

Условия существования петель скрытой обратной связи в рекомендательных системах

А. А. Пилькевич¹, А. С. Хританков²
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 на шаге t выбирает элементы (a_t^1, \dots, a_t^l) из конечного набора M . Обозначим за t очередной момент выдачи рекомендаций. 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 $\mathbf{a}_t = (a_t^1, \dots, a_t^l)$ пользователь возвращает отклик $\mathbf{c}_t = (c_t^1, \dots, c_t^l)$, $c_t^i \in \{0, 1\}$. Предполагается, он выбирает элементы случайно и независимо, пропорционально $\mu_t(a)$. Значит отклик имеет распределение Бернулли : $c_t^i \sim \text{Bern}(\sigma(\mu_t(a_t^i)))$, где $\sigma(x) = \frac{1}{1+e^{-x}}$ — 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 : $\mu_{t+1} - \mu_t = \delta_t c_t - \delta_t (1 - c_t)$, где $\delta_t \sim U[0, 0.01]$.

Оптимизационной задачей рекомендательной системы является минимизация 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

В данной задаче рекомендательная система использует алгоритм Thompson Sampling [3] для задачи бернуллиевского бандита. Бандитами являются отклики пользователя c_t^i на очередную 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 бэта-распределению $Beta(1, 1) = U[0, 1]$. Апостериорное распределение для элемента $a^i \in M$ описывается бэта-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.

2.4 Цель

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

3 Теоретическое обоснование

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

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

Доказательство. При достаточно большом t бандин 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 δ_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:

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

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

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

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

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

Тогда остаётся посчитать:

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

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

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

$$E\sigma(\mu_t) = 0.$$

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

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

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

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

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

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

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

$$\begin{aligned} E\sigma(\mu_t) &= \frac{1}{16w}(y+2)^2 \Big|_{E\mu_t-w}^2 + \frac{1}{2w} \Big|_2^{E\mu_t+w} = \\ \frac{1}{16w} (16 - (E\mu_t - w + 2)^2) + \frac{1}{2w} (E\mu_t + w - 2) &= \\ \frac{1}{w} - \frac{(E\mu_t - w + 2)^2}{16w} + \frac{E\mu_t + w}{2w} - \frac{1}{w} &= \\ -\frac{1}{16w} (E^2\mu_t - 2(w-2)E\mu_t + (w-2)^2) + \frac{E\mu_t + w}{2w} &> \frac{1}{2} \Rightarrow \\ E^2\mu_t - 2(w-2)E\mu_t + (w-2)^2 - 8(E\mu_t + w) + 8w &< 0, \\ E^2\mu_t - 2(w+2)E\mu_t + (w-2)^2 &< 0, \\ (E\mu_t - (w+2))^2 - (w+2)^2 + (w-2)^2 &< 0, \\ (E\mu_t - (w+2))^2 - 8w &< 0, \\ \begin{cases} E\mu_t < w + 2 + \sqrt{8w}, \\ E\mu_t > w + 2 - \sqrt{8w}, \end{cases} &\rightarrow \text{рост.} \end{aligned}$$

6. $E\mu_t - w > 2$. Тогда:

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

4 Вычислительный эксперимент

4.1 Цель

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

4.2 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 сэмпляются начальные значения интереса $\{\mu_0^i\}_{i=1}^{|M|}$. Параметры априорного распределения $\{\alpha_0^i, \beta_0^i\}_{i=1}^{|M|}$ также сэмпляются 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 рассматривается случайная модель генерации рекомендации. То есть случайным образом выбираются l элементов для очередной рекомендации.

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

4.3 Псевдокод

Вход: M, l, T, w, p

BanditLoopExperiment.prepare()

для t от 1 до T

$r_t \leftarrow \text{TSBandit.predict}()$

$c_t \leftarrow \text{make_response_noise}(r_t, w, p)$

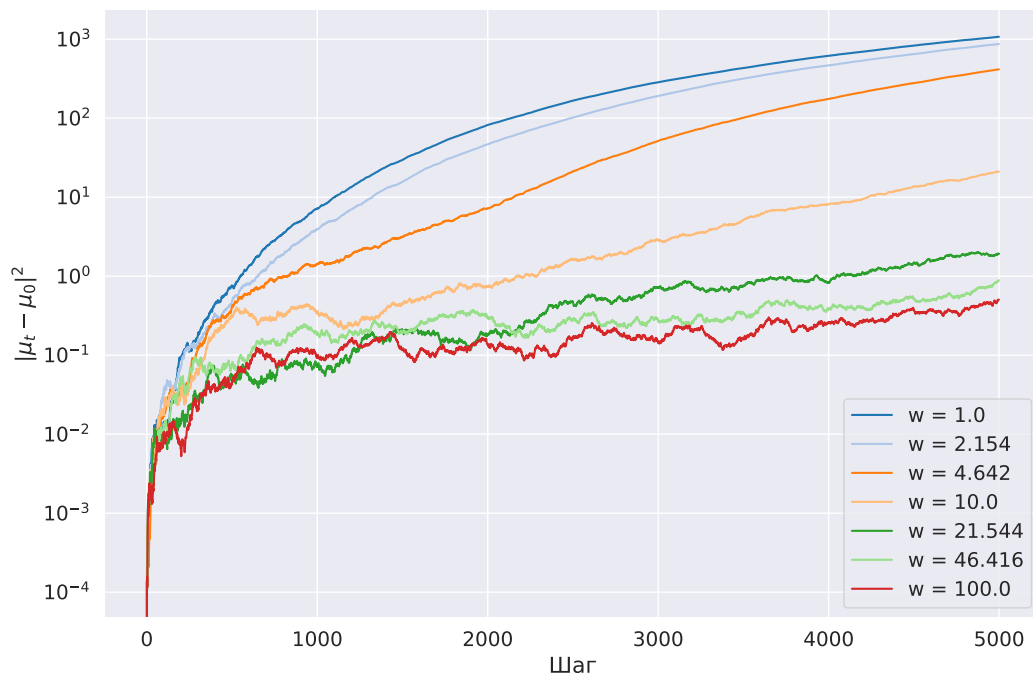
TSBandit.update(c_t)

Model.interest_update(c_t)

save_iter(t, c_t, μ_t)

5 Результаты

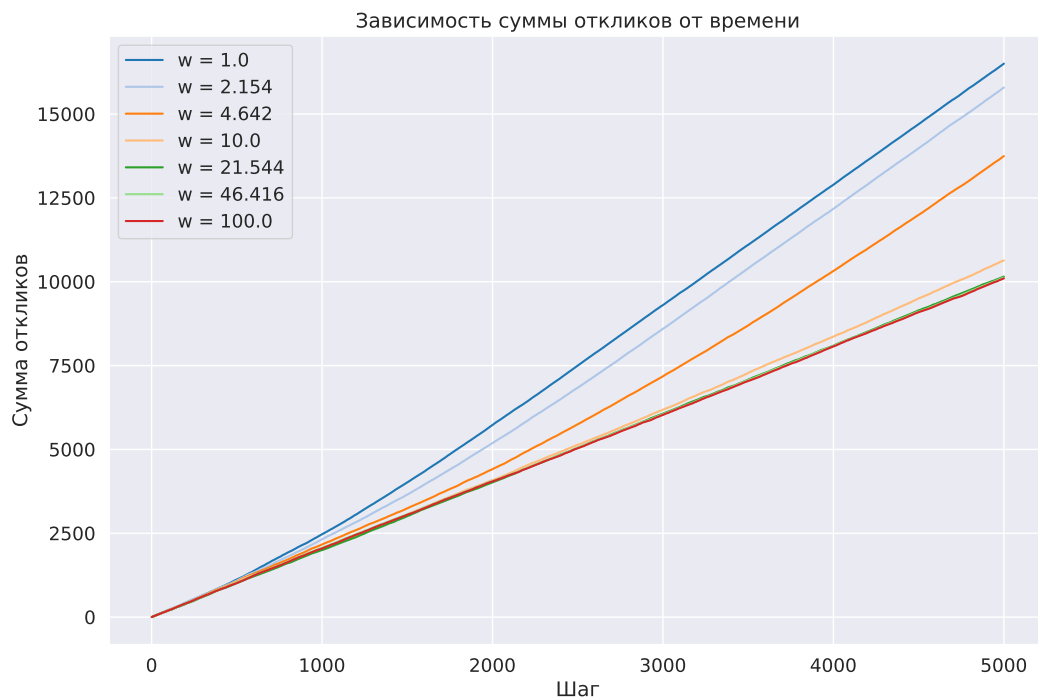
Наблюдается эффект неограниченного роста интереса. Наличие петель скрытой обратной связи иллюстрируется следующим графиком:



111

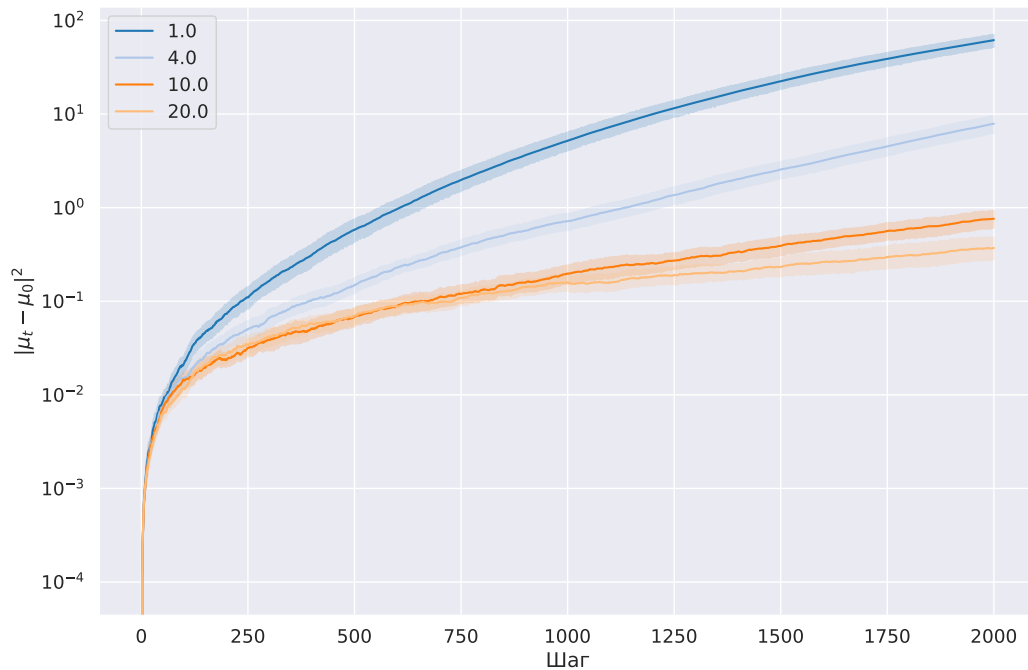
112 Это согласуется с определением петель.

113 Из графика суммы наград видно, что с определённого момента кривые начинают ид-
 114 ти параллельно максимально возможной сумме. Это свидетельствует о наличии петель
 115 обратной связи.

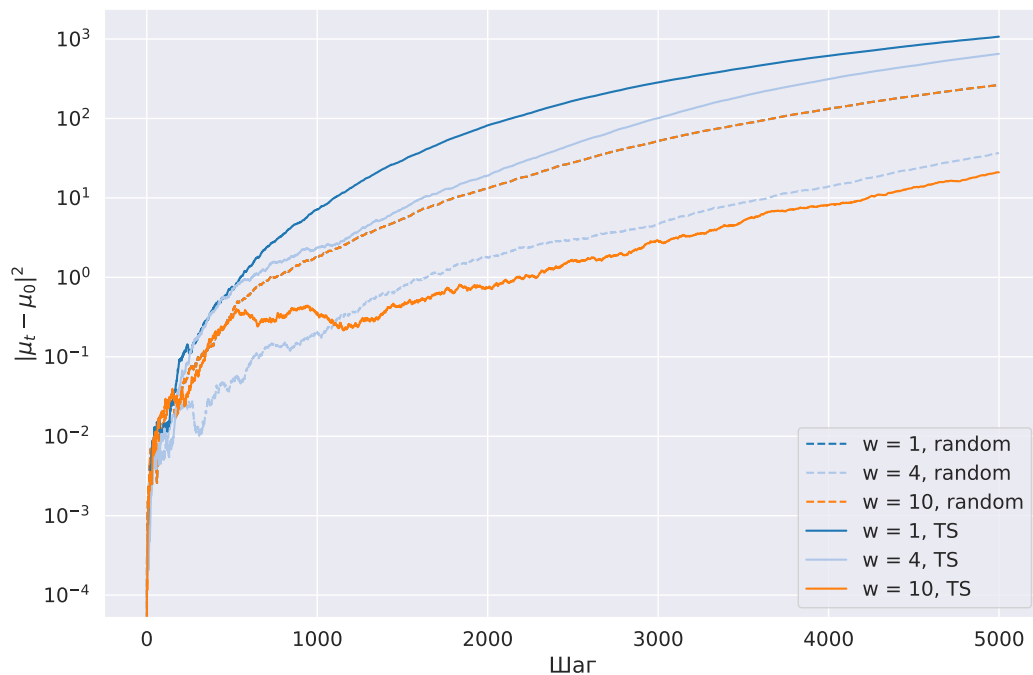


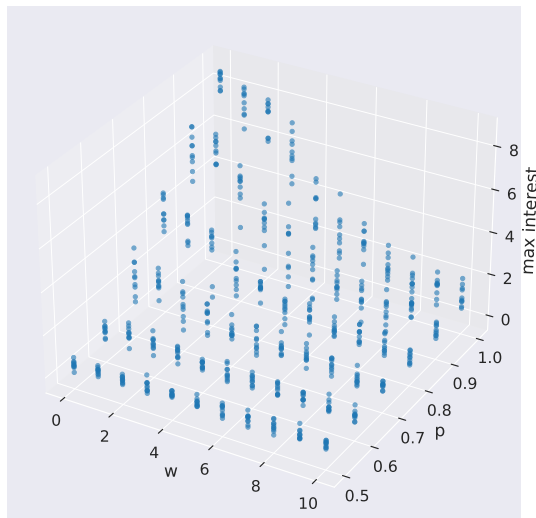
116

117 Разброс значений нормы интересов.

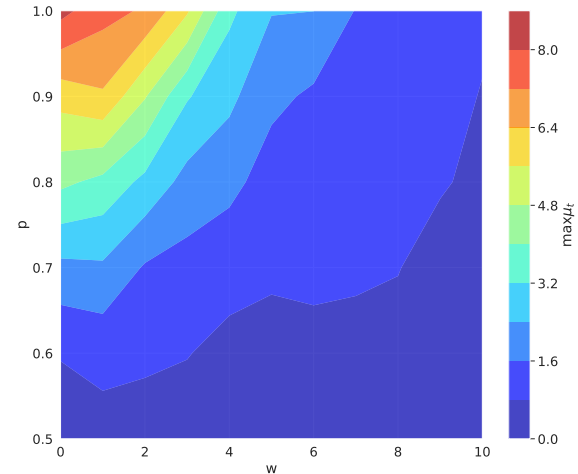


118
 119 Для случайной модели тоже наблюдается образование петли. Она более хаотична, но
 120 тренд неограниченного роста интереса всё равно присутствует.





(a)



(б)

Весь эксперимент и исходники расположены в гите:

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

6 Заключение

Была 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 сигмойды 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