

# 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:

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

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  $p, w, u$  и экспериментальное подтверждение полученных соотношений. Также делается уточнений условий из [1].

## 3 Теоретическое обоснование

**Утверждение 1.** При фиксированном  $p = 1$  и любых параметрах шума  $w$  возникает петля скрытой обратной связи.

**Доказательство.** При достаточно большом  $t$  бандит имеет очень узкое распределение, поэтому точно известно, что он буде рекомендовать. Для случая нормы интересов:

$$\|\mu_t - \mu_0\|_2^2 = \sum_{i=1}^M (\mu_t^i - \mu_0^i)^2,$$

с ростом  $t$  сновной вклад будут давать только  $l$  объектов.

Тогда рассмотрим поведение интереса для произвольного фиксированного объекта  $a \in M$ . Обновление интереса происходит согласно:  $\mu_t - \mu_{t-1} = \delta_t c_t - \delta(1 - c_t)$ .

Случайные величины  $\delta_t, c_t$  независимы, поэтому:  $E\delta_t c_t = E\delta_t E c_t$ . Для удобства будем считать, что у нас  $c_t \sim \text{Bern}_{\pm}(\sigma(s_t \cdot \mu_t(a_t) + q_t))$  Тогда:

$$\begin{aligned} E(c_t | s_t = x, q_t = y) &= 2\sigma(x \cdot E\mu_{t-1} + y) - 1, \\ E(E(c_t | s_t, q_t = y)) &= p \cdot (2\sigma(E\mu_{t-1} + y) - 1) + (1 - p) \cdot (2\sigma(-E\mu_{t-1} + y) - 1). \end{aligned}$$

Далее для простоты считается, что  $\sigma(x) \approx \text{ReLU}(x)$ . Поэтому:

$$\begin{aligned} \int_{-w}^w \text{ReLU}(E\mu_{t-1} + y) \cdot \frac{1}{2w} dy &= \begin{cases} 0, & \text{если } -E\mu_{t-1} > w, \\ E\mu_{t-1}, & \text{если } E\mu_{t-1} > w, \\ \frac{(E\mu_{t-1} + w)^2}{4w}, & \text{если } |E\mu_{t-1}| \leq w. \end{cases} \\ \int_{-w}^w \text{ReLU}(-E\mu_{t-1} + y) \cdot \frac{1}{2w} dy &= \begin{cases} -E\mu_{t-1}, & \text{если } -E\mu_{t-1} > w, \\ 0, & \text{если } E\mu_{t-1} > w, \\ \frac{(-E\mu_{t-1} + w)^2}{4w}, & \text{если } |E\mu_{t-1}| \leq w. \end{cases} \end{aligned}$$

Отсюда можно получить рекуррентное соотношение:

$$E\mu_t - E\mu_{t-1} = -E\delta_t + 2E\delta_t \cdot \begin{cases} (p - 1) \cdot E\mu_{t-1}, & \text{если } -E\mu_{t-1} > w, \\ p \cdot E\mu_{t-1}, & \text{если } E\mu_{t-1} > w, \\ \frac{1}{4w} (4pwE\mu_{t-1} + (E\mu_{t-1} - w)^2), & \text{если } |E\mu_{t-1}| \leq w. \end{cases}$$

Решением рекуррентного соотношения являются:

$$E\mu_t = \begin{cases} (1 + 2(p - 1)E\delta_t)^t \cdot \mu_0 - E\delta_t \sum_{i=0}^{t-1} (1 + 2(p - 1)E\delta_t)^i, & \text{если } -E\mu_{t-1} > w, \\ (2pE\delta_t + 1)^t \cdot \mu_0 + \sum_{i=0}^{t-1} (2pE\delta_t + 1)^i \cdot (1 - 2p)E\delta_t, & \text{если } E\mu_{t-1} > w, \\ \dots, & \text{если } |E\mu_{t-1}| \leq w. \end{cases}$$

Рассмотрим в каком случае мы не покинем условие  $-E\mu_{t-1} > w$ . Для этого должно выполняться:

$$E\mu_t - E\mu_{t-1} = -E\delta_t + 2E\delta_t \cdot (p-1)E\mu_{t-1} < 0 \Rightarrow \\ E\mu_{t-1} < \frac{1}{2(p-1)}.$$

Аналогично получаем, что для выполнения  $E\mu_{t-1} > w$  требуется:

$$E\mu_t - E\mu_{t-1} = -E\delta_t + 2E\delta_t \cdot pE\mu_{t-1} > 0 \Rightarrow \\ E\mu_{t-1} > \frac{1}{2p}.$$

Очевидно, при отсутствии шума, то есть при  $p = 1$  условия всегда верны. Поэтому будет наблюдаться монотонное поведения интереса.

## 4 Вычислительный эксперимент

### 4.1 Описание данных и работы модели

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  $p, w, u$ . Далее случайным образом сэмпятся начальное значение интереса  $\{\mu_0^i\}_{i=1}^{|M|}$  и параметры априорного 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 по правилу  $\alpha_{t+1} = \alpha_t + c_t, \beta_{t+1} = \beta_t + 1 - c_t$ , и обновления интереса согласно  $\mu_{t+1} - \mu_t = \delta_t c_t - \delta_t (1 - c_t)$ .

Также рассматривается вариант эксперимента, когда рассматривается рандомная модель генерации рекомендации. То есть случайным образом выбираются  $l$  элементов для очередной рекомендации.

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 петель скрытой обратной связи. Как определялось раньше, петля скрытой обратной связи выражается так:  $\lim_{t \rightarrow \infty} \|\mu_t - \mu_0\|_2 = \infty$ .

### 4.2 Цель

Целью эксперимента является наблюдение петель скрытой обратной связи для определённых параметров шума. Проверяется гипотеза о возникновении петель при параметрах шума, найденных из теоретических соотношений. Важной частью эксперимента является сравнения поведения рекомендательной системы с шумом в ответах пользователя и без.

### 4.3 Псевдокод

**Вход:**  $M, l, T, w, p$   
 BanditLoopExperiment.prepare()  
 для  $t$  от 1 до  $T$   
      $r_t \leftarrow \text{TSBandit.predict}()$   
      $c_t \leftarrow \text{make\_response\_noise}(r_t, w, p)$

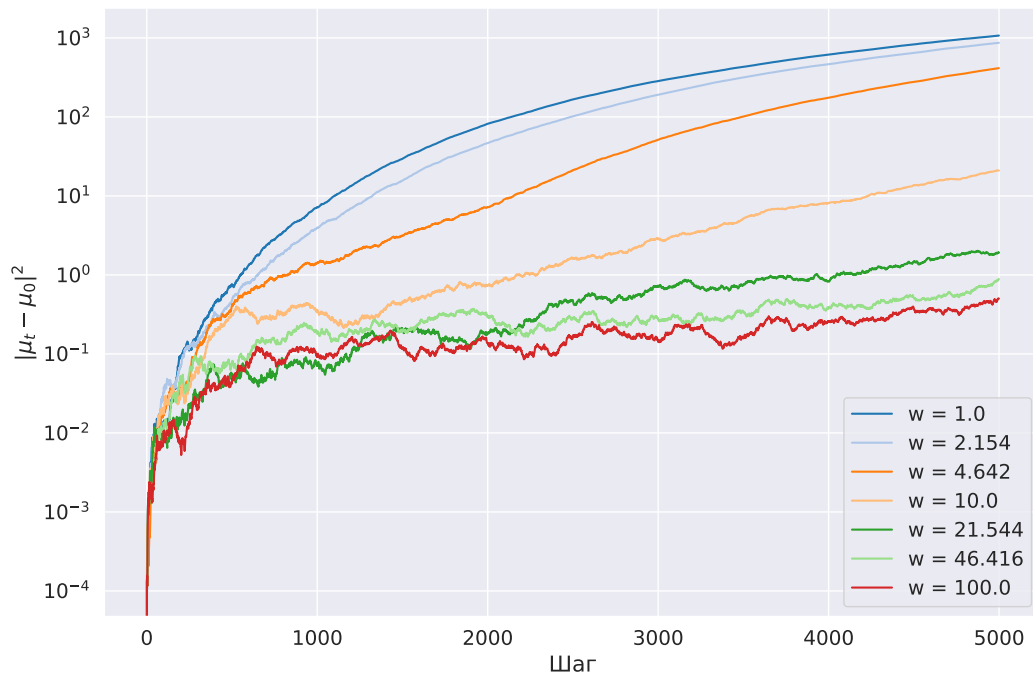
```

100     TSBandit.update( $c_t$ )
101     Model.interest_update( $c_t$ )
102     save_iter( $t, c_t, \mu_t$ )

```

## 103 5 Результаты

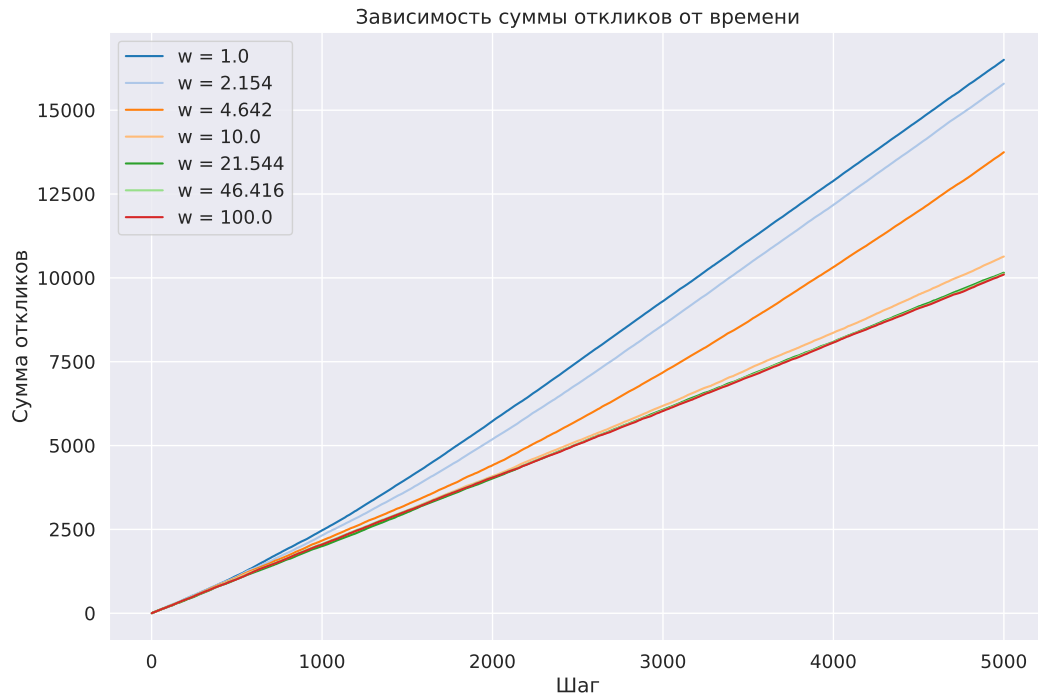
104 При определённом диапазоне параметров шума удалось про наблюдать эффект неогра-  
 105 ниченного роста интереса. Наличие петель скрытой обратной связи иллюстрируется сле-  
 106 дующим графиком:



107

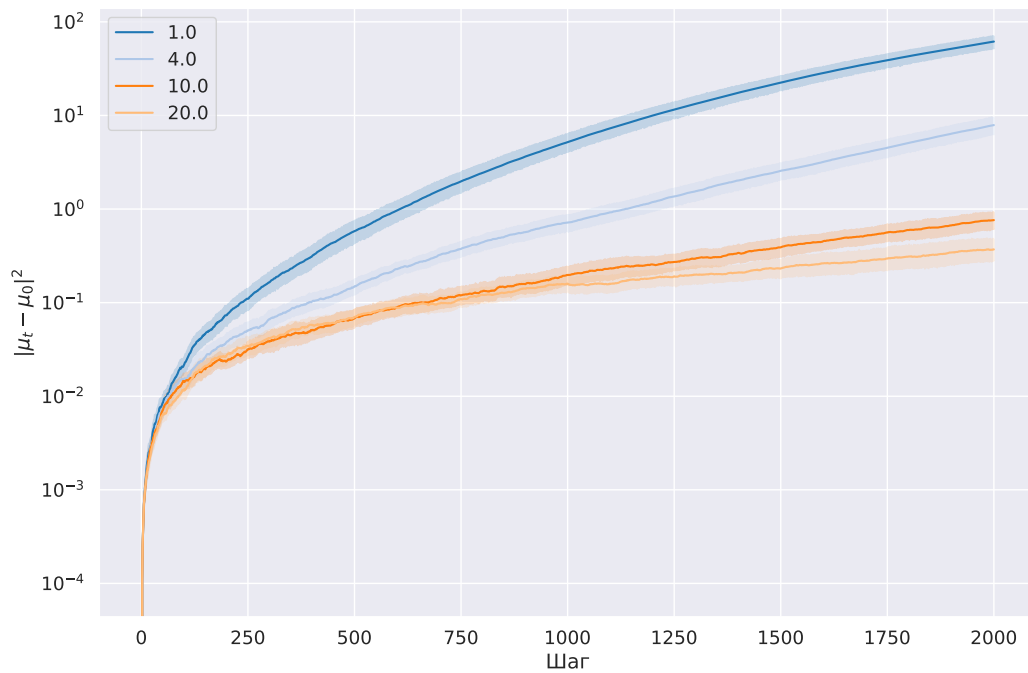
108 Как видим, это согласуется с определением петель.

109 Из графика суммы ревардов видим, что с определённого момента кривые начинают  
 110 идти параллельно максимально возможной сумме. Это также свидетельствует о наличии  
 111 петель обратной связи.



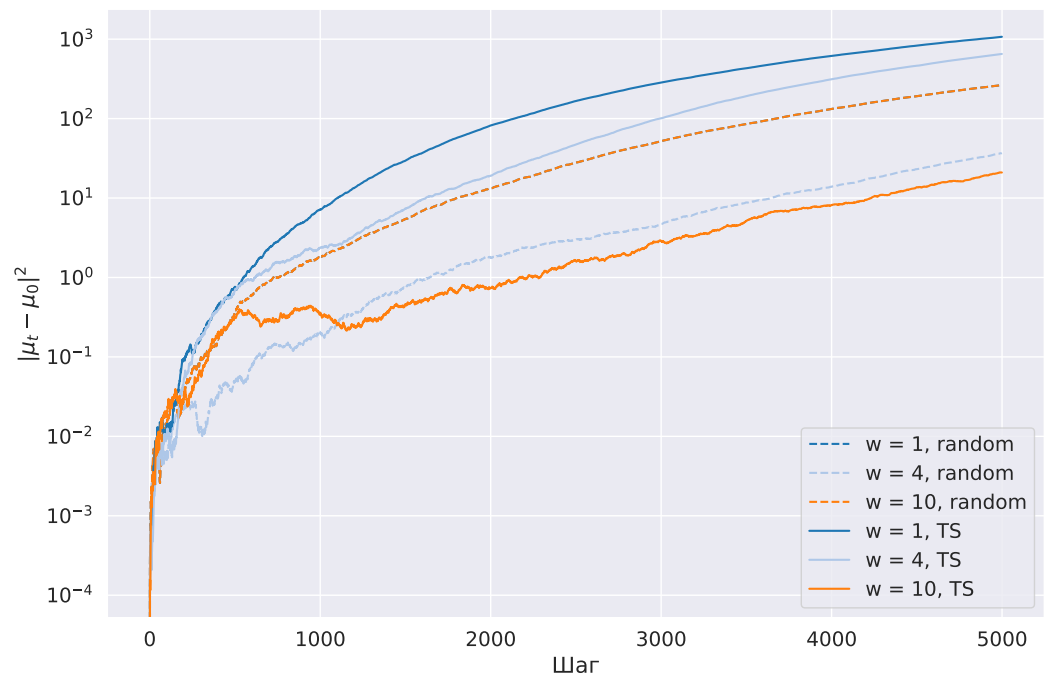
112

113      Разброс значений нормы интересов.

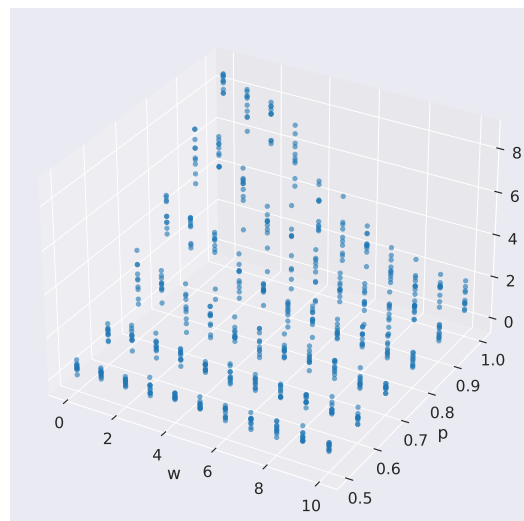


114

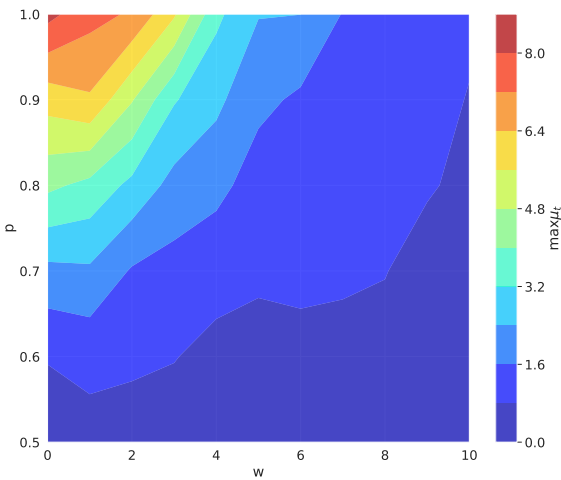
115      Для случайной модели не наблюдается явного образования петли. Она более хаотична,  
 116 но тренд неограниченного роста интереса всё равно присутствует.



117



(a)



(б)

## 6 Заключение

## 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*