UDC 004.8:519.86:336.76

N. O. KOZUB, O. S. KORNIIENKO
Kherson National Technical University

MODELS AND ALGORITHMS OF AN INTELLIGENT INFORMATION TECHNOLOGY FOR PERSONALIZED INVESTMENT PORTFOLIO OPTIMIZATION BASED ON ADAPTIVE RISK PROFILING

The paper develops an intelligent information technology for personalized investment portfolio optimization based on adaptive risk profiling. The relevance of the study is determined by the growing demand for digital decision-support tools that can account for individual investor characteristics and operate with heterogeneous financial instruments within a unified portfolio construction process. In many practical solutions, personalization is reduced either to selecting one of several predefined portfolio templates or to assigning an investor to a discrete risk category. Such approaches are insufficiently flexible and do not provide a formal connection between the estimated risk profile and the optimization constraints that determine the final portfolio structure.

The purpose of the study is to develop and formally describe an intelligent information technology for personalized portfolio optimization and to justify its algorithmic support and validation approach. The proposed technology is represented as the tuple $IT = \langle D, M, Algo, C, V \rangle$, where D denotes market and user data with preprocessing procedures, M denotes intelligent models, Algo denotes algorithmic portfolio-construction components, C denotes a parameterized system of optimization constraints, and V denotes visualization and validation modules. Adaptive questionnaire responses are transformed into a user feature vector and then mapped by a machine-learning model to a continuous risk profile. Personalization is implemented through a formal dependence between the user risk value and the feasible portfolio set.

Two key algorithmic modules are proposed. The first forms a candidate asset universe by reducing redundancy in the initial instrument set on the basis of correlation or covariance dependencies, which decreases dimensionality and improves the stability of optimization results. The second parameterizes optimization constraints as a function of the continuous risk profile, including bounds on asset-class weights, the share of anchor instruments, and concentration restrictions. Portfolio construction is formulated as an optimization problem in which the feasible set is determined by personalized constraints and the reduced candidate asset universe. The paper also considers computational complexity in offline and online contours, portfolio-quality metrics, stability metrics with respect to small risk-profile changes, and an ablation-based validation procedure.

The practical value of the proposed approach lies in combining explainability, reproducibility, controllable personalization, and computational feasibility within one integrated technology. The obtained results show that the proposed formalization provides a coherent foundation for building applied portfolio recommendation systems that are more stable, interpretable, and adaptable to individual investor characteristics.

Keywords: *personalized portfolio optimization, adaptive risk profiling, intelligent information technology, candidate asset universe, portfolio constraints, portfolio stability, correlation analysis, ablation validation.*

Н. О. КОЗУБ, О. С. КОРНІЄНКО
Херсонський національний технічний університет

МОДЕЛІ ТА АЛГОРИТМИ ІНТЕЛЕКТУАЛЬНОЇ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ПЕРСОНАЛІЗОВАНОЇ ОПТИМІЗАЦІЇ ІНВЕСТИЦІЙНОГО ПОРТФЕЛЯ НА ОСНОВІ АДАПТИВНОГО РИЗИК-ПРОФІЛЮВАННЯ

У статті розроблено інтелектуальну інформаційну технологію персоналізованої оптимізації інвестиційного портфеля на основі адаптивного ризик-профілювання. Актуальність дослідження зумовлена зростанням попиту на цифрові засоби підтримки інвестиційних рішень, здатні враховувати індивідуальні характеристики інвестора та працювати з гетерогенними фінансовими інструментами в межах єдиного процесу формування портфеля. У багатьох прикладних рішеннях персоналізація зводиться або до вибору одного з кількох шаблонних портфелів, або до віднесення інвестора до дискретної категорії ризику. Такі підходи є недостатньо гнучкими та не забезпечують формального зв'язку між оціненим ризик-профілем і тими оптимізаційними обмеженнями, що визначають кінцеву структуру портфеля.

Метою дослідження є розроблення та формальний опис інтелектуальної інформаційної технології персоналізованої оптимізації портфеля, а також обґрунтування її алгоритмічного забезпечення й підходу до валідації. Запропоновану технологію подано у вигляді кортежу $IT = \langle D, M, Algo, C, V \rangle$, де D – ринкові й користувацькі *дані з процедурами попередньої обробки, M – інтелектуальні моделі, Algo – алгоритмічні компоненти побудови портфеля, C – параметризована система оптимізаційних обмежень, V – модулі візуалізації та валідації. Відповіді адаптивного опитувальника перетворюються на вектор ознак користувача, який далі за допомогою моделі машинного навчання відображається у неперервний ризик-профіль. Персоналізація реалізується через формальну залежність між значенням ризику користувача та множиною допустимих портфельних рішень.*

Запропоновано два ключові алгоритмічні модулі. Перший формує кандидатну множину активів шляхом зменшення надлишковості вихідного набору інструментів на основі кореляційних або коваріаційних залежностей, що знижує розмірність задачі та підвищує стійкість результатів оптимізації. Другий параметризує оптимізаційні обмеження як функцію неперервного ризик-профілю, зокрема межі ваг класів активів, частку опорних інструментів і концентраційні обмеження. Побудову портфеля формалізовано як оптимізаційну задачу, у якій множина допустимих розв'язків визначається персоналізованими обмеженнями та зменшеною кандидатною множиною активів. У роботі також розглянуто обчислювальну складність в офлайн- та онлайн-контурах, метрики якості портфеля, метрики стійкості щодо малих змін ризик-профілю та процедуру валідації на основі абляційного аналізу.

Практична цінність підходу полягає у поєднанні пояснюваності, відтворюваності, керованої персоналізації та обчислювальної придатності в межах єдиної інтегрованої технології. Отримані результати показують, що запропонована формалізація створює цілісну основу для побудови прикладних систем портфельного рекомендування, які є більш стійкими, інтерпретованими та адаптованими до індивідуальних характеристик інвестора.

Ключові слова: *персоналізована оптимізація портфеля, адаптивне ризик-профілювання, інтелектуальна інформаційна технологія, кандидатна множина активів, портфельні обмеження, стійкість портфеля, кореляційний аналіз, абляційна валідація.*

Statement of the problem

The rapid development of digital financial services has increased the demand for intelligent decision-support technologies capable of generating personalized investment recommendations. In the field of portfolio construction, this problem is especially relevant because investors differ in their financial goals, risk tolerance, behavioral preferences, and acceptable portfolio structure. At the same time, modern portfolio solutions must operate with heterogeneous asset classes that differ in expected return, volatility, liquidity, denomination currency, and mutual statistical dependencies [1].

In many practical systems, personalization is implemented in a simplified form. Most often, users are either assigned to one of several discrete risk categories or offered a choice among predefined template portfolios. Although such approaches are easy to implement, they do not provide sufficient flexibility and do not ensure a formal relationship between the estimated user profile and the optimization constraints that determine the final portfolio composition. As a result, recommendations may be weakly interpretable, poorly controllable, and unstable under small changes in user inputs.

Another important problem is the redundancy of the initial asset universe. If highly correlated or statistically similar instruments are included simultaneously in the optimization process, the dimensionality of the problem increases, while the resulting portfolio may become less stable and harder to explain. Therefore, the relevant scientific and applied task is to develop an intelligent information technology in which adaptive risk profiling, reduction of asset redundancy, parameterization of optimization constraints, and portfolio validation are integrated into a single formalized process. This determines the direction of the present study.

Analysis of recent research and publications

The theoretical basis of portfolio construction is formed by classical portfolio theory, in which the relationship between expected return and risk is expressed through the optimization of portfolio weights using expected returns and covariance matrices [2; 3]. These approaches remain fundamental for modern recommendation systems because they provide a formal mechanism for transforming market data into portfolio decisions. At the same time, further studies in portfolio management have shown the importance of robust optimization, concentration control, and practical constraints that improve the stability and applicability of portfolio solutions in real-world settings [4; 5; 6].

A separate direction of research concerns the use of intelligent methods for investor profiling. In our previous work, an AI-based adaptive investor survey service was proposed for determining an individual risk profile, which creates a practical basis for personalized portfolio recommendation [7]. More broadly, machine-learning approaches make it possible to map questionnaire-based and behavioral data into a quantitative estimate of risk tolerance, which can then be used in decision-support systems [8; 9]. However, in many applied implementations personalization is still limited to assigning users to several broad risk classes instead of constructing a formal dependence between the assessed profile and the optimization problem.

An additional aspect that requires consideration is the structure of the candidate asset universe. Empirical results presented in our earlier study on the diversification saturation point in ETF portfolios demonstrated the practical importance of identifying redundancy among highly correlated instruments and reducing the search space before optimization [10]. At the same time, portfolio solutions for Ukrainian investors require adaptation to the local financial context, including the integration of domestic government bonds together with global market instruments within a single recommendation framework [11].

Thus, the existing body of research provides a strong foundation in portfolio theory, robust optimization, and risk profiling, but does not fully formalize their integration into an intelligent information technology with explicit links between adaptive risk profiling, candidate asset selection, parameterized optimization constraints, and experimental validation. This gap determines the novelty and relevance of the present study.

Purpose of the study

The purpose of this study is to develop and formally describe an intelligent information technology for personalized investment portfolio optimization based on adaptive risk profiling, as well as to justify its algorithmic support and the approach to its experimental validation.

To achieve this purpose, the following tasks are addressed: to formalize the proposed technology as a system of interconnected components; to develop an algorithm for constructing a candidate asset universe based on correlation dependencies; to develop an algorithm for parameterizing optimization constraints as a function of a continuous risk profile; and to define quality and stability metrics, including an ablation-based procedure for evaluating the contribution of individual algorithmic modules.

Presentation of the main research material

Let us define the proposed intelligent information technology for personalized investment portfolio optimization based on adaptive risk profiling as the tuple:

$$IT = \langle D, M, Algo, C, V \rangle,$$

where D denotes data and preprocessing procedures, M denotes intelligent models, $Algo$ denotes algorithmic components of portfolio construction, C denotes a formalized constraint system (including risk-profile-parameterized constraints), and V denotes visualization and experimental validation modules. This representation makes it possible to separate and formally describe the key stages of the process–from forming the user profile to obtaining the portfolio decision and verifying it. The overall technology structure and data flows between components are shown in Fig. 1.

Let U be the set of users (investors) and A the set of available financial instruments. For consistent statistical analysis, assume a discrete set of time points T corresponding to the chosen frequency (e.g., monthly). For each asset $a \in A$ and each time point $t \in T$, define the price $P(a, t)$ and the return $r(a, t)$, which can be specified as simple or logarithmic depending on the implementation assumption. Based on historical $r(a, t)$, we estimate asset return and risk parameters $\mu(a)$ and $\sigma(a)$, as well as inter-asset dependencies in the form of the covariance matrix Σ and the correlation matrix Corr. Introducing Σ and Corr is essential, because they are used for algorithmic reduction of a redundant asset set prior to optimization.

https://doi.org/10.32782/mathematical-modelling/2026-9-1-16 ISSN 2618-0332 (print), 2618-0340 (online)

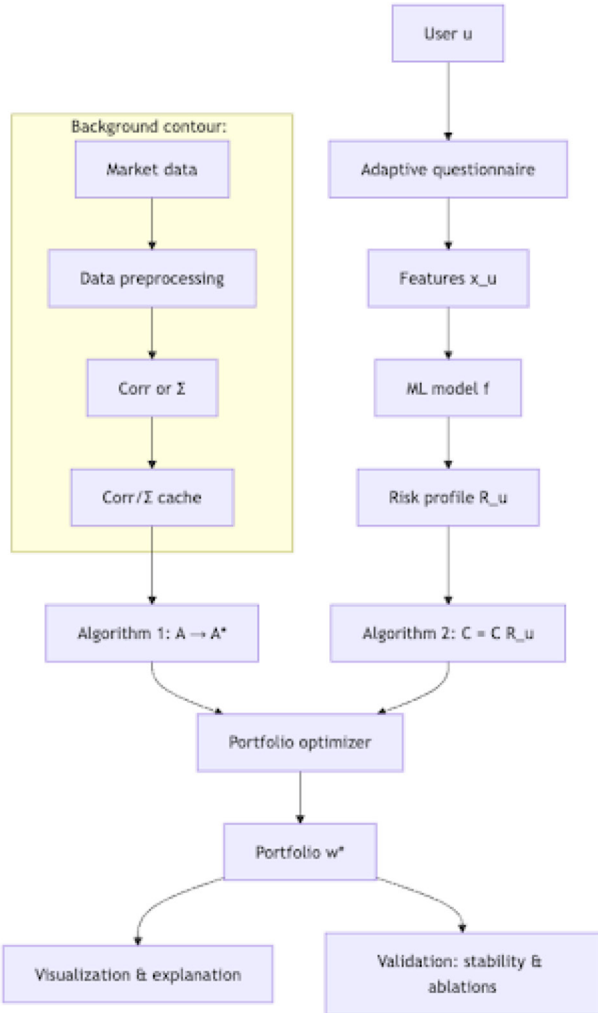

Fig. 1. Overall scheme of the intelligent information technology for personalized portfolio optimization. The background contour of market-data preprocessing (including Corr/Σ) runs in batch mode with subsequent caching; personalized portfolio construction is performed online for a specific user

A separate layer of formalization concerns user data. Let a user $u \in U$ complete an adaptive questionnaire, resulting in a feature vector $x(u) \in Rm$ [7]. The vector $x(u)$ is obtained by applying response normalization and encoding rules and, if needed, integrating additional features that characterize behavioral attitudes or user constraints [8]. Next, the user's risk profile is estimated on a continuous scale:

$$R(u) = f(x(u)), \quad R(u) \in [0, 1],$$

where $f(\cdot)$ is a regression model that maps individual risk tolerance to a scalar score [9]. Using continuous R(u) allows moving from discrete "risk classes" to parameterized personalization: the risk profile becomes a control parameter for constraint construction and the structure of the portfolio decision.

The algorithmic component *Algo* in the proposed technology includes at least two key modules. The first module constructs a candidate asset universe A^* ($A^* \subseteq A$) based on the correlation or covariance matrix and a redundancy-reduction criterion; this reduces problem dimensionality and improves stability of portfolio weights. The second module constructs the constraint system $C(R(u))$ as a function of the risk profile, i.e., it defines the feasible set of portfolios not by general "one-size-fits-all" rules, but in a personalized way depending on $R(u)$. Based on the resulting A^* and $C(R(u))$, an optimization problem is defined whose solution yields the user's portfolio weight vector $w^*(u) \in R|A^*|$.

https://doi.org/10.32782/mathematical-modelling/2026-9-1-16 ISSN 2618-0332 (print), 2618-0340 (online)

Under this view, optimization is a coherent part of the information technology: its inputs, parameters, and constraints are generated by upstream modules, and the result is then interpreted and validated.

Component C formalizes constraints that determine the admissible solution class. These include constraints on asset-class weights, bounds on the share of an "anchor" asset (notably government bonds), concentration constraints (upper bounds on individual instrument weights or aggregate concentration measures), and, if needed, a cardinality constraint limiting the maximum number of instruments in the portfolio. Importantly, a substantial part of these constraints is defined as a function of R(u), ensuring a formal link between risk assessment and portfolio formation.

Component V supports result presentation and experimental validation. In applied intelligent systems it is important not only to display the resulting weights but also to explain the reasons behind the decision–e.g., through active constraints $C(R(u))$ and asset-selection rules used to form A^*. Experimental validation includes evaluating portfolio quality via traditional metrics, as well as analyzing stability of portfolio weights under a small change of the risk profile $R(u)$ by ε. Additionally, ablation experiments are conducted, where the technology is evaluated in the full configuration and with certain algorithmic modules disabled; this enables quantitative assessment of the contribution of A^* formation and $C(R)$ parameterization to final decision quality.

Given these definitions, the technology can be viewed as the following transformation chain: user u produces $x(u)$, then $R(u)$ is obtained; next A^* and $C(R(u))$ are formed; the optimization problem is solved and yields $w^*(u)$, which is then explained and validated in module V. This formalization provides the basis for a detailed description of algorithms and the optimization problem statement, as well as for reproducible experimental verification of the proposed information technology.

Algorithm 1. Constructing the candidate asset set ($A \to A^*$) with caching

In portfolio optimization problems, the full instrument set A may contain assets with very similar statistical properties and high mutual correlation. A large number of "near-duplicate" instruments increases the dimensionality of the optimization problem and often makes results unstable: even small changes in input data or parameters may lead to substantial changes in portfolio weights, while the economic meaning of the recommendation changes little. For an applied information technology, this is undesirable because it reduces interpretability and reproducibility. Therefore, before optimization we propose an algorithmic stage that reduces the search space by transforming A into a subset A^* ($A^* \subseteq A$).

The algorithmic modules for candidate universe construction and constraint parameterization, as well as their interaction with cached statistical artifacts, are schematically presented in Fig. 2.

The input to the algorithm consists of the asset set A and estimated dependencies between assets in the form of a correlation matrix Corr or a covariance matrix Σ. In addition, a threshold parameter $\tau \in [0, 1]$ is specified, determining when two assets are considered overly similar (e.g., if $|Corrij| \geq \tau$). If needed, a desired output size k or an admissible range $[k_{\min}, k_{\max}]$ can be specified to constrain the candidate set size.

A graph-based representation is used for selection. Consider an undirected graph $G = (V, E)$ where each vertex corresponds to an asset from A. An edge between two vertices is added if the corresponding assets have absolute correlation at least τ. Under this representation, connected components can be interpreted as groups of assets that are mutually substitutable or close in return behavior. Next, one representative is selected from each connected component–the most "central" asset within the group. A practical centrality criterion is minimizing the mean absolute correlation with other assets in the same component. After selecting representatives from all components, the subset A^* is formed and used at the next optimization step.

The computational cost depends on the usage mode of statistical artifacts. Building the correlation matrix $Corr$ for $n = |A|$ assets requires computing pairwise correlations and has order $O(n2)$ in the number of pairs (for a fixed window length). Constructing the threshold graph may also be $O(n2)$ in the worst case, because all pairs must be examined. Finding connected components

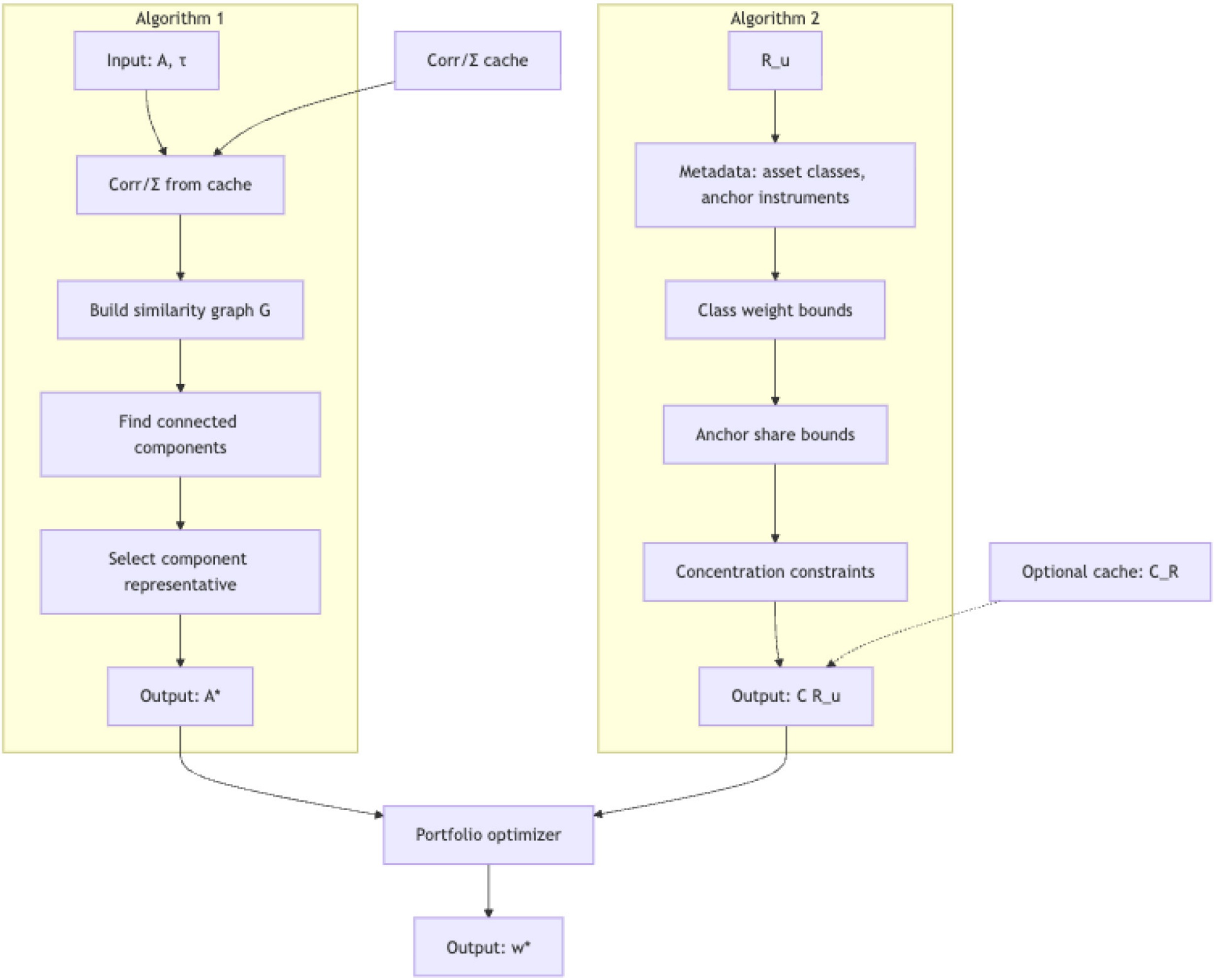

Fig. 2. Detailed view of the key algorithmic modules. Algorithm 1 forms the candidate asset universe $A \to A^*$ using cached Corr/Σ; Algorithm 2 parameterizes the constraint system $C(R)$ based on the user risk profile, with optional caching for discretized values of R

runs in $O(|V| + |E|)$ and for dense graphs may approach $O(n2)$. Selecting representatives by directly computing mean correlations within components also does not exceed $O(n2)$ in total.

However, in an applied system it is reasonable to compute Corr in the background and refresh it in batch mode (e.g., daily or upon market-data updates), then cache the result. During inference, Corr is assumed to be already available, and costs shift to graph construction and component search. For a fixed threshold τ, caching the graph itself or even the final result A^* is possible, enabling the candidate universe to be formed once and reused for many users. Thus, the $O(n2)$ estimate for building Corr reflects the cost of background preparation, while the online stage of forming A^* may be substantially faster due to cached intermediate results.

Algorithm 1 does not aim to “select the best assets” in a financial sense; rather, it is designed to construct a compact and structured search space for the optimization module, reduce redundancy among instruments, and improve stability of the portfolio decision. The next section describes Algorithm 2 for parameterizing constraints $C(R)$, which formalizes personalization depending on the user’s risk profile.

Algorithm 2. Parameterizing optimization constraints as a function of the risk profile: C = $C(R)$

In applied portfolio recommendation systems, personalization is often reduced to a discrete choice among a few predefined “profiles” or templates. This approach is simple but has two significant

https://doi.org/10.32782/mathematical-modelling/2026-9-1-16 ISSN 2618-0332 (print), 2618-0340 (online)

drawbacks. First, it poorly reflects individual differences within a single risk class because it relies on coarse discretization. Second, it does not define a formal link between the estimated risk profile and the constraints actually applied in the optimization problem: bounds on asset-class weights, the share of "anchor" instruments, concentration constraints, etc. In the proposed information technology, personalization is implemented differently: the user risk profile $R(u)$ is defined on a continuous scale in [0, 1] and is used as a control parameter for constructing the constraint system, i.e., $C = C(R(u))$.

The constraint-parameterization algorithm takes as input the risk profile R(u) and the metadata required to define rules (e.g., asset-class partitioning, identification of "anchor" instruments, admissible global bounds). The output is a concrete set of constraints C(R(u)) defining the feasible set of portfolios for the user–i.e., which portfolio solutions are considered admissible when solving the optimization problem. The constraint system is designed to reflect intuitive risk-preference logic while remaining formal and reproducible.

It is convenient to organize $C(R(u))$ as a set of parameterized bounds. First, class allocation bounds are defined. If the instrument set is partitioned into classes (e.g., equities, bonds, crypto), then for each class a lower and upper bound for its total portfolio share is specified. These bounds are chosen as functions of $R(u)$: as $R(u)$ increases, the upper bound for risky classes increases, while the minimum requirement for conservative classes may decrease. In practice, continuous or piecewise-linear functions are convenient because they ensure controlled behavior of bounds under small changes in $R(u)$.

Next, an "anchor" asset (or anchor class) constraint is introduced to fix the share of instruments that provide stabilization (e.g., government bonds). For low values of $R(u)$, a higher minimum anchor share is appropriate; for high values of $R(u)$, this minimum can decrease, but not necessarily to zero if the technology targets practical stability and controllability of risk. Upper bounds on the anchor share may also be imposed to avoid trivial solutions and preserve diversification.

A separate class of constraints concerns concentration. Even a correctly specified optimization problem may yield overly concentrated portfolios due to statistical properties of parameter estimation or local advantages of specific instruments. Therefore, $C(R(u))$ includes maximum-weight constraints (and, if needed, minimum weights) or aggregate concentration limits. These rules improve interpretability and reduce the risk of obtaining a solution that is difficult to explain or reproduce in practice.

In terms of computational complexity, constraint parameterization is "lightweight," because it reduces to evaluating a finite number of functions or applying a rule set to compute bounds. In typical scenarios, complexity does not depend on the number of assets; it is mainly driven by the number of classes and constraint types. This makes the algorithm suitable for large-scale deployment. Additionally, results $C(R)$ can be cached for discretized values of R, e.g., with step 0.01, since the parameter set is small and the dependence on R is smooth. In such cases, constraints can be retrieved from cache during inference, further reducing costs for large user volumes.

A key property of $C(R(u))$ is controlled stability of personalization. Since $R(u)$ is continuous, it is natural to require that small changes in $R(u)$ do not cause abrupt jumps in the feasible set of portfolios and, consequently, do not lead to disproportionate changes in portfolio weights. This property can be experimentally verified via sensitivity metrics that measure the difference between portfolio weights constructed for $R(u)$ and for $R(u) + \varepsilon$. Beyond stability, $C(R)$ substantially strengthens interpretability: for a specific user, one can directly explain which bounds were applied, why they correspond to the user's risk profile, and how they affected the final optimizer decision.

Thus, Algorithm 2 formalizes personalization as construction of an optimization constraint system based on the risk profile, ensuring reproducibility, explainability, and a controlled response of portfolio recommendations to changes in user parameters.

https://doi.org/10.32782/mathematical-modelling/2026-9-1-16 ISSN 2618-0332 (print), 2618-0340 (online)

Portfolio optimization problem statement

After forming the candidate asset universe A^* and parameterizing the constraint system $C(R(u))$, portfolio construction reduces to solving an optimization problem whose result is the portfolio weight vector $w^*(u)$. Let A^* contain $n^* = |A^*|$ instruments; then the portfolio is represented by a weight vector $w = (w_1, w_2, \ldots, w_n^*)$, where w_i is the share of instrument i in the portfolio. For correct interpretation, a full-investment constraint is typically imposed: the sum of weights equals 1. For most practical scenarios, short selling is assumed absent, leading to non-negativity constraints on weights.

The expected portfolio return can be represented as a linear combination of expected asset returns estimated from a historical window. Let μ be the vector of expected returns for assets in A^*; then expected portfolio return is μTw. Portfolio risk in the classical Markowitz approach [2] is measured via variance given by the quadratic form $wT\Sigma w$, where Σ is the covariance matrix of returns for assets in A^*. Under this formulation, the optimization problem can be posed as minimizing risk for a given expected return or maximizing expected return for a given risk level [3]. For an applied information technology, it is convenient to use a generalized single-objective formulation with a risk–return trade-off parameter that enables controlled adjustment of the optimizer's "risk aversion" at the technology level [5].

One practical formulation is minimizing a function of the form:

$$J(w) = \lambda(wT\Sigma w) - (1 - \lambda)(\mu Tw),$$

where $\lambda \in [0, 1]$ specifies the trade-off between risk and return. For small λ, the optimizer is more return-oriented; for large λ, it is more risk-oriented. Importantly, within the proposed technology personalization may be implemented not only through changing λ, but primarily through constraints $C(R(u))$ that define the feasible solution set for a specific user.

The constraint system $C(R(u))$ includes basic feasibility constraints and personalized structural bounds. Basic constraints include full investment $\sum_i w_i = 1$ and $w_i \geq 0$. Personalized constraints include bounds on total shares by asset classes [12], bounds on the anchor asset share, and concentration constraints, e.g., $w_i \leq w_{\max}$. In general, the feasible set can be written as $\Omega(R(u)) = \{w : w$ satisfies $C(R(u))\}$. The optimization problem is then:

$$w^*(u) = \text{argmin\_}\{w \in \Omega(R(u))\}J(w).$$

If the technology additionally accounts for a practical requirement of limiting the number of instruments in the portfolio (cardinality), the constraint $|\{i : w_i > 0\}| \leq k$ is introduced. This constraint significantly complicates the problem and often requires heuristic or approximate methods. At the current stage, cardinality can be treated as an extension, while the baseline implementation relies on forming A^* (Algorithm 1), which already reduces dimensionality and brings solutions closer to practically convenient portfolios.

Thus, the optimization problem provides a formal mechanism for converting outputs of upstream modules (A^* and $C(R(u))$) into a concrete portfolio weight allocation. The next section considers computational complexity of key stages and quality metrics, including stability metrics and an ablation procedure that enables quantitative assessment of the contribution of algorithmic components to recommendation quality.

Computational complexity assessment and quality metrics

Computational costs of the proposed information technology should be analyzed under two execution modes: a background (batch/offline) contour for market-data preparation and an online contour for personalized portfolio construction for a specific user. This separation is essential in applied systems because it decouples expensive statistical artifact preparation from frequently executed low-latency recommendation steps.

In the background contour, the main computational stage is estimation of statistical dependencies among assets. For $n = |A|$ assets, constructing the correlation matrix *Corr* or covariance matrix Σ requires pairwise computations and has order $O(n2)$ in the number of pairs (for a fixed window length). In practice, this step is performed periodically (e.g., daily or upon market-data updates) and cached, so its cost does not directly affect online recommendation latency.

In the online contour, assuming cached *Corr*/Σ, candidate universe construction A^* (Algorithm 1) reduces to operations over already prepared dependencies. In the worst case, building the similarity graph under the threshold rule is also $O(n2)$, since all pairs must be examined; however, connected-component search runs in $O(|V| + |E|)$ and is substantially cheaper for sparse graphs. If the threshold τ is fixed, caching the graph or even the final A^* further reduces repeated costs when serving many users. Algorithm 2 (constraint parameterization $C(R)$) is computationally lightweight: it evaluates a small number of functions or rules to obtain class and concentration bounds and is practically independent of the number of assets; it can also be cached for discretized values of R. The optimization stage depends on the chosen formulation and solver; within the technology, it is important that optimization is performed in a reduced space $n^* = |A^*|$, i.e., after dimensionality reduction.

Quality evaluation of the proposed technology should account for both traditional portfolio metrics and criteria specific to an applied information system: decision stability, interpretability, and reproducibility. The basic group of metrics includes expected portfolio return and risk. Let w be the weight vector, μ the expected-return vector, and Σ the covariance matrix. Then expected portfolio return is μTw, and risk (variance) is $wT\Sigma w$ [6]. If the Sharpe ratio is used [13], it can be computed as the ratio of expected excess return to standard deviation; however, for comparing technology configurations it is sufficient to use a consistent computation protocol on the same dataset.

A key metric for the personalized technology is stability with respect to the risk profile. Since personalization uses a continuous parameter $R(u)$, it is important to measure how much the portfolio changes under a small change in R. Let $w(R)$ be the optimal portfolio for a given risk value R; stability can then be measured via norms of weight differences, e.g., $L1$ or $L2$:

$$Stability\_\{L1\}(R, \varepsilon) = ||w(R + \varepsilon) - w(R)||\_1 \text{ or}$$

$$Stability\_\{L2\}(R, \varepsilon) = ||w(R + \varepsilon) - w(R)||\_2.$$

Smaller values correspond to smoother and more controllable personalization, which is important for user trust and practical deployment. In the minimal experimental stage of this paper, stability is also assessed at the constraint level via ΔC, which reflects smoothness of the parameterization $C(R)$.

In addition to stability, portfolio concentration is useful as an indicator of interpretability and practical usability. A simple option is the maximum weight $\max_i w_i$ and the number of instruments with non-zero weight. Aggregate concentration indices can also be used, but for comparing technology configurations it is sufficient to apply a consistent metric reflecting whether the optimizer concentrates weight into one instrument or distributes it more evenly.

Finally, an ablation evaluation procedure is important to quantitatively determine the contribution of algorithmic modules to final results. In this procedure, configurations with the full set of modules are compared to configurations with individual components disabled (e.g., without forming A^* or without parameterizing $C(R)$). Ablation results are compared using portfolio metrics, stability metrics, and concentration indicators. This provides an engineering-relevant answer to which component improves stability, controllability, and recommendation quality.

Experiments and ablations

The experimental verification aims to quantitatively demonstrate two key aspects of the proposed technology: (1) redundancy in the initial asset set and the effect of search-space reduction via Algorithm 1; (2) controllability of personalization via smooth parameterization of the constraint

https://doi.org/10.32782/mathematical-modelling/2026-9-1-16 ISSN 2618-0332 (print), 2618-0340 (online)

system $C(R)$ in Algorithm 2. In addition, results are considered in an ablation sense as a "before/after" comparison for individual modules (without forming A^* and without parameterizing $C(R)$).

Results of preliminary correlation analysis

At the first stage, correlation dependencies among assets were analyzed using monthly returns. The obtained values confirm substantial redundancy within the equity ETF subset [10], which justifies algorithmic construction of a candidate asset universe. Specifically, the mean absolute correlation within the equity group is 0.784, and between equity and bond instruments it is 0.578. For certain asset pairs, very high correlations are observed (e.g., SHY–SHV: 0.993, VWO–EEM: 0.991, AGG–IEF: 0.989), corresponding to "near-duplicate" instrument behavior and potentially increasing solution instability without prior search-space reduction. Meanwhile, Ukrainian domestic government bonds (OVDP) exhibit near-zero correlation with equity ETFs: mean absolute correlation between equity and OVDP is 0.038, and between bond instruments and OVDP is 0.045. This supports treating OVDP as an "anchor" asset class within the technology and the idea of controlled anchor inclusion depending on the user's risk profile via the constraint system $C(R)$ [11]. In addition, crypto assets show a moderate correlation level with equities (mean absolute correlation 0.384), which is considered when setting upper bounds on crypto allocation in $C(R)$.

Demonstration results for Algorithm 1 ($A \to A^*$)

Algorithm 1 was demonstrated on an equity subset with $n = 17$ assets at threshold $\tau = 0.9$. The initial set exhibited high redundancy: the maximum absolute correlation between asset pairs reached 0.991, and the mean absolute correlation within the group was 0.784. Applying Algorithm 1 (building a threshold similarity graph at τ, finding connected components, and selecting the "central" representative of each component) reduced the candidate set from 17 to 6 assets, with 6 connected components and the largest component size equal to 11, indicating a large cluster of mutually substitutable instruments. After reduction, the mean absolute correlation within A^* decreased to 0.654, and the maximum absolute correlation decreased to 0.828. This confirms redundancy reduction in the candidate space and creates prerequisites for more stable and interpretable optimization at the next stage. In this example, A^* included the following component representatives: EEM, EWJ, VNQ, XLE, XLP, XLU. Numerical reduction results are provided in Table 1.

Table 1

Results of candidate set reduction by Algorithm 1 (τ = 0.9, equity group)

Parameter	Value
$\lvert A\rvert$	17
$\lvert A^*\rvert$	6
Number of connected components	6
Largest component size	11
mean |corr|	0.784 → 0.654
max |corr|	0.991 → 0.828

Demonstration results for Algorithm 2 ($C = C(R)$)

To demonstrate Algorithm 2, we constructed a parameterized constraint system $C(R)$ in which bounds on asset-class shares (equity, bonds), the "anchor" class (OVDP), and the maximum share of crypto assets are defined as smooth (linear) functions of the risk profile $R \in [0, 1]$. At this stage, parameters $C(R)$ are treated as a technology policy that formalizes personalization and ensures a reproducible link between the risk profile and admissible portfolio structure. The choice of OVDP as the "anchor" is empirically justified by the decorrelation of OVDP relative to equities shown in Section 7.1. In the minimal experiment, values $R = 0.2$, 0.5, 0.8 and a small change $\varepsilon = 0.05$ were

https://doi.org/10.32782/mathematical-modelling/2026-9-1-16 ISSN 2618-0332 (print), 2618-0340 (online)

considered. For each R, the constraint systems $C(R)$ and $C(R + \varepsilon)$ were computed and the change in constraint parameters was evaluated using the metric ΔC defined as the sum of absolute changes in bounds. It was obtained that $\Delta C = 0.1175$ for all three R values at the specified ε, which is consistent with smooth parameterization $C(R)$ and confirms controlled constraint variation under a small risk-profile change. Example bounds are given in Table 2.

Table 2

Example of parameterized constraints $C(R)$ and their change under $R \to R + \varepsilon$ ($\varepsilon = 0.05$)

R	OVDP (anchor)	Equity	Crypto, max	ΔC(eps = 0.05)
0.2	0.50..0.82	0.14..0.40	0.03	0.1175
0.5	0.35..0.70	0.28..0.55	0.075	0.1175
0.8	0.20..0.58	0.41..0.70	0.12	0.1175

Conclusions

The paper proposes an approach to a formalized description of an intelligent information technology for personalized investment portfolio optimization based on adaptive risk profiling. The technology is represented as the tuple $IT = \langle D, M, Algo, C, V\rangle$, which separates data and preprocessing, intelligent models, algorithmic modules, the constraint system, and visualization/validation tools, and provides a holistic representation of portfolio construction as a sequence of formal transformations.

Two key algorithmic modules of the technology are developed. The first module forms the candidate asset universe $A \to A^*$ based on correlation (or covariance) dependencies, with support for background preparation and caching of statistical artifacts. The second module formalizes personalization via parameterization of the optimization constraint system $C = C(R)$ as a function of the user's continuous risk profile, ensuring a reproducible and explainable link between estimated risk and admissible portfolio structure. The portfolio optimization problem statement is also provided, integrating outputs of upstream modules into a unified mechanism for obtaining the portfolio weight vector.

A technology-quality evaluation approach is proposed that, in addition to traditional portfolio metrics, includes stability metrics with respect to risk-profile changes and ablation experiments to quantify the contribution of algorithmic components. This approach allows the technology to be assessed not only as a method of obtaining a portfolio solution, but also as an applied intelligent system with requirements for controllability, interpretability, and reproducibility of personalization.

Future research should extend the experimental base, compare the proposed approach with alternative methods for candidate-universe formation and constraint parameterization, and investigate cardinality constraints (limits on the number of instruments in a portfolio) and their efficient implementation within the proposed technology [14].

Bibliography

1. Bodie Z., Kane A. J., Marcus A. J. *Investments*. New York : McGraw-Hill Education, 2014.
2. Markowitz H. Portfolio selection. *The Journal of Finance*. 1952. Vol. 7, No. 1. P. 77–91. DOI: https://doi.org/10.1111/j.1540-6261.1952.tb01525.x
3. Merton R. C. An analytic derivation of the efficient portfolio frontier. *Journal of Financial and Quantitative Analysis*. 1972. Vol. 7, No. 4. P. 1851–1872. DOI: https://doi.org/10.2307/2329621
4. Fabozzi F. J., Kolm P. N., Pachamanova D. A., Focardi S. M. *Robust portfolio optimization and management*. Hoboken : Wiley, 2007. DOI: https://doi.org/10.1002/9781119202172
5. Grinold R. C., Kahn R. N. *Active portfolio management: A quantitative approach for producing superior returns and controlling risk*. New York : McGraw-Hill, 2000.
6. Meucci A. *Risk and asset allocation*. Berlin : Springer, 2005.
7. Kozub N. O., Korniienko O. S., Dorenskyi O. P. Method and technological solution of an AI-based adaptive investor survey service for determining an individual risk profile. *Централь-*

ноукраїнський науковий вісник. Технічні науки. 2025. Вип. 11(42), ч. ІІ. DOI: https://doi.org/10.32515/2664-262X.2025.11(42).2.3-10
8. Hastie T., Tibshirani R., Friedman J. *The elements of statistical learning: data mining, inference, and prediction*. 2nd ed. New York : Springer, 2009. DOI: https://doi.org/10.1007/978-0-387-84858-7
9. Breiman L. Random forests. *Machine Learning*. 2001. Vol. 45. P. 5–32. DOI: https://doi.org/10.1023/A:1010933404324
10. Корнієнко О. С., Козуб Н. О. Емпірична оцінка «точки насичення» диверсифікації в ETF-портфелях для практичних рекомендацій. *Сучасні комп'ютерні системи та технології : матеріали VIII Всеукр. наук.-практ. інтернет-конф. студентів, аспірантів та молодих вчених* (24 листопада 2025 р., м. Херсон, м. Хмельницький) / за ред. А. А. Григорової. Херсон : ФОП Вишемирський В. С., 2025. С. 154. DOI: https://doi.org/10.5281/zenodo.17711825
11. Козуб Н. О., Корнієнко О. С. Сучасні IT-рішення для інвестиційних портфелів: локалізація для України. *Вісник Херсонського національного технічного університету*. 2024. DOI: https://doi.org/10.35546/kntu2078-4481.2024.4.37
12. Tobin J. Liquidity preference as behavior toward risk. *The Review of Economic Studies*. 1958. Vol. 25, No. 2. P. 65–86. DOI: https://doi.org/10.2307/2296205
13. Sharpe W. F. Mutual fund performance. *The Journal of Business*. 1966. Vol. 39, No. 1. P. 119–138. DOI: https://doi.org/10.1086/294846
14. Black F., Litterman R. Global portfolio optimization. *Financial Analysts Journal*. 1992. DOI: https://doi.org/10.2469/faj.v48.n5.28

References

1. Bodie, Z., Kane, A. J., & Marcus, A. J. (2014). Investments (10th ed.). New York : McGraw-Hill Education [in English].
2. Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77–91. https://doi.org/10.1111/j.1540-6261.1952.tb01525.x [in English].
3. Merton, R. C. (1972). An analytic derivation of the efficient portfolio frontier. *Journal of Financial and Quantitative Analysis*, *7*(4), 1851–1872. https://doi.org/10.2307/2329621 [in English].
4. Fabozzi, F. J., Kolm, P. N., Pachamanova, D. A., & Focardi, S. M. (2007). *Robust portfolio optimization and management*. Wiley, 2007. https://doi.org/10.1002/9781119202172 [in English].
5. Grinold, R. C., & Kahn, R. N. (2000). Active portfolio management: *A quantitative approach for producing superior returns and controlling risk* (2nd ed.). New York : McGraw-Hill [in English].
6. Meucci, A. (2005). Risk and asset allocation. Berlin; Heidelberg: Springer [in English].
7. Kozub, N. O., Korniienko, O. S., & Dorenskyi, O. P. (2025). Method and technological solution of an AI-based adaptive investor survey service for determining an individual risk profile. *Central Ukrainian Scientific Bulletin. Technical Sciences*, *11*(42), Part II. https://doi.org/10.32515/2664-262X.2025.11(42).2.3-10 [in English].
8. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: *Data mining, inference, and prediction* (2nd ed.). Springer. [in English].
9. Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32. https://doi.org/10.1023/A:1010933404324 [in English].
10. Korniienko, O. S., & Kozub, N. O. (2025). Empirychna otsinka "tochky nasychennia" dyversyfikatsii v ETF-portfeliakh dlia praktychnykh rekomendatsii [Empirical estimation of the "saturation point" of diversification in ETF portfolios for practical recommendations]. In A. A. Hryhorova (Ed.), *Suchasni kompiuterni systemy ta tekhnolohii: materialy VIII Vseukrainskoi naukovo-praktychnoi internet-konferentsii studentiv, aspirantiv ta molodykh vchenykh* [*Modern computer systems and technologies*: *Proceedings of the VIII All-Ukrainian*

Scientific and Practical Internet Conference of students, postgraduate students, and young scientists] (p. 154). Kherson. https://doi.org/10.5281/zenodo.17711825 [in Ukrainian].
11. Kozub, N. O., & Korniienko, O. S. (2024). Suchasni IT-rishennia dlia investytsiinykh portfeliv: lokalizatsiia dlia Ukrainy [Modern IT solutions for investment portfolios: localization for Ukraine]. *Visnyk Khersonskoho natsionalnoho tekhnichnoho universytetu* [*Bulletin of Kherson National Technical University*], *4*(91), 286–290. https://doi.org/10.35546/kntu2078-4481.2024.4.37 [in Ukrainian].
12. Tobin, J. (1958). Liquidity preference as behavior toward risk. *The Review of Economic Studies*, 25(2), 65–86 https://doi.org/10.2307/2296205 [in English].
13. Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business*, *39*(1), 119–138. https://doi.org/10.1086/294846 [in English].
14. Black, F., & Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal*, *48*(5), 28–43. https://doi.org/10.2469/faj.v48.n5.28 [in English].

Korniienko Oleksandr Serhiiovych – Postgraduate Student at the Department of Software and Technologies of the Kherson National Technical University. E-mail: aleksandrkornienko19992106@gmail.com, ORCID: 0009-0008-6234-784X.

Kozub Nataliia Oleksandrivna – Candidate of Technical Sciences, Associate Professor of the Kherson National Technical University. E-mail: actinis1@gmail.com, ORCID: 0000-0002-0406-0161.

Корнієнко Олександр Сергійович – аспірант кафедри програмних засобів і технологій Херсонського національного технічного університету. E-mail: aleksandrkornienko19992106@gmail.com, ORCID: 0009-0008-6234-784X.

Козуб Наталія Олександрівна – к.т.н., доцент, доцент кафедри програмних засобів і технологій Херсонського національного технічного університету. E-mail: actinis1@gmail.com, ORCID: 0000-0002-0406-0161.

Дата першого надходження статті до видання: 30.03.2026
Дата прийняття статті до друку після рецензування: 05.05.2026
Дата публікації (оприлюднення) статті: 01.07.2026