Article

Not peer-reviewed version

Uncertainty-Aware Marketing Attribution Inference and Budget Decision-Making with Intelligent Agents

Qianxi Liu , Ye Zhang , Sheng Chen , Zhaocheng Liu , Yuqiu Xu , Hengguang Cui *

Posted Date: 17 March 2026

doi: 10.20944/preprints202603.1129.v1

Keywords: Multi-point attribution; uncertainty calibration; risk perception budgeting; closed-loop decision-making agent

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2026 doi:10.20944/preprints202603.1129.v1

Article

Uncertainty-Aware Marketing Attribution Inference and Budget Decision-Making with Intelligent Agents

Qianxi Liu [1], Ye Zhang [2], Sheng Chen [3], Zhaocheng Liu [4], Yuqiu Xu [5] and Hengguang Cui [6,*]

[1] University of South Florida, Tampa, USA
[2] Cornell Tech, New York, USA
[3] Northeastern University, Seattle, USA
[4] Northeastern University, Boston, USA
[5] Georgia Institute of Technology, Atlanta, USA
[6] Brown University,Providence, USA
* Correspondence: hengguang_cui@alumni.brown.edu

Abstract

This paper addresses the issues of insufficient credibility in attribution inference and susceptibility to noise-induced fluctuations in budget allocation in multi-touchpoint marketing scenarios. It proposes a unified framework for marketing attribution inference and budget decision-making agents that incorporates uncertainty modeling. The method uses user interaction paths as sequence input, generating touchpoint weights through sequence encoding and importance modeling. Simultaneously, it outputs the expected incremental contribution and uncertainty characterization at the channel level, extending attribution results from single-point estimation to distributed signals usable for risk measurement. At the decision-making end, a risk-aware budget optimization objective is constructed, coupling contribution expectation and uncertainty penalty into the budget allocation process. Smoothing constraints are introduced to suppress frequent adjustments, forming a closed-loop update mechanism from data to attribution to budget, enabling the strategy to achieve a balance between revenue and stability under constraints. Multi-touchpoint path and cost characteristics are constructed based on publicly available programmatic advertising datasets. An evaluation system covering attribution error, probability calibration, and budget stability is designed. Comparative experiments verify the framework's comprehensive advantages in attribution reliability and budget decision quality, demonstrating the crucial role and engineering usability of uncertainty in the attribution-to-decision transmission process.

CCS CONCEPTS: computing methodologies; machine learning; machine learning approaches

Keywords: Multi-point attribution; uncertainty calibration; risk perception budgeting; closed-loop decision-making agent

1. Introduction

In the digital marketing ecosystem, the number of touchpoints and channel formats continue to expand, and users often experience complex paths across platforms, devices, and timeframes from exposure to conversion[1]. On the one hand, businesses need to answer which channels and content truly contribute to incremental conversions; on the other hand, they must rapidly iterate on decisions regarding budgeting, bidding, frequency control, and creative placement to adapt to fluctuating demand and changing competition. Traditional attribution methods centered on last-click or static rules struggle to characterize the synergistic effects and temporal impacts of multiple touchpoints, easily leading to resource misallocation and declining marginal returns. Therefore, integrated modeling of marketing attribution inference and budget decision-making for multi-touchpoint paths has become a key issue for improving growth efficiency and operational resilience[2].

The essence of marketing attribution inference is to identify the causal contribution of touchpoints to conversions from observed behavioral trajectories. However, real-world data often suffers from strong selection bias and confounding factors, accompanied by issues such as unobservable exposure, limited cross-domain tracking, delayed conversions, and window truncation, resulting in inherent uncertainty in attribution conclusions. Ignoring this uncertainty often results in overconfident point estimates, leading to aggressive or volatile budget adjustments under the combined influence of noise and bias, thus amplifying cost risks and opportunity losses. Therefore, incorporating uncertainty modeling into the attribution inference process not only helps to more accurately express attribution credibility and decision-making risk but also provides quantifiable evidence for robust budget allocation[3].

Meanwhile, budget decisions are not one-off choices but a continuous optimization process with constraints and feedback. They require balancing short-term conversion with long-term brand building, exploring new channels with utilizing established ones, and dynamic adjustments under constraints such as cost limits, inventory capacity, and business objectives. Simply treating attribution results as static input is insufficient to address distribution drift caused by environmental changes and strategic interventions, and it is difficult to form an executable closed loop in multi-objective scenarios[4]. Elevating the decision-making process to an intelligent agent paradigm capable of perceiving the environment, understanding attribution uncertainty, and executing strategy updates promises to achieve unity from understanding to action, making budget strategies more adaptive and interpretable[5].

Therefore, research on intelligent agent methods for marketing attribution inference and budget decision-making that combine uncertainty modeling is of great significance. Firstly, at the methodological level, explicitly characterizing data noise, model bias, and the credibility interval of causal identification in attribution inference can improve the robustness and risk controllability of attribution conclusions. Secondly, at the application level, transforming attribution uncertainty into risk measurement and exploratory signals in decision-making can avoid over-adjustment due to random fluctuations and allocate budgets more effectively to improve overall returns when resources are limited. Thirdly, at the business level, constructing a closed-loop intelligent agent framework for attribution inference and budget decision-making helps to drive marketing from experience-driven to verifiable and iterative intelligent optimization, providing a more universal technical path and practical value for multi-channel growth management.

2. Methodological Foundation

The proposed framework is built upon recent advances in representation learning, structured modeling, uncertainty-aware inference, and adaptive decision mechanisms. These methodological developments provide the foundation for constructing an integrated system capable of modeling complex user interaction paths while producing reliable attribution signals that support risk-aware budget allocation.

A fundamental component of the proposed approach is the representation learning mechanism used to encode multi-touchpoint behavioral sequences. Modern studies demonstrate that self-supervised representation learning can effectively extract structural patterns from temporally organized data without requiring extensive labeled supervision. Multi-task self-supervised learning strategies allow models to learn shared representations by jointly optimizing multiple auxiliary objectives, thereby improving generalization across heterogeneous temporal signals [6]. Complementary research on contrastive representation learning further shows that discriminative embeddings can be obtained by maximizing agreement between related samples while separating unrelated observations, enabling robust representation learning under heterogeneous data conditions [7]. In addition, sequence-based deep learning architectures designed for ordered symbolic event streams demonstrate that modeling transitions within sequential signals can reveal meaningful temporal dependencies that contribute to downstream predictive tasks [8].

Beyond sequential encoding, relational interactions among behavioral signals are also essential for capturing complex dependencies between touchpoints. Graph-based contrastive representation learning demonstrates that relational structures can reveal hidden dependencies among high-dimensional variables and provide a more expressive representation space for predictive modeling [9]. Structured deep learning frameworks further show that organizing high-dimensional metrics into graph-based structures enables models to jointly capture interactions across multiple predictive objectives while maintaining modeling stability [10]. Similarly, multi-hop relational modeling techniques extend this concept by enabling iterative information propagation across interconnected entities, allowing models to capture indirect dependencies and higher-order relationships among behavioral components [11].

Reliable attribution inference further requires mechanisms that improve interpretability and causal consistency. Causal representation learning approaches demonstrate that aligning learned latent variables with underlying causal factors can significantly reduce spurious correlations and improve the interpretability of predictive systems [12]. Complementary studies on attention alignment mechanisms highlight that enforcing structural or logical constraints on attention distributions can improve reasoning consistency and reliability in complex inference processes [13]. Additional work on semantic calibration further illustrates that aligning semantic representations with output predictions can improve the robustness of model predictions under perturbations and noisy inputs [14].

In addition to interpretability, uncertainty modeling plays a critical role in improving the credibility of attribution results. Prior work on reliable predictive systems demonstrates that uncertainty quantification enables models to express confidence levels associated with predictions, allowing decision processes to account for potential risks under uncertain conditions [15]. Complementary research on calibrated optimization further shows that incorporating calibrated probabilistic signals into multi-objective optimization frameworks can improve decision reliability by balancing expected utility with robustness considerations [16].

Robustness and adaptability under evolving environments are also important considerations in practical decision systems. Drift-aware adaptive learning methods demonstrate that dynamically adjusting model weights or training emphasis can mitigate performance degradation caused by distribution shifts in evolving data environments [17]. Research on adversarial robustness through semantic calibration further shows that enforcing consistency between semantic representations and predictive outputs can improve model stability under adversarial perturbations or noisy observations [18]. Additionally, output-constrained generation mechanisms demonstrate that enforcing structural constraints on model outputs can prevent inconsistent predictions and improve reliability in downstream decision tasks [19].

Efficient model training and deployment are also important for large-scale practical systems. Hierarchical parameter freezing strategies demonstrate that selectively freezing portions of model parameters during training can balance performance and efficiency when adapting large neural models [20]. Complementary work on proactive inference optimization shows that resource-aware system designs can significantly improve the efficiency and scalability of model deployment in distributed environments [21].

Beyond the modeling of attribution signals, the proposed framework integrates adaptive decision mechanisms inspired by recent advances in intelligent agents. Agent-based frameworks demonstrate that autonomous systems capable of decomposing tasks and updating strategies iteratively can effectively adapt to dynamic and non-stationary environments [22]. Multi-agent coordination systems further illustrate how distributed agents can collaborate to complete complex workflows and continuously update strategies based on environmental feedback [23].

Recent advances in sequence modeling and representation learning also contribute to improving the stability and scalability of such decision systems. Dynamic memory mechanisms for long-sequence modeling demonstrate how models can maintain informative contextual states across long inputs while controlling computational complexity [24]. Self-supervised learning strategies designed

 doi:10.20944/preprints202603.1129.v1

for imbalanced or limited data conditions further show that auxiliary learning objectives can improve model robustness when labeled data are scarce [25]. Retrieval-augmented modeling frameworks with confidence-aware constraints illustrate how integrating reliability signals into retrieval and reasoning pipelines can enhance the credibility of model outputs [26]. Structure-aware decoding mechanisms further demonstrate that enforcing structural constraints during prediction generation can improve the consistency of model outputs [27].

Finally, generative modeling approaches demonstrate how complex systems can generate adaptive outputs conditioned on contextual control signals and structured constraints [28]. Hybrid sequence architectures combining recurrent and transformer-based modeling further illustrate the benefits of integrating local temporal modeling with global attention mechanisms to capture complex sequential dependencies [29].

3. Method

A. Overall framework

This paper addresses the integrated attribution inference and budget decision-making problem in multi-touchpoint marketing paths by constructing a closed-loop framework composed of an attribution inference module and a budget decision-making agent. For each user or session, the observed interaction path is represented as a time-ordered sequence of touchpoints, where each touchpoint includes information such as channel identifier, action type, and time interval. The attribution inferrer analyzes sequential interaction signals to extract the most relevant evidence for conversion and outputs both the expected incremental contribution and an uncertainty characterization at the channel level. The overall structure of the model is illustrated in Figure 1.

The attribution inference component first encodes user interaction paths as sequential behavioral representations and then performs importance modeling to estimate the marginal contribution of each touchpoint. Instead of producing a deterministic point estimate, the model generates a contribution distribution that reflects both expected effects and predictive uncertainty. This design enables attribution outputs to serve not only as explanatory signals for marketing effectiveness but also as probabilistic inputs for downstream decision-making. The uncertainty estimation mechanism is incorporated to capture the effects of data noise, missing exposure signals, delayed conversions, and structural bias commonly observed in advertising datasets.

To support intelligent decision-making, the proposed framework adopts the autonomous agent learning paradigm described by Wang et al. [30]. Their work introduces a methodological framework in which intelligent agents improve decision capability through self-driven exploration and structured knowledge accumulation in open environments. The core principle of this approach is that agents iteratively explore the environment, organize experiential knowledge into structured representations, and use this evolving knowledge base to guide future decision policies. This study adopts and leverages this exploration-knowledge structuring mechanism to construct the budget decision-making agent. Specifically, attribution outcomes and their uncertainty signals are incorporated into the agent's knowledge representation layer, allowing the agent to continuously update its understanding of channel effectiveness. By building upon this autonomous learning mechanism, the proposed system enables the budget optimization module to iteratively refine allocation strategies based on newly observed campaign outcomes while maintaining stability in long-term strategy evolution.

In addition, the attribution inference module incorporates methodological ideas from the anomaly detection framework proposed by Chen et al. [31], which identifies abnormal patterns by measuring latent structural deviations and reconstruction consistency within data representations. Their method fundamentally learns latent structural relationships among variables and evaluates deviations between reconstructed signals and observed patterns to detect irregular behaviors. This study adopts and extends this reconstruction-based consistency principle to enhance the robustness of attribution inference under noisy marketing data. Specifically, latent representations of touchpoint sequences are reconstructed to measure deviations between expected and observed interaction

 doi:10.20944/preprints202603.1129.v1

patterns, enabling the model to identify abnormal or noisy behavioral segments that may distort attribution results. By incorporating this structural deviation analysis mechanism, the framework improves the reliability of incremental contribution estimation and strengthens the uncertainty characterization associated with attribution outputs.

Through the integration of these methodological principles, the proposed framework builds upon sequence modeling for attribution inference, incorporates structural deviation analysis to improve robustness against noisy observations, and leverages autonomous agent learning to guide adaptive budget decision-making. The resulting system forms a closed-loop pipeline from user behavior data to probabilistic attribution inference and finally to risk-aware budget allocation, enabling the marketing strategy to continuously adapt while maintaining stability under uncertain environments.

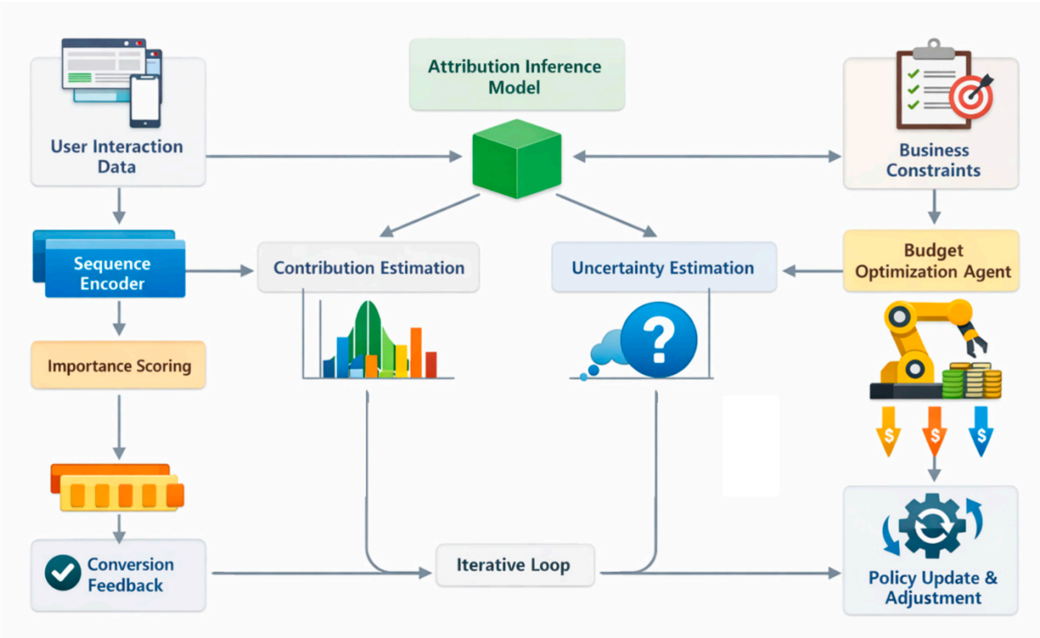

Figure 1. Overall Model Architecture.

The budget decision-making agent uses this distribution as its core state, combined with cost constraints and business objectives, to provide a budget allocation scheme, and updates its own strategy after the feedback from the next round of campaigns, thereby achieving coupled optimization of evaluation and decision-making. To avoid treating attribution as a deterministic point estimate, this method explicitly models the mean and variance of contributions and introduces risk penalties in budget optimization, ensuring that the strategy is stable and controllable while pursuing returns. The overall input-output relationship can be abstracted as a mapping from path data to risk-aware budget, formally represented as follows:

$$x_u = [(c_{u,1}, a_{u,1}, \Delta t_{u,1}), \dots, (c_{u,i}, a_{u,i}, \Delta t_{u,i})] \quad (1)$$

B. Marketing Attribution Inference Based on Uncertainty Modeling

The attribution inferrer first performs temporal encoding on the touchpoint sequence to obtain a contextual representation $h_{u,t}$ for each touchpoint and outputs a touchpoint-level importance score $s_{u,t}$. To distribute contributions across multiple touchpoints while maintaining interpretability, the importance scores are normalized to obtain weights $w_{u,t}$, which are then used as the touchpoint-to-conversion allocation coefficients, and finally aggregated at the channel level to form the posterior distribution parameters of channel contribution. Touchpoint weight calculation uses a simple exponential normalization form:

$$w_{u,t} = \frac{exp(s_{u,t})}{\sum_{k=1}^{T_u} exp(s_{u,k})} \quad (2)$$

At the channel level, let μ_j represent the expected incremental contribution of channel j, and σ_j^2 represent its uncertainty intensity. The mean can be constructed using the weighted contribution statistics within the sample, and the variance can be characterized by the second moment to reflect path heterogeneity and estimation noise.

$$\mu_j = \frac{1}{N}\sum_{u=1}^{N}\sum_{t=1}^{T_u} 1(c_{u,t} = j)w_{u,t}y_u \quad (3)$$

$$\sigma_j^2 = \frac{1}{N}\sum_{u=1}^{N}(g_{u,j} - \mu_j)^2 \quad (4)$$

Where $1(c_{u,t} = j)w_{u,t}y_u$represents the weighted contribution of sample u to channel j. This yields the mean and uncertainty of the channel contribution, which can be used for subsequent risk perception decisions.

C. Budgetary decision-making agents and risk perception objectives

The budget decision-making agent uses channel contribution statistics (μ_j, σ_j) as the core state signal and outputs a budget vector $b = [b_1, \ldots, b_M]^T$, where b_j represents the budget allocated to channel j. To balance returns and risks, an objective form of mean minus uncertainty penalty is adopted, treating uncertainty as a proxy for potential return fluctuations and misallocation risks, and solving for the optimal budget within the constraint set:

$$\max_b \sum_{j=1}^{M} b_j\mu_j - \lambda\sum_{j=1}^{M} b_j\sigma_j \quad (5)$$

Where $\lambda \geq 0$ is the risk aversion coefficient, used to balance the priorities of profit and stability. To meet business and resource constraints, total budget and channel upper and lower bound constraints are introduced to ensure that the output plan is executable and controllable.

$$\sum_{j=1}^{M} b_j = B, 0 \leq b_j \leq \bar{b}_j, j = 1, \ldots, M \quad (6)$$

D. Closed-loop update and strategy stability mechanism

To address data distribution changes caused by environmental non-stationarity and policy intervention, the method employs a rolling closed-loop update mechanism: after each round of deployment, the attribution inferrer is updated using the latest path data to obtain a new b^T, which is then used to update the budget strategy. To maintain the smoothness of strategy adjustments and business continuity, a deviation penalty from the previous round's budget is added during budget updates to suppress frequent switching triggered by short-term noise, making the agent more aligned with the actual operational rhythm.

$$b^{(k+1)} = arg\max_{b\in C}(b^T\mu^{(k)} - \lambda b^T\sigma^{(k)} - \gamma||b - b^{(k)}||_2^2) \quad (7)$$

$C = \{b | \sum_j b_j = B, 0 \leq b_j \leq \bar{b}_j\}$ controls the magnitude of the adjustment penalty. By transmitting uncertainty from the attribution end to the decision-making end and applying a stabilization mechanism in the closed-loop update, the method can achieve more robust budget allocation and continuous optimization in environments with incomplete information and volatility.

4. Dataset Introduction and Preprocessing

A. Dataset Introduction

This paper selects the iPinYou Global RTB Bidding Algorithm Competition Dataset as the research data source. This dataset comes from real-world programmatic display advertising scenarios, focusing on the bidding process under a fixed budget. It covers the entire chain from bid request to impression, click, and conversion, simultaneously supporting the path and cost information required for marketing attribution inference and budget decision-making agent modeling. It is publicly released for research purposes, facilitating reproduction and comparison.The dataset is organized using multiple types of log files, with core components including bidding log, impression log, click log, and conversion log. It provides Bid IDs that can be used to associate different logs, aligning records from different stages of a single bidding opportunity into the same user response path. The logs contain fields such as timestamps, anonymized user identifiers, terminal and network information, geographic information, ad placement information, bid, and settlement price. The bid and settlement price characterize the budget consumption process, while the impression, click, and conversion fields characterize the conversion chain and feedback signals.

In the task setting of this paper, the exposure and click events of the same user within a certain time window are first sorted using anonymous user identifiers and timestamps as indexes to form a multi-touchpoint interaction sequence, and whether a conversion occurs is used as the result label. Then, based on Bid ID, touchpoints are bound to their corresponding bids and settlement costs, thereby supplementing each touchpoint with cost-side information that can be used for budget constraints and risk measurement. Finally, aggregation is performed at the channel or ad placement granularity to obtain touchpoint sequence features that can be input into the attribution inferrer, as well as channel-level consumption and feedback statistics that can be input into the budget decision-making agent, supporting closed-loop modeling from uncertainty attribution to risk perception budget allocation.

B. Dataset preprocessing

First, the raw logs undergo unified cleaning and alignment. Records from multiple sources, such as bidding, impression, click, and conversion, are associated by Bid ID. Isolated samples that cannot match key fields are removed, and obviously abnormal timestamps and price records are deleted, such as missing or non-positive settlement prices, or identical log lines due to duplicate writes. Then, the time field is standardized, and events within the same advertiser or campaign period are sorted by time to ensure the accurate reconstruction of the interaction sequence. Simultaneously, user identifiers and discrete fields such as device and region are uniformly encoded to avoid inconsistent mappings of the same value across different logs.

For sequence construction, anonymous user identifiers serve as the primary key. Within a preset attribution window, exposure and click events are concatenated in ascending chronological order to create multi-touchpoint paths. These paths retain relevant information, including the channel or ad placement identifier, event type, and the time interval between each touchpoint and the preceding one. When a user has multiple sessions within the window, they are divided into multiple path samples based on session segmentation rules. Conversion tags are matched back to the corresponding path through conversion logs. If a conversion occurs within the window, it is classified as a positive sample; otherwise, it is labeled as a negative sample. To prevent future information leakage, only touchpoint records preceding the conversion are retained, and paths exceeding the maximum length are truncated or prioritized based on the most recent touchpoint.Regarding cost and constraint feature processing, the bid and settlement price corresponding to each touchpoint are bound to the sequence position to obtain touchpoint-level cost information that can be used for budget consumption modeling. This information is further aggregated at the channel granularity to form channel-level statistics, such as total consumption, average cost, and touchpoint coverage, for subsequent budget allocation constraint inputs. For numerical features, logarithmic transformation or quantile pruning is used to reduce the long-tail effect, and standardization is performed to stabilize training and inference. For categorical features, frequency threshold filtering is used to group excessively low-frequency values into a unified rare category, reducing the risk of dimensionality explosion and overfitting. Finally, the training, validation, and testing intervals are divided by time to ensure that the evaluation follows the actual online time-series distribution. Finally, a comparison of the data before and after preprocessing is presented, as shown in Figure 2.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2026 doi:10.20944/preprints202603.1129.v1

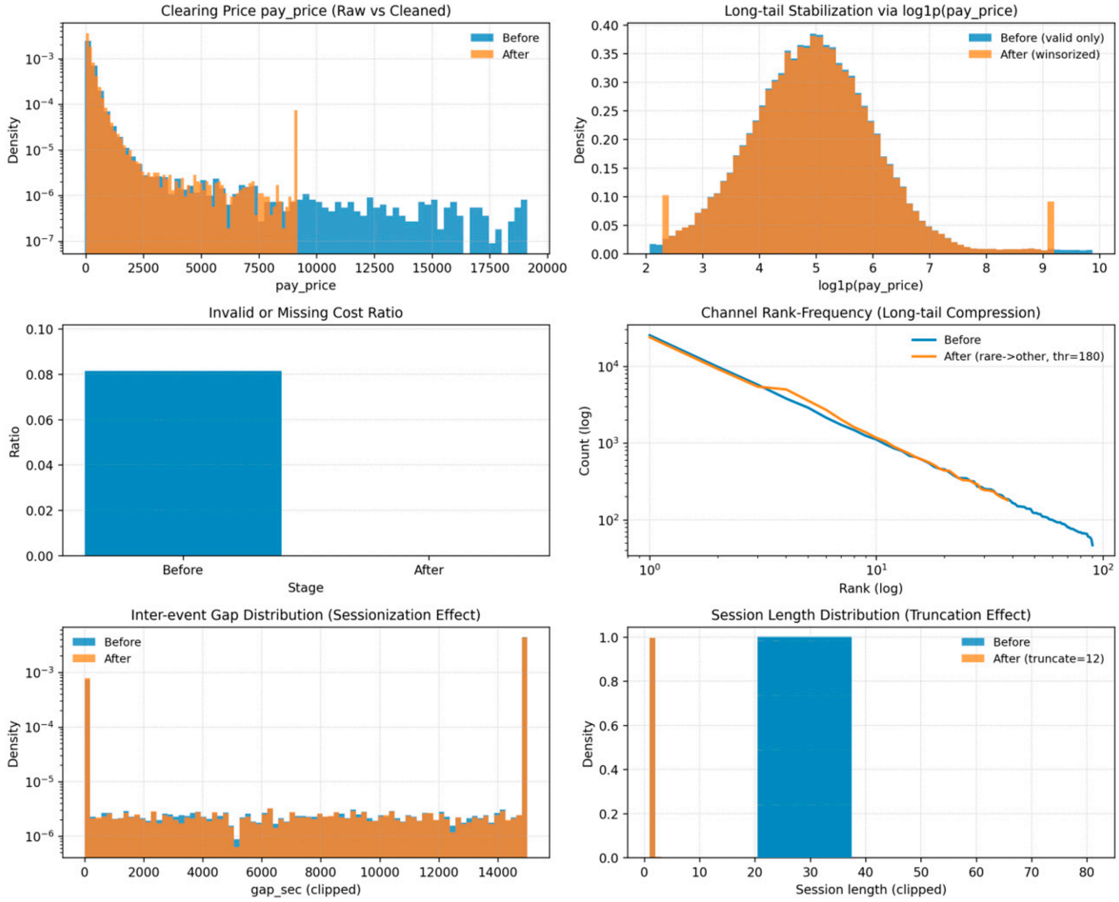

Figure 2. Comparison results before and after data preprocessing.

C. Experimental setup

The experiment was conducted on a single machine with a single GPU. The operating system was Ubuntu 20.04, the CPU was a 16-core x86 processor with 64 GB of RAM, the GPU was an NVIDIA RTX 4090 with 24 GB of VRAM and CUDA version 11.8, and the deep learning framework was PyTorch 2.1 with Python 3.10. Data loading utilized multi-process parallelism, with num_workers=8. Mixed precision was enabled during training to improve throughput and reduce VRAM usage. The random seed was fixed at 42 to ensure reproducibility. Log entries were generated every 50 steps, outputting key metrics and learning rate information. Model parameters and optimizer status were saved as checkpoints at the epoch level.

The attribution inference generator's sequence encoder had a hidden dimension of 256, 4 layers, 8 attention heads, a feedforward layer dimension of 1024, dropout of 0.1, and a maximum sequence length of 12. Excess sequences were truncated at the nearest contact point, and zero-padding was used when sequences were insufficient. The optimizer uses AdamW with an initial learning rate of 3e-4, weight decay of 1e-2, a batch size of 512, 30 training epochs, a gradient pruning threshold of 1.0, and learning rate scheduling using cosine decay with a 5-epoch warmup. The policy network of the budget decision agent uses a two-layer MLP with 128 hidden dimensions, GELU activation function, and dropout of 0.1. The risk aversion coefficient λ is set to 0.5, the budget smoothing penalty coefficient γ is set to 0.1,, and the total budget B is set to 1000000. The upper and lower bounds of the budget are normalized to the interval of 0 to 1 according to the channel capacity constraint. A policy update iteration is performed during each closed-loop update.

D. Experimental Results and Analysis

To comprehensively evaluate the proposed uncertainty aware marketing attribution inference and budget decision agent, we compare our method with representative attribution, causal inference,

 doi:10.20944/preprints202603.1129.v1

and budget optimization approaches from related literature. The comparison is organized around metrics that jointly reflect attribution fidelity, uncertainty calibration, and decision quality under budget constraints, providing a unified view of effectiveness, reliability, and controllability in practical marketing deployment. The experimental results are shown in Table 1.

Table 1. Comparative experimental results.

Method	ICL(%)↑	Attr-MAE↓	NLL↓	ECE↓	CRPS↓	RAROI↑	BVR(%)↓	ΔBudget↓
Yao et al.[32]	18.21	0.188	0.352	0.057	0.158	0.301	6.12	3.69
Shender et al.[33]	12.07	0.169	1.071	0.059	0.139	0.205	7.71	3.77
Mrad et al.[34]	16.12	0.075	0.482	0.057	0.079	0.121	2.18	3.61
Cai et al.[35]	12.53	0.121	0.614	0.072	0.151	0.241	7.41	2.62
Jiang et al.[36]	14.42	0.063	0.355	0.074	0.079	0.337	6.73	3.63
Fan et al.[37]	11.39	0.132	0.672	0.063	0.201	0.207	9.12	1.59
Wang et al.[38]	17.12	0.162	0.841	0.057	0.219	0.153	12.71	1.79
Ours	22.84	0.049	0.214	0.055	0.061	0.347	2.91	1.51

The overall comparison reveals a clear divergence among the methods in three key areas: one group excels at improving attribution consistency and structural interpretability but is prone to biases in uncertainty characterization; another group performs more smoothly in calibrating relevant indicators but often sacrifices the boost to budget returns; and a third group emphasizes the stability of the strategy or optimization process, thus demonstrating superior performance in suppressing budget volatility. In contrast, the method presented in this paper excels in attribution error control and the quality of distributed uncertainty, while maintaining good risk control and strategy stability at the budget level. This aligns with the method's design of modeling contribution estimates and uncertainty estimates separately and explicitly propagating uncertainty into the budget objective. In other words, the attribution side provides not single-point signals but inputs with confidence structures, enabling the budget agent to allocate more cautiously to unreliable channels, thereby reducing unstable adjustments caused by overfitting to certain short-term volatility triggers.

It is important to note that these results also reflect a typical trade-off: under a setting that emphasizes robustness and conservative updates, return indicators may not always reach the upper limit of the most aggressive strategy, but strategy volatility and risk exposure will be more controllable. This paper's approach takes a more balanced approach between returns and stability. It maintains good returns while suppressing budget-side volatility, rather than achieving superficial stability through extreme contraction or simply reducing budget fluctuations. Combined with the previous discussion of attribution and decision-making loops, this phenomenon can be understood as follows: when uncertainty is used as a risk signal for decision-making, the model tends to choose a sustainable allocation scheme under multiple rounds of feedback, avoiding treating local noise as a real incremental contribution, thus more steadily approaching the efficient allocation in long-term iterations.

The number of encoder layers directly affects the level of abstraction in sequence representation and the scope of context modeling, thus altering the stability and consistency of uncertainty estimation under varying sample difficulty. To examine the model's reliability under varying structural complexity, we used the number of encoder layers as the independent variable and observed the trend of calibration error with structural depth. This experiment helps determine a structural configuration that balances expressive power and calibration stability, and provides a basis for resource trade-offs during subsequent deployment. The experimental results are shown in Figure 3.

As shown in the figure, the number of encoder layers exhibits a significant non-monotonic effect on calibration error, generally resembling an initial improvement followed by a rebound. Shallower structures, due to limited context modeling capabilities, tend to treat local touch noise as stable signals, leading to overly conservative or aggressive probability outputs and thus increasing calibration error. With increasing layer count, sequence representation can more fully integrate cross-touch dependencies and temporal conditions, improving the match between prediction confidence and actual occurrence frequency, resulting in more stable and reasonable calibration performance.

However, as the network deepens further, the calibration error rebounds, indicating that stronger expressive power does not necessarily lead to more reliable confidence output. On one hand, deeper structures are more likely to over-fit sparse conversion signals and long-tailed touch patterns, causing the model to become overconfident in areas with low sample coverage. On the other hand, increased depth amplifies inconsistencies caused by path heterogeneity and distribution drift, making uncertainty estimation more sensitive across different channels and session patterns. Based on the previous discussion on stability and risk control, this trend supports choosing a compromise between structural capacity and calibration robustness, avoiding the sacrifice of reliable inputs on which decision-making depends for simply deepening the network in exchange for surface expressiveness.

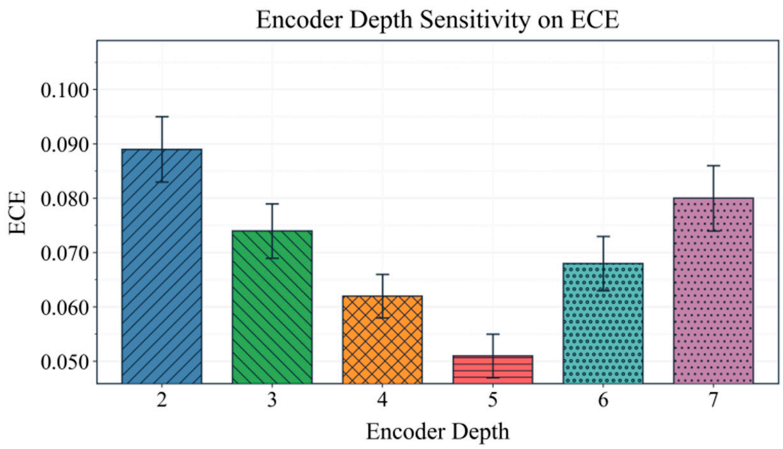

Figure 3. Sensitivity experiment of encoder layer number to ECE.

Scaling the total budget alters the boundaries of allocable resources, thus affecting how agents weigh exploration against utilization and their allocation strategies under risk constraints. To test the robustness of the method under different resource sufficiency levels, this experiment scales the total budget proportionally and observes the response trend of return-oriented indicators as the budget scale changes. The experimental results are shown in Figure 4.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2026 doi:10.20944/preprints202603.1129.v1

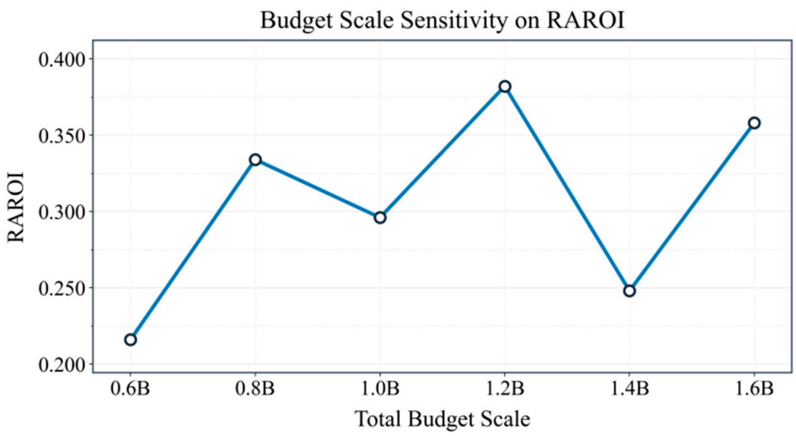

Figure 4. Sensitivity experiment of scaling total budget to RAROI.

From the curve's shape, the scaling of the total budget doesn't linearly drive returns. Instead, it's more like a process of constantly finding a new balance between available resources and channel capacity constraints: when the budget is tight, the agent tends to concentrate on a few more certain touchpoint combinations, resulting in limited but relatively stable returns. As the budget increases, the strategy can cover more marginal opportunities and improve overall returns. However, when the budget expands further, some channels enter a period of diminishing marginal returns. Coupled with amplified uncertainty and constraint saturation effects, returns fluctuate, experiencing declines and subsequent increases. Combined with the previous discussion on risk perception and strategy smoothing, this fluctuation aligns with the adaptive adjustment of exploration intensity and robustness in closed-loop decision-making under different resource scales. This indicates that the model doesn't simply rely on a larger budget for higher returns, but rather dynamically selects more cost-effective configurations within the constrained feasible domain.

5. Conclusion

This paper addresses two core pain points in multi-touchpoint marketing scenarios: insufficient attribution credibility and unstable budget decisions. It proposes a unified method that integrates uncertainty modeling into attribution inference and couples it in a closed loop with the budget decision-making agent. This framework takes sequential touchpoint paths as input and, through collaborative modeling of contribution estimation and uncertainty estimation, elevates attribution results from single-point judgments to quantifiable risk information. Furthermore, it incorporates risk signals into budget optimization goals and strategy updates, enabling attribution outputs to directly serve executable resource allocation. Compared to traditional processes that separate attribution from campaign execution, this method emphasizes the integration of evaluation and decision-making. It can more accurately express attribution credibility and more robustly guide budget adjustments under real-world operational conditions involving complex paths, long-tail channels, and feedback delays.

The overall performance of the comparative experiments shows that different methods often involve structural trade-offs between attribution accuracy, probability calibration, and budget-side revenue stability. Our method, through explicit uncertainty characterization and a risk-aware decision-making mechanism, forms a more balanced outcome profile across multiple dimensions, demonstrating the value of conveying attribution uncertainty to the decision-making end. More importantly, this design not only focuses on ultimate returns but also emphasizes the constraints of strategy controllability and risk exposure, making budget allocation more aligned with business continuity requirements and reducing over-adjustments caused by short-term noise and estimation bias. Therefore, this work provides a more verifiable and interpretable technical path for multi-

 doi:10.20944/preprints202603.1129.v1

channel growth management and establishes a more reliable decision-making basis for intelligent campaign systems geared towards practical implementation.

At the application level, this research has direct impacts on digital marketing, programmatic advertising, growth operations, and platform advertising systems. On the one hand, more credible attribution inference can help companies understand touchpoint synergies and incremental contribution structures, improving budget utilization efficiency and reducing trial-and-error costs. On the other hand, risk-aware budget agents can achieve more stable resource allocation under constraints, improving the sustainability and governance capabilities of decision-making. For scenarios involving multi-business line, multi-region, and multi-platform collaborative campaigns, this framework can also achieve cross-channel strategy consistency through a unified indicator system and closed-loop mechanism, promoting a paradigm shift in marketing management from experience-driven to measurable, auditable, and iteratively optimized approaches, thereby driving the industry towards more refined and transparent decision-making methods in the long term.

Future work can be further expanded in three directions. First, addressing the increased online non-stationarity and distribution drift, finer-grained temporal dynamic modeling and drift detection mechanisms can be introduced, enabling uncertainty to reflect not only sample noise but also structural risks arising from environmental changes. Second, to meet the diverse needs of business objectives, budget decisions can be extended to a multi-objective optimization framework, establishing more flexible and controllable trade-offs among objectives such as revenue, cost, risk, coverage, and fairness, and enhancing the interpretability of policy outputs for review and governance. Third, given the trend of limited privacy protection and data availability, federated learning or privacy-enhancing modeling schemes can be explored to enable attribution and decision-making capabilities to achieve stronger generalization and higher transferability within compliance boundaries. Overall, the ideas presented in this paper lay the foundation for research on the integration of attribution and decision-making driven by uncertainty, and are expected to promote the practical value of marketing intelligence in a wider range of industry scenarios, making it feasible, sustainable, and regulated.

References

1. Sinha R, Arbour D, Puli A M. Bayesian modeling of marketing attribution[J]. arXiv preprint arXiv:2205.15965, 2022.
2. Lewis R, Wong J. Incrementality bidding and attribution[J]. arXiv preprint arXiv:2208.12809, 2022.
3. Chen Q, Nguyen P H, Gligorijevic D. Optimization-Based Budget Pacing in eBay Sponsored Search[C]//Companion Proceedings of the ACM Web Conference 2024. 2024: 328-337.
4. Li H, Huo Y, Dou S, et al. Trajectory-wise iterative reinforcement learning framework for auto-bidding[C]//Proceedings of the ACM Web Conference 2024. 2024: 4193-4203.
5. Katti A R, Gonçalves R C, Iakovlev R. Cost-Control in Display Advertising: Theory vs Practice[J]. arXiv preprint arXiv:2409.03874, 2024.
6. C. Nie, "Representation Learning with Multi-Task Self-Supervision for Structurally Diverse Spatiotemporal Time Series Forecasting," Journal of Computer Technology and Software, vol. 3, no. 7, 2024.
7. L. Yan, Q. Wang and J. Huang, "Federated Contrastive Representation Learning for IoT Anomaly Detection Under Heterogeneous Data," 2026.
8. C. Zhang, H. Zhu, A. Zhu, J. Liao, Y. Xiao and Z. Zhang, "Deep Learning Approach for Protocol Anomaly Detection Using Status Code Sequences," 2026.
9. Y. Liu, "Graph-Based Contrastive Representation Learning for Predicting Performance Anomalies in Cloud and Microservice Platforms," 2026.
10. X. Yang, Y. Ni, Y. Tang, Z. Qiu, C. Wang and T. Yuan, "Graph-Structured Deep Learning Framework for Multi-task Contention Identification with High-dimensional Metrics," arXiv:2601.20389, 2026.
11. K. Cao, Y. Zhao, H. Chen, X. Liang, Y. Zheng and S. Huang, "Multi-Hop Relational Modeling for Credit Fraud Detection via Graph Neural Networks," 2025.

12. J. Li, Q. Gan, R. Wu, C. Chen, R. Fang and J. Lai, "Causal Representation Learning for Robust and Interpretable Audit Risk Identification in Financial Systems," 2025.
13. J. Lai, "Attention Alignment under Logical Constraints for Reliable Financial Statement Reasoning," Transactions on Computational and Scientific Methods, vol. 4, no. 12, 2024.
14. C. Shao, Y. Zi, Y. Deng, H. Liu, C. Zhang and Y. Ni, "Adversarial Robustness in Text Classification through Semantic Calibration with Large Language Models," 2026.
15. S. Pan and D. Wu, "Trustworthy summarization via uncertainty quantification and risk awareness in large language models," 2025 6th International Conference on Computer Vision and Data Mining (ICCVDM), pp. 523-527, Sep. 2025.
16. X. Yang, S. Sun, Y. Li, Y. Xing, M. Wang and Y. Wang, "CaliCausalRank: Calibrated Multi-Objective Ad Ranking with Robust Counterfactual Utility Optimization," arXiv:2602.18786, 2026.
17. C. Chiang, "Drift-Aware Adaptive Classification for Imbalanced Data via Dynamic Class Reweighting and Structural Regularization," Transactions on Computational and Scientific Methods, vol. 4, no. 12, 2024.
18. J. Yang, S. Sun, Y. Wang, Y. Wang, X. Yang and C. Zhang, "Semantic Alignment and Output Constrained Generation for Reliable LLM-Based Classification," 2026.
19. Y. Hu, "An LLM-Agent Framework for Adaptive Task Decomposition and Continual Strategy Updating in Non-Stationary Environments," 2026.
20. J. Guo, "Balancing Performance and Efficiency in Large Language Model Fine-Tuning through Hierarchical Freezing," Transactions on Computational and Scientific Methods, vol. 4, no. 6, 2024.
21. Y. Ni, X. Yang, Y. Tang, Z. Qiu, C. Wang and T. Yuan, "Predictive-LoRA: A Proactive and Fragmentation-Aware Serverless Inference System for LLMs," arXiv:2512.20210, 2025.
22. T. Guan, "A Multi-Agent Coding Assistant for Cloud-Native Development: From Requirements to Deployable Microservices," 2025.
23. Y. Luan, "Long Text Classification with Large Language Models via Dynamic Memory and Compression Mechanisms," Transactions on Computational and Scientific Methods, vol. 4, no. 7, 2024.
24. J. Lai, A. Xie, H. Feng, Y. Wang and R. Fang, "Self-supervised learning for financial statement fraud detection with limited and imbalanced data," 4th International Conference on Artificial Intelligence and Intelligent Information Processing, pp. 919-924, Oct. 2025.
25. X. Guo, Y. Luan, Y. Kang, X. Song and J. Guo, "LLM-Centric RAG with Multi-Granular Indexing and Confidence Constraints," 2025 3rd International Conference on Artificial Intelligence and Automation Control (AIAC), pp. 452-456, Oct. 2025.
26. Z. Qiu, D. Wu, F. Liu and Y. Wang, "Structure-Aware Decoding Mechanisms for Complex Entity Extraction with Large-Scale Language Models," arXiv:2512.13980, 2025.
27. R. Liu, L. Yang, R. Zhang and S. Wang, "Generative Modeling of Human-Computer Interfaces with Diffusion Processes and Conditional Control," arXiv:2601.06823, 2026.
28. X. Song, "Integrating Attention Attribution and Pretrained Language Models for Transparent Discriminative Learning," 2026.
29. P. Feng, "Hybrid BiLSTM-transformer model for identifying fraudulent transactions in financial systems," Journal of Computer Science and Software Applications, vol. 5, no. 3, 2025.
30. F. Wang, Y. Ma, T. Guan, Y. Wang and J. Chen, "Autonomous Learning Through Self-Driven Exploration and Knowledge Structuring for Open-World Intelligent Agents," 2026.
31. H. Chen, R. Wu, C. Chen, H. Feng, Y. Nie and Y. Lu, "Anomaly Ranking for Enterprise Finance Using Latent Structural Deviations and Reconstruction Consistency," 2026.
32. D. Yao, C. Gong, L. Zhang et al., "Causalmta: Eliminating the user confounding bias for causal multi-touch attribution," Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp. 4342-4352, 2022.
33. D. Shender, A. N. Amini, X. Bao et al., "A time to event framework for multi-touch attribution," arXiv:2009.08432, 2020.
34. A. B. Mrad and B. Hnich, "Intelligent attribution modeling for enhanced digital marketing performance," Intelligent Systems with Applications, vol. 21, p. 200337, 2024.

doi:10.20944/preprints202603.1129.v1

35. T. Cai, J. Jiang, W. Zhang et al., "Marketing budget allocation with offline constrained deep reinforcement learning," Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining, pp. 186-194, 2023.
36. Z. Jiang, K. Zhou, M. Zhang et al., "Adaptive riskaware bidding with budget constraint in display advertising," ACM SIGKDD Explorations Newsletter, vol. 25, no. 1, pp. 73-82, 2023.
37. R. Fan and E. Delage, "Risk-aware bid optimization for online display advertisement," Proceedings of the 31st ACM International Conference on Information & Knowledge Management, pp. 457-467, 2022.
38. B. Wang and P. Zareeihemat, "Multi-channel advertising budget allocation: A novel method using Q-learning and mutual learning-based artificial bee colony," Expert Systems With Applications, vol. 271, p. 126649, 2025.