

Design and Implementation of a Hybrid Recommender System Based on Collaborative, Evolutionary, and Content-Based Techniques

José Luis Junior Gutiérrez Agüero¹ , Christian Daniel von Lücken Martínez² 

¹ Polytechnic School, National University of Asunción
jlgutierrez2000@fpuna.edu.py

² Polytechnic School, National University of Asunción
clucken@pol.una.py

Abstract Recommender systems (RS) are computational tools that predict which items may be of interest to a user, and they are present in many areas of everyday digital life. One of the main challenges faced by RS is the cold-start problem, which arises when the lack of interaction history prevents the generation of reliable recommendations for new users or items. This work presents the design, implementation, and evaluation of an extension of the HP-MOEA (Hybrid Probabilistic Multiobjective Evolutionary Algorithm), applied to the domain of recommender systems. The original algorithm combines user-based collaborative filtering with multiobjective evolutionary algorithms to mitigate this problem. The proposed extension incorporates a content-based component that weights the novelty objective function through thematic affinity profiles between users and items. Experimental results using the MovieLens ml-100k dataset show that the modified algorithm maintains competitive performance in convergence metrics such as generational distance (GD) and its inverted version (IGD), while achieving a 3.4% improvement in hypervolume with respect to the original algorithm, statistically significant and consistent across all evaluated population sizes, reflecting a broader Pareto front and better distributed solutions. The proposal demonstrates that integrating content information into the novelty function improves solution diversity without sacrificing convergence, offering a concrete contribution to the treatment of the cold-start problem.

Keywords Recommender System, Collaborative Filtering, Cold Start, Multiobjective Evolutionary Algorithms

Diseño e Implementación de un Sistema de Recomendación híbrido basado en técnicas colaborativas, evolutivas y de contenido

José Luis Junior Gutiérrez Agüero¹ , Christian Daniel von Lücken Martínez² 

¹ Facultad Politécnica, Universidad Nacional de Asunción
jlgutierrez2000@fpuna.edu.py

² Facultad Politécnica, Universidad Nacional de Asunción
clucken@pol.una.py

Resumen Los sistemas de recomendación (SR) son herramientas computacionales que predicen qué ítems serán de interés para un usuario, y se encuentran presentes en múltiples lugares a lo largo de nuestra vida digital. Uno de los principales desafíos que enfrentan los SR es el problema del inicio frío, que surge cuando la falta de historial de interacciones impide generar recomendaciones confiables para nuevos usuarios o ítems. El presente trabajo presenta el diseño, implementación y evaluación de una extensión del algoritmo HP-MOEA (Hybrid Probabilistic Multiobjective Evolutionary Algorithm), aplicado al dominio de los sistemas de recomendación. El algoritmo original combina filtrado colaborativo basado en usuarios con algoritmos evolutivos multiobjetivo para mitigar este problema. La extensión propuesta incorpora un componente basado en contenido que pondera la función objetivo de novedad mediante perfiles de afinidad temática entre usuarios e ítems. Los resultados experimentales utilizando el dataset MovieLens ml-100k muestran que el algoritmo modificado mantiene un rendimiento competitivo en métricas de convergencia como la distancia generacional (GD) y su versión inversa (IGD), al tiempo que logra una mejora del 3.4 % en hipervolumen respecto al algoritmo original, estadísticamente significativa y consistente en todos los tamaños de población evaluados, reflejando una frontera de Pareto más amplia y soluciones mejor distribuidas. La propuesta demuestra que la integración de información de contenido en la función de novedad mejora la diversidad de soluciones sin sacrificar convergencia, ofreciendo un aporte concreto en el tratamiento del problema del inicio frío.

Palabras clave Sistema de Recomendación, Filtrado Colaborativo, Inicio Frío, Algoritmos Evolutivos Multiobjetivo

1. Introducción

Los sistemas de recomendación (SR) son herramientas que predicen qué ítems serán de interés para un usuario a partir de sus preferencias históricas o del comportamiento de usuarios similares (Aggarwal, 2016). Además de personalizar la experiencia, se estima que entre el 10 % y el 25 % de los ingresos de compañías de *streaming* y comercio electrónico provienen directamente de la efectividad de sus SR (Koren et al., 2009).

Los SR enfrentan desafíos como el inicio frío (*cold start*) y la falta de diversidad en las recomendaciones (Lika et al., 2014). En particular, los ítems nuevos (*newcomers*) tienden a ser ignorados por los métodos tradicionales, reduciendo la capacidad del sistema de explorar nuevas alternativas (Wei et al., 2021). Para abordar estas limitaciones, diversos trabajos han explorado la formulación de la recomendación como un problema de optimización multiobjetivo, empleando algoritmos evolutivos (MOEA) para encontrar compromisos entre objetivos en conflicto como precisión, novedad y diversidad (Zuo et al., 2015; Jain et al., 2020; Cai et al., 2020).

En esta línea, el algoritmo HP-MOEA (Wei et al., 2021) combina filtrado colaborativo basado en usuarios con un esquema evolutivo que integra NSGA-II y SMS-EMOA para optimizar simultáneamente ganancia y novedad. Sin embargo, su cálculo de novedad se limita a contabilizar la cantidad de *newcomers* incluidos en la lista, sin considerar su relevancia para el usuario.

El presente trabajo propone una extensión del HP-MOEA que incorpora un componente basado en contenido al objetivo de novedad, ponderando los *newcomers* según la afinidad temática con el perfil de cada usuario. Esta modificación busca mejorar el balance entre ganancia y novedad, favoreciendo la inclusión de ítems nuevos relevantes y abordando de manera más efectiva el problema de inicio frío. La evaluación experimental sobre MovieLens ml-100k muestra un incremento promedio del 3.4 % en hipervolumen respecto al algoritmo original, estadísticamente significativo y consistente en todos los tamaños de población evaluados, sin deterioro práctico de las métricas de convergencia (GD) y cobertura (IGD).

El resto del artículo se organiza como sigue: la Sección 2 presenta los preliminares sobre SR y optimización multiobjetivo; la Sección 3 revisa los trabajos relacionados; la Sección 4 describe la metodología, incluyendo la modificación propuesta; la Sección 5 presenta los resultados experimentales; y la Sección 6 concluye el trabajo.

2. Preliminares

2.1. Sistemas de recomendación

Los sistemas de recomendación constituyen una clase de herramientas computacionales que intentan predecir qué ítems serán de interés para un usuario, a partir de la información disponible en el sistema o entorno donde se está aplicando y de los patrones observados en las interacciones previas de los usuarios (Aggarwal, 2016). Su finalidad principal es personalizar la experiencia de uso, incrementando la relevancia del contenido mostrado y mejorando con ello la satisfacción, la fidelización y la interacción dentro del sistema (Koren et al., 2009).

Existen dos modelos básicos de SR:

1. **Basado en contenido.** Recomienda ítems en función de sus características descriptivas (género, director, actores), comparándolas con el perfil de preferencias del usuario. Puede recomendar ítems nuevos sin depender de valoraciones de otros usuarios, aunque tiende a la sobre-especialización (Aggarwal, 2016).
2. **Filtrado colaborativo.** Se fundamenta en que usuarios con comportamientos similares en el pasado tenderán a coincidir también en el futuro (Aggarwal, 2016; Dong et al., 2022). Sus variantes principales son: basado en usuarios, que recomienda ítems consumidos por usuarios similares; basado en ítems, que recomienda ítems similares a los ya consumidos; y basado en modelos, que utiliza técnicas de aprendizaje automático para generalizar patrones en la matriz de calificaciones usuario-ítem.

Al depender de la información histórica de usuarios e ítems, la incorporación de nuevos elementos al sistema da lugar al problema del inicio frío (cold start), donde la ausencia de interacciones previas dificulta la estimación de preferencias y limita la precisión de las recomendaciones iniciales (Aggarwal, 2016; Lika et al., 2014). Este problema se manifiesta en tres variantes: de usuario, cuando no existe historial de interacciones; de ítem, cuando no se han recibido evaluaciones; y de sistema, cuando la plataforma completa carece de datos suficientes (Lika et al., 2014).

2.2. Optimización multiobjetivo con algoritmos evolutivos

La optimización multiobjetivo (MOO) busca resolver problemas con varios objetivos simultáneos en conflicto. A diferencia de la optimización monoobjetivo, no existe una única solución óptima, sino un conjunto de soluciones que representan diferentes compromisos entre los objetivos (Kubat, 2015).

Estos problemas se formulan como:

$$\begin{aligned}
& \text{mín } f_m(x), \quad m = 1, \dots, M, \\
& \text{s.a. } g_j(x) \leq 0, \quad j = 1, \dots, J, \\
& h_k(x) = 0, \quad k = 1, \dots, K, \\
& x_i^L \leq x_i \leq x_i^U, \quad i = 1, \dots, N.
\end{aligned} \tag{1}$$

donde $x = (x_1, \dots, x_N)$ es el vector de variables de decisión, $f_m(x)$ las funciones objetivo, $g_j(x)$ y $h_k(x)$ las restricciones de desigualdad e igualdad, y x_i^L, x_i^U los límites de cada variable (Kubat, 2015).

Una solución x domina a otra x' si no es peor en ningún objetivo y es estrictamente mejor en al menos uno. El conjunto de soluciones no dominadas constituye el frente de Pareto, que representa los compromisos óptimos entre objetivos (Kubat, 2015).

Algoritmos evolutivos multiobjetivo Los algoritmos evolutivos (AE) son métodos de optimización inspirados en la evolución darwiniana, que utilizan selección natural, cruce y mutación para guiar la búsqueda de soluciones. Se clasifican como metaheurísticas estocásticas: no garantizan la solución óptima, pero resultan eficaces para aproximarse a ella en problemas de alta complejidad (Kubat, 2015). Los Algoritmos Evolutivos Multiobjetivo (MOEA) extienden este paradigma para aproximar el frente de Pareto, aprovechando la naturaleza poblacional para explorar simultáneamente diferentes regiones del espacio de búsqueda y mantener un equilibrio entre convergencia y diversidad (Blank y Deb, 2020).

Un AE típico opera sobre una población de individuos, donde cada individuo codifica una solución candidata mediante un cromosoma. En cada generación se aplican operadores de cruce (combinación de padres para generar hijos) y mutación (alteración aleatoria de genes para mantener diversidad), seguidos de un mecanismo de selección guiado por una función de aptitud que favorece la supervivencia de los individuos más prometedores (Kubat, 2015).

Entre los MOEA más influyentes se destacan NSGA-II y SMS-EMOA, que constituyen la base del algoritmo propuesto en este trabajo.

NSGA-II Deb et al. (2002) emplea tres mecanismos fundamentales: un ordenamiento no dominado rápido que clasifica la población en frentes de Pareto sucesivos; una distancia de hacinamiento que mide la densidad relativa dentro de cada frente para favorecer la diversidad; y una selección elitista que combina padres e hijos y retiene los mejores individuos según frente y hacinamiento. Esto le permite lograr un balance entre convergencia y diversidad con complejidad $O(MN^2)$, donde M es el número de objetivos y N el tamaño de la población (Deb et al., 2002).

SMS-EMOA Beume et al. (2007) integra directamente el hipervolumen dominado como criterio de selección. Opera en estado estacionario: en cada iteración genera un solo descendiente y elimina del peor frente al individuo con menor contribución exclusiva al hipervolumen. Este enfoque garantiza que la población mantenga o aumente el volumen dominado, favoreciendo tanto la convergencia como una distribución uniforme sobre la frontera. Ha mostrado ventajas en calidad de aproximación, especialmente en problemas con dos o tres objetivos, aunque con mayor costo computacional por el cálculo del hipervolumen (Beume et al., 2007).

2.3. Trabajos relacionados

La aplicación de MOEA a sistemas de recomendación ha progresado en la última década. A continuación se describen los trabajos más representativos.

Personalized Recommendation Based on Evolutionary Multi-Objective Optimization (Zuo et al., 2015) formalizó la recomendación como problema multiobjetivo, combinando Probabilistic Spreading con NSGA-II para optimizar precisión y cobertura, codificando múltiples usuarios en un solo individuo.

Multi-objective Item Evaluation for Diverse as Well as Novel Item Recommendations (Jain et al., 2020) introdujo el marco NewCrossPMOEA con similitud basada en el coeficiente de Bhattacharyya y un operador de cruce multi-generacional, optimizando precisión y diversidad mediante NSGA-II.

A Hybrid Recommendation System with Many-objective Evolutionary Algorithm (Cai et al., 2020) combinó filtrado colaborativo y basado en contenido con RVEA para optimizar cuatro objetivos simultáneamente, superando a NSGA-II y MOEA/D en distribución de soluciones.

An Improved Matrix Factorization Based Model for Many-objective Optimization Recommendation (Cui et al., 2021) propuso MaORA-IMF, un modelo de dos capas con factorización de matrices regularizada y NSGA-III para optimizar precisión, recall, diversidad y novedad.

A Hybrid Probabilistic Multiobjective Evolutionary Algorithm for Commercial Recommendation Systems (Wei et al., 2021) propuso HP-MOEA, combinando filtrado colaborativo con versiones modificadas de NSGA-II y SMS-EMOA para optimizar ganancia y novedad. Este trabajo constituye la base directa del algoritmo propuesto en el presente trabajo, donde se extiende el objetivo de novedad incorporando información de contenido.

3. Metodología y resultados

3.1. Descripción conceptual

El HP-MOEA optimiza ganancia y novedad simultáneamente, permitiendo introducir ítems sin historial de calificaciones (*newcomers*) en las listas de recomendación. Sin embargo, el cálculo original de novedad se limita a contabilizar la cantidad de *newcomers* incluidos, sin considerar si estos guardan relación con las preferencias del usuario. En consecuencia, el algoritmo puede recomendar ítems nuevos irrelevantes, afectando la utilidad real de la recomendación. Esta limitación es especialmente crítica en datasets como MovieLens, donde cada película cuenta con metadatos de contenido (géneros cinematográficos) que no son aprovechados en la versión original.

Para superar esta limitación, se construyen perfiles de usuario a partir de sus calificaciones históricas y la información de contenido de los ítems. Cada usuario u posee un vector de preferencias P_u , inferido de la frecuencia y valoración otorgada a géneros previamente consumidos. La relevancia entre un usuario u y un ítem i se define como:

$$\text{Rel}(u, i) = \sum_{g \in G} P_u(g) \cdot C_i(g) \quad (2)$$

donde G es el conjunto de géneros, $P_u(g)$ la preferencia de u por el género g , y $C_i(g)$ la pertenencia binaria del ítem i a dicho género. Este valor cuantifica la afinidad temática entre un *newcomer* y los intereses del usuario.

La propuesta consiste en reemplazar la función objetivo de novedad del HP-MOEA por una versión ponderada que incorpora esta afinidad. La idea de utilizar perfiles contruidos a partir de metadatos proviene de los SR basados en contenido (Aggarwal, 2016); sin embargo, su aplicación al cálculo del objetivo de novedad en el HP-MOEA constituye el aporte central de este trabajo. Para un usuario u y su lista de recomendación L_u , la novedad ponderada se define como:

$$\text{Novelty}^*(u) = \frac{1}{|G|} \sum_{i \in L_u \cap N} \text{Rel}(u, i) \quad (3)$$

donde N es el conjunto de *newcomers* y $|G|$ el número total de géneros. Adicionalmente, se redefine el individuo máximo de novedad: en lugar de seleccionar *newcomers* de forma indiscriminada, se priorizan aquellos con mayor relevancia según el perfil de cada usuario, asegurando que el extremo del frente de Pareto asociado a la novedad represente la mejor configuración posible en términos de afinidad temática.

La modificación introduce cambios únicamente en el preprocesamiento de datos y en el cálculo del objetivo de novedad, con complejidades adicionales de $O(m)$ y $O(1)$ respectivamente. La complejidad total del HP-MOEA modificado se mantiene en $O(T \cdot N \cdot m^2 \cdot L)$.

3.2. Implementación

Para la validación experimental se utilizó el dataset *MovieLens ml-100k* (Harper y Konstan, 2015), mantenido por el GroupLens Research Group de la Universidad de Minnesota y el MIT. Este conjunto contiene 100.000 calificaciones (escala 1–5) de 943 usuarios sobre 1.682 películas, e incluye atributos de contenido como género, año de estreno y datos demográficos de los usuarios.

La implementación se realizó en Python 3.12 utilizando la biblioteca pymoo para los algoritmos evolutivos, NumPy, pandas y matplotlib. La configuración experimental siguió los parámetros del trabajo original (Wei et al., 2021): tamaño de lista $L = 10$, población $N = 100$, generaciones $T = 3000$, probabilidad de mutación $p_m = 1/L$ y punto de transición de NSGA-II a SMS-EMOA en $\mu = 500$.

El filtrado colaborativo basado en usuarios constituye el primer componente del HP-MOEa. Para determinar el valor óptimo del parámetro de vecinos k se dividió el dataset en 80 % entrenamiento y 20 % prueba, evaluando el MSE para $k \in \{5, 10, 15, 20, 25, 30, 35\}$. Se observó que al aumentar k el MSE de prueba crece a partir de $k = 10$, evidenciando sobreajuste. Se seleccionó $k = 10$ por ofrecer el mejor equilibrio entre ajuste y generalización.

Para validar la corrección de la implementación se comparó el hipervolumen obtenido con el reportado por en el trabajo original (Wei et al., 2021): la implementación propia alcanzó un promedio de 0.51738 ($\sigma = 0,0008$) frente al 0.51739 ($\sigma = 0,0011$) original. La diferencia de 0.00001, inferior al 0.002 % y muy por debajo de la desviación estándar, confirma la fidelidad de la replicación.

3.3. Métricas

Se emplearon tres indicadores de calidad para evaluar los frentes de Pareto obtenidos:

- **Hipervolumen (HV):** mide la porción del espacio objetivo dominada por el conjunto de soluciones. Constituye la métrica principal, utilizada también en el trabajo original (Wei et al., 2021).
- **Generational Distance (GD):** mide convergencia mediante la distancia promedio desde las soluciones obtenidas hacia el frente verdadero (Kubat, 2015).
- **Inverted Generational Distance (IGD):** evalúa simultáneamente convergencia y cobertura midiendo la distancia en sentido inverso (Kubat, 2015).

Las dos últimas métricas no fueron utilizadas por Wei et al. y se incorporan en este trabajo para una evaluación más completa. La implementación base obtuvo $GD = 0.0009$ ($\sigma = 0,0004$) e $IGD = 0.0047$ ($\sigma = 0,0003$), valores que confirman buena convergencia y cobertura del frente de Pareto y sirven como línea base para evaluar el impacto de la modificación propuesta.

3.4. Resultados

El algoritmo original y la versión modificada se ejecutaron 30 veces cada uno con diferentes semillas aleatorias para $N = 100$. El hipervolumen promedio del algoritmo modificado fue de 0.5926 ($\sigma = 0,0010$) frente a 0.5732 ($\sigma = 0,0008$) del original, ambos computados con la misma configuración, lo que representa un incremento del 3.39 %. El IGD se mantuvo estable (0.0047 en ambas variantes), mientras que el GD registró un leve incremento (de 0.0009 a 0.0013, $\sigma = 0,0007$), de magnitud despreciable frente a la ganancia en hipervolumen. Esto indica que el frente de Pareto obtenido conserva una proximidad y distribución equivalentes respecto a la frontera de referencia, pero domina una región significativamente mayor del espacio objetivo. La Figura 1 ilustra este comportamiento para la corrida mediana en hipervolumen con $N = 100$.

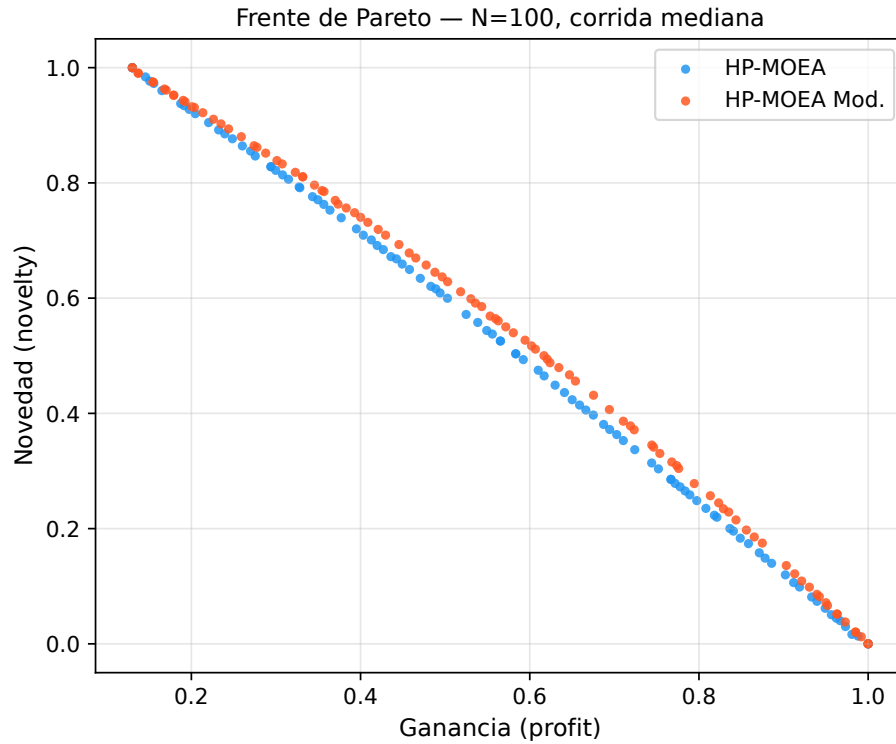


Figura 1. Frentes de Pareto superpuestos de la corrida mediana en hipervolumen para $N = 100$ (mismas semillas). El frente modificado (naranja) domina una mayor región que el original (azul).

Para validar la robustez de la mejora, se repitieron los experimentos con poblaciones de 50, 150 y 200 individuos, evaluando cada configuración con 30 corridas independientes y 30 corridas adicionales con semillas compartidas entre ambas versiones para comparación controlada. Los resultados se presentan en el Cuadro 1.

Población	Diferentes semillas				Mismas semillas			
	HV (Orig.)	Desv.	HV (Mod.)	Desv.	HV (Orig.)	Desv.	HV (Mod.)	Desv.
50	0.5636	0.0042	0.5828	0.0023	0.5644	0.0040	0.5832	0.0033
100	0.5732	0.0008	0.5926	0.0010	0.5734	0.0007	0.5929	0.0008
150	0.5761	0.0005	0.5963	0.0005	0.5761	0.0006	0.5963	0.0005
200	0.5777	0.0004	0.5984	0.0004	0.5778	0.0004	0.5981	0.0008

Cuadro 1. Hipervolumen promedio y desviación estándar para distintos tamaños de población.

La versión modificada supera consistentemente al algoritmo original en todos los tamaños de población, tanto con semillas diferentes como compartidas. El incremento relativo es estable a lo largo de todo el rango, entre 3.3 % y 3.5 %, lo que evidencia que la mejora no depende del tamaño de población. La consistencia observada entre los resultados con semillas compartidas e independientes confirma que la mejora no es atribuible a la variabilidad estocástica, sino al efecto de la ponderación por contenido en el objetivo de novedad. La baja desviación estándar en todas las configuraciones, indica que la modificación no introduce inestabilidad en la convergencia del algoritmo.

3.5. Significancia estadística

Para descartar que las diferencias observadas se deban al azar, se aplicaron pruebas no paramétricas sobre las 30 corridas de cada configuración. Sobre las semillas compartidas se empleó la prueba de Wilcoxon pareada (comparación par a par entre original y modificado bajo la misma semilla) y sobre las semillas independientes la prueba de Mann-Whitney U. Los p -valores se corrigieron por comparaciones múltiples mediante el método de Holm-Bonferroni, y el tamaño del efecto se cuantificó con el estadístico de Vargha-Delaney $\hat{A}_{12}$, definido aquí como la probabilidad de que una corrida del algoritmo original presente un valor de la métrica superior al de una corrida del modificado. En el Cuadro 2, los casos significativos ($p < 0,05$) se destacan en negrita; un valor $\hat{A}_{12} = 0,00$ en HV indica separación completa a favor del modificado, mientras que en GD e IGD un $\hat{A}_{12} < 0,5$ favorece al original.

El Cuadro 2 resume los resultados. Para el hipervolumen, la diferencia es significativa en todos los tamaños de población y bajo ambos esquemas de semilla ($p < 0,001$ tras la corrección), con $\hat{A}_{12} = 0,00$: existe separación completa, es decir, toda corrida del modificado domina en hipervolumen a toda corrida del original. Este es el resultado más contundente y respalda el aporte central del trabajo. Para GD, el modificado presenta un incremento leve pero detectable a partir de $N = 100$ ($\hat{A}_{12} < 0,5$, favorable al original), coherente con el ligero estiramiento del frente hacia regiones de mayor novedad ponderada; no obstante, la diferencia absoluta resulta despreciable. Para IGD no se observa diferencia significativa en la práctica totalidad de los casos, lo que confirma que la cobertura del frente se mantiene estable. En conjunto, la modificación mejora de forma concluyente el hipervolumen sin comprometer la convergencia ni la distribución de soluciones sobre la frontera de Pareto.

Métrica	Prueba	$N = 50$		$N = 100$		$N = 150$		$N = 200$	
		p	$\hat{A}_{12}$	p	$\hat{A}_{12}$	p	$\hat{A}_{12}$	p	$\hat{A}_{12}$
HV	Wilcoxon	<0.001	0.00	<0.001	0.00	<0.001	0.00	<0.001	0.00
	Mann-Whitney	<0.001	0.00	<0.001	0.00	<0.001	0.00	<0.001	0.00
GD	Wilcoxon	0.056	0.36	0.035	0.28	0.056	0.28	<0.001	0.15
	Mann-Whitney	0.501	0.45	0.011	0.27	<0.001	0.18	<0.001	0.21
IGD	Wilcoxon	1.000	0.53	0.293	0.37	0.227	0.37	0.013	0.27
	Mann-Whitney	1.000	0.56	1.000	0.49	0.097	0.31	0.227	0.34

Cuadro 2. Significancia estadística (original vs. modificado): p -valores corregidos por Holm-Bonferroni y tamaño de efecto Vargha-Delaney $\hat{A}_{12}$. Wilcoxon pareada sobre semillas compartidas y Mann-Whitney U sobre independientes.

4. Conclusiones

Este trabajo presentó una extensión al algoritmo HP-MOEA (Wei et al., 2021) aplicado a Sistemas de Recomendación, incorporando información basada en contenido al objetivo de novedad para mejorar el tratamiento de ítems *newcomers*. La modificación pondera los ítems nuevos según su afinidad temática con los perfiles de usuario, integrando así características de los SR basados en contenido dentro de un marco evolutivo multiobjetivo.

Los resultados experimentales sobre MovieLens ml-100k muestran que la versión modificada logra un incremento del 3.4% en hipervolumen manteniendo GD e IGD en valores estables, lo que indica que la mejora en cobertura del espacio objetivo no compromete la convergencia ni la distribución de soluciones sobre la frontera de Pareto. La mejora se mantuvo consistente y estadísticamente significativa en todos los tamaños de población evaluados, confirmando la robustez del enfoque.

La incorporación de información de contenido en la función objetivo de novedad permite abordar de manera más efectiva el problema de inicio frío, sin sacrificar el rendimiento global del algoritmo ni incrementar significativamente su complejidad computacional.

4.1. Trabajos futuros

Entre las líneas de extensión se identifican: la incorporación de objetivos adicionales como diversidad, cobertura o serendipia hacia un esquema *many-objective*; la evaluación de algoritmos alternativos como MOEA/D o NSGA-III; el uso de fuentes de contenido más ricas mediante *embeddings* obtenidos con técnicas de *deep learning*; la paralelización del algoritmo evolutivo mediante tecnologías de aceleración por GPU; y la validación en datasets de mayor escala como MovieLens-20M para analizar escalabilidad.

Referencias

- Aggarwal, Charu C. (2016). *Recommender Systems: The Textbook*. Cham, Switzerland: Springer. ISBN: 978-3-319-29657-9. DOI: [10.1007/978-3-319-29659-3](https://doi.org/10.1007/978-3-319-29659-3). URL: <https://doi.org/10.1007/978-3-319-29659-3>.
- Beume, N., B. Naujoks y M. Emmerich (2007). «SMS-EMOA: Multiobjective selection based on dominated hypervolume». En: *European Journal of Operational Research* 181.3, págs. 1653-1669. DOI: [10.1016/j.ejor.2006.08.008](https://doi.org/10.1016/j.ejor.2006.08.008).
- Blank, J. y K. Deb (2020). «pymoo: Multi-Objective Optimization in Python». En: *IEEE Access* 8.
- Cai, Xingjuan, Zhaoming Hu, Peng Zhao, WenSheng Zhang y Jinjun Chen (2020). «A hybrid recommendation system with many-objective evolutionary algorithm». En: *Expert Systems with Applications* 159. ISSN: 0957-4174. DOI: <https://doi.org/10.1016/j.eswa.2020.113648>. URL: <https://www.sciencedirect.com/science/article/pii/S0957417420304723>.
- Cui, Z., P. Zhao, Z. Hu, Z. Cai, X. Zhang y W. Chen (2021). «An improved matrix factorization based model for many-objective optimization recommendation». En: *Information Sciences* 579. DOI: <https://doi.org/10.1016/j.ins.2021.07.077>.
- Deb, K., A. Pratap, S. Agarwal y T. Meyarivan (2002). «A fast and elitist multiobjective genetic algorithm: NSGA-II». En: *IEEE Transactions on Evolutionary Computation* 6.2.
- Dong, Zhenhua, Zhe Wang, Jun Xu, Ruiming Tang y Jirong Wen (2022). «A Brief History of Recommender Systems». En: *Proceedings of ACM Conference (Conference'17)*. Accepted by DLP-KDD 2022. ACM.
- Harper, F. Maxwell y Joseph A. Konstan (dic. de 2015). «The MovieLens Datasets: History and Context». En: *ACM Transactions on Interactive Intelligent Systems (TiiS)* 5.4, 19:1-19:19. DOI: [10.1145/2827872](https://doi.org/10.1145/2827872). URL: <http://dx.doi.org/10.1145/2827872>.
- Jain, Ankush, Pramod Kumar Singh y Dhar Joydip (2020). «Multi-objective item evaluation for diverse as well as novel item recommendations». En: *Expert Systems with Applications* 139. DOI: [10.1016/j.eswa.2019.112857](https://doi.org/10.1016/j.eswa.2019.112857).
- Koren, Y., R. Bell y C. Volinsky (oct. de 2009). «Matrix factorization techniques for recommender systems». En: *Computer* 42.8, págs. 30-37.
- Kubat, M. (2015). *An Introduction to Machine Learning*. Springer. Cap. 13 The Genetic Algorithm.
- Lika, Bleta, Kostas Kolomvatsos y Stavros Hadjiefthymiades (2014). «Facing the cold start problem in recommender systems». En: *Expert Systems With Applications* 41.4. DOI: [10.1016/j.eswa.2013.09.005](https://doi.org/10.1016/j.eswa.2013.09.005).
- Wei, Guijin, Qinglong Wu y MengChu Zhou (2021). «A Hybrid Probabilistic Multiobjective Evolutionary Algorithm for Commercial Recommendation Systems». En: *IEEE Transactions On Computational Social Systems* 8.3. DOI: [10.1109/tcss.2021.3055823](https://doi.org/10.1109/tcss.2021.3055823).
- Zuo, Y., M. Gong, J. Zeng, L. Ma y L. Jiao (2015). «Personalized recommendation based on evolutionary multi-objective optimization [research frontier]». En: *IEEE Comput. Intell. Mag.* 10.1.