Two Decades of Recommender Systems: From Foundational Models to State-of-the-Art Advancements (2004-2024)

Bahareh Rahmatikargar, Pooya Moradian Zadeh, Ziad Kobti
{rahmatib,pooya,kobti}@uwindosr.ca
School of Computer Science, University of Windsor
Windsor, Ontario, Canada

ABSTRACT

This study explores the evolution of recommender systems from 2004 to 2024, highlighting advancements in machine learning and their societal impact. The analysis covers 60 highly cited papers, tracing the journey from foundational models to state-of-the-art hybrid models and the integration and adoption of deep learning techniques.

KEYWORDS

Recommender Systems, Evolution of Recommender Systems, Machine Learning, Deep Learning, Social Network Analysis, Bibliometric Analysis, Highly Cited Papers

ACM Reference Format:
Bahareh Rahmatikargar, Pooya Moradian Zadeh, Ziad Kobti. 2024. Two Decades of Recommender Systems: From Foundational Models to State-of-the-Art Advancements (2004-2024). *ACM/IMS J. Data Sci.* x, x, Article xxx (August 2024), 2 pages. https://doi.org/XXXXXXX.XXXXXXX

1 PROBLEM STATEMENT

Recommender systems have undergone significant transformations over the past two decades, driven by rapid advancements in machine learning [2], and evolving societal demands. Our objective is to analyze the evolution of recommender systems using a combination of content analysis, natural language processing (NLP), and social network analysis techniques. We aim to study the trending topics across different phases of their evolution and map these trends to societal changes outside this domain. Understanding these trends can play an important role in guiding the future direction of these technologies.

2 METHODS

We selected 60 highly cited papers, other than surveys and handbooks, in the field of recommender systems from 2004 to 2024, divided into four distinct periods: 2004-2010, 2010-2015, 2015-2020, and 2020-2024. These papers, obtained from Google Scholar, were analyzed using social network analysis and bibliometric perspectives.

Social Network Analysis: The selected papers were analyzed using IBM Cloud's NLP analytics tool [3], which extracted their categories. A social graph was constructed by linking papers to their categories. We then measured degree, betweenness, and PageRank centralities, to assess their influence and evolution over time. This analysis explored the relationships between different categories and identified how key themes and technologies have risen or declined in prominence across different periods.

Bibliometric Analysis We also conducted bibliometric and content analysis by extracting detailed information on the papers' types, approaches, methodologies, challenges, evaluation metrics, and their explainability, as well as their ability to handle side information.

We complemented our bibliometric and social network analyses with a trend analysis, focusing on how advancements in AI, specifically machine learning techniques, from 2004 to 2024 have driven the evolution of recommender systems. This analysis highlighted the increasing complexity and integration of these systems across various domains, influenced by emerging AI technologies.

3 RESULTS

Social Network Analysis: 2004-2010 (Foundational Years): During this period, categories such as Artificial Intelligence (AI) and Information Security exhibited high degrees of centrality and betweenness. Alongside AI and Information Security, emerging categories such as Internet & Networking and Computing & General Technology began to gain prominence.

2010-2015 (Growth and Integration): The results show that the research domain expanded into new areas, particularly education and entertainment. As a result, categories such as Online Education, Music & Audio, and Video Gaming began to rise in centrality. This shift reflects the increasing complexity and application of these models, which started integrating contextual and social data. These findings align with advancements in online video, and online education platforms as well as e-commerce and social networking websites. The analysis indicates a significant increase in the importance of these categories, highlighting their growing influence as they became more embedded in digital platforms (such as Spotify and Netflix) and educational platforms (such as Coursera, edX, and Khan Academy), enhancing user experiences.

2015-2020 (Deep Learning and Specialization): The adoption of deep learning techniques led to a noticeable shift in the centrality of the categories. Categories such as AI, Information Security, and Biological Sciences gained prominence, indicating the diversification of their applications into specialized fields like healthcare and

Author's address: Bahareh Rahmatikargar, Pooya Moradian Zadeh, Ziad Kobti, {rahmatib,pooya,kobti}@uwindosr.ca,
School of Computer Science, University of Windsor, Windsor, Ontario, Canada.

ACM 2831-3194/2024/8-ARTxxx
https://doi.org/XXXXXXX.XXXXXXX

3D graphics. In this period AI maintained a strong central position while new categories related to deep learning and advanced analytics rose in importance.

2020-2024 (Industry 4.0, Real-Time Interaction): Our results show that the integration of recommender systems with other technologies and real-life applications, especially in the Industry 4.0 era, along with its challenges and opportunities, coupled with the growing emphasis on explainability, has led to the rise of categories such as Explainable AI, Data Storage & Warehousing, and Automotive Technology. These categories have gained significant centrality, reflecting the critical role of recommender systems in sectors like energy management and autonomous vehicles. The analysis highlights the increasing importance of transparency, accountability, and real-time data processing in these fields, driven by the need for recommender systems that can operate effectively in high-stakes, real-time environments.

Bibliometric Analysis: 2004-2010: Highly cited papers from this period focused on foundational methodologies, primarily utilizing collaborative filtering and content-based filtering to address issues such as data sparsity and recommendation accuracy [5, 6]. These works were key in laying the foundation for machine learning techniques, which were later advanced, especially through matrix factorization to enhance recommendations.

2010-2015: This period saw the emergence of hybrid models that combined collaborative filtering, content-based filtering, and early forms of deep learning [9, 10]. Papers highlighted the need for context-aware recommendations, leveraging side information to enhance user engagement. The rise of platforms like Coursera and LinkedIn Learning, as well as streaming services like Netflix and Spotify, was reflected in the focus of these papers, which emphasized the integration of RS into these expanding digital ecosystems.

2015-2020: Deep neural networks, particularly convolutional neural networks (CNNs) and recurrent neural networks (RNNs), became central in the field, enabling more personalized and context-sensitive models, including advanced sequential, session-based, and session-aware models [8, 11]. Highly cited papers explored deep learning for extracting complex patterns from large datasets, advancing specialized RS applications like personalized medicine and 3D content. The integration of NLP paved the way for conversational approaches that interact with users to refine recommendations in real-time.

2020-2024: Recent papers have focused on integrating large language models (LLMs) and NLP, revolutionizing recommendations through more sophisticated conversational interfaces that leverage real-time user feedback [1, 12]. The growing emphasis on explainability and fairness in AI-driven systems is evident, with research highlighting the importance of transparency in decision-making processes, particularly in critical applications such as energy management and autonomous vehicles. Advances in graph neural networks (GNNs), attention mechanisms, and reinforcement learning have improved these systems, expanded the capability of these systesm and enable them to provide more refined recommendations and better handling of bias and interoperability. Additionally, GNNs and community detection have proven valuable in social network analysis and identifying patterns in key applications [4, 7].
Table 1 presents a summary of the bibliometric analysis results, highlighting the key trends across different periods

4 SIGNIFICANCE

In this work, we looked into the evolution of recommender systems through differnet lenses of bibliometric, content, and social network analyses. We believe this research offers a novel perspective by linking these technologies to social events and and rapid progress in AI. This approach reveals the mutual impact between recommender systems and societal changes, demonstrating how the continuous advancements in AI and machine learning have both shaped and been shaped by these systems, thereby helping to predict trends and understand the evolving nature of both the technology and its societal implications.

Period	Popular RS Types	Contributions
2004-2010	Collaborative Filtering, Content-based Filtering	Tackled data sparsity and cold start issues with matrix factorization
2010-2015	Hybrid, Context-aware, Social	Enhanced accuracy through contextual and social integration using early deep learning techniques
2015-2020	Sequential, Session-based, Session-aware, Cross-domain, Graph-based	Improved temporal handling, complex relationships, and cross-domain challenges with advanced deep learning techniques
2020-2024	Conversational, Cross-domain, On-device, Federated, Explainable, Fair RS	Advanced privacy, explainability, and real-time processing with sophisticated techniques like large language models (LLMs).

Table 1: Summary of Key Trends and Contributions in Recommender Systems

REFERENCES

[1] Sunhao Dai, Ninglu Shao, Haiyuan Zhao, Weijie Yu, Zihua Si, Chen Xu, Zhongxiang Sun, Xiao Zhang, and Jun Xu. 2023. Uncovering chatgpt's capabilities in recommender systems. In *Proceedings of the 17th ACM Conference on Recommender Systems*. 1126–1132.
[2] Aminu Da'u and Naomie Salim. 2020. Recommendation system based on deep learning methods: a systematic review and new directions. *Artificial Intelligence Review* 53, 4 (2020), 2709–2748.
[3] IBM. 2024. IBM Natural Language Understanding Service. https://www.ibm.com/demos/live/natural-language-understanding/self-service/home Accessed: 2024.
[4] Farzaneh Jouyandeh, Sarvnaz Sadeghi, Bahareh Rahmatikargar, and Pooya Moradian Zadeh. 2021. Fake news and covid-19 vaccination: a comparative study. In *Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. 525–531.
[5] Yehuda Koren. 2008. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. 426–434.
[6] Andriy Mnih and Russ R Salakhutdinov. 2007. Probabilistic matrix factorization. *Advances in neural information processing systems* 20 (2007).
[7] Bahareh Rahmatikargar, Pooya Moradian Zadeh, and Ziad Kobti. 2022. Social Isolation Detection in Palliative Care Using Social Network Analysis. In *2022 22nd IEEE International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*. IEEE, 905–912.
[8] Fei Sun, Jun Liu, Jian Wu, Changhua Pei, Xiao Lin, Wenwu Ou, and Peng Jiang. 2019. BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer. In *Proceedings of the 28th ACM international conference on information and knowledge management*. 1441–1450.
[9] Aaron Van den Oord, Sander Dieleman, and Benjamin Schrauwen. 2013. Deep content-based music recommendation. *Advances in neural information processing systems* 26 (2013).
[10] Hao Wang, Naiyan Wang, and Dit-Yan Yeung. 2015. Collaborative deep learning for recommender systems. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*. 1235–1244.
[11] Chao-Yuan Wu, Amr Ahmed, Alex Beutel, Alexander J Smola, and How Jing. 2017. Recurrent recommender networks. In *Proceedings of the tenth ACM international conference on web search and data mining*. 495–503.
[12] Jizhi Zhang, Keqin Bao, Yang Zhang, Wenjie Wang, Fuli Feng, and Xiangnan He. 2023. Is chatgpt fair for recommendation? evaluating fairness in large language model recommendation. In *Proceedings of the 17th ACM Conference on Recommender Systems*. 993–999.