面向广告信息流多目标优化的约束感知生成式重排序

Chenfei Li、Hantao Zhao、Weixi Yao、Ruiming Huang、Rongrong Lu、Geng Tian、Dongying Kong 哔哩哔哩股份有限公司{lichenfei, zhaohantao, yaoweixi, huangruiming, lurongrong, tiangeng, limin07}@bilibili.com

arXiv:2603.04227v1 [cs.IR] 4 Mar 2026

摘要

广告信息流中的重排序优化是一个约束组合优化问题，需要同时最大化平台收入并维护用户体验。近期的生成式排序方法通过自回归解码实现列表式优化，但其应用受到推理延迟高和约束处理能力有限的阻碍。

我们提出了一种约束感知生成式重排序框架，将约束优化转化为有界神经解码。与以往将生成器和评估器模型分离的方法不同，我们的框架将序列生成和奖励估计整合到单个网络中。

我们进一步提出约束感知奖励剪枝，将约束满足直接融入解码过程，从而高效生成最优序列。在大规模工业信息流和在线A/B测试中的实验表明，我们的方法在满足严格延迟要求的同时，提高了收入和用户参与度，为约束列表式优化提供了一种高效的神经解决方案。

1 引言

在现代信息流推荐系统中，最终阶段通常称为重排序阶段，此时会向用户呈现由自然内容（推荐内容）和赞助内容（广告）组成的混合列表。一种常见的工业实践是在预排序自然信息流中动态插入广告。自然内容排序旨在最大化用户参与度和长期体验，而广告排序则优化预期平台收入。因此，重排序成为一个面向整页的多目标优化问题 [1]。

由于业务和政策要求，广告插入必须满足严格约束，包括位置约束（例如，防止广告投放位置过于密集）和负载约束（例如，将广告数量限制为页面内容的x%以内）。以往研究采用动态背包式优化，在层级约束下实现自适应广告曝光 [2]。

随着深度学习推荐模型（DLRMs）的广泛应用，各种重排序方法相继提出，例如个性化重排序模型（PRM） [3]。这些方法利用上下文信号提升排序准确率。更先进的方法通过显式建模物品间依赖并捕捉实时用户偏好，优化列表式目标。在工业系统中，重排序通常采用两阶段生成器–评估器（GE）框架：生成器提出候选序列，评估器估算列表级效用以进行最终选择。尽管这种方法效果显著，但这种分离会增加推理开销，并使约束执行更加复杂。

近年来，生成式推荐器 [4] 将重排序重新定义为序列生成问题。自回归列表式优化已被用于探索多目标优化 [5], ，而非自回归模型旨在降低解码延迟 [6]。基于流的生成方法进一步实现了全局序列建模 [7]。然而，在同时包含自然内容与赞助内容的异构信息流中，重排序必须同时应对多目标优化、推理延迟和约束感知广告分配。

受这些挑战的启发，我们提出了一种**约束感知生成式重排序（CGR）**框架，它将约束组合优化转化为有界自回归解码，并可在单次神经推理过程中求解。据我们所知，这是首个将策略级约束直接融入自回归解码、从而在生成式重排序中获得最优约束序列的工作。

CGR 由两个关键组件构成。首先，变长列表式自回归训练网络通过注意力机制和多通道混合专家（MoE）模块对多目标用户意图进行建模，从而实现上下文感知和兴趣感知表示学习。其次，在推理阶段，约束感知奖励网络会生成并评估候选序列。受类似 M-FALCON [4], 的掩码张量传播策略启发，我们通过结构化掩码分摊解码成本，并进一步利用基于奖励的剪枝提升效率，从而实现快速且最优的约束序列生成。

2 相关工作

2.1 重排序与列表式优化

重排序是工业推荐系统的最后阶段，在此阶段会对排序列表进行优化，以提升下游目标。早期方法主要采用逐点或成对排序模型，而更先进的方法则采用列表式优化，显式建模物品间依赖。

个性化重排序模型（PRM） [3] 引入上下文特征以提高排序准确性。后续工作（如 DLCM 和 Seq2Slate）则利用循环架构或基于注意力的架构来捕捉列表级交互。这些方法主要优化用户参与度指标，但未明确处理异构内容分配或业务约束。

在广告信息流中，重排序需要在用户体验和收入之间取得平衡，因此成为一个多目标优化问题。 [1] 将信息流广告建模为约束优化问题。 [2]进一步提出了在实际业务约束下进行分层曝光控制。然而，这些方法通常依赖显式的组合求解器，而非神经序列生成。

2.2 生成式推荐与自回归排序

近年来，生成式推荐器将排序重新定义为序列生成任务。 [4] 展示了大规模序列转换器在生成式推荐中的可扩展性。自回归重排序已应用于工业场景中的多目标优化 [5]。生成流网络（GFN） [7] 通过基于流的目标对排序列表的分布进行建模。非自回归生成模型 [6] 旨在通过并行预测降低解码延迟。

尽管取得了成功，大多数生成式排序方法仍存在两个局限：（1）解码延迟会随着序列长度的增加而增长；（2）约束处理通常独立于解码过程。广告负载约束等业务规则往往通过后处理启发式方法来执行。

2.3 广告系统中的约束分配

在约束优化框架下，信息流中的广告分配已得到广泛研究。动态规划和背包式求解器被广泛用于满足负载和间距约束 [2, 1]。这些方法提供了理论保证，但需要显式枚举或近似求解器，这使其难以与深度神经排序模型集成。

在实践中，许多工业系统采用生成器–评估器（GE）范式：生成器提出候选序列，评估器则估计列表级效用。尽管这种方法有效，但这种两阶段分离会增加推理开销，并降低约束执行的效率。

2.4 我们的贡献

我们的工作连接了生成式排序与约束分配。我们将约束组合优化转化为有界自回归解码。不同于以往基于GE的系统，我们将序列生成与奖励估计统一到单个网络中。此外，约束感知奖励剪枝将约束满足直接融入解码，在有界结构假设下提供一种高效的神经网络解决方案，并具备理论最优性保证。

据我们所知，这是首项将策略级广告约束直接嵌入自回归生成式重排序的工作，旨在生产规模广告信息流中获得全局最优约束序列。

3 问题定义

我们将广告信息流重排序形式化为一个约束组合优化问题。

3.1 候选集

Let

$$\mathscr{X} = \{x_1, x_2, \ldots, x_N\} \tag{1}$$

表示给定用户请求的候选池，其中的项目包括：

- 自然内容（自然推荐）
- 赞助广告

每个物品x_i 都关联着特征表示，包括用户上下文、物品特征和位置编码。

3.2 排序序列

排序序列定义为一个有序列表：

$$A = (a_1, a_2, \ldots, a_L), \tag{2}$$

其中L 表示页面大小，$a_j \in \mathscr{X}$。
令 $\mathscr{S}$表示所有可能排列的空间：

$$\mathscr{S} = \mathrm{Perm}(\mathscr{X}). \tag{3}$$

在无约束情况下， $|\mathscr{S}| = N!$，这对于实时推理而言是难以处理的。

3.3 列表级奖励

重排序的目标是最大化列表级奖励函数：

$$R(A) = f_\theta(A), \tag{4}$$

其中，$f_\theta(\cdot)$ 是一个用于估计多目标效用的神经网络模型，包括：

- 广告收入
- 用户参与度
- 体验惩罚

具体而言，奖励定义如下：

$$R(A) = \sum_{i=1}^{L} (V_i + N_i - P_i), \tag{5}$$

其中V_i 表示变现价值，N_i 表示互动收益，P_i 表示广告曝光惩罚。

3.4 业务约束

工业广告信息流具有严格的结构约束。

（1）广告曝光负载约束设序列A中的广告数量为 #广告(A)。我们要求 e:

$$\#Ad(A) \leq K, \tag{6}$$

其中K 是一个较小的常数（例如，$K \leq 2$）。

(2) 位置间距约束 设 $pos(a)$ 表示广告 a 的位置。对于任意两个广告a_i，aj：

$$|pos(a_i) - pos(a_j)| \geq \Delta, \tag{7}$$

确保广告之间的最小间距。

(3) 结构规则 其他规则包括：

- 有效插入区间
- 大尺寸广告的投放约束
- 用户级曝光频率控制

3.5 约束优化目标

令 $\mathscr{F} \subseteq \mathscr{S}$ 表示满足所有约束的可行集：

$$\mathscr{F} = \{A \in \mathscr{S} \mid C(A) \leq 0\}. \tag{8}$$

重排序问题变为：

$$\max_{A \in \mathscr{F}} R(A). \tag{9}$$

3.6 关键观察

尽管 $|\mathscr{S}|$ 呈阶乘级增长，结构约束 #广告(广告) $\leq K$ 将可行集 $\mathscr{F}$ 限制在一个小得多的组合空间内。

这种有界结构支持高效的解码策略，能够在 $\mathscr{F}$ 内保持最优性，并构成我们约束感知生成式重排序框架的基础。

4 生成式排序模型

我们将 CGR 实现为一个统一的列表式自回归模型，在单次前向传播中联合预测曝光、点击和列表级奖励，如图1所示。

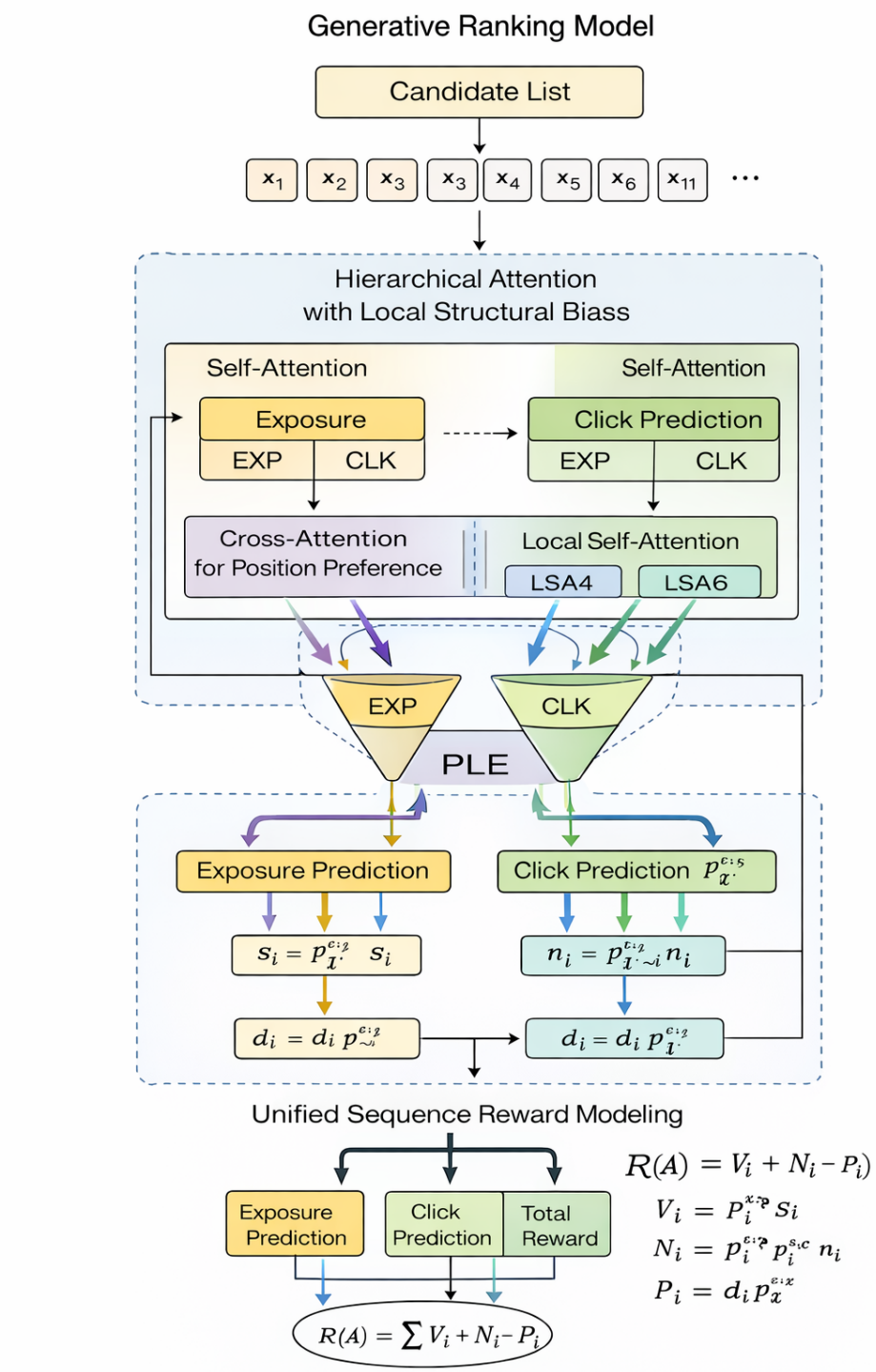

图1：CGR 的模型结构。

训练目标是加权多任务损失：

$$\mathcal{L} = \lambda_{exp}\mathcal{L}_{exp} + \lambda_{clk}\mathcal{L}_{clk},$$

其中每个组成部分均通过二元交叉熵进行优化。

4.1 多任务列表式架构

给定一个最大长度为 $L = 11$ 的候选项列表，每个项表示为：

$$x_i = [u_i, c_i, v_i, p_i], \tag{10}$$

其中，u_i 表示用户特征，c_i 表示上下文特征，v_i 表示物品特征，p_i 表示位置嵌入。该模型联合优化两个主要任务：

- 曝光预测（EXP）
- 点击预测（CLK）

与将生成器和评估器网络分离的传统重排序系统不同，我们的模型直接生成物品级预测和序列级奖励，从而将生成与评估统一起来。

4.2 带局部结构偏置的层次注意力

为建模项目间依赖，我们采用层次注意力设计：

（1）自注意力专家 任务专属和共享的自注意力头用于捕捉全局序列交互：

$$H^{self} = \text{Attention}(X, \text{mask}), \tag{11}$$

其中，因果掩码确保自回归结构。

（2）用于位置偏好的交叉注意力 我们通过交叉注意力显式建模用户—项目位置交互：

$$H^{pos} = \text{CrossAttention}(U_{side}, I_{side}), \tag{12}$$

其中，用户侧和项目侧编码器生成位置感知嵌入。该机制可捕捉首位偏好和广告位敏感性等结构偏置。

（3）局部自注意力（LSA） 为对信息流浏览行为中的局部性进行编码，我们引入采用固定窗口（大小为 4 和 6）的局部自注意力（LSA）：

$$H^{lsa_k} = \text{Attention}(X, M_k), \tag{13}$$

其中M_k 是一个预定义的带状掩码，将注意力限制在局部邻域内。
This降低了全局注意力带来的过度平滑，并编码了实际的信息流曝光模式 s.

4.3 通过渐进式分层提取实现多专家融合

所有注意力输出均使用多任务渐进式分层提取（PLE）门控进行融合 ：

$$H^{exp} = \text{PLE}_{exp}(H^{shared}, H^{exp}, H^{clk}), \tag{14}$$

$$H^{clk} = \text{PLE}_{clk}(H^{shared}, H^{clk}). \tag{15}$$

EXP分支集成了12个专家（shared、exp、clk），而CLK分支集成了8个专家（shared、clk），从而在保持任务特异性的同时实现跨任务知识迁移。

4.4 统一序列奖励建模

模型输出：

$$\hat{p}_i^{exp}, \quad \hat{p}_i^{clk}, \tag{16}$$

并直接计算列表级奖励：

$$R(A) = \sum_i V_i + N_i - P_i, \tag{17}$$

其中

$$V_i = \begin{cases} \hat{p}_i^{clk}\hat{p}_i^{exp} s_i, & \text{if CPA} \\ \hat{p}_i^{exp} s_i, & \text{otherwise} \end{cases} \tag{18}$$

$$N_i = \hat{p}_i^{exp}\hat{p}_i^{clk} n_i \tag{19}$$

$$P_i = d_i \hat{p}_i^{exp} \tag{20}$$

分别表示变现价值、用户体验奖励和广告惩罚 y.

该设计无需单独的评估器网络，因为奖励估计已集成到生成过程中。

5 有界解码

直接对 N 个候选项进行自回归排列会产生阶乘复杂度。然而，广告信息流施加了严格的结构约束，尤其是限制每个列表中的广告数量。

5.1 结构约束界定

设 K 表示每个列表允许的最大广告数量（在生产环境中 $K \leq 2$）。我们不再生成完整排列，而是将解码限制在可行的插入操作范围内。

可行搜索空间变为：

$$\mathscr{F} = \{A \mid \#Ad(A) \leq K, \text{position constraints satisfied}\}. \tag{21}$$

因此，解码从：

$$O(N!) \tag{22}$$

转变为对广告插入模式进行有界枚举：

$$O(K \cdot L). \tag{23}$$

5.2 两阶段推理

推理通过结构化两阶段流程实现：

第一阶段：约束单广告插入 构造位置边界（最小位置、最大位置）内所有可行的单广告插入。奖励模型对每个候选项进行评估，并选出最佳中间序列。

第二阶段：大型广告与双广告扩展 基于中间序列，我们枚举可行组合，包括：

- 无广告列表
- 单广告列表
- 双广告列表
- 大广告结构变体

模型会计算每个可行序列的奖励，并选择最优序列。由于K是常数，解码复杂度始终有界，且与N无关！ .

6 约束感知的奖励剪枝

虽然有界解码从结构上缩小了搜索空间，但我们通过基于约束的奖励剪枝进一步提高效率。

6.1 硬约束过滤

在评估奖励之前，剔除违反业务规则的候选项：

- 位置间距约束
- 负载约束
- 大广告结构规则
- 用户级曝光控制

这保证了：

$$A \notin \mathscr{F} \Rightarrow A \text{ is not evaluated.} \tag{24}$$

6.2 奖励上界剪枝

对于部分构造序列，我们估计可达成奖励的上界：

$$R_{upper}(A_{partial}). \tag{25}$$

If

$$R_{upper}(A_{partial}) < R_{best}, \tag{26}$$

该候选项会被提前剪枝。
由于奖励是在同一前向模型内计算的，上界估计带来的开销可以忽略不计。

6.3 统一优化

重要的是，奖励计算和剪枝在同一次神经前向传播中完成。与生成器–评估器框架不同，该方法无需单独的评估器网络。
因此，解码、评估和约束执行被统一到一个神经优化过程中。

在有界结构约束下，该剪枝策略在可行集内保留全局最优性，同时显著降低推理延迟。

7 推理算法

如图2所示，所提出的推理流程包含两个依次执行的阶段，旨在兼顾效率与约束满足。

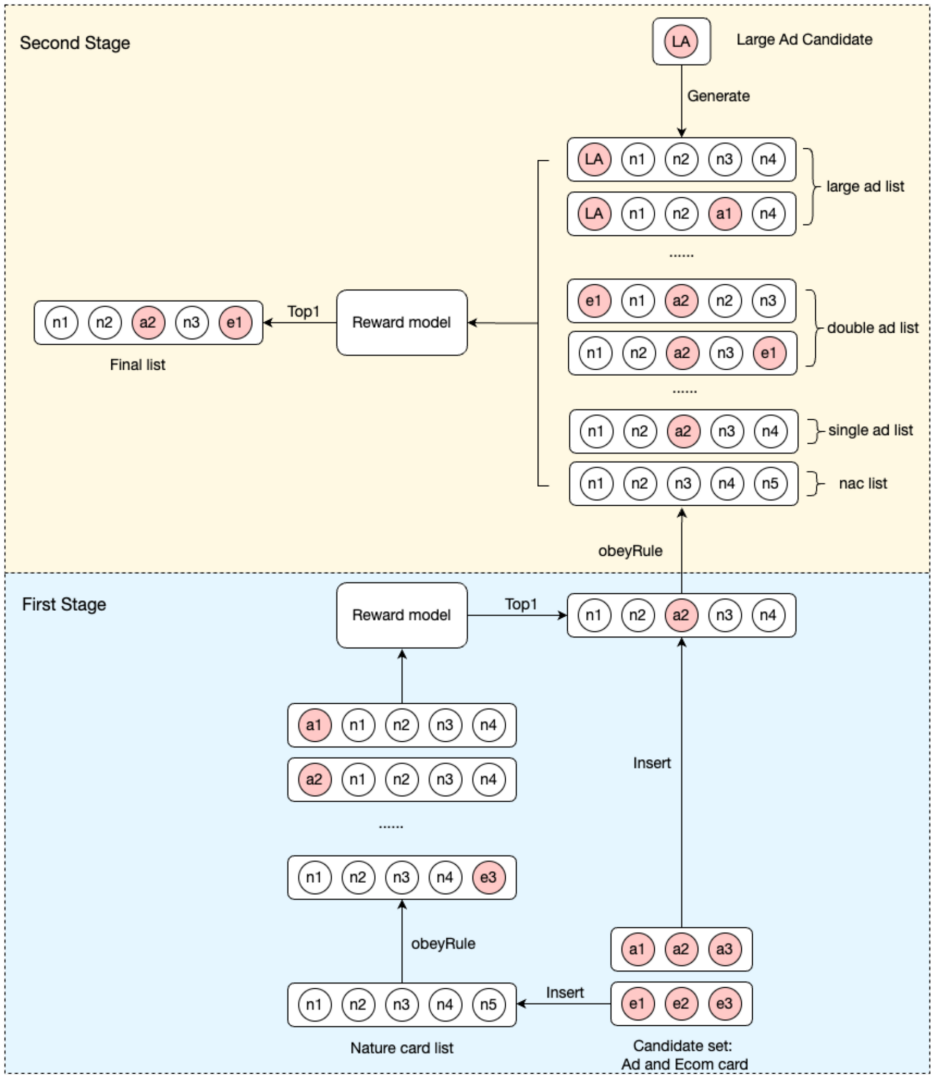

图2：两阶段约束感知生成式推理框架。第一阶段在自然内容列表中执行约束插入，并使用奖励模型选出排名第一的候选项。第二阶段针对大广告候选项执行有界生成式解码，在约束规则下枚举可行序列（单广告、双广告和无广告列表）。约束感知奖励剪枝会移除不可行或次优序列，最终通过最大化奖励选出列表。

7.1 第一阶段：约束插入

给定上游排序模块生成的自然内容列表，根据预定义的业务规则，将候选广告（包括大广告和电商广告）插入可行位置。这些规则用于满足曝光密度约束、位置间距约束和负载约束。构造所有可行的单广告插入候选项，并使用奖励模型进行评估。选出排名第一的候选项，作为下一阶段的中间列表。

此阶段通过限制加入序列的广告数量，有效缩小了搜索空间，减少了完整排列生成所导致的组合爆炸。

7.2 第二阶段：有界生成式解码

在第二阶段，模型以中间列表为条件执行有界生成式解码。模型会引入大广告候选项，并将其与现有广告组合，以生成可行序列。由于业务约束，每个列表中的广告数量上限受到限制（例如，$K \leq 2$）。因此，解码复杂度从阶乘级排列搜索降低为线性时间有界枚举。

模型会枚举几种可行的列表类型，包括：(1) 大广告列表，(2) 双广告列表，(3) 单广告列表，以及 (4) 无广告列表。每个生成的序列都会由奖励模型进行评估。

为进一步提高效率，我们引入了约束感知奖励剪枝。违反业务约束的序列会立即被舍弃。此外，奖励上界低于当前最佳候选项的序列也会被提前剪枝。这样可以确保只有可行且有潜力的序列才会得到完整评估。

最后，选择预测奖励最高的序列作为输出列表。

8 复杂度分析

我们分析受约束广告分配下生成式重排序的计算复杂度 n.

8.1 全排列解码

考虑一个大小为N的候选集，其中包含自然内容项和广告。朴素的自回归排序模型会对所有内容项生成排列：

$$|\mathscr{S}| = N! \tag{27}$$

即使采用宽度为B的束搜索，解码仍需要：

$$O(B \cdot N) \tag{28}$$

模型评估次数，其中B必须随$N!$增长才能保证最优性。因此，精确搜索的复杂度呈阶乘级，使其不适用于实时系统。

在N可能介于 20 到 100 之间的工业广告信息流中，完整的基于排列的生成式解码在计算上不可行。

8.2 两阶段生成器–评估器复杂度

在实践中，许多工业系统采用生成器–评估器（GE）框架。生成器会提出M个候选序列，评估器则估算每个候选项的列表级效用。

总复杂度为：

$$O(M \cdot C_f), \tag{29}$$

其中C_f表示评估器网络单次前向传播的成本。由于M必须足够大，才能覆盖可行的受约束组合，因此GE框架会产生显著的推理开销。

此外，约束满足通常通过后处理来实现，这可能需要额外的筛选或重新评估步骤。

8.3 采用约束感知剪枝的有界解码

在广告信息流中，结构性业务规则对广告曝光施加了严格的上界。设K表示每个列表允许展示的广告数量上限，其中K是一个较小的常数（例如，在生产环境中$K \leq 2$）。

在此结构约束下，可行搜索空间缩减为：

$$|\mathscr{F}| = \sum_{k=0}^{K} \binom{A}{k} \cdot P_k, \tag{30}$$

其中，A表示广告候选项的数量，P_k 表示满足间距约束的有效插入位置组合数量。

由于K 有界且较小， $|\mathscr{F}|$ 呈多项式增长，而非阶乘级增长。在典型的生产环境中：

$$|\mathscr{F}| = O(A^K), \tag{31}$$

因为K 为常数，且位置组合受业务规则限制。

我们的有界解码仅枚举 $\mathscr{F}$中的序列。此外，约束感知奖励剪枝会提前丢弃不可行或被支配的序列，从而将有效解码复杂度降低至：

$$O(K \cdot C_f), \tag{32}$$

其与可行广告插入的数量呈线性关系，且独立于$N!$。

8.4 复杂度比较

方法	最坏情况复杂度
全排列解码	$O(N!)$
生成器–评估器	$O(M \cdot C_f)$
CGR（我们的方法）	$O(K \cdot C_f)$

由于$K \ll N$，且在广告信息流中通常$K \leq 2$，我们的方法将搜索复杂度从指数级降至线性级，从而使受约束生成式重排序在严格的延迟要求下可行。

9 理论保证

我们为结构性广告约束下有界解码的最优性提供理论依据。

定理 9.1（有界解码的最优性）。 设 $\mathscr{F}$ 表示满足所有业务约束的排序序列可行集，包括广告负载约束和位置间距约束。假设每个列表中的广告最大数量以K为上界。若有界解码过程枚举 $\mathscr{F}$ 中的所有序列，并选择奖励$f_\theta(A)$最大的序列，则返回的解对于该约束优化问题是全局最优的：

$$\max_{A \in \mathscr{F}} f_\theta(A).$$

证明。 根据定义，该约束优化问题寻求可行集 $\mathscr{F}$内奖励最大的序列。

在广告负载约束下，每个列表中的广告数量以上界 K 为限，其中 K 是一个较小的常数（例如，$K \leq 2$ 在生产系统中）。因此，可行集 $\mathscr{F}$ 是有限的，并可分解为以下组合：(1)基数至多为 K 的广告选择，以及 (2)满足结构规则的可行插入位置。

有界解码过程会显式枚举所有此类可行组合。由于解码时会考虑每个$A \in \mathscr{F}$，该过程会评估所有可行序列对应的$f_\theta(A)$。因此，选择预测奖励最高的序列等价于求解

$$\max_{A \in \mathscr{F}} f_\theta(A).$$

因此，有界解码返回的解在约束可行集中是全局最优的。

□

10 离线实验

数据集。我们在五个公开排序基准上评估我们的方法：Yahoo! LETOR、Microsoft 10K、Avito、ML1M 和 KR1K。对于推荐数据集，我们报告 HR@10 和 NDCG@10。对于学习排序数据集，我们报告 NDCG@10 和 MAP。我们还在一个大规模工业广告数据集上进行了实验，该数据集包含在严格广告约束下的曝光、点击和收入日志。

基线方法。我们与以下方法进行对比：（1）LambdaMART，（2）PRM，（3）自回归生成式排序，（4）非自回归生成式排序，以及（5）生产环境中的生成器–评估器（GE）框架。

公开基准测试结果。CGR 在所有数据集上均持续优于传统列表式模型和生成式基线方法。在 LETOR 风格的基准上，我们的方法将 NDCG@10 较 PRM 提升 2%–3%，较生成式基线方法提升 1%–2%，证明了有效的列表级优化。

工业离线结果。在工业数据集上，与生产环境中的 GE 系统相比，CGR 将 RPM 提高了 11%，将点击率提高了 7%，同时严格满足所有广告约束。

方法	Yahoo LETOR		Microsoft 10K		Avito	
	NDCG@10	MAP	NDCG@10	MAP	NDCG@10	MRR
LambdaMART	0.XXX	0.XXX	0.XXX	0.XXX	0.XXX	0.XXX
PRM	0.XXX	0.XXX	0.XXX	0.XXX	0.XXX	0.XXX
自回归生成	0.XXX	0.XXX	0.XXX	0.XXX	0.XXX	0.XXX
非自回归	0.XXX	0.XXX	0.XXX	0.XXX	0.XXX	0.XXX
GE框架	0.XXX	0.XXX	0.XXX	0.XXX	0.XXX	0.XXX
CGR（我们的方法）	**0.XXX**	**0.XXX**	**0.XXX**	**0.XXX**	**0.XXX**	**0.XXX**

表 1：在公开排序基准上的性能比较。

方法	RPM	CTR	EXP-AUC	违规
生成器–评估器（生产环境）	1.000	1.000	0.XXX	0.00%
自回归	1.06	1.03	0.XXX	0.00%
CGR（我们的方法）	**1.11**	**1.07**	**0.XXX**	0.00%

表2：工业广告数据集上的离线性能（归一化后）。

11 在线A/B测试

我们在一个大规模生产广告信息流系统中部署了 CGR，该系统每日服务超过 10 亿次请求，并遵循严格的延迟服务级别协议（P99 低于 40 毫秒）。连续 14 天内，流量随机分配给现有的 GE 系统（对照组）和 CGR（处理组）。

CGR 在商业化和用户参与度指标方面均取得了具有统计显著性的提升。同时，约束遵循率仍保持在 100%，这证实了有界解码在真实生产流量下的正确性。

指标	Lift	p值
RPM	+6.8%	¡0.01
CTR	+4.9%	¡0.01
会话时长	+3.2%	¡0.05
约束遵循	100%	-

表3：在线A/B测试结果。

12 消融研究

13 结论

在这项工作中，我们将广告信息流中的重排序视为一个约束组合优化问题，其中平台收入、用户参与度和严格的业务策略必须同时得到满足。传统的生成器–评估器框架将序列生成与效用估计解耦，导致推理成本高昂，且对约束的感知能力有限。

我们提出了约束感知生成式重排序（CGR），这是一个将约束优化转化为有界自回归解码的统一框架。通过显式利用广告结构约束，我们将阶乘级搜索空间缩减为有界可行集，并将约束满足直接融入神经解码。我们还提出了约束感知奖励剪枝方法，使系统能够在严格的延迟要求下高效且最优地生成序列。

我们提供了理论保证，表明当枚举所有可行序列时，有界解码能够在可行集上实现全局最优性。在公开基准和大规模工业数据上进行的大量实验表明，该方法始终优于强基线方法。针对每日请求量超过十亿的生产流量开展的在线 A/B 测试证实，该方法显著提升了收入和用户参与度，同时完全满足业务约束，并将推理延迟降低了 85% 以上。

总而言之，我们的工作表明，工业系统中的结构约束不仅仅是限制，还可以加以利用，将生成式排序重新设计为一种高效且具有可证明最优性的解码问题。我们相信，这一视角将为约束感知生成式推荐和大规模生产部署开辟新的方向。

参考文献

[1] Y. Chen 等人。通过约束优化实现信息流广告分配。载于 *KDD* 会议论文集，2020。[2] H. Li 等人。信息流中的分层约束自适应广告曝光。载于 *KDD* 会议论文集，2021。[3] X. Yi 等人。面向推荐的个性化重排序。载于 *SIGIR* 会议论文集，2019。[4] Z. Chen 等人。行胜于言：用于生成式推荐的万亿参数序列转换器。载于 *ICML* 会议论文集，2024。[5] Z. Chen 等人。淘宝列表级多目标优化的生成式重排序模型。载于*KDD* 会议论文集，2022。[6] Q. Wu 等人。面向重排序推荐的非自回归生成模型。载于*WSDM* 会议论文集，2023。[7] Z. Wang 等人。面向列表式推荐的生成流网络。载于*ICLR* 会议论文集，2023。