Министерство науки и высшего образования Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

ТЕОРЕТИЧЕСКОЕ ИССЛЕДОВАНИЕ ЭВОЛЮЦИОННОГО АЛГОРИТМА (1+1) НА ДИНАМИЧЕСКИХ ПСЕВДОБУЛЕВЫХ ЗАДАЧАХ ОПТИМИЗАЦИИ

Автор: Торопин Константин Игоревич _______________

Направление подготовки: 01.03.02 Прикладная математика и информатика

Квалификация: Бакалавр

Руководитель ВКР: Буздалов М.В., к.т.н. _______________

Санкт-Петербург, 2020 г.

Обучающийся Торопин Константин Игоревич
Группа M3437 Факультет ИТиП

Направленность (профиль), специализация
Математические модели и алгоритмы в разработке программного обеспечения

ВКР принята «____» ____________ 20___ г.

Оригинальность ВКР ____%

ВКР выполнена с оценкой ____

Дата защиты «____» ____________ 20___ г.

Секретарь ГЭК Павлова О.Н. _______________

Листов хранения ____

Демонстрационных материалов/Чертежей хранения ______________________

Министерство науки и высшего образования Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО

УТВЕРЖДАЮ

Руководитель ОП

проф., д.т.н. Парфенов В.Г. ______________

«____» ___________ 20___ г.

ЗАДАНИЕ
НА ВЫПУСКНУЮ КВАЛИФИКАЦИОННУЮ РАБОТУ

Обучающийся Торопин Константин Игоревич

Группа M3437 **Факультет** ИТиП

Квалификация: Бакалавр

Направление подготовки (специальность): 01.03.02 Прикладная математика и информатика

Направленность (профиль): Математические модели и алгоритмы в разработке программного обеспечения

Тема ВКР: Теоретическое исследование эволюционного алгоритма (1+1) на динамических псевдобулевых задачах оптимизации

Руководитель Буздалов М.В., к.т.н.,

2 Срок сдачи студентом законченной работы до: «____» ___________ 20___ г.

3 Техническое задание и исходные данные к работе

4 Содержание выпускной квалификационной работы (перечень подлежащих разработке вопросов)

а) Постановка задачи

б) Описание

5 Перечень графического материала (с указанием обязательного материала)

Графические материалы и чертежи работой не предусмотрены

6 Исходные материалы и пособия

7 Дата выдачи задания «____» ___________ 20___ г.

Руководитель ВКР ______________

Задание принял к исполнению ______________«____» ___________ 20___ г.

Министерство науки и высшего образования Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО

АННОТАЦИЯ
ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЫ

Обучающийся: Торопин Константин Игоревич

Наименование темы ВКР: Теоретическое исследование эволюционного алгоритма (1+1) на динамических псевдобулевых задачах оптимизации

Наименование организации, где выполнена ВКР: Университет ИТМО

ХАРАКТЕРИСТИКА ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЫ

1 Цель исследования: Разработка теоретической базы для анализа эволюционных алгоритмов с динамически меняемыми функциями

2 Задачи, решаемые в ВКР:

а) анализ возвращения алгоритма к оптимуму;
б) нахождение плато в зависимости от динамики;
в) проработка теории для анализа времени работы алгоритмов
г) проверка теории на практике

3 Число источников, использованных при составлении обзора: 0

4 Полное число источников, использованных в работе: 0

5 В том числе источников по годам:

Отечественных			**Иностранных**		
Последние 5 лет	От 5 до 10 лет	Более 10 лет	Последние 5 лет	От 5 до 10 лет	Более 10 лет
0	0	0	0	0	0

6 Использование информационных ресурсов Internet: нет

7 Использование современных пакетов компьютерных программ и технологий:

Пакеты компьютерных программ и технологий	**Раздел работы**
Интегрированная среда разработки PyCharm	**??**

8 Краткая характеристика полученных результатов

Расширена теоретическая база эволюционных алгоритмов на динамических псевдобулевых задачах оптимизации .

9 Гранты, полученные при выполнении работы

Грантов и других форм государственной поддержки и субсидирования в процессе в процессе выполнения не предусматривалось.

10 Наличие публикаций и выступлений на конференциях по теме выпускной работы

Публикации и выступления, связанные с данной работой, отсутствуют.

Обучающийся	Торопин К. И.	______________

Руководитель	Буздалов М.В.	______________

«_____» ____________ 20___ г.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

В практике применения эволюционных алгоритмов нередко встречаются динамические задачи оптимизации, в рамках которых функция выживания изменяется со временем. В теоретическом анализе времени работы эволюционных алгоритмов не так много работ уделяется этим задачам. И в этих работах центральным ставится вопрос может ли быть достигнут оптимум между изменениями функции приспособленности, а то, как в промежутке меняется значение приспособленности, не особо рассматривается, за вычетом приспособленностей, крайне близких к оптимуму. В частности не решается вопрос что должно происходить с различными способами адаптации параметров, которые применяются для повышения эффективности алгоритмов. Поэтому задача анализа алгоритма (1 + 1) на динамических псевдобулевых задачах оптимизации является актуальной.

ГЛАВА 1. ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Эволюционные алгоритмы могут применяться для поиска экстремума функции. Однако в теории эволюционных алгоритмов существует проблема, которая заключается в том, что функция в процессе поиска экстремума может изменятся. Примером такой задачи может быть поиск оптимального маршрута в городе - искомый маршрут в зависимости от времени может отличаться. Поэтому возникает потребность проанализировать алгоритмы учитывая динамическое преобразование функции. Для начала был взят алгоритм (1 + 1) на задаче $OneMax$.

Данный алгоритм был выбран потому что это наиболее понятная и простая задача для анализа и часто является отправной точкой для исследования.

OneMax - задача поиска минимума функции количества несовпадающих бит с некоторым вектором f_{max}, длина которого n.

Алгоритм (1 + 1) решает её следующим образом - берётся случайный вектор длины n и далее итеративно мутируется - каждый бит независимо меняется с некоторой вероятностью $\frac{r}{n}$ (r рассматривается в данном исследовании $o(1)$ по отношению к n). Если количество совпадающих бит увеличивается, берётся мутируемый. Добавим следующую динамику для задачи - пусть каждые T итераций алгоритма вектор f_{max} меняет один случайный бит.

В рамках исследования были выявлены две ситуации - когда T слишком мала, то алгоритм "застревает"на некотором значении, расстояние до f_{max} колеблются на уровне некоторого $\dot{n}$. Значение c будем называть точкой стабилизации или плато. То есть в этой ситуации с некоторого момента примерное отношение бит которые не совпадают с вектором f_{max} ко всем битам будет равна c. Вторая ситуация - T оказывается достаточно большой, что алгоритм доходит до нужного оптимума, однако медленнее чем если бы динамики не было. Помимо этого был рассмотрена ситуация когда алгоритм не заканчивается на достижении оптимума (что вполне можно увидеть на реальных задачах). Тогда будет разумным вопрос - какая вероятность, что при изменении одного бита алгоритм успеет найти новый оптимум до следующего изменения.

По первому случаю был установлен факт, что значение c не зависит от параметра T (если $n > 1/c$), это значит что если мы фиксируем какое-то значение T, то насколько сильно мы бы не увеличивали n, значение c будет оставаться тем же. Были найдены некоторые оценки на значение c, а так же теоретически обоснован факт почему алгоритм фиксируется и почти не уходит от этого значения.

Что касается второго случая - асимптотика EA-алгоритмов обычно оценивается с помощью некоторых теорем которые несут общее название drift-analysis. Были взяты теоремы Additive Drift и Variable Drift и рассмотрены с точки зрения того, что математическое ожидание изменения функции на некоторых шагах отличается от обычного. Если рассматривать данную конкретную задачу, то в ней каждые K шагов математическое ожидание будет отличаться на некоторое значение $\leqslant 1$. Эти теоремы в дальнейшим могут помочь исследовать скорость нахождения первого оптимума и в более сложных алгоритмах в которых будет введена динамика. С помощью видоизмененных теорем была доказано, что если T больше некоторого значения, то асимптотика алгоритма не изменится, то есть останется $O(n\dot{l}og(n))$

По последнему вопросу были рассмотрены ситуации в которых вероятность нахождения нового оптимума быстрее чем придёт новое изменение равна $o(1)$ и $1 - o(1)$. Так же и найдены границы для T чтобы вероятность была больше или меньше заданного значения.

Экспериментальные данные показывают, что плато находится на тех значениях, которые предсказывает теория.

ГЛАВА 2. ТЕОРЕТИЧЕСКОЕ ИССЛЕДОВАНИЕ

2.1. Постановка задачи

2.2. Поиск плато

Введём обозначения - для конкретной I

$$Bin(m,n) \equiv C_m^n \cdot p^m \cdot q^{n-m} = C_m^n \cdot p^m \cdot q^{n-m}$$
$$A(i) \equiv Bin(i, I)$$
$$Q(i) \equiv Bin(i, n-I)$$
$$= \frac{I}{n}$$

Так как вероятность изменения значения вектора ровно на i вероятности того что среди ещё не найденных бит изменится ровно k бит, а в найденных $k-i$ для всех возможных k.

$$P(\Delta Y = i | Y_t = I) = \sum_{j=1}^{\min(I-i, n-I-i)} A(i+j)Q(i)$$

Давайте рассматривать только такие I что

$$c > 0.5$$
$$c \cdot r < 1$$

Вероятностный смысл будет высказан позже Заметим что

$$P(\Delta Y = i | Y_t = I) = \sum_{j=1}^{I-i} A(i+j)Q(i)$$

Тогда значение $E[_\Delta Y, X_t \;=\; I]$можно расписать в виде: E$[_\Delta Y, X_t \;=\; I]$= $\sum_{i=1}^{I} i \cdot \sum_{j=1}^{I-i} A(i+j)Q(i) = \sum_{i=1}^{I} Q(i) \sum_{j=i}^{I} j \cdot A(j-i)$

$$\begin{aligned} 5Q(0)A(1)+ \quad & 2*Q(0)A(2) && +3*Q(0)A(3) && +\ldots && +I*Q(0)A(I)+ \\ & Q(1)A(2) && +2*Q(1)A(3) && +\ldots && +(I-1)*Q(1)A(I)+ \end{aligned}$$

$$+\dots \qquad +A(I)Q(n-I)+$$
$$+A(I)Q(n-I)+$$

Так как $\sum_{j=1}^{I} j \cdot A(j)$ это определение матожидание случайной величины A, что равно pI Заметим что первая строчка равна $\tau^* \equiv Q(0) \cdot \sum_{j=1}^{I} j \cdot A(j) = Q(0) \cdot pI = q^{n-I} \cdot c \cdot r$ Это будет нашей нижней оценкой на $E[_\Delta Y, X_t = I]$Далее оценим сумму остальных строчек. Давайте оценим сверху для каждой строчки во сколько изменяется элемент в строчке относительно предыдущего: Для строчки i: $i \geqslant 2i \geqslant 1$

$$\frac{(j+1) * Q(i) * A(i+j+1)}{j * Q(i) \cdot A(i+j)}$$
$$= \frac{j+1}{j} \cdot \frac{(I-(i+j+1))!}{I! \cdot (i+j+1)!} \cdot p^{I-(i+j+1)} q^{i+j+1} \frac{I! \cdot (i+j)!}{(I-(i+j))!} \cdot \frac{1}{p^{I-(i+j)} q^{i+j}}$$
$$= \frac{j+1}{j} \cdot \frac{I-i-j}{i+j+1} \cdot (\frac{q}{p} = \frac{r \cdot n}{n \cdot (n-r)} = \frac{r}{n-r})$$
$$= \frac{j+1}{j \cdot (i+j+1)} \cdot \frac{c*(n-r) + c*r - i - j - 1}{n-r} \cdot r \leqslant$$
$$\leqslant \frac{1}{2} \cdot \frac{c*(n-r)}{n-r} \cdot r = \frac{c \cdot r}{2}$$

Значит каждая строчка i:

$$a_i \equiv \sum_{j=1}^{i-i} (j \cdot Q(i) \cdot A(i+j)) =$$
$$= \sum_{j=1}^{inf} (Q(i) \cdot A(i+1) \cdot (\frac{c \cdot r}{2})^0)$$
$$= Q(i) \cdot A(i+1) \cdot \frac{1}{1 - \frac{c \cdot r}{2}}$$

Аналогично проделаем с данным рядом:

$$\frac{Q(i+1) \cdot A(i+2)}{Q(i) \cdot A(i+1)} = \frac{(n-I-i) \cdot (I-i-1)}{(i+1) \cdot (i)} \cdot (\frac{r}{n-r})^2$$
$$= \frac{(n-r)*(1-c) + r - r*c - i}{n-r} \cdot \frac{c*n - r*c - r*c - i - 1}{n-r} \cdot \frac{1}{(i+1) \cdot (i)}$$
$$\leqslant (1-c) \cdot c \cdot 1/2 + (\frac{(r - r*c - i) \cdot c \cdot 1/2}{n-r} = \theta(1))$$

С другой стороны, если брать $n >= r/c$

$$c \cdot 1/2 \cdot \frac{n - nc - I}{n - r} \leqslant$$
$$c \cdot 1/2 \cdot \frac{n - r - I}{n - r} < c \cdot 1/2$$

Давайте рассмотрим оценку $c/2$, тогда получается, что вся остаточная часть меньше или равна по аналогии:

$$Q(1)A(2) \cdot \frac{1}{1 - c/2} \cdot \frac{1}{1 - c \cdot r/2} = \frac{(n - I) * I * (I - 1)}{2} \cdot p^3 \cdot q^{n-3} \cdot constants \leqslant$$
$$\leqslant c^2 \cdot r^3 \cdot q^{n-3} = (q^{n-I} \cdot c \cdot r) \cdot (q^{I+3} \cdot c \cdot r^2) = \tau^* \cdot (q^{I+3} \cdot c \cdot r^2)$$

Значение $constants < 1$ его можно использовать, чтобы брать более качественные оценки на c_{lower} о котором будет говорится позже.

Оценки τ^*:

Введем значение $\tau \equiv \exp(-r + r * c) \cdot c \cdot r$

$$\tau^* = q^{n-I} \cdot c \cdot r = (1 - r/n)^n \cdot (1 - c) \cdot c \cdot r < \exp(-r + r * c) \cdot c \cdot r = \tau$$
$$\tau^* = (1 - r/n)^{n-r} * (1 - c) \cdot c \cdot r \cdot (1 - r/n)^r * (1 - c)$$
$$> \exp(-r + r * c) \cdot c \cdot r \cdot (1 - r^2/2n)^{1-c} >$$
$$\exp(-r + r * c) \cdot c \cdot r \cdot (1 - r^2/2n) > \exp(-r + r * c) \cdot c \cdot r \cdot (1 - o(1)) = \tau \cdot (1 - o(1))$$

Заметим так же:

$$q^{I+3} = (1 - r/n)^{c+3} < \text{ (6)} < \exp(-r \cdot c) \cdot (1 - r/n)^3 < \exp(-r * c)$$

Обозначим $k \equiv \exp(-r * c) \cdot c \cdot r^2$

Тогда сумма всей таблицы оценивается сверху как:

$\tau^* \cdot (1 + q^{I+3} \cdot c \cdot r^2) < \exp(-r + r * c) \cdot c \cdot r \cdot (1 + \exp(-r * c) \cdot c \cdot r^2) < \tau \cdot (1 + k)$

Из чего сделаем вывод что E$[_\Delta Y, X_t = I] \in (\tau \cdot (1 - o(1)), \tau \cdot (1 + k))$

Рассмотрим $E[_\Delta Z, X_t = I]$.
Для начала давайте рассматривать задачу в которой каждый шаг с вероятностью $1/T$ меняется один бит в векторе $f_{\max}$
Связь с истинной задачей будет рассматриваться позже. Так как вероятность что случайно выбранный бит будет из тех, которые уже достигли оптимума равняется $1 - I/n = (1 - c)$ Тогда

E$[_\Delta Z, X_t = I]$= (1 - c)/K

Заметим что E$[_\Delta Z, X_t = I]$c, E$[_\Delta Y, X_t = I]$убывает.
Второй факт доказывается тем что при $p > 1/2$ вероятность изменить бит $> \frac{1}{2}$, а количество необходимых бит убывает. (Можно доказать строже) Тогда рассмотрим для фиксированного K такие два значения c_{upper} и c_{lower} для которых мы можем быть уверены что:

$$E[_\Delta(X_t)|X_t = n \cdot c_{upper}] > 0 E[_\Delta(X_t)|X_t = n \cdot c_{lower}] < 0$$

Грубо говоря вне этого промежутка алгоритм будет "тянуть"его назад.

$$E[_\Delta(X_t)|X_t = n \cdot c] = [_\delta(Y_t)|X_t = n \cdot c] - \frac{(1-c}{K}$$
$$\text{E}[_\Delta Y, X_t = I] \in (\tau - o(1), \tau \cdot (1 + k))$$
$$0 \leqslant [_\Delta(X_t)|X_t = n \cdot c] < \tau - o(1) - \frac{(1-c}{K}$$
$$K \cdot (\tau - o(1)) > (1 - c)$$
$$K > \frac{(1-c)}{\tau \cdot (1 - o(1))}$$

Рассмотрим $g_{lower}(c) = \frac{(1-c)}{\tau} = \frac{(1-c)}{\exp(-r+r*c)\cdot c\cdot r}$

Тогда $c_{lower} = g_{lower}^{-1}(K) + \epsilon$ где $\epsilon \to 0$ для больших n (TODO формальнее)

Аналогично $g_{upper}(c) = \frac{(1-c)}{\tau\cdot(1+k)} = \frac{(1-c)}{\exp(-r+r*c)\cdot c\cdot r\cdot(1+\exp(-r*c)\cdot c\cdot r^2)}$
$c_{upper} = g_{upper}^{-1}(K)$

Заметим что величины c_{upper} и c_{lower} не зависят от n, а значит мы можем сделать вывод что плато будет начинаться в одинаковых значениях вне зависимости от величины n.

Далее оценим величину:

$$P({}_{\Delta}Y \geqslant j | Y_t = I)$$

Установим следующий факт:

Теорема 1.

$$r, \eta > 0, I > n/2 :$$
$$P({}_{\Delta}Y \geqslant j | Y_t = I) \leqslant \frac{r}{(1+\eta)^j}$$

Доказательство. TODO переписать доказательство

Тогда используя Negative-drift теорему мы можем доказать что уйти на любое значение от плато займёт экспоненциальное время.

Теорема 2. Пусть $X_{t_0} = c : c \in (c_{lower}, c_{upper})$ Тогда найдётся такое большое n, что:

$$1.\forall k : k > 1/2, k \cdot n < c_{lower} \cdot n - 1,$$
$$T_a = \min t > t_0 | X_t \leqslant k \cdot n \exists l :$$
$$P[T_a \leqslant e^{l*n}] < e^{-l*n}$$
$$2.\forall k : k \leqslant 1, k \cdot n > c_{upper} \cdot n + 1,$$
$$T_a = \min t > t_0 | X_t \geqslant k \cdot n \exists l :$$
$$P[T_a \leqslant e^{l*n}] < e^{-l*n}$$

Доказательство. Из факта выше следует, что для всех $I > 1/2$
$P({}_\Delta Y \geqslant j | Y_t = I) \leqslant \frac{r}{(1+\eta)}$
Так же заметим что Z_t не может превышать по модулю 1

$$P({}_\Delta Z \geqslant 1) \leqslant \frac{1}{K} P({}_\Delta Z \geqslant j > 1) = 0$$

Таким образом:
$P(|{}_\Delta X| \geqslant j | X_t = I) \leqslant \frac{\max(r, 1/K)}{(1+\eta)}$

1. Рассмотрим второй случай, так как $E[{}_\Delta X_t | X_t = I]$убывает и $E[{}_\Delta X | X_t = c_{upper} \cdot n] \equiv \delta < 0$
И по монотонности для всех

$$I \in (c_{upper} \cdot n, k \cdot n) E[{}_\Delta X | X_t = I] < \delta < 0$$

Таким образом все условия Negative-Drift теоремы выполнены.

2. Аналогично, только рассмотрим вместо X_t случайную величину $W_t = -X_t$, тогда

$$E[{}_\Delta W | W_t = c_{lower} \cdot n] \equiv \delta < 0$$
$$I \in (k \cdot n, c_{lower} \cdot n)$$
$$E[{}_\Delta W | W_t = I] \equiv \delta < 0$$

TODO: показать какие конкретно l и n_0 и что вообще с этим можно делать - построение честного вероятностного удалось сделать с помощью марковский цепей.
Однако только снизу и если брать более слабую теорию где не $EY > EZ$, а $PY > PZ$, что мало чего даёт.

Наверное всё же нужно просто доказать что мы постоянно будем возвращаться в это плато, но хотелось бы чего-то лучше.

2.3. Оценки скорости нахождения первого оптимума

Рассмотрим классические Drift-теоремы с точки зрения задачи.

Теорема 3. Пусть $(X_t)_{t \in N}$, последовательность случайных величин с $S \in [0, \infty)$ состояниями
$X \in$Z
s$_{\min} = \inf(S \backslash \{0\})$
Пусть $T := \inf t > 0 | X_t = 0$.
Для любого $t \geqslant 0$:
$\delta_t(s) = E[X_t - X_{t+1} | t\%K \neq K - 1, X_t = s]$
$\Delta_t(s) = E[X_t - X_{t+1} | t\%K = K - 1, X_t = s]$
Если существуют такие константы ϵ и δ, что:

$$\epsilon \cdot K > \delta$$

$$\delta_t(s) > \epsilon$$

$$\Delta_t(s) > \epsilon - \delta$$

Тогда $E[T] \leqslant \frac{E[X_0]}{\delta - \epsilon / K}$

Доказательство. Теорему можно свести к классической теореме об Аддитивном дрифте. Рассмотрим случайную величину $Y_i = X_{i \cdot K}$

$$E[Y_0] = E[X_0]$$

$$Y_{i+1} = X_i \cdot k + \sum_{t=i \cdot K}^{i \cdot (K+1)-1} (X_{t+1} - X_t) + \delta$$

$$E[Y_i - Y_{i+1}] \geqslant K \cdot \epsilon - \delta > 0$$

Последнее неравенство из условия.
Тогда если I - первое такое i, что $Y_i = 0$, $E[I] \leqslant \frac{E[X_0]}{K \cdot \epsilon - \delta}$.
А значит $E[T] \leqslant E[I] \cdot K \leqslant \frac{E[X_0]}{\epsilon - \frac{\delta}{K}}$

Теперь рассмотрим вариативный дрифт:

Теорема 4. Пусть $(X_t)_{t\in N}$, последовательность с $S \in [0,\infty)$ состояниями, число K $s_{\min} = \inf(S\backslash\{0\})$ Пусть $T := \inf t > 0|X_t = 0$. И для любого $t \geqslant 0$:
$\delta_t(s) = E[X_t - X_{t+1}|t\%K \neq K-1, X_t = s]$
$\Delta_t(s) = E[X_t - X_{t+1}|t\%K = K-1, X_t = s]$
Если существуют такая константа ϵ и монотонно возрастающая функция $h : R^+ \to R^+$ что:

$$\delta_t(s) > h(s)$$
$$\Delta_t(s) > h(s) - \epsilon$$
$$K > \frac{\epsilon}{h(s_{\min})}$$

Тогда:

$$E[T] \leqslant \frac{1}{1-\frac{\epsilon}{K\cdot h(s_{\min})}} \cdot (\frac{s_{\min}}{h(s_{\min})} + E[\int_{s_{\min}}^{X_0} \frac{1}{h(\sigma)} d\sigma])$$

Доказательство. Воспользуемся доказательством классического вариативного дрифта.

Если рассматривать функцию g(s) := $\begin{cases} \frac{s_{\min}}{h(s_{\min})} + \int_{s_{\min}}^{s} \frac{1}{h(\sigma)} d\sigma, & \text{if } s \geqslant s_{\min} \\ \frac{s_{\min}}{h(s_{\min})}, & \text{if } 0 \leqslant s \leqslant s_{\min} \end{cases}$ То по некоторым свойствам (здесь можно подробно расписать, но это уже есть в самой теореме) получается что если брать $Y_t = g(X_t)$:
Для $t \mod K \neq K-1$:

$$\begin{aligned} E[Y_t - Y_{t+1}|Y_t = g(s)] &= E[g(X_t) - g(X_{t+1})|g(X_t) = g(s)] \\ &\geqslant E[\frac{X_t - X_{t+1}}{h(X_t)}] = \frac{\delta_t(s)}{h(s)} \geqslant 1 \end{aligned}$$

Для $t \mod K = K-1$:

$$E[Y_t - Y_{t+1}|Y_t = g(s)] \geqslant \frac{\Delta_t(s)}{h(s)} \geqslant 1 - \frac{\epsilon}{h(s_{\min})}$$

Тогда можем применяя (3) выше со следующими условиями:

$$\epsilon_{new} \equiv 1$$
$$\delta_{new} \equiv \frac{\epsilon}{h(s_{\min})}$$
$$K_{new} \equiv K$$
$$K > \frac{\epsilon}{h(s_{\min})} > \delta_{new}$$
$$\epsilon_{new} \cdot K_{new} > \delta_{new}$$

Значит:

$$E[T] \leqslant \frac{1}{1 - \frac{\epsilon}{K \cdot h(s_{\min})}} \cdot (\frac{s_{\min}}{h(s_{\min})} + E[\int_{s_{\min}}^{X_0} \frac{1}{h(\sigma)} d\sigma])$$

Применим данную теорему к рассматриваемой задаче:

Теорема 5. Если Алгоритм $OneMax(K)$ имеет $K = O(n)$, $K > \frac{n \cdot \exp(r)}{r} + O(n)$, то $E[T] = O(n \cdot \log(n))$

Доказательство.

$$\delta_t(s) \geqslant A(1)Q(0) = i \cdot p \cdot q^{n-1} = i \cdot r/n \cdot (1 - r/n)^{n-1}$$
$$\geqslant \frac{i}{n} \cdot r \cdot \exp(-r) - \phi(n)$$
$$\phi(n) = o(1/n)$$
$$\Delta_t(s) = \delta_t(s) - \frac{n - s}{n} \geqslant \delta_t(s) - 1$$
$$\epsilon = 1$$
$$s_{min} = 1$$
$$h(s) = r \cdot \frac{s}{n} \cdot \exp(-r) - \phi(n)$$
$$h(s_{min}) = \frac{r \cdot \exp(-r)}{n} - \phi(n)$$
$$K > \frac{1}{h(s_{\min})}$$

$$\exists c \in (0,1) : \frac{1}{1 - \frac{1}{K \cdot h(s_{\min})}} < c$$

Значит мы можем применить теорему (4):

$$E[T] \leqslant \frac{1}{1 - \frac{\epsilon}{K \cdot h(s_{\min})}} \cdot (\frac{s_{\min}}{h(s_{\min})} + E[\int_{s_{\min}}^{X_0} \frac{1}{h(\sigma)} d\sigma]) =$$
$$= c \cdot (O(n \cdot \log(n))) = O(n \cdot \log(n))$$

2.4. Добавочные леммы

Леммы про свойства экспоненты:

Лемма 6. $(1 - r/n)^n < e^{-r}$

Лемма 7. $(1 - r/n)^{n-r} > e^{-r}$

Лемма 8. $(1 - r/n)^r > 1 - 2r/n$

Доказательство. Пусть a - такое что $a = 2^k, a \leqslant r, 2 * a > r$, где $k \in \mathbb{N}$

$$(1 - r/n)^r \geqslant (1 - r/n)^a = (1 - 2r/n + r^2/n^2)^{a/2} > (1 - 2r/n)^{a/2}$$
$$> (1 - 4r/n)^{a/4} > 1 - a * r/n > 1 - 2 * r^2/n$$

Лемма 9. $(1 - r/n)^n > e^{-r} \cdot (1 - o(1/n))$

Доказательство.

$$(1 - r/n)^n > e^{-r} \cdot (1 - r/n)^r > e^{-r} \cdot (1 - 2 \cdot r^2/n) > e^{-r} - o(r^2/n) = e^{-r} - o(1)$$

ГЛАВА 3. ПРОВЕРКА ГИППОТЕЗ НА РЕАЛЬНЫХ ДАННЫХ

Для того чтобы посмотреть насколько быстро будет возвращаться алгоритм после первого достижения оптимума были проведён следующий эксперимент:

Алгоритм запускается несколько раз на одних и тех же n и T. И рассматривается количество раз, когда алгоритм успевает найти новый оптимум.

μ - матожидание первого срабатывания.

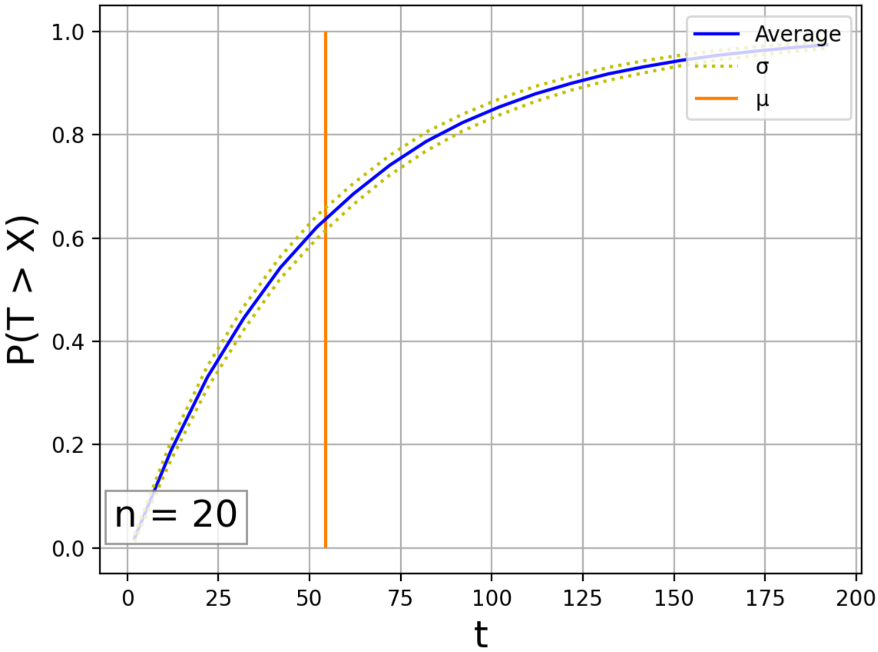

Рисунок 1 – Как часто алгоритм найти успевает новый оптимум при n = 20

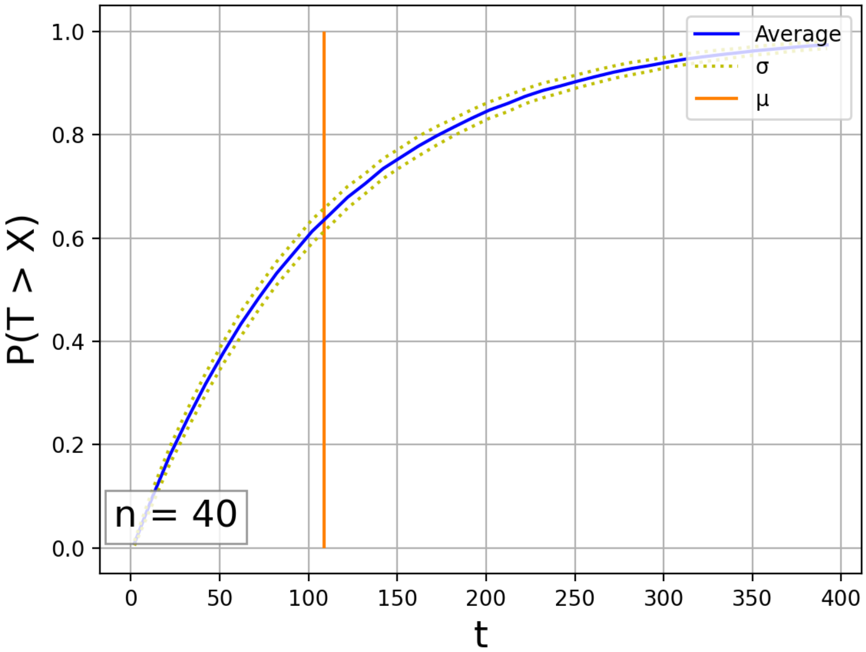

Рисунок 2 – Как часто алгоритм найти успевает новый оптимум при n = 40

Для оценки скорости алгоритма так же рассматриваются разные значения T

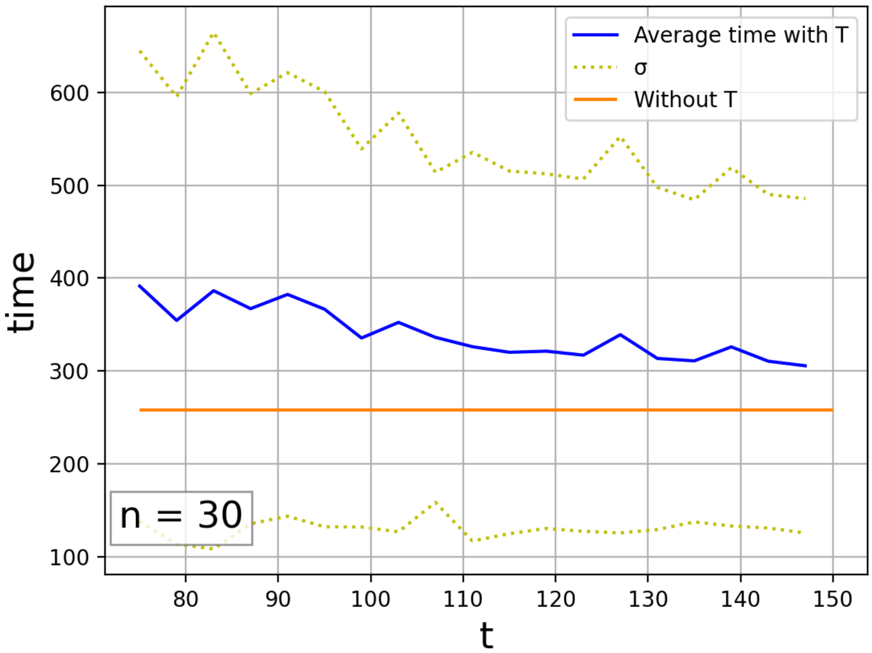

Рисунок 3 – Время работы алгоритма с n = 30

В качестве демонстрации основной гипотезы о независимости значения плато от n поставлены несколько экспериментов с разными n, но с одной и той же $T = 20$.

На графиках изображено усредненное поведение функции d - количество разных бит у f и $f_{\max}$ за 20 запусков.

Как видно из экспериментов они все останавливаются примерно на одном уровне по отношению к n.

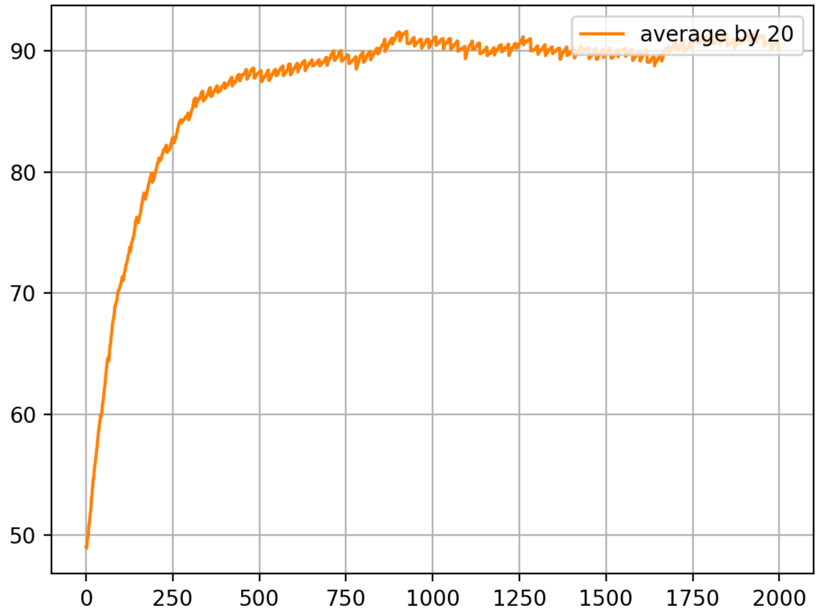

Рисунок 4 – Обнаружение плато с n = 100

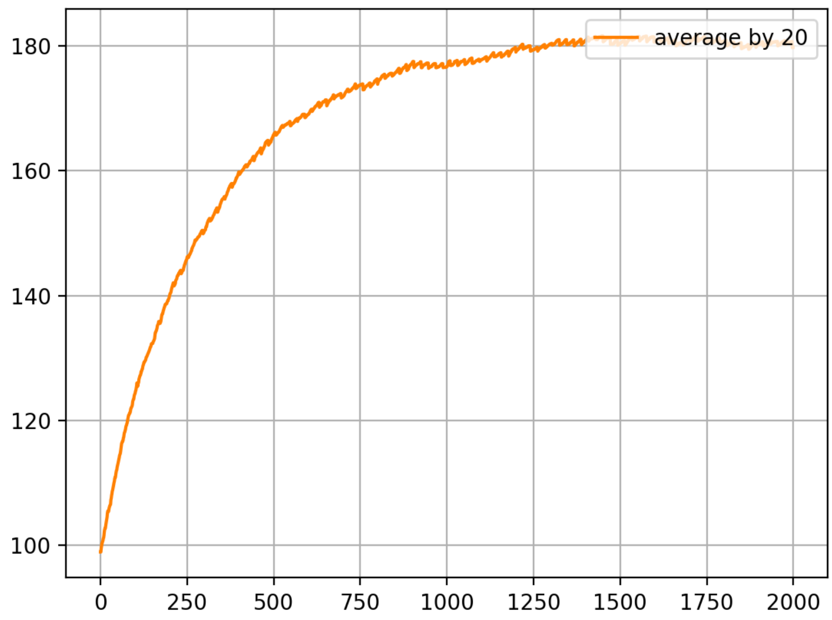

Рисунок 5 – Обнаружение плато с n = 200

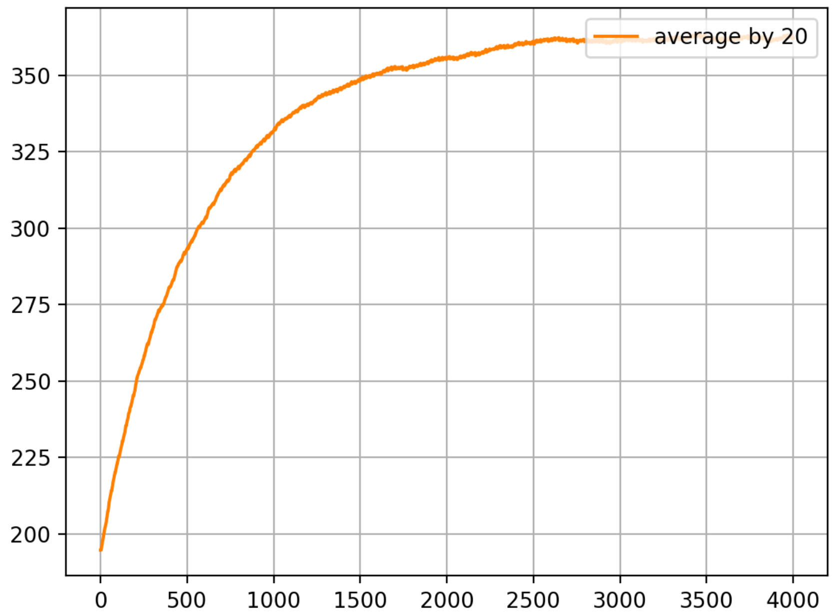

Рисунок 6 – Обнаружение плато с n = 400

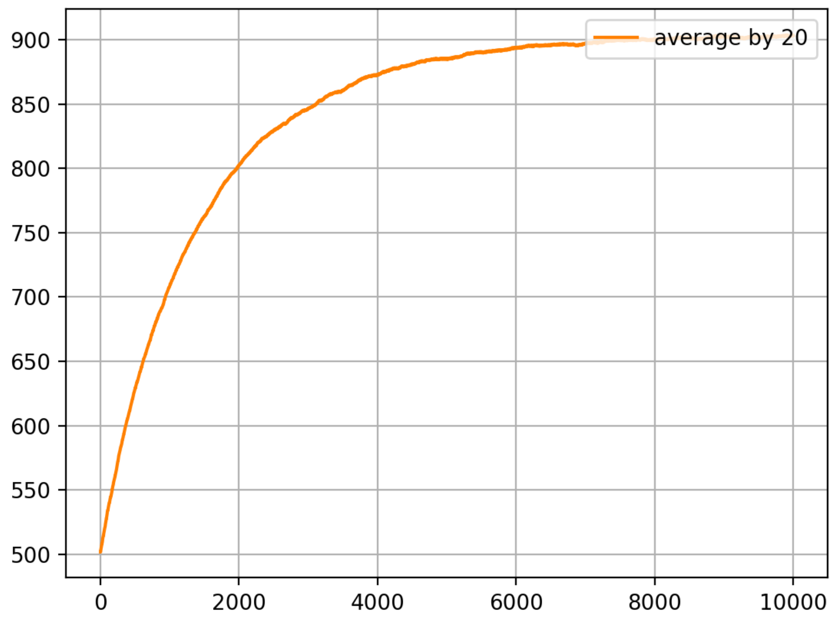

Рисунок 7 – Обнаружение плато с n = 1000

Для анализа функций g_{upper} и g_{lower} приведены следующие графики:

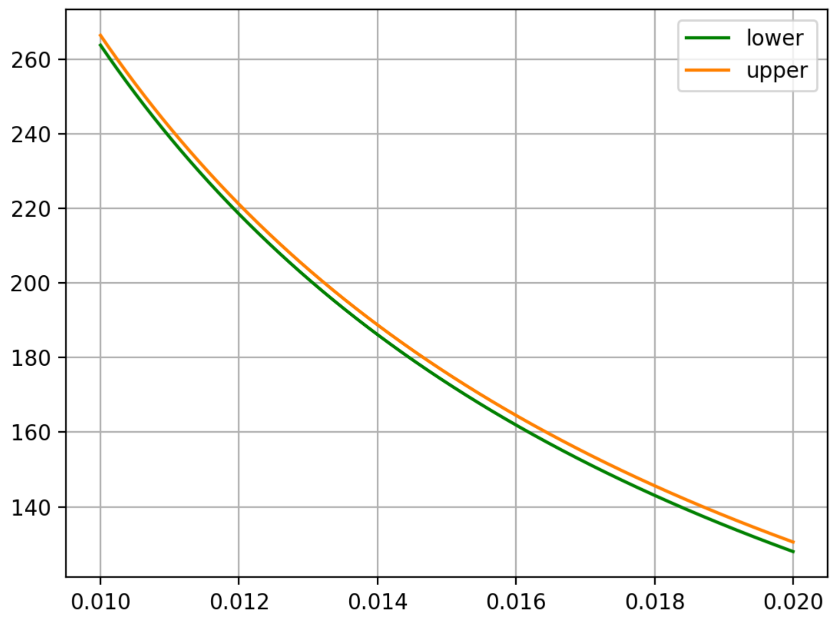

Рисунок 8 – Зависимость с и T

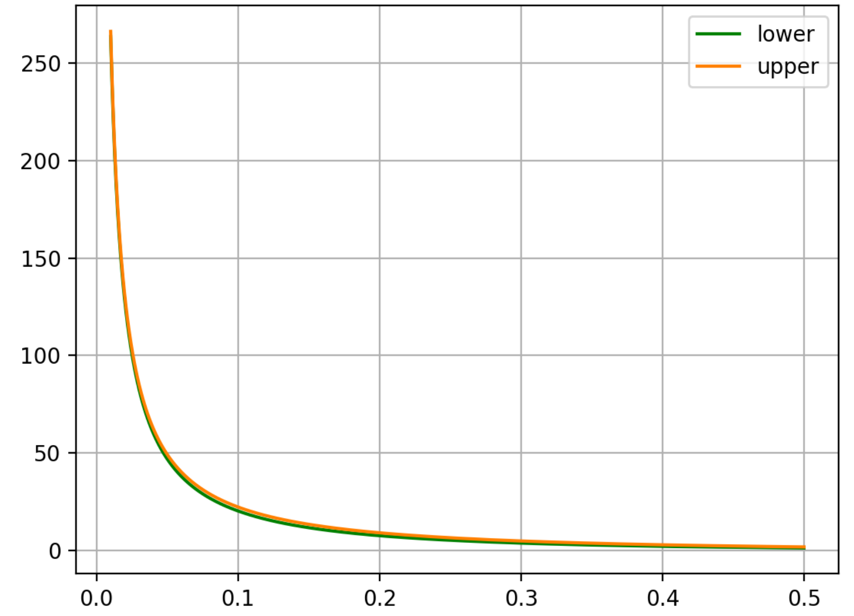

Рисунок 9 – Зависимость c и T

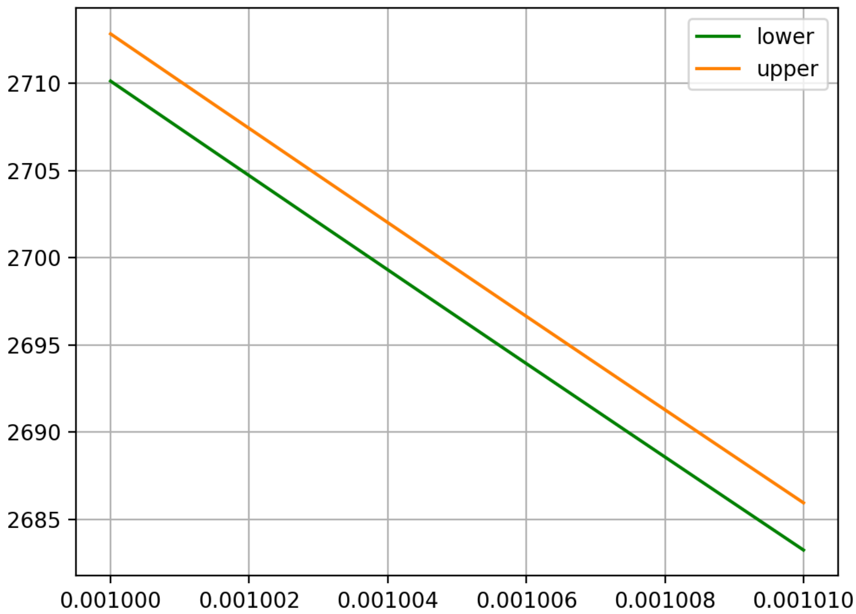

Рисунок 10 – Зависимость c и T

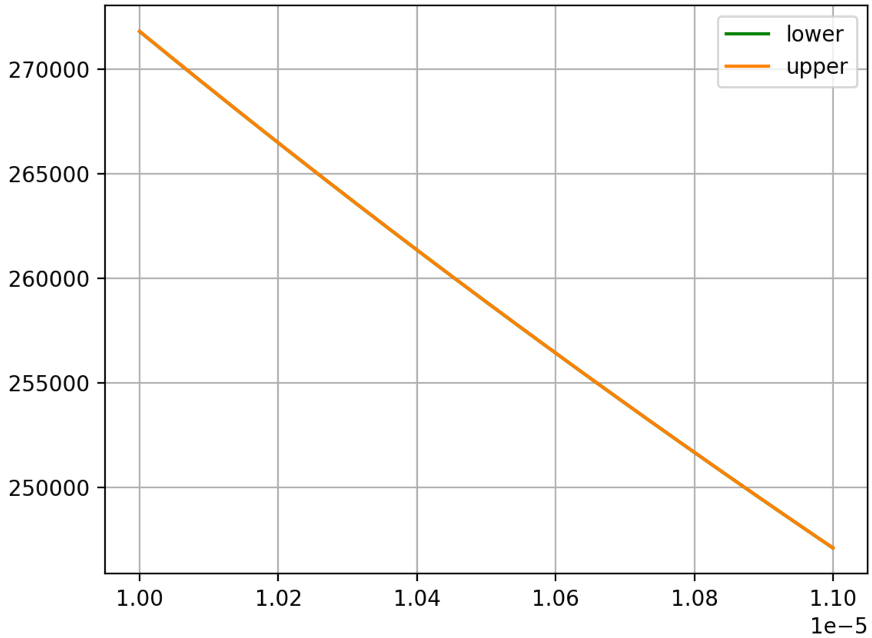

Рисунок 11 – Зависимость c и T

ЗАКЛЮЧЕНИЕ

В результате было обнаружено и доказано несколько теорем, которые помогут в дальнейшим анализировать эволюционные алгоритмы с динамическими изменениями. Предсказаны оценки какая точка стабилизация будут находиться для данного T. То есть в каком месте алгоритм будет находиться в независимости от n. Выведены пару drift-теорем, которые могут помочь дальнейшим исследованиям динамических оптимизаций, в частности теоремы применены к данной задаче в следствие чего была установлены такие значение T, при которых алгоритм будет находить оптимумы за то же асимптотическое время. Так же найдены такие T, в которых алгоритм будет с большой вероятностью в этом оптимуме останавливаться - вероятность его покинуть < какого-то значения или в каких вероятность будет равна o(1) Оценки были подтверждены экспериментальными данными.