threat to the coverage of research questions. In Sect. 3.1, we proposed four types of important research questions, focusing on the development status, generic processing flow, challenges, and directions. These research questions covered important activities in the evolution and development of the field.

7 Conclusion

In this work, we conducted a systematic mapping study of review-based user profiles in order to provide references for other researchers and promote the development of the field. Specifically, 51 papers were carefully selected from 2478 papers for standardization and systematic research. Based on the status of RBUP, we focus on the generic analysis process of RBUP and identify challenges and potential research directions. This generic process framework can help researchers discover important topics in the field and provide a basis for researchers to propose new frameworks.

Through mapping study, we found that RBUP, as an emerging research hot spot, has received widespread attention in recent years. The results show that although traditional methods have been continuously improved, they are not sufficient to unleash the full potential of large-scale user reviews, especially the use of heterogeneous data for multi-dimensional user profiling. Implicit information in reviews can be mined by more advanced deep learning techniques to improve the accuracy of user profiling models. Therefore, in-depth review of data mining and utilization is also a direction worthy study.

Acknowledgements This work is supported by the National Key R&D Program of China (Nos. 2017YFC0803300, 2017YFC0803307) and the National Natural Science of Foundation of China (Nos. 61902010, 91546111, 91646201).

Appendix A: RBUP papers

(S1) Jiang M., Song D., Liao L., Zhu F.: A Bayesian recommender model for user rating and review profiling. Tsinghua Science and Technology. 2015).

(S2) Eskandanian F., Mobasher B., Burke R.: A Clustering Approach for Personalizing Diversity in Collaborative Recommender Systems. In: 2017)

(S3) Cardiff J., Cardiff J., Rosso P.: A comparative evaluation of personality estimation algorithms for the twin recommender system. In: International Workshop on Search & Mining User-generated Contents, vol. 2011)

(S4) Missaoui S.: A Language Modeling Approach for the Recommendation of Tourism-Related Services. In: Acm Sigapp Symposium on Applied Computing Sac, vol. 2017)

(S5) Gao Y., Yu W., Chao P., Rong Z., Zhou A., Yang X.: A Restaurant Recommendation System by Analyzing Ratings and Aspects in Reviews. 2015).

(S6) Abdullah N., Yue X., Geva S.: A Recommender System for Infrequent Purchased Products based on User Navigation and Product Review Data. In: Web Information Systems Engineering-wise Workshops-wise International Symposium Wiss, vol. 2010)

(S7) Roshchina A., Cardiff J., Rosso P.: A User Profile Construction in the TWIN Personality-based Recommender System. 2011).

(S8) Dragoni M.: A Three-Phase Approach for Exploiting Opinion Mining in Computational Advertising. IEEE Intelligent Systems. 32. 21 (2017).

(S9) Sano M.: An Affect-Based Text Mining System for Qualitative Analysis of Japanese Free Text. In: 2004)

(S10) Yang P., Hui F.: Opinion-based User Profile Modeling for Contextual Suggestions. In: Conference on the Theory of Information Retrieval, vol. 2013)

(S11) Kardaras D.K., Kaperonis S., Barbounaki S., Petrounias I., Bithas K.: An Approach to Modelling User Interests Using TF-IDF and Fuzzy Sets Qualitative Comparative Analysis. 2018).

(S12) Angioni M., Devola A., Locci M., Mura F., Tuveri F., Varchetta M.: An Evaluation Method for the Performance Measurement of an Opinion Mining System. 2018)

(S13) Agreste S., Meo P.D., Ferrara E., Piccolo S., Provetti A.: Analysis of a Heterogeneous Social Network of Humans and Cultural Objects. IEEE Transactions on Systems Man & Cybernetics Systems. 45. 559 (2014).

(S14) Sun L., Guo J., Zhu Y.: Applying uncertainty theory into the restaurant recommender system based on sentiment analysis of online Chinese reviews. World Wide Web-internet & Web Information Systems. 1 (2018).

(S15) Jian P., Choo K.K., Ashman H.: Astroturfing Detection in Social Media: Using Binary n-Gram Analysis for Authorship Attribution. In: 2016 IEEE Trustcom/BigData SE/ISPA, vol. 2016)

(S16) Musat C.C., Liang Y., Faltings B.: Recommendation using textual opinions. In: International Joint Conference on Artificial Intelligence, vol. 2013)

(S17) Dragoni M.: Computational advertising in social networks: an opinion mining-based approach. In: 1798. ACM,2018)

(S18) Rossetti M., Stella F., Zanker M.: Analyzing user reviews in tourism with topic models. Information Technology & Tourism. 16. 5 (2016).

(S19) Dong R., Smyth B.: From More-Like-This to Better-Than-This: Hotel Recommendations from User Generated Reviews. 2016).

(S20) Wu S., Guo W., Xu S., Huang Y., Wang L., Tan T.: Coupled Topic Model for Collaborative Filtering With User-Generated Content. IEEE Transactions on Human-Machine Systems. 46. 908 (2017).

(S21) Tutubalina, E., Nikolenko, S.: Demographic prediction based on user reviews about medications. Computacion y Sistemas 21(2), 227C241 (2017)

(S22) Tutubalina E., Nikolenko S.: Exploring convolutional neural networks and topic models for user profiling from drug reviews. Multimedia Tools & Applications. 1 (2017)

(S23) Yang Y., Hu S., Yi C., Du Q., Leung H.F., Lau R.Y.: Exploring Reviews and Ratings on Reviews for Personalized Search. In: Revised Selected Papers of the Icwl International Workshops on Current Developments in Web Based Learning, vol. 2015)

(S24) Chen W., Zheng H.T., Mao X.X.: Extracting Deep Semantic Information for Intelligent Recommendation. In: International Conference on Neural Information Processing, vol. 2017)

(S25) Xu Y., Bing L.: Sentiment Classification Incorporating User Profile. In: International Conference on Information Science & Control Engineering, vol. 2017)

(S26) Zhong Y., Yuan N.J., Zhong W., Zhang F., Xie X.: You Are Where You Go:Inferring Demographic Attributes from Location Check-ins. 2015).

(S27) Abdillah O., Adriani M.: Mining User Interests through Internet Review Forum for Building Recommendation System. In: IEEE International Conference on Advanced Information Networking & Applications Workshops, vol. 2015)

(S28) Abdullah N., Yue X., Geva S., Chen J.: Infrequent Purchased Product Recommendation Making Based on User Behaviour and Opinions in E-commerce Sites. In: IEEE International Conference on Data Mining Workshops, vol. 2011)

(S29) Ramkumar V., Rajasekar S., Swamynathan S.: Scoring products from reviews through application of fuzzy techniques. Expert Systems with Applications. 37. 6862 (2010).

(S30) Levi A., Mokryn O., Diot C., Taft N. Finding a needle in a haystack of reviews:cold start context-based hotel recommender system demo. Acm Conference on Recommender Systems. 2012.

(S31) Cozza V., Petrocchi M., Spognardi A.: Mining implicit data association from Tripadvisor hotel reviews. 2018)

(S32) Malmi E., Weber I.: You Are What Apps You Use: Demographic Prediction Based on User's Apps. 2016).

(S33) Zhang Y.: Incorporating Phrase-level Sentiment Analysis on Textual Reviews for Personalized Recommendation. 2015).

(S34) K. L., Y. C., C. S., M. L.: Leveraging Online Word of Mouth for Personalized App Recommendation. IEEE Transactions on Computational Social Systems. 5. 1061 (2018).

(S35) Huiming, W. and L. Nianlong. Collaborative filtering enhanced by user free-text reviews topic modelling. in 2014 International Conference on Information and Communications Technologies (ICT 2014). 2014.

(S36) Liu, H., et al., Combining user preferences and user opinions for accurate recommendation. Electronic Commerce Research and Applications, 2013. 12(1): p. 14–23.

(S37) Yang, J. and B. Yecies, Mining Chinese social media UGC: a big-data framework for analyzing Douban movie reviews. Journal of Big Data, 2016. 3(1): p. 3.

(S38) Korfiatis, N. and M. Poulos, Using online consumer reviews as a source for demographic recommendations: A case study using online travel reviews. Expert Systems with Applications, 2013. 40(14): p. 5507–5515.

(S39) Pradhan, S. and V. Gay. Introducing Patient and Dentist Profiling and Crowdsourcing to Improve Trust in Dental Care Recommendation Systems. 2014. Berlin, Heidelberg: Springer Berlin Heidelberg.

(S40) Zhao, W.X., et al., Exploring demographic information in social media for product recommendation. Knowledge and Information Systems, 2016. 49(1): p. 61–89.

(S41) Ma, Y., G. Chen and Q. Wei, Finding users preferences from large-scale online reviews for personalized recommendation. Electronic Commerce Research, 2017. 17(1): p. 3–29.

(S42) Cao, N., et al. Sentimental Preference Extraction from Online Reviews for Recommendation. 2015. Cham: Springer International Publishing.

(S43) Chen, L. and F. Wang, Preference-based clustering reviews for augmenting e-commerce recommendation. Knowledge-Based Systems, 2013. 50: p. 44–59.

(S44) Barragns-Martnez, A.B., et al., A hybrid content-based and item-based collaborative filtering approach to recommend TV programs enhanced with singular value decomposition. Information Sciences, 2010. 180(22): p. 4290–4311.

(S45) S., M., et al., KASR: A Keyword-Aware Service Recommendation Method on MapReduce for Big Data Applications. IEEE Transactions on Parallel and Distributed Systems, 2014. 25(12): p. 3221–3231.
(S46) Decker, R. and M. Trusov, Estimating aggregate consumer preferences from online product reviews. International Journal of Research in Marketing, 2010. 27(4): p. 293–307.
(S47) Peng, J., K.R. Choo and H. Ashman, Bit-level n-gram based forensic authorship analysis on social media: Identifying individuals from linguistic profiles. Journal of Network and Computer Applications, 2016. 70: p. 171–182.
(S48) Peng, J., R.K.K. Choo and H. Ashman. Astroturfing detection in social media: Using binary n-gram analysis for authorship attribution. 2016: Institute of Electrical and Electronics Engineers Inc.
(S49) S., A., J. S. and B. M. Comparison of Personalised Systems. in 2013 6th International Conference on Emerging Trends in Engineering and Technology. 2013.
(S50) Chen, L., G. Chen and F. Wang, Recommender systems based on user reviews: the state of the art. User Modeling and User-Adapted Interaction, 2015. 25(2): p. 99–154.
(S51) Suganeshwari, G. and S.P. Syed Ibrahim. A Survey on Collaborative Filtering Based Recommendation System. 2016. Cham: Springer International Publishing.

References

1. Ahn, H.J., Kang, H., Lee, J.: Selecting a small number of products for effective user profiling in collaborative filtering. Expert Syst. Appl. **37**(4), 3055–3062 (2010). https://doi.org/10.1016/j.eswa.2009.09.025
2. Ampatzoglou, A., Bibi, S., Avgeriou, P., Verbeek, M., Chatzigeorgiou, A.: Identifying, categorizing and mitigating threats to validity in software engineering secondary studies. Inf. Softw. Technol. **106**, 201–230 (2019). https://doi.org/10.1016/j.infsof.2018.10.006
3. Ananthapadmanaban, K.R., Srivatsa, S.K.: Personalization of user profile: creating user profile ontology for tamilnadu tourism. Int. J. Comput. Appl. **23**(8), 42–47 (2011)
4. Budgen, D., Kitchenham, B., Charters, S., Turner, M., Brereton, P., Linkman, S.: Preliminary results of a study of the completeness and clarity of structured abstracts. In: International Conference on Evaluation & Assessment in Software Engineering (2007)
5. Catherine, R., Cohen, W.: Personalized recommendations using knowledge graphs: a probabilistic logic programming approach. In: ACM Conference on Recommender Systems (2016)
6. Chen, W., Zheng, H.T., Mao, X.X.: Extracting deep semantic information for intelligent recommendation. In: International Conference on Neural Information Processing, pp. 134–144. Springer (2017)
7. Cicchetti, A., Ciccozzi, F., Pierantonio, A.: Multi-view approaches for software and system modelling: a systematic literature review. Softw. Syst. Model. **18**, 3207 (2019)
8. Cufoglu, A.: User profiling-a short review. Int. J. Comput. Appl. https://doi.org/10.5120/18888-0179 (2014)
9. Dong, R., Smyth, B.: From more-like-this to better-than-this: hotel recommendations from user generated reviews. In: UMAP16, Halifax, NS, Canada. ACM (2016)
10. Dong, X., Li, T., Li, X., Song, R., Ding, Z.: Review-based user profiling: a systematic mapping study Enterprise Business-Process and Information Systems Modeling, pp. 229–244. Springer, Cham (2019)
11. Hojaji, F., Mayerhofer, T., Zamani, B., Hamou-Lhadj, A., Bousse, E.: Model execution tracing: a systematic mapping study. Softw. Syst. Model. **18**, 3461 (2019)
12. Horkoff, J., Aydemir, F.B., Cardoso, E., Li, T., Maté, A., Paja, E., Salnitri, M., Piras, L., Mylopoulos, J., Giorgini, P.: Goal-oriented requirements engineering: an extended systematic mapping study. Requir. Eng. **24**(2), 133–160 (2017). https://doi.org/10.1007/s00766-017-0280-z
13. Jiang, M., Song, D., Liao, L., Zhu, F.: A bayesian recommender model for user rating and review profiling. Tsinghua Sci. Technol. **20**(6), 634–643 (2015)
14. Kai, P., Feldt, R., Mujtaba, S., Mattsson, M.: Systematic mapping studies in software engineering. In: International Conference on Evaluation & Assessment in Software Engineering (2008)
15. Kanoje, S., Girase, S., Mukhopadhyay, D.: User profiling trends, techniques and applications. Int. J. Adv. Found. Res. Comput. **1**, 2348–4853 (2014)
16. Keele, S., et al.: Guidelines for performing systematic literature reviews in software engineering. Technical report, Technical report, Ver. 2.3 EBSE Technical Report. EBSE (2007)
17. Kitchenham, B.A., Budgen, D., Brereton, O.P.: Using mapping studies as the basis for further research: a participant-observer case study. Inf. Softw. Technol. **53**(6), 638–651 (2011)
18. Kitchenham, B.A., Pfleeger, S.L., Pickard, L.M., Jones, P.W., Hoaglin, D.C., Emam, K.E., Rosenberg, J.: Preliminary guidelines for empirical research in software engineering. IEEE Trans. Softw. Eng. **28**(8), 721–734 (2002)
19. Lancaster, G.A., Susanna, D., Williamson, P.R.: Design and analysis of pilot studies: recommendations for good practice. J. Eval. Clin. Pract. **10**(2), 307–312 (2010)
20. Levi, A., Mokryn, O., Diot, C., Taft, N.: Finding a needle in a haystack of reviews: cold start context-based hotel recommender system. In: ACM Conference on Recommender Systems (2012)
21. Li, C., Chen, G., Wang, F.: Recommender systems based on user reviews: the state of the art. User Model. User-Adap. Inter. **25**(2), 99–154 (2015)
22. Luo, Y.: The comparison of personalization recommendation for e-commerce. Phys. Proc. **25**, 475–478 (2012)
23. Markov, I., Gómez-Adorno, H., Posadas-Durán, J.P., Sidorov, G., Gelbukh, A.: Author profiling with doc2vec neural network-based document embeddings. In: Mexican International Conference on Artificial Intelligence, pp. 117–131. Springer (2016)
24. Moon, C., Jones, P., Samatova, N.F.: Learning entity type embeddings for knowledge graph completion. In: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, CIKM '17, pp. 2215–2218. ACM, New York, NY, USA (2017)
25. Natarajan, S., Moh, M.: Recommending news based on hybrid user profile, popularity, trends, and location. In: 2016 International Conference on Collaboration Technologies and Systems (CTS), pp. 204–211 (2016). Doi: https://doi.org/10.1109/CTS.2016.0050
26. Papadimitriou, C.H., Raghavan, P., Tamaki, H., Vempala, S.: Latent semantic indexing: a probabilistic analysis. J. Comput. Syst. Sci. **61**(2), 217–235 (2000)
27. Park, D.H., Kim, H.K., Choi, I.Y., Kim, J.K.: A literature review and classification of recommender systems research. Int. Proc. Econ. Dev. Res. **39**(11), 10059–10072 (2012)

28. Peng, J., Detchon, S., Choo, K.K.R., Ashman, H.: Astroturfing detection in social media: a binary n-gram-based approach. Concur. Comput. Pract. Exp. **29**(17), e4013 (2017)
29. Perozzi, B., Al-Rfou, R., Skiena, S.: Deepwalk: Online learning of social representations. In: ACM Sigkdd International Conference on Knowledge Discovery & Data Mining (2014)
30. Ramkumar, V., Rajasekar, S., Swamynathan, S.: Scoring products from reviews through application of fuzzy techniques. Expert Syst. Appl. **37**(10), 6862–6867 (2010)
31. Sarwat, M., Levandoski, J.J., Eldawy, A., Mokbel, M.F.: Lars*: An efficient and scalable location-aware recommender system. IEEE Trans. Knowl. Data Eng. **26**(6), 1384–1399 (2014)
32. Stewart, S., Davies, J.: User profiling techniques : a critical review. In: 19th Annual BCS-IRSG Colloquium on IR Aberdeen, UK. 8th-9th April 1997 (1997)
33. Suganeshwari, G., Ibrahim, S.: A survey on collaborative filtering based recommendation system. In: Proceedings of the 3rd International Symposium on Big Data and Cloud Computing Challenges, vol. 49, 503–518 (2016). https://doi.org/10.1007/978-3-319-30348-242
34. Szvetits, M., Zdun, U.: Systematic literature review of the objectives, techniques, kinds, and architectures of models at runtime. Softw. Syst. Model. **15**(1), 31–69 (2016)
35. Wang, D., Peng, C., Zhu, W.: Structural deep network embedding. In: ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (2016)
36. Williams, G., Mahmoud, A.: Modeling user concerns in the app store: A case study on the rise and fall of yik yak. In: 2018 IEEE 26th International Requirements Engineering Conference (RE), pp. 64–75 (2018)
37. Wohlin, C.: Guidelines for snowballing in systematic literature studies and a replication in software engineering. In: International Conference on Evaluation & Assessment in Software Engineering (2014)
38. Wolny, S., Mazak, A., Carpella, C., Geist, V., Wimmer, M.: Thirteen years of sysml: a systematic mapping study. Softw. Syst. Model. (2019). https://doi.org/10.1007/s10270-019-00735-y
39. Wu, S., Guo, W., Xu, S., Huang, Y., Wang, L., Tan, T.: Coupled topic model for collaborative filtering with user-generated content. IEEE Trans. Hum. Mach. Syst. **46**(6), 908–920 (2016)
40. Wu, S.L.: Enabling searching of user ratings and reviews using user profile location, and social networks (2011). US Patent 7,895,177
41. Xue, B., Fu, C., Shaobin, Z.: A study on sentiment computing and classification of sina weibo with word2vec. In: 2014 IEEE International Congress on Big Data, pp. 358–363. IEEE (2014)
42. Yang, Z., Zhi, L., Zhi, J., Chen, Y.: A systematic literature review of requirements modeling and analysis for self-adaptive systems. In: International Working Conference on Requirements Engineering: Foundation for Software Quality (2014)
43. Zhong, Y., Yuan, N.J., Zhong, W., Zhang, F., Xie, X.: You are where you go: Inferring demographic attributes from location check-ins. In: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, pp. 295–304. ACM (2015)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Xin Dong is a M.S. student of Faculty of Information Technology, Beijing University of Technology, China. She has published two international conference papers in the area of software requirements analysis during her graduate studies. Her main research interests are in the field of knowledge graph and recommendation system.

Tong Li holds a lecturer position in the Faculty of Information Technology at the Beijing University of Technology, China. His research interests are in the areas of software engineering, conceptual modeling and data mining.He is now hosting a National Natural Science Foundation of China, a subtask of a National Key Research and Development Program of China, and a Beijing Education Science Planning Funding. He has been an author or co-author of more than 50 papers in peer-reviewed journals, conferences, or workshops. He is an expert of ISO/IEC JTC 1/ SC 27/ WG 4 and is serving as a co-editor of ISO 24392.

Rui Song is a M.S. student of Faculty of Information Technology, Beijing University of Technology,China. She received her bachelor degree (2018) in Computer Science and Technology from Beijing University of Technology, China. Her main research interests are in the field of data mining, machine learning, and big data analysis.

Zhiming Ding is a Professor of the Faculty of Information Technology, Beijing University of Technology, China. He received his Ph.D. degree from Institute of Computing Technology, Chinese Academy of Sciences (2002) respectively. His main research interests include database systems, spatial-temporal data management, machine learning and big data analysis. He has published more than 110 papers and 4 books in the field of data management, and obtained 4 invention patents and 7 software copyrights.