Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 4 (2023)

Bridging the Information Gap with Buddy Network: Strategies for Cold Start in Recommender Systems

Namrathaa Malu

Department of CS & IT, iNurture, NagindasKhandwala College, Malad w., Mumbai 400064, INDIA

Abstract - In the realm of modern digital entertainment, the concept of a "Buddy Network" emerges as a central theme when granting access to contacts during the installation of Over-The-Top (OTT) platforms. This permission serves as a pivotal juncture where user experience and data-driven recommendations converge. By establishing connections between user accounts and contact numbers, OTT platforms unveil a wealth of insights encompassing viewing history, user ratings, personal details, and age. Age, a critical variable within this buddy network, plays a significant role in identifying potential "Buddies" among a user's contacts—individuals of similar ages who are predisposed to sharing cinematic preferences. This Buddy Network principle forms the foundation for a sophisticated, personalized recommendation system. By computing average movie ratings based on buddy input, the platform curates tailor-made movie suggestions, directly sourced from within the user's Buddy Network. This intricate process transforms the seemingly routine permission request into a dynamic tool for enriching the user's cinematic journey. Seamlessly blending technology and social dynamics, it offers pertinent and relatable movie recommendations, enhancing the user's digital experience by strengthening the bonds within the Buddy Network and bridging the information gap for the Cold-start problem in the Recommender System.

Index Terms – *cold start, collaborative filtering, recommender system*

I. Background

Recommender systems serve the purpose of predicting a user's preferences, interests, and rankings for specific products or services, subsequently offering tailored recommendations for potential purchases[1]. These systems operate by predicting product choices from a vast inventory, drawing insights from a user's historical purchasing behavior and in some instances their current context. The increasing demand for recommender systems can be attributed to the overwhelming volume of available data in today's digital landscape [2].

Recommender systems are typically classified into three primary categories: Content-based, Collaborative filtering-based, and Hybrid systems [3]. Content-based filtering (CBF) relies on product attributes and user characteristics to suggest items akin to the user's previous purchase [4]. Collaborative filtering (CF), the most widely adopted approach, relies on the preferences of other users who exhibit similar tastes [5]. This is achieved by computing similarities based on the user's historical interactions.

Collaborative filtering can be further subdivided into two subcategories: Memory-based and Model-based [6]. Memory-based or User-based CF takes into account a user's prior product ratings to generate item recommendations, employing similarity metrics and correlations. Conversely, Model-based or item-based collaborative filtering predicts product ratings for recommendations by constructing models utilizing machine learning techniques like clustering, neural networks, decision trees, and more [6][7].

In collaborative filtering, a user's historical purchase behavior encompasses their feedback, which can be either implicit or explicit [7]. Implicit feedback is typically binary and is deduced from observing the user's actions, such as product purchases, song listening, webpage visits, content downloads, and more [8]. Explicit feedback, on the other hand, involves users assigning discrete ratings to each purchased product. Users and products can be classified as either "warm" (where previous interactions are known) or "cold" [9] (where previous interactions are unknown, often involving new users or products without an interaction history). Recommendations for the latter group present the well-recognized challenge of cold-start recommendations.

When collaborative filtering (CF) faces the cold-start problem, providing valuable recommendations to users becomes a challenge because of the absence of prior interaction history [10]. In these scenarios, Content-based filtering (CBF) is often used as an alternative method for making recommendations. However, CBF algorithms suggest products based on their content similarity, which can lead to a lack of diversity among the recommendations[9][12]. This can give rise to several issues, such as when a user finds one recommendation unhelpful, there's a strong likelihood that they may perceive all the other similar recommendations as equally unhelpful.

II. Motivation

The motivation for tackling the cold start problem in recommender systems is underpinned by several compelling factors:

- Ensuring precise recommendations from initial user interactions significantly augments user satisfaction, fostering a positive perception of the system [3][11].
- Effective recommendations for new users contribute to higher user retention rates, thereby ensuring the sustained success of the platform.
- Enhanced recommendations have a direct impact on key performance metrics, such as increased sales, click-through rates, and conversion rates, driving revenue and facilitating business expansion.
- Leveraging available data, even in cases of sparse or non-existent historical data, optimizes resource utilization and enhances the system's functionality [12].

III. Proposed Methodology

In the realm of contemporary digital entertainment, the integration of Over-The-Top (OTT) platforms introduces a compelling solution to the cold start problem in recommender systems. When we embark upon the installation of OTT applications designed for movie consumption, a discerning solicitation emerges the request for access to our contact list. This seemingly innocuous prompt bears profound implications, serving as the foundation of a robust recommender system that addresses the challenge of recommending content to new users.

At its essence, this solicitation serves as a gateway, a prerequisite for navigating the immersive world of cinematic offerings. It is within this permission that the seeds of a seamless user experience are sown, offering an avenue through which the platform can strategically enhance its services. The rationale behind this permission request is rooted in the symbiotic relationship between a user's contacts and their OTT account. Each OTT account, invariably, is intricately linked to a contact number, establishing a foundational connection.

From this interconnected web, the ability to discern existing subscribers within the OTT ecosystem is bestowed upon the platform. This connection within the Buddy Network opens doors to a treasure trove of information, extending beyond mere contact numbers. It allows for the acquisition of a user's viewing history, their rated preferences, personal identifiers such as name, and demographic data, including age. The age variable, in particular, emerges as a potent tool, a bridge to potential social interactions.

Harnessing the age variable, the platform embarks on a journey to discern peers—individuals within the user's contact list whose age falls within a margin of ±5 years. This deliberate pursuit is grounded in the notion that individuals of similar ages are predisposed to forming friendships more effortlessly. Hence, they can be identified as prospective "buddies" within the context of the OTT platform.

As we are acutely aware, the influence of one's social circle is undeniable, and this extends to the realm of cinematic choices. Individuals who share commonalities in age tend to exhibit congruent tastes in movies. The amalgamation of this insight with the data pertaining to watched movies and user ratings becomes the cornerstone of a personalized recommendation system.

The platform diligently calculates an average rating for each movie, taking into account the collective input of users' ratings. Subsequently, a curated selection of the top 5-10 movies, tailor-made to align with the user's cinematic predilections, is proffered. This recommendation, sourced directly from the user's buddies, carries with it a heightened level of relevance and resonance, underpinned by the presumption of shared affinities.

In essence, this meticulous orchestration of data access, user engagement, and personalized recommendation transcends the mundane. It is a testament to the dynamic synergy that exists at the intersection of technology and human connection. This recommendation approach mitigates the cold start problem by providing relevant content based on shared affinities within the Buddy Network.

An Illustrative Example

Let us consider a hypothetical scenario where a new user, referred to as Mike, has recently registered on the Netflix platform with an age of 25 years. Following user consent, we have been granted permission to access Mike's contact information and subsequently, we have meticulously acquired and collated contact information into a structured dataset, which has subsequently undergone data mining and extraction processes to isolate and retrieve the essential variables outlined below:

Table 1 Data Sets of Mike's Contacts

Contact Name	Contact No.	Registered on OTT	Age
John	9123423445	Yes	26
Milly	89343345345	No	37
Camillie	9243454567	Yes	29
Williams	9345781230	Yes	23
Nemar	9121203407	Yes	56
Julie	9457889980	No	45
Steve	9567211212	No	28
Peter	9876543210	Yes	27

Table 2 Data points of Mike's Contacts afterthe removal of unregistered members of the OTT

Contact Name	Contact No.	Registered on OTT	Age
John	9123423445	Yes	26
Camillie	9243454567	Yes	29
Williams	9345781230	Yes	23
Nemar	9121203407	Yes	56
Peter	9876543210	Yes	27

From the dataset at our disposal, our primary objective is to discern individuals who can be categorized as "buddies" under their age proximity to Mike, who is 25 years old. In pursuit of this objective, we will isolate users whose ages fall within a predefined range, specifically within plus or minus 5 years of Mike's age (i.e., individuals aged between 20 and 30 years). Extensive research supports the notion that individuals of similar ages tend to establish social connections more readily, and these connections frequently exhibit shared interests and preferences [13].

Subsequently, through a meticulous process of dataset analysis and the judicious removal of extraneous data, we arrive at a refined dataset that aligns with the specified criteria. This refined dataset serves as a foundational resource for our further investigations.

Table 3 Data points of people who can be considered Mike's buddies

Contact Name	Contact No.	Registered on OTT	Age
John	9123423445	Yes	26
Camillie	9243454567	Yes	29
Williams	9345781230	Yes	23
Peter	9876543210	Yes	27

Now, let us proceed to examine the cinematic and web series preferences of the selected group of individuals encompassing these four users. To facilitate this analysis, we will refer to the dataset presented below:

Table 4 Data points of Mike's buddies watch history with ratings

Name/Movie	M1	M2	M3	M4	M5
John	3	4	1	-	5
Camillie	2	-	-	4	3
Williams	4	1	-	2	4
Peter	4	3	2	1	3

Now, in our pursuit of identifying the highest-rated movie within this dataset, a systematic approach is essential. The initial step involves the aggregation of all individual ratings for each movie. Subsequently, these collected ratings are utilized to calculate the average rating score for every film featured in the dataset. This process is fundamental to ascertaining the relative quality and reception of each movie.

With each movie now associated with an average rating score, the task of identifying the top-rated options becomes more straightforward. Whether one seeks to pinpoint the top movies with the most favorable ratings, this computed data enables us to make precise and informed recommendations based on user preferences and collective feedback.

Table 5 Mean value for each movie

M1	M2	M3	M4	M5
3.25	2.66	1.5	2.33	3.75

Movie M1 and M5 have garnered the highest ratings among all the movies in our dataset. This distinct achievement positions them as prime candidates for recommendation to our new users. In light of the positive reception these films have received from their social connections, it is plausible to anticipate that the user will exhibit a similar affinity for the recommended content. By offering these top-rated cinematic selections, we aim to ensure a gratifying and engaging user experience, setting a promising trajectory for their journey within our platform, and effectively mitigating the cold-start problem.

The practice of employing a mobile number as the primary registration point for user enrollment is a prevalent and widely adopted approach across various applications and services. Leveraging this standard, we have the opportunity to implement a Buddy Network within these applications to provide users with personalized recommendations tailored to their preferences, thereby enhancing the user experience and fostering user engagement.

IV. Results

1. Efficient Sign-Up Process:

When a user registers or signs up for the system, they won't encounter the burden of filling out a time-consuming or extensive form that requires them to provide personal information or specify preferences [2].

Instead, immediately after completing the sign-up process, the system takes swift action to generate and present personalized recommendations or suggestions. This approach enhances user convenience and efficiency, as they can swiftly access valuable content without a lengthy onboarding procedure.

2. Valuable Recommendations from Personal Network:

The recommendations generated by our system are intricately woven into the user's network, particularly within the context of their buddies. This deliberate design principle holds significant potential in terms of the caliber and pertinence of the suggestions proffered. The foundational concept here is that recommendations derived from one's immediate social sphere possess a pronounced capacity to resonate with the user's specific requirements and anticipations. This resonance stems from a shared synergy of preferences and interests prevailing among individuals who constitute the same social network. As a consequence, the recommendations thus crafted are finely attuned to cater to the user's distinct tastes and preferences.

3. Addressing the Lack of Explanation in Collaborative Filtering:

In the realm of user-to-user collaborative filtering (CF), the conventional approach often relies on generating recommendations grounded in similarities with anonymous users. Regrettably, this practice frequently results in recommendations lacking a transparent rationale, as highlighted by scholarly references [14] [15]. This absence of a valid explanation behind the recommendations can indeed be regarded as a notable limitation.

Nonetheless, our system employs a highly effective strategy when presenting recommendations, particularly when suggesting items that have garnered high ratings from within the user's social network. In such instances, the system simplifies the communication by delivering a concise message: "Some of your friends liked this item, are you interested?"

In this context, the user may not necessitate further elaboration or justification for the recommended item. The implicit endorsement from their own friends within their social circle imbues the recommendation with a compelling influence and trust factor. Consequently, the recommendation becomes inherently more convincing and intuitive, harnessing the power of social validation to enhance the user's engagement and satisfaction.

Shortcomings

Rare Possibility of No Contacts on Netflix:

In the worst-case scenario, if none of the user's contacts are linked to OTT platform, the system might face a situation where it lacks any data to provide recommendations. However, it's essential to note that this situation is relatively rare, as many individuals in a user's network are likely to have OTT accounts or engage with the platform in some way.

V. Conclusion

In conclusion, the proposed approach represents a novel and potentially more effective strategy for tackling the inherent challenges associated with user onboarding and addressing the cold start problem within recommender systems. By harnessing a user's pre-existing network of contacts, with due consideration for privacy, we unlock a valuable data source commonly referred to as the "Buddy Network," which significantly enhances the recommendation process.

The system's remarkable capacity to identify a user's buddies, encompassing individuals like friends, family members, colleagues, and acquaintances, offers a unique avenue for delivering personalized recommendations. This approach capitalizes on the well-established principle that individuals who share similar preferences often develop friendships and exhibit shared interests [16]. This not only streamlines the onboarding experience but also elevates the user's propensity to actively engage with the platform.

In essence, this innovative approach to user onboarding and recommendation leverages the influential dynamics of the Buddy Network to orchestrate a more seamless and user-centric experience. By upholding user privacy and harnessing the collective wisdom embedded within a user's network, we aspire to redefine how recommender systems effectively address the vexing cold start problem, paving the way for a more intuitive and engaging user journey.

References

[1] Bernardi, Lucas &Kamps, Jaap&Kiseleva, Julia & Mueller, Melanie. (2015). The Continuous Cold Start Problem in e-Commerce Recommender Systems. CEUR Workshop Proceedings. 1448. [2] R.Ho, K.W.Mai, and M.A.Horowitz, "The future of wires," Proc. IEEE, vol. 89, no 4, 2001, pp. 490-504.

[2] Jie Lu, Dianshuang Wu, Mingsong Mao, Wei Wang, Guangquan Zhang, Recommender system application developments: A survey, Decision Support Systems,Volume 74,2015, Pages 12-32

[3] Masthoff, J.: Group recommender systems: Combining individual models. In: Recommender Systems Handbook, pp. 677–702 (2011) [4] C.J. Alpert, A. Devgan, J.P. Fishburn, and S.T. Quay, "Interconnect synthesis without wire tapering, " IEEE Trans. Computer-Aided Design Integrated Circuits and Systems 20 (1), 2001, pp. 90–104.https://en.wikipedia.org/wiki/Interpersonal_attraction

[5] Fayyaz, Z.; Ebrahimian, M.; Nawara, D.; Ibrahim, A.; Kashef, R. Recommendation Systems: Algorithms, Challenges, Metrics, and Business Opportunities. Appl. Sci. 2020, 10, 7748.

[6] Prem Melville, VikasSindhwani, Recommender Systems Encyclopedia of Machine Learning, 2010, ISBN : 978-0-387-307688.

[7] A. Krzywicki, W. Wobcke, Y.S. Kim, X. Cai, M. Bain, A. Mahidadia, P. Compton, Collaborative Filtering for people-to-people recommendation in online dating: Data analysis and user trial, International Journal of Human-Computer Studies,Volume 76,2015,Pages 50-66,ISSN 1071-5819

[8] Gope, Jyotirmoy& Jain, Sanjay. (2017). A survey on solving cold start problem in recommender systems. 133-138. 10.1109/CCAA.2017.8229786.

[9] Panteli, A.; Boutsinas, B. Addressing the Cold-Start Problem in Recommender Systems Based on Frequent Patterns. Algorithms 2023, 16, 182.

[10] de Carvalho, L.C.; Rodrigues, F.; Oliveira, P. A Hybrid Recommendation Algorithm to Address the Cold Start Problem. In Hybrid Intelligent Systems; Springer International Publishing: Cham, Switzerland, 2020

[11] Zhang, Q., Lu, J. & Jin, Y. Artificial intelligence in recommender systems. Complex Intell. Syst. 7, 439–457 (2021).

[12] Jieun Son, Seoung Bum Kim, Content-based filtering for recommendation systems using multiattribute networks, Expert Systems with Applications,Volume 89,2017,Pages 404-412

[13] Campbell K, Holderness N, Riggs M. Friendship chemistry: An examination of underlying factors☆. SocSci J. 2015 Jun;52(2):239-247. doi: 10.1016/j.soscij.2015.01.005. PMID: 26097283; PMCID: PMC4470381.

[14] Li, J.; Jing, M.; Lu, K.; Zhu, L.; Yang, Y.; Huang, Z. From Zero-Shot Learning to Cold-Start Recommendation. In Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence, Honolulu, Hawaii, USA, 27 January–1 February 2019; Volume 33, pp. 4189–4196

[15] Yang, F.; Liu, F.; Liu, S. Collaborative Filtering Based on a Variational Gaussian Mixture Model. Future Internet 2021, 13, 37.