UDC 621.391.83(045)
DOI:10.18372/1990-5548.88.20958

[1]Olena Arsirii,
[2]Dmitriy Ivanov

DATA MINING OF B2B TRANSACTIONS USING A HIGH-UTILITY TEMPORAL APPROACH

Odesa Polytechnic National University, Odesa, Ukraine
E-mails: [1]e.arsiriy@gmail.com ORCID 0000−0001−8130−9613
[2]ivanovdima9988@gmail.com ORCID 0009-0009-3958-5310

Abstract—*Modern B2B e-commerce systems generate vast arrays of transactional data, yet classical association rule mining methods face several critical limitations, notably the "financial blindness" of binary models and the "temporal blindness" of static approaches. These constraints hinder the discovery of patterns with genuine economic value and overlook the cyclical nature of wholesale procurement.Methods. This paper proposes a high-utility temporal approach to B2B transaction analysis. The approach is centered on the UP-Growth (Utility Pattern Growth) mathematical model, adapted to operate with a utility function based on the rational transaction context, defined as the product of individual price and wholesale purchase volume. To address the anti-monotonicity problem, the concept of transaction-weighted utility is introduced. The integration of the temporal factor is achieved through the analysis of inter-purchase interval vectors, utilizing the coefficient of variation as a criterion for cycle stability. As a result, a holistic dual-stage pipeline architecture for the intelligent analysis of B2B customer transactional data has been developed. In the first stage, the pipeline performs the extraction of High-Utility Itemsets, while the second stage involves calculating the trigger date to establish a proactive offer window. A comparative analysis with the classical FP-Growth algorithm confirmed the superior efficiency of the proposed approach in identifying economically significant patterns, validated using real-world data from the "Baza Vzuttya"*

Keywords—B2B E-commerce, data mining, high-utility itemset mining, predictive analytics, customer behavior analysis, temporal data mining, personalized content, recommendation systems.

I. INTRODUCTION

The development of e-commerce and the digitalization of B2B processes are accompanied by the rapid accumulation of customer transaction data. Due to intensifying competition in this market segment, classical B2C sales management models applied to B2B platforms are losing their effectiveness. Consequently, companies are increasingly shifting toward proactive engagement with partners through the creation of personalized content. The global B2B e-commerce market, valued at $11.54 trillion in 2024 and projected to reach $60.62 trillion by 2034, significantly exceeds the B2C segment [1] – [3]. At such scales, even a slight increase in personalization accuracy through intelligent data analysis generates a substantial economic effect.

Traditional approaches to personalization, borrowed from the B2C segment, are primarily based on classical association rule mining methods such as Apriori, FP-Growth, or Eclat [4], [5]. These methods focus on identifying frequency-based relationships in arrays of homogeneous transactions, which is characteristic of emotional and impulsive purchases by individual consumers. However, the specifics of B2B procurement require a fundamental revision of these models. Unlike B2C, a business customer is driven by rational motives, the need for return on investment (ROI), and complex internal decision-making procedures (Buying Center) [6].

An analysis of existing solutions has revealed two critical problems that remain overlooked by classical algorithms. First is the "financial blindness" of binary models: when association mining methods consider only purchase frequency, they ignore the actual value and volume of transactions, which are decisive factors in the B2B segment. The sale of a hundred small accessories cannot be equated to the purchase of a large batch of high-value goods (e.g., full size-runs of shoes), as their value to the business is incomparable. Second is the "temporal blindness" of static models, which record the existence of a link between item transactions but fail to account for the cyclical nature and time intervals between purchases. In a B2B environment, where orders are seasonal and planned, it is crucial to understand not only "what" the customer bought but also "when" the need for the next delivery will arise.

Thus, there is a scientific and practical task to develop models and methods for the intelligent analysis of B2B customer transactional data that integrate financial parameters (price and volume) with the temporal dynamics of this data. Such integration forms the basis of the developed unified high-utility temporal approach. The application of this approach will allow for the transformation of raw transactional data into proactive personalized content based on the calculation of precise Restock Prediction dates and the formation of individual commercial offers with high economic significance.

II. Literature Review and Problem Statement

As stated in [7], a B2B e-commerce system is a comprehensive online platform or a set of interconnected technological solutions designed for commercial transactions, information exchange, and business process automation between two or more legal entities (enterprises). It is well established that B2B (Business-to-Business) transactions involve the exchange of goods, services, or information between business clients, and their effectiveness largely depends on the degree of content personalization on B2B platforms [8], [9].

In articles [10], [11], it is noted that increasing content personalization in B2B e-commerce systems requires a transition from emotion-oriented B2C methods to rationally justified solutions that directly impact the efficiency, profitability, and operational costs of the client. Intelligent analysis of B2B platform customer transactional data serves as a key tool that allows not only for recording the fact of purchase but also for identifying the logic and sequence of procurement decisions.

According to [12] – [14], content creation technologies for B2B e-commerce involve the generation of several content types:

- product content (descriptions, related products, commercial offers);
- informational content (analytical reports, advice, documentation);
- recommender content (personalized selections of products and offers).

Regarding the capabilities of association rule mining models for B2B customer transactional data (invoices, wholesale orders), [12] points out that tools traditionally focused on frequent itemset mining, such as Apriori, FP-Growth, and Eclat [15], are commonly used for product content generation. In a B2B system, these models must be adapted to form *product content* in the form of *comprehensive offers* by integrating financial and quantitative metrics [12]. An example is the *formation of "Wholesale Bundles"*, where transactional analysis identifies association rules dominated by the *total value* or *margin* of the set instead of frequency (Support):

```
If {Product A, Product B}→{Product C}
```

This allows the system to offer not just related items, but comprehensive procurement solutions that, from the client's perspective, optimize logistics and reduce overall costs (e.g., suggesting specialized fasteners and tools when ordering a large batch of construction materials). Another example of product content creation is *price offer personalization*, where association rules can be applied to determine individual terms:

```
If {Client in "Premium", Orders Product A monthly}→{10% Discount on Product B}
```

The system generates not merely a general promotion, but a personalized investment plan for the client, demonstrating how the inclusion of associated goods allows for more favorable wholesale pricing.

Studies [12], [13] also explore the potential of using sequential patterns for creating B2B *recommender content*. Sequential Pattern Mining (SPM) works with transactional-behavioral and temporal data, identifying cycles and event sequences. This enables the prediction of the B2B client's next logical step, which serves as the basis for recommender content [12]. For instance, Restock Prediction: SPM identifies the typical time interval between orders of specific product sets.

```
If {Winter Footwear Collection Order} →{Leather Care Products Order (after 6 weeks)}
```

In this scenario, the system generates a recommendation notification one week before the predicted date, reminding the client of the need to replenish stock. This increases loyalty as the system acts as an assistant in the client's inventory management.

Another example of *recommendations* involves behavioral scenarios of the Buying Center. Sequential analysis is applied to the actions of various employees of the purchasing company:

```
If {Engineer reviews technical specifications}→{Financier requests final price}
```

In other words, once an engineer has reviewed the technical specifications, the system automatically generates a recommendation notification for the financier in the form of an ROI calculation or a specialized price list for that product, thereby accelerating the sales cycle.

In article [14], the use of transactional-behavioral data for creating *informational content* in B2B systems is discussed. Informational content in B2B includes not only technical specifications but also terms of service, reports, and analytics [14]. Thus, association rule mining of transactional data is utilized to provide personalized analytics. For instance, the system compares a customer's purchasing behavior with typical patterns within their industry or segment:

```
If {Customer does not buy Product B, but competitors buy Product B with A} →{Send analytical report on Product B and its benefits}
```

Such automatically generated informational content is rationally justified, as it appeals to missed opportunities or the best practices of competitors. Furthermore, sequential pattern analysis can identify behavioral anomalies that precede customer churn, which is vital for risk management:

```
If {Sharp volume decrease, Downloading of old invoices}→{Search for new suppliers (Forecast)}
```

This triggers personalized communication aimed at customer retention, such as offering special, individual cooperation terms.

Thus, intelligent analysis adapted to B2B metrics, such as value, volume, time, and behavioral scenarios, is essential for transforming a B2B e-commerce system from a simple catalog into an intelligent assistant that provides rationally justified, proactive, and personalized content.

However, [16] emphasizes that the vast majority of B2C e-commerce platforms utilize static content generation methods based on fixed rules or rely on universal recommendations for individual customers. Such an approach fails to account for procurement patterns, the specific business needs of wholesale clients, seasonality, or inter-item dependencies. This limits service efficiency and reduces the likelihood of repeat purchases. Addressing these issues requires a systematic analysis of existing association rule mining models and methods to modify them for enhancing the personalization of product, informational, and recommender content within B2B e-commerce systems.

An analysis of modern association rule mining models indicates that most existing approaches [17] – [19] are oriented toward identifying frequently recurring item combinations without considering their quantitative and monetary significance. This is a critical drawback for B2B systems, as these methods are based on the concept of "frequent itemsets," whereas for business transactions, the decisive factors are financial value and volume.

The fundamental toolkit for association identification is based on three classical approaches:

- The *Apriori algorithm* [17], which uses iterative candidate generation, leading to significant computational costs in the context of large B2B datasets.
- The *FP-Growth algorithm* [18], based on the construction of a prefix tree (FP-Tree). While it effectively solves the candidate generation problem, in its classical form, it remains "binary", meaning it only records the presence of an item, ignoring its individual price (p) and quantity (q).
- The *Eclat algorithm* [19], which utilizes vertical data representation but also has limitations regarding the inclusion of continuous numerical transaction parameters.

Studies [17] – [19] confirm the existence of "*financial blindness*" in these methods: they are unable to differentiate between a transaction involving low-value related goods and a high-value strategic procurement. Furthermore, classical approaches demonstrate "*temporal blindness*", as they treat the transactional database as a static array, ignoring time intervals between purchases and seasonal cyclicity, which form the basis of rational B2B customer behavior.

Research dedicated to the development of High-Utility Itemset Mining (HUIM) methods, specifically the UP-Growth algorithm, proposes solutions for incorporating item utility [20]. However, the question of integrating value-based indicators with temporal characteristics (cyclicity) for proactive content personalization remains insufficiently explored [21]. It is precisely the need to bridge the gap between frequency-based analysis and value-temporal forecasting that determines the relevance of this research.

III. MATHEMATICAL FOUNDATIONS OF THE HIGH-UTILITY TEMPORAL APPROACH

The study is based on the mathematical framework of Data Mining, specifically the concepts of Association Rule Mining and Sequential Pattern Mining.

Formally, the problem is defined over a set of transactions $D = \{T_1, T_2, ..., T_m\}$, where each transaction T consists of a set of items (SKUs) $I = \{i_1, i_2, ..., i_n\}$. The traditional approach, implemented in the *FP-Growth* algorithm, relies on a binary model $i \in \{0,\ 1\}$, which establishes the presence of an item and involves the calculation of probabilistic characteristics such as *Support* and *Confidence*. However, for the B2B segment, this approach is limited due to "financial blindness" regarding the significance of operations.

This paper proposes an extension of the mathematical framework through a transition to High-Utility Itemset Mining (HUIM). This allows for replacing the simple frequency count $count(i_k)$ with a utility function $U(i_k, T_j)$, which, within the value-weighted model, is defined as: $U(i_k, T_j) = p(i_k) \times q(i_k, T_j)$ where $p(i_k)$ is the individual unit price for a specific customer, and $q(i_k, T_j)$ is the quantity of the item in the transaction. The utility of an itemset X in transaction T_j is defined as the sum of the utilities of all its elements: $U(X, T_j) = \sum_{i_k \in X} U(i_k, T_j)$.

To incorporate the temporal aspect, the value-weighted mathematical model is augmented with a transaction timestamp $t(T_j)$. This enables the formation of an inter-purchase interval vector $\Delta t = \{t_2 - t_1, t_3 - t_2, ..., t_n - t_{n-1}\}$ for each high-utility pattern. The temporal stability of a pattern is determined by the coefficient of variation $CV(\Delta t)$, which provides a mathematical basis for calculating the trigger date for proactive content generation.

Thus, the mathematical foundation of the proposed approach combines the identification of sets high-value items (High-Utility Itemsets) using the UP-Growth method with the prediction of temporal windows for their actualization, ensuring rationally justified personalization of B2B content.

IV. The Aim and Objectives of the Research

The objective of this study is to develop a high-utility temporal approach to enhance the efficiency of intelligent B2B customer transactional data analysis within e-commerce systems. To achieve this objective, the following research tasks have been defined:

- to conduct a systematic analysis of existing association rule mining methods within the context of B2B e-commerce transactional datasets;
- to develop a mathematical model for value-weighted transactions and a high-utility itemset mining method based on the UP-Growth (Utility Pattern Growth) algorithm for the utility-based analysis of B2B transactions;
- to develop a mathematical model for sequences of value-weighted B2B transactions and a method for their temporal verification through the analysis of inter-purchase interval vectors, utilizing the coefficient of variation as a criterion for cycle stability;
- to perform an analysis of the practical results regarding the developed models and methods for intelligent B2B transaction analysis, which form the basis of the proposed high-utility temporal approach.

V. The Research Materials and Methods

A. Systematic analysis of existing association rule mining methods within the context of B2B e-commerce transactional datasets

The effectiveness of proactive sales support systems in the B2B segment is directly determined by the ability of analytical methods to transform accumulated transactional data into actionable knowledge. The purpose of conducting a systematic analysis of existing analytical methods is to identify their *fundamental methodological limitations* regarding association rule mining, which preclude their direct application in building proactive recommender systems within a B2B environment.

Association rule mining methods were originally developed within the context of B2C retail tasks. Their analytical framework is built upon the concept of a transaction as a binary market basket – the fact of the presence or absence of an item. The problem is formally described as follows: given a set of item positions $I = \{i_1, i_2, ..., i_n\}$ and a transaction database $D = \{T_1, T_2, ..., T_m\}$, where each transaction $T_j \subseteq I$. It is necessary to find association rules of the form $X \to Y$, where $X \subset I, Y \subset I$, and $X \cap Y = \varnothing$. The key metric around which the entire apparatus of classical methods [17] – [19] is structured is *Frequency Support*. For an itemset $X \subseteq I$, it is defined as the proportion of transactions containing this set in the total number of transactions in the database D:

$$Supp(X) = \frac{\left|\{T \in D | X \subseteq T\}\right|}{|D|}, \tag{1}$$

where X is the investigated association itemset; D is the enterprise's full transaction database for the analyzed period; $|D|$ is the total number of transactions in the database; T is an individual database element (transaction, invoice, or waybill); $\left|\{T \in D | X \subseteq T\}\right|$ is the subset of transactions in which the set X is present as a whole.

It follows from (1) that every transaction carries the same algorithmic weight, regardless of its financial capacity. This approach is acceptable in retail – the average consumer's receipt does not differ significantly, and purchases are made chaotically and frequently. However, transferring this logic to a B2B transaction array reveals four critical gaps (limitations) between the mathematical assumptions of the method and the nature of the subject domain.

The first limitation is the "financial blindness" of binary models. A transaction in the B2B segment is not a simple list of goods but a multi-parameter function:

$$T = f(i_k, q_k, p_k), \tag{2}$$

where the transaction T reflects the relationship between the item identifier (i_k), its quantity (q_k), and the negotiated price (p_k).

When a transaction is converted into a binary vector to meet the requirements of classical methods, the parameters q_k and p_k are discarded. Consequently, the very factors that determine the actual financial utility of the operation for the enterprise are lost.

The variance of transaction amounts in B2B is extremely high, meaning that the purchase of small consumables and the acquisition of a large batch of strategic goods are algorithmically perceived by the classical method as equivalent events. As a result, the system identifies patterns based on the criterion of frequency rather than economic significance. This leads to the phenomenon of "financial blindness", which systematically undervalues rare but strategically important wholesale purchases. To illustrate the consequences of such an approach, let us compare two types of procurement patterns (Table I). As seen from the data in Table I, the classical method recognizes Pattern A as *significant* due to its high frequency, whereas Pattern B – which generates 14 times more revenue – is rejected as statistically insignificant *noise*.

TABLE I. COMPARATIVE EVALUATION OF PATTERNS UNDER FREQUENCY-BASED AND UTILITY-BASED APPROACHES

Comparison Criteria	Pattern A (Consumables)	Pattern B (Technological equipment)
Itemset composition	{Protective gloves, Cleaner}	{Diesel generator, Lubricant}
Number of occurrences	148	6
Average set value, c.u.	180	62 000
Frequency support Supp, %	29,6 % (Significant)	1.2 % (Noise)
Total financial result, UAH.	1 118 880	15 624 000

The second limitation is the "combinatorial explosion" during weighted analysis. An attempt to directly replace frequency support with financial weighting encounters a computational obstacle. In the B2B segment, so-called "whale-receipts" are common – large-scale orders containing dozens of different items. If a single such receipt independently exceeds the minimum financial support threshold, the method is obliged to consider all possible subsets of items within that transaction. The number of such subsets N is defined as:

$$C(k) = 2^k - 1, \quad (3)$$

where k is the number of unique items (SKUs) within a single transaction; C is the total number of subsets to be generated and verified.

A comparison of values (3) for typical B2C and B2B transaction lengths is presented in Table II.

TABLE II. DEPENDENCE OF COMPUTATIONAL COMPLEXITY ON B2C AND B2B TRANSACTION LENGTH

Number of items per transaction k	Number of combinations $2^k - 1$	Typical segment	Computational implications
5	31	B2C (regular basket)	Instant processing
10	1 023	B2C (large purchase)	Negligible load
20	1 048 575	B2B (small order)	Significant RAM load
30	1 073 741 823	B2B (typical invoice)	Out of Memory
50	> 10^15	B2B (bulk shipment)	Processing impossible

Furthermore, the integration of financial weights violates the fundamental *Downward Closure Property* (*Anti-monotonicity*), which serves as the basis for the effective pruning of irrelevant search branches. In the classical approach, the support of any subset $Y \subset X$ is always no less than the support of the superset $X : Supp(Y) \geq Supp(X)$. In financially weighted analysis, this property does not hold, making it impossible to apply standard *pruning mechanisms* and giving the search an undirected nature.

The third limitation is the "temporal blindness" of static models. Classical methods treat the transaction database as a static snapshot, where the chronology of events is irrelevant to the algorithm. The result is a rule of the form $X \rightarrow Y$, which merely establishes the fact of item co-occurrence but fails to answer the question of exactly when a customer's recurring need arises.

In the B2B segment, purchases are rational, planned, and governed by the internal business cycles of the purchasing enterprise: production processes, equipment maintenance schedules, budget cycles, and seasonality. If a customer procures an itemset X every 45 days, a proposal made on the 10th or 80th day after the previous purchase will be equally irrelevant. The failure to account for the time interval between transactions Δt deprives the system of the ability to distinguish between these situations, leading to the generation of recommendation triggers at irrelevant points in time. Receiving a non-systematic stream of offers, counterparties quickly develop indifference toward platform communications – an effect known in the literature as *Notification Fatigue*.

The fourth limitation is the lack of a differentiated customer interaction strategy. Even assuming the first three limitations are addressed, classical methods overlook a fundamentally critical aspect of B2B commerce: different counterparties possess radically different value to the enterprise and require distinct interaction strategies. The traditional approach to customer segmentation – *RFM-analysis* (Recency, Frequency, Monetary) operates on retrospective metrics: time since the last purchase, transaction frequency, and total volume. These indicators describe the past but do not predict future purchasing behavior.

Consequently, the recommendation system applies the same *lead time* and *communication window* width to all customers, regardless of their profile. A premium client with a large average check and a strictly regular procurement cycle receives an identical communication format as a small counterparty with chaotic orders – which is suboptimal from a resource profitability standpoint. For an enterprise, it is vital not only to identify a stable pattern and determine the moment of contact but also to adapt the parameters of this contact to the specific customer's financial profile. A high-value counterparty requires a precise hit at the moment of need (*narrow window, high precision*), whereas a less valuable one allows for broader coverage with a higher tolerance. The absence of such differentiation leads to the *dissipation of communication resources* and a decrease in the overall profitability of the sales department.

Conceptually, the four described limitations are interconnected and incrementally cumulative. An attempt to eliminate the first (adding financial weight) triggers the second (combinatorial explosion). Addressing the third (the temporal dimension) requires a fundamentally different data processing architecture. Finally, eliminating the fourth (differentiation of interaction) becomes possible only after the system has learned to identify both the pattern and the moment of contact for each customer, thus requiring the results of all previous steps as input data.

The alignment of these identified limitations allows for the formulation of four requirements for the method under development. *First*, replace the frequency-based significance criterion with a financially weighted one, where the weight of each transaction is proportional to the product of item quantity and value. *Second*, provide mechanisms to prevent combinatorial explosion when analyzing long B2B invoices without losing economically significant patterns. *Third*, integrate a temporal component – an assessment of the stability of time intervals between purchases to generate proactive recommendations within the window of maximum probability of a customer's recurring need. *Fourth*, introduce counterparty segmentation based on the cyclical and financial characteristics of purchasing behavior to differentially determine the parameters of the communication contact.

These requirements form the scientific basis for developing a *method for utility-weighted analysis* and customer classification, the mathematical apparatus of which is described in the following subsections.

B. Development of a mathematical model for utility-weighted transactions and the UP-Growth (Utility Pattern Growth) method

Let us consider an extended formal definition of a B2B transaction. In contrast to the classical model, where a transaction is a subset of items $T \subseteq I$, in the developed mathematical model, a *utility-weighted transaction* is represented by a *tuple* that includes not only the item identifier but also quantitative and value parameters. A transaction element is defined by the following expression:

$$t_j = \langle i_k, q_k, p_k \rangle, \tag{4}$$

where i_k is the item identifier (SKU); q_k is the quantity of the item (cases, units, or stock-keeping units) purchased within this transaction; p_k is the individual price per unit for a given counterparty, accounting for agreed-upon discounts, delivery terms, and volume coefficients.

Based on parameters q_k and p_k, a *utility function* is introduced, characterizing the financial contribution of a specific item within a transaction. The utility of an element i_k in transaction T is defined by the following expression:

$$u(i_k, T) = q_k, p_k, \tag{5}$$

where $u(i_k, T)$ is the utility of item i_k in transaction T.

Transaction Utility is the total financial value of the entire order, calculated as the sum of the contributions of all item positions according to the following expression:

$$U(T) = \sum\nolimits_{i_k \in T} u(i_k, T) = \sum\nolimits_{i_k \in T} q_k, p_k, \tag{6}$$

where $U(T)$ is the total utility (monetary value) of transaction *T*; $i_k \in T$ is an item belonging to the set of elements in transaction *T*. The value $U(T)$ from equation (6) becomes the weight of the transaction during the construction of the *utility tree*.

In the classical *FP-Growth* method, each transaction contributes an identical increment of "+1" to the node counter, regardless of its value. In the proposed *UP-Growth* method, instead of "+1", the node accumulates $U(T)$ – the actual monetary value of the order. Thus, the transaction becomes *utility-weighted*. However, the intuitive idea of replacing frequency support with financial utility encounters a serious theoretical obstacle. In the classical *FP-Growth* method, the *effective pruning* of irrelevant search branches is based on the *Downward Closure Property* (*Anti-monotonicity*). This property states that the support (1) of any subset $Y \subset X$ is always no less than the support of the superset $X: Supp(Y) \geq Supp(X)$. This allows for the immediate dismissal of entire branches of the search space without iterating through every combination.

In utility-weighted analysis, this property does not hold. A small itemset *Y* may occur predominantly in low-value transactions, whereas a broader set $X \supset Y$ may appear in high-value ones. In such a case, the actual financial utility $U(X)$ may exceed $U(Y)$, which destroys anti-monotonicity and eliminates the possibility of safely pruning search branches. The consequence is an uncontrolled growth in the number of candidates, as shown in equation (3) of Section V, Part A of this article.

To address this problem, the concept of *Transaction-Weighted Utilization* (TWU) is introduced. The TWU of an itemset *X* is calculated as the total value of all transactions in database *D* that contain this set, serving as a conservative *upper bound* for the actual utility of the set, as reflected in the following expression:

$$\mathrm{TWU}(X) = \sum_{\substack{T \in D \\ X \subseteq T}} U(T), \qquad (7)$$

where $\mathrm{TWU}(X)$ is the transaction-weighted utility of itemset *X*; $\{T \in D | X \subseteq T\}$ is the subset of transactions in database *D* containing all elements of set *X* simultaneously; $U(T)$ is the total transaction value according to expression (6).

The key theoretical property of TWU lies in the restoration of anti-monotonicity within the weighted space: for any subset $Y \subset X$, the condition $\mathrm{TWU}(Y) \geq \mathrm{TWU}(X)$ holds. This allows for the safe pruning of irrelevant branches: if $\mathrm{TWU}(X) < \mathrm{min\_utility}$, no superset of *X* can be a high-utility pattern, and further analysis is not conducted. Consequently, the method maintains the computational efficiency of FP-Growth while operating within a financially significant search space.

The process of constructing the Utility Pattern Tree (UP-Tree) according to the developed *UP-Growth* (*Utility Pattern Growth*) *method* occurs in two passes over the transaction database.

During the ***first pass***, the $\mathrm{TWU}(i_k)$ is calculated for each item i_k according to expression (7). Items for which $\mathrm{TWU}(i_k) < \mathrm{min\_utility}$ are excluded from further analysis. The remaining items are sorted in descending order of their TWU. This priority order is denoted as $\prec_U$ and determines the structure of the tree branches.

During the ***second pass***, each transaction *T* is inserted into the tree: items are ordered according to $\prec_U$, followed by a traversal of the corresponding branch. The fundamental difference from FP-Growth is that node i_k does not increment – "+1" but instead accumulates $U(T)$ according to expression (6). Formally, the node value after inserting *n* transactions passing through it is defined by the following expression:

$$NodeValue(i_k) = \sum_{j=1}^{n} U(T_j) \cdot 1[i_k \in path(T_j)], \quad (8)$$

where $NodeValue(i_k)$ is the accumulated node value for item i_k; $U(T_j)$ is the total value of the *j*-th transaction according to expression (6); $1[\cdot]$ is an indicator function equal to "1" if item i_k is present in the path of transaction T_j, and "0" otherwise.

After constructing the UP-Tree, the method recursively identifies all itemsets whose total financial value exceeds the min_utility threshold. For each item (from least to most valuable according to $\prec_U$), a *conditional* utility base is formed, from which a *conditional* UP-Tree is built. An itemset *X* is recognized as a *High-Utility Itemset* **(**HUI**)** and included in the result if the following condition is met:

$$U(X) = \sum_{\substack{T \in D \\ X \subseteq T}} u(X,T) \geq \mathrm{min\_utility}, \qquad (9)$$

where $U(X)$ is the actual total utility of itemset *X* across the entire transaction database; $\$u(X,T) = \sum_{i_k \in X} u(i_k, T)$ is the utility of itemset *X* in a specific transaction *T*; min_utility is the established threshold of minimum financial significance in UAH.

To illustrate the method's operation, let us consider a test example where database D contains five transactions $\{T_1, ..., T_5\}$ with items $\{A, B, C, D, E\}$ and a set threshold min_utility = 130 UAH. The input data and the calculated TWU values for each item are presented in Tables III and IV.

The data in these tables show that item E, which is present in three transactions, is excluded because the cumulative value of the transactions where it occurs (125 UAH) does not reach the min_utility threshold. The remaining items are sorted in descending order of their TWU: min_utility $B(240) \succ C(165) \succ D(155) \succ A(150)$. In this specific order, the items are inserted into the UP-Tree, which ensures maximum branch overlapping and minimizes the size of the data structure. The step-by-step process of constructing the UP-Tree is presented in Table V.

At the conditional pattern base generation step, the analysis is performed bottom-up according to $\prec_U$ (from A to B). For the itemset $\{B, C\}$, the actual utility is $U(T_2)+U(T_3)+U(T_5)= =55+50+60=165$ UAH, which exceeds the threshold of 130 UAH. Consequently, the set $\{B, C\}$ is recognized as a *high-utility pattern* and is included in the output. Compared to the classical FP-Growth, where this same set might be displaced by a set containing the cheaper but more frequent item E, the UP-Growth method guarantees the priority of financially significant associations.

To provide a clear understanding of the introduced enhancement, the fundamental differences between the classical FP-Growth method and the proposed UP-Growth are systematized in Table VI. As shown in Table VI, the change is targeted and methodologically justified. The sole transformation is the replacement of the unit increment of the counter with the actual transaction value, while preserving the entire tree architecture and the conditional pattern base mechanism.

TABLE III. INPUT DATA FOR THE TEST EXAMPLE (MIN_UTILITY = 130 UAN)

TID	Transaction content (SKU, quantity, price, UAH)	$U(T)$, UAH
T_1	$\{A(1,10), B(1,20), D(1,5), E(1,5)\}$	40
T_2	$\{B(2,20), C(1,10) D(1,5)\}$	55
T_3	$\{A(1,10), B(1,20), C(1,10), E(2,5)\}$	50
T_4	$\{B(1,20), E(3,5)\}$	35
T_5	$\{A(2,10), B(1,20), C(1,10), D(2,5)\}$	60

TABLE IV. TWU CALCULATION AND ITEM SELECTION

Item (SKU)	Transactions containing the item	TWU, UAH	Status
B	$T_1+T_2+T_3+T_4+T_5$	240	✓ Included
C	$T_2+T_3+T_5$	165	✓ Included
D	$T_1+T_2+T_5$	155	✓ Included
A	$T_1+T_3+T_5$	150	✓ Included
E	$T_1+T_3+T_4$	125	✗ Excluded (125 < 130)

TABLE V. STEP-BY-STEP PROCESS OF UP-TREE CONSTRUCTION

TID	Ordered sequence (by $\prec_U$)	Tree operation (Update node values)
T_1	$B \to D \to A$	New branch: $B(40) \to D(40) \to A(40)$
T_2	$B \to C \to D$	B:40+55=95; New branch $C(55) \to D(55)$
T_3	$B \to C \to A$	B:95+50=145; C:55+50=105; New branch $C(105) \to A(50)$
T_4	B	B:145+35=180;
T_5	$B \to C \to D \to A$	B:180+60=240; C:105+60=165; D:55+60=115; New branch $D(115) \to A(60)$

TABLE VI. COMPARATIVE CHARACTERISTICS OF FP-GROWTH AND UP-GROWTH METHODS

Method component	FP-Growth (Classical)	UP-Growth (Proposed)
Unit of measurement in the tree node	Occurrence counter (+1)	Transaction value, $U(T)$, UAH
Itemset significance criterion	Frequency support $Supp(X)$	Total utility $U(X) \geq \text{min\_utility}$
Anti-monotonic upper bound	*Supp* (anti-monotonic by definition)	TWU (anti-monotonic in weighted space)
Item ordering in the tree	By descending frequency	By descending TWU
Consideration of item quantity q_k	No	Yes
Consideration of item price p_k	No	Yes
Target result	Frequent Itemsets	High-Utility Itemsets
Priority in B2B	Mass-market low-cost items	Wholesale orders with maximum revenue

This implies that the UP-Growth method inherits the computational efficiency of FP-Growth while reconfiguring the pattern significance criterion in alignment with the financial logic of B2B commerce.

Thus, the developed mathematical model of a utility-weighted transaction (4) – (9) and the UP-Growth (Utility Pattern Growth) method eliminate the "financial blindness" of classical FP-Growth by modifying one key parameter: UP-Tree nodes accumulate the monetary value of transactions rather than the frequency of occurrences. The TWU mechanism restores the anti-monotonicity property and ensures computational stability even when analyzing long B2B transactions.

The output of the method is an ordered list of high-utility patterns – item combinations that generate the major share of revenue on the B2B platform. This list serves as the input data for the second component of the proposed approach: the temporal verification method, described further in Section C.

C. Development of a mathematical model for sequences of utility-weighted B2B transactions and a method for their temporal verification

The use of the UP-Growth method provides an answer to the question of *"what B2B customers buy together"* by identifying financially significant item combinations. However, to create effective recommendation content, it is equally important to answer the question of "when to offer it." A recommendation made outside the cycle of a customer's actual need is not only useless but also erodes the counterparty's trust. Furthermore, an excess of irrelevant offers (*Notification Fatigue*) leads to the ignoring of any messages from the supplier, including timely and useful ones.

This section describes the development of the second component of the proposed utility-weighted temporal approach to B2B transaction mining, namely *temporal verification*. Its task is to determine the regularity or randomness of a high-utility pattern identified by the UP-Growth method. If its appearance in the UP-Tree for a specific counterparty is regular, it becomes necessary to calculate the optimal moment for the next offer as a constituent part of the recommendation content.

The formal basis of temporal analysis is the concept of a B2B transaction sequence (2). A *sequence* s_j for customer C_j is a time-ordered chain of transactions that reflects the chronology of their purchasing activity according to the following expression:

$$s_j = \left\langle T_{\tau_1}, T_{\tau_2}, \ldots, T_{\tau_m} \right\rangle, \quad \tau_1 < \tau_2 < \tau_m, \qquad (10)$$

where s_j is the transaction sequence of customer C_j; T_{τ_m} is the transaction conducted at time τ_i; m is the total number of the customer's transactions during the analyzed period.

A *Sequential Pattern* (*SP*) in B2B describes the time-ordered occurrence of events or transactions that repeat with sufficient regularity across a set of customer sequences. If $\alpha = \left\langle T_{\tau_\alpha} \right\rangle$ and $\beta = \left\langle T_{\tau_\beta} \right\rangle$ are two sub-sequences where $\tau_\alpha < \tau_\beta$, then the sequential pattern is formulated as:

$$\text{SP} = \alpha \Rightarrow \beta, \qquad (11)$$

That is, if at time τ_α the customer conducted a transaction or a set of transactions α, there is a high probability that at a subsequent time $\tau_\beta > \tau_\alpha$, they will conduct a transaction or a set of transactions β. The ⇒ symbol emphasizes specifically the temporal, rather than logical, sequence, in contrast to the → operator used in association rules.

The difference between an association rule $AR = A \rightarrow B$ and a sequential pattern (11) is fundamental: the former describes the simultaneous presence of items within a single transaction, while the latter describes their appearance in different transactions separated by time. For creating personalized content, sales management through B2B sequential patterns is the primary analytical tool, as they model the cycle of "purchased → need restored → repurchased".

The quality of sequential patterns is evaluated using metrics similar to those for association rules, but applied to the sequence database $S = \{s_1, s_2, ..., s_n\}$. The support of a sequential pattern $S_{SP}(\alpha \Rightarrow \beta)$ is defined as the fraction of sequences in which α precedes β:

$$S_{SP}(\alpha \Rightarrow \beta) = \frac{\left|\{s \in S \mid \alpha \text{ precedes } \beta \text{ in sequence } s\}\right|}{|S|}. \quad (12)$$

The *confidence* of a sequential pattern $C_{SP}(\alpha \Rightarrow \beta)$ indicates the probability of β occurring, given that α has already occurred:

$$C_{SP}(\alpha \Rightarrow \beta) = \frac{S_{SP}(\alpha \Rightarrow \beta)}{S_{SP}(\alpha)}. \quad (13)$$

The *lift* of a sequential pattern $L_{SP} = (\alpha \Rightarrow \beta)$ evaluates the non-triviality of the connection by comparing the observed probability with the expected probability under independent events:

$$L_{SP}(\alpha \Rightarrow \beta) = \frac{S_{SP}(\alpha \Rightarrow \beta)}{S_{SP}(\alpha) \cdot S_{SP}(\beta)}. \quad (14)$$

A value of $L_{SP} > 1$ indicates that the occurrence of α genuinely increases the probability of β beyond a random level, implying a causal-temporal relationship between the events. In a B2B context, this is a necessary condition for including a pattern in the proactive recommendation system.

The identification of a sequential pattern $\alpha \Rightarrow \beta$ only confirms the existence of a temporal connection between events. For proactive forecasting, more specific information is required: namely, how many days after α the event β typically occurs. To formalize this characteristic, the concept of the *Inter-Purchase Time* (*IPT*) vector is introduced.

Suppose that for a high-utility pattern *X* identified by the UP-Growth method and a customer C_j the following sequence of purchase dates τ for this set is recorded:

$$T(X, C_j) = \{\tau_1, \tau_2, ..., \tau_n\}, \ \tau_1 < \tau_2 < \cdots < \tau_n, \quad (15)$$

where n is the number of times the set *X* was purchased by customer C_j during the analyzed period; the minimum value $n \geq 3$ is a condition for applying further analysis, as at least three purchases are necessary to obtain at least two intervals.

The IPT $\delta = (\delta_1, \delta_2, ..., \delta_{n-1})$ vector is calculated according to the following expression:

$$\delta_i = t_{i+1} - t_i, \ i = 1, 2, ..., n-1, \quad (16)$$

where δ_i is the duration of the *i*-th interval between adjacent purchases of set *X* by customer C_j in days; the value $\delta_i > 0$ by definition, as transactions are ordered chronologically.

The mean interval value μ characterizes the typical (average) purchase cycle of the set for a given customer and is calculated as follows:

$$\mu = \frac{1}{n-1} \sum_{i=1}^{n-1} \delta_i, \quad (17)$$

where μ is the average purchase cycle, in days; $n - 1$ is the number of observed intervals.

The standard deviation σ characterizes the variability of the cycle – how much individual purchases deviate from the mean value:

$$\sigma = \sqrt{\frac{1}{n-2} \sum_{i=1}^{n-1} (\delta_i - \mu)^2}, \quad (18)$$

where σ is the standard deviation of the intervals, in days. Expression (18) uses an unbiased estimate of the variance with a divisor of (n – 2) as the number of observations equals $n - 1$ intervals, and the estimate μ is a free parameter.

The mean μ (17) and the deviation σ (18) do not, by themselves, answer the question of whether a cycle is stable enough for forecasting. A deviation of σ = 10 days for an average cycle of μ = 30 days indicates significant instability, whereas the same deviation of σ = 10 days for μ = 120 days represents a minor fluctuation. For a comparable assessment of cycle stability across different durations, a dimensionless indicator – the *Coefficient of Variation* (*CV*) is applied, defined by the following expression:

$$\mathrm{CV}(X, \ C_j) = \frac{\sigma}{\mu}, \quad (19)$$

where $\mathrm{CV}(X, \ C_j) \in [0, \infty)$ is the relative variability of the inter-purchase intervals of set *X* for customer C_j; σ and μ are determined by expressions (18) and (17), respectively.

The Coefficient of Variation is the central criterion for temporal verification. Specifically, if $\mathrm{CV}(X, \ C_j) \leq \gamma$, where γ is the established stability threshold, the purchasing cycle is considered regular enough to be included in the proactive trigger system. If $\mathrm{CV}(X, C_j) > \gamma$, the purchases occur irregularly, and reliable forecasting of the next need is impossible. The interpretation of CV values regarding B2B management decisions is presented in Table VII. The threshold value γ is not fixed and can be adjusted depending on industry specifics and the enterprise's tolerance for forecast error. In most practical applications, the recommended value is $\gamma \in [0.3; 0.5]$.

TABLE VII. INTERPRETATION OF COEFFICIENT OF VARIATION VALUES FOR B2B PROCUREMENT

CV Range	Cycle Characteristics	Typical B2B Scenario	Method Decision
$0.00 \le \mathrm{CV} \le 0.15$	Strictly deterministic	Monthly procurement of consumables under a long-term contract	Generate trigger
$0.15 \le \mathrm{CV} \le 0.30$	Stable with minor stochastic noise	Quarterly procurement of the core product range with a ± 2-week margin of error	Generate trigger
$0.30 \le \mathrm{CV} \le 0.50$	Moderately stochastic	Budget-dependent seasonal procurement or purchases with shifted timelines	Generate trigger with an expanded window $k < 0.5$
$\mathrm{CV} > 0.50$	Highly stochastic / Chaotic	Ad-hoc or situational orders without pronounced regularity	Exclude from proactive triggers

For patterns that have passed the verification according to expression (19), the temporal verification method calculates the optimal moment for activating the communication trigger – specifically, the date on which the system should initiate an offer to the customer. It is taken into account that the offer must arrive not on the day of the expected purchase, but with a certain lead time: the customer needs time for review, approval, and order processing.

The trigger date $T_{trigger}$ is determined by the following expression:

$$T_{trigger} = T_{last} + \mu - k \cdot \sigma, \tag{20}$$

where $T_{trigger}$ is the calculated offer activation date; T_{last} is the date of the last actual purchase of set X by customer C_j; μ is the average cycle according to expression (17), in days; σ is the standard deviation according to expression (18), in days; $k \in [0;2]$ is the *prevention coefficient*, which determines the depth of the advance notification.

The economic significance of the coefficient k is defined as follows. At $k = 0$, the trigger is activated exactly at the moment of the average expected date of the next purchase, that is, when the need has already arisen. At $k = 1$, the trigger fires σ days before this moment; this is a standard advance notification window designed for customers with a typical order approval cycle. At $k = 2$, the offer is sent 2σ days in advance, which is appropriate for high-value and long-term procurement where the decision-making process takes several weeks (e.g., procurement of industrial equipment). The choice of k is made by the enterprise based on industry specifics and empirical data regarding the average lead time from receiving an offer to signing an order.

Expression (20) also defines the "active offer window" – specifically, the time interval $[T_{trigger}; T_{last} + \mu + \sigma]$, during which the probability of purchasing set X is maximal. Offers sent before $T_{trigger}$ fall into the "premature zone", where the need is not yet relevant. Offers sent after the upper bound of the window fall into the "delayed zone" where the customer has most likely already turned to a competitor or postponed the purchase. The concept of the proactive offer window is illustrated in Table VIII and Fig. 1.

TABLE VIII. DEFINITION OF THE PROACTIVE OFFER WINDOW ZONES

Zone Type	Interval	Characteristics
Premature	before $T_{trigger}$	Need is not yet relevant; the offer is ignored
Active (optimal)	$[T_{trigger}; T_{last} + \mu + \sigma]$	Maximum probability of offer acceptance
Late (delayed) zone	after $T_{last} + \mu + \sigma$	The customer has either already purchased from a competitor or postponed the order

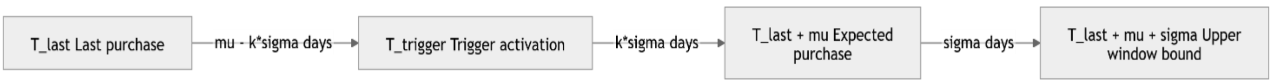

Fig. 1. Diagram of the proactive offer window using re-order probability distribution

The combination of the UP-Growth method (Section 2.2) with temporal verification forms a cohesive analytical pipeline, where each stage performs a clearly defined function. The sequential interaction of both stages is shown in Fig. 2. The first stage answers the question "what to offer" by identifying financially significant product combinations. The second stage answers the question "to whom and when" by identifying the customers for whom a given pattern is regular and determining the optimal moment for the offer.

The implementation of the two-stage architecture shown in Fig. 2 combining UP-Growth generation at the first stage with temporal verification at the second stage solves the computational complexity problem described in (3) in Section 1 of this sub-chapter. In other words, the temporal analysis (Stage 2) is applied not to millions of possible combinations, but only to a relatively small list of high-utility patterns that have passed the financial selection at Stage 1. This makes the method practically feasible for real-world transactional datasets.

VI. Results of the Research and their Discussion

Practical validation of the developed models and data mining methods, which form the basis for the proposed value-weighted temporal approach, was conducted using transactional and behavioral data from B2B e-commerce customers of the “Baza Obuvi” (bazaobuvi.com.ua) system [25].

The specificity of the research object lies in the fact that the atomic unit of analysis (SKU) is a case of footwear (size run), which contains a fixed set of pairs of the same article in various sizes (e.g., from 36 to 41). Furthermore, the transaction includes financial parameters specific to B2B trade (individual price per case, number of cases, seasonal delivery terms). Below are examples of the generated association rules and scenarios built on sequential patterns, accompanied by explanations of their economic significance.

`Association Rule` `(Product): {Women's sneakers Art. 403, q=5 cases, p=4600 UAH/case; Men's sneakers Art. 510, q=5 cases, p=5200 UAH/case}→{Insole set "Comfort", q=100 pcs, p=15 UAH/pc}`

The constructed rule has a high *Leverage level* (10), as the sale of the cross-sell product is tied to a high-value core purchase, where the total transaction utility $U(T)$ = (5×4600) + (5×5200) + (100×15) = 50500 UAH.

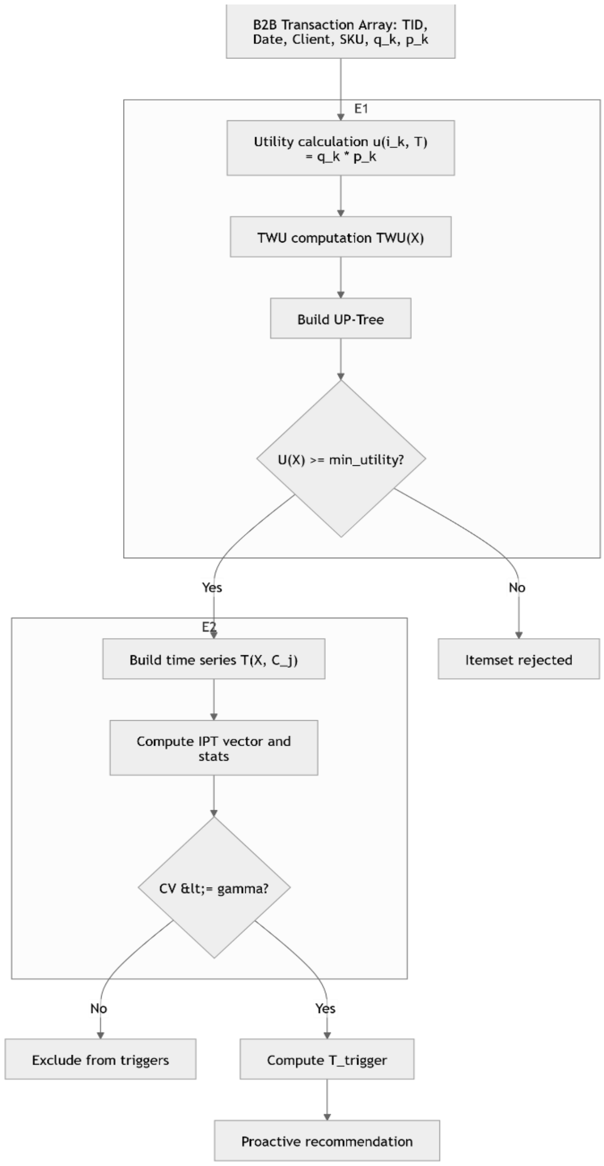

Fig. 2. Complete architecture of the value-weighted temporal analysis method

`Association Rule` `(Product + Service): {Footwear order q > 1000 pairs, U(T) > 500000 UAH}→{Credit line provision, 60-day term}`

The constructed rule links the purchase volume with financial content (crediting), representing an example of payment term personalization based on rational justification (volume q and price p).

To demonstrate the substantial meaning of expressions (15) – (20), let us consider two typical scenarios characteristic of B2B e-commerce systems in the wholesale trade sector.

Scenario 1. ***Regular Seasonal Procurement***

The customer C_1 purchased set X = {winter boots Art. 501, salt protector} on the following dates:

t_1 = 15.10.2022, t_2 = 12.10.2023, t_3 = 18.10.2024. Interval vector: δ_1 = 362 days, δ_2 = 371 days.

Calculation: μ = 366.5 days, $\sigma \approx$ 6.4 days, CV $\approx$ 0.017.

The CV << γ value indicates a *strictly deterministic* annual cycle (Table VIII). At $k = 1$, the trigger date for the next season is: T_{trigge} = 18.10.2024 +367 – 6 ≈ 14.10.2025. The system will proactively generate and offer the customer a wholesale bundle from the winter assortment.

Scenario 2. *Irregular Procurement*

The customer C_2 purchased set Y = {men's sneakers Art. 403} on the following dates: t_1 = 05.02.2023, t_2 = 18.05.2023, t_3 = 11.01.2024. Interval vector: δ_1 = 102 days, δ_2 = 238 days

Calculation: μ = 170 day, $\sigma \approx$ 96 day, CV $\approx$ 0.56

The $CV > \gamma = 0.5$ value indicates the *Highly stochastic / Chaotic* annual cycle (see Table VIII). Set Y for this customer C_2 is excluded from proactive triggers. Possible explanation: the purchase of sneakers depends on market conditions (availability of promotions, shifts in demand) and is not subject to a stable cycle.

These two scenarios clearly demonstrate the key advantage of temporal verification. It distinguishes systemic business cycles from situational orders, allowing the sales department to focus resources exclusively on interactions where a proactive offer has a real chance of success.

For further discussion, let us present a summary of the proposed value-weighted temporal approach to B2B transaction mining.

The implementation of the mathematical representation formalized by expressions (3) – (20) eliminates three limitations of the classical FP-Growth method identified in Section 1 of this sub-chapter, namely:

- financial blindness by replacing the frequency criterion with the utility metric $U(X)$;
- combinatorial explosion through the TWU-mechanism and two-stage filtering;
- temporal blindness through the analysis of IPT-vectors and the CV criterion.

A summary comparison of the value-weighted temporal approach and the classical FP-Growth method is presented in Table IX.

TABLE IX. COMPARATIVE CHARACTERISTICS OF THE CLASSICAL FP-GROWTH METHOD AND HIGH-UTILITY TEMPORAL ANALYSIS OF B2B TRANSACTIONS

Characteristic	FP-Growth	High-Utility Temporal Analysis
Pattern significance criterion	Frequency (Supp)	Financial utility $U(X)$, UAN
Accounting for q_k and p_k	No	Yes
Protection against combinatorial explosion	Only through *antimonot. Supp*	TWU + Stage 1 value filter
Analysis of temporal cycles	No	Yes (IPT, μ, σ, CV)
Customer-level personalization	No	Yes (individual cycle C_j)
Offer timing	Not defined	$T_{trigger} = T_{last} + \mu + k\sigma$
Output type	Static rule $A \rightarrow B$	Dynamic trigger $T_{trigger}$
Applicability to B2B	Limited	Target

VII. CONCLUSION

The study addresses the relevant scientific task of improving data mining methods for content personalization in B2B e-commerce systems. Unlike traditional approaches focused on purchase frequency, the proposed High-Utility Temporal Approach accounts for both the financial significance and the temporal dynamics of business transactions.

The main results of the study are as follows:

1) Justification of classical model limitations. It was established that using Apriori and FP-Growth algorithms in the B2B segment leads to “financial blindness” (ignoring order value) and “temporal blindness” (lack of cycle forecasting). This justifies the necessity of transitioning from binary models to Utility Mining.

2) Development of a value-weighted model based on UP-Growth. The mathematical framework of the UP-Growth algorithm was adapted, allowing for the replacement of the support metric $Supp(X)$ with a utility function TWU(X) that considers individual price and wholesale batch volume. This ensured the identification of High-Utility Itemsets, which form the basis for rational product content.

3) Integration of temporal analysis. A method for identifying cyclic patterns based on inter-purchase time (IPT) vectors was implemented. The use of the coefficient of variation (CV) as a cycle stability criterion formalized the calculation of the "trigger date" for generating proactive Restock Prediction recommendations.

4) Practical validation. Experimental results using real-world data from the “Baza Obuvi” platform confirmed that the proposed two-stage pipeline (utility extraction and time-window calculation) increases the relevance of personalized content, transforming it from purely informational into an effective tool for supporting rational decision-making by B2B customers.

Thus, the conducted research establishes the necessary theoretical foundation for transforming B2B e-commerce systems from passive catalogs into intelligent decision support systems, thereby increasing customer loyalty and the platform's economic efficiency.

REFERENCES

[1] Global Ecommerce Statistics: Trends to Guide Your Store in 2025 – Available from: https://www.shopify.com/enterprise/blog/global-ecommerce-statistics – [Accessed: Nov 2025].

[2] Insider Intelligence – Available from: https://www.emarketer.com/insights/mobile-commerce- shopping-trends-stats/ [Accessed: Nov 2025].

[3] T. S. Panasiuk, & I. R. Zubenko, (2024). Perspektyvy rozvytku b2c elektronnoi komertsii v sviti [Prospects for the development of b2c e-commerce in the world]. https://doi.org/10.5281/zenodo.14603373 [in Ukrainian]

[4] O. O. Arsirii, D. V. Ivanov, & O. Yu. Babilunha, Analiz modelei, metodiv ta tekhnolohii stvorennia personalizovanoho kontentu v systemam B2B elektronnoi komertsii [Analysis of models, methods and technologies for creating personalized content in B2B e-commerce systems]. *Informatyka. Kultura. Tekhnika [Informatics. Culture. Technology]*, 2, 273–278, 2025. https://ict.op.edu.ua/index.php/journal/uk/issue/view/3 [in Ukrainian]

[5] Y.-H. Hu, Y.-L. Chen, & K. Tang, “Mining sequential patterns in the B2B environment,” *Journal of Information Science*, 35(6), 677-694, 2009. https://doi.org/10.1177/0165551509103600 (Original work published 2009)

[6] D. Kahneman, “Thinking, Fast and Slow,” *Farrar, Straus and Giroux August 2014 Statistical Papers*, 55(3), 2011. https://doi.org/10.1007/s00362-013-0533-y

[7] E. Turban, J. Outland, D. King, et al., “Electronic Commerce 2018: A Managerial and Social Networks Perspective,” *Springer International Publishing*, 2018, Cham. https://doi.org/10.1007/978-3-319-58715-8

[8] N. M. Shveda, & O. I. Krauze, (2024). Elektronna komertsiia: suchasnyi stan ta stratehii rozvytku [E-commerce: current state and development strategies]. *Mizhnarodnyi naukovyi zhurnal "Internauka". Seriia: Ekonomichni nauky [International Scientific Journal "Internauka". Series: Economic Sciences]*, 2 (82). URL: https://doi.org/10.25313/2520-2294-2024-2-9639. [in Ukrainian]

[9] Gunasekaran, A., et al. (2002) E-Commerce and Its Impact on Operations Management. International. Journal of Production Economics, 75, 185–197. https://doi.org/10.1016/S0925-5273(01)00191-8

[10] Fred Wiersema, “The B2B Agenda: The current state of B2B marketing and a look ahead,” *Industrial Marketing Management*, vol. 42, Issue 4, pp. 470–488, May 2013. https://doi.org/10.1016/j.indmarman.2013.02.015. (https://www.sciencedirect.com/science/article/pii/S0019850113000345)

[11] G. L. Lilien, “The b2b Knowledge Gap,” *International Journal of Research in Marketing*, 75, pp. 196–210, 2016. https://doi.org/10.1016/j.ijresmar.2016.01.003

[12] Y.-H. Hu, Y.-L. Chen, & K. Tang, “Mining sequential patterns in the B2B environment,” *Journal of Information Science*, 35(6), pp. 677–694, 2009. https://doi.org/10.1177/0165551509103600 (Original work published 2009)

[13] E. Cramer-Flood, (2022). Global Ecommerce Forecast & Growth Projections. – Available from: https://www.insiderintelligence.com/content/globalecommerce-forecast-2022 – [Accessed: Nov 2025]

[14] Rakesh Agrawal, Tomasz Imieliński, and Arun Swami, “Mining association rules between sets of items in large databases,” *In Proceedings of the 1993 ACM SIGMOD international conference on Management of data (SIGMOD '93)*, Association for Computing Machinery, New York, NY, USA, 1993, pp. 207–216. https://doi.org/10.1145/170035.170072

[15] O. Arsirii, & D. Ivanov, “A Model and Method of Transactional-behavioral Data Mining for B2B Content Personalization,” *Electronics and Control Systems*, 1(87), pp. 14–23, 2026. https://doi.org/10.18372/1990-5548.87.20881

[16] Francesco Ricci & Lior Rokach, & Bracha Shapira, Recommender Systems Handbook, 2010. https://doi.org/10.1007/978-0-387-85820-3_1

[17] J. Han, J. Pei, Y. Yin, et al., “Mining Frequent Patterns without Candidate Generation: A Frequent-Pattern Tree Approach,” *Data Mining and Knowledge Discovery*, 8, 53–87, 2004. https://doi.org/10.1023/B:DAMI.0000005258.31418.83

[18] M. J. Zaki, "Scalable algorithms for association mining," *in IEEE Transactions on Knowledge and Data Engineering*, vol. 12, no. 3, pp. 372–390, May-June 2000, https://doi.org/10.1109/69.846291

[19] Grow your B2B sales with 6 smart B2B product bundling strategies: Available from: https://turis.app/b2b-ecommerce/grow-b2b-sales-6-b2b-product-bundle-strategies/ – [Accessed: Nov 2025

[20] Vincent Tseng, & Cheng-Wei Wu, & Bai-En Shie, & Philip Yu, “UP-Growth: An efficient algorithm for high utility itemset mining,” *Proceedings of the ACM*

SIGKDD International Conference on Knowledge Discovery and Data Mining, 2010, pp. 253–262. https://doi.org/10.1145/1835804.1835839

[21] W. Gan, J. C. -W. Lin, P. Fournier-Viger, H. -C. Chao, V. S. Tseng and P. S. Yu, "A Survey of Utility-Oriented Pattern Mining," in *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 4, pp. 1306–1327, 1 April 2021, https://doi.org/10.1109/TKDE.2019.2942594. https://ieeexplore.ieee.org/document/8845637

[22] Grow your B2B sales with 6 smart B2B product bundling strategies: Available from: https://turis.app/b2b-ecommerce/grow-b2b-sales-6-b2b-product-bundle-strategies/ – [Accessed: Nov 2025].

[23] Investopedia: SKU: What it is and How it Works by Andrew Bloomenthal Updated June 15, 2025, Reviewed by Samantha Silberstein, Fact checked by Ariel Courage: Available from: https://www.investopedia.com/terms/s/stock-keeping-unit-sku.asp – [Accessed: Jun 2025]

[24] O. O. Arsirii, O. M. Liubomska, O. V. Rudenko, & D. V. Ivanov, (2024). Hibrydna rekomendatsiina systema dlia pidtrymky UI/UX dyzaineriv [Hybrid recommendation system to support UI/UX designers]. *Odeska politekhnika [Odesa Polytechnic]*, 29–35. https://doi.org/10.15276/ict.01.2024.30 https://ics60.aait.od.ua/zbirnik2024.pdf [in Ukrainian]

[25] *Systema elektronnoi komertsii B2B "Baza vzuttia" [B2B E-commerce system "Baza Obuvi"]*. (n.d.). Available at: https://bazaobuvi.com.ua/ua/ [Accessed: Nov 2025] [in Ukrainian]

Received: January 29, 2026
Accepted: February 19, 2026
Published: April 17, 2026

Arsirii Olena. ORCID 0000-0001-8130-9613. Doctor of Science (Eng), Professor, Head of the Department of Information Systems.
Odesa Polytechnic National University, Odessa, Ukraine.
Education: Odessa Polytechnic Institute, Odessa, Ukraine, (1981).
Research interests: Information technology; artificial intelligence; decision support systems; machine learning; neural networks.
Publications: more than 100 papers.
E-mail: e.arsiriy@gmail.com

Ivanov Dmitriy. ORCID 0009-0009-3958-5310. Postgraduate Student.
Department of Information Systems, Odesa Polytechnic National University, Odessa, Ukraine.
Education: Odesa Polytechnic National University, Odessa, Ukraine, (2016).
Research interests: Data mining, e-commerce, B2B systems.
Publications: 6.
E-mail: ivanovdima9988@gmail.com

О. О. Арсірій, Д. В. Іванов. Інтелектуальний аналіз транзакційних даних B2B-клієнтів на основі ціннісно-зваженого темпорального підходу

Сучасні системи B2B-комерції генерують значні масиви транзакційних даних, проте класичні методи інтелектуального аналізу асоціативних правил мають низку критичних обмежень, зокрема «фінансову сліпоту» бінарних моделей та «темпоральну сліпоту» статичних підходів. Це унеможливлює виявлення паттернів, що мають реальну економічну цінність, та ігнорує циклічність оптових закупівель. У роботі запропоновано ціннісно-зважений темпоральний підхід до аналізу B2B-транзакцій. В основу підходу покладено математичну модель UP-Growth (Utility Pattern Growth), яку адаптовано для роботи з функцією корисності, що базується на раціональному контексті транзакції у вигляді добутку індивідуальної ціни та обсягу оптової закупівлі. Для подолання проблеми антимонотонності введено концепцію транзакційно-зваженої корисності. Інтеграція темпорального фактора реалізована через аналіз векторів міжзакупівельних інтервалів із використанням коефіцієнта варіації як критерію стабільності циклу. В результаті розроблено цілісну архітектуру двоетапного конвеєра інтелектуального аналізу транзакційних даних B2B-клієнтів. На першому етапі конвеєр здійснює екстракцію високовартісних товарних наборів (High-Utility Itemsets), а на другому розрахунок дати тригера для формування вікна проактивної пропозиції. Порівняльний аналіз із класичним алгоритмом FP-Growth підтвердив вищу ефективність запропонованого підходу у виявленні економічно значущих закономірностей, що підтверджено на прикладі реальних даних платформи «База взуття».

Ключові слова: B2B електронна комерція, інтелектуальний аналіз, виявлення високовартісних патернів, прогнозування, аналіз поведінки клієнтів, темпоральний аналіз даних, персоналізація контенту, рекомендаційні системи.

Арсірій Олена Олександрівна. ORCID 0000-0001-8130-9613. Доктор технічних наук. Професор.
Завідувачка кафедра інформаційних систем, Національний університет “Одеська Політехніка”, Одеса, Україна.
Освіта: Одеський політехнічний інститут, Одеса, Україна, (1981).
Напрямок наукової діяльності: Інформаційні технології; штучний інтелект; системи підтримки рішень; машинне навчання; нейронні мережі.
Кількість публікацій: більше 100 наукових робіт.
E-mail: aprodeus@gmail.com

Іванов Дмитро В’ячеславович. ORCID 0009-0009-3958-5310. Аспірант.
Кафедра інформаційних систем, Національний університет “Одеська Політехніка”, Одеса, Україна..
Освіта: Одеський Національний політехнічний університет, Одеса, Україна, (2016).
Напрямок наукової діяльності: Інтелектуальний аналіз даних, електронна комерція, B2B-системи.
Кількість публікацій: 6.
E-mail: ivanovdima9988@gmail.com