

# 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. Решается задача поиска условий возникновения положительной обратной связи для системы с алгоритмом Thomson Sampling Multi-armed Bandit с учётом наличия шума в выборе пользователя. Под положительной обратной связью подразумевается неограниченный рост интереса пользователя к предлагаемым объектам. Без шума известно, что всегда существуют условия неограниченного роста. 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. Мы будем рассматривать 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 во многих публикациях описывается как нежелательное явление. Частные и часто рассматриваемые случаи скрытой обратной связи являются echo chamber и filter bubbles. 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 требуемых условий модели исходя из теоретических свойств алгоритма TS путём нахождения рекуррентного соотношения для регрессов. Также рассматривается вариант нахождение этих условий чисто из экспериментов. Наибольший интерес представляет математическое описание искомых условий с дальнейшим экспериментальным подтверждением полученных соотношений. Для проверки результатов используется имитационная модель, использующая синтетические данные.

Уже существует модель этого эффекта в случае отсутствия шума в действиях пользователя, что не реализуется на практике. Подобное исследование проводилось в статье [1] на примере различных моделей 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 (GBR). Важным отличием нашей работы является факт рассмотрения более сложных и приближенных условий модели, таких как шум в выборе пользователя и другой алгоритм рекомендательной системы.

## 2 Постановка задачи

### 2.1 Данные

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 ф-й  $\mu_t : M \rightarrow R$ . Чем больше значение  $\mu_t(a)$ , тем заинтересованнее пользователь.

После очередной рекомендации  $a_t = (a_t^1, \dots, a_t^l)$  пользователь возвращает "отклик"  $c_t = (c_t^1, \dots, c_t^l)$ . В отсутствии шума в ответах пользователя отклик будет иметь распределение Бернулли :  $c_t^i \sim \text{Bern}(\sigma(\mu_t(a_t^i)))$ , где  $\sigma$  — сигмоида.

Эволюция интереса во времени описывается как  $\mu_{t+1} > \mu_t$ , 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 :  $\mu_{t+1} - \mu_t = \delta_t I\{c_t = 1\} - \delta_t I\{c_t = 0\}$ , где  $\delta_t \sim U[0, 0.01]$

### 2.2 Thompson Sampling for the Bernoulli Bandit

В данной задаче рекомендательная система будет использовать алгоритм Thompson Sampling. Бандитами являются отклики пользователя на очередную 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. Согласно байесовской теории апостериорное распределение для элемента  $a \in M$  будет описываться  $\text{Beta}(\alpha + c, \beta + 1 - c)$ . А параметры будут обновляться по закону :  $\alpha_{t+1} = \alpha_t + c_t, \beta_{t+1} = \beta_t + 1 - c_t$ .

Оптимизационной задачей алгоритма является максимизация суммы ревардов:

$$\sum_{t=1}^T \sum_{i=1}^l c_t^i \rightarrow \max_{???}$$

### 2.3 Шум

Шум откликов будет описываться следующим образом:

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

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

### 2.4 Цель

Целью работы является анализ решения алгоритма TS и нахождение начальных условий  $\mu_0^i, \pi_0$  и параметра шума  $p$ .

## 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*