Index Policies for Promotion Strategies in Reward-based Crowdfunding

Dong Li, Department of Management Science, Lancaster University
Chenguang Wang, Baibing Li, Loughborough Business School, Loughborough University

Workshop on restless bandits, index policies and applications in reinforcement learning
Grenoble, France, Nov 20-21, 2023
Agenda

- An introduction to reward-based crowdfunding
- The problem and the model
- Relaxation and decomposition
- Optimal policy for single project problems
- Indexability and index values
- Numerical experiments
An Introduction to Reward-based Crowdfunding

Reward-based crowdfunding is an innovative online financing alternative.

- The fundraisers launch their projects on a crowd-funding platform
- A typical project includes
  1. Information about their products/technology
  2. A pre-specified funding goal
  3. Length of the campaign
  4. A set of reward options that backers can purchase
- The backers choose which project to support, and then decide which reward option to purchase.
- Many platforms follow an **All-or-Nothing (AoN)** scheme.
- The platforms charge a percentage commission of the total funds raised from **successful** projects.
Low Success Rate

69-89% Failed (Clifford 2016)

Most researches focus on factors of individual projects

- Fundraisers’ experience and expertise
- Funding goal, duration and reward options
- Information description
A crowdfunding platform may boost the chance of success of a campaign by highlighting it on the platform’s homepage.

**Our Research Question**

How would platform managers maximise the total revenue by dynamically assigning limited promotion slots to projects?
The Problem Statement

- \( J \) crowdfunding projects seek financial investment from time \( 0 \) to \( T-1 \). Each project has a funding goal \( G_j \).

- Discretise the time horizon into sufficiently small intervals \( t \), Assume that customers visit the platform according to a Bernoulli process with a probability \( \lambda \in (0,1) \) in each time \( t \).

- Upon arrival, each customer either chooses to back one project, say project \( j \), with probability \( p_j \) or leave without any purchases.

- Having decided which project to support, the backer chooses one reward option to purchase before leaving.

- At each time period \( t \), the platform chooses one project to promote on its homepage.

- Our aim is to allocate the promotion slot to projects over time to maximise the revenue.
Modelling of Customers’ Choices

Random utility function: customers’ perceived valuation on project $j$

$$u_j(g_j, a_j) = z_j(g_j, a_j) + \epsilon_j = m_j + \beta_1 a_j + \beta_2 \left(1 - \frac{g_j}{G_j}\right) - \beta_3 + \epsilon_j$$

Multinomial logit model (MNL) – customer’s backing probability

$$p_j(g_j, a_j) = \frac{e^{z_j(g_j, a_j)}}{1 + \sum_{k=1}^{J} e^{z_k(g_k, a_k)}}, 1 \leq j \leq J$$

Non-purchase probability

$$p_0 = \frac{1}{1 + \sum_{k=1}^{J} e^{z_k(g_k, a_k)}}$$

Customer’s pledging

A customer will purchase a reward $r_j$ with a known probability $F_j(r_j)$ where $r_j \in \{1, \ldots, R_j\}$

- $G_j$: the funding goal of project $j$
- $g_j$: the shortfall to the funding goal of project $j$
- $m_j$: the overall attraction of project $j$
- $a_j \in \{0,1\}$: promotion indicator
- $\beta_1$: promotion power
- $\beta_2$: herding effect
- $\beta_3$: side effect, e.g., market saturation
The Model - a Dynamic Program

- States: \( g = (g_1, \ldots, g_J) \), a vector of shortfalls for all projects. Denote the state space at time \( t \) by \( \Omega_t \)
- Action: \( a = (a_1, \ldots, a_J) \), a vector of actions for all projects.
  Action space: \( A = \{a: a_j \in \{0,1\}, \sum_j a_j = 1\} \)
- A policy \( \pi: \Omega_t \to A, \forall 0 \leq t \leq T - 1 \), a decision rule to choose the project for promotion after observing the state at each time epoch
- Immediate reward in each time \( t \) under policy \( \pi \):
  \[ h_t(g, \pi(g)) = \lambda \sum_{j=1}^J p_j(g, \pi(g)) \sum_{r_j=1}^{R_j} r_j F_j(r_j) \]
- Our **objective** is to find a policy that maximises the overall revenue
The Bellman Equation

Denote by $V_t(g)$ the value function, i.e., the maximal expected fund still obtainable from time $t$ onwards, given the system occupies state $g$ at time $t$.

$$V_t(g) = \max_{a \in A} \left\{ \lambda \sum_{j=1}^{J} p_j(g, a) \sum_{r_j=1}^{R_j} F_j(r_j) (r_j + V_{t+1}(\tilde{g})) + (1 - \lambda \lambda p_0(g, a) V_{t+1}(g)) \right\}$$

where $\tilde{g} = g - r$ and $r$ is a $J$-dimensional vector that takes value of $r_j$ on the $j$-th component and zero elsewhere.

$$V_T(g) = \sum_{j=1}^{J} h_T(g_j), \text{ where } h_T(g_j) = \begin{cases} -(G_j - g_j), & \text{if } g_j > 0 \\ 0, & \text{if } g_j \leq 0 \end{cases}$$
Whittle’s Restless Bandits Method- In a Nutshell

**Restless Bandits**
- Each fundraising project is a restless bandit, which always evolves regardless being promoted or not.

**Relaxation and Decomposition**
- Relaxation 1: allow multiple projects to be promoted simultaneously, but require on average the resource consumed is not more than one.
- Relaxation 2: associate a non-negative Lagrangian multiplier $W$ (a fee for promotion) to the constraint and incorporated it into the objective function.
- Decomposition: these relaxations allow the problem to be decomposed into a collection of single bandit/project problems.

**Indexability and index values**
- Prove the *indexability* to each project
- Calculate the index values (or fair charges) for each project in each state

**Index policies**
- Always choose to promote the project with the largest index value.
Relaxations

Relaxation 1: $\tilde{A} = \{a: a_j \in \{0,1\}\}, \tilde{\pi}: \Omega_t \rightarrow \tilde{A}, \forall 0 \leq t \leq T - 1$. We require

$$E \left[ \sum_{t=0}^{T-1} \left( 1 - \sum_{j=1}^{J} \tilde{\pi}_t (g(t))_j \right) \right] \geq 0$$

Relaxation 2: Associate a non-negative Lagrangian multiplier $W$ to the constraint above, and add it to the objective function (1)

$$V_0(G) = \max_{\tilde{\pi}} E \left[ \sum_{t=0}^{T-1} h_t (g(t), \tilde{\pi}_t (g(t))) + \sum_{j=1}^{J} h_{T'} (g_j(T')) + W \sum_{t=0}^{T-1} \left( 1 - \sum_{j=1}^{J} \tilde{\pi}_t (g(t))_j \right) \right]$$ (2)

However, due to the MNL, the problem (2) is not yet decomposable.

$$h_t(g, \pi(g)) = \lambda \sum_{j=1}^{J} p_j(g, \pi(g)) \sum_{r=1}^{R_j} r_j F_j(r_j)$$
Approximation of MNL by BNL

We further relax the problem by approximating the MNL-based backing probabilities with the following $J$ Binomial Logit Models (BNLs), one for each project $j$:

$$p_{aj}^{a_j}(g_j) = \frac{\exp(m_j + \beta_1 a_j + \beta_2 (1 - g_j/G_j))}{1 + \exp(m_j + \beta_1 a_j + \beta_2 (1 - g_j/G_j))}$$

It can be understood that each project faces the entire arrival stream, of which each arriving customer makes a binary choice of either backing this project or not, based on the BNL model above.

Problem (2) can now be decomposed by project.
Single Project Problems

\[ v_0^w(g) = \max_{\pi} \left\{ \sum_{t=0}^{T-1} \left( \lambda \pi_t(g) \sum_{r=1}^{R} rF(r) - W \pi_t(g) \right) \right\} \]

where we still use \( \pi \) for the single project policy.

In each single-project problem, the project has a dedicated promotion space, and the action is whether or not to use the space for promotion at each decision epoch.

If the action is to promote \( (\pi(g) = 1) \), the project will be highlighted on the homepage with a cost of \( W \). If the decision is not to promote \( (\pi(g) = 0) \), the project will not be highlighted and no cost is incurred.
Monotonicity of the Optimal Policy to the Single Project Problem
– under the condition of sufficiently long duration

Proposition 1 (Monotonicity of the optimal policy) For any $W \geq 0$, the optimal policy $\pi^*$ satisfies:

- $\pi^*_t, W(g) \geq \pi^*_t, W(g - r), \forall g \in \Omega_t, 0 \leq t \leq T - 1$
- $\pi^*_t, W(g) \geq \pi^*_{t+1}, W(g), \forall g \in \Omega_t, 0 \leq t \leq T - 1$

Table 1: The Setting of an Example

<table>
<thead>
<tr>
<th>Project</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\beta_3$</th>
<th>$T$</th>
<th>$\lambda$</th>
<th>$G$</th>
<th>$m$</th>
<th>$F(r = 1)$</th>
<th>$F(r = 2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.0</td>
<td>1.5</td>
<td>60</td>
<td>0.7</td>
<td>8</td>
<td>0.01</td>
<td>0.73</td>
<td>0.27</td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>1.5</td>
<td>0.7</td>
<td>10</td>
<td>0.1</td>
<td>0.7</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
</tr>
</tbody>
</table>
For any $W$ and $t$, define the optimal promotion set as

$$B_t(W) = \{ g : \pi_t^*,W(g) = 1, g \in \Omega_t \}$$

From Proposition 1, we have

**Indexability:** $B_t(W) \subseteq B_t(W')$ for any $W \geq W'$, $\forall 0 \leq t \leq T - 1$
Demonstration of Indexability of the Example- project 2

\[ W = 0.208 \]

\[ W = 0.205 \]

\[ W = 0.201 \]

\[ W = 0.197 \]
Index Values

Whittle’s Index: for an indexable project, the Whittle’s index is defined as

\[ w(g, t) = \arg \max_{g \in B_t(W)} \{ g \} \]

**Proposition 2**: The Whittle index is evaluated as follows:

\[ w(g, t) = \lambda (p^1(g) - p^0(g)) \left( \sum_{r=1}^{R} F(r) r + \Delta v_{t+1}^{p_0}(g) \right), \]

where \( \Delta v_{t}^{p_0}(g) = \sum_{r=1}^{R} F(r) v_{t}^{p_0}(g - r) - v_{t}^{p_0}(g) \) is the marginal future revenue of an additional purchase under a non-promotion policy \( p_0 \).

**A closed-form index value approximation**: for each state \( g \) at time \( t \), we assume that the herding effect from \( t + 1 \) onwards and for all the future states remains the same as it is evaluated at time \( t \) for state \( g \), then \( v_{t}^{p_0}(g) \) can be approximated by

\[ \hat{v}_{t}^{p_0}(g) = \lambda p^0(g) \sum_{t}^{T-1} \tilde{r} = (T - t) \lambda p^0(g) \tilde{r}. \]

And thus

\[ \hat{w}(g, t) = \lambda (p^1(g) - p^0(g)) \left[ \tilde{r} + \lambda \tilde{r} (T - t - 1) \left( \sum_{r=1}^{R} F(r) (p^0(g - r) - p^0(g)) \right) \right]. \]
Lemma 1: The index value $w(g, t)$ (i) increases in state $g$; and (ii) decreases in time $t$. 
## Numerical Experiments - The Policies

<table>
<thead>
<tr>
<th>Policy Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smallest/largest shortfall first</td>
<td>Always promotes the project with the smallest/largest percentage shortfall</td>
</tr>
<tr>
<td>(SSF/LSF):</td>
<td></td>
</tr>
<tr>
<td>Smallest/largest utility first</td>
<td>Always promotes the project with the smallest/largest utility</td>
</tr>
<tr>
<td>(SUF/LUF):</td>
<td></td>
</tr>
<tr>
<td>Greedy policy</td>
<td>Always promotes the unfinished project with the highest funding goal</td>
</tr>
<tr>
<td>(GP):</td>
<td></td>
</tr>
<tr>
<td>Conservative policy</td>
<td>Always promotes the unfinished project with the lowest funding goal</td>
</tr>
<tr>
<td>(CP):</td>
<td></td>
</tr>
<tr>
<td>Myopic policy</td>
<td>Always promotes the project that leads to the highest immediate reward</td>
</tr>
<tr>
<td>(MP):</td>
<td></td>
</tr>
<tr>
<td>Index policy</td>
<td>Always promotes the project with the largest index value/approximate index value</td>
</tr>
<tr>
<td>(IP/IPx):</td>
<td></td>
</tr>
</tbody>
</table>
## The Settings

### Global parameters

<table>
<thead>
<tr>
<th>( \beta_1 )</th>
<th>( \beta_2 )</th>
<th>( \beta_3 )</th>
<th>( \lambda )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.01</td>
<td>2.5</td>
<td>0.7</td>
</tr>
</tbody>
</table>

### Project parameters

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Project ( j )</th>
<th>( j = 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( m )</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>( G )</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>( F(r = 1) )</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>( F(r = 2) )</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>( \bar{r}_b )</td>
<td>1.45</td>
</tr>
<tr>
<td><strong>Smaller difference btw projects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( m_n )</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>( F_n(r = 1) )</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>( F_n(r = 2) )</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>( \bar{r}_n )</td>
<td>1.46</td>
</tr>
<tr>
<td><strong>Larger difference btw projects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( m_e )</td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td>( F_e(r = 1) )</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>( F_e(r = 2) )</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>( \bar{r}_e )</td>
<td>1.44</td>
</tr>
</tbody>
</table>
## The Results

### – Percentage Revenue Gap between the IP and Other Policies

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Variation</th>
<th>SSF</th>
<th>LSF</th>
<th>SUF</th>
<th>LUF</th>
<th>GP</th>
<th>CP</th>
<th>MP</th>
<th>IPx</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>-</td>
<td>33.922</td>
<td>140.081</td>
<td>41.428</td>
<td>25.297</td>
<td>25.301</td>
<td>16.358</td>
<td>28.122</td>
<td>-0.061</td>
</tr>
<tr>
<td><strong>Global Parameters Sensitivity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration</td>
<td>5% ↑</td>
<td>33.344</td>
<td>70.658</td>
<td>39.744</td>
<td>6.798</td>
<td>4.798</td>
<td>18.731</td>
<td>6.875</td>
<td>0.057</td>
</tr>
<tr>
<td></td>
<td>5% ↓</td>
<td>35.022</td>
<td>219.287</td>
<td>40.707</td>
<td>61.510</td>
<td>61.531</td>
<td>13.582</td>
<td>59.542</td>
<td>0.502</td>
</tr>
<tr>
<td>Promotion Power</td>
<td>5% ↑</td>
<td>32.933</td>
<td>110.590</td>
<td>39.939</td>
<td>10.641</td>
<td>10.708</td>
<td>19.162</td>
<td>10.940</td>
<td>-0.097</td>
</tr>
<tr>
<td></td>
<td>5% ↓</td>
<td>35.958</td>
<td>181.861</td>
<td>42.956</td>
<td>51.898</td>
<td>58.308</td>
<td>14.339</td>
<td>47.905</td>
<td>0.342</td>
</tr>
<tr>
<td>Herding Effect</td>
<td>5% ↑</td>
<td>34.819</td>
<td>145.683</td>
<td>41.509</td>
<td>24.653</td>
<td>24.747</td>
<td>16.966</td>
<td>25.823</td>
<td>-0.172</td>
</tr>
<tr>
<td></td>
<td>5% ↓</td>
<td>35.174</td>
<td>156.000</td>
<td>40.804</td>
<td>24.246</td>
<td>28.102</td>
<td>16.159</td>
<td>25.216</td>
<td>0.739</td>
</tr>
<tr>
<td><strong>Project Parameters Sensitivity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attraction</td>
<td>Smaller difference</td>
<td>33.511</td>
<td>158.343</td>
<td>40.395</td>
<td>28.194</td>
<td>27.898</td>
<td>16.746</td>
<td>28.174</td>
<td>0.252</td>
</tr>
<tr>
<td></td>
<td>Larger difference</td>
<td>35.579</td>
<td>150.206</td>
<td>41.155</td>
<td>21.339</td>
<td>23.592</td>
<td>16.640</td>
<td>23.017</td>
<td>-0.580</td>
</tr>
<tr>
<td>Pledge</td>
<td>Smaller difference</td>
<td>34.343</td>
<td>149.283</td>
<td>40.302</td>
<td>31.015</td>
<td>30.080</td>
<td>16.642</td>
<td>28.170</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>Larger difference</td>
<td>30.599</td>
<td>140.589</td>
<td>36.521</td>
<td>18.565</td>
<td>19.772</td>
<td>12.545</td>
<td>19.592</td>
<td>0.102</td>
</tr>
</tbody>
</table>
Thank you for attending, any questions?

dong.li@lancaster.ac.uk