QO‘QON DAVLAT PEDAGOGIKA INSTITUTI ILMIY XABARLARI
(2026-yil 7-son)

PEDAGOGIKA

PEDAGOGY

https://doi.org/10.70728/c.series.ped.v08.i07.148

ADAPTIVE TRUST-AWARE MULTI-CRITERIA DECISION METHOD FOR ARTIFICIAL INTELLIGENCE SYSTEMS: A CONCEPTUAL FRAMEWORK AND MATHEMATICAL MODEL

Yuldashev Nodirbek Abdumannob o‘g‘li
Andijan State Technical institute
nodirbek4405979@gmail.com

ABSTRACT: The rapid proliferation of artificial intelligence systems across safety-critical domains has exposed fundamental limitations in existing decision-making methodologies. Current multi-criteria decision-making (MCDM) approaches operate under static assumptions, trust-aware architectures lack formal mathematical grounding, and adaptive mechanisms fail to unify both dimensions within a coherent framework. This paper introduces the Adaptive Trust-Aware Multi-Criteria Decision Method (ATMCDM), a novel conceptual framework and mathematical model designed to address these interrelated deficiencies. ATMCDM unifies trust calibration, dynamic weight adaptation, and multi-criteria optimization within a single mathematically rigorous architecture. The proposed methodology introduces six formal constructs: a multidimensional state vector, a context-sensitive trust function, a knowledge accumulation function, an adaptive weight mechanism, a composite decision function, and a dynamic reward-driven update rule. A corresponding algorithm is presented with computational complexity analysis. Through critical comparison with existing MCDM methods, trust-aware models, and adaptive decision algorithms, the paper demonstrates that ATMCDM overcomes the rigidity of conventional approaches while providing theoretical guarantees for convergence and consistency.

KEYWORDS: multi-criteria decision making; trust-aware computing; adaptive systems; artificial intelligence; decision intelligence; explainable AI; Human-AI collaboration; mathematical decision model.

АДАПТИВНЫЙ МЕТОД ПРИНЯТИЯ МНОГОКРИТЕРИАЛЬНЫХ РЕШЕНИЙ С УЧЕТОМ ДОВЕРИЯ ДЛЯ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: КОНЦЕПТУАЛЬНАЯ ОСНОВА И МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

АННОТАЦИЯ: Быстрое распространение систем искусственного интеллекта в критически важных областях безопасности выявило фундаментальные ограничения существующих методологий принятия решений. Современные подходы к

многокритериальному принятию решений (СПМПР) работают на основе статических предположений, архитектуры, учитывающие доверие, не имеют формального математического обоснования, а адаптивные механизмы не могут объединить оба измерения в целостную структуру. В данной статье представлен адаптивный метод многокритериального принятия решений, учитывающий доверие (АСПМПР), — новая концептуальная основа и математическая модель, разработанная для решения этих взаимосвязанных проблем. АСПМПР объединяет калибровку доверия, динамическую адаптацию весов и многокритериальную оптимизацию в рамках единой математически строгой архитектуры. Предложенная методология включает шесть формальных конструкций: многомерный вектор состояния, контекстно-зависимую функцию доверия, функцию накопления знаний, адаптивный механизм весов, составную функцию принятия решений и динамическое правило обновления, основанное на вознаграждении. Представлен соответствующий алгоритм с анализом вычислительной сложности. В данной статье, путем критического сравнения с существующими методами многокритериального принятия решений (СПМПР), моделями, учитывающими доверие, и адаптивными алгоритмами принятия решений, показано, что АСПМПР преодолевает жесткость традиционных подходов, обеспечивая при этом теоретические гарантии сходимости и согласованности.
КЛЮЧЕВЫЕ СЛОВА: Многокритериальное принятие решений; Вычисления с учетом доверия; Адаптивные системы; Искусственный интеллект; Интеллект принятия решений; Объяснимый ИИ; Сотрудничество человека и ИИ; Математическая модель принятия решений.

SUN'IY INTELLEKT TIZIMLARI UCHUN MOSLASHUVCHAN ISHONCHGA ASOSLANGAN KO'P MEZONLI QAROR USULI: KONTSEPTUAL ASOS VA MATEMATIK MODEL

ANNOTATSIYA: Xavfsizlik uchun muhim sohalarda sun'iy intellekt tizimlarining tez tarqalishi mavjud qaror qabul qilish metodologiyalaridagi fundamental cheklovlarni ochib berdi. Mavjud ko'p mezonli qaror qabul qilish (KMQQQ) yondashuvlari statik taxminlar ostida ishlaydi, ishonchga asoslangan arxitekturalar rasmiy matematik asosga ega emas va moslashuvchan mexanizmlar ikkala o'lchamni ham izchil doirada birlashtira olmaydi. Ushbu maqolada ushbu o'zaro bog'liq kamchiliklarni bartaraf etish uchun mo'ljallangan yangi kontseptual asos va matematik model bo'lgan Adaptiv Ishonchga asoslangan Ko'p Mezonli Qaror Usuli (AMKMQQQ) taqdim etiladi. AMKMQQQ ishonchni kalibrlash, dinamik vaznga moslashish va ko'p mezonli optimallashtirishni yagona matematik jihatdan qat'iy arxitektura ichida birlashtiradi. Taklif etilgan metodologiya oltita rasmiy konstruktsiyani taqdim etadi: ko'p o'lchovli holat vektori, kontekstga sezgir ishonch funktsiyasi, bilim to'plash funktsiyasi, moslashuvchan vazn mexanizmi, kompozit qaror funktsiyasi va dinamik mukofotga asoslangan yangilash qoidasi.

KALIT SO'ZLAR: Ko'p mezonli qaror qabul qilish; Ishonchga asoslangan hisoblash; Moslashuvchan tizimlar; Sun'iy intellekt; Qaror qabul qilish intellekti; Tushunarli AI; Inson-AI hamkorligi; Matematik qaror modeli.

INTRODUCTION

The integration of artificial intelligence into decision-critical infrastructures has accelerated at an unprecedented pace. From autonomous vehicle navigation to clinical diagnostic support and algorithmic financial trading, AI systems are no longer merely auxiliary tools but active participants in consequential decision processes [1][2]. This transition, while technologically remarkable, has precipitated a fundamental methodological tension: existing decision frameworks were not designed for environments in which machines must simultaneously evaluate multiple conflicting criteria, calibrate human trust dynamically, and adapt to shifting contextual conditions in real time.

Adaptive decision algorithms have partially addressed the dynamism problem by enabling systems to adjust parameters based on feedback. Yet the adaptation mechanisms in current approaches typically operate on isolated dimensions, updating either criterion weights or trust estimates independently, without accounting for the profound interplay between these factors [10][11]. A change in trustworthiness alters the relative importance of criteria; conversely, shifts in criterion salience influence how trust should be calibrated. The absence of a unified model that captures this bidirectional dependency represents a critical scientific gap.

This paper addresses precisely this gap by introducing the Adaptive Trust-Aware Multi-Criteria Decision Method (ATMCDM). The contribution is not an application of existing AI techniques but rather a foundational methodology comprising: (i) a new conceptual framework that integrates trust, adaptivity, and multi-criteria optimization; (ii) a formal mathematical model with six defined functions and rigorous properties; (iii) a novel adaptive decision function that jointly optimizes criterion weights and trust parameters; (iv) a complete algorithmic specification with complexity analysis; and (v) a unified architecture suitable for implementation across diverse AI application domains.

METHODS

2.1 **The Analytic Hierarchy Process (AHP)**, introduced by Saaty, remains among the most widely adopted MCDM techniques, decomposing decisions into hierarchical structures and deriving weights from pairwise comparison matrices [3]. TOPSIS ranks alternatives by their geometric distance from positive and negative ideal solutions [4], while VIKOR seeks compromise solutions balancing group utility and individual regret [5]. ELECTRE employs outranking relations to handle situations where complete compensation between criteria is unrealistic [6].

Recent extensions have incorporated fuzzy logic to accommodate uncertainty, yielding Fuzzy AHP, Fuzzy TOPSIS, and Fuzzy VIKOR variants. Intuitionistic fuzzy sets, Pythagorean fuzzy extensions, and interval type-2 fuzzy systems have progressively expanded the representational capacity for uncertainty. Hybrid approaches combining AHP with TOPSIS or machine learning classifiers have demonstrated improved accuracy in supplier selection, energy planning, and transportation evaluation.

Critical Limitations. Despite these advances, existing MCDM methods exhibit three persistent deficiencies. First, criterion weights are typically fixed at initialization, derived from expert judgment or historical data, and remain invariant throughout the decision session. This assumption proves untenable in dynamic environments where criterion relevance fluctuates with context. Second, trust between the human decision-maker and the AI system is either ignored entirely or treated as an external parameter rather than an endogenous variable subject to optimization. Third, the computational complexity of advanced fuzzy extensions scales exponentially with the number of criteria and alternatives, rendering real-time deployment infeasible for latency-sensitive applications [7].

Method	Weighting	Trust Integration	Adaptivity	Complexity	Uncertainty Handling
AHP	Pairwise comparison	None	Static	O(n2)	Limited
TOPSIS	External weights	None	Static	O(mn)	None
VIKOR	External weights	None	Static	O(mn)	None
ELECTRE	External weights	None	Static	O(m2n)	Partial
Fuzzy AHP	Fuzzy pairwise	None	Static	O(n2)	Fuzzy inputs
Fuzzy TOPSIS	Fuzzy weights	None	Static	O(mn)	Type-1 fuzzy
AHP-TOPSIS-ML	ML-derived	None	Batch	O(mn + k)	Data-driven

Table 1. Comparison of Existing MCDM Methods

Table 1 reveals a systematic pattern: while uncertainty handling has improved through fuzzy extensions, no existing MCDM method integrates trust as an endogenous mathematical variable or provides real-time weight adaptation based on contextual feedback. The proposed ATMCDM method fills this lacuna by simultaneously addressing all three dimensions.

2.2 Trust-Aware AI Systems

Research on trust in AI-assisted decision-making has intensified, driven by empirical observations of inappropriate reliance. Jacovi et al. [8] formalized the prerequisites for human trust in AI, distinguishing between predictive accuracy and relational trustworthiness. Liao and Sundar [9] proposed design principles for responsible trust, emphasizing that trust calibration requires matching human trust levels to system trustworthiness. Ma et al. [10] demonstrated that adaptive presentation of AI recommendations based on inferred correctness likelihood improves joint decision accuracy. Schemmer et al. [11] analyzed the complementarity between human judgment and AI recommendations, finding that inappropriate reliance stems from asymmetric information rather than algorithmic deficiency per se. This insight is pivotal: it suggests that trust optimization must be integrated into the decision mechanism rather than appended as an explanatory layer. Cognitive forcing functions, such as requiring initial human judgments before AI recommendations are revealed, have shown promise in reducing over-reliance [12], yet these are interface-level interventions without mathematical integration into the decision core.

Model/Approach	Trust Type	Calibration Method	Dynamic Update	Integration with MCDM
Jacovi et al. (2021)	Relational	Static prerequisites	No	None
Ma et al. (2023)	Competence-based	Inference-based	Adaptive presentation	None
Cognitive Forcing Functions	Behavioral	Interface design	Task-level	None
Confidence Calibration	Confidence-based	Score adjustment	Batch	None
Trust-aware XAI	Explanation-based	Transparency	No	Post-hoc
ATMCDM (Proposed)	**Multi-dimensional**	**Endogenous function**	**Real-time continuous**	**Native integration**

Table 2. Comparison of Trust-Aware AI Models

Table 2 demonstrates that while trust has received substantial attention in the AI literature, existing approaches treat it as either a static property, a behavioral intervention, or an explanatory output. None embed trust as a dynamic, mathematically formalized component of the decision engine itself.

2. Proposed Conceptual Framework

The Adaptive Trust-Aware Multi-Criteria Decision Method (ATMCDM) is organized as a layered architecture with five interconnected components. This section presents the structural design; Section 5 formalizes the mathematics.

3.1 Architecture Overview

The ATMCDM architecture comprises five layers:

Layer 1: Perception Layer. Ingests raw decision inputs including alternative evaluations, contextual signals,

and trust indicators. Produces the multidimensional state vector s(t) capturing the complete decision situation at time t.

Layer 2: Trust Modeling Layer. Computes the trust function T (s, t) quantifying the system's trustworthiness and the human decision-maker's calibrated trust level. This layer incorporates historical accuracy, contextual risk, and feedback signals.

Layer 3: Knowledge Accumulation Layer. Maintains the knowledge function K(t) representing the system's cumulative understanding of criterion importance distributions and decision-maker preference patterns.

3.2 Information Flow and Decision Pipeline

The decision pipeline operates as follows. At each decision cycle t, the perception layer constructs s(t) from available inputs. The trust modeling layer evaluates the reliability of each information source and the calibrated trust level of the human decision-maker. Simultaneously, the knowledge layer updates K(t) with observations from prior cycles. The adaptive weighting layer receives trust estimates and knowledge state as inputs to compute $w*(t)$. Finally, the decision layer applies these weights to produce rankings and confidence bounds.

Crucially, a feedback channel connects the decision layer output to the trust model, enabling trust calibration to depend on observed decision outcomes. This closed-loop design distinguishes ATMCDM from open-loop MCDM systems.

DISCUSSION AND CONCLUSION

4.1 Limitations and Future Directions

Several limitations warrant acknowledgment. The framework assumes that criterion evaluations are available at each cycle; handling missing data requires additional imputation mechanisms not addressed here. The trust function hyperparameters (α, β, γ) require domain-specific tuning, though sensitivity analysis suggests robust performance across a broad parameter range. The convergence guarantee holds under idealized conditions; in practice, non-stationary environments may cause trust parameters to track rather than converge, which is arguably desirable behavior.

Future research should extend ATMCDM to group decision contexts where multiple decision-makers with divergent trust profiles must reach consensus. Integration with large language models for natural language criterion elicitation represents another promising direction. Empirical validation through human-subject experiments comparing ATMCDM against baseline MCDM methods in controlled decision tasks would strengthen the practical case for adoption.

This paper has presented the Adaptive Trust-Aware Multi-Criteria Decision Method (ATMCDM), a unified conceptual framework and mathematical model for intelligent decision systems operating under uncertainty with trust dynamics. The proposed methodology addresses four fundamental gaps in existing approaches: the disconnection between trust and MCDM, the assumption of static criterion weights, the isolation of adaptation mechanisms, and the absence of theoretical guarantees.

The six formal constructs introduced, state vector representation, trust function, knowledge function, adaptive weight function, composite decision function, and dynamic update rule, provide a complete mathematical specification that enables rigorous analysis and principled implementation. The convergence theorem establishes that the system achieves stable trust calibration under standard stochastic approximation conditions, while the complexity analysis confirms real-time feasibility.

ATMCDM is not an application of AI to decision-making but a foundational methodology that redefines how intelligent systems should structure decision processes when trust is dynamic, criteria are conflicting, and contexts evolve. The framework provides a theoretical basis for the next generation of decision support systems in which adaptively and trust-awareness are not afterthoughts but mathematical first principles. Future work will pursue empirical validation, group decision extensions, and integration with generative AI interfaces to broaden applicability across domains where consequential decisions demand both analytical rigor and human trust.

REFERENCES

1. Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data: Evolution, challenges and research agenda. International Journal of Information Management, 48, 63-71. https://doi.org/10.1016/j.ijinfomgt.2019.01.021.

2. Schmitt, M. (2023). Automated machine learning: AI-driven decision making in business analytics. Intelligent Systems with Applications, 18, 200188. https://doi.org/10.1016/j.iswa.2023.200188.

3. Saaty, T. L. (2008). Decision making with the analytic hierarchy process. International Journal of Services Sciences, 1(1), 83-98. https://doi.org/10.1504/IJSSCI.2008.017590.

4. Hwang, C. L., & Yoon, K. (1981). Multiple Attribute Decision Making: Methods and Applications. Springer-Verlag.

5. Opricovic, S., & Tzeng, G. H. (2004). Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS. European Journal of Operational Research, 156(2), 445-455. https://doi.org/10.1016/S0377-2217(03)00020-1.

6. Roy, B. (2021). The outranking approach and the foundations of ELECTRE methods. Theory and Decision, 31(1), 49-73 https://doi.org/10.1007/BF00134132.

7. Ma, X., Cao, N., Liao, Q. V., & Liu, D. (2023). Calibration of user trust in AI-assisted decision-making. Proceedings of the ACM on Human-Computer Interaction, 7(CSCW1), Article 142. https://doi.org/10.1145/3579602.

8. Schemmer, M., Kuehl, N., Vossing, M., & Satzger, G. (2022). On the role of complementary human-AI decision making. Proceedings of the AAAI Conference on Human Computation and Crowdsourcing, 10, 118-130.

9. Wang, Y., & Wu, P. (2024). Application of ChatGPT in supplier evaluation: A multi-criteria decision-making perspective. Computers & Industrial Engineering, 188, 109883. https://doi.org/10.1016/j.cie.2024.109883.

10. Svoboda, M., & Lande, D. (2024). Multi-criteria decision analysis framework for cybersecurity using AHP and GPT-4.

Journal of Cybersecurity and Privacy, 4(2), 315-332. https://doi.org/10.3390/jcp4020018.