Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 1s (2025)

iROAS-Based Dynamic Bidding Strategy Using Multi-Armed Bandits in Retail Media Networks

Saurabh Mittal

North Carolina State University, Raleigh, NC

saurabhmittalmnit@gmail.com

Article History:

Received: 19-11-2024

Revised: 24-12-2024

Accepted: 21-01-2025

Abstract:

Within the competitive and complicated environment of the retail media networks (RMN), advertisers are experiencing pressure to be more precise in ad spending. Conventional indicators such as Return on Ad Spend (ROAS) are normally inflated to reflect the performance of a campaign by ignoring organic conversions, which encourages practitioners to use incremental ROAS (iROAS) as a more precise metric. This paper will discuss the combination of iROAS and Multi-Armed Bandit (MAB) algorithms to form a dynamic bidding strategy that will evolve in real time. Using causal inference to estimate iROAS and MAB models to trade off exploration and exploitation, advertisers can efficiently optimize budgets and make ongoing progress in media performance. The paper provides a theoretical background, a scalable architecture of implementation, and addresses the problem of delayed incentives, cold-start environments, preservation of privacy, and exploration of many-objective optimization. Overall, it provides a causally responsive method of automated bidding, which causes responsible and value-based advertising in RMNs.

Keywords: iROAS; Multi-Armed Bandits; Retail Media Networks; Dynamic Bidding; Causal Inference

1. Introduction

Digital advertising has greatly changed the manner in which retailers are able to budget and optimize campaigns, especially in Retail Media Networks (RMNs), as illustrated in Figure 1. The platforms have emerged as an important point of contact in the overall marketing strategy, enabling the marketing brands to directly advertise on their platforms. Competition is growing tougher, and inventory is becoming tough to find; therefore, maximizing the efficiency of the ad spending takes center stage. Another fundamental measure of advertising performance is the incremental Return on Ad Spend (iROAS), which attempts to quantify the effectiveness of advertising by quantifying the revenue increment actually due to the ad exposure and not the total increment in revenue. The dependence on the past ROAS measures tends to exaggerate the success of the campaign because it does not consider the baseline or organic sales that would otherwise have been experienced without the intervention of advertising. Therefore, iROAS offers a better and more sophisticated indicator of the performance of the campaign [1-3].

Advertisers in the world of RMNs are not only faced with the difficulty of ensuring accurate measurement of advertising performance but also with the difficulty of dynamically budgeting and allocating the budgets in real time across a range of bidding alternatives. The growing, competitive, and fragmented world requires quick and adaptive decision-making in the allocation of the budget. Multi-armed bandit (MAB) algorithms make a viable solution here. Sequential decision-making problems are best addressed using MAB frameworks, which are derived from reinforcement learning, in which the objective is to balance between exploration (trying out less certain ad placements) and exploitation (capitalizing on known successful channels). When applied to RMNs, a bandit arm in this case might be associated with any of the ad placements or campaigns, and the algorithm will be able to dynamically reallocate the budget resources depending on the iROAS feedback [4-6]. Combining iROAS with MAB algorithms is a paradigm shift in bidding strategies since it is not based on the strict rules or heuristics anymore, but rather on a more intelligent and data-driven dynamic one. Although most modern retail advertising models may be based on historical averages or attribution models to use in making decisions about bids, they usually have a negative response to the dynamic changes in consumer behavior or competition. MAB strategies can learn continuously and optimize budget allocation through incorporating live feedback on the incremental effectiveness of ad spend to improve overall marketing effectiveness. Notably, the iROAS reward signal makes the algorithm not optimize just based on an engagement metric or total revenue, but rather the actual value being brought due to the advertising activity [7-9].

Modern retail analytics platforms can be used to capture high-frequency, granular data important to estimate iROAS and Multi-Armed Bandit (MAB) optimization. Real-time attribution, causal inference, and randomized experiments enable advertisers to isolate the incremental effect of every impact [10-12]. Ingesting this information into adaptive MAB systems creates a feedback loop that can continuously improve the ad allocation to be effective in order to respond quickly to a changing inventory, consumer needs, and diminishing returns [13-18]. The traditional models are unable to perform at this level and speed, so dynamic iROAS-based bidding and the MAB algorithms are much needed to perform campaigns efficiently, accountably, and sustainably.

The joint value of measuring iROAS and using MAB strategies is becoming widely accepted by retailers and advertisers as a means of performance, as well as transparency and accountability. Solutions that are capable of providing measurable incremental lift and dynamically adjusting strategies will be at a competitive advantage [19-21]. In this paper, a framework using iROAS measurement with MAB-based bidding is suggested in Retail Media Networks by using causal models and algorithmic decision-making, as well as real-time data, in order to optimize campaign scalability. The second part will discuss the theoretical basis of iROAS and its benefits over traditional ROAS, which will be the foundation of MAB integration.

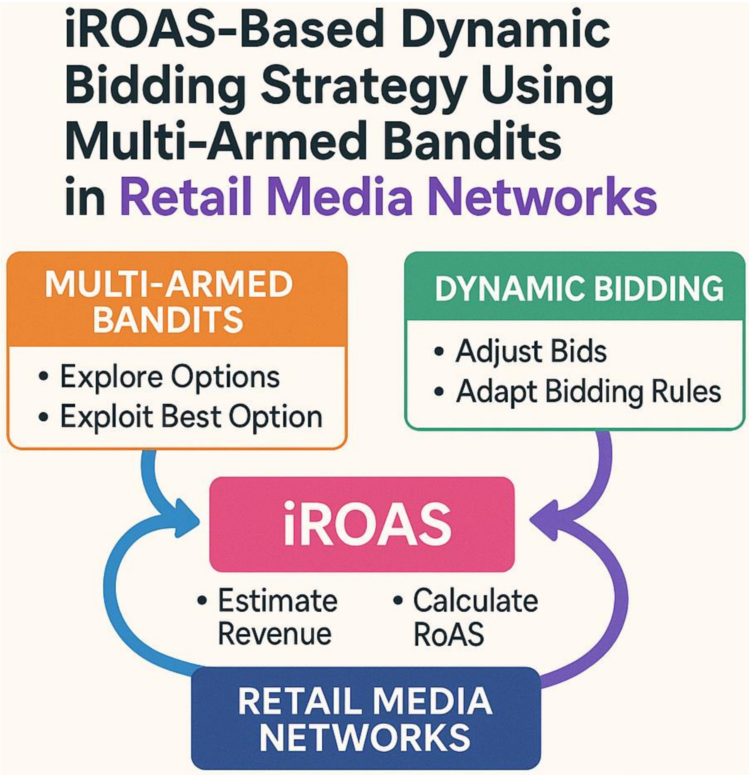

Figure 1: Diagram illustrating the iROAS-based dynamic bidding strategy

2. Understanding iROAS: Beyond Traditional Metrics

As the need to precisely evaluate the effectiveness of advertising increases, incremental Return on Ad Spend (iROAS) comes out as one of the major improvements in comparison to the old evaluation models. iROAS, in contrast with the traditional ROAS, which calculates total revenue divided by the amount of money spent on the advertisement, estimates the incremental revenue that is directly attributed to an advertisement, excluding organic conversions or any baseline sales that would have happened without being exposed to the ad. This optimization essentially redefines value to advertisers, especially when operating within dynamic and competitive marketplaces such as Retail Media Networks (RMNs), where allocation of resources needs to be optimized at scale and in real time [1, 2].

Conventional ROAS may be deceptive in high traffic or brand-dense settings, where organic interest in consumers is already high. To illustrate, in a case when a product is already trending, an advertising campaign may receive unnecessary credit for conversions, which would have taken place on their own. This advertising value inflation results in misallocation of budgets and marginal returns. iROAS addresses this issue by using causal models of inference, such as randomized controlled trials (RCTs), uplift modeling, and counterfactual analysis to isolate the actual effect of an advertising campaign. These approaches give a much better representation of ad effectiveness and, as such, allow smarter decisions when bidding and allocating when combined with algorithmic models such as multi-armed bandits [3-5]. iROAS heavily depends on experimental or quasi-experimental designs in its calculation. A typical method is the ghost ad methodology, where the users are divided according to treatment and control. To one group, an advertisement is shown, as compared to the other group, and the behavioral difference is attributed to the ad. In other cases, propensity score matching may be used in situations where they cannot be randomized and they are able to determine the incremental effect between treated and untreated users by matching them on observable characteristics. These methods give a statistically valid basis for incremental lift estimation, and their impact is reflected in the iROAS metric [6-7].

These mechanisms are strong, but scaling them can be very technically challenging. Because the campaigns in RMNs need to be optimized on thousands of SKUs, locations, and user groups, near-real-time generation of iROAS estimates needs a strong data backbone that is able to ingest and analyze user-level events via many channels. This is where new developments in the stream processing systems and real-time causal inference algorithms have brought significant latency improvement among ad impressions and iROAS feedback to be integrated into the quick response system(s), such as MABs [8-10]. In addition, iROAS provides more accountability and transparency to the advertiser. Unlike opaque attribution models, which have tended to use heuristic attribution rules, like the last-touch or linear attribution, iROAS provides a causally based estimate of impact. This will be of great importance in the justification of budgets to fund departments, cross-channel strategies, and proper performance reviews. RMNs are becoming increasingly competitive in gathering advertiser dollars as brands seek finer performance numbers [11, 12].

From a strategic standpoint, adopting iROAS as a core bidding metric can also mitigate risks associated with overspending on ineffective channels. In many digital ecosystems, it's common for advertisers to see high ROAS figures in retargeting campaigns simply because those users were already inclined to convert. While ROAS suggests high efficiency, iROAS often reveals diminishing or even negative incremental returns in such cases. This discrepancy underscores the value of transitioning to an iROAS-centric approach, which aligns better with the true objectives of performance marketing, driving net-new value rather than capturing conversions that were likely to occur regardless [13-15]. Furthermore, iROAS enables better budget pacing and allocation. Advertisers can more confidently shift spend towards campaigns and placements that demonstrate high incremental efficiency, rather than chasing total revenue or click-through rates. This is particularly relevant in MAB frameworks where the reward signal (i.e., iROAS) determines which ad placements or strategies to prioritize. Using iROAS instead of ROAS reduces the risk of feedback loops that reinforce suboptimal spending patterns, a common issue in many current programmatic platforms [16, 17]. iROAS also enhances multi-touch attribution models by introducing a layer of causal reasoning. While multi-touch attribution assigns fractional credit to multiple ad interactions across the customer journey, iROAS can act as a calibration layer that filters out interactions that did not causally influence the conversion. This allows for more accurate path analysis and ultimately supports more effective bidding strategies, especially in sequential decision environments such as those managed by MABs [18, 19]. The future relevance of iROAS becomes even more pronounced when considering the evolving regulatory landscape and privacy concerns. With third-party cookies being phased out and deterministic tracking becoming harder, probabilistic models like iROAS offer a compliant alternative for measuring ad effectiveness without requiring user-level tracking across the web. Because iROAS methodologies often rely on aggregate behaviors and experimental designs rather than persistent identifiers, they align well with privacy-first marketing approaches. This positions iROAS not just as a superior metric in terms of accuracy, but also in terms of sustainability and regulatory compliance [20-22].

In summary, iROAS offers a fundamentally more truthful lens through which to evaluate advertising performance. By focusing on causal impact rather than correlation, and by integrating with experimental or quasi-experimental methods, it allows advertisers in RMNs to base decisions on the *actual* value generated by campaigns. When this powerful metric is paired with the adaptive capabilities of multi-armed bandits, the subject of our next section creates a closed-loop system that continuously learns and improves upon itself. This combination of precision measurement and adaptive bidding creates the foundation for a more intelligent, effective, and accountable retail advertising ecosystem.

To further contextualize the limitations of traditional ROAS and highlight the superiority of iROAS in retail advertising, Table 1 provides a structured comparison of their core characteristics across key dimensions.

Table 1. Comparison Between Traditional ROAS and Incremental ROAS (iROAS)

Criteria	Traditional ROAS	Incremental ROAS (iROAS)
Definition	Total revenue generated divided by ad spend	Incremental revenue directly caused by ads divided by ad spend
Captures Organic Sales?	Yes, includes organic sales	No, excludes baseline/organic sales
Causal Measurement?	No, correlational	Yes, based on causal inference techniques
Use in Budget Optimization	May lead to over-investment in low-value ads	Supports efficient allocation based on true impact
Sensitivity to Attribution Errors	High	Lower, due to experiment-based estimation
Computation Complexity	Simple, readily available	Complex, requires experimentation or modeling

Criteria	Traditional ROAS	Incremental ROAS (iROAS)
Suitability for RMNs	Low	High

3. Multi-Armed Bandits in Retail Advertising: A Framework for Real-Time Decision-Making

Moving on to the advantages of iROAS in contrast with the traditional measures, its usage towards real-time environments should be coupled with the utilization of dynamic learning algorithms, in particular, Multi-Armed Bandits (MABs) [4-6]. When iROAS is used together with MABs, the feedback mechanism is formed and keeps the bidding decisions updated in a continuous way based on incremental value that balances exploration and exploitation. In this framework, all arms are constructions of compositions of advertisement positions, targeting choices, or advertising categories, and the reward signal is replaced by iROAS to display real advertising achievement rather than superficial interaction rates [23-26]. Advertisers can efficiently allocate a budget with flexibility through the knowledge of the decision with the greatest cumulative incremental returns, which makes Retail Media Networks more efficient and responsible.

Applying MAB in RMNs involves defining the action space, the context (if using contextual bandits), and the reward function. In iROAS-driven frameworks, the reward function is informed by real-time causal inference mechanisms that compute the lift attributable to advertising. For example, when an ad is served on a product detail page or via a search results slot, the system can compare the performance of similar users who were not exposed to the ad to estimate the incremental effect. This value becomes the observed reward that updates the bandit's policy, influencing future bidding decisions on that slot or format [27, 28].

There are many variants of Multi-Armed Bandit (MAB) algorithms that can be used with Retail Media Networks (RMNs), starting with a simple epsilon-greedy strategy to complex Upper Confidence Bound (UCB), Thompson Sampling, and Contextual Bandits. The epsilon-greedy algorithms trade off exploration and exploitation, randomly sampling the underutilized ad placements and prioritizing those high-performing, but this may result in under-exploration. UCB algorithms maintain dynamically explorative and expected returns based on probabilistic upper bounds, and Thompson Sampling is based on a Bayesian model, drawing samples on the posterior distribution to exploit noisy or uncertain iROAS estimates [4, 5, 29]. In RMNs, CMABs are more beneficial because they use rich contextual cues-time of day, device type, purchase history, or weather, etc., when selecting an arm and therefore can be conditioned on the present context and thus utilize this information to maximize iROAS given each user interaction [6, 8, 18].

The inherent flexibility of MAB-based frameworks is one of the largest strengths of these frameworks. In contrast to fixed rules or heuristic-based bid strategies, which assume that there is stationarity in user behavior, MABs are continually being trained and updated on the dynamics of the market. In other words, seasonality or a promotional effort can see one of the product categories soar. The main problem with a traditional system is that it may perform poorly; it is based on outdated values of averages, whereas an iROAS-based MAB model will see the incremental returns increase and will allocate bids based on this increase. This flexibility renders MABs the best when it comes to real-time optimization of campaigns within RMNs [9, 10, 13]. The use of iROAS as the reward signal in the MABs also achieves budget efficiency. Advertisers have a tendency to squander money on impressions that can bring in revenue but fail to increase consumer behavior incrementally. Because MABs are optimized using the reward function, iROAS would ensure that all the dollars spent are allocated to ad placements that will produce net-new conversions. The system gradually develops to reject "false-positive" placements, those that perform well when measured using ROAS, but offer low incremental value when measured using iROAS [7, 10, 15].

The other important benefit is that MAB-based bidding strategies are scalable. RMNs can generally include an enormous number of ad options, such as sponsored placement on search results and display placement in various categories and platforms. The manual management of these options and the management of these options by means of the use of manual rule-based systems is not just inefficient but impossible to manage at scale. MABs automate this process by continuously evaluating each potential arm and reallocating resources as needed. This rate of automation helps thousands of micro-decisions per second, ideally in a high-frequency bidding setting such as an RMN [13, 14].

Although MAB systems based on iROAS are highly beneficial, there are a number of practical considerations that have to be made. To begin with, the quality and latency of iROAS estimates are paramount because erratic or slow data may lead the MAB to make suboptimal strategies. New products or ad placements that have limited data can also get cold-started, which can be reduced by adopting hybrid strategies to combine supervised learning and bandit strategies or bootstrap iROAS estimates of similar settings [14-16]. The privacy limitations, especially when low third-party tracking is used, also determine the system design; whereas aggregate-level and experiment-based iROAS computations allow privacy-compliant bidding without affecting the intelligence [19, 20, 22]. On the whole, the combination of the MAB algorithms with the iROAS measurement can reshape the RMN campaign management so that it represents real-time, adaptive, and causally informed bidding, which correlates better with the real-life business results. As addressed in the next section, this combined framework can be put into practice and evaluated in practice.

Based on the diversity of retail advertisement settings and the intention of the campaigns, different versions of MAB algorithms would be more or less complementary in different settings. Table 2 gives actual work instances of popular MAB models in Retail Media Networks.

Table 2. Applicability of MAB Algorithm Variants in Retail Media Environments

MAB Variant	Best Use Case	Strengths	Limitations
Epsilon-Greedy	Early-stage testing of new ad placements	Simple, easy to implement	May over-explore or under-exploit
Upper Confidence Bound (UCB)	Well-established SKUs with moderate traffic	Balances exploration and exploitation effectively	Assumes reward distributions are stationary
Thompson Sampling	Medium-to-large campaigns with reward uncertainty	Probabilistic, adapts well to noise	Requires Bayesian prior estimation
Contextual Bandits	Real-time bidding with user/device-level features	Personalization uses rich context data	More complex data pipeline and infrastructure needed
Bayesian Neural Bandits	Dynamic pricing and creative selection	Learns non-linear relationships, high accuracy	Computationally intensive, slower to converge
Multi-objective Bandits	Campaigns with multiple advertiser goals (e.g., reach + iROAS)	Optimizes across conflicting objectives	Requires multi-dimensional reward modeling

4. Implementation and Evaluation of iROAS-MAB Framework in RMNs

Having already given the conceptual and algorithmic foundation of both iROAS and Multi-Armed Bandits (MABs), the next most important thing to do is to explore how this hybrid strategy can be applied to Retail Media Networks (RMNs), as shown in Figure 2. Implementation is not only a technical process but a very strategic process that involves a coordinated progression as regards data infrastructure, modeling architecture, experimentation protocols with regard to performance evaluation. The practical use of iROAS-based MAB bidding must overcome several practical issues and ensure that the system is scalable and able to deliver real-time feedback and satisfy the regulatory demands [4-6].

Firstly, the implementation of the iROAS-MAB system is impossible without a well-developed data engineering pipeline that will be able to receive, process, and synchronize large streams of data regarding user interaction. These are advertisements, clicks, conversions, user characteristics, session recordings, and product category updates. Such streams of data need to be handled in close real-time with event-driven frameworks like Apache Kafka or Flink. More importantly, the system should be capable of identifying the treatment and control groups dynamically, and such information is important to estimate the iROAS in real time. It is through this infrastructure that the bandit algorithms can be regularly updated with new performance feedback; hence, the bidding strategies can be modified based on the current user behavior and market trends [10, 14, 30]. The calculation of iROAS in the production settings should not be limited to a simple uplift calculation. Incremental effects can also be more effectively estimated using Bayesian hierarchical models, counterfactual regression methods, and synthetic control groups under the different conditions of data sparseness. As an example, hierarchical shrinkage can share statistical strength with related items or campaigns in categories or SKUs with low conversion volume and enhance the strength of iROAS estimates. It is also possible to update iROAS with rolling windows or exponentially weighted moving averages to reflect the current market conditions and smooth out short-term noise [3, 12, 25].

After the iROAS estimates are created, they can be used as a reward signal for MAB algorithms, which can dynamically optimize bids based on inventory types. A policy engine is provided with the real-time iROAS estimates, calculates expected returns, and recalculates budgets by choosing the best arms-combinations of placement, strategy, or creative balancing exploration and exploitation [4, 6, 8]. Value-based models used to determine Bid prices include iROAS, competition, and constraints. To operationalize this framework, low-latency model serving, a contextual bandit feature store, monitoring, and compliance logging, alongside integration with real-time bidding engines, are needed to provide milliseconds-responding functionality [13, 14, 18]. Performance assessment is based on state-of-the-art experimentation, which involves interleaving, multi-cell holdouts, cumulative iROAS, incremental conversions, time-to-optimal policy, and exploring efficiency, which gives a more holistic picture of system responsiveness to different levels of traffic and budgets [11, 21, 23].

One major concern in implementation is the cold-start problem, which occurs when new products, campaigns, or segments lack historical data. To address this, hybrid approaches can be deployed where supervised learning models provide initial reward estimates based on similar contexts or metadata, which are then refined by the bandit's learning process. Alternatively, warm-start strategies can initialize new arms with posterior distributions drawn from related arms, enabling more stable early performance [16, 17, 24]. Another key aspect is the system's resilience to noisy or delayed iROAS signals. Since iROAS is derived from experimental or observational data, it is inherently more variable than deterministic metrics like click-through rate. This necessitates the use of robust MAB algorithms that can accommodate reward uncertainty. Techniques such as bootstrapped Thompson Sampling or Bayesian UCB with uncertainty penalization allow the bandit to act conservatively in the face of noisy data, gradually increasing exploration once confidence in iROAS values improves [5, 6,18].

Bringing to business, the implementation of an iROAS-MAB system must comply with the strategic objectives and advertiser expectations. The operators of the platforms are forced to teach advertisers to focus on incrementality more than gross returns, to adjust KPIs, and reset success metrics not based on short-term revenue, but on long-term value [19, 20, 26]. It is necessary to build trust in algorithmic decisions, which can be facilitated by transparent reporting, confidence interval of iROAS, audit trail, and policy visualization. The regulatory and ethical aspects are essential: data should be privacy-friendly to meet frameworks like GDPR, CCPA, and it can be ensured by means of anonymization, aggregate groups, or differential privacy. It is possible to add fairness restrictions to facilitate fair exposure between products, brands, or segments [21, 22, 27]. An iROAS-MAB system can improve campaign efficiency, budget use, and incremental revenue when properly applied, which results in a dynamic and data-driven framework that allows real-time bidding to achieve the true causal effect and business value.

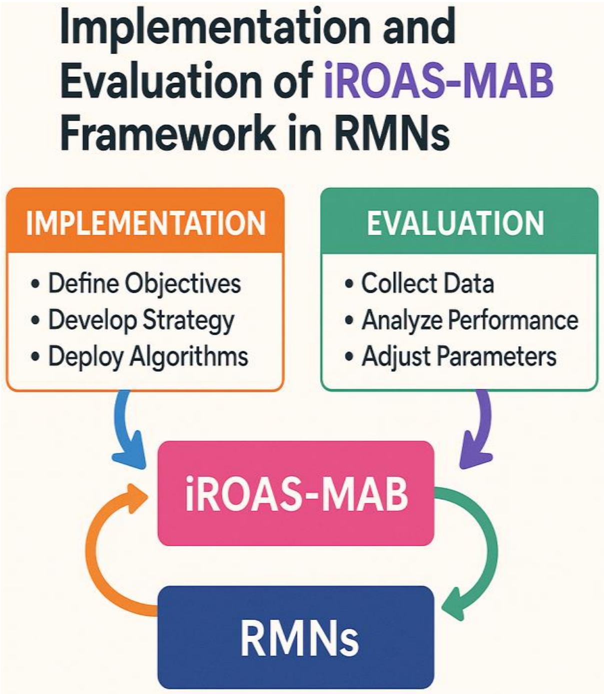

Figure 2: Flowchart illustrating the implementation and evaluation process of the iROAS-MAB framework in Retail Media Networks (RMNs)

5. Future Directions and Research Challenges

The combination of iROAS-based bidding and Multi-Armed Bandit (MAB) algorithms on Retail Media Networks (RMNs) has led to quantifiable campaign efficiency and incremental revenue. However, issues such as real-time causal estimation, sophisticated bandit algorithms, multi-objective optimization, delayed reward management, privacy-preserving learning, multi-touch journey modeling, and optimizing at a level of creativity are still noted as key research problems. Simulation-based evaluation and interpretability complementary efforts are necessary to deploy safely and be adopted by the organization. Table 3 provides a summary of key future trends and study areas for iROAS-MAB in RMNs.

Table 3: Future Research for iROAS-MAB

Future Direction / Challenge	Description / Opportunity	Potential Approaches / Techniques
Real-time causal estimation	Improve accuracy, latency, and scalability of iROAS computation	Hybrid experimental/quasi-experimental frameworks; adaptive propensity scoring; synthetic controls
Advanced bandit algorithms	Enhance performance in high-dimensional action spaces	Bayesian Neural Bandits, Deep Contextual Bandits, Meta-Learning Bandits; rich contextual signals (product embeddings, user behavior, inventory constraints)
Multi-objective optimization	Balance iROAS, customer lifetime value, diversity, and advertiser constraints	Pareto front optimization, scalarization, and multi-objective bandits
Delayed reward problem	Handle latency between ad exposure and conversion	Temporal Difference Learning, LSTM-based bandits, sequential credit assignment
Privacy-preserving learning	Maintain performance under privacy constraints.	Differential privacy mechanisms, Bayesian robustness, distributionally robust optimization, federated bandits
Multi-touch journey modeling	Capture value across sequential user interactions	Markov Decision Processes (MDPs), Recurrent Neural Networks (RNNs), policy gradient methods
Creative-level optimization	Dynamically personalize ad creatives alongside placements	Multi-level or hierarchical bandits selecting both placement and creative variation
Simulation-based evaluation	Safely test policies before real-world deployment	Offline simulators using historical data, user models, and traffic emulation

Interpretability & adoption	Facilitate organizational acceptance and understanding	Explainable AI tools, transparent dashboards, training, and education for marketers

6. Conclusion

In conclusion, the integration of incremental Return on Ad Spend (iROAS) with Multi-Armed Bandit (MAB) algorithms within Retail Media Networks (RMNs) represents a powerful evolution in performance marketing, one that is both intelligent in its adaptiveness and grounded in causal rigor. This approach shifts the strategic focus from superficial engagement metrics to a more substantive goal: measuring and optimizing incremental value generated by advertising. Unlike traditional heuristics or rule-based bidding systems, the iROAS-MAB framework enables continuous learning, data-driven adaptation, and real-time decision-making that better aligns with actual business outcomes. However, fully realizing the potential of this paradigm requires sustained innovation and collaboration across several domains. Advances in algorithmic development, particularly in contextual, multi-objective, and privacy-preserving bandit variants, must keep pace with the increasing complexity of user journeys and retail inventories. Similarly, causal inference methods used to estimate iROAS need to evolve toward more scalable, robust, and low-latency techniques to enable seamless integration into real-time systems. Furthermore, the supporting data infrastructure must be capable of ingesting, processing, and synchronizing high-frequency behavioral signals across platforms while maintaining accuracy, compliance, and interpretability. Equally important is the human element. For broad adoption to occur, explainability and transparency must be embedded into the framework so that marketing teams, media planners, and executives can confidently interpret model decisions, align them with strategic goals, and trust the systems managing increasingly large portions of their ad budgets. This necessitates the development of human-centered tools, intuitive reporting dashboards, and comprehensive governance frameworks to support responsible AI deployment. Ultimately, the future of retail advertising lies not simply in capturing consumer attention but in quantifying and optimizing the true value that advertising generates. In this transformation, iROAS-MAB frameworks are poised to serve as the foundational engine combining causality, adaptability, and automation to enable smarter, more accountable advertising at scale.

References

[1] Lewis, R. A., & Rao, J. M. (2015). The unfavorable economics of measuring the returns to advertising. *The Quarterly Journal of Economics*, *130*(4), 1941-1973.

[2] Dalessandro, B., Perlich, C., Stitelman, O., & Provost, F. (2012, August). Causally motivated attribution for online advertising. In *Proceedings of the sixth international workshop on data mining for online advertising and internet economy* (pp. 1-9).

[3] Rojas-Carulla, M., Schölkopf, B., Turner, R., & Peters, J. (2018). Invariant models for causal transfer learning. *Journal of Machine Learning Research*, *19*(36), 1-34.

[4] Mierswa, I., & Wurst, M. (2005, October). Efficient case-based feature construction. In *European Conference on Machine Learning* (pp. 641-648). Berlin, Heidelberg: Springer Berlin Heidelberg.

[5] Scott, S. L. (2010). A modern Bayesian look at the multi-armed bandit. *Applied Stochastic Models in Business and Industry*, *26*(6), 639-658.

[6] Chapelle, O., & Li, L. (2011). An empirical evaluation of Thompson sampling. *Advances in neural information processing systems*, *24*.

[7] Reverdy, P., Srivastava, V., & Leonard, N. E. (2016). Satisficing in multi-armed bandit problems. *IEEE Transactions on Automatic Control*, *62*(8), 3788-3803.

[8] Li, L., Jamieson, K., Rostamizadeh, A., Gonina, E., Ben-Tzur, J., Hardt, M., ... & Talwalkar, A. (2020). A system for massively parallel hyperparameter tuning. *Proceedings of machine learning and systems*, *2*, 230-246.

[9] Bottou, L., Peters, J., Quiñonero-Candela, J., Charles, D. X., Chickering, D. M., Portugaly, E., ... & Snelson, E. (2013). Counterfactual reasoning and learning systems: The example of computational advertising. *The Journal of Machine Learning Research*, *14*(1), 3207-3260.

[10] Johansson, F., Shalit, U., & Sontag, D. (2016, June). Learning representations for counterfactual inference. In *International Conference on machine learning* (pp. 3020-3029). PMLR.

[11] Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. (2012, April). The role of social networks in information diffusion. In *Proceedings of the 21st International Conference on World Wide Web* (pp. 519-528).

[12] Athey, S., & Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, *31*(2), 3-32.

[13] Qiao, W., Haghani, A., & Hamedi, M. (2012). Short-term travel time prediction considering the effects of weather. *Transportation research record*, *2308*(1), 61-72.

[14] McMahan, H. B., Holt, G., Sculley, D., Young, M., Ebner, D., Grady, J., ... & Kubica, J. (2013, August). Ad click prediction: a view from the trenches. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1222-1230).

[15] Chen, H., He, X., An, C., & Nguyen, T. Q. (2020). Adaptive image coding efficiency enhancement using deep convolutional neural networks. *Information Sciences*, *524*, 298-317.

[16] Swaminathan, A., & Joachims, T. (2015, June). Counterfactual risk minimization: Learning from logged bandit feedback. In *International Conference on machine learning* (pp. 814-823). PMLR.

[17] Krishnamurthy, A., Agarwal, A., & Dudik, M. (2016). Contextual semibandits via supervised learning oracles. *Advances In Neural Information Processing Systems*, *29*.

[18] Li, L., Chu, W., Langford, J., & Schapire, R. (2012). *U.S. Patent Application No. 12/836,188*.

[19] Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and trends® in theoretical computer science*, *9*(3–4), 211-407.

[20] Dwork, C., Naor, M., Pitassi, T., & Rothblum, G. N. (2010, June). Differential privacy under continual observation. In *Proceedings of the forty-second ACM symposium on Theory of Computing* (pp. 715-724).

[21] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016, August). " Why should I trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1135-1144).

[22] Balseiro, S., Golrezaei, N., Mahdian, M., Mirrokni, V., & Schneider, J. (2019). Contextual bandits with cross-learning. *Advances in Neural Information Processing Systems*, *32*.

[23] Galbreth, M. R., Ghosh, B., & Shor, M. (2012). Social sharing of information goods: Implications for pricing and profits. *Marketing Science*, *31*(4), 603-620.

[24] Dudík, M., Langford, J., & Li, L. (2011). Doubly robust policy evaluation and learning. *arXiv preprint arXiv:1103.4601*.

[25] Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests.

[26] Chu, W., Li, L., Reyzin, L., & Schapire, R. (2011, June). Contextual bandits with linear payoff functions. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics* (pp. 208-214). JMLR Workshop and Conference Proceedings.

[27] Li, L., Chu, W., Langford, J., Moon, T., & Wang, X. (2012, May). An unbiased offline evaluation of contextual bandit algorithms with generalized linear models. In *Proceedings of the Workshop on Online Trading of Exploration and Exploitation 2* (pp. 19-36). JMLR Workshop and Conference Proceedings.

[28] Chen, M., Beutel, A., Covington, P., Jain, S., Belletti, F., & Chi, E. H. (2019, January). Top-k off-policy correction for a REINFORCE recommender system. In *Proceedings of the twelfth ACM international conference on web search and data mining* (pp. 456-464).

[29] Hill, D. N., Moakler, R., Hubbard, A. E., Tsemekhman, V., Provost, F., & Tsemekhman, K. (2015, August). Measuring causal impact of online actions via natural experiments: Application to display advertising. In *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1839-1847).

[30] Alam, M. A., Nabil, A. R., Mintoo, A. A., & Islam, A. (2024). Real-time analytics in streaming big data: techniques and applications. *Journal of Science and Engineering Research*, *1*(01), 104-122.