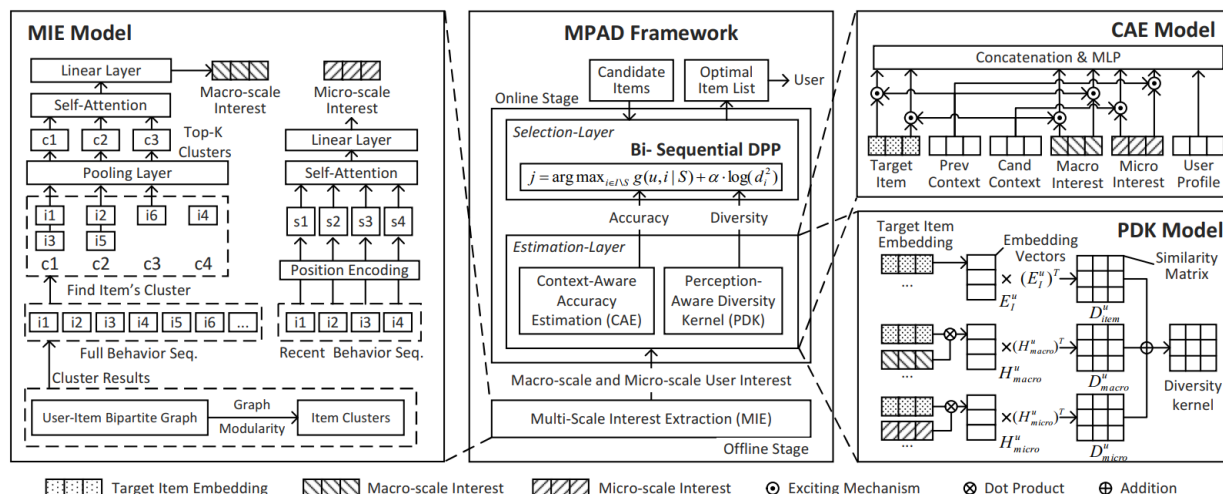


# 【2023 KDD】MPAD

## Multi-factor Sequential Re-ranking with Perception-Aware Diversification

### 模型设计



#### ▼ selection layer

##### ▼ BS-DPP

从候选物品集中基于准确度和多样性选出部分物品

#### ▼ estimation layer

##### ▼ 多尺度兴趣提取（MIE）

将用户和物品聚类，基于聚类编码表征用户多尺度兴趣

##### ▼ 上下文感知的准确度评估（CAE）

基于浏览上下文和多尺度兴趣将rank阶段的point-wise准确度分数细化为上下文感知的准确度分数

##### ▼ 知觉感知内核（PDK）

基于物品id和用户的多尺度兴趣计算物品相似度（衡量多样性）

## Bi-Sequential Item Selection

$$\mathcal{P}_1 : \arg \max_{S \in I} F(\text{Acc}(u, S), \text{Div}(u, S)) = \log \det(K_S^u), \quad (2)$$

从I中选出子集S，使得核函数 $K_S^u$ 达到最大

$$K_S^u(i, j) = g(u, i) \cdot D(i, j) \cdot g(u, j), \quad (3)$$

$g(u, i)$ 是衡量用户u和物品i之间的point-wise准确度分数，即相关性分数

$D(i, j)$ 是衡量物品i和物品j之间的相似度，即多样性分数

公式（3）是传统的DPP的核函数定义，在此基础上，1）将相关性与浏览上下文相关联；2）将多样性与用户兴趣相关联

$$K_S^u(i, j) = g(u, i|S) \cdot D(i, j|E_u) \cdot g(u, j|S), \quad (4)$$

其中，S是展示给用户的历史物品， $E_u$ 是用户的兴趣表征。因此（2）转化为：

$$h(u, S) = \sum_{i \in S} g(u, i|S) + \alpha \cdot \log \det(D_S^u), \quad (5)$$

$\alpha$ 是权衡相关性和多样性的一个可调节参数

公式（2）通过贪心算法（从候选集中贪心地选择一个能使边际收益最大的商品加入到最终的结果子集中）解决：

$$j = \arg \max_{i \in I \setminus S} \log \det(K_{S \cup \{i\}}^u) - \log \det(K_S^u) \quad (6a)$$

$$= \arg \max_{i \in I \setminus S} h(u, S \cup \{i\}) - h(u, S) \quad (6b)$$

$$= \arg \max_{i \in I \setminus S} g(u, i|S) + \alpha \cdot \left( \log \det(D_{S \cup \{i\}}^u) - \log \det(D_S^u) \right) \quad (6c)$$

$$= \arg \max_{i \in I \setminus S} g(u, i|S) + \alpha \cdot \log(d_i^2). \quad (6d)$$

S初始化为空集，每次选择一个物品i到S中，i要满足（6a）条件

---

### Algorithm 1 Bi-Sequential Item Selection in MPAD

---

```

1: Initialization:
2:  $D_S, \epsilon, c_i = [], d_i^2 = D_{ii}, j = \arg \max_{i \in I} \log(d_i^2), S = \{j\}.$ 
3: Iteration:
4: while  $|S| < k$  and  $d_i^2 < \epsilon$  do
5:   for  $i \in I \setminus S$  do
6:      $e_i = (D_{ji} - \langle c_j, c_i \rangle) / d_j.$ 
7:     Update  $d_i^2 = d_i^2 - e_i^2, c_i = [c_i \ e_i].$ 
8:     Update  $g(u, i|S)$  with the proposed preference model.
9:   end for
10:  Obtain  $j = \arg \max_{i \in I \setminus S} g(u, i|S) + \alpha \cdot \log(d_i^2).$ 
11:  Update subset  $S = S \cup \{j\}.$ 
12: end while

```

---

### Multi-Scale Interest Extraction

提取用户微观兴趣（recent interest）（具体到物品级别）和宏观兴趣（长期兴趣）（比如时尚、运动等）

#### 1. 模块化的图聚类（宏观macro）

基于用户-物品二部图的二部模块化分组聚类，通过Lovain算法迭代优化，度量函数如下：

$$Q = \frac{1}{E} \sum_{i,j} (A_{ij} - P_{ij}) \delta(c_i, c_j), \quad (7)$$

其中，E代表边的数量，A交互邻接矩阵， $P_{i,j}$ 代表经过不同类分图后物品i和物品j之间期待的边， $\delta(c_i, c_j)$ 代表

#### 2. 微观级别的用户兴趣

基于物品的建模

$$\tilde{e}_{i_x} = \text{Concat}\{e_{i_x}, t_{i_x}\}, \forall i_x \in I_{\text{micro}}^u, \quad (11)$$

t是时间衰减的可学习embedding，代表交互时间到现在的一个时间间隔。最后使用注意力机制。

$$\text{Head}_i = \text{Att}(Q_i, K_i, V_i),$$

$$h_{\text{micro}} = \text{Concat}(\text{Head}_1, \dots, \text{Head}_h) W^O$$

如果ij属于同个类，则为1，否则为0。  
Q越大，代表社区划分效果越好。类  
似下面的计算方式：

$$Q = \frac{1}{2m} \sum_{vw} [A_{vw} - \frac{k_v k_w}{2m}] \delta(c_v, c_w)$$

k代表节点的度

把一个类下的所有物品的表征聚合作  
为这个聚类的表征

$$h_m^u = \text{Aggregate} \{e_{i_x}, \forall i_x \in c_m\},$$

取top-M个类，在之间做多头自注意力  
机制得到最后的表征：

$$\text{Att}(Q_m, K_m, V_m) = \text{Softmax}(\alpha Q_m K_m^T) V_m,$$

## Context-Aware Accuracy Estimation

$$h_{\text{prev}} = \text{Aggregate}(e_i, i \in [k-1]),$$

$$h_{\text{cand}} = \text{Aggregate}(e_i, i \in [N]).$$

基于SENet提出的激励机制（低成本的attention），计算上下文对目标物品的影响：

$$W_{\text{prev}} = \sigma(W_2 \cdot \delta(W_1 \cdot h_{\text{prev}})), \quad (15)$$

$$h_{\text{prev}}^{i_t} = W_{\text{prev}} \otimes e_{i_t},$$

$$\mathbf{h}_{\text{all}}^u = \text{Concat} \left( \mathbf{e}_{i_t}, \mathbf{h}_{\text{prev}}^{i_t}, \mathbf{h}_{\text{cand}}^{i_t}, \mathbf{h}_{\text{pr,lo}}^u, \mathbf{h}_{\text{pr,sh}}^u, \mathbf{h}_{\text{ca,lo}}^u, \mathbf{h}_{\text{ca,sh}}^u \right), \quad (17a)$$

$$\hat{Y}_{II}(u, i) = \text{Softmax} \left( \text{MLP} \left( \mathbf{h}_{\text{all}}^u \right) \right), \quad (17b)$$

### Perception-Aware Diversity Kernel

对微观、宏观和物品本身计算核函数并组合到一起

$$D_{\text{macro}}^u(i, j|E_u) = a_l^2 \exp \left[ -\frac{\mathbf{h}_{i,\text{macro}}^u \otimes \mathbf{h}_{j,\text{macro}}^u}{b_l^2} \right], \quad (18)$$

其中h是物品表征和宏观兴趣表征的内积

$$D_{\text{micro}}^u(i, j|E_u) = a_s^2 \exp \left[ -\frac{\mathbf{h}_{i,\text{micro}}^u \otimes \mathbf{h}_{j,\text{micro}}^u}{b_s^2} \right], \quad (19)$$

$$D_{\text{item}}(i, j|E_u) = a_s^2 \exp \left[ -\frac{\mathbf{e}_i^u \otimes \mathbf{e}_j^u}{b_s^2} \right], \quad (20)$$

$$D^u(i, j|E_u) = D_{\text{item}}^u(i, j|E_u) + \beta_1 \cdot D_{\text{macro}}^u(i, j|E_u) + \beta_2 \cdot D_{\text{micro}}^u(i, j|E_u), \quad (21)$$