

# Условия существования петель скрытой обратной связи в рекомендательных системах

А. А. Пилькевич<sup>1</sup>, А. С. Хританков<sup>2</sup>  
anton39reg@mail.ru; anton.khritankov@gmail.com

В работе исследуется эффект петель скрытой обратной связи в рекомендательных системах. Под положительной обратной связью подразумевается неограниченный рост интереса пользователя к предлагаемым объектам. Решается задача поиска условий возникновения положительной обратной связи для системы с алгоритмом Thomson Sampling Multi-armed Bandit с учётом наличия шума в выборе пользователя. В задачах без шума известно, что существуют условия неограниченного роста. Но отсутствие не реализуются в реальных системах. Экспериментально проверяются полученные условия в имитационной модели.

**Ключевые слова:** *machine learning; hidden feedback loops; echo chamber; filter bubble*

**DOI:**

## 1 Введение

Рекомендательные системы являются важной составляющей социальных сетей, веб-поиска и других сфер [...]. Рассматривается эффект петель скрытой обратной связи, который подразумевает рост качества предсказаний, как результат учёта принятых решений. Эффект петель скрытой обратной связи в реальных и модельных задачах в публикациях [...] описывается как нежелательное явление. Частные и часто рассматриваемые случаи скрытой обратной являются echo chamber и filter bubbles [1]. До сих пор нет какой-либо строгой формализации условий возникновения этих эффектов при условиях приближенных к реальности.

Целью данной работы является нахождение условий существования петель обратной связи в рекомендательной системе с алгоритмом Thomson Sampling в условиях зашумлённости выбора пользователя. Зашумлённость выбора рассматривается, как смещение первоначального интереса к исходному объекту или категории. Предлагается способ отыскание условий модели исходя из теоретических свойств алгоритма TS. Для описания условия предлагается рекурентное соотношение для регардов. Также рассматривается вариант нахождения этих условий чисто из экспериментов. Наибольший интерес представляет математическое описание искомых условий с дальнейшим экспериментальным подтверждением полученных условий. Для проверки результатов используется имитационная модель, использующая синтетические данные.

Уже существует модель [1] этого эффекта в случае отсутствия шума в действиях пользователя, что не реализуется на практике. Подобное исследование проводилось в статье [1] на примере различных моделей ( Oracle, Optimal Oracle, UCB, TS ) в задаче многорукого бандита. Им удалось показать условия существования неограниченного роста интереса пользователя. В работе [2] изучалась схожая постановка задачи и были получены условия возникновения, но рассматривалась линейная модель и градиентный бустинг. Важным отличием данной работы является факт рассмотрения более сложных условий модели, таких как шум в выборе пользователя и другой алгоритм рекомендательной системы.

## 2 Постановка задачи

### 2.1 Модель рекомендательной системы

Рекомендательная система выбирает элементы  $(a^1, \dots, a^l)$  из конечного набора  $M$ . Обозначим за  $t$  очередной момент выдачи рекомендаций. Истинный *интерес* пользователя к элементу  $a \in M$  описывается неизветсной функцией  $\mu_t : M \rightarrow \mathbb{R}$ . При этом, считается, что чем больше значени  $\mu_t(a)$ , тем заинтересованние пользователь в рекомендации  $a$ .

После очередного набора рекомендаций  $a_t = (a_t^1, \dots, a_t^l)$  пользователь возвращает *отклик*  $c_t = (c_t^1, \dots, c_t^l)$ . В отсутствие шума в ответах пользователя, он выбирает элементы случайно и независимы, пропорционально  $\mu_t(a)$ . Значит отклик будет имеет распределение Бернулли :  $c_t^i \sim \text{Bern}(\sigma(\mu_t(a_t^i)))$ , где  $\sigma$  — сигмоида.

Потребуем от пользователя, что интерес во времени описывается как  $\mu_{t+1} \geq \mu_t$ , если  $c_t = 1$ ,  $\mu_{t+1} < \mu_t$  иначе. Тогда эффект обратной связи выражается как  $\lim_{t \rightarrow \infty} \|\mu_t - \mu_0\|_2 = \infty$ . Обновление интереса происходит по правилу :  $\mu_{t+1} - \mu_t = \delta_t c_t - \delta_t (1 - c_t)$ , где  $\delta_t \sim U[0, 0.01]$

Оптимизационной задачей рекомендательной системы является минимизация регрета. Максимальная сумма ревардов :

$$\max_{c_t^i} \sum_{t=1}^T \sum_{i=1}^l c_t^i = T \cdot l.$$

Тогда задача ставится так :

$$T \cdot l - \sum_{t=1}^T \sum_{i=1}^l c_t^i \rightarrow \min_b,$$

где  $b$  — используемый алгоритм в рекомендательной системе.

### 2.2 Алгоритм

В данной задаче рекомендательная система будет использовать алгоритм Thompson Sampling для задачи Бернуллиевского бандита. Бандитами являются отклики пользователя на очередую рекомендацию, а средним ревардов:  $\sigma(\mu_t(a_t^i))$ .

В начальный момент времени определены вероятности Бернуллиевских случайных величин для элементов  $M$  равные  $\pi_0(\theta_1), \dots, \pi_0(\theta_m)$ . Задаётся априорное распределение для  $\theta_i$  равное Бэта распределению  $\text{Beta}(1, 1) = U[0, 1]$ . Апостериорное распределение для элемента  $a^i \in M$  будет описываться бэта распределением:  $\text{Beta}(\alpha_t^i, \beta_t^i)$ . А параметры будут обновляться по закону :  $\alpha_{t+1} = \alpha_t + c_t$ ,  $\beta_{t+1} = \beta_t + 1 - c_t$ .

### 2.3 Учёт шума в поведении пользователя

Шум откликов будет описываться следующим образом:

$$\begin{aligned} c_t^i &\sim \text{Bern}(\sigma(s_t^i \cdot \mu_t(a_t^i) + w \cdot q_t^i)), \\ P(s_t^i = 1) &= p, \\ P(s_t^i = -1) &= 1 - p, \\ q_t^i &\sim \text{Beta}(Q, Q). \end{aligned}$$

Величина  $p$  должна быть равна доле шума в ответах.

Также рассматривается случай марковского процесса:

$$\begin{aligned} P(s_t^i = 1 | s_{t-1}^i = 1) &= \min(1, p + u), \\ P(s_t^i = 1 | s_{t-1}^i = -1) &= \max(0, p - u). \end{aligned}$$

## 2.4 Цель

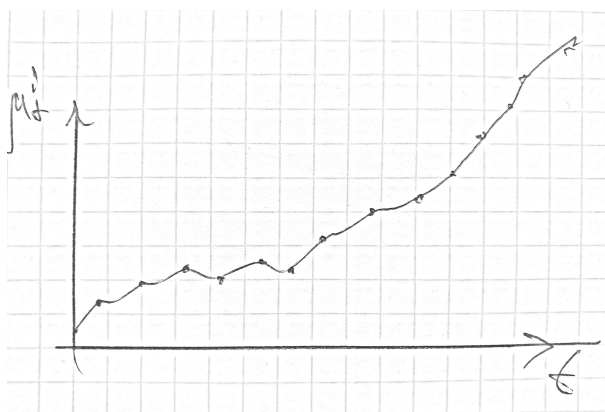
Целью работы является теоретический анализ условий сходимости TS для различных параметров шума  $p, Q, w, u$  и экспериментальное подтверждение полученных соотношений. Также делается уточнений условий из [1].

## 3 Вычислительный эксперимент

Перед началом эксперимента фиксируются следующие параметры:  $T$  — кол-во итераций рекомендательной системы,  $|M|$  — кол-во рассматриваемых объектов для рекомендации,  $l$  — кол-во элементов в одной выдаче. Также фиксируются параметры шума  $p, w, Q, u$ . Далее случайным образом сэмпляются начальное значение интереса  $\{\mu_0^i\}_{i=1}^{|M|}$  и параметры апериорного распределения  $\{\alpha_0^i, \beta_0^i\}_{i=1}^{|M|}$ .

Генерация элементов очередной рекомендации производится на основе текущего апостериорного распределения. Выбираются элементы с наибольшим значением. Получение отклика от пользователя заключается в генерации случайных величин на основе рекомендации. Далее идёт обновление параметров апостериорного распределения, и алгоритм начинает генерировать следующую рекомендацию.

В каждый момент выдачи фиксируются значения интереса  $\mu_t^i$ , сумма откликов  $c_t^i$  и параметры апостериорного распределения. По полученным данным строятся графики для определения наличия скрытой обратной связи. Наличие скрытой обратной связи иллюстрируется следующим графиком:



Важной частью эксперимента является сравнения поведений рекомендательной системы с шумом в ответах пользователя и без.

## Литература

- [1] Ray Jiang, Silvia Chiappa, Tor Lattimore, András György, Pushmeet Kohli Degenerate Feedback Loops in Recommender Systems// CoRR, 2019, Vol. abs/1902.10730, URL: <https://arxiv.org/abs/1902.10730>.
- [2] Khritankov, Anton Hidden Feedback Loops in Machine Learning Systems: A simulation Model and Preliminary Results// Springer, 2021, P. 54–65,
- [3] Daniel Russo, Benjamin Van Roy, Abbas Kazerouni, Ian Osband A Tutorial on Thompson Sampling// CoRR, 2017, Vol. abs/1707.02038, URL: <https://arxiv.org/abs/1707.02038>.
- [4] Shipra Agrawal, Navin Goyal Analysis of Thompson Sampling for the multi-armed// CoRR, 2011, Vol. abs/1111.1797, URL: <https://arxiv.org/abs/1111.1797>.

Поступила в редакцию