International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 31, No. 2, February, 2026, pp. 335-343
Journal homepage: http://iieta.org/journals/isi

A Hybrid Bundle Recommender System Using Modified FP-Growth and Compound Critiquing for E-Commerce

Rachmi Helfianur, Z. K. A. Baizal*

School of Computing, Telkom University, Bandung 40257, Indonesia

Corresponding Author: baizal@telkomuniversity.ac.id

https://doi.org/10.18280/isi.310202

Received: 15 August 2025
Revised: 1 November 2025
Accepted: 19 February 2026
Available online: 28 February 2026

Keywords:
bundle recommender system, hybrid recommender system, Compound Critiquing, conversational recommender system, e-commerce

ABSTRACT

Bundle recommender systems enhance user experience in e-commerce by suggesting personalized product bundles. Traditional systems, however, often generate monotonous recommendations, limiting user exploration. This study presents a hybrid bundle recommender system that combines MFP-Growth and Compound Critiquing to address these challenges. MFP-Growth extracts association rules from transaction data to generate candidate bundles based on support, confidence, and lift values. Compound Critiquing enables users to provide multi-attribute feedback, refining recommendations dynamically. We evaluated the system using an e-commerce book transaction dataset. Results show that the proposed system improves recommendation quality, with MAP@10 increasing from 0.181436 to 0.181556 and nDCG@10 rising from 0.327025 to 0.327216 compared to the baseline. Additionally, MFP-Growth achieves 30% faster computation and reduced memory usage. The proposed system offers better computational efficiency, increased recommendation diversity, and enhanced novelty, providing a promising solution for real-time, personalized recommendations. These results demonstrate that integrating MFP-Growth with Compound Critiquing can significantly improve both recommendation accuracy and system performance, making it a suitable approach for e-commerce applications.

1. INTRODUCTION

In recent years, e-commerce has become a primary channel for digital transactions. Despite continuous technological advances, significant limitations remain, particularly in delivering relevant and personalized shopping experiences [1]. One challenge stems from user behavior: customers often purchase multiple related products in a single transaction, yet most recommender systems are still limited to suggesting individual items and are not capable of identifying and recommending product bundles frequently bought together. Consequently, users are required to perform manual searches, which are time-consuming and may lead to the omission of relevant products [2-4].

Bundle recommenders address this issue by recommending a set of complementary items as a single package, improving search efficiency while supporting marketing strategies such as cross-selling and up-selling [3]. However, conventional bundle recommendation still has limitations: there is little personalization of the resulting bundles because they are often created manually without considering each user's unique preferences [3]. Prior studies have attempted to tackle this through approaches such as collaborative bundle recommendation and personalized bundle generation [5].

However, this system still faces several unresolved challenges. First, the interaction problem is sparse: the number of user interactions with bundles is much lower than with individual products, making it difficult for systems to accurately learn user preferences [3, 6]. Second, there is the output space complexity problem: the number of possible item combinations in a bundle is extremely large, making it hard for the system to select an optimal bundle. Third, many bundle recommender systems are still static and cannot adapt to user-specific contexts and needs, which affects recommendation quality and relevance [7-10].

In addition to these technical limitations, most existing bundle recommendation systems are still static, as bundles are formed entirely based on historical transaction patterns. This condition gives rise to the problem of bundle monotony, namely, the tendency of the system to produce bundles with items that are too homogeneous, for example, coming from the same category or attribute. As a result, user exploration is limited, and the level of recommendation personalization decreases.

A conversational bundle recommender has been proposed to address these challenges [3], in which the system and user interact progressively through real-time feedback. One study adopting this approach is Bundle MCR (Multi-Round Conversational Recommendation) [3], which introduces a multi-turn interaction model between the user and the system. In this model, the system can choose to ask about user preferences or gradually recommend items in sequence [3]. The Bundle MCR framework is built on the Bundle BERT architecture and is trained using a two-stage strategy: clozestyle pre-training and reinforcement learning fine-tuning [3]. Although highly flexible and adaptive, this approach requires complex computing infrastructure and substantial resources. Compound Critiquing is one of the more adaptive innovations [11], which allows users to provide feedback on multiple attributes at once, speeding up the adjustment process and improving accuracy. For example, in the book domain, users can provide inputs such as “latest edition” or “different author” to filter recommendations based on content preferences and interaction history.

Nevertheless, the effectiveness of a recommender system still depends heavily on its ability to identify association patterns among products. A suitable approach for this is ARM, which is used to find items that frequently co-occur in users’ transaction histories. ARM has proven capable of capturing complex product relationships and is widely used in recommender systems. Common ARM algorithms include Apriori and FP-Growth [11]. Of the two, FP-Growth is known to be more efficient because it eliminates the need to generate all candidate itemsets explicitly. However, this algorithm still faces challenges in execution time and memory usage when applied to large-scale datasets [12]. The MFP-Growth algorithm was developed to overcome these limitations [12] by improving the FP-tree structure, optimizing the header table, and eliminating the need to repeatedly construct conditional subtrees, resulting in faster performance, lower memory consumption, and higher-quality association rules [12]. In addition to these technical limitations, most existing bundle recommendation systems are still static, as bundles are generated entirely based on historical transaction patterns. This creates the problem of bundle monotony, a system's tendency to generate bundles with items that are too homogeneous, such as those from the same category or attribute. As a result, user exploration is limited, and the level of recommendation personalization decreases.

This study proposes integrating MFP-Growth with Compound Critiquing to build a dynamic, efficient, and personalized bundle recommender system. MFP-Growth rapidly extracts product association patterns, while Compound Critiquing ensures the system adapts to user preferences. This combination is expected to improve recommendation relevance, accelerate search, broaden product exploration, and enhance user satisfaction on e-commerce platforms.

2. RELATED WORK

Bundle recommender systems are a branch of recommender systems that focus on suggesting sets of complementary items or products as a package [5]. Unlike single-item recommendations, bundle recommendations consider combinations of items that deliver added value when consumed together (whether in terms of utility, attribute compatibility, or price advantages). This concept is widely applied across domains such as e-commerce, entertainment services, and online games, to increase basket value and user satisfaction [6].

In general, research on bundle recommendation can be grouped into three main approaches. First, association-based methods rely on ARM to discover relationships among items in historical transaction data [13]. Their key strength is high interpretability, since the resulting rules are easy for both developers and users to understand. However, they tend to be static because once extracted, the rules are not easily adapted to changing user preferences or new trends.

Second, collaborative filtering (CF) based methods leverage similarities among users or items to recommend bundles [14]. These methods enable more personalized recommendations by considering similarity patterns derived from user ratings or interaction data. Nevertheless, CF faces challenges related to data sparsity (when user–item interactions are limited) and the cold-start problem, especially for new items or users.

Third, learning-to-rank and deep learning-based approaches use machine learning models to optimize the ranking of recommended bundles. These methods typically integrate multiple data sources, ranging from content features and user behaviour to item metadata, to produce adaptive recommendations. While they excel at capturing complex patterns, they require substantial computational resources, long training times, and sometimes sacrifice interpretability [14, 15].

Some studies also combine these approaches to form hybrid bundle recommender systems. Such systems exploit the strengths of each method to offset their weaknesses, for example, by using ARM to obtain initial bundle associations and then refining them with CF or neural networks for deeper personalization. Although promising, these hybrids face challenges in system integration and in balancing the trade-off between computational efficiency and recommendation quality [12].

Based on the literature review, this study proposes a hybrid bundle recommendation system that integrates MFP-Growth and Compound Critiquing. MFP-Growth is used to efficiently form bundle candidates through optimizing the FP-tree structure, while Compound Critiquing allows the system to dynamically adapt to user preferences through multi-attribute feedback. This hybrid approach is expected to alleviate the issue of bundle monotony and provide efficient, adaptive, and personalized bundle recommendations for use in e-commerce environments.

3. MODIFIED FP-GROWTH ALGORITHM

ARM is a core method in data mining aimed at discovering interesting relationships among items in large datasets, particularly transactional data. ARM focuses on identifying items that frequently co-occur and forming rules in the format $X \rightarrow Y$, where X and Y are subsets of the item universe and are disjoint ($X \cap Y = \emptyset$) [12]. In this context, X is called the antecedent, and Y is the consequent [12]. These associative rules are important for uncovering hidden patterns in data and have been widely applied across sectors such as retail, telecommunications, IoT systems, patient symptom analysis, and even transportation scheduling [12]. The two primary metrics used to evaluate such rules are:

- Support measures how often X and Y occur together in the data.

$$Support\ (X \rightarrow Y) = \frac{X \cup Y}{DB} \tag{1}$$

- Confidence reflects the strength of the rule, namely the probability of Y occurring when X occurs.

$$Confidence\ (X \rightarrow Y) = \frac{support\ X \cup Y}{support\ (X)} \tag{2}$$

An itemset is called frequent if its support exceeds a userdefined minimum threshold [13]. By first identifying frequent itemsets, the system then derives associative rules based on minimum support and confidence thresholds, eliminating any rules that do not meet these minima.

The proposed MFP-Growth algorithm is designed to reduce overall execution time and memory usage [16]. In conventional FP-Growth, the search proceeds by traversing the header table in descending order of item frequency, followed by scanning the current FP-tree to obtain the conditional pattern base for each item [17]. The system then repeatedly reconstructs conditional FP-trees without fully optimizing pruning strategies. Typically, pattern growth proceeds through a sequence of suffix patterns derived from frequently occurring itemsets. Figure 1 illustrates the recursive mining process in FP-Growth [12]. This recursive construction of FP-trees has been shown to require more time and memory [18]. A new structure is therefore needed to avoid these costly operations and improve processing efficiency.

The proposed MFP-tree structure is a prefix tree with an organized node layout and an address table for frequent items that consists of three parts: the first is a node list for each frequent item, while the other two store item information and the total frequency count for each node in the list. In addition, this structure preserves data relationships across itemsets.

4. COMPOUND CRITIQUING

A Conversational Recommender System (CRS) is an intelligent recommendation approach that enables interactive dialogue between the system and the user [19]. Unlike conventional, one-way recommenders, a CRS leverages multi-turn conversations to gradually elicit user preferences. This interaction allows the system to adapt to explicit user input, producing recommendations that are more personalized and contextually relevant [3, 20]. One widely used interaction mechanism in CRS is critiquing, which enables users to provide explicit feedback on the system's suggestions [20].

Critiquing is an approach where users express feedback or "critiques" about the recommendations [11]. This feedback takes the form of preference statements that constrain or redirect the results. Examples include "I want a cheaper version" or "I want a different author." This technique gives users control and allows the system to iteratively refine its recommendations through adaptive dialogue [21-23].

Critiquing categories in the recommender literature include three main types:

1. Unit Critiquing: The user critiques one attribute at a time, for example, only changing the price or only switching the genre.
2. Dynamic Critiquing: The system automatically selects which attribute to ask about, based on historical preferences and context. The dialogue path is adaptive.
3. Compound Critiquing: The user provides more than one critique or preference in a single statement. For example, "I want a cheaper book and from a different genre."

Compound Critiquing extends critiquing by allowing multiple attribute inputs in a single interaction [11]. This accelerates the refinement process because the system can process several preferences at once. It also improves dialogue efficiency by reducing the number of turns needed to reach suitable recommendations. The study [11] showed that Compound Critiquing leads to higher-quality recommendations and significantly increases user satisfaction compared with single- attribute (unit) critiquing. In practice, the system filters and modifies the initial bundle by evaluating all user inputs in parallel and then recommends a new bundle that better matches the user's needs [21, 22].

5. METHODOLOGY

5.1 System design

This study presents a system workflow that describes the overall operation of the proposed system. Figure 1 shows the workflow of the bundle recommender system.

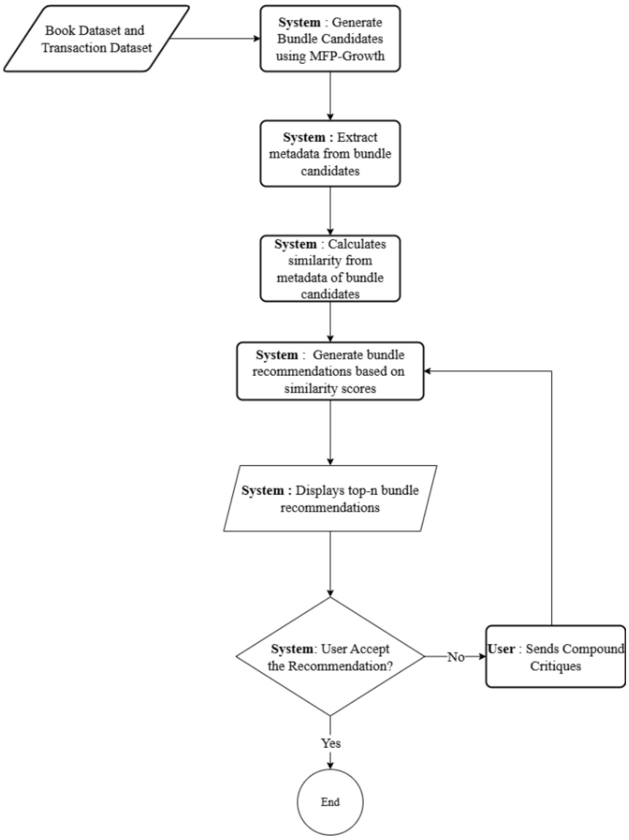

Figure 1. Flowchart

Bundle recommendations are generated from users' explicit preferences by leveraging the MFP-Growth algorithm and the Compound Critiquing technique. The process begins with collecting user transaction data and book metadata, such as ISBN, genre, publication year, price, author, and publisher. Next, MFP-Growth is used to extract association rules among items by considering metrics such as support, confidence, lift, and metadata similarity scores, producing a set of candidate bundles. The system then computes metadata similarity scores for the candidate bundles. These similarity scores serve as the basis for forming the Top N bundle recommendations.

When users receive the initial recommendations, they can provide feedback through Compound Critiquing. Users may submit more specific preferences (in the form of compound critiques), such as wanting cheaper bundles, newer editions, different authors, or specific genres. The system's next task is to interpret these critiques, update the scores and composition of the bundles, and re-filter results based on the critiqued metadata dimensions. This makes the recommendations more dynamic and personalized. New recommendations are then generated and shown again, taking all user critiques into account. This process can be repeated iteratively until the system finds bundles that best match user needs. Finally, the system stores the interaction outcomes as preference references for future recommendation sessions. With this workflow, the system is not only reactive to purchase history but also proactive in understanding and adapting to users' explicit preferences.

5.2 Data description

This study uses a dataset of book purchase transactions from the Amazon e-commerce platform. The dataset was chosen because it captures real user consumption patterns in selecting and buying books and enables analysis of bundle-based purchasing behavior. It serves as the basis for mining association rules, developing the recommendation model, and evaluating the system's effectiveness.

Dataset characteristics:

- Total Transactions: 94,004 user transaction sessions
- Number of Unique Books: 101 titles
- Average Items per Transaction: 1.98
- Metadata: title, author, publisher, publication year, genre, and price

The book transaction dataset used has an average number of items per transaction of 1.98, which reflects realistic purchasing behaviours in the book domain. Even though the number of items per transaction is relatively small, joint purchasing patterns can still be identified through the accumulation of large numbers of transactions, so the Association Rule Mining-based approach remains relevant.

In addition to transaction data, the dataset includes metadata for each book. This metadata not only adds context to the modelling process, but it is also essential for Compound Critiquing. For example, attributes such as genre and price are used when users provide specific preferences like "prefer a different genre" or "prefer a lower price." By combining transaction data with metadata, the system can apply rule-based processing while also considering content similarity, so the resulting bundle recommendations are statistically relevant and aligned with individual user preferences.

5.3 Modified FP-Growth algorithm

The implementation of the MFP-Growth algorithm follows a system design workflow that explains how the algorithm is applied in this study. Figure 2 shows the implementation process of the MFP-Growth algorithm.

The implementation of the MFP-Growth algorithm begins by converting the binary-encoded DataFrame into a list of transactions and applying a minimum support (min_support) threshold for each item, which is the minimum frequency an item must reach to be considered relevant. The system then checks whether the support count of each item meets the min_support value. Items that do not meet the threshold are discarded, while those that do are prioritized according to a chosen order, for example, by highest frequency. Next, an MFP-tree structure is constructed. This is a modified version of the FP-tree that stores transaction data compactly while preserving associations among items. MFP-Growth then extracts purchasing patterns using a bottom-up approach, starting from more specific patterns and moving toward more general ones. In the following stage, the system builds a header table that lists items along with node-link pointers to their corresponding nodes in the MFP-tree, enabling traversal without rebuilding conditional FP-trees. This process yields frequent patterns that can be used as the basis for generating product bundle recommendations.

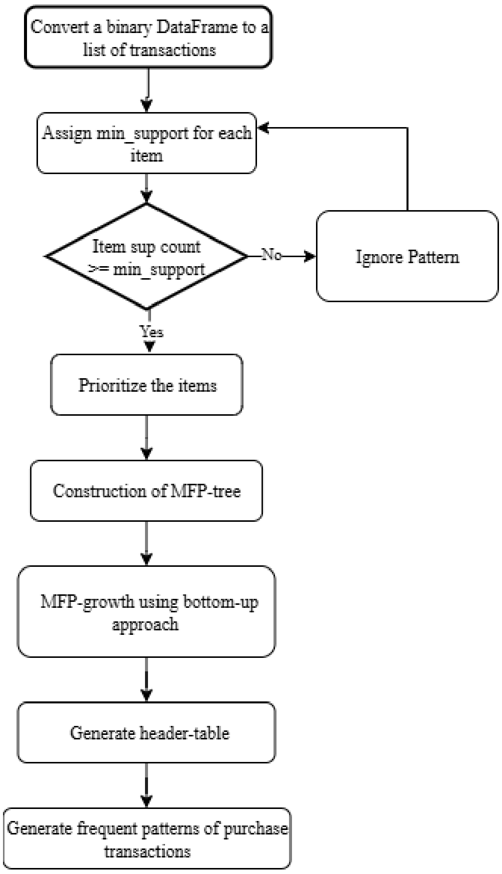

Figure 2. Process Modified FP-Growth algorithm

```
Algorithm MFP-Growth (df, min_support_ratio) ⟶ a table of frequent itemsets
// Mine frequent itemsets with MFP-Growth from a transaction DataFrame.

Input:
df: a transaction DataFrame that can be converted into a list of transactions.
min_support_ratio is a minimum support ratio ($0 < r \leq 1$).
Output:
T is a table (DataFrame) of pairs (itemsets, support).

transactions ← convert_to_transactions(df);
total_rows ← |transactions|
min_support ← int (min_support_ratio × total_rows)
item_support ← get_item_support(transactions)
items ← sort_items_by_support(item_support,
    min_support)
frequent_itemsets ← ∅
for each item ∈ items do
support ← item_support[item]
append(((item), support/total_rows)) to
    frequent_itemsets
projected_db ← build_projected_db (transactions,
    item, min_support, item_support)
recursive_mine([item], projected_db, min_support,
```

```
        total_rows, frequent_itemsets)
12:   end for
13:   return DataFrame(frequent_itemsets, columns =
        ["itemsets", "support"])
```

The implementation of the MFP-Growth algorithm for generating bundle recommendations begins by converting the list of transactions into a DataFrame (df) and computing the number of transactions (total_rows). In line 2, the algorithm converts the minimum support ratio (for example, 0.01) into an absolute minimum transaction count. Next, items are filtered using the criterion support ≥ min_support. Line 6 starts the iteration over each frequent item as the initial prefix (a 1-itemset). For each item, the system retrieves its occurrence count and appends the itemset together with its support ratio to the results list. In line 9, the system builds a projected database that retains only items whose support meets the minimum; within each projected transaction, only the portion after the prefix is kept, and any subsequent duplicate occurrences are ignored to reduce the search space when extending itemsets. The system then performs recursive mining to discover longer itemsets. Finally, it returns the resulting set of itemsets, typically as a table of itemsets with their support values.

In this study, FP-Growth was used as a comparative baseline. FP-Growth was chosen because MFP-Growth is a direct extension of the FP-Growth algorithm, allowing for a fair evaluation of the improvements in efficiency and quality of the proposed association rules.

Other approaches, such as deep learning-based methods, were not used as baselines because they require significantly different data structures, learning models, and computational resources. Furthermore, the primary focus of this study was improving the efficiency and adaptability of the Association Rule Mining-based bundle recommendation system, as well as ease of interpretation of the results, which are not the primary focus of deep learning approaches. Therefore, using FP-Growth as a baseline was deemed appropriate for evaluating the methodological contributions of MFP-Growth in an objective and controlled manner.

5.4 Generating bundle recommendation

This section outlines the workflow used to construct bundle recommendations. Figure 3 shows the bundle construction workflow.

The Generate Bundle Recommender stage is performed after MFP-Growth produces association rules among book items, yielding a set of candidate bundles. Each candidate bundle is then assigned a metadata score based on similarity across attributes such as genre, publication year, author, publisher, and price, ensuring that recommendations are not only statistically relevant but also context aware. The resulting similarity scores are ranked using a combined score that merges association strength with the metadata score, producing the Initial Bundle Recommender as the Top N bundles. This initial set then serves as the basis for the Compound Critiquing stage, where users can provide feedback to iteratively refine the recommendations for greater personalization.

5.5 Compound Critiquing

The developed bundle recommender integrates Compound Critiquing to enable two-way interaction between users and the system. After the initial recommendations are generated using the MFP-Growth algorithm, users can provide explicit feedback on various aspects of a bundle, such as price, publication year, author, publisher, or genre. This mechanism is iterative: each round of feedback filters out irrelevant items, adjusts the recommendations while preserving valid association rules, and produces new alternatives until the user's preferences are satisfied.

The critiquing criteria include:

a. Price, with the cheaper and same price rules, using a ±20% tolerance margin to ensure fair comparisons.
b. Author, with options to retain or switch the author.
c. Publication Year, using weighted median, sigmoid scaling, and variance-based penalties to ensure temporal fit.
d. Publisher, to maintain or broaden publisher variety.
e. Genre, with a 40% overlap threshold to distinguish exploration from continuity.

In this study, 11 critiquing types are available, each reflecting a user preference over a specific attribute:

a. "cheaper" – Want an item at a lower price
b. "same_price" – Want an item at the same price
c. "latest" – Want an item with a more recent publication year
d. "older" – Want an older or classic item
e. "same_genre" – Want an item in the same genre
f. "same_author" – Want an item by the same author
g. "same_publisher" – Want an item by the same publisher
h. "same_year" – Want an item with the same publication year
i. "another_author" – Want a different author
j. "another_publisher" – Want a different publisher
k. "another_genre" – Want a different genre

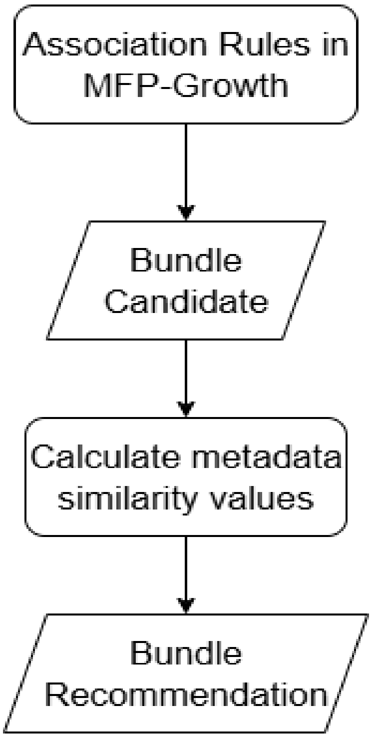

Figure 3. Generating bundle recommender

6. EVALUATION

To obtain a comprehensive evaluation of the performance of the developed bundle recommender system, we created two test scenarios as follows:

1. Measuring the accuracy metrics of MFP-Growth utilization in bundle recommendations by comparing it with the use of FP-growth. The metrics used are Precision@K, Recall@K, F1@K, MAP@K, and nDCG@K [22].
2. Measuring the diversity, novelty, and coverage of the recommendation results. Diversity is used to assess the diversity of the resulting bundles. Novelty is used to measure the novelty of the recommended bundle results. Coverage measures the extent to which the system covers the entire available item space.
3. Measuring the efficiency of time and memory usage in utilizing the MFP-Growth algorithm with the FP-Growth algorithm as a comparison.
4. Measuring the variation of bundle recommendations based on user preferences for items in the bundle.

6.1 Evaluation scenario 1

This scenario is used to assess the accuracy of bundle recommendations. The FP-Growth algorithm was used as a benchmark because it is a commonly used algorithm for generating association rules. The accuracy test results for both algorithms are shown in Table 1.

Table 1. Test results for scenario 1

Metric	FP-Growth	MFP-Growth
Precision@10	0.363199	0.363199
Recall@10	0.185305	0.185305
F1-Score@10	0.242275	0.242275
nDCG@10	0.327025	0.327216
MAP@10	0.181436	0.181556

The evaluation results in Table 1 show that the Precision@10, Recall@10, and F1-Score@10 values generated by FP-Growth and MFP-Growth are identical. This indicates that the modifications made to MFP-Growth do not change the number of relevant items recommended but rather affect the order in which items are placed in the recommendation list.

Performance differences are visible on rank-based metrics. The nDCG@10 value increased from 0.327025 to 0.327216, and the MAP@10 value increased from 0.181436 to 0.181556. Although these increases are relatively small numerically, the MAP and nDCG metrics are sensitive to changes in the order of items in the top positions. Therefore, these improvements indicate that MFP-Growth is more consistent in ranking relevant items than FP-Growth. These improvements are important in the context of bundled recommendation systems, as the quality of the user experience is significantly influenced by the order in which recommendations are initially displayed.

6.2 Evaluation scenario 2

To understand the breadth and variety of recommendations provided by the system, additional metrics focused on the exploration aspect were used. The test results are shown in Table 2.

The results in Table 2 show that the system produces a very high diversity value (±0.95), which indicates that the items in the recommended bundle have a good level of diversity. A novelty value above 5.7 indicates that the system is capable of recommending items that are not only popular but also encourages exploration of less frequently appearing book titles. However, the relatively low coverage value (±3.5%) indicates that the rule-based system tends to recommend only a subset of items with strong associations. This is a common limitation of the Association Rule Mining approach, where low-frequency items tend not to appear in recommendations.

Table 2. Test results for scenario 2

Metric	FP-Growth	MFP-Growth
Total Bundles Evaluated	28.43	28.43
Avg. Diversity	0.94	0.95
Avg. Novelty	5.76	5.75
Coverage	0.03	0.03

6.3 Evaluation scenario 3

Running time is an important metric in measuring the performance of an algorithm. Table 3 presents the running time and memory usage between FP-Growth and MFP-Growth.

Table 3. Time and memory efficiency test results

Algorithm	Processing Time (s)	Memory Usage (MB)
FP-Growth	41,897.36	344.69
MFP-Growth	151.50	37.22

The results show that MFP-Growth requires significantly less processing time, only about 151 seconds, compared to FP-Growth, which requires over 41,000 seconds (equivalent to over 11 hours). This represents a more than 270-fold increase in time efficiency. Similarly, in terms of memory, MFP-Growth consumes only about 37 MB, or about 89% less than FP-Growth, which uses 344 MB.

These findings confirm that although MFP-Growth produces identical rule output, its internal structure and approach significantly improve processing efficiency, making it more ideal for use in real-time or large-scale systems.

6.4 Evaluation scenario 4

In this final scenario, users can provide feedback on the bundle recommendations generated by the system. In Compound Critiquing, there are several feedback options (critical words) that users can choose from regarding book bundle recommendations. Table 4 displays the results of the critique testing.

Table 4 shows the effectiveness of user feedback on bundle recommendations. The metrics used are confidence, lift, similarity, diversity, novelty, bundle size, and total price. Confidence and Lift values indicate the relatedness between items purchased by users. Diversity indicates the variety of each recommended bundle. Novelty represents the novelty of the recommended bundle. Overall, compound critique has proven capable of accommodating multi-attribute preferences with increased diversity and novelty in recommendations but requires consideration of the impact of reduced association strength and metadata suitability, particularly for combinations that encourage extreme exploration.

Table 4. Test results testing for each critique

Compound Critique	Confidence	Lift	Similarity Score	Diversity	Novelty	Bundle Size (Book)	Total Bundle Price
Cheaper	-0.4787	-0.1705	-0.5542	+0.3218	+0.9532	0	-Rp58,500
Same Price	-0.3788	+0.0415	-0.3653	+0.2336	+0.0872	+1	+Rp33,500
Latest	-0.4787	-0.1705	-0.5542	+0.3218	+0.9532	0	-Rp58,500
Older	-0.3974	+0.3267	-0.3351	+0.1910	-0.0588	+2	+Rp352,500
Same Genre	-0.3974	+0.3267	-0.3351	+0.1910	-0.0588	+2	+Rp352,500
Same Author	-0.2428	+0.2552	-0.1341	+0.0370	+0.0325	+1	+Rp175,500
Same Publisher	-0.2428	+0.2552	-0.1341	+0.0370	+0.0325	+1	+Rp175,500
Same Years	-0.2428	+0.2552	-0.1341	+0.0370	+0.0325	+1	+Rp175,500
Another Author	-0.2828	-0.2229	-0.4592	+0.3343	-0.1760	+1	+Rp142,500
Another Publisher	-0.0474	+0.1045	-0.1906	+0.0188	+0.1139	0	+Rp51,500
Another Genre	-0.4634	-0.2749	-0.5414	+0.4024	+0.8091	0	-Rp89,000
Cheaper & Another Author	-0.4787	-0.1705	-0.5542	+0.3218	+0.9532	0	-Rp58,500
Cheaper & Another Genre	-0.4634	-0.2749	-0.5414	+0.4024	+0.8091	0	-Rp89,000
Same Price & Same Genre	-0.3788	+0.0415	-0.3653	+0.2336	+0.0872	+1	+Rp33,500
Same Price & Another Author	-0.3888	-0.2418	-0.5221	+0.3320	+0.5551	0	+Rp22,000
Latest & Same Genre	-0.4787	-0.1705	-0.5542	+0.3218	+0.9532	0	-Rp58,500
Older & Cheaper	-0.3928	-0.2524	-0.5430	+0.3330	+0.5599	0	-Rp92,000
Same Genre & Same Author	-0.2428	+0.2552	-0.1341	+0.0370	+0.0325	+1	+Rp175,500
Same Genre & Another Publisher	-0.0474	+0.1045	-0.1906	+0.0188	+0.1139	0	+Rp51,500
Same Year & Same Price	-0.3888	-0.2418	-0.5221	+0.3320	+0.5551	0	+Rp22,000

The Compound Critiquing testing, which combines the dimensions of price, genre, author, publisher, and year, has proven effective in creating adaptive and rich interactions. Exploration-based critiquing (Cheaper, Cheaper & Another Genre, Latest) increases diversity by >40% and novelty significantly but is accompanied by a decrease in confidence and similarity due to a shift to less historically related items. In contrast, thematic relevance-based critiquing (Same Author, Same Publisher, Same Year) increases lift by ±25–32% and maintains high similarity, making it ideal for retention and satisfaction of loyal users. The Same Price option offers a compromise that adds variety without significantly sacrificing relevance.

6.5 Interpretation of results

The research results consistently show that the hybrid approach combining MFP-Growth and Compound Critiquing provides a significant contribution to improving the quality of recommender systems, especially in the aspects of personalization and dynamic user interaction. In terms of efficiency, MFP-Growth has been shown to drastically reduce execution time from ±41,897 seconds to ±151.5 seconds and reduce memory usage from ±344.7 MB to ±37.2 MB, without sacrificing the number or quality of generated association rules. This advantage is achieved through an improved MFP-tree structure with header tables and node-links, which allows traversal without constructing a new conditional FP-tree, in line with findings from previous research.

In terms of recommendation quality, MFP-Growth provides stability in metrics such as Precision@10, Recall@10, and F1@10, with slight improvements in MAP@10 and nDCG@10, as well as high diversity (±0.95) and novelty above 5.7. The Compound Critiquing mechanism, which combines the dimensions of price, genre, author, publisher, and year, has proven effective in creating adaptive and rich interactions. Exploration-based critiquing (Cheaper, Cheaper & Another Genre, Latest) increases diversity by >40% and novelty significantly but is accompanied by a decrease in confidence and similarity due to a shift to less historically related items. In contrast, thematic relevance- based critiquing (Same Author, Same Publisher, Same Year) increases lift by ±25–32% and maintains high similarity, making it ideal for retention and satisfaction of loyal users. The Same Price option offers a compromise that adds variety without significantly sacrificing relevance.

Compared to traditional association methods or deep learning approaches, the combination of MFP-Growth and Compound Critiquing offers comparable, and even superior, performance on some metrics, with lower computational requirements and better interpretability. This not only strengthens the system's relevance and flexibility but also ensures real-time adaptability to changing user preferences, making it an effective solution for real-world applications such as bookselling and e-commerce platforms.

Although the improvements in MAP@10 and nDCG@10 values are relatively small, these metrics are highly sensitive to changes in the order of top-ranked items. Thus, the increase in MAP@10 from 0.181436 to 0.181556 and nDCG@10 from 0.327025 to 0.327216 indicates improved consistency in placing relevant items at the top of recommendations.

In addition to accuracy, the proposed system demonstrates significant improvements in computational efficiency, reflected in lower execution time and memory usage compared to FP-Growth. In terms of recommendation quality, the improvements in diversity and novelty values indicate that the system can reduce the tendency for monotonous recommendations and expand product exploration. These results confirm that the proposed approach offers practical benefits in the context of bundled recommendation systems.

Overall, the experimental results demonstrate that the hybrid MFP-Growth and Compound Critiquing approach improves the quality of recommendation systems, particularly in terms of efficiency, diversity, novelty, and interaction adaptability. While the numerical accuracy gains are relatively small, the system's primary benefit lies in improving the user's experience through more varied and dynamically adaptable recommendations.

This approach emphasizes that the evaluation of a bundled recommendation system depends not solely on improvements in accuracy metrics, but also on the system's ability to support product exploration and adaptive interactions.

7. CONCLUSIONS

This study proposes a bundle recommendation system based on a hybrid approach by integrating MFP-Growth and Compound Critiquing to improve the efficiency, personalization, and adaptability of the recommendation system in the book e-commerce domain.

Experimental results show that MFP-Growth significantly improves computational efficiency compared to conventional FP-Growth. The improved MFP-tree structure, equipped with header tables and node links, allows for traversal without constructing a new conditional FP-tree, reducing execution time from ±41.897 seconds to ±151.50 seconds and memory usage from ±344.7 MB to ±37.2 MB. These findings support the first hypothesis that MFP-Growth can accelerate execution and save computational resources without compromising the quality of the resulting association rules.

In terms of recommendation quality, the integration of MFP-Growth with Compound Critiquing produces more diverse and contextual book bundles than systems relying solely on traditional association rules. Exploration-based critiquing (Cheaper, Cheaper & Another Genre, Latest) increased diversity by over 40% and novelty significantly, although it was accompanied by a decrease in confidence and similarity scores. Meanwhile, thematic relevance-based critiquing (Same Author, Same Publisher, Same Year) increased lift by approximately 25–32% and maintained high similarity, making it suitable for retaining loyal users. The Same Price option is a compromise solution that provides variety without significantly sacrificing relevance.

Overall, the evaluation results prove the second hypothesis: the combination of MFP-Growth and Compound Critiquing can create an effective direct interaction mechanism between users and the system, allowing recommendations to be updated adaptively and in real time. This approach offers high recommendation quality, computational efficiency, and better interpretability than traditional methods and deep learning-based approaches, making it suitable for application to real-world book e-commerce platforms. Even small increases may indicate more consistent placement of relevant items at higher ranks, which is critical in practical e-commerce scenarios where users typically focus on the first few recommended bundles. The improvements observed, therefore, reflect a meaningful improvement in recommendation quality rather than random fluctuations. In addition to ranking performance, the proposed approach demonstrates improved computational efficiency and higher diversity in the generated bundles compared to the FP-Growth baseline, contributing to a more flexible and adaptive recommendation process. Due to the computational cost of repeatedly executing FP-Growth in an end-to-end experimental setting, formal statistical significance testing could not be fully conducted within the revision timeframe. To avoid overstating the empirical findings, this limitation is explicitly acknowledged, and comprehensive statistical validation using repeated runs and non-parametric tests such as the Wilcoxon signed-rank test is identified as an important direction for future work.

The Compound Critiquing mechanism allows the system to adaptively tailor recommendations based on the user's multi-attribute preferences. Exploration-based critiquing increases the diversity and novelty of recommendations, while thematic relevance-based critiquing maintains the strength of associations and metadata relevance.

While the results are promising, this study has several limitations. The dataset used is relatively small and limited to a single domain (books), so the results cannot be directly generalized to other e-commerce domains with different transaction characteristics. Furthermore, the rule-based system has limited coverage because it only recommends items with strong associations.

As further research, the system can be developed using larger-scale datasets, conducting cross-domain validation, and integrating more contextual and complex critiquing mechanisms to improve the system's generalization and flexibility.

REFERENCES

[1] Deo, A., Jaisinghani, R., Gupta, S., Khan, S.S., Soni, A., Gehlot, K. (2023). Stratified advance personalized recommendation system based on deep learning. Ingénierie des Systèmes d’Information, 28(1): 189-196. https://doi.org/10.18280/isi.280121

[2] He, Z., Zhao, H., Yu, T., Kim, S., Du, F., McAuley, J. (2022). Bundle MCR: Towards conversational bundle recommendation. In Proceedings of the 16th ACM Conference on Recommender Systems, Seattle, WA, USA, pp. 288-298. https://doi.org/10.1145/3523227.3546755

[3] Zhu, T., Harrington, P., Li, J., Tang, L. (2014). Bundle recommendation in ecommerce. In SIGIR '14: Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval, Gold Coast, Queensland, Australia, pp. 657-666. https://doi.org/10.1145/2600428.2609603

[4] Deng, Q., Wang, K., Zhao, M., Zou, Z., Wu, R., Tao, J., Fan, C., Chen, L. (2020). Personalized bundle recommendation in online games. In CIKM '20: Proceedings of the 29th ACM International Conference on Information & Knowledge Management, pp. 2381-2388. https://doi.org/10.1145/3340531.3412734

[5] Wei, Y., Liu, X., Ma, Y., Wang, X., Nie, L., Chua, T.S. (2023). Strategy-aware bundle recommender system. In Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, Taipei, Taiwan, pp. 1198-1207. https://doi.org/10.1145/3539618.3591771

[6] Bai, J., Zhou, C., Song, J., Qu, X., An, W., Li, Z., Gao, J. (2019). Personalized bundle list recommendation. In WWW '19: The World Wide Web Conference, San Francisco, CA, USA, pp. 60-71. https://doi.org/10.1145/3308558.3313568

[7] Du, X., Qian, K., Ma, Y., Xiang, X. (2025). Enhancing item-level bundle representation for bundle recommendation. ACM Transactions on Recommender Systems, 3(3): 1-21. https://doi.org/10.1145/3637067

[8] Shalaby, A. hamada, el letif, D.M., Rezk, D., Hagag, P. (2015). Suggested marketing strategy using apriori and FP-growth algorithms in retail sales in Egypt.

International Journal of Computers & Technology, 14(11): 6190-6200. https://doi.org/10.24297/ijct.v14i11.1807

[9] Sun, Z., Feng, K., Yang, J., Fang, H., Qu, X., Ong, Y.S., Liu, W. (2024). Revisiting bundle recommendation for intent-aware product bundling. ACM Transactions on Recommender Systems, 2(3): 1-34. https://doi.org/10.1145/3652865

[10] Fatiyah, A.C., Baizal, Z.K.A., Wibowo, A.T. (2022). Compound critiquing for improving query refinement on conversational recommender system. In 2022 10th International Conference on Information and Communication Technology (ICoICT), Bandung, Indonesia, pp. 340-345. https://doi.org/10.1109/ICoICT55009.2022.9914858

[11] Shawkat, M., Badawi, M., El-Ghamrawy, S., Arnous, R., El-Desoky, A. (2022). An optimized FP-growth algorithm for discovery of association rules. The Journal of Supercomputing, 78(4): 5479-5506. https://doi.org/10.1007/s11227-021-04066-y

[12] Uzun-Per, M., Gürel, A.V., Can, A.B., Aktas, M.S. (2021). An approach to recommendation systems using scalable association mining algorithms on big data processing platforms: A case study in airline industry. In 2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA), Kocaeli, Turkey, pp. 1-6. https://doi.org/10.1109/INISTA52262.2021.9548413

[13] Padhy, N., Suman, S., Priyadarshini, T.S., Mallick, S. (2024). A recommendation system for e-commerce products using collaborative filtering approaches. Engineering Proceedings, 67(1): 50. https://doi.org/10.3390/engproc2024067050

[14] Bunga, R., Batista, F., Ribeiro, R. (2021). From implicit preferences to ratings: Video games recommendation based on collaborative filtering. In Proceedings of the 13th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, pp. 209-216. https://doi.org/10.5220/0010655900003064

[15] Fitrianah, D., Zain, S.Y. (2021). Analysis of consumer purchase patterns on handphone accessories sales using fp-growth algorithm. In International Conference on Engineering, Technology and Social Science (ICONETOS 2020), pp. 149-158. https://doi.org/10.2991/assehr.k.210421.022

[16] Putri, I.D., Hardianto, R. (2024). Application of the FP-growth algorithm in consumer purchasing pattern analysis. Journal of Computer Scine and Information Technology, 10(2): 44-49. https://doi.org/10.35134/jcsitech.v10i2.99

[17] Irawan, K.M., Wulansari, T.T., Sari, N.W.W. (2021). Market basket analysis method on sales data using FP-growth algorithm. Multica Science and Technology, 1(2): 55-60. https://doi.org/10.47002/mst.v1i2.239

[18] Baizal, Z.K.A., Widyantoro, D.H., Maulidevi, N.U. (2020). Computational model for generating interactions in conversational recommender system based on product functional requirements. Data & Knowledge Engineering, 128: 101813. https://doi.org/10.1016/j.datak.2020.101813

[19] Jannach, D., Manzoor, A., Cai, W., Chen, L. (2021). A survey on conversational recommender systems. ACM Computing Surveys (CSUR), 54(5): 1-36. https://doi.org/10.1145/3453154

[20] Sun, Y., Zhang, Y. (2018). Conversational recommender system. In SIGIR '18: The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, Ann Arbor, MI, USA, pp. 235-244. https://doi.org/10.1145/3209978.3210002

[21] Jannach, D. (2023). Evaluating conversational recommender systems: A landscape of research. Artificial Intelligence Review, 56(3): 2365-2400. https://doi.org/10.1007/s10462-022-10229-x

[22] Abiodun, A.G., Ukandu, O.K., Emmanuel, A.S., Udechukwu, C.S., Oladapo, M.O., Nadasan, T., Bakare, O. (2025). Detection of heart disease using binary classification machine learning model. Ingenierie des Systemes d'Information, 30(5): 1111-1122. https://doi.org/10.18280/isi.300501

[23] Chen, L., Pu, P. (2012). Critiquing-based recommenders: Survey and emerging trends. User Modeling and User-Adapted Interaction, 22(1): 125-150. https://doi.org/10.1007/s11257-011-9108-6