Navigating the Landscape of Recommender Systems: An Extensive Literature Review

Boralessa G.W.Y.S
Department of Information And Communication Technology
South Eastern University of Sri Lanka
Oluvil, Sri Lanka
geethyapa@gmail.com

Roshan G. Ragel
Department of Computer Engineering, Faculty of Engineering
University of Peradeniya
Peradeniya, Sri Lanka
ragelrg@gmail.com

Shafana M.S
Department of Information And Communication Technology
South Eastern University of Sri Lanka
Oluvil, Sri Lanka
zainashareef@seu.ac.lk

Ahamed Sabani M.J
Department of Information And Communication Technology
South Eastern University of Sri Lanka
Oluvil, Sri Lanka
mjasabani@seu.ac.lk

Abstract*—This electronic *Recommender Systems (RS) are algorithms that try to suggest relevant items to customers, and they have grown significantly over the past few years as web service giants like "YouTube" and "Netflix" exploded in popularity. The main recommendation techniques that will be covered in this study are content-based, hybrid, and collaborative filtering. The basics and feasible strategies for improving the relevance and competency of RS are also covered in this work, along with the limitations and constraints of the present recommendation techniques, such as the cold-start problem, stability vs. plasticity problem, sparsity concerns, etc. in Overall this research gives a thorough review of the state-of-the-art recommender system approaches being used in a variety of application fields. The most recent and relevant researches in the RS field were selected for a systematic review utilizing highly referenced literature found on Google Scholar. This study seeks to offer scholars and commercial developers a clear guide to recommender systems through a thorough analysis.

***Keywords*—Recommendation Approaches, Recommender Systems, hybrid recommenders, Review, Survey**

I. Introduction

The information explosion refers to the rapid increase in publicly available data or information and the consequences of this abundance. As available data expands, handling information becomes more complex, resulting in information overload when a system cannot process big data. Lately, recommender systems (RS) have emerged as a critical answer to this problem [1]. As per the original definition, RS is a system that collects and delivers people's suggestions in the form of inputs to appropriate recipients [2]. In recent years, the discipline has adopted a broader and more generic definition, referring to recommender systems as those systems as "software programs that employ various ways to produce and deliver ideas for objects and other entities to users" [3]. With the publication of the original research paper by Resnick et al., (1994), which focused on collaborative filtering, both in research and commercial applications, recommender systems grew in popularity swiftly [4]. Since then, the RS field has grown and progressed significantly. The growth can be proved with the vast number of publications, university courses related to the field and the ACM Recommender Systems Conference solely dedicated to RS [5]. Since we cannot go through a website's items manually one by one, an RS can assist a firm in improving its consumers' experience by suggesting new things that they wouldn't have found otherwise.

However, the extensive literature and methodologies for new researchers pose a problem: they have no idea which papers or which RS methodologies should be picked. Even well-experienced academics might face difficulties due to the rise of annual publications [6]. To fill that void, we conduct a survey in the RS field aiming for the betterment of developers and researchers keen on the field to identify promising research fields, be aware of the current trends, and motivate them to seek answers to the challenges and limitations in RS. In a larger sense, this research has two goals: to begin, a summary of the techniques employed in various recommender application domains will be provided. Second, to convey the current difficulties and problems to potential researchers/practitioners looking for new research possibilities and interested in pursuing their study in the field of RSs'. Following is the layout of the remainder of the paper. The research approach and methodology are in Section 2. The evaluation techniques outline three significant kinds of RS fields in Section 3. This section also highlights the most important concerns and challenges in the field of RS and several study directions that are expected to become the focus of RS research shortly. The final concluding observations are presented in Section 4.

II. Methodology

We used a systematic methodology to conduct our research to attain our goals. To secure the recommendation system somehow, an in-depth examination of alternative learning techniques was performed. This review was based on papers in the field of RS that were connected to the topic. Figure 1 shows the methodology that was used. This study aims to investigate the existing literature to acquire a better understanding of recent advancements and issues in the RS field. This method lays out the basic steps for identifying, interpreting, and assessing research articles, making it easier to find supporting data. As part of the search plan, rigorous professional planning and validation of search strings were carried out. The research articles are primarily from Google Scholar, are recent, and have many citations. Because Google Scholar is a comprehensive, source of research articles, it was picked. To study the issues in this title, the search results were filtered to include peer-reviewed and high-quality database journals and reputable conferences such as IEEE Xplore, Springer, Wiley, ACM, and Elsevier. In response to the research topic, the search keywords were carefully selected. The search keywords must be tweaked multiple times to get nearly all the relevant documents. As a result, numerous search strings containing various word combinations were used to locate relevant documents. "Recommendation systems," "challenges," "approaches," and

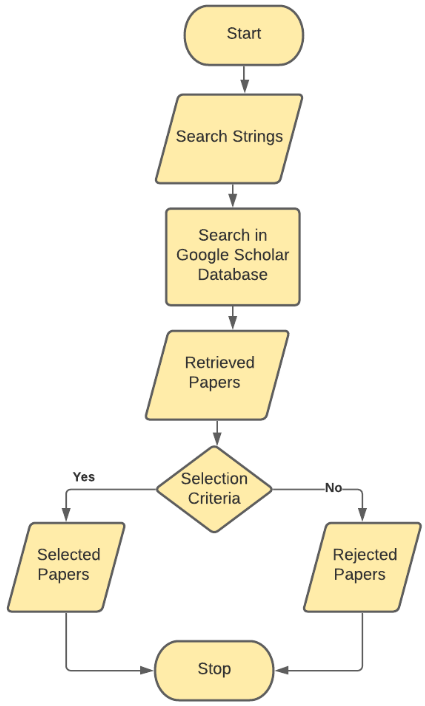

Figure 2-Paper Selection Methodology

"issues" are all terms that can be used interchangeably. The search engines of the digital library were used to conduct an automated search using these search terms. The criteria for selecting papers were then utilised to narrow down the most relevant research in this field. The remainder of the article highlights the RS approaches and challenges we discovered during the review.

III. Recommendation Approaches

When it comes to the grandeur reasons that led corporations to employ RS technology, [7], which is often known as one of the good works in the area of study, explains that most companies think about high sales and diversity of products and favourable customer satisfaction. As for the functioning of RS, there is a need for data that can be classified into two kinds.

a) *Information about the individual* - this section contains information about things (keywords, categories, etc.) And people (preferences, profiles, etc.).

b) *Interactions between the user and the item* - this includes data such as ratings, purchases, likes, and so on.

The research of Melville and Sindhwani (2017) is quite significant for its contribution in identifying the approaches, challenges and elements of recommender systems. The authors [8] identified three categories of RS approaches. Based on a survey of literature studies, the current research categorised the recommendation techniques into three based on their frequency of application which figure 2 depicts.

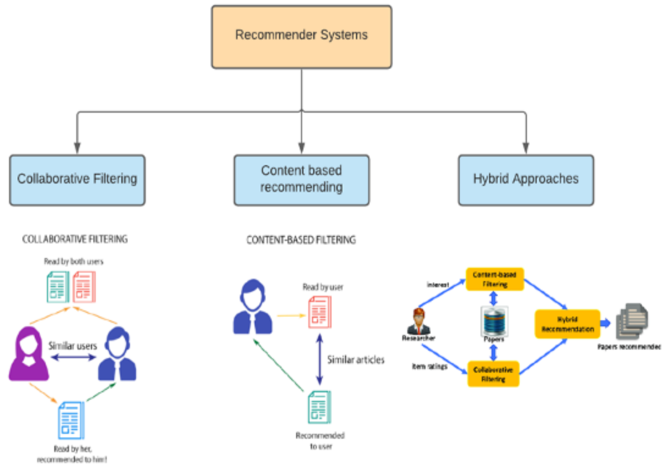

Figure 1-Types of Recommendation Approaches

A. *Collaborative Filtering (CF)*

Resnick and Varian (1997) identified and defined collaborative filtering in 1997, it has been a highly explored technique [9]. Collaborative recommender systems combine item ratings or suggestions and suggest new products after identifying commonalities according to the user's preferences. CF is widely known for situations like complicated objects where taste differences account for a lot

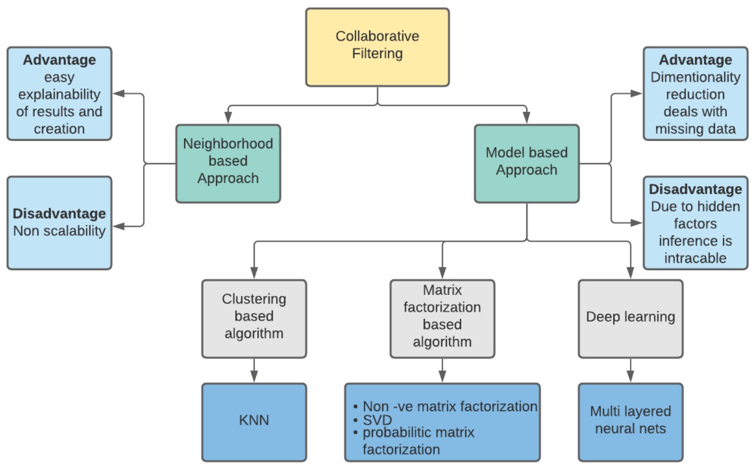

Figure 3-Types of Collaborative Filtering Techniques

of the diversity in preferences and its' accomplishments due to the reason several highly modified annual recommender algorithms [10]. Collaborative filtering is considered the most prominent recommendation technique [11]. Neighbourhood based and model-based approaches are two subsets of Collaborative Filtering methodologies. Memory-based approaches are another term for neighbourhood-based strategies. After a thorough review, we filtered the main two types of collaborative filtering techniques, and their pros and cons. Figure 3 also depicts the filtered algorithms frequently used in the model-based approach.

1) Neighbourhood Based Approach:

This approach can be further explained using two kinds of techniques. This first technique is based on user clusters. Here, one user's interactions will be considered and used as the primary source for forecasting other similar users' tastes in the system. However, the following approach takes a different route of predicting. In this way, item clusters will be considered despite user clusters. The second technique focuses on how one user preferred a set of items and using them to forecast how that same user would respond to different comparable things.

2) Model-Based Approach:

Model-based techniques generate suggestions for user ratings by calculating the parameters of statistical models [8]. The goal is to develop models that can make predictions using several techniques like data mining. This approach shows superior advancements over the neighbourhood approach. Significantly, the ability to cater to high amounts of items to a large audience makes this approach far more efficient and unique.

B. Content-Based (CB) Recommending

Systems that use CB recommends recommendations to a particular user using sensitive data like user profile and users' preferred items. The whole concept is based on the foundation that "if he like this, he will surely like this in the near future!". The fact that these CB based systems differ from other approaches is because of the methods it follows to create user profiles. It would either contact the users differently to know about their likes & dislikes or track the user's previous choices.

However, these systems sometimes could be very challenging, especially regarding the new user problem and suggestions of obvious recommendations because they wouldn't be entitled to a defined profile automatically. This can prevail if the users specially request a new profile. Nonetheless, adding new things to the system is simple. We must allocate them to a group based on their characteristics.

C. Hybrid Approach

Several hybrid techniques that combine the characteristics of content-based and collaborative recommenders have been presented to exploit both capabilities [8]. Single recommendation systems are integrated as sub-components in hybrid recommendation systems. This hybrid methodology was developed to address a flaw in traditional recommendation systems. Researchers in this discipline has focused on two primary issues: the cold-start difficulty and the stability vs plasticity conundrum. Usually, in systems based on CF and CB, these problems frequently occur because they use user information to provide recommendations. The cold-start problem's occurrence is its difficulty of manual data collection. In contrast, the stability vs plasticity problem occurs when users try to change their profile which was maintained for a long time, collecting sensitive information. The former problem can be tackled with a hybrid method. Temporal discount is one solution for the latter situation, which reduces the significance of previous evaluations. These problems are the direct reasons that led developers to pay attention to hybrid recommendation approaches. Other than CF and CB, several other evidence-based and demographic practices are now being used to make these hybrid recommenders much more efficient. By combining these different strategies, hybrid recommendation systems can give outcomes that surpass single component systems. Combining multiple approaches, such as CB and CF, is the most popular hybridising methodology. A paper [12] that significantly impacted the hybrid field established taxonomy by introducing "Weighted, Mixed, Meta-level Feature combination, Feature augmentation, Switching and, Cascade" as seven categories. Hybrid approaches have proven to be the most successful so far, even though these tactics only consider direct similarities between users or products [13].

D. Challenges In the RS field

1) Cold Start Problem:

As the name suggests, when fresh users and fresh items are placed in the system, it would be practically impossible to provide recommendations because it doesn't have any ratings or reviews, making it difficult to predict user preferences, resulting in less accuracy. This can be explained with an example of an online book store. A newly published book cannot be recommended by the system for a certain user until it starts receiving ratings or reviews. In this type of situation, finding users with commonalities is very difficult.

2) Sparsity:

Sparsity issue often occurs when users of a system avoid rating or writing reviews for their purchased items through the system due to laziness or other factors. This situation breaks down the concept of recommending as it would cause serious side effects like data sparsity issues. This is the most serious flaw in the CF method. Using some extra domain knowledge, this problem can be resolved [14].

3) Scalability:

Usually, ratings and reviews in a system are considered user-item interactions in the RS field. As RSs' always have to deal with large datasets to provide recommendations, the system's scalability is regarded as a major worry. This is where specially developed large-scaled algorithms come to the rescue. But this solution has two sides. A confusing situation might occur because some RS algorithms are at their best for tiny portions of data; however, when applied to massive datasets, they may produce inefficient or worst outcomes. To address this problem, advanced large-scale assessment methodologies are required.

4) Privacy:
In general, an individual must provide his personal information to a recommendation system to receive more beneficial services, but doing so raises concerns about the user's sensitive details. Due to privacy concerns, users may be extra cautious or sometimes reject exposing their sensitive information to such systems. Companies should take necessary steps to restore faith in their users to prevent this issue.

5) Diversity:
Usually, even in a physical store, a person would express interest to buy a product if the store offers variations of products that are rich in diversity. Just like that, if the system can recommend a list of items that are rich in diversity, users may express more interest. Of course, it would not show many results unless the user defines a desire for it or unless it is clearly stated by the user with a specific set of preferences. Users may want to experiment with new and different settings when the RS is first used as a knowledge discovery tool. There hasn't been much research done on this subject so far. However, the need of the hour is to develop solutions that may meet the goal of item diversity while also ensuring the accuracy of recommendations.

6) Multimodal input/output handling:
Users' input is typically not restricted to text in the emerging paradigm of recommender systems known as Conversational Recommender Systems (CRS). Instead, it may take many forms, including images, sounds, and emojis. Which means CRS agents should be able to not only analyze multimodal inputs but also to provide multimodal outputs. This presents a challenge to the RS since it introduces a whole new degree of complexity to the paradigm, such as the capacity to grasp non-textual inputs. These different input types require unique strategies in order to extract their semantic meanings which lead to further model training [15]. Because this is a new field in the world of RS, there is just a small amount of literature on solutions for the challenge. Most notably, a well-known publication in the field [16] investigates the notion of using neural networks to include visual semantic meaning.

Table 1-Open Challenges and Their Solutions of Recommendation Approaches

CHALLENGE	CF	CB	HYBRID APPROACH	SOLUTION	REFERENCES
Cold start problem	√		√	• Lack of ratings of users can be prevailed in a new system by mapping user-item relationships with use of association rule. • A proposal to modify the RS algorithm with item mining and pruning rules, which is applied to mining of movie swarms.	[17] [18] [19]
Sparsity	√		√	• Imputation with high quality by making assumptions about the data production process • To generate extra pseudo ratings, use the few existing ratings or particular item attributes.	[8] [3]
Stability Vs plasticity problem	√	√	√	• Reduce the impact of previous ratings by gradually discounting them. • SI techniques, such as PSO, can obtain feature weights for the user, allowing the matching function to be tailored to the user's preferences.	[20]
Accuracy		√	√	• Fuzzy logic conjunction with CF. • Using mathematical constructs to combine CF and CBF. *Ex – Bayesian networks*	[21] [22]
Gray sheep problem	√		√	• Proposed a method for increasing the size of neighbours in order to use their ratings in computing the prediction function and then improving the recommendations	[23]

IV. Conclusion

Information retrieval and filtering study evolved into recommender systems research, and its gradual development has made it into a powerful and challenging field of research on its own. This research aims to educate researchers and practitioners about various RS approaches, identify the flaws and strengths of each technique, and propose challenges and answers for future systems. Although the interests of industry and academic research are frequently in conflict, there are opportunities to produce significant academic research that may help the industry. Many RS applications are built by industry developers with the assistance of the research and development division, although academic researchers do systematic exploratory experiments. We can design better applications by bridging the gap between academic researchers and industry. In this paper, we looked at the current recommendation approaches' limitations and challenges, as well as alternative extensions and solutions proposed by academic pioneers to improve recommendation capabilities, concluded that when designing recommender system solutions, six grand issues (cold start problem, sparsity, scalability, privacy, and diversity, Multimodal input/output handling) have the most impact. Also, even though RS technologies have been around for over five decades, we believe that some of the issues that were present at the start of the field can still be found today with minor modifications, based on the early state of the art we studied and compared to the most recent state of the art. Therefore, we believe that RS technologies have a long way to go by overcoming significant limitations such as over complexity and cold start problem where it is tough to provide a recommendation as in case of new user/ new item. We think that the issues addressed in this paper will assist in forwarding the debate about choosing RS technologies when developing systems and ethical perspective in the recommender systems community about the future RS technologies with no limitations.

V. References

[1] J. O'donovan and B. Smyth, "The Technical Writer's Handbook. Mill Valley," in *Proceedings of the 10th international conference on Intelligent user interfaces - IUI '05, 2005.M. Young*, 1989.

[2] P. Resnick and H. R. Varian, "Recommender systems," Commun," *Commun. ACM*, vol. 40, no. 3, pp. 56–58, 1997.

[3] E. Çano and M. Morisio, "Hybrid recommender systems: A systematic literature review," *Intell. Data Anal.*, vol. 21, no. 6, pp. 1487–1524, 2017.

[4] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens: An open architecture for collaborative filtering of netnews," in *Proceedings of the 1994 ACM conference on Computer supported cooperative work - CSCW '94*, 1994.

[5] J. A. J Konstan and J. Riedl, "Recommender systems: from algorithms to user experience, " User Model," *User Model. User Model. User-Adapt Interact*, vol. 22, no. 1–2, pp. 101–123, 2012.

[6] S. Jbeel, M. Langer, B. Genzmehr, C. Gipp, and A. Breitinger, "Research paper recommender system evaluation: A quantitative literature survey," in *Proceedings of the International Workshop on Reproducibility and Replication in Recommender Systems Evaluation - RepSys '13*, 2013.

[7] F. Ricci, L. Rokach, and B. Shapira, "Recommender Systems: Introduction and Challenges," in *Recommender Systems Handbook*, Boston, MA: Springer US, 2015, pp. 1–34.

[8] P. Melville and V. Sindhwani, "Recommender Systems," in *Encyclopedia of Machine Learning and Data Mining*, Boston, MA: Springer US, 2017, pp. 1056–1066.

[9] D. Asanov, "Algorithms and methods in recommender systems,", 2011.

[10] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, "Evaluating collaborative filtering recommender systems," *ACM Trans. Inf. Syst.*, vol. 22, no. 1, pp. 5–53, 2004.

[11] J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen, "Collaborative Filtering Recommender Systems," in *The Adaptive Web*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 291–324.

[12] R. Burke, "Hybrid Web Recommender Systems," in *The Adaptive Web*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 377–408.

[13] M. Sharma and S. Mann, "A survey of recommender systems: approaches and limitations," *International Journal of Innovations in Engineering and Technology*, vol. 22, pp. 8–14, 2013.

[14] P. Melville, R. Mooney, and R. Nagarajan, "Content-Boosted Collaborative Filtering for Improved Recommendations," *Proceedings of the National Conference on Artificial Intelligence*, vol. 23, pp. 187–192, 2002.

[15] D. H. Tran, Q. Z. Sheng, W. E. Zhang, S. A. Hamad, N. L. D. Khoa, and N. H. Tran, "Deep conversational recommender systems: Challenges and opportunities," *Computer (Long Beach Calif.)*, vol. 55, no. 4, pp. 30–39, 2022.

[16] L. Liao, Y. Ma, X. He, R. Hong, and T.-S. Chua, "Knowledge-aware multimodal dialogue systems," in *Proceedings of the 26th ACM international conference on Multimedia*, 2018.

[17] S. Maneeroj and A. Takasu, "Hybrid Recommender System Using Latent Features," in *2009 International Conference on Advanced Information Networking and Applications Workshops*, 2009.

[18] L. H. Son, "HU-FCF: A hybrid user-based fuzzy collaborative filtering method in Recommender Systems," *Expert Syst. Appl.*, vol. 41, no. 15, pp. 6861–6870, 2014.

[19] M. H. Mohamed, M. H. Khafagy, and M. H. Ibrahim, "Recommender Systems Challenges and Solutions Survey," in *2019 International Conference on Innovative Trends in Computer Engineering (ITCE)*, 2019.

[20] M. Taghavi, J. Bentahar, K. Bakhtiyari, and C. Hanachi, "New insights towards developing recommender systems," *Comput. J.*, vol. 61, no. 3, pp. 319–348, 2018.

[21] G. Fenza, E. Fischetti, D. Furno, and V. Loia, "A hybrid context aware system for tourist guidance based on collaborative filtering," in *2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011)*, 2011.

[22] L. M. de Campos, J. M. Fernández-Luna, J. F. Huete, and M. A. Rueda-Morales, "Combining content-based and collaborative recommendations: A hybrid approach based on Bayesian networks," *Int. J. Approx. Reason.*, vol. 51, no. 7, pp. 785–799, 2010.

[23] A. E. L. Fazziki, O. E. Aissaoui, Y. E. M. E. Alami, Y. E. L. Allioui, and M. Benbrahim, "A new collaborative approach to solve the gray-sheep user's problem in recommender systems," in *2019 Third International Conference on Intelligent Computing in Data Sciences (ICDS)*, 2019.