Information Processing and Management 54 (2018) 529–544

Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/infoproman

Social network data to alleviate cold-start in recommender system: A systematic review

Lesly Alejandra Gonzalez Camacho*, Solange Nice Alves-Souza

Departamento de Engenharia de Computação e Sistemas Digitais, Escola Politécnica da Universidade de São Paulo (EPUSP), Av. Prof. Luciano Gualberto, Travessa 3, 158 - Butantã, São Paulo - SP, 05508-010, Brazil

A R T I C L E I N F O

Keywords:
Cold start
Social network
Collaborative filtering
Recommender system
Systematic literature review

A B S T R A C T

Recommender Systems are currently highly relevant for helping users deal with the information overload they suffer from the large volume of data on the web, and automatically suggest the most appropriate items that meet users needs. However, in cases in which a user is new to Recommender System, the system cannot recommend items that are relevant to her/him because of lack of previous information about the user and/or the user-item rating history that helps to determine the users preferences. This problem is known as cold-start, which remains open because it does not have a final solution. Social networks have been employed as a good source of information to determine users preferences to mitigate the cold-start problem. This paper presents the results of a Systematic Literature Review on Collaborative Filtering-based Recommender System that uses social network data to mitigate the cold-start problem. This Systematic Literature Review compiled the papers published between 2011–2017, to select the most recent studies in the area. Each selected paper was evaluated and classified according to the depth which social networks used to mitigate the cold-start problem. The final results show that there are several publications that use the information of the social networks within the Recommender System; however, few research papers currently use this data to mitigate the cold-start problem.

1. Introduction

In the last few years, the number of services and users have rapidly increased, partly due to the popularity of cloud computing services. By offering a flexible computational architecture, they allow having more and more industrial modern services (Deng, Huang, & Xu, 2014). The prevalence of social network services, mobile devices and large-scale service-oriented systems has produced a vast volume of data. This quantity of data overloads the users with information and hinders finding services that can accomplish their functional requirements (Almohsen & Al-Jobori, 2015; Deng et al., 2014). In this context, Recommender Systems (RSs) have an increasingly important role in helping users to overcome data or service overloading situations, suggesting the most suitable data/service automatically. This happens, for example, in e-commerce services, in which the main idea is to recommend not already acquired items which may be considered relevant to users (Deng et al., 2014; Maniktala, Sachdev, Bansal, & Susan, 2016). Examples in which RSs have been used include movies (Christou, Amolochitis, & Tan, 2016; Moreno, Segrera, López, Muñoz, & Sánchez, 2016), venues of interest (Khalid, Khan, Khan, & Zomaya, 2014; Yin, Cui, Sun, Hu, & Chen, 2014), television programs (Zhang, Chen, & Yin, 2013), online news (Lin, Xie, Guan, Li, & Li, 2014), virtual study groups (Salehi, Nakhai Kamalabadi, &

* Corresponding author at: Departamento de Engenharia de Computação, EPUSP, Av. Prof. Luciano Gualberto, Travessa 3, 158 - Butantã, São Paulo - SP, 05508-010, Brazil.

E-mail addresses: alejandragc@usp.br (L.A. Gonzalez Camacho), ssouza@usp.br (S.N. Alves-Souza).

https://doi.org/10.1016/j.ipm.2018.03.004
Received 5 September 2017; Received in revised form 23 January 2018; Accepted 9 March 2018
Available online 26 March 2018
0306-4573/

Ghaznavi Ghoushchi, 2013) and others. RSs use diverse information collected from the user to identify her/his preferences, as well as demographic information such as gender, age, location, etc. to propose new items that may be useful to her/him. For that, the Recommender System (RS) processes a large volume of data to make more adequate suggestions to users.

Despite the advances in RSs, some problems such as *data sparsity* and *cold-start* still remain open, because there is not a solution that meets the different needs (Al-Hassan, Lu, & Lu, 2015; Bobadilla, Ortega, Hernando, & Bernal, 2012; Huang, Chen, & Chen, 2016). Data sparsity is characterized by a low number of ratings for available items, making it difficult to find a relationship between users and items (Khalid et al., 2014). Cold-start is caused by the lack of both user data and item rating history, which are used as a mechanism to infer users preferences and perform the recommendation (Khalid et al., 2014; Sun, Wang, Cheng, & Fu, 2015). In this paper, we are interested in evaluating only the contributions related to the cold-start problem.

There is an active line of research to solve the difficulties associated with the above problems, and a variety of techniques have been proposed, such as the use of machine learning methods, approximation theory and various heuristics applied to different areas of recommendation (Al-Hassan et al., 2015; Alhamid, Rawashdeh, Dong, Hossain, & Saddik, 2016; Bobadilla et al., 2012; De Campos, Fernández-Luna, Huete, & Rueda-Morales, 2010; Katakis, Tsapatsoulis, Mendez, Triga, & Djouvas, 2014; Khalid et al., 2014; Xu, Fu, & Gu, 2016). In each of these approaches, different sources are explored to increase information about users - items to reduce data sparsity and to mitigate the cold-start problem (Barjasteh, Forsati, Masrour, Esfahanian, & Radha, 2015; Sun et al., 2015).

Social networks are sources of information that provide valuable data to establish the users preferences regarding items, besides identifying relationships of trust between users and the influence of one user on others. All such information can be extremely useful to make recommendations to users more accurately and objectively (Deng et al., 2014; Derczynski et al., 2015; Jiang et al., 2015), helping to alleviate the cold-start problem.

The Systematic Literature Review(SLR) presented herein was motivated by the results from previous works (Prando, 2016; Prando, Contratres, Alves-Souza, & deSouza, 2017), which showed that social networks data is a good information source for RSs to mitigate the cold-start issue.

Some filtering techniques used in the recommendation process in RSs are Collaborative filtering-based(CF), Content-based(CB) (De Campos et al., 2010; Huang et al., 2016), demographic and hybrid. CF-RSs identify groups of people with similar preferences to target user and recommend items these people like. Conversely, CB-RSs determine users' preferences and recommend similar items users previously preferred or browsed. Demographic Filter is based on the principle that users with certain common attributes, such as gender, age, education level, among others, also have common preferences. Hybrid Filter combines two or more filtering approaches to process recommendations (Bobadilla, Ortega, Hernando, & Gutiérrez, 2013; Jain, Grover, Thakur, & Choudhary, 2015; Kaššák, Kompan, & Bieliková, 2016).

While CB-RSs consider the behavior of previous users for recommendation (Huang et al., 2016), CF-RSs explore the assessments of similar users to the target user, increasing the amount of information for the recommendation process. Thus, the CF technique has been the most commonly approach used in RSs.

This paper introduces a SLR of researches, published in the last seven years, about CF-RSs that use data from social networks as main sources of information to alleviate cold-start problems.

This paper is organized as follows. Section 2 presents the main aspects of previous works in the literature review regarding correlate themes. Section 3 details the process followed to carry out the SLR here presented. Section 4 shows the SLR results and analyzes them. Section 5 introduces the main aspects of the papers classified as relevant to the SLR. Finally, the conclusion is shown in Section 6.

2. Related works

Themes on RSs that use social networks as an information source for the recommendation process are quite current, given the large number of papers in recent years (Alhamid et al., 2016; Christou et al., 2016; Katakis et al., 2014; Lim & Finkelstein, 2012; Maniktala et al., 2016; Xu et al., 2016). Moreover, literature review papers have been presented to evidence the interest in these subjects.

Rastogi and Singh (2016) presented an overview of RSs that use social network data to improve the quality of recommendations. This paper proposed categorizing socio-contextual information of social networks into explicit and implicit user-item information and analyzes the main aspects of generic RSs. However, the research was not a SLR because it did not show the steps followed for compiling articles, such as keyword definition, inclusion and exclusion criteria, and databases searched. In addition to that, it was not the focus of the research to identify articles that address the cold-start problem.

Aznoli and Navimipour (2017) made an SLR about the mechanisms of RSs used in cloud computing. In this research, RS filtering techniques were classified into four main categories: collaborative, demographic-based, knowledge-based and hybrid filtering. It also presented a comparison of these techniques in terms of scalability, availability, accuracy, and trust attributes. However, this research neither considered the cold-start problem nor the use of social network information in RSs.

Rahayu et al. (2017) presented a systematic review on RSs for the e-Portfolio domain, classifying articles according to the type of recommendation (Personalization or Business) and techniques employed. However, this review did not specifically research articles that mitigated the cold-start problem or made use of social networks. The authors proposed associate weights with questions to evaluate the selected articles. This strategy was also adopted by our SLR.

The main contribution of the SLR here presented is the collection of papers that address the use of social network data to mitigate the cold-start problem in RSs. Thus, this SLR is specific and, to the best of our knowledge, it differs from other reviews and briefly allows to show the trends of RSs in the adoption of social information as an important strategy to improve the quality of the recommendations.

Table 1
Research questions for SLR.

Research Questions	
Q1	Is there any improvement for the cold-start problem with the use of CF techniques?
Q2	Is there improvement in the recommendation when using data extracted from social networks, considering the data extraction from a friends network?
Q3	Does using data from friends on social networks as the only or main source of data improve the recommendation in situations of cold-start?

3. Procedure for systematic literature review

In the field of knowledge and especially in Computer Engineering, a SLR has great importance to increase the quality of the research. SLR consists of creating an analytical methodology to identify, select, evaluate and synthesize the main scientific research that allows elaborating a more objective bibliographic review. According to (Kitchenham & Charters, 2007), SLR generates a particular value to the state of the art because it not only establishes the protocols of compilation of information, but also details them to allow its later reproduction. The SLR presented here was based on recent approaches to SLR in computing (Kitchenham & Charters, 2007).

3.1. Research questions

Questions about the use of social networks as the only or main source to mitigate the cold-start problem, besides the results of previous works Prando (2016); Prando et al. (2017), were the motivators of this SLR.

The main research question we tried to answer was "Does using data from friends on social networks as the only or main source of data improve recommendation in cold-start situations?". This question corresponds to Q3 in Table 1. On the other hand, using social network information from friends network implies using the collaborative filtering technique. Hence, this was the technique chosen for searching papers and two other questions (Q1 and Q2 in Table 1) associated with the main question were formulated (Table 1).

Questions Q1, Q2 e Q3 conducted to look for papers that:

(i) Question Q1: apply CF techniques to mitigate the cold-start problem in RSs.
(ii) Question Q2: apply social network information and CF-based technique for recommendation improvement.
(iii) Question Q3: use social network information to mitigate the cold-start problem in CF-based RSs.

In order to start the SLR, the issues presented in Table 1 were defined in the first place, originating the following keywords: (i) recommender system or recommendation system, (ii) collaborative filtering, (iii) cold-start and (iv) social network. Thus, search strings appropriate for each database were elaborated and an example of such strings is presented in Table 2.

3.2. SLR Flow

Fig. 1 shows the flow diagram, which summarizes the different steps followed to conduct the SLR, which are:

Stage 1 Definition of the research questions and keywords: at this stage the research questions (Table 1) and the keywords (Table 2) were defined.
Stage 2 Elaboration of search strings and choice of the search databases: at this step the search strings were defined according to the format of the search databases selected. Scopus, IEEE, ACM and Web of Science were selected due to their credibility, adequacy to the computing area, besides being paid by the university, allowing full access to the content of their articles.
Stage 3 Elaboration of inclusion and exclusion criteria: criteria were developed single out papers that would be selected for complete reading. These criteria (Table 3) are discussed as follows.
Stage 4 Search of papers: at this step, we searched the papers using the search strings (Table 2) elaborated based on the research questions and the keywords.
Stage 5 Pre-selection: only the title and abstract of the whole papers retrieved in stage 4 were read at this step. Following the inclusion and exclusion criteria, the papers were either selected for full reading or discarded.
Stage 6 Complete reading of selected papers: the papers selected in step 5 were thoroughly read and evaluated for their relevance within the scope of the research. Criteria were defined to evaluate the relevance of papers (Rahayu et al., 2017). For this

Table 2
Search strings for SLR.

Search strings	
S1	("*Recommender System*" OR "*Recommendation system*") AND "*Collaborative filtering*" AND "*Cold-start*"
S2	("*Recommender System*" OR "*Recommendation system*") AND "*Collaborative filtering*" AND ("*social networking*" OR "*social network*")
S3	("*Recommender System*" OR "*Recommendation system*") AND "*Collaborative filtering*" AND "*Cold-start*" AND ("*social networking*" OR "*social network*")

Table 3
Inclusion and exclusion criteria for SLR.

Inclusion criteria	Exclusion criteria
Paper was published in a journal between 01/01/2011 and 07/12/2017 Paper was published in conference between 01/01/2013 - 07/12/2017 Paper presents a proposal to mitigate the cold-start problem in RSs. Paper incorporates social network data in CF-based RSs. Paper proposes a RS to solve cold-start using social network information and CF-based.	Paper was published in a journal with JCR less than 1. Paper is not written in English.

evaluation, metrics (Table 6) were defined based on the research questions (Table 1), which allowed the classification of the papers.

Stage 7 Classification: Finally, the papers were evaluated and ranked as highly, partially, or not relevant to the established research questions.

The inclusion-exclusion criteria (Table 3) employed in the SLR were used for the initial selection of the relevant papers. Thus, for this selection, only the title and abstract of the papers found in step 4 (Fig. 1) were read. For inclusion - exclusion criteria the following aspects were added:

- To estimate the quality and relevance of the paper within the area of Recommender Systems (RSs), only papers from publications with an impact factor equal to or greater than 1 were considered. The impact factor chosen for this evaluation was the Journal Citation Reports (JCR).
- Considering the previous aspect, to obtain the most recent advances in the area, only papers published in journals in the period from 2011 to 2017 were considered. Result is shown in Table 4 and refers the Scopus, ACM and IEEE research databases.

However, due to the low number of publications found in the end of stage 5 (Fig. 1), following these criteria and considering that the main research question would be a really current topic, the search was expanded to conference proceedings for the period of 2013 - 2017, in order to have only the most recent publications and specifically for the search string S2 and S3 (Table 2), since they are directly connected to the main research question. Considering this extension, the Web of Science database was included in the research process. At this point of the SLR, the decision was to use a different database to search conference papers. These criteria were included in Table 3.

The texts selected in stage 5 (Fig. 1) were grouped according to the research topic addressed by the paper. This grouping was performed only to separate the paper regarding the main subject treated by it, not meaning any classification of importance for the SRL. The topics considered for this division were:

- Cold-start, CF-RSs
- Social Networking, CF-RSs
- Cold-start, Social Networking, CF-RS
- Cold-start, trust Social Networking, CF-RS

The papers that approach the trust concept in social networks generally try to identify the sub-network of friends with greater connection among them. These papers advocate the idea that those within this subnet are more similar or suffer greater influence among themselves.

4. SLR Results

Table 4 shows the number of papers found by each search string S1, S2, S3 (Table 2) executed in each databases. The total number of papers per databases corresponds to the sum of all papers found with all search strings, as shown in Eq. 1:

$$TOTAL = S1 + S2 + S3 \tag{1}$$

Table 5 presents the general results of the SLR. This table shows the number of papers found in each of the databases used to search, for each of the stages of Fig. 1. The last column has the final total of more important papers for the SLR.

The Scopus database was the first database consulted, which brings together publications from a wide range of different newspapers and magazines (Springer, ACM, IEEE and Elsevier). Those papers that were repeated in other databases were discarded, which explains the greater number of papers found in the Scopus database in relation to the other databases. It was not the purpose of this SLR to qualify research databases for the total papers found that were in line with the research. Here the total of each database is given only as part of the process followed by SRL.

Fig. 2 shows the distribution per publication year (between 2011–2017) of 46 selected papers to complete reading and evaluation, respectively at stages 6 and 7 (Fig. 1). The greatest number of relevant papers, 14, appears in 2016 and a more sharply rise may be observed from 2015 to 2016. This can indicate that the interest in using data from social networks to solve the cold-start issue in recommender systems is raising. Results for 2017 cannot be decisive because the year was already in course and the search was

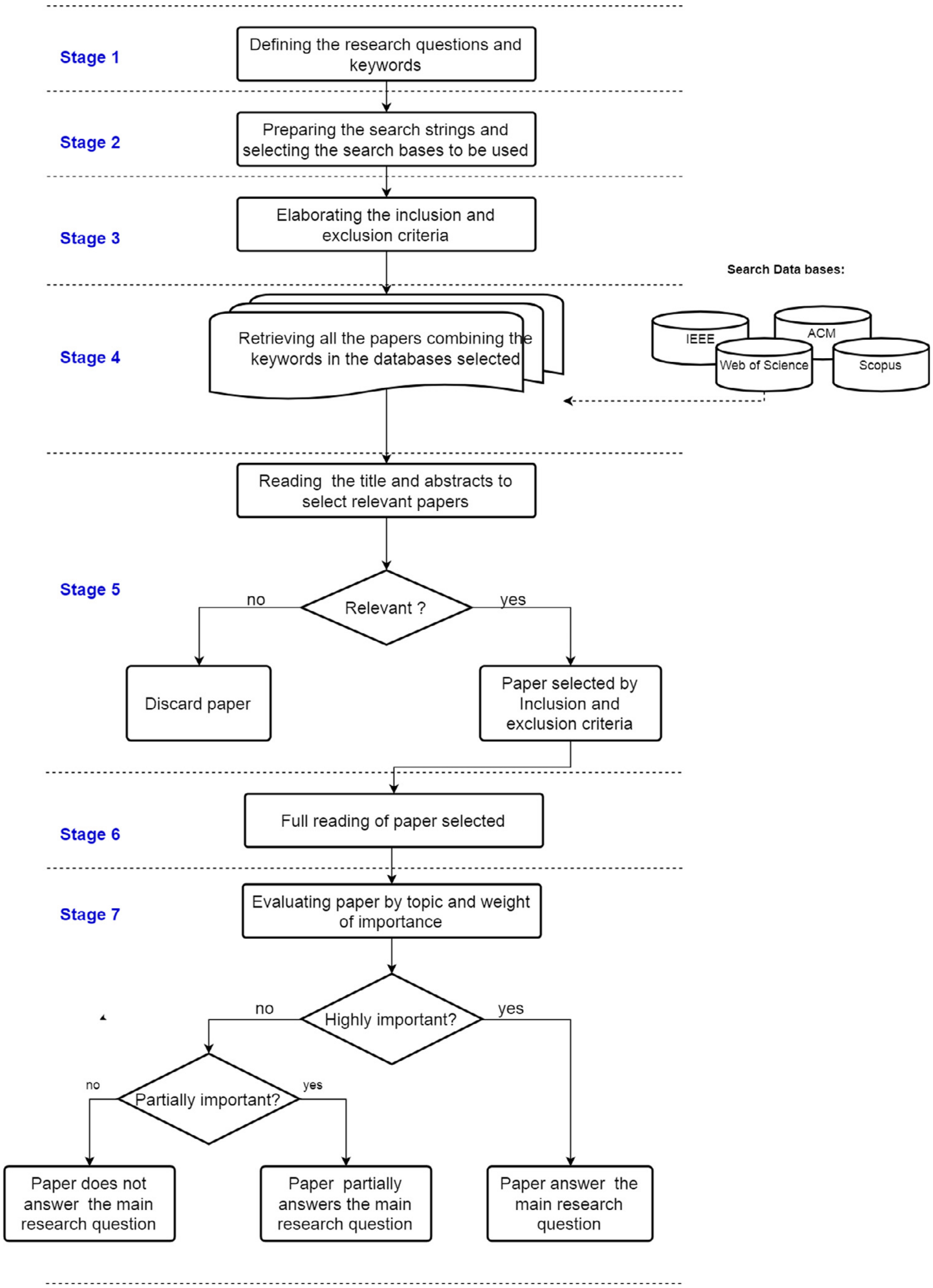

Fig. 1. SLR flow diagram.

performed until the first half of 2017 (Table 3).

The 46 papers selected for complete reading were found in 37 publication sites, which clearly show the interest and concern of the scientific community for the use of social networks to mitigate cold-start problems in RS. Table A.15 in Appendix A shows the distribution of the 46 papers selected over publication venue.

For evaluating and classifying papers in terms of their importance for SLR, three questions were elaborated (Table 6) based on the research questions (Table 1). For the score, the following values were established: (i) 2 - the article presents a solution to the Question Ai (QAi, where $i \in 1, 2, 3$) considered, (ii) 1 - the article partially solves the QAi considered, (iii) 0 - the article does not solve the QAi considered. This score allowed classifying the papers (stage 7 - Fig. 1). Each paper was evaluated for each question - QA1, QA2 and QA3 (Table 6), being the sum of paper grades its final score. Therefore, the papers were classified as:

Table 4
Number of papers found in each research databases.

String	Research Datasbases			
	Journals (2011–2017)			Journals /conferences (2013–2017)
	Scopus	IEEE	ACM	Web of Sciences
S1	170	51	11	No apply
S2	174	93	13	47
S3	43	34	2	28
TOTAL	387	178	26	75

Table 5
Number of papers found in stages 4 through 7 of flow of Fig. 1.

Research Databases	Papers found (stage 4)	Papers pre-selected (stage 5)	Complete reading (stage 6)	Final synthesis
Scopus	387	40	15	3
IEEE	178	23	16	2
ACM	26	2	2	0
Web of Science	75	20	13	5
Total	666	85	46	10

Table 6
Paper quality assessment questions.

Evaluation Questions to rank papers in stage 7 (Fig. 1)	Grade
QA1 - Does the research solve the cold-start problem using CF? QA2 - Does the research employ social network information to improve recommendation on CF-based RSs? QA3 - Does the research use social network information to mitigate cold-start problem in CF-based RSs?	yes - value 2, partially - value 1, no - value 0

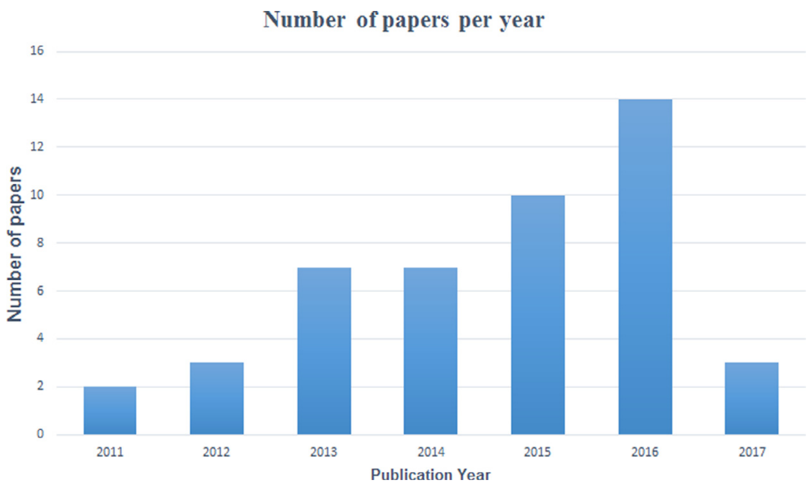

Fig. 2. Number of papers per year.

- not relevant (*score* < =2).
- partially relevant (3 < =*score* < =4).
- completely relevant (5 < =*score* < =6).

Considering the 46 articles (Table 5) selected for full reading and the score defined for their classification, Tables 7, 8 and 9 present the results found. These results were divided according to topics of importance relative to questions 1, 2 and 3, respectively (Table 6). Articles presented in Tables 7, 8 and 9 were organized chronologically (from the most recent to the oldest, using the paper year for it) and alphabetically.

Table 7 shows the 30 (65.21%) most relevant articles that proposed a solution to the cold-start problem applying different techniques in different domains. 3 articles (6.52%) partially addressed the cold-start problem. The remaining 13 articles (28.26%) did

Table 7
QA1 papers classification.

Relevance Classification	Number of papers	%	Paper Reference
Completely	30	65.21	Al-Hassan et al. (2015); Barjasteh, Forsati, Ross, Esfahanian, and Radha (2016); Bobadilla et al. (2012); Chen et al. (2013, 2016); Christou et al. (2016); Deng et al. (2014); Gogna and Majumdar (2015); Ha et al. (2013); Huang et al. (2016); Jiang et al. (2015); Kwon and Hong (2011); Lian et al. (2016); Lin et al. (2014); Liu et al. (2013); Liu, Chen, Xiong, Ding, and Chen (2012); Mirbakhsh and Ling (2015); Mohammadi and Andalib (2017); Moreno et al. (2016); Rosli et al. (2015); Salehi et al. (2013); Senthilkumar and Ponnusamy (2016); Vizine Pereira and Hruschka (2015); Wang et al. (2016); Wu et al. (2013); Xiushan and Dongfeng (2017); Zhang et al. (2014, 2017); Zhang et al. (2013); Zhao et al. (2016)
Partially	3	6.52	Hannech et al. (2016); Lalwani et al. (2015); Thilagam (2016)
Not	13	28.26	Alhamid et al. (2016); Katakis et al. (2014); Khalid et al. (2014); Lim and Finkelstein (2012); Ma, Zhou, Lyu, and King (2011); Maniktala et al. (2016); Qian, Zhang, Zhang, and Duan (2013); Sun et al. (2015); Wenjuan and Zhubing (2015); Xu et al. (2016); Yin et al. (2014); Yu, Xu, Yang, and Guo (2016); Zhao, Wang, Yu, and Gao (2014)

Table 8
QA2 papers classification.

Relevance Classification	Number of papers	%	Paper Reference
Completely	19	41.30	Alhamid et al. (2016); Ha et al. (2013); Jiang et al. (2015); Khalid et al. (2014); Lalwani et al. (2015); Lian et al. (2016); Liu et al. (2013); Maniktala et al. (2016); Rosli et al. (2015); Sun et al. (2015); Wang et al. (2016); Xiushan and Dongfeng (2017); Xu et al. (2016); Yin et al. (2014); Yu et al. (2016); Zhang et al. (2017); Zhang et al. (2013); Zhao et al. (2014, 2016)
Partially	15	32.60	Chen et al. (2013, 2016); Deng et al. (2014); Hannech et al. (2016); Huang et al. (2016); Katakis et al. (2014); Lim and Finkelstein (2012); Lin et al. (2014); Ma et al. (2011); Mohammadi and Andalib (2017); Senthilkumar and Ponnusamy (2016); Thilagam (2016); Wenjuan and Zhubing (2015); Wu et al. (2013); Zhang et al. (2014)
Not	12	26.08	Al-Hassan et al. (2015); Barjasteh et al. (2016); Bobadilla et al. (2012); Christou et al. (2016); Gogna and Majumdar (2015); Kwon and Hong (2011); Liu et al. (2012); Mirbakhsh and Ling (2015); Moreno et al. (2016); Qian et al. (2013); Salehi et al. (2013); Vizine Pereira and Hruschka (2015)

Table 9
QA3 papers classification.

Relevance Classification	Number of papers	%	Paper Reference
Completely	10	21.73	Ha et al. (2013); Jiang et al. (2015); Lian et al. (2016); Liu et al. (2013); Rosli et al. (2015); Wang et al. (2016); Xiushan and Dongfeng (2017); Zhang et al. (2017); Zhang et al. (2013); Zhao et al. (2016)
Partially	10	21.73	Chen et al. (2013, 2016); Deng et al. (2014); Hannech et al. (2016); Lalwani et al. (2015); Lin et al. (2014); Mohammadi and Andalib (2017); Senthilkumar and Ponnusamy (2016); Wu et al. (2013); Zhang et al. (2014)
Not	26	56.52	Al-Hassan et al. (2015); Alhamid et al. (2016); Barjasteh et al. (2016); Bobadilla et al. (2012); Christou et al. (2016); Gogna and Majumdar (2015); Huang et al. (2016); Katakis et al. (2014); Khalid et al. (2014); Kwon and Hong (2011); Lim and Finkelstein (2012); Liu et al. (2012); Ma et al. (2011); Maniktala et al. (2016); Mirbakhsh and Ling (2015); Moreno et al. (2016); Qian et al. (2013); Salehi et al. (2013); Sun et al. (2015); Thilagam (2016); Vizine Pereira and Hruschka (2015); Wenjuan and Zhubing (2015); Xu et al. (2016); Yin et al. (2014); Yu et al. (2016); Zhao et al. (2014)

not consider the cold-start problem and were therefore classified as not relevant to Q1.

Table 8 contains the 19 articles (41.30%) that used information extracted from social networks in CF-RSs to improve the recommendation. 15 articles (32.60%) were classified as partially relevant because they used databases that simulated a trust social relationship among users, but the data were not really extracted from user's social networks. Finally, 12 articles (26.08%) were classified as non-relevant because they did not meet Q2, since they did not use social network data to improve the recommendation.

Table 9 shows the 10 articles (21.73%) that proposed a solution to the cold-start problem using data from social networks, in which data from the friends network were actually used for the recommendation. 10 more articles (21.73%) were classified as partially relevant, and 26 (56.52%) were not relevant. These results show that, although social network information is being used in RSs, there are still few studies that use it to mitigate cold-start problems. The main aspects of these 20 papers are discussed in the next section. It should be noted that these 20 papers are included within those classified as completely or partially relevant in Tables 7 and

Table 10
Techniques employed in recommendation process for the selected papers.

Technique	Number of Papers	%	References
MF and extensions	10	50	Chen et al. (2016); Deng et al. (2014); Jiang et al. (2015); Lian et al. (2016); Liu et al. (2013); Senthilkumar and Ponnusamy (2016); Wang et al. (2016); Xiushan and Dongfeng (2017); Zhang et al. (2017, 2013)
hybrid	4	20	Ha et al. (2013); Lalwani et al. (2015); Lin et al. (2014); Zhao et al. (2016)
k-means and extensions	4	20	Chen et al. (2013); Rosli et al. (2015); Wu et al. (2013); Zhang et al. (2014)
Fuzzy Geographic clustering	1	5	Mohammadi and Andalib (2017)
Social network ontological representation	1	5	Hannech et al. (2016)

8. Therefore, the remaining papers in Tables 7 and 8 different from the 20 described were classified as not relevant by QA3, which is directly connected to the main research question Q3 (Table 1), and are therefore, not detailed in this SLR.

5. Key aspects of main papers

Tables 10, 11 and 12 summarize the techniques used by the 20 papers classified as most important (Tables 13 and 14) for the main research question Q3 (Table 1) employed in the SLR.

The remaining columns of Table 10 represent the number of papers, in which the technique is used, the percentage correspondent and the references that use the technique. The majority of the papers either used techniques that already exist in the literature (ex. k-means) or proposed extensions to them or a combination of algorithms to obtain a new hybrid algorithm (ex. singular value decomposition and deep learning). In a more detailed view, 10 papers employed the matrix factorization (MF) technique as a singular value decomposition (SVD) or probabilistic matrix factorization (PMF) or new proposals based on these techniques. 4 papers utilized hybrid technique and 4 used k-means or extensions of it. According to this result, techniques based on the matrix factorization model seems to be the dominant approach employed.

The matrix factorization (MF) model transforms the characteristics of users and items into latent factor space and predicts the rating of users concerning items by computing the similarity between the user's interest and the target item (Desrosiers & Karypis, 2011; Zhang et al., 2013).

k-Means clustering is a partitioning method that partitions the data set of items into several subsets, in which the items forming the subset are as close to each other as possible, according to a given distance measure (Amatriain, Jaimes, Oliver, & M. Pujo, 2011).

Similarity measures (Table 11) are used to calculate the similarities between two users or two items, finding the most similar users or items (Sun et al., 2015). In RS, users preferences for items, or preferences of all users for an item, are represented as a vector to calculate the similarity between users, or between items, with the two-dimensional user-rating matrix. The selected studies applied a

Table 11
Similarity measures employed in the main papers.

Reference	similarity measures			
	Cosine similarity	Pearson correlation	jaccard similarity	Trust similarity
Mohammadi and Andalib (2017)		✓		✓
Xiushan and Dongfeng (2017)				✓
Zhang et al. (2017)				✓
Chen et al. (2016)	✓			✓
Hannech et al. (2016)			✓	✓
Lian et al. (2016)	✓			
Senthilkumar and Ponnusamy (2016)				✓
Wang et al. (2016)				✓
Zhao et al. (2016)		✓		
Jiang et al. (2015)	✓		✓	✓
Lalwani et al. (2015)	✓			
Rosli et al. (2015)	✓	✓	✓	✓
Deng et al. (2014)	✓			✓
Lin et al. (2014)	✓			
Zhang et al. (2014)	✓	✓		
Chen et al. (2013)				✓
Ha et al. (2013)		✓		
Liu et al. (2013)		✓		
Wu et al. (2013)		✓		
Zhang et al. (2013)	✓			✓
Number of Papers	9	7	3	12

Table 12
Evaluation measures used in the main papers.

Reference	Evaluation measure						
	MAE	RMSE	Precision	Coverage	Recall	F1-measure	Other
Mohammadi and Andalib (2017)	✓						
Xiushan and Dongfeng (2017)	✓	✓					
Zhang et al. (2017)	✓	✓					
Chen et al. (2016)			✓	✓		✓	
Lian et al. (2016)			✓		✓		
Hannech et al. (2016)							✓
Senthilkumar and Ponnusamy (2016)	✓	✓					
Wang et al. (2016)			✓		✓		✓
Zhao et al. (2016)			✓		✓		
Jiang et al. (2015)	✓	✓	✓		✓		
Lalwani et al. (2015)				✓			✓
Rosli et al. (2015)	✓			✓			
Deng et al. (2014)		✓	✓	✓		✓	
Lin et al. (2014)						✓	
Zhang et al. (2014)	✓	✓					
Chen et al. (2013)	✓		✓	✓	✓	✓	
Wu et al. (2013)							✓
Ha et al. (2013)	✓						
Liu et al. (2013)	✓	✓					
Zhang et al. (2013)		✓		✓			✓
Number of Papers	10	8	7	6	5	4	5

Table 13
Features of papers that use social network in CF-RSs to alleviate cold-start problem.

Reference	Social Network	Environment application	Technique
Xiushan and Dongfeng (2017)	Tencent Weibo	Variety of items	Trust and Behavior based Singular Value Decomposition (TBSVD) (Trust Matrix factorization)
Zhang et al. (2017)	Yelp	Venue, Movies	kernel-based attribute-aware matrix factorization
Lian et al. (2016)	Jiepang, Sina Wibo	Location	Implicit-feedback based Content-aware Collaborative Filtering (ICCF)
Wang et al. (2016)	Douban	Movies, books, Music.	Extend-Bayesian Personalized Ranking (BPR)
Zhao et al. (2016)	Sina Weibo	Jingdong E-commerce	Modified Gradient Boosting tree algorithm, Feature-based matrix factorization
Jiang et al. (2015)	Tencent Weibo	Web-post	Hybrid Random Walk (HRW)
Rosli et al. (2015)	Facebook	Movies	K-mean clustering
Ha et al. (2013)	Facebook	TV advertisement	Frequent Pattern Network (FPN)
Liu et al. (2013)	Douban, Last.FM	Movies, books, Music	Bayesian probabilistic matrix factorization.
Zhang et al. (2013)	Sina Weibo	TV Programs	Probabilistic matrix factorization, Latent Dirichlet Allocation (LDA)

wide range of similarity measures. The most common among such techniques include Cosine similarity, Pearson correlation and Jaccard coefficient (Amatriain et al., 2011; Rehman, Khalid, & Madani, 2017). Cosine similarity is the most used measurement method and represents the characteristics of items or preferences of users. This is performed as vectors of a n-dimensional space, computing their similarity as the cosine of the angle they form. Pearson correlation calculates the linear relationship between two variables and it is calculated based on the covariance between these variables and their standard deviation. In RS, Pearson correlation is used to find correlations between items or preferences of users. Jaccard calculated the similarity taking the intersection between the occurrence of two variables (ex. user's ratings for an item) divided by their union (Amatriain et al., 2011).

Considering the 20 papers selected (Tables 13 and 14), 12 papers applied extensions of similarity measure to weight the trust degree between users. Some of them also proposed an empirical trust similarity measure that associates the social network data and the degree of friendship among the users, in order to improve the recommender algorithm. These researches believes that among user's friends, there is a set for which the friendship is strongest and the similarities and/or influence among this set of friends is greater than with the other ones. Moreover, for the trust approach, some papers focused on discovering the most influential friend, since they consider that his/her ratings have a greatest weight in the recommendation process.

In order to measure the prediction accuracy of RSs, many measures have been proposed in the literature. For this analysis, Mean Average Error (MAE), Root Mean Squared Error (RMSE), Coverage, Precision, F1-measure and Recall were the most popular measures encountered to evaluate the accuracy of a recommendation method (Table 12). Besides, different measures were found to evaluate the accuracy of the system. However, they are less popular among researchers and, therefore, were not detailed here.

MAE and RMSE compute the error between an item real rating and the predicted rating made by a recommendation method through the average distances between them. The smaller the value is, the better the method performance (Chen, Wan, Chung, & Sun, 2013; Jiang et al., 2015).

Table 14
Features of papers that use partially social network in CF-RSs to alleviate cold-start problem.

Reference	Dataset	Environment application	Technique
Mohammadi and Andalib (2017)	Epinions	Variety of items	Fuzzy geographic clustering
Chen et al. (2016)	Epinions	Variety of items	SVD Matrix Factorization, modified Random Walk
Hannech et al. (2016)	Academic research dataset	Academic research	Social network ontological representation
Senthilkumar and Ponnusamy (2016)	Epinions	Variety of items	Matrix factorization
Lalwani et al. (2015)	Facebook, MovieLents	Movies	Louvains Community Detection (CD), Map reduce framework
Deng et al. (2014)	Epinions	Variety of items	Modified Random Walk
Lin et al. (2014)	Popular news service web sites	Online News	Probabilistic matrix factorization
Zhang et al. (2014)	Movielens and Netflix	Movies	K-means clustering
Chen et al. (2013)	Epinions	Variety of items	K-means clustering, PageRank algorithm
Wu et al. (2013)	Academic Web sites, Movielens	Movies and academic papers	K-means div-clustering

Coverage is a measurement that assesses the importance of recommendation through the number of the set of available items covered in the recommendation and the number of recommendations that can be generated to all potential users. If the coverage is low, it indicates that the prediction is less valuable to the users due to the limited data available to make a prediction (Rosli, You, Ha, Chung, & Jo, 2015).

Precision measures the recommended items that were effectively relevant to a user. Recall measure is the portion of relevant items recommended. Lastly, F1 measure refers to the harmonic average of precision and recall and it is used to assess the overall effectiveness of a RS (Chen et al., 2013).

Tables 13 and 14 summarize other interesting characteristics of the 20 papers. Table 13 presents the features of the 10 papers that used the social networks information to mitigate the cold-start problem in RSs. Table 14 shows the features of the 10 papers that partially fulfilled the research questions of the SLR.

5.1. Researches that use social network information to mitigate cold-start problem in CF-based RSs

The general idea of RSs is to model user-item interactions as factors representative of latent characteristics of users and items in the system, such as preference types and item types. This model is trained using rating data to predict user ratings for some new items (Ricci, Rokach, & Shapira, 2011). Unlike the traditional RSs, which assume that all users are independent, the surveys presented here explore the different social interactions or connections about users and incorporate these user-item and user-user interactions into the RSs to make the recommendation (Lalwani, Somayajulu, & Krishna, 2015). These surveys make direct use of information from the user's social networks as complementary information to fill the lack of data on users preferences. Similarly, the results of the research (Ha, Oh, & Jo, 2013; Jiang et al., 2015; Lian et al., 2016; Rosli et al., 2015; Zhang et al., 2013; Zhao et al., 2016) showed that RSs that use information from social networks have the ability to reduce the cold-start problem and, therefore, significantly improve the recommendation (Wang, Lu, Ester, Wang, & Chen, 2016).

Regardless of the domain applied, being able to define the user's relationship with his / her set of friends becomes a very important factor when making the recommendation. The context of the recommendation, or the context in which the user is located, determines the weights of relationships, or their implicit evaluations. Knowing how to translate this information into latent features is critical for recommendation. Latent features (or latent factors) consist in assigning values to evaluate relationships. For example, assuming a social relationship as strong, its evaluation could be 1 and, in case it is weak, its evaluation is 0. These strategies are assumed to achieve a better interpretation of the information that comes from social networks, and to differentiate its degree of importance within RSs.

Some researches such as Zhao et al. (2016) e Zhang et al. (2013) determine the level of users relationship based on the number of posts shared in ones social network with a particular friend, or in a group that simply shares the same information type. This type of behavior is assumed as a very strong bond of friendship and has great relevance to evaluate items to recommend to the cold-start user.

For instance, Zhang et al. (2013) proposed a recommendation model based on the Probabilistic Matrix Factoring algorithm with regularization of social information and items similarity to suggest TV programs. The assessment of the relation trust degree, which is made based on retweets between users, gives the social regularization. They also introduce the regularity of item similarity based on the similarities of items related to the micro-blog content, which emphasizes the preferences of the user. These two regularizations are inserted into the Probabilistic Matrix Factor (PMF) to calculate the top-N of the most relevant items for the user. Data from Sina Weibo, the Chinese social network, were used to test the proposal and results were compared with other algorithms. According to the results presented, the proposal was effective to make recommendations to new users.

Zhao et al. (2016) generates recommendations for an e-commerce using latent features composed of information from users' social networks and their friends network. Using recurrent neural networks, RS specifically learns the characteristics of users and products from the data collected from e-commerce sites. They employ a method based on Gradient Boosting Trees to map the user's social network information into latent feature representations, in order to relate user-user and item-user relationship as values 0–1 within the correlation matrix. The CF-RS was tested with data from the Jing-Dong e-commerce site using users transaction records (composed of user ID, product ID and purchase timestamp). In addition, data from Sina Weibo, China's social network, was used to retrieve tweets from active users. The authors employed 8 assessment methods to evaluate results to recommended products, including for cold-start users in different scenarios for performance validation.

Ha et al. (2013) makes television advertisements recommendation using CF techniques and social information to find frequent preference patterns of items through association rules that differentiate and classify the (direct or indirect) relationship of the user with his/her friends group and set of items. The FPN (Frequent Pattern Network) algorithm is proposed to alleviate the data sparsity and cold-start problems, considering minimum levels of confidence according to the weights of the relationships between users. Two models are implemented in RS, one based on CF techniques and the other on FPN. The CF model is used when there is a users preference history and then the RS suggests personalized advertisement and updates the users model using the user feedbacks. The FPN model is used when the minimum item evaluation number ($a = 15$) is not reached; therefore, the system then recommends TV advertisements based on association rules generated from FPNs. For the experiment, data from Facebook and Advertising Information Center (ADIC) of Korea were used. The Mean Absolute Error (MAE) was employed to measure the accuracy of the system and Hit Ratio (HR) to evaluate the recommendation quality of the proposed method. The results were compared with systems that applied the techniques of CF and FPN. They showed that the combined use of these techniques improved the recommendation for different users with little or no information about their preferences or with a large or small friends network.

Jiang et al. (2015) proposes the Hybrid Random Walk (HRW) method based on the random walk algorithm for transferring knowledge from social networks. The level of preference of an item is estimated based on the weight of the user's social relationship (strong and weak links) with his/her set of friends. The CF-RS groups the set of friends of the user by labels or tags, using item preferences of the group to facilitate the recommendation. The authors made an experiment with a dataset crawled in January 2011 from Tencent Weibo, a Chinese social network. According to their experimental results, the proposed HRW produced better recommendations for cold-start users, compared to other algorithms based on Random Walk such as TrustWalker and ItemRank.

Rosli et al. (2015) compares a database of movie ratings with likes and co-likes about movies, made by a user and his/her friends network on the *facebook pages*, to establish a similarity metric to organize users in the same interest group and to predict movies more accurately. To evaluate the proposal, an experiment was build involving 50 users that access the RS using their Facebook login IDs. The authors compare the proposed method and RS with others, which employ 3 algorithms based on KNN and K-Means that incorporated demographic information to improve the recommendation. According to the experimental result, the cold-start problem was successfully solved with significant increases in accuracy up to 5%.

Lian et al. (2016) proposes a location-based RS, investigating the location history and representing all the users characteristics and items into latent factor characteristics. Using weighting of user-item characteristics on matrix factorization, a common point between the two characteristics is defined to detect the preference category of items and subsequently recommend items associated with that category. The Implicit-feedback-based Content-aware Collaborative Filtering (ICCF) framework was introduced to transform the information extracted from social networks into latent features and insert them into the matrix factorization, where their product indicates an user preference score for location category. The dataset used to test the RS was created by data extracted from "Jiepang", a location-based site, Sina Weibo's micro-blog social network and tweets posted by users. According to the results, this approach is more assertive to make recommendations for cold-start users when compared to other algorithms based on content-aware collaborative filtering frameworks, such as LibFM.

Wang et al. (2016) classifies items according to the social ties of users with their friends. It adapts the coefficient of Jaccard as an intrinsic feature of the social network topology, to calculate the strength of affinity between users and, based on a threshold, to determine which could be categorized as strong or weak. These data are used within an extension of the "Bayesian Personalized Ranking (BPR) framework", to establish a customized ranking model of items that could be of interest to the user. Subsequently, the RS classifies the strong and weak ties of a social network by applying the Expectation-Maximization (EM) algorithm, based on the stochastic gradient descent algorithm. The experiments are conducted in 4 different datasets (DBLP, Ciao, Douban, Epinions). The Douban social network is a Chinese social network used to record information and to create content related to movies, books, music, recent events and activities in Chinese cities. DBLP, Ciao and Epinions have information about the rating of items and their content information.

Liu, Wu, and Liu (2013) proposes a framework that integrates social relationships and item content into a Bayesian Probabilistic Matrix Factorization (BPMF) to make recommendations. Initially, an algorithm is proposed to integrate social relationship information within BPMF and to alleviate the sparsity and cold-start problem. This algorithm differs from traditional collaborative filtering methods based on regularization and factorization, because it considers social information and item content. Besides, the user characteristics vector is different from all users and not the same as traditional BPMF. The algorithm is tested in 3 different datasets (Epinions, Douban and Last.fm) that contain information about rating of item, friendship links between users and item content information, such as Last.fm social network, which focuses on music. As a result, an accuracy improvement is obtained in terms of MAE and Root Mean Squared Error (RMSE).

Xiushan and Dongfeng (2017) proposes the Trust and Behavior based Singular Value Decomposition algorithm (TBSVD). TBSVD is a matrix factorization algorithm that combines social trust and social behavior in micro-blogs to determine users preferences. First, implicit trust is calculated based on user interaction behaviors within social networks. Basically, the number of retweets, @mentions and comments users make with other people are quantified to determine the most reliable. Explicit trust is considered as direct trust of social connections among users, such as trust relations in Epinions and following relations in micro-blogs. The combination of implicit and explicit trust information generates an array called the Extended Trust Matrix, which contains not only binary values, but also indicates different degrees of trust between users and allows making better recommendation of items. The experiments were conducted for users with historical rating of items and users cold-start (which presents a history of rating with less than 5 items). The dataset used for testing was extracted from the Tencent Weibo social network. The proposed method was compared with four other methods classed as classics (item-CF, Matrix Factorization (MF), trust-MF, trust-SVD) to evaluate the performance in the recommendation. The results showed that the proposed method (TBSVD) performs much better than traditional methods in terms of MAE and RMSE, including cold-start user cases.

Zhang, Chow, and Xu (2017) proposes a model called kernel-based attribute-aware matrix factorization (KAMF) for custom rating prediction of items. This model combines the item attribute information with item-user rating matrix to mitigate the data sparsity and cold-start problems. According to the authors, the KAMF model can considerably increase the predictive ability of user-item rating to assume possible nonlinear interactions among users, items and their attributes. This is different from other models that always assume the interaction is linear. KAMF deals intrinsically with the cold-start problem, using attributes of items and taking advantage of social ties between users. KAMF can measure the strength of the social relationship between users by the social tie between them and the number of their common friends. The experiments were conducted using Movielents dataset and the Yelp social network site. The results were compared with the competing predictive models (user-based CF, Professional Matrix Factorization (PMF), Topic Regression Matrix Factorization (TRMF) and Context-aware Factorization Machine (CFM)), and showed that KAMF improves the recommendation performance in cold-start situations.

5.2. Researches that partially use Social network in CF-based RSs to Alleviate the Cold-start Problem

Ten papers, described below, were classified as partially relevant in this SLR, because although they have an interesting approach to mitigate the cold-start problem, most of them did not use data from real social networks to corroborate the efficiency of their approaches. However, they present good proposals for treating social network information in RS.

Lin et al. (2014) proposes a new framework for recommending online news called 'Premise', which is an unified model of different techniques (content-based methods, collaborative filtering and information diffusion models) and considers the opinions of a specific group of users cataloged as "experts", to mitigate data sparsity and cold-start issues in the recommendations. The recommendation is made based on reading choices of the most influential users in a virtual social network, built from "follow" adoption relationships among the users. The strategy proposed in this research can mitigate the cold-start problems and data sparsity in terms of Precision and Recall. However, the results from this RS are questionable, since the social network used is not real. It is created from both history of the old users of the system and the type of relationship presumed between them.

Wu, Wang, Peng, and Li (2013) develops a RS based on the div-clustering method, which groups users and items according to their similarity, using the K-means algorithm to solve the cold-start issue. The recommendations are made considering the characteristics of the cluster which the user belongs to and the characteristics of the clusters associated with the items the user preferred in the past. To ensure reliability in the recommendations, the degree of user activity within the network was considered, aiming to identify the most active users. The idea is to provide collaborative recommendations appropriated to the interests of the user, with active users' help. However, the proposed RS does not use explicit information from social networks to improve the recommendation accuracy. According to the author, in spite of preserving the users privacy, it creates a profile for each user with information on their interests, activities, etc.

Chen et al. (2013) proposes a recommendation method that integrates an user model with the trust and distrust networks, to identify trusted users or experts, who can be a reference for useful recommendations to cold-start users. The recommendation process is divided into two steps: construction of the model, and the recommendation. In the first step, experienced users (non-cold-start users) are clustered using the k-means algorithm to construct a user model that brings together, in the same cluster, those with similar item preferences. Then, in the recommendation step, the evaluation of an item is calculated based on the identification of the clusters closely related to the unclassified item and the cold-start user profile. The identification of experienced users was determined by a score metric of reputation calculated by the PageRank algorithm. RS has been tested with the Epinions dataset which contains trust network information.

Zhang et al. (2014) proposes a method called Bi-clustering and Fusion (BIFU), that simultaneously groups item and user dimensions in the user-item matrix through k-means algorithm, to attack the cold-start problem. In order to reduce the dimensionality of the item-user matrix, the system initially removes empty items and user profiles that negatively contribute to the recommendation. Subsequently, it elaborates a dense area in the user-items rating matrix with the most popular items and their associated rating. It then allocates them to the upper left corner of the rating matrix and aggregates similar items and users based on bi-clustering, to reduce data dispersion. The results of the experiment showed better performance in recommending items to cold-start users. However, Movielents and Netflix were the datasets used for the tests, which do not represent data from social networks and their network of friends, but only an item rating history.

Chen, Hendry, Huang, and Chen, Rung-Ching,Hendry, Huang (2016) uses trusted information to improve recommendations. This crawled research dataset explores how neighbors can affect the concept of trust value among social network users and also proposes the metric, "GT, GlobalTrust", to model the confidence values within the RS to improve the accuracy of the recommendation. A matrix factorization based on the Singular Value Decomposition (SVD) is used to obtain a trust network constructed with the GT value. The results of recommendations are obtained through a modified random walk algorithm called the GlobalTrustWalker. The main contribution of this work is baseing the value of GT as a probabilistic value and not as a binary value (1/0). The system was tested with the Epinions dataset.

Deng et al. (2014) presents the RelevantTrustWalker (Random walk extension) algorithm, a trust-based RS to identify the user trust ties with his or her group of friends, and their influence on the user according to their similarity. In this RS, trust ties are refined, considering the trust relationships and similarity between users, since the target user and the trusted users may differ in interests, preferences and perception. The RelevantTrustWalker algorithm is used to recommend items. Unlike the classical approaches, which search the target node at random, the RelevantTrustWalker searches the node (item) based on the relevance of the trust. This RS was also tested with the Epinions dataset.

Hannech, Adda, Mcheick, and Science (2016) explores the social network to identify the user with more number of social ties and to classify as a key user (central user). For authors, ”the key user” has a larger network of friends and can represent the community of which she/he is part of. To determine the ”key user”, an ontological representation of interest domains is accomplished to structure the network of a set of communities, to find in each of them a leader who can represent the community. This will facilitate the recommendation of items to new users, according to the area of interest. For the experiment, a social network with different teams of a research department was modeled. In the modeling process, two characteristics were considered: the type of actor (Teacher, student, etc.) and the type of relationship (personal or professional). As preliminary results, the effectiveness of a node in its community was evaluated based on the centrality of the proximity, the degree of intermediation and the score of benefit. However, no positive results were reported to improve the accuracy of the recommendation for cold-start users, since the experiment was not concluded.

Senthilkumar and Ponnusamy (2016) creates a trust-based RS, where trustworthy interpersonal and personal aspects are considered to improve the quality of recommendations and to mitigate data sparsity and cold-start issues. For that, the system incorporates data classified as local trust and global trust. Local trust data references information from past user trust experiences with his closest neighbors in a restricted way. Global trust, on the other hand, collects information from trusted interactions within the entire social network, where aspects such as the degree of reputation of a user in the social network are analyzed. The experiment was performed using the Epinions database and the results revealed that the proposed approach improves the performance of RS for the cold-start case in relation to other methods.

Mohammadi and Andalib (2017) bases its RS on finding within social networks people who are considered opinion leaders, to help improving the recommendation processes and to mitigate the cold-start problem. According to the authors, ”opinion leaders” and information interpersonal communication networks directly influence individual decisions of each person. Hence, in the proposed model, the opinion leaders is identified based on the highest number of trust between users. An algorithm based on Fuzzy Geographic Clustering is used to select a new neighbor user, evaluating data such as user profiles, beliefs and social tags in the social network. The system analyzes the degree of friendship by calculating the intensity and closeness of the people. The intensity is calculated based on the impact that opinion leaders have on users. Proximity is assessed as the trust degree that an active user has on opinion leaders. The system is tested with the Epinions data set and, its results improve the accuracy of recommendation for cold start users.

Among the selected articles, even though they were not classified as completely relevant to mitigate the problem of cold-start, or in the use of the social network data, some deserve to be highlighted. For example: Huang et al. (2016) uses information from Google to address the shortage of new item ranking. The Google information is used to recalculate the similarity between two items to solve the *data sparsity* and *cold -start* issues in the movie recommendation. The authors propose a metric to classify a new item, which compares the similarity of local system items with semantic similarity, acquired from the number of results returned by the Google search engine, for a specific set of items. The result allows adjusting the similarity, initially calculated by conventional techniques (cosine similarity) with local data of items, in order to increase the probability of finding items more similar to the interests of the user.

Khalid et al. (2014) presents a new strategy for predicting interest venue, based on users preferences such as past check-in, location, and collaborative social opinions (preferences of other individuals). RS recommends to the user, or to groups of users, a broader set of options regarding places that might be of interest to them and in which they could not visited in the past. The approach of this research is interesting in the sense that it is able to incorporate information from social networks as support to determine a user profile. However, this strategy is only tested to attenuate data sparsity and does not show any results for the cold-start problem.

Thilagam (2016) uses the features of Social Network Graph (SNG) in conjunction with User-Rating Matrix (URM) to create association rules to mitigate the data sparsity and cold-start issues. The proposed RS presents three steps: 1) Algorithms are applied to identify neighbors closer to the user, based on a threshold of co-rated items. Those that exceed the threshold are used in calculating similarity of preferences. 2) The interaction of users in social network is analyzed to obtain the intensity of interaction with the set of friends. 3) The weighted sum of user interaction and its similarity determine the set of neighbors closest to the preferences of the target user. The experiment was implemented using as datasets the MovieLens and a ”synthetic” social network, i.e., a not real network. The question remains whether this approach would be valid for real data from social networks. In addition, the paper also does not present specific results for the cold-start problem.

6. Conclusion

This research identified and compiled papers using social network data to mitigate the cold-start problem. According to the SLR produced, for the period from 2011 to 2017, there was an increase in the number of papers published in newspapers or magazines of high citation ($JCR = >1$) and conferences, which propose the use of social networks and / or mitigate the cold-start problem in RSs.

The SLR revealed that although there are studies that use social network data in the RSs, there are few papers that use this information to ease the cold-start problem. Being able to measure the level of trust and the influence of friends on the user has become a subject of great interest in RSs, since users could have a greater affinity for consuming the items their closest friends consumed in the past. (Wang et al., 2016). In this context, 10 research papers were found. Through the application of different techniques, it was possible to extract and represent the social ties of the users and the semantic contents shared in the social networks, as latent characteristics in RSs, to establish evaluations of preference of items that improved the precision of the Recommendations to the user.

This SLR showed that social network information is a good source, that can be extensively explored to fill users' data shortages with no purchase history or no preference selection in RSs, in order to improve user experience with more assertive recommendations.

This work was based on CF-RSs in general, and did not differentiate between model-based and memory-based approaches, which are part of a sub-classification of CF-RSs. Likewise, the papers classified here were explicitly selected when they mentioned mitigating the cold-start user problem and not the cold-start item problem. So other SLR could be extended to these topics.

Acknowledgment

Authors are grateful for the support given by Sao Paulo Research Foundation (FAPESP). Grant #2014/04851-8

Appendix A

Table A.15
Number of the selected studies over publication venue.

Publication venue	Type	Number
Information Sciences	Journal	2
Decision Support Systems	Journal	3
Expert Systems with Applications	Journal	1
Knowledge-Based Systems	Journal	3
Neurocomputing	Journal	1
Network and Computer Applications	Journal	1
IEEE Intelligent Systems	Journal	2
IEEE Access	Journal	1
IEEE Transactions on Consumer Electronics	Proceding	1
IEEE Transactions on Cybernetics	Proceding	1
IEEE Transactions on Emerging Topics in Computing	Proceding	1
IEEE Transactions on Human-Machine Systems	Proceding	2
IEEE Transactions on Knowledge and Data Engineering	Proceding	4
IEEE Transactions on Learning Technologies	Proceding	1
IEEE Transactions on Services Computing	Proceding	1
IEEE Transactions on Software Engineering	Proceding	1
IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics	Proceding	1
Knowledge and Information Systems	Journal	1
Cluster Computing -The Journal of Networks Software Tools and Applications	Journal	1
Information Processing and Management	Journal	1
Multimedia Tools and Applications	Journal	1
Tsinghua Science and Technology	Journal	1
ACM Transaction on Knowledge Discovery from Data	Proceding	1
ACM Transactions on Information Systems	Proceding	2
ACM International Conference on Information and Knowledge Management (CIKM 2016)	Conference	1
Annual IEEE India Conference (INDICON-2015)	Conference	1
IEICE Transactions on Information and Systems	Conference	1
IEEE International Conference on Big Data, IEEE Big Data 2015	Conference	1
IEEE International Conference on Data Mining, ICDM	Conference	1
Information Technology, Electronics & Mobile Communication Conference IEEE IEMCON-2016	Conference	1
International Conference on Advanced Computing and Communication Systems (ICACCS)	Conference	1
International Conference on Computer Science & Education (ICCSE 2015)	Conference	1
International Conference on Recent Trends in Information Technology	Conference	1
International Conference on Cloud Computing and Big Data	Conference	1
International Conference on Web Research (ICWR)	Conference	1

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ipm.2018.03.004.

References

Al-Hassan, M., Lu, H., & Lu, J. (2015). A semantic enhanced hybrid recommendation approach: a case study of e-Government tourism service recommendation system. *Decision Support Systems, 72*, 97–109. http://dx.doi.org/10.1016/j.dss.2015.02.001.

Alhamid, M. F., Rawashdeh, M., Dong, H., Hossain, M. A., & Saddik, A. E. (2016). Exploring latent preferences for context-Aware personalized recommendation systems. *IEEE Transactions on Human-Machine Systems, 46*, 615–623. http://dx.doi.org/10.1109/THMS.2015.2509965.

Almohsen, K. A., & Al-Jobori, H. (2015). Recommender systems in light of big data. *International Journal of Electrical and Computer Engineering, 5*, 1553–1563.

Amatriain, X., Jaimes, A., Oliver, N., & M. Pujo, J. (2011). Data mining methods for recommender systems. In F. Ricci, L. Rokach, & B. Shapira (Eds.). *Recommender systems handbook*. US: Springer.

Aznoli, F., & Navimipour, N. J. (2017). Cloud services recommendation: Reviewing the recent advances and suggesting the future research directions. *Journal of Network and Computer Applications, 77*, 73–86.

Barjasteh, I., Forsati, R., Masrour, F., Esfahanian, A.-H., & Radha, H. (2015). Cold-Start item and user recommendation with decoupled completion and transduction. *Proceedings of the 9th ACM Conference on Recommender Systems,* 91–98.

Barjasteh, I., Forsati, R., Ross, D., Esfahanian, A.-H., & Radha, H. (2016). Cold-Start recommendation with provable guarantees: a decoupled approach. *IEEE Transactions on Knowledge and Data Engineering, 28*, 1462–1474. http://dx.doi.org/10.1109/TKDE.2016.2522422.

Bobadilla, J., Ortega, F., Hernando, A., & Bernal, J. (2012). A collaborative filtering approach to mitigate the new user cold start problem. *Knowledge-based systems, 26*, 225–238. http://dx.doi.org/10.1016/j.knosys.2011.07.021.

Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. *Knowledge-based systems, 46*, 109–132.

Chen, C. C., Wan, Y. H., Chung, M. C., & Sun, Y. C. (2013). An effective recommendation method for cold start new users using trust and distrust networks. *Information Sciences. 224*, 19–36. http://dx.doi.org/10.1016/j.ins.2012.10.037.

Chen, R. C., Hendry, Huang, C. Y., & ChenRung-Ching; Hendry; Huang, C.-Y. (2016). A domain ontology in social networks for identifying user interest for personalized recommendations. *Journal of Universal Computer Science, 22*, 319–339.

Christou, I. T., Amolochitis, E., & Tan, Z.-H. H. (2016). AMORE: Design and implementation of a commercial-strength parallel hybrid movie recommendation engine. *Knowledge and Information Systems, 47*, 671–696. http://dx.doi.org/10.1007/s10115-015-0866-z.

De Campos, L. M., Fernández-Luna, J. M., Huete, J. F., & Rueda-Morales, M. A. (2010). Combining content-based and collaborative recommendations: A hybrid approach based on bayesian networks. *International Journal of Approximate Reasoning, 51*, 785–799.

Deng, S., Huang, L., & Xu, G. (2014). Social network-based service recommendation with trust enhancement. *Expert Systems with Applications, 41*, 8075–8084. http://dx.doi.org/10.1016/j.eswa.2014.07.012.

Derczynski, L., Maynard, D., Rizzo, G., van Erp, M., Gorrell, G., Troncy, R., et al. (2015). Analysis of named entity recognition and linking for tweets. *Information Processing & Management, 51*, 32–49. http://dx.doi.org/10.1016/j.ipm.2014.10.006.

Desrosiers, C., & Karypis, G. (2011). A comprehensive survey of neighborhood-based recommendation methods. In F. Ricci, L. Rokach, & B. Shapira (Eds.). *Recommender systems handbook*. US: Springer.

Gogna, A., & Majumdar, A. (2015). A comprehensive recommender system model: Improving accuracy for both warm and cold start users. *IEEE Access, 3*, 2803–2813. http://dx.doi.org/10.1109/ACCESS.2015.2510659.

Ha, I., Oh, K. J., & Jo, G. S. (2013). Personalized advertisement system using social relationship based user modeling. *Multimedia Tools and Applications, 74*, 1–19. http://dx.doi.org/10.1007/s11042-013-1691-6.

Hannech, A., Adda, M., Mcheick, H., & Science, C. (2016). *Cold-start recommendation strategy based on social graphs 1. 2016 ieee 7th annu. inf. technol. electron. mob. commun. conf*.

Huang, T. C. K., Chen, Y. L., & Chen, M. C. (2016). A novel recommendation model with Google similarity. *Decision Support Systems, 89*, 17–27.

Jain, S., Grover, A., Thakur, P., & Choudhary, S. (2015). *Trends, problems and solutions of recommender system. International conference on computing, communication & automation*955–958.

Jiang, M., Cui, P., Chen, X., Wang, F., Zhu, W., & Yang, S. (2015). Social recommendation with cross-domain transferable knowledge. *IEEE Transactions on Knowledge and Data Engineering, 27*, 3084–3097. http://dx.doi.org/10.1109/TKDE.2015.2432811.

Kaššák, O., Kompan, M., & Bieliková, M. (2016). Personalized hybrid recommendation for group of users: Top-N multimedia recommender. *Information Processing and Management, 52*(3), 459–477. http://dx.doi.org/10.1016/j.ipm.2015.10.001.

Katakis, I., Tsapatsoulis, N., Mendez, F., Triga, V., & Djouvas, C. (2014). Social voting advice applications-definitions, challenges, datasets and evaluation. *IEEE Transactions Cybernetics. 44*, 1039–1052. http://dx.doi.org/10.1109/TCYB.2013.2279019.

Khalid, O., Khan, M. U. S., Khan, S. U., & Zomaya, A. Y. (2014). Omnisuggest: A Ubiquitous cloud based context aware recommendation system for mobile social networks. *IEEE Transactions on Services Computing,* 1. http://dx.doi.org/10.1109/TSC.2013.53.

Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering version 2.3. *Engineering, 45*(4ve), 1051. http://dx.doi.org/10.1145/1134285.1134500.

Kwon, H.-J. J., & Hong, K.-S. S. (2011). Personalized smart TV program recommender based on collaborative filtering and a novel similarity method. *IEEE Transactions on Consumer Electronics, 57*, 1416–1423. http://dx.doi.org/10.1109/TCE.2011.6018902.

Lalwani, D., Somayajulu, D. V. L. N., & Krishna, P. R. (2015). A community driven social recommendation system. *Proc. - 2015 IEEE Int. Conf. Big Data, IEEE Big Data 2015,* 821–826. http://dx.doi.org/10.1109/BigData.2015.7363828.

Lian, D., Ge, Y., Zhang, F., Yuan, N. J., Xie, X., Zhou, T., & Rui, Y. (2016). Content-aware collaborative filtering for location recommendation based on human mobility data. *IEEE International Conference on Data Mining (ICDM), 2016-Janua*, 261–270. http://dx.doi.org/10.1109/ICDM.2015.69.

Lim, S. L., & Finkelstein, A. (2012). StakeRare: Using social networks and collaborative filtering for large-scale requirements elicitation. *IEEE Transactions on Software Engineering, 38*, 707–735. http://dx.doi.org/10.1109/TSE.2011.36.

Lin, C., Xie, R., Guan, X., Li, L., & Li, T. (2014). Personalized news recommendation via implicit social experts. *Information Sciences, 254*, 1–18. http://dx.doi.org/10.1016/j.ins.2013.08.034.

Liu, J., Wu, C., & Liu, W. (2013). Bayesian probabilistic matrix factorization with social relations and item contents for recommendation. *Decision Support Systems, 55*, 838–850. http://dx.doi.org/10.1016/j.dss.2013.04.002.

Liu, Q., Chen, E., Xiong, H., Ding, C. H. Q., & Chen, J. (2012). Enhancing collaborative filtering by user interest expansion via personalized ranking. *IEEE Transactions on Systems, Man, and Cybernetics, Part B. 42*, 218–233.

Ma, H., Zhou, T. C., Lyu, M. R., & King, I. (2011). Improving recommender systems by incorporating social contextual information. *ACM Transactions on Information Systems, 29*, 1–23.

Maniktala, M., Sachdev, S., Bansal, N., & Susan, S. (2016). Finding the most informational friends in a social network based recommender system. *12th IEEE International Conference Electronics, Energy, Environment, Communication, Computer, Control (E3-C3), INDICON 2015,* 1–6. http://dx.doi.org/10.1109/INDICON.2015.7443226.

Mirbakhsh, N., & Ling, C. X. (2015). Improving Top-N Recommendation for Cold-Start Users via Cross-Domain Information. *ACM Transactions on Knowledge Discovery, 9*, 1–19. http://dx.doi.org/10.1145/2724720.

Mohammadi, S. A., & Andalib, A. (2017). *Using the opinion leaders in social networks to improve the cold start challenge in recommender systems. 2017 3th International Conference on Web Research*62–66. http://dx.doi.org/10.1109/ICWR.2017.7959306.

Moreno, M. N., Segrera, S., López, V. F., Muñoz, M. D., & Sánchez, Á. L. (2016). Web mining based framework for solving usual problems in recommender systems. A case study for movies' recommendation. *Neurocomputing, 176*, 72–80. http://dx.doi.org/10.1016/j.neucom.2014.10.097.

Prando, A. (2016). *Um Sistema de Recomendação para E-commerce Utilizando Redes Sociais em ambiente Big Data. Master´s. thesis*. Instituto de Pesquisas Tecnológicas do Estado de São Paulo IPT.

Prando, A., Contratres, F., Alves-Souza, S., & deSouza, L. (2017). *Content-based recommender system using social networks for cold-start users. Proceedings of the 9th*

*International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*181–189. http://dx.doi.org/10.5220/0006496301810189.
Qian, F., Zhang, Y. Y., Zhang, Y. Y., & Duan, Z. (2013). Community-based user domain model collaborative recommendation algorithm. *Tsinghua Science and Technology, 18*, 353–359. http://dx.doi.org/10.1109/TST.2013.6574673.
Rahayu, P., Sensuse, D. I., Purwandari, B., Budi, I., Khalid, F., & Zulkarnaim, N. (2017). *A systematic review of recommender system for e-portfolio domain. Proceedings of the 5th International Conference on Information and Education Technology*. New York, New York, USA: ACM Press21–26.
Rastogi, P., & Singh, V. (2016). Systematic evaluation of social recommendation systems : Challenges and future. *International Journal of Advanced Computer Science and Applications, 7*(4), 158–166.
Rehman, F., Khalid, O., & Madani, S. A. (2017). A comparative study of location-based recommendation systems. *The Knowledge Engineering Review, 32*, e7.
Ricci, F., Rokach, L., & Shapira, B. (2011). *Recommender systems handbook.* Springer US.
Rosli, A. N., You, T., Ha, I., Chung, K.-Y. Y., & Jo, G.-S. S. (2015). Alleviating the cold-start problem by incorporating movies facebook pages. *Cluster Comput. 18*, 187–197. http://dx.doi.org/10.1007/s10586-014-0355-2.
Salehi, M., Nakhai Kamalabadi, I., & Ghaznavi Ghoushchi, M. B. (2013). An effective recommendation framework for personal learning environments using a learner preference tree and a GA. *IEEE Transactions on Learning Technologies, 6*, 350–363.
Senthilkumar, K. T., & Ponnusamy, R. (2016). *Diffusing multi-aspects of local and global social trust for personalizing trust enhanced recommender system. ICACCS 2016 - 3rd International Conference on Advanced Computing and Communication Systems bringing to table, futur. technol. from arround globe*http://dx.doi.org/10.1109/ICACCS.2016.7586387.
Sun, J., Wang, G., Cheng, X., & Fu, Y. (2015). Mining affective text to improve social media item recommendation. *Information Processing. Management, 51*(4), 444–457. http://dx.doi.org/10.1016/j.ipm.2014.09.002.
Thilagam, P. S. (2016). *Alleviating data sparsity and cold start in recommender systems using social behaviour. 2016 Fifth International Conference Recent Trends Information Technology*http://dx.doi.org/10.1109/ICRTIT.2016.7569532.
Vizine Pereira, A. L., & Hruschka, E. R. (2015). Simultaneous co-clustering and learning to address the cold start problem in recommender systems. *Knowledge-Based Systems, 82*, 11–19.
Wang, X., Lu, W., Ester, M., Wang, C., & Chen, C. (2016). *Social recommendation with strong and weak ties. Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. ACM Press5–14. http://dx.doi.org/10.1145/2983323.2983701.
Wenjuan, W., & Zhubing, L. (2015). *A personalized recommendation strategy based on trusted social community. 2015 10th International Conference on Computer Science. Education (ICCSE)*. IEEE496–499.
Wu, H., Wang, X., Peng, Z., & Li, Q. (2013). Div-clustering: Exploring active users for social collaborative recommendation. *Journal of Network and Computer Applications, 36*, 1642–1650. http://dx.doi.org/10.1016/j.jnca.2013.02.016.
Xiushan, X., & Dongfeng, Y. (2017). *A novel matrix factorization recommendation algorithm fusing social trust and behaviors in micro-blogs. 2017 IEEE 2nd International Conference on Cloud Computing and Big Data Analysis*. IEEE283–287. http://dx.doi.org/10.1109/ICCCBDA.2017.7951925.
Xu, G., Fu, B., & Gu, Y. (2016). Point-of-Interest recommendations via a supervised random walk algorithm. *IEEE Intelligent Systems, 31*, 15–23. http://dx.doi.org/10.1109/MIS.2016.4.
Yin, H., Cui, B., Sun, Y., Hu, Z., & Chen, L. (2014). LCARS A spatial item recommender system. *ACM Transactions on Information Systems, 32*, 11:1–11:37. http://dx.doi.org/10.1145/2629461.
Yu, Z., Xu, H., Yang, Z., & Guo, B. (2016). Personalized travel package with multi-Point-of-Interest recommendation based on crowdsourced user footprints. *IEEE Transactions on Human-Machine Systems Trans. Human-Machine Systems, 46*, 151–158.
Zhang, D., Hsu, C.-H. H., Chen, M., Chen, Q., Xiong, N., & Lloret, J. (2014). Cold-start recommendation using Bi-clustering and fusion for large-scale social recommender systems. *IEEE Transactions on Emerging Topics in Computing, 2*, 239–250. http://dx.doi.org/10.1109/TETC.2013.2283233.
Zhang, J. D., Chow, C. Y., & Xu, J. (2017). Enabling kernel-based attribute-aware matrix factorization for rating prediction. *IEEE Transactions on Knowledge and Data Engineering, 29*, 798–812. http://dx.doi.org/10.1109/TKDE.2016.2641439.
Zhang, Y., Chen, W., & Yin, Z. (2013). Collaborative filtering with social regularization for TV program recommendation. *Knowledge-based systems, 54*, 310–317. http://dx.doi.org/10.1016/j.knosys.2013.09.018.
Zhao, K., Wang, X., Yu, M., & Gao, B. (2014). User recommendations in reciprocal and bipartite social networks–An online dating case study. *IEEE Intelligent Systems, 29*. http://dx.doi.org/10.1109/MIS.2013.104.
Zhao, W. X., Li, S., He, Y., Chang, E. Y., Wen, J. R., & Li, X. (2016). Connecting social media to E-Commerce: Cold-Start product recommendation using microblogging information. *IEEE Transactions on Knowledge and Data Engineering, 28*, 1147–1159. http://dx.doi.org/10.1109/TKDE.2015.2508816.