

Условия существования петель скрытой обратной связи в рекомендательных системах

А. А. Пилькевич¹, А. С. Хританков²
anton39reg@mail.ru; anton.khritankov@gmail.com

В работе исследуются петли 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. Учитывается наличие шума в выборе пользователя. Рекомендательная система использует алгоритм Thomson Sampling Multi-armed Bandit. В задачах без шума известно, что существуют условия неограниченного роста. Но отсутствие шума не реализуется в реальных системах. Экспериментально проверяются полученные условия в имитационной модели.

Ключевые слова: *machine learning; hidden feedback loops*

DOI:

1 Введение

Рекомендательные системы являются важной составляющей социальных сетей, веб-поиска и других сфер [...]. Рассматриваются петли скрытой обратной связи, которые 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 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. Под условиями подразумеваются параметры шума и параметры рекомендательной системы. Для описания условий предлагается выражение для математического ожидания интереса. Также рассматривается вариант нахождения этих условий чисто из экспериментов. Целью является математическое описание искомых условий с дальнейшим экспериментальным подтверждением полученных условий. Для проверки результатов используется имитационная модель, использующая синтетические данные.

Существует ранее описанная модель [1] петель в случае отсутствия шума в действиях пользователя. Подобное исследование проводилось в статье [1] на примере различных моделей (Oracle, Optimal Oracle, UCB, TS) в задаче многорукого бандита. 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. Но отсутствие шума в ответах пользователей в работах [1,2] не реализуется на практике. Важным отличием данной работы является факт рассмотрения более сложных условий модели, таких как шум в выборе пользователя и другой алгоритм рекомендательной системы.

В работе предлагается анализ роста интереса пользователя. Рассматривается математическое ожидание изменения интереса. Полученные условия проверяются в вычислительном эксперименте.

2 Петли скрытой обратной связи

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

2.1 Модель рекомендательной системы

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

После очередного набора рекомендаций $a_t = (a_t^1, \dots, a_t^l)$ пользователь возвращает *отклик* $c_t = (c_t^1, \dots, c_t^l)$, $c_t^i \in \{0, 1\}$. Предполагается, он выбирает элементы c_t^i случайно и независимо, пропорционально $\mu_t(a)$. Значит отклик имеет распределение Бернулли :

$$c_t^i \sim \text{Bern}(\sigma(\mu_t(a_t^i))), \text{ где } \sigma(x) = \frac{1}{1 + x} \text{ — сигмоида.}$$

Предполагаем, что интерес пользователя во времени описывается как

$$\begin{cases} \mu_{t+1} \geq \mu_t, & \text{если } c_t = 1, \\ \mu_{t+1} < \mu_t, & \text{иначе.} \end{cases}$$

Тогда петля обратной связи выражается как

$$\lim_{t \rightarrow \infty} \|\mu_t - \mu_0\|_2 = \infty.$$

Обновление интереса происходит по правилу :

$$\mu_{t+1} - \mu_t = \delta_t c_t - \delta_t (1 - c_t), \text{ где } \delta_t \sim U[0, 0.01].$$

Оптимизационной задачей рекомендательной системы является задача минимизации потерь. Максимальная сумма наград :

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

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 рекомендательной системы

В данной задаче рекомендательная система использует алгоритм Thompson Sampling [3] для задачи бернуллиевского бандита. Бандитами являются отклики пользователя c_t^i на очередую рекомендацию. Средняя награда равна: $\sigma(\mu_t(a_t^i))$.

В начальный момент времени определены вероятности бернуллиевских случайных величин c_t^i для элементов M равные $\pi_0(\theta_1), \dots, \pi_0(\theta_m)$. Задаётся априорное распределение для θ_i равное бэ́та-распределению $Beta(1, 1) = U[0, 1]$. Апостериорное распределение для элемента $a^i \in M$ описывается бэ́та-распределением: $Beta(\alpha_t^i, \beta_t^i)$. Параметры после очередной рекомендации обновляются 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 описывается следующим образом:

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

Наличие q_t^i позволяет описать несмещённый аддитивный шум, то есть отклонение от истинного интереса пользователь. А s_t^i описывает кардинальное изменение интереса на противоположный.

3 Теоретическое обоснование

Утверждение 1. Пусть $p = 1$. Тогда при $w \geq 0 : \lim_{t \rightarrow \infty} \|\mu_t - \mu_0\|_2 = \infty$.

Или же при фиксированном $p = 1$ и любых параметрах шума w возникает петля скрытой обратной связи.

Доказательство. При достаточно большом t бандит имеет ограниченную дисперсию апостериорного распределения, поэтому точно известно, что он будет 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.

Рассмотрим изменение 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 δ_t, c_t независимы, поэтому:

$$\mathbb{E} \delta_t c_t = \mathbb{E} \delta_t \mathbb{E} c_t$$

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

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

В случае $\mathbb{E}(\mathbb{E}(c_t | s_t, q_t)) > 0$ петля будет возникать, так как рост интереса в среднем положителен.

Далее для простоты считается, что $\sigma(x) \approx (\frac{x}{4} + \frac{1}{2}) \cdot I[-2, 2] + I[2, \infty]$ и $p = 1$. Задача в этом случае записывается так:

$$\mathbb{E}(\mathbb{E}(c_t | s_t, q_t = y)) \approx 2 \left(\frac{\mathbb{E} \mu_{t-1} + y}{4} + \frac{1}{2} \right) - 1.$$

82 Теперь петля возникает при условии: $\mathbf{E}\sigma(x) > \frac{1}{2}$.

Тогда остаётся посчитать:

$$\begin{aligned}\mathbf{E}\sigma(\mu_t) &\approx \int_{-\infty}^{\infty} \left(\frac{\mathbf{E}\mu_t + y}{4} + \frac{1}{2} \right) I\{-2 < \mathbf{E}\mu_t + y < 2\} f(y) dy \\ &\quad + \int_{-\infty}^{\infty} I\{2 < \mathbf{E}\mu_t + y\} f(y) dy = \\ &\quad \int_{-2}^2 \left(\frac{z}{4} + \frac{1}{2} \right) f_s(z) dz + \int_2^{\infty} f_s(z) dz,\end{aligned}$$

83 где $f_s(z)$ плотность $U[\mathbf{E}\mu_t - w, \mathbf{E}\mu_t + w]$. Таким образом у нас возникает 6 случаев.

1. $\mathbf{E}\mu_t + w < -2$. Тогда, очевидно:

$$\mathbf{E}\sigma(\mu_t) = 0.$$

2. $\mathbf{E}\mu_t - w < -2 < \mathbf{E}\mu_t + w < 2$. Тогда:

$$\begin{aligned}\mathbf{E}\sigma(\mu_t) &= \frac{1}{16w} (y+2)^2 \Big|_{-2}^{\mathbf{E}\mu_t+w} = \frac{1}{16w} (\mathbf{E}\mu_t + w + 2)^2 > \frac{1}{2}, \\ &\quad (\mathbf{E}\mu_t + w + 2)^2 > 8w, \\ &\quad \begin{cases} \mathbf{E}\mu_t > -w - 2 + \sqrt{8w}, \\ \mathbf{E}\mu_t < -w - 2 - \sqrt{8w}, \end{cases} \rightarrow \text{рост.}\end{aligned}$$

3. $\mathbf{E}\mu_t - w < -2, \mathbf{E}\mu_t + w > 2$.

$$\begin{aligned}\mathbf{E}\sigma(\mu_t) &= \frac{1}{16w} (y+2)^2 \Big|_{-2}^2 + \frac{1}{2w} (\mathbf{E}\mu_t + w - 2) = \\ &\quad \frac{1}{w} + \frac{\mathbf{E}\mu_t + w}{2w} - \frac{1}{w} = \frac{\mathbf{E}\mu_t + w}{2w} > \frac{1}{2} \Rightarrow \\ &\quad \mathbf{E}\mu_t > 0, w > 2 \rightarrow \text{рост.}\end{aligned}$$

4. $\mathbf{E}\mu_t - w > -2, \mathbf{E}\mu_t + w < 2$. Тогда:

$$\begin{aligned}\mathbf{E}\sigma(\mu_t) &= \frac{1}{16w} (y+2)^2 \Big|_{\mathbf{E}\mu_t-w}^{\mathbf{E}\mu_t+w} > \frac{1}{2}, \\ &\quad (\mathbf{E}\mu_t + w + 2)^2 - (\mathbf{E}\mu_t - w + 2)^2 > 8w, \\ &\quad (2\mathbf{E}\mu_t + 4) \cdot 2w > 8w, \\ &\quad \mathbf{E}\mu_t > 0 \rightarrow \text{рост.}\end{aligned}$$

5. $E\mu_t - w > -2, E\mu_t + w > 2$. Тогда:

$$\begin{aligned}
 E\sigma(\mu_t) &= \frac{1}{16w} (y+2)^2 \Big|_{E\mu_t-w}^2 + \frac{1}{2w} \Big|_{E\mu_t-w}^{E\mu_t+w} = \\
 &= \frac{1}{16w} (16 - (E\mu_t - w + 2)^2) + \frac{1}{2w} (E\mu_t + w - 2) = \\
 &= \frac{1}{w} - \frac{(E\mu_t - w + 2)^2}{16w} + \frac{E\mu_t + w}{2w} - \frac{1}{w} = \\
 &= -\frac{1}{16w} (E^2\mu_t - 2(w-2)E\mu_t + (w-2)^2) + \frac{E\mu_t + w}{2w} > \frac{1}{2} \Rightarrow \\
 &E^2\mu_t - 2(w-2)E\mu_t + (w-2)^2 - 8(E\mu_t + w) + 8w < 0, \\
 &E^2\mu_t - 2(w+2)E\mu_t + (w-2)^2 < 0, \\
 &(E\mu_t - (w+2))^2 - (w+2)^2 + (w-2)^2 < 0, \\
 &(E\mu_t - (w+2))^2 - 8w < 0, \\
 &\begin{cases} E\mu_t < w+2 + \sqrt{8w}, \\ E\mu_t > w+2 - \sqrt{8w}, \end{cases} \rightarrow \text{рост.}
 \end{aligned}$$

6. $E\mu_t - w > 2$. Тогда:

$$E\sigma(\mu_t) = 1 > \frac{1}{2}.$$

4 Вычислительный эксперимент

Целью эксперимента является подтверждение существования петель скрытой обратной связи для произвольных параметров шума w . Важной частью эксперимента является сравнения поведений рекомендательной системы с шумом в ответах пользователя и без.

4.1 Описание данных и работы модели

Перед началом эксперимента фиксируются следующие параметры: T — число итераций рекомендательной системы, $|M|$ — число рассматриваемых объектов для рекомендации, l — число элементов в одной выдаче. Также фиксируются параметры шума p, w, u . Далее случайным образом сэмпятся начальные значения интереса $\{\mu_0^i\}_{i=1}^{|M|}$. Параметры априорного распределения $\{\alpha_0^i, \beta_0^i\}_{i=1}^{|M|}$ также сэмпятся случайно.

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. Обновление параметров апостериорного распределения происходит по правилу $\alpha_{t+1} = \alpha_t + c_t, \beta_{t+1} = \beta_t + 1 - c_t$. Интерес обновляется согласно $\mu_{t+1} - \mu_t = \delta_t c_t - \delta_t (1 - c_t)$.

Также рассматривается вариант эксперимента, когда используется случайная модель генерации рекомендации. В этом случае l элементов для очередной рекомендации выбираются случайным образом.

В каждый момент выдачи t фиксируются значения интереса μ_t^i , сумма откликов c_t^i и параметры апостериорного распределения. По полученным данным строятся графики для определения наличия петель скрытой обратной связи (см. рис. ...). Как определялось раньше, петля скрытой обратной связи выражается так: $\lim_{t \rightarrow \infty} \|\mu_t - \mu_0\|_2 = \infty$.

4.2 Псевдокод проведения эксперимента

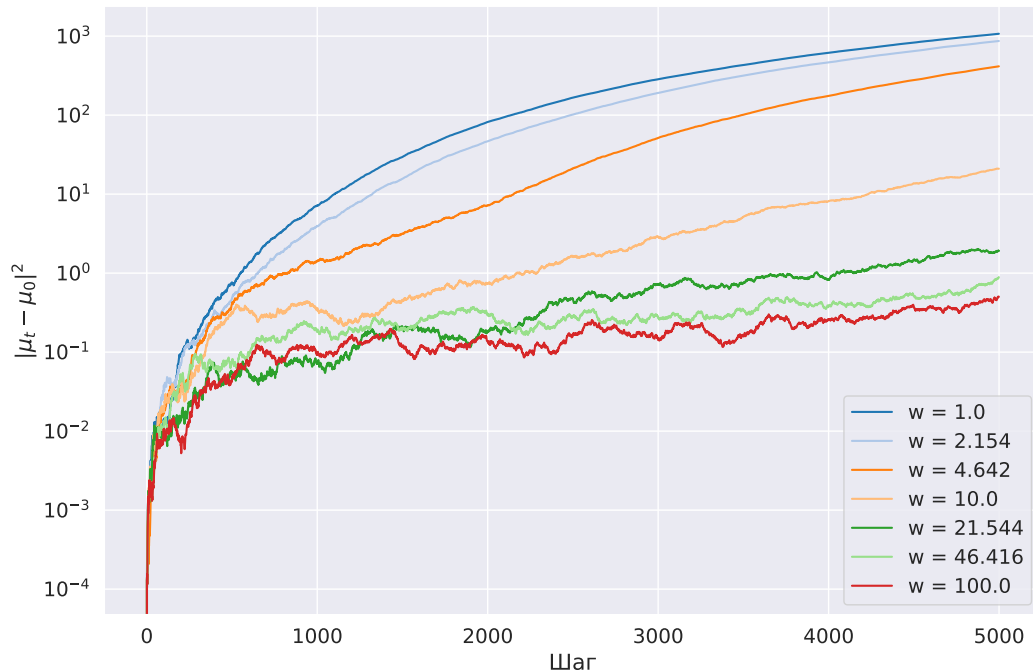
```

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

```

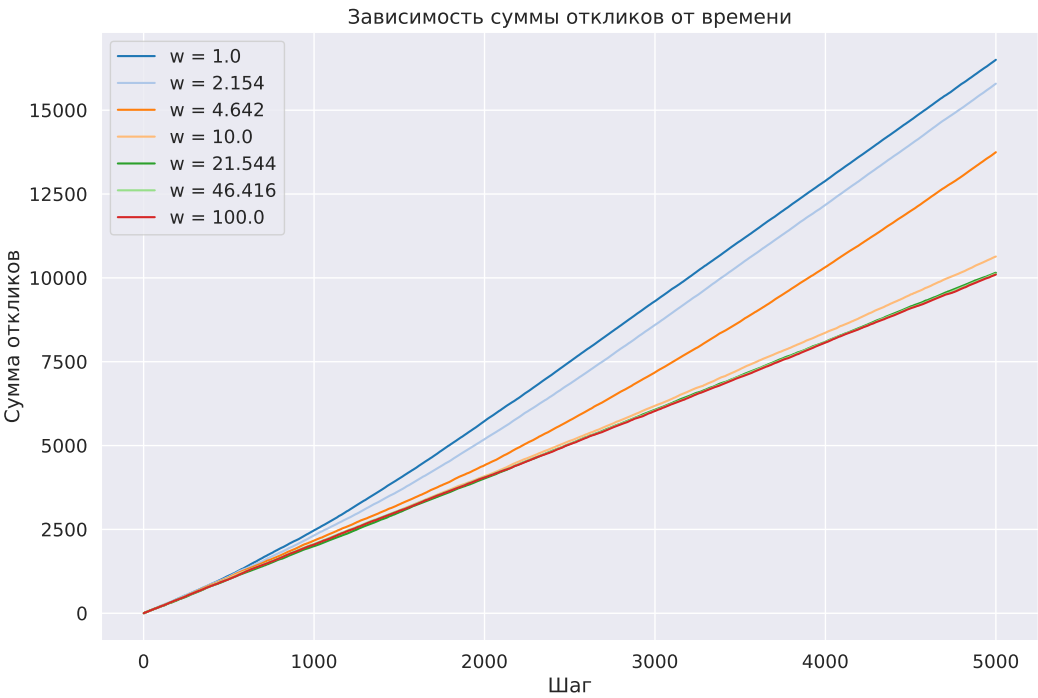
5 Результаты

Наблюдается эффект неограниченного роста интереса. Наличие петель скрытой обратной связи иллюстрируется следующим графиком:



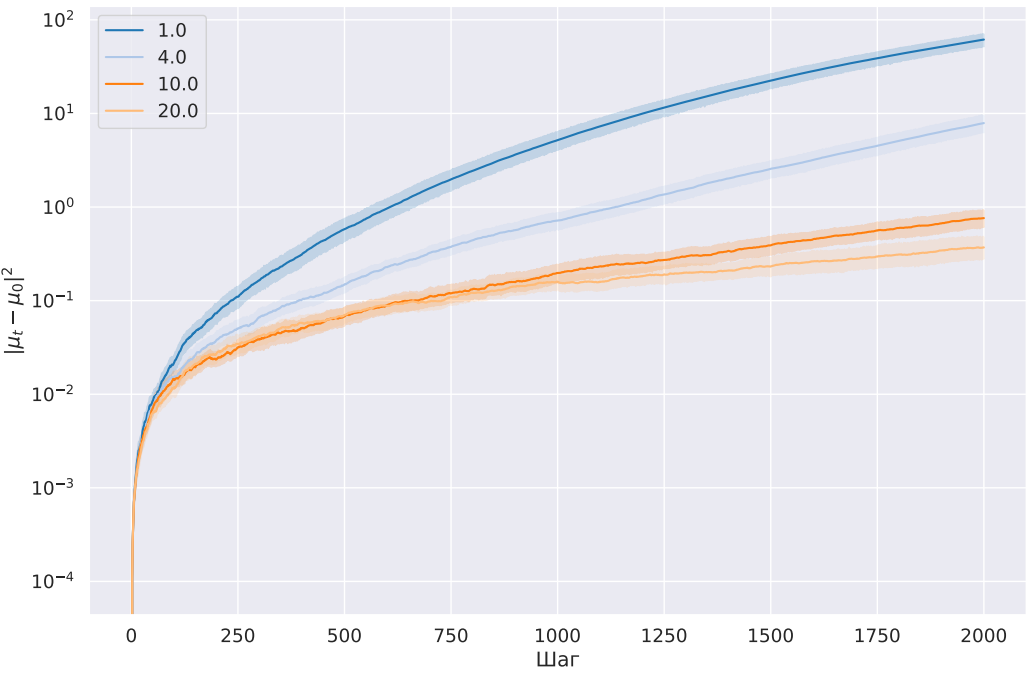
Это согласуется с определением петель.

Из графика суммы наград видно, что с определённого момента кривые начинают идти параллельно максимально возможной сумме. Это свидетельствует о наличии петель обратной связи.



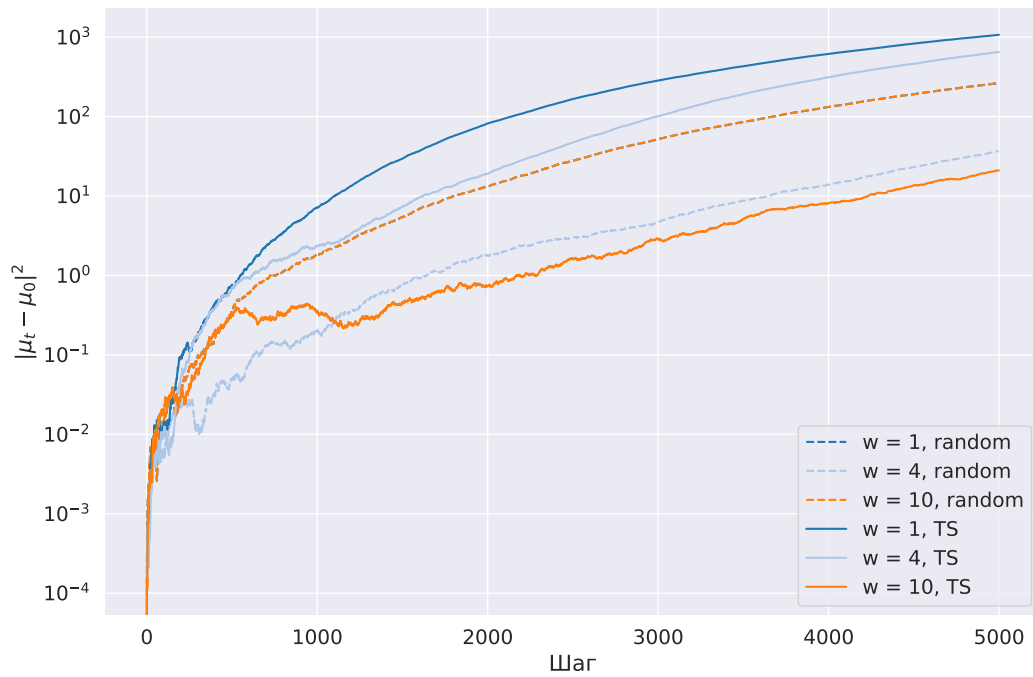
123

124 Разброс значений нормы интересов.

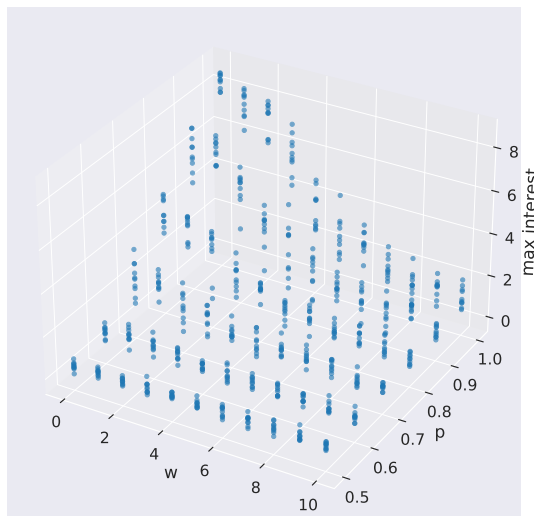


125

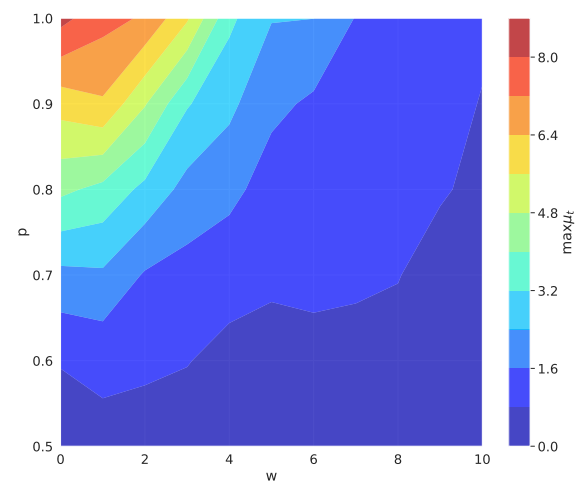
126 Для случайной модели тоже наблюдается образование петли. Она более хаотична, но
127 тренд неограниченного роста интереса всё равно присутствует.



128



(a)



(б)

129 Весь эксперимент и исходники расположены в гите:

130 <https://github.com/Intelligent-Systems-Phystech/2021-Project-74>.

6 Заключение

Поставлена задача существования петель скрытой обратной связи при наличии шума в ответах пользователя. Для текущей модели шума было получено, что при любых параметрах возникают петли. Это также подтверждается в эксперименте.

В дальнейшем требуется проверить гипотезу о возникновении петель при любом несмещённом аддитивном шума. Также стоит рассмотреть другие модели шума. Так как после определённого значения интереса из-за сигмной любое влияние шума сводилось на нет.

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