Math-Net.Ru

Общероссийский математический портал

Д. В. Пильщиков, Об одном теоретико-вероятностном подходе к обоснованию надежности метода Хеллмана, *Матем. вопр. криптогр.*, 2019, том 10, выпуск 1, 83–114

DOI: <https://doi.org/10.4213/mvk278>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.224.60.19

26 апреля 2023 г., 05:46:06

**Об одном теоретико-вероятностном подходе
к обоснованию надежности метода Хеллмана**

Д. В. Пильщиков

Лаборатории ТВП, Москва

Получено 18.IV.2018

Аннотация. Оценка надежности метода Хеллмана сводится к оценке среднего значения случайного числа $\xi(m, t, N)$ различных элементов множества X в таблице из m цепочек по t итераций функции $F: X \rightarrow X$. Предложена вероятностная модель, в рамках которой получены оценки уклонения среднего значения величины $\xi(m, t, N)/(mt)$ от его приближения. Выявлены свойства функции F , существенно влияющие на надежность метода. Оценка среднего значения величины $\xi(m, t, N)$ проводится с помощью подходящего процесса Гальтона–Ватсона.

Ключевые слова: метод Хеллмана, вероятностные модели, ветвящиеся процессы

**On a probabilistic approach to the estimation of reliability
of the Hellman method**

D. V. Pil'shchikov

TVP Laboratories, Moscow

Abstract. The evaluation of the reliability of the Hellman method reduces to estimating the mean value of a random number $\xi(m, t, N)$ of different elements of the set X in a table containing m records of t iterations of function $F: X \rightarrow X$. We suggest a probabilistic model, within which estimates of the deviation of the mean value of $\xi(m, t, N)/(mt)$ from its approximation are obtained. The properties of the F function that significantly affect the reliability of the method are revealed. The estimation of the mean value of $\xi(m, t, N)$ is carried out using the appropriate Galton–Watson process.

Key words: Hellman method, probabilistic models, branching processes

Citation: *Matematicheskie Voprosy Kriptografii*, 2019, v. 10, № 1, pp. 83–114 (Russian)

© Академия криптографии Российской Федерации, 2019 г.

1. Введение

Ряд задач прикладной криптографии [1–8] связан с задачей восстановления хотя бы одного из элементов ключевого набора $\mathcal{K} = \{x_i \mid i = 1, \dots, D\}$ по набору данных

$$\mathcal{D} = \{a_i \mid a_i = G(x_i), \quad i = 1, \dots, D\},$$

где $G : X \rightarrow X$, $|X| = N$, есть некоторая труднообратимая функция.

Для ее решения широкое распространение получили переборные алгоритмы типа балансировки времени-памяти-данных [1–3]. Один из этих методов [1] был предложен в 1980 году Мартином Хеллманом (далее метод Хеллмана) применительно к вскрытию ключей блочного шифра.

На предварительном этапе метода Хеллмана проводятся цепочки вычислений, результаты которых записываются в таблицы. Решение задачи восстановления хотя бы одного элемента ключевого набора $\mathcal{K}$ происходит на оперативном этапе с использованием созданных таблиц.

В связи с практической реализацией метода Хеллмана теоретический интерес представляют оценки его временной сложности и надежности. Надежность метода Хеллмана характеризуется вероятностью восстановления хотя бы одного из элементов ключевого набора $\mathcal{K}$ при его случайном выборе.

Получению оценок надежности метода Хеллмана в рамках того или иного теоретико-вероятностного подхода посвящен ряд работ [1, 8–10]. В данных работах оценка надежности метода Хеллмана сводится к оценке среднего значения случайной величины $\xi(m, t, N)$, моделирующей число различных элементов множества X , задействованных при построении одной таблицы из m записей для цепочек вычислений длины t .

В [10] с помощью эвристических предположений найдена оценка $\delta(\alpha)$ среднего значения величины $\xi(m, t, N)/(mt)$, которая позволяет оценивать надежность метода Хеллмана с приемлемой для практических приложений точностью. Она имеет вид

$$\delta(\alpha) = \frac{\text{th}(\sqrt{\alpha})}{\sqrt{\alpha}}, \quad \text{где} \quad \alpha = \frac{mt^2}{2N}, \quad (1)$$

а $\text{th}(x)$ — гиперболический тангенс.

Несмотря на это, в проблеме оценки надежности метода Хеллмана имеется ряд нерешенных теоретических задач. Например, не описана в строгих математических терминах мера близости надежности и ее оценки, не выявлено влияние на надежность различных свойств труднообратимой функции G , не известна область значений параметров метода Хеллмана, в которой оценка (1) остается справедливой.

Исследования, проводимые в данной статье, направлены на решение этих проблем. Описан теоретико-вероятностный подход, в рамках которого без привлечения эвристических предположений удастся получить оценки отклонения среднего значения величины $\xi(m, t, N)/(mt)$ от $\delta(\alpha)$ (теорема 2). Кроме того, найденные соотношения позволяют выявить свойства труднообратимой функции G , существенно влияющие на надежность, и выявить область значений параметров, в которой оценка (1) остается справедливой.

В используемом теоретико-вероятностном подходе задача оценки среднего значения величины $\xi(m, t, N)$ сводится к задаче оценки средних значений некоторых характеристик подходящего процесса Гальтона–Ватсона (теорема 1). Для решения последней задачи имеется разработанный математический аппарат (см. [11–15]). Некоторые положения использованного теоретико-вероятностного подхода излагались автором ранее в [14–18].

2. Описание решаемой математической задачи

Для построения одной таблицы метода Хеллмана выбираются *подстановка редукции*

$$\pi: X \rightarrow X$$

и *множество начальных точек*

$$S \subset X, \quad |S| = m.$$

Для каждого начального элемента $x_{i,0} \in S$, $i = 1, \dots, m$, путем t -кратного последовательного применения функции $F = \pi(G(x))$ вычисляются *цепочки элементов*

$$x_{i,k+1} = F(x_{i,k}), \quad k = 0, 1, \dots, t-1.$$

Записи $(x_{i,0}, x_{i,t})$, $j = 1, \dots, m$, состоящие из начал и концов цепочек, размещаются в *таблицу*. Множество всех аргументов функции F , вовлеченных в процесс построения цепочек, образует *носитель таблицы*

$$\Xi = \{x_{i,k}, \quad i = 1, \dots, m, \quad k = 0, \dots, t-1\}.$$

Оперативный этап метода Хеллмана позволяет с помощью записей таблицы восстановить все решения уравнений $G(x) = a$, $a \in X$, лежащие в носителе таблицы Ξ .

Изучаемое число $\xi(m, t, N)$ различных элементов множества X , задействованных при построении одной таблицы, определяется мощностью носителя

$$\xi(m, t, N) = |\Xi|.$$

Далее будем полагать, что отображение G фиксированно и выполнены следующие теоретико-вероятностные предположения относительно остальных параметров, определяющих вероятностное пространство:

- подстановка π получена в результате случайной равновероятной выборки из множества Ω_R всевозможных подстановок на множестве X ,
- множество начальных точек S получено в результате случайной равновероятной выборки из множества Ω_S всевозможных подмножеств множества X мощности m ,
- выбор подстановки редукции производится независимо от выбора множества начальных точек.

Рассмотрим индикаторы $\rho(x) = I\{x \notin \Xi\}$, $x \in X$. Ввиду представления

$$\frac{\xi(m, t, N)}{mt} = \frac{1}{N} \sum_{x \in X} (1 - \rho(x)) \frac{N}{mt}$$

изучение среднего значения величины $\xi(m, t, N)$ сводится к изучению средних значений величин $\rho(x)$, $x \in X$.

Перейдем к их исследованию.

3. О связи задач оценки средних значений величины $\rho(x)$ и некоторой характеристики случайного шлейфа

Зафиксируем произвольный элемент $x_0 \in X$ и представим с помощью случайных элементов π, S величину $\rho(x_0)$. Для этого введем ряд определений.

С любым отображением $F: X \rightarrow X$ вида $F(x) = \pi(G(x))$, $\pi \in \Omega_R$, свяжем *граф отображения*. Его вершинами являются элементы множества X . Из каждой вершины $x \in X$ выходит ровно одна дуга. Ее конец находится в вершине $F(x)$.

Удалим из графа отображения дугу с началом в вершине x_0 и изменим ориентацию дуг на противоположную. *Шлейфом* $T(\pi)$ вершины x_0 назовем максимальное поддерево полученного графа с корнем в вершине x_0 .

Множество вершин шлейфа $T(\pi)$, высота которых меньше t , обозначим символом $\vartheta(T(\pi), t)$.

Для произвольного дерева T далее будем использовать следующие обозначения:

- $\mu(T, t)$ — число вершин высоты t ,
- $v(T, t)$ — число вершин высоты меньше t ,
- $v_d(T, t)$ — число вершин высоты меньше t и кратности d .

С помощью введенных понятий получаем следующее представление:

$$\rho(x_0) = I\{x_0 \notin \Xi\} = I\{\vartheta(T(\pi), t) \cap S = \emptyset\}.$$

Учитывая, что для любого фиксированного множества $M \subset X$, $|M| \leq N-m$, верна цепочка равенств

$$\mathbf{E}I\{M \cap S = \emptyset\} = \mathbf{P}\{M \cap S = \emptyset\} = \frac{C_{N-|M|}^m}{C_N^m},$$

находим

$$\begin{aligned} \mathbf{E}\rho(x_0) &= \mathbf{E}(\mathbf{E}(I\{\vartheta(T(\pi), t) \cap S = \emptyset\} | \pi)) = \\ &= \mathbf{E}I\{v(T(\pi), t) \leq N-m\} \cdot \frac{C_{N-v(T(\pi), t)}^m}{C_N^m}. \end{aligned}$$

Таким образом, изучение среднего значения величины $\rho(x_0)$ сводится к изучению среднего значения некоторой характеристики случайного шлейфа $T(\pi)$.

4. О связи задач оценки средних значений величины $\rho(x)$ и некоторых характеристик случайного генеалогического дерева

Под *генеалогическим деревом* будем понимать ориентированное дерево, у которого вершинами высоты i , $i=0, 1, \dots$, являются $(i+1)$ -граммы целых чисел из промежутка $0, \dots, N$, причем корнем является однограмма (0) . Если вершина $(j_0, \dots, j_i)$ генеалогического дерева имеет высоту i и кратность d ($d < N$), то все исходящие из нее дуги заканчиваются в вершинах $(j_0, \dots, j_i, 0), (j_0, \dots, j_i, 1), \dots, (j_0, \dots, j_i, d-1)$.

Множество всех конечных генеалогических деревьев обозначим символом $\mathcal{H}$. С каждым шлейфом свяжем генеалогическое дерево.

Пусть множество X некоторым образом упорядочено. Рассмотрим шлейф $T(\pi)$ вершины x_0 . Сопоставим корню однограмму (0) , а вершинам высоты i , $i=1, 2, \dots$, сопоставим $(i+1)$ -граммы неотрицательных чисел, соблюдая при этом следующее правило:

— если из вершины $x \in X$ шлейфа исходит d дуг в вершины $x_1, x_2, \dots, x_d$, $x_1 < x_2 < \dots < x_d$, и вершине x сопоставлена $(i+1)$ -грамма $(j_0, \dots, j_i)$, то вершинам $x_1, x_2, \dots, x_d$ сопоставляются $(i+2)$ -граммы

$$(j_0, \dots, j_i, 0), (j_0, \dots, j_i, 1), \dots, (j_0, \dots, j_i, d-1)$$

соответственно.

Тем самым задается отображение φ множества шлейфов элемента x_0 в множество генеалогических деревьев.

Вероятностную меру на множестве генеалогических деревьев зададим по формуле

$$\mathbf{P}_0(H) = \mathbf{P} \{ \varphi(T(\pi)) = H \}, \quad H \in \mathcal{H}.$$

Случайное генеалогическое дерево с этим распределением обозначим символом $\mathcal{T}_0$.

Легко заметить, что с учетом введенных обозначений верно следующее представление

$$\mathbf{E} \rho(x_0) = \mathbf{E} I \{ v(\mathcal{T}_0, t) \leq N - m \} \frac{C_{N-v(\mathcal{T}_0, t)}^m}{C_N^m}.$$

5. Определение двойственного генеалогического дерева

Опишем еще одно распределение на множестве генеалогических деревьев, которое обозначим символом $\mathbf{P}_1$ и назовем *двойственным к распределению $\mathbf{P}_0$* .

Напомним, что процессом Гальтона–Ватсона $\mu(i), i = 0, 1, \dots$, называют процесс размножения и гибели однотипных частиц в целочисленные моменты времени $i = 0, 1, 2, \dots$. В начальный момент времени $i = 0$ существует одна частица. Каждая частица, существующая в момент времени i , независимо от предыстории и остальных частиц, с вероятностью p_d порождает d , $d = 0, 1, \dots$, частиц-потомков, а сама исчезает. Частицы-потомки размножаются в момент времени $i + 1$.

Символами $v(t)$ и $v_d(t)$, $d = 0, 1, \dots$, обозначим число частиц, существовавших в процессе $\mu(i)$, $i = 0, 1, \dots$, до момента времени t , и число частиц, существовавших в процессе $\mu(i)$, $i = 0, 1, \dots$, до момента времени t и имеющих d непосредственных потомков.

Рассмотрим в качестве распределения числа непосредственных потомков частицы в процессе $\mu(i)$, $i = 0, 1, \dots$, распределение числа сыновей корневой вершины дерева $\mathcal{T}_0$:

$$p_d = \mathbf{P} \{ \mu(1) = d \} = \mathbf{P} \{ \mu(\mathcal{T}_0, 1) = d \}, \quad d = 0, \dots, N.$$

Легко заметить, что величины p_d , $d = 0, \dots, N$, задаются формулами

$$p_d = \frac{n_d}{N}, \quad n_d = |S_d|, \quad S_d = \{ x \in X \mid |G^{-1}(x) \setminus x_0| = d \}, \quad d = 0, 1, \dots, N.$$

Среднее значение числа непосредственных потомков одной частицы в таком ветвящемся процессе при любом отображении G и любом элементе x_0 описывается формулой

$$\mathbf{E} \mu(1) = 1 - \frac{1}{N}$$

(т. е. процесс $\mu(i), i=0, \dots$, является докритическим).

Если ассоциировать вершины высоты $i, i=0, 1, \dots$, генеалогического дерева с частицами, существовавшими в процессе Гальтона–Ватсона в момент времени i , и считать, что частица, находящаяся в начале дуги, породила частицу, находящуюся в конце дуги, то каждой вырождающейся реализации процесса Гальтона–Ватсона будет соответствовать одно из деревьев множества $\mathcal{H}$. Приписав дереву вероятность соответствующей реализации, получим вероятностную меру $\mathbf{P}_1$ на множестве $\mathcal{H}$. Случайное генеалогическое дерево с этим распределением обозначим символом $\mathcal{T}_1$.

Многие характеристики случайной величины $\mu(1)$ зависят от выбора x_0 . Некоторые из них можно оценить через соответствующие характеристики случайной величины $\tilde{\mu}$, распределение которой уже не зависит от x_0 и имеет вид

$$\mathbf{P} \{\tilde{\mu}=d\} = \frac{|\{x \in X \mid |G^{-1}(x)|=d\}|}{N}, \quad d=0, \dots, N.$$

Легко заметить, что моменты $\mu(1)$ не превосходят соответствующие моменты $\tilde{\mu}$, а число различных значений, которое величина $\mu(1)$ принимает с ненулевой вероятностью, не превышает величины $R+1$, где R — соответствующее число для $\tilde{\mu}$.

6. Формулировка теоремы о близости средних значений одинаковой характеристики двойственных случайных деревьев $\mathcal{T}_0, \mathcal{T}_1$

Обозначим через $\mathcal{F}_t$ σ -алгебру событий, которую задают генеалогические деревья, описывающие реализацию первых t поколений процесса Гальтона–Ватсона $\mu(i), i=0, 1, \dots$.

Оценку близости средних значений одинаковой характеристики двойственных случайных деревьев $\mathcal{T}_0, \mathcal{T}_1$ дает следующая теорема.

Теорема 1. Пусть $\eta : \mathcal{H} \rightarrow \mathbb{R}$ — $\mathcal{F}_t$ -измеримая неотрицательная функция, а R — число различных значений случайной величины $\tilde{\mu}$

$$R = \sum_{d=0}^N I \{ \{x \in X \mid |G^{-1}(x)| = d\} \neq \emptyset \}.$$

Если $\eta \leq C$, то верна оценка

$$|\mathbf{E} \eta(\mathcal{T}_0) - \mathbf{E} \eta(\mathcal{T}_1)| \leq \frac{CRt}{N} \left(2 + \frac{1}{1-1/R} \left({}^{1-1/R}\sqrt{e} + 1 \right) \right).$$

В доказательстве данной теоремы используется ряд вспомогательных утверждений.

7. Свойства плотности меры P_0 относительно меры P_1

Элементарные события алгебры $\mathcal{F}_t$ будем кодировать генеалогическими деревьями, высота которых не превосходит t . Множество всех таких деревьев обозначим символом $\mathcal{H}_t$.

Событие, закодированное деревом $H \in \mathcal{H}_t$, образуется всеми деревьями $H' \in \mathcal{H}$, удовлетворяющими соотношению $C_t(H') = H$, где через $C_t(H')$ обозначено поддерево дерева H' , состоящее из вершин, высота которых не превосходит t .

Обозначим через $\mathbf{P}_{0,t}$ и $\mathbf{P}_{1,t}$ сужения мер $\mathbf{P}_0$ и $\mathbf{P}_1$ на σ -алгебру $\mathcal{F}_t$ (т. е. для любых $H \in \mathcal{H}_t$ верны равенства $\mathbf{P}_{0,t}(H) = \mathbf{P}(C_t(\mathcal{T}_0) = H)$ и $\mathbf{P}_{1,t}(H) = \mathbf{P}(C_t(\mathcal{T}_1) = H)$).

Пользуясь независимостью размножения частиц в процессе Гальтона–Ватсона $\mu(i), i=0, 1, \dots$, легко показать, что верна формула (везде далее полагаем, что $0^0 = 1$)

$$\mathbf{P}_{1,t}(H) = \prod_{d=0}^N p_d^{v_d(H,t)} = \frac{1}{N^{v(H,t)}} \prod_{d=0}^N n_d^{v_d(H,t)}$$

Заметим, что мера $\mathbf{P}_{0,t}$ абсолютно непрерывна относительно меры $\mathbf{P}_{1,t}$. Следовательно, на всех таких $H \in \mathcal{H}_t$, что $\mathbf{P}_{1,t}(H) > 0$, определена плотность $\gamma_t(H)$ меры $\mathbf{P}_{0,t}$ относительно меры $\mathbf{P}_{1,t}$.

В следующем утверждении приводятся точные формулы для величин $\mathbf{P}_{0,t}(H)$ и $\gamma_t(H)$.

Утверждение 1. Для любого дерева $H \in \mathcal{H}_t$ верны формулы

$$\begin{aligned}
 1) \mathbf{P}_{0,t}(H) &= \prod_{d=0}^N I\{v_d(H, t) \leq n_d\} \times \\
 &\quad \times \frac{(N - v(H, t))!}{N!} \prod_{d=0}^N \frac{n_d!}{(n_d - v_d(H, t))!}, \\
 2) \mathbf{P}_{0,t}(H) &= \mathbf{P}_{1,t}(H) \cdot \gamma_t(H), \text{ где} \\
 \gamma_t(H) &= \prod_{d=0}^N I\{v_d(H, t) \leq n_d\} \times \\
 &\quad \times \prod_{d: v_d(H, t) \geq 1} \prod_{i=1}^{v_d(H, t)-1} \left(1 - \frac{i}{n_d}\right) \prod_{i=1}^{v(H, t)-1} \left(1 - \frac{i}{N}\right)^{-1}.
 \end{aligned}$$

Доказательство. Вычислим величину $\mathbf{P}_{0,t}(H)$. По определению она равна доле подстановок π среди всех подстановок из Ω_R , которые порождают шлейф $T(\pi)$, удовлетворяющий соотношению

$$C_t(\varphi(T(\pi))) = H.$$

Любую подстановку из этого множества будем называть **правильной**. Подсчитаем число правильных подстановок. Для этого укажем разбиение множества правильных подстановок на равномошные подмножества, после чего подсчитаем их мощность и число.

Для краткости обозначим величины $v_d(H, t)$ символами v_d , а величину $v(H, t)$ символом v . При $v_d \leq n_d$, $d=0, \dots, N$, с каждым набором

$$Sp = \{Sp(d), d=0, \dots, N\}$$

упорядоченных множеств

$$Sp(d) = \{y_{d,i} \in S_d, i=0, \dots, v_d - 1\}$$

из v_d различных элементов множества S_d , $d=0, \dots, N$, свяжем подмножество правильных подстановок $\pi(Sp)$. Оно строится с помощью следующего алгоритма.

Зададим порядок обхода дерева H , при котором вершины меньшей высоты обходятся раньше вершин большей высоты. Обойдем все вершины дерева H от корня до вершин высоты $t-1$ включительно, помечая их элементами множества X и заполняя некоторые переходы подстановки π^{-1} .

Перед началом обхода корневую вершину (0) пометим элементом x_0 .

Находясь в вершине $(j_0, \dots, j_i)$, помеченной элементом $x \in X$ и имеющей d сыновей $(j_0, \dots, j_i, 0), \dots, (j_0, \dots, j_i, d-1)$, заполним переход подстановки π^{-1} из точки x первым элементом в упорядоченном множестве $Sp(d)$ и удалим его оттуда. Пусть это был элемент y и $G^{-1}(y) = \{x_1, \dots, x_d\}$, причем $x_1 < x_2 < \dots < x_d$. Пометим вершины $(j_0, \dots, j_i, 0), \dots, (j_0, \dots, j_i, d-1)$ элементами $x_1, \dots, x_d$ соответственно (если $d = 0$, то этап пометки не выполняется).

После данного обхода все вершины дерева H окажутся помеченными различными элементами множества X , а в подстановке π^{-1} окажутся определенными $v = \sum_{d=0}^N v_d$ переходов.

Множество подстановок $\pi(Sp)$ составим из всех подстановок π , которые при обращении будут содержать все определенные переходы. По построению для любой подстановки π из множества $\pi(Sp)$ выполняется равенство $C_t(\varphi(T(\pi))) = H$, поэтому все подстановки из $\pi(Sp)$ правильные.

Легко заметить, что верно и обратное. Для любой правильной подстановки π можно построить ровно один такой набор упорядоченных множеств Sp , что выполняется включение $\pi \in \pi(Sp)$. Для этого надо совершить аналогичный обход дерева H , в котором шаг построения перехода подстановки π^{-1} по первому элементу множества S_d надо заменить на шаг добавления первого элемента в множество S_d по известному переходу подстановки π^{-1} .

Таким образом, при $v_d \leq n_d$, $d = 0, \dots, N$, объединение непересекающихся множеств $\pi(Sp)$ по всевозможным наборам Sp порождает все множество правильных подстановок. Если же хотя бы одно из условий $v_d \leq n_d$, $d = 0, \dots, N$, не выполняется, то построить правильную подстановку невозможно.

При $v_d \leq n_d$, $d = 0, \dots, N$, число всевозможных наборов упорядоченных множеств $Sp = \{Sp(d), d = 0, \dots, N\}$ описывается величиной

$$\prod_{d=0}^N \frac{n_d!}{(n_d - v_d)!}.$$

Мощность произвольного множества $\pi(Sp)$ равна числу $(N - v)!$ подстановок из Ω_R с v фиксированными переходами. Таким образом, при $v_d \leq n_d$, $d = 0, \dots, N$, число всех правильных подстановок равно

$$(N - v)! \prod_{d=0}^N \frac{n_d!}{(n_d - v_d)!}.$$

Доля правильных подстановок среди всех подстановок из Ω_R описывается формулой

$$\mathbf{P}_{0,t}(H) = \frac{(N-v)!}{N!} \cdot \prod_{d=0}^N \frac{n_d!}{(n_d - v_d)!}.$$

Если же условие $v_d \leq n_d$, $d = 0, \dots, N$, нарушено, то $\mathbf{P}_{0,t}(H) = 0$. Отсюда получаем справедливость первого пункта утверждения.

Для доказательства второго пункта достаточно заметить, что если выполнено условие $v_d \leq n_d$, $d = 0, \dots, N$, то $\mathbf{P}_{1,t}(H) > 0$, и, следовательно, определено частное

$$\begin{aligned} \gamma_t(H) &= \frac{\mathbf{P}_{0,t}(H)}{\mathbf{P}_{1,t}(H)} = \frac{(N-v)!N^v}{N!} \cdot \prod_{d=0}^N \frac{n_d!}{(n_d - v_d)!n_d^{v_d}} = \\ &= \prod_{d: v_d \geq 1} \prod_{i=1}^{v_d-1} \left(1 - \frac{i}{n_d}\right) \prod_{i=1}^{v-1} \left(1 - \frac{i}{N}\right)^{-1}. \quad \square \end{aligned}$$

8. Некоторые оценки плотности

В данном разделе приведем результаты, позволяющие оценивать функцию плотности.

Утверждение 2. Пусть имеются два набора из k , $k > 0$, натуральных чисел

$$(v_1, \dots, v_k), \quad (n_1, \dots, n_k).$$

Построим два набора из $v = \sum_{d=1}^k v_d$ чисел

$$A = \bigcup_{d=1}^k \left\{ \frac{1}{n_d}, \frac{2}{n_d}, \dots, \frac{v_d}{n_d} \right\}, \quad B = \left\{ \frac{1}{N}, \frac{2}{N}, \dots, \frac{v}{N} \right\}, \quad \text{где } N = \sum_{d=1}^k n_d.$$

(В наборе A одинаковые числа не отождествляются.) Тогда существует биекция ψ множества A на множество B , для которой при всех $a \in A$ выполняется неравенство $a \geq \psi(a)$.

Доказательство. Выпишем элементы множеств A и B в порядке убывания $a_1, \dots, a_v$, $b_1, \dots, b_v$. Для доказательства утверждения достаточно показать, что справедливы неравенства $a_i \geq b_i$, $i = 1, \dots, v$.

Множества

$$\left\{ \frac{1}{n_d}, \frac{2}{n_d}, \dots, \frac{v_d}{n_d} \right\}, \quad d = 1, \dots, k,$$

обозначим символами A_d , $d = 1, \dots, k$. Максимальный элемент множества A_d обозначим символом $a_d(1)$, $d = 1, \dots, k$.

Далее последовательно при $i = 2, 3, \dots, v$ будем удалять в одном из множеств $A_d, d = 1, \dots, k$, элемент, совпадающий с максимальным элементом объединения $\bigcup_{d=1}^k A_d$. Максимальный элемент множества A_d на шаге $i, i = 2, 3, \dots, v$, обозначим символом $a_d(i), d = 1, \dots, k$ (если на шаге i выполняется $A_d = \emptyset$, то положим $a_d(i) = 0$).

Легко заметить, что при любых $i = 1, \dots, v$ выполнены следующие соотношения:

$$1) a_i = \max_{d=1, \dots, k} \{a_d(i)\},$$

$$2) b_i = \sum_{d=1}^k q_d a_d(i), \quad q_d = n_d / N, \quad d = 1, \dots, k.$$

Отсюда получаем неравенство

$$b_i = \sum_{d=1}^k q_d \cdot a_d(i) \leq \sum_{d=1}^k q_d \cdot \max_{d \in 1, k} \{a_d(i)\} = \sum_{d=1}^k q_d \cdot a_i = a_i,$$

которое и доказывает утверждение. $\square$

Везде далее элементы $a \in A$ и $b \in B$, связанные биекцией $b = \psi(a)$, будем называть **парными элементами**.

Далее нам понадобится следующий результат.

Следствие 1. Пусть имеется два набора из $k, k > 0$, натуральных чисел

$$(v_1, \dots, v_k), \quad v_1 + \dots + v_k = v, \quad (n_1, \dots, n_k), \quad n_1 + \dots + n_k = n,$$

и натуральное число N .

1) Если $N \geq n$, то верно неравенство

$$\left(\sum_{d=1}^k \sum_{i=1}^{v_d} \frac{i}{n_d} - \sum_{i=1}^v \frac{i}{N} \right) \geq 0.$$

2) Если $N \geq n$ и $v_d < n_d, d = 1, \dots, k$, то верно неравенство

$$\prod_{d=1}^k \prod_{i=1}^{v_d} \left(1 - \frac{i}{n_d} \right) \prod_{i=1}^v \left(1 - \frac{i}{N} \right)^{-1} \leq 1.$$

3) Если $N \geq n, v_d < n_d, d = 1, \dots, k$, и для некоторого $c > 0$ выполнено ограничение $v \leq Nc/(1+c)$, то верно неравенство

$$1 - \prod_{d=1}^k \prod_{i=1}^{v_d} \left(1 - \frac{i}{n_d} \right) \prod_{i=1}^v \left(1 - \frac{i}{N} \right)^{-1} \leq (1+c) \left(\sum_{d=1}^k \sum_{i=1}^{v_d} \frac{i}{n_d} - \sum_{i=1}^v \frac{i}{N} \right).$$

Доказательство. Учитывая неравенства $i/n \geq i/N$, $i = 1, \dots, v$, по утверждению 2 получаем, что каждому элементу множества

$$A = \bigcup_{d=1}^k \left\{ \frac{1}{n_d}, \dots, \frac{v_d}{n_d} \right\}$$

соответствует не превосходящий его парный элемент множества

$$B' = \left\{ \frac{1}{N}, \dots, \frac{v}{N} \right\}.$$

Оцениваемое в п. 1 выражение представляется суммой слагаемых вида $i/n_d - j/N$, где i/n_d и j/N — парные элементы множеств A и B' . Каждое из этих слагаемых неотрицательно, и, следовательно, вся сумма неотрицательна.

Оцениваемое в п. 2 выражение представляется в виде произведения положительных множителей вида $(1 - i/n_d)/(1 - j/N)$, где i/n_d и j/N — парные элементы множеств A и B' . Каждый из этих множителей не превышает 1, и, следовательно, все произведение не превышает 1.

Заметим далее, что верно следующее представление этих множителей

$$\left(1 - \frac{i}{n_d}\right) \left(1 - \frac{j}{N}\right)^{-1} = 1 - x, \quad x = \left(\frac{i}{n_d} - \frac{j}{N}\right) \left(1 - \frac{j}{N}\right)^{-1}.$$

Из условия $v/N \leq c/(1+c)$ находим оценки

$$\left(1 - \frac{j}{N}\right)^{-1} \leq \left(1 - \frac{v}{N}\right)^{-1} \leq 1 + c.$$

Следовательно, для любой пары парных элементов верна двусторонняя оценка

$$0 \leq x \leq (1 + c) \left(\frac{i}{n_d} - \frac{j}{N}\right).$$

По индукции несложно показать, что для произвольных величин x_i , удовлетворяющих ограничениям $0 \leq x_i < 1$, $i = 1, \dots, k$, справедлива оценка

$$1 - \prod_{i=1}^k (1 - x_i) \leq \sum_{i=1}^k x_i$$

(шаг индукции следует из оценок $1 - (1 - y_1)(1 - y_2) = y_1 + y_2 - y_1 y_2 \leq y_1 + y_2$ при $y_1 = x_k$ и $y_2 = 1 - \prod_{i=1}^{k-1} (1 - x_i)$).

Применяя данное неравенство к выражению, оцениваемому в п. 3, находим, что оно ограничено сверху суммой слагаемых вида $(1 + c)(i/n_d - j/N)$ по всем парам парных элементов. $\square$

9. Оценки средних значений некоторых функционалов от процессов Гальтона–Ватсона

Приведем некоторые свойства критических и докритических процессов Гальтона–Ватсона $\mu(i), i = 0, \dots$, которые нам понадобятся в дальнейшем. Часть этих свойств известна и приводится далее с доказательствами для полноты изложения. Символом $x^{(k)}$ будем обозначать k -ю факториальную степень числа x .

Утверждение 3. Пусть $\mu(i), i = 0, 1, \dots$, — процесс Гальтона–Ватсона, начинающийся с одной частицы, в котором вероятность каждой частицы породить $d, d = 0, 1, \dots$, непосредственных потомков равна p_d . Если два первых факториальных момента

$$A = \mathbf{E} \mu(1), \quad B = \mathbf{E} \mu^{(2)}(1)$$

— числа непосредственных потомков одной частицы конечны, причем $A \leq 1$, то при любых $i = 1, 2, \dots$ и $m \leq N/2$ верны соотношения

- 1) $\mathbf{E} v_d(i) \leq p_d i, \quad d = 0, 1, \dots,$
- 2) $\mathbf{E} v(i) \leq i,$
- 3) $\mathbf{E} v_d(i) = p_d \mathbf{E} v(i), \quad d = 0, 1, \dots,$
- 4) $\mathbf{E} \sum_{d, p_d > 0} \frac{v_d^{(2)}(i)}{p_d} = \mathbf{E} v^{(2)}(i),$
- 5) $\mathbf{E} v^{(2)}(i) \leq \frac{B i^3}{3} + i^2,$
- 6) $\mathbf{P} \left\{ \{v(i) \geq h\} \cup \bigcup_{d, p_d > 0} \{v_d(i) \geq h_d\} \right\} \leq \frac{i}{h} + \sum_{d, p_d > 0} p_d \frac{i}{h_d},$
- 7) $0 \leq \left| \mathbf{E} \left(1 - \frac{m}{N}\right)^{v(i)} - \mathbf{E} I \{v(i) \leq N - m\} \frac{C_{N-v(i)}^m}{C_N^m} \right| \leq$
 $\leq \frac{m}{(N-m)N} \cdot \left(\frac{B i^3}{3} + i^2 \right) + e^{-m/2}.$

Доказательство. Получим несколько серий рекуррентных соотношений между факториальными моментами величин $v(i), v_d(i), i = 0, 1, \dots$

Зафиксируем d и рассмотрим совместную производящую функцию

$$f_i(y, z) = \mathbf{E} y^{\mu(i)} z^{v_d(i)}$$

величин $\mu(i), v_d(i)$.

Символами $f_{i,y}(y, z)$ и $f_{i,z}(y, z)$ обозначим ее частные производные по переменным y и z соответственно

$$f_{i,y}(y, z) = \mathbf{E} \mu(i) y^{\mu(i)-1} z^{v_d(i)}, \quad f_{i,z}(y, z) = \mathbf{E} v_d(i) y^{\mu(i)} z^{v_d(i)-1}.$$

Пользуясь свойствами процесса Гальтона–Ватсона, легко показать, что выполняются соотношения

$$\begin{aligned} f_1(y, z) &= p_d z y^d + \sum_{k, k \neq d} p_k y^k, \\ f_i(y, z) &= f_1(f_{i-1}(y, z), z), \quad i = 2, 3, \dots \end{aligned} \quad (2)$$

Дифференцируя соотношения (2) по z в точке $y=1, z=1$, получаем при $i=2, 3, \dots$ равенства

$$\begin{aligned} f_{i,z}(1, 1) &= \left. \frac{\partial f_i(y, z)}{\partial z} \right|_{y=1, z=1} = \left. \frac{\partial f_1(f_{i-1}(y, z), z)}{\partial z} \right|_{y=1, z=1} = \\ &= f_{1,y}(f_{i-1}(1, 1), 1) f_{i-1,z}(f_{i-1}(1, 1), 1) + f_{1,z}(f_{i-1}(1, 1), 1). \end{aligned}$$

Учитывая соотношения

$$\begin{aligned} f_{1,z}(1, 1) &= p_d, \quad f_{1,y}(1, 1) = A, \\ f_i(1, 1) &= 1, \quad f_{i,z}(1, 1) = \mathbf{E} v_d(i), \quad i = 1, 2, \dots, \end{aligned}$$

находим серию рекуррентных соотношений

$$\mathbf{E} v_d(i) = A \cdot \mathbf{E} v_d(i-1) + p_d, \quad i = 2, 3, \dots \quad (3)$$

Дважды дифференцируя соотношения (2) по z в точке $y=1, z=1$, аналогичным образом находим при $i = 2, 3, \dots$ серию рекуррентных соотношений

$$\mathbf{E} v_d^{(2)}(i) = B \cdot (\mathbf{E} v_d(i-1))^2 + 2dp_d \mathbf{E} v_d(i-1) + A \cdot \mathbf{E} v_d^{(2)}(i-1). \quad (4)$$

Рассмотрим теперь совместную производящую функцию

$$f_i(y, z) = \mathbf{E} y^{\mu(i)} z^{v(i)}$$

величин $\mu(i), v(i), i = 2, 3, \dots$. Для нее верны соотношения

$$f_1(y, z) = z \left(\sum_d p_d y^d \right), \quad f_i(y, z) = f_1(f_{i-1}(y, z), z), \quad i = 2, 3, \dots$$

Дважды дифференцируя последние соотношения по переменной z в точке $y=1, z=1$, получаем при $i = 2, 3, \dots$ серию рекуррентных соотношений

$$\mathbf{E} v^{(2)}(i) = B \cdot (\mathbf{E} v(i-1))^2 + 2A \cdot \mathbf{E} v(i-1) + A \cdot \mathbf{E} v^{(2)}(i-1). \quad (5)$$

Используя серии рекуррентных соотношений (3)–(5), покажем справедливость утверждения.

Учитывая, что $A \leq 1$, из соотношений (3) находим серию неравенств

$$\mathbf{E} v_d(i) - \mathbf{E} v_d(i-1) \leq p_d, \quad i = 2, 3, \dots$$

Суммируя их по индексу i , получаем неравенство

$$\mathbf{E} v_d(i) \leq \mathbf{E} v_d(1) + (i-1) p_d = i p_d.$$

Таким образом, первый пункт утверждения доказан.

Просуммировав последнее неравенство по d , получим доказательство второго пункта утверждения

$$\mathbf{E} v(i) = \sum_d \mathbf{E} v_d(i) \leq \sum_d i p_d = i.$$

Суммируя по d соотношения (3), находим серию соотношений

$$\mathbf{E} v(i) = A \cdot \mathbf{E} v(i-1) + 1, \quad i = 2, 3, \dots$$

Умножим уравнения этой серии на p_d и вычтем их из уравнений (3). Получим следующую серию рекуррентных соотношений

$$\mathbf{E} (v_d(i) - p_d v(i)) = A \cdot \mathbf{E} (v_d(i-1) - p_d v(i-1)), \quad i = 2, 3, \dots$$

Следовательно, при $i = 1, 2, \dots$ верно равенство

$$\mathbf{E} (v_d(i) - p_d v(i)) = A^{i-1} \cdot \mathbf{E} (v_d(1) - p_d v(1)).$$

Учитывая, что $v_d(1) - p_d v(1) = 0$, получаем доказательство третьего пункта утверждения

$$\mathbf{E} v_d(i) = p_d \mathbf{E} v(i), \quad i = 1, 2, \dots$$

Заменяя $\mathbf{E} v_d(i)$ на $p_d \mathbf{E} v(i)$ в соотношениях (4), при $p_d > 0$ находим

$$\begin{aligned} \mathbf{E} \frac{v_d^{(2)}(i)}{p_d} &= A \cdot \mathbf{E} \frac{v_d^{(2)}(i-1)}{p_d} + \\ &+ B \cdot p_d (\mathbf{E} v(i-1))^2 + 2 d p_d \mathbf{E} v(i-1), \quad i = 2, 3, \dots \end{aligned}$$

Просуммируем эти соотношения по d и вычтем соотношения (5). Получим

$$\begin{aligned} & \mathbf{E} \sum_{d, p_d > 0} \frac{v_d^{(2)}(i)}{p_d} - \mathbf{E} v^{(2)}(i) = \\ &= A \cdot \left(\mathbf{E} \sum_{d, p_d > 0} \frac{v_d^{(2)}(i-1)}{p_d} - \mathbf{E} v^{(2)}(i-1) \right) = \\ &= A^{i-1} \cdot \left(\mathbf{E} \sum_{d, p_d > 0} \frac{v_d^{(2)}(1)}{p_d} - \mathbf{E} v^{(2)}(1) \right), \quad i = 2, 3, \dots \end{aligned}$$

Учитывая, что

$$\sum_{d, p_d > 0} \frac{v_d^{(2)}(1)}{p_d} - v^{(2)}(1) = 0,$$

получаем доказательство четвертого пункта утверждения.

Из соотношений (5) получаем серию оценок

$$\mathbf{E} v^{(2)}(i) - \mathbf{E} v^{(2)}(i-1) \leq B \cdot (\mathbf{E} v(i-1))^2 + 2 \mathbf{E} v(i-1), \quad i = 2, 3, \dots$$

Суммируя их по i и используя оценку $\mathbf{E} v(i-1) \leq (i-1)$, получаем доказательство пятого пункта утверждения

$$\mathbf{E} v^{(2)}(i) \leq B \cdot \sum_{j=1}^{i-1} j^2 + 2 \sum_{j=1}^{i-1} j \leq \frac{B i^3}{3} + i^2, \quad i = 2, 3, \dots$$

Доказательство шестого пункта утверждения следует из цепочки неравенств

$$\begin{aligned} & \mathbf{P} \left\{ \{v(i) \geq h\} \cup \{v_d(i) \geq h_d\} \right\} \leq \\ & \leq \mathbf{P}\{v(i) \geq h\} + \sum_{d, p_d > 0} \mathbf{P}\{v_d(i) \geq h_d\} \leq \\ & \leq \frac{\mathbf{E} v(i)}{h} + \sum_{d, p_d > 0} \frac{\mathbf{E} v_d(i)}{h_d} \leq \\ & \leq \frac{i}{h} + \sum_{d, p_d > 0} p_d \frac{i}{h_d}. \end{aligned}$$

Докажем седьмой пункт утверждения. Обозначим

$$S^* = \mathbf{E} \left(\left(1 - \frac{m}{N}\right)^{v(i)} - I \{v(i) \leq N - m\} \frac{C_{N-v(i)}^m}{C_N^m} \right).$$

Воспользуемся представлением

$$\frac{C_{N-v(i)}^m}{C_N^m} = \prod_{j=0}^{v(i)-1} \left(1 - \frac{m}{N-j}\right)$$

и заметим, что при любом $j < N - m$ верны неравенства

$$1 - \frac{m}{N} \geq 1 - \frac{m}{N-j}.$$

Следовательно, при $v(i) \leq N - m$ выполняется неравенство

$$\left(1 - \frac{m}{N}\right)^{v(i)} \geq \frac{C_{N-v(i)}^m}{C_N^m}$$

и верна оценка снизу $S^* \geq 0$. Докажем оценку сверху.

При $m \leq N/2$ справедливо неравенство

$$S^* \leq \mathbf{E} S_0 \cdot I \left\{v(i) \leq \frac{N}{2}\right\} + \mathbf{E} S_1 \cdot I \left\{v(i) > \frac{N}{2}\right\},$$

где

$$S_0 = \left(1 - \frac{m}{N}\right)^{v(i)} - \prod_{j=0}^{v(i)-1} \left(1 - \frac{m}{N-j}\right), \quad S_1 = \left(1 - \frac{m}{N}\right)^{v(i)}.$$

Заметим, что для всех неотрицательных j , ограниченных сверху величиной $N/2$, справедлива оценка

$$\frac{i}{N-j} \leq \frac{2}{N},$$

а для произвольного набора чисел $x_1, \dots, x_k$ из интервала $[0, 1)$ справедлива оценка

$$1 - \prod_{j=1}^k (1 - x_j) \leq \sum_{j=1}^k x_j.$$

Отсюда при $v(i) \leq N/2$ получаем следующую цепочку верхних оценок величины S_0 :

$$\begin{aligned} S_0 &= \left(1 - \frac{m}{N}\right)^{v(i)} \left(1 - \prod_{j=0}^{v(i)-1} \left(1 - \frac{m}{(N-m)} \cdot \frac{j}{(N-j)}\right)\right) \leq \\ &\leq 1 - \prod_{j=0}^{v(i)-1} \left(1 - \frac{m}{(N-m)} \cdot \frac{j}{(N-j)}\right) \leq \\ &\leq \frac{m}{(N-m)} \cdot \sum_{j=0}^{v(i)-1} \frac{j}{N-j} \leq 2 \frac{m}{(N-m)} \cdot \sum_{j=0}^{v(i)-1} \frac{j}{N} = \\ &= \frac{m}{(N-m)} \cdot \frac{v^{(2)}(i)}{N}. \end{aligned}$$

Для величины S_1 при $v(i) > N/2$ справедлива оценка

$$S_1 \leq e^{-v(i)m/N} \leq e^{-m/2}.$$

Объединяя две предыдущие оценки, находим искомую оценку

$$S^* \leq \frac{m}{(N-m)N} \cdot \mathbf{E} v^{(2)}(i) + e^{-m/2} \leq \frac{m}{(N-m)N} \cdot \left(\frac{B i^3}{3} + i^2\right) + e^{-m/2}. \quad \square$$

Следующее утверждение имеет вспомогательный характер и используется в доказательстве утверждения 5.

Утверждение 4. Пусть имеются две числовые последовательности $x_0, x_1, \dots, y_0, y_1, \dots$, а также дифференцируемая на отрезке $[a, b]$ функция $g: \mathbb{R} \rightarrow \mathbb{R}$. Если для некоторых положительных чисел ε , Δ , C и некоторого целого числа i , $i \geq 1$, выполнены условия

$$x_j, y_j \in [a, b] \quad \text{при} \quad j = 0, \dots, i-1, \quad (6)$$

$$-\frac{1}{\Delta} \leq g'(x) \leq C \quad \text{при} \quad x \in (a, b), \quad (7)$$

$$\max_{0 \leq j < i} \left| \left(\frac{x_{j+1} - x_j}{\Delta} - g(x_j) \right) - \left(\frac{y_{j+1} - y_j}{\Delta} - g(y_j) \right) \right| < \varepsilon, \quad (8)$$

то верна оценка

$$|x_i - y_i| \leq (|x_0 - y_0| + \varepsilon \Delta i) e^{C \Delta i}.$$

Доказательство. Обозначим

$$\rho_j = \frac{x_{j+1} - x_j}{\Delta} - g(x_j) \quad \text{и} \quad \delta_j = \frac{y_{j+1} - y_j}{\Delta} - g(y_j), \quad j = 0, \dots, i-1.$$

Используя формулу Лагранжа

$$g(x_j) - g(y_j) = g'(\xi_j)(x_j - y_j), \quad \xi_j \in (x_j, y_j),$$

при любом $j = 0, \dots, i-1$ получаем представление

$$\begin{aligned} x_{j+1} - y_{j+1} &= x_j - y_j + \Delta (g(x_j) - g(y_j)) + \Delta (\rho_j - \delta_j) = \\ &= (x_j - y_j)c_j + \Delta d_j, \end{aligned}$$

где

$$c_j = (1 + \Delta \cdot g'(\xi_j)), \quad d_j = (\rho_j - \delta_j).$$

Последовательно выражая $(x_{j+1} - y_{j+1})$ через $(x_j - y_j)$ при $j = i-1, \dots, 0$, получаем формулу

$$x_i - y_i = (x_0 - y_0) \prod_{k=0}^{i-1} c_k + \Delta \sum_{j=0}^{i-1} d_j \prod_{k=j+1}^{i-1} c_k.$$

Из условий утверждения следуют оценки

$$0 < c_j \leq e^{\Delta g'(\xi_j)} \leq e^{\Delta C}, \quad |d_j| \leq \varepsilon, \quad j = 0, \dots, i-1.$$

Отсюда находим

$$\begin{aligned} |x_i - y_i| &\leq |x_0 - y_0| \prod_{k=0}^{i-1} c_k + \Delta \sum_{j=0}^{i-1} |d_j| \prod_{k=j+1}^{i-1} c_k \leq \\ &\leq |x_0 - y_0| e^{\Delta \sum_{j=0}^{i-1} g'(\xi_j)} + \Delta \varepsilon \sum_{j=0}^{i-1} e^{\Delta \sum_{k=j+1}^{i-1} g'(\xi_k)} \leq \\ &\leq (|x_0 - y_0| + \Delta i \varepsilon) e^{\Delta i C}. \quad \square \end{aligned}$$

В следующем утверждении производящие функции случайных величин $v(i)$, $i = 0, 1, \dots$, и $\mu(1)$ обозначены символами

$$f_i(x) = \mathbf{E} x^{v(i)} \quad \text{и} \quad \Phi(x) = \mathbf{E} x^{\mu(1)}.$$

Утверждение 5. Пусть u, v, B, C — положительные числа,

$$\varepsilon = \frac{vt}{u} \cdot \max \left\{ \sqrt{\frac{v}{B}} \left(\frac{C}{2B} + \frac{B}{3} \right) + \left(1 - \frac{\Phi''(1)}{B} \right) + v, \sqrt{\frac{2v}{B}} + \sqrt{\frac{2}{vB}} (1 - \Phi'(1)) \right\},$$

$$\delta_t(u, v, B) = \sqrt{\frac{2}{B} \cdot \frac{v}{u^2}} \operatorname{th} \sqrt{\frac{B}{2} v t^2}.$$

Если $v < \min \{1, 1/(2B)\}$, $\Phi'(1) \leq 1$, $\Phi''(1) \leq B$, $\Phi'''(1) \leq C$,

$$\delta_t(u, v, B) \leq \sqrt{\frac{2}{B} \cdot \frac{v}{u^2}} - \varepsilon,$$

то выполняется оценка

$$\left| \frac{1}{u} (1 - f_t(1-v)) - \delta_t(u, v, B) \right| \leq \varepsilon.$$

Доказательство. Для доказательства утверждения опишем две числовые последовательности $x_i, y_i, i = 0, 1, \dots$. Затем покажем, что они удовлетворяют условиям (6)–(8), и оценим близость их членов с помощью утверждения 4.

Пусть последовательность $x_i, i = 0, 1, \dots$, задается формулой

$$x_i = \frac{1}{u} (1 - f_i(1-v)).$$

Заметим, что она начинается в нуле и возрастает.

Для производящей функции $f_i(x)$ известны следующие рекуррентные соотношения

$$f_i(x) = x \Phi(f_{i-1}(x)), \quad i = 1, 2, \dots$$

Отсюда находим представление

$$\begin{aligned} \frac{x_{i+1} - x_i}{\Delta} &= \frac{1}{u\Delta} (f_i(1-v) - (1-v)\Phi(f_i(1-v))) = \\ &= \frac{1}{u\Delta} (1 - x_i u - \Phi(1 - x_i u)) + \frac{v}{u\Delta} \Phi(1 - x_i u) = \\ &= \frac{1}{u\Delta} H(x_i u), \end{aligned}$$

$$H(x) = 1 - x - (1-v) \Phi(1-x).$$

Применяя формулу Тейлора к функции $H(x)$, находим

$$\begin{aligned}\frac{x_{i+1}-x_i}{\Delta} &= \tilde{g}(x_i) + \frac{(1-v)}{3!} \frac{u^2}{\Delta} \Phi'''(\xi_i) \cdot x_i^3, \quad \xi_i \in (1-x_i u, 1), \\ \tilde{g}(x) &= \frac{v}{u\Delta} - \frac{1-(1-v)\Phi'(1)}{\Delta} x - \frac{(1-v)\Phi''(1)}{2} \frac{u}{\Delta} x^2.\end{aligned}\quad (9)$$

Положим $\Delta = 1/t$ и рассмотрим функцию

$$q(x) = \sqrt{\frac{2z}{B}} \cdot \operatorname{th}\left(x \sqrt{\frac{Bz}{2}}\right), \quad z = \frac{v}{u\Delta}, \quad \overline{B} = B \frac{u}{\Delta}.$$

Она является решением дифференциального уравнения

$$q'(x) = g(q(x)), \quad g(x) = z - \frac{\overline{B}}{2} x^2,$$

с начальным условием $q(0) = 0$.

Построим последовательность

$$y_i = q\left(\frac{i}{t}\right), \quad i=0, 1, \dots$$

Анализ функции $q(x)$ позволяет заключить, что данная последовательность начинается в нуле, монотонно возрастает и лежит в промежутке $[a, b]$, где

$$a = 0, \quad b = \sqrt{\frac{2z}{B}}.$$

Положим $C = 0$ и заметим, что при $x \in [a, b]$ выполнено условие

$$-\frac{1}{\Delta} \leq g'(x) \leq C.$$

По формуле Тейлора для любого $i=0, 1, \dots$ получаем представление

$$\begin{aligned}\frac{y_{i+1}-y_i}{\Delta} &= \frac{q(i+1/t) - q(i/t)}{\Delta} = \\ &= q'\left(\frac{i}{t}\right) + q''(\lambda_i) \frac{\Delta}{2} = \\ &= g(y_i) + \frac{q''(\lambda_i)}{2} \Delta, \quad \lambda_i \in (i\Delta, (i+1)\Delta).\end{aligned}\quad (10)$$

Заметим, что при любом $i=0, 1, \dots$ верна двусторонняя оценка

$$0 \leq -\frac{q''(\lambda_i)}{2} \leq \sqrt{\frac{2}{27} \overline{B} z^3}.$$

В самом деле, дифференцируя соотношение $q'(x) = g(q(x))$, находим, что область значений функции $-q''(x)$ при $x \geq 0$ совпадает с областью значений функции

$$-g'(x)g(x) = \overline{B}x \left(z - \frac{\overline{B}}{2}x^2 \right) \quad \text{при } x \in [a, b].$$

Исследуя последнюю на экстремум, нетрудно показать, что при $x \in [a, b]$ выполняется двустороннее неравенство

$$0 \leq -\frac{q''(x)}{2} \leq \sqrt{\frac{2}{27} \overline{B} z^3}.$$

Из (9) и (11) следует, что при $i=0, 1, \dots$

$$\begin{aligned} & \left| \left(\frac{x_{i+1} - x_i}{\Delta} - g(x_i) \right) - \left(\frac{y_{i+1} - y_i}{\Delta} - g(y_i) \right) \right| = \\ & = \left| \left(\frac{x_{i+1} - x_i}{\Delta} - \tilde{g}(x_i) \right) - \left(\frac{y_{i+1} - y_i}{\Delta} - g(y_i) \right) + (\tilde{g}(x_i) - g(x_i)) \right| = \\ & = |S_{1,i}(x_i) + S_{2,i}(x_i) + S_3(x_i) - S_4(x_i)|, \end{aligned}$$

где

$$\begin{aligned} S_{1,i}(x) &= \frac{(1-v)}{3!} \frac{u^2}{\Delta} \Phi'''(\xi_i) x^3, & S_{2,i}(x) &= -\frac{q''(\lambda_i)}{2} \Delta, \\ S_3(x) &= \frac{u}{2\Delta} (B - (1-v)\Phi''(1)) x^2, & S_4(x) &= \frac{1 - (1-v)\Phi'(1)}{\Delta} x. \end{aligned}$$

Заметим, что если $x \in [a, b]$, то все эти слагаемые неотрицательны и верны следующие оценки:

$$\begin{aligned} S_{1,i}(x) &\leq \frac{u^2}{3!\Delta} C b^3 \leq \varepsilon_1, & \varepsilon_1 &= \frac{v^{3/2}}{\Delta u} \frac{C}{2B^{3/2}}, \\ S_{2,i}(x) &\leq \sqrt{\frac{2}{27} \overline{B} z^3} \Delta \leq \varepsilon_2, & \varepsilon_2 &= \frac{v^{3/2}}{\Delta u} \frac{\sqrt{\overline{B}}}{3}, \\ S_3(x) &\leq \frac{u}{2\Delta} (B - (1-v)\Phi''(1)) b^2 \leq \varepsilon_3, & \varepsilon_3 &= \frac{v}{u\Delta} \left(1 - \frac{\Phi''(1)}{B} + v \right), \\ S_4(x) &\leq \frac{1 - (1-v)\Phi'(1)}{\Delta} b \leq \varepsilon_4, & \varepsilon_4 &= \frac{v^{1/2}}{u\Delta} (1 - \Phi'(1) + v) \sqrt{\frac{2}{B}}. \end{aligned}$$

Отсюда получаем, что если выполняются условия

$$x_j \in [a, b], \quad j = 0, 1, \dots, i-1, \quad (11)$$

то справедливы неравенства

$$\max_{0 \leq j < i} \left| \left(\frac{x_{j+1} - x_j}{\Delta} - g(x_j) \right) - \left(\frac{y_{j+1} - y_j}{\Delta} - g(y_j) \right) \right| \leq \max \{ \varepsilon_1 + \varepsilon_2 + \varepsilon_3, \varepsilon_4 \} = \varepsilon,$$

и согласно утверждению 4

$$|x_i - y_i| \leq \varepsilon i \Delta.$$

Из последней оценки следует справедливость неравенства $x_i \leq y_i + \varepsilon i \Delta$. Поэтому если $y_i \leq b - \varepsilon i \Delta$, то условие (11) выполняется и при $j = i$. Применяя индукцию, находим, что оценка $|x_j - y_j| \leq \varepsilon j \Delta$ справедлива для $j = 0, 1, \dots, i$, где i — максимальное целое число, удовлетворяющее ограничению $y_i = q(i/t) \leq b - \varepsilon i \Delta$. Полагая $i = t$, находим, что если $q(1) \leq b - \varepsilon$, то $|x_t - y_t| \leq \varepsilon$. $\square$

10. Оценка близости средних значений через плотность мер

Рассмотрим измеримое пространство $(\Omega, \mathcal{F})$, $\Omega = \{\omega\}$, на котором заданы две вероятностные меры $\mathbf{P}$ и $\tilde{\mathbf{P}}$. Мере $\tilde{\mathbf{P}}$ будем полагать абсолютно непрерывной относительно меры $\mathbf{P}$, через $\gamma(\omega)$ обозначим плотность меры $\tilde{\mathbf{P}}$ относительно меры $\mathbf{P}$, через $\mathbf{E}$ будем обозначать среднее значение по мере $\mathbf{P}$, тогда среднее значение любой случайной величины κ по мере $\tilde{\mathbf{P}}$ будет записываться в виде $\mathbf{E} \kappa \gamma$.

Утверждение 6. Пусть $\eta(\omega)$, $\Gamma_1(\omega)$ и $\Gamma_2(\omega)$ — неотрицательные случайные величины, а плотность $\gamma(\omega)$ на некотором множестве $M \in \mathcal{F}$ представляется в виде

$$\gamma(\omega) I\{\omega \in M\} = \frac{\gamma_1(\omega)}{\gamma_2(\omega)} I\{\omega \in M\}.$$

Если $\eta(\omega) \leq C$ и при $\omega \in M$ выполнены неравенства

$$\begin{aligned} \gamma_1(\omega) &\geq 1, & \gamma_2(\omega) &\geq 1, \\ \Gamma_1(\omega) &\geq \gamma_1(\omega) - 1, & \Gamma_2(\omega) &\geq 1 - \frac{1}{\gamma_2(\omega)}, \end{aligned}$$

то верна оценка

$$|\mathbf{E} \eta \gamma - \mathbf{E} \eta| \leq C \cdot (\mathbf{P}(\overline{M}) + \mathbf{E}(\Gamma_1 + \Gamma_2)).$$

Доказательство. Далее символом I_M будем обозначать индикатор события $\{\omega \in M\}$.

Оцениваемую в утверждении величину $|\mathbf{E} \eta \gamma - \mathbf{E} \eta|$ обозначим символом S^* . Используем представление

$$S^* = |S_0 - S_1|,$$

где

$$S_0 = \mathbf{E} I\{\gamma > 1\} I_M \cdot \eta(\gamma - 1) + \mathbf{E} I_{\overline{M}} \cdot \eta \gamma,$$

$$S_1 = \mathbf{E} I\{\gamma \leq 1\} I_M \cdot \eta(1 - \gamma) + \mathbf{E} I_{\overline{M}} \cdot \eta.$$

Заметим, что как величина S_0 , так и величина S_1 неотрицательны. Следовательно, $S \leq C \cdot \max\{S_2, S_3\}$, где

$$S_2 = \mathbf{E} I\{\gamma > 1\} I_M \cdot (\gamma - 1) + \tilde{\mathbf{P}}\{\overline{M}\},$$

$$S_3 = \mathbf{E} I\{\gamma \leq 1\} I_M \cdot (1 - \gamma) + \mathbf{P}\{\overline{M}\}.$$

Получим оценки слагаемых, образующих величины S_2 и S_3 .

Верны цепочки оценок

$$\mathbf{E} I\{\gamma > 1\} I_M \cdot (\gamma - 1) = \mathbf{E} I\{\gamma_1 > \gamma_2\} I_M \cdot \left(\frac{\gamma_1}{\gamma_2} - 1\right) \leq$$

$$\leq \mathbf{E} I_M \cdot (\gamma_1 - 1) \leq \mathbf{E} \Gamma_1,$$

$$\mathbf{E} I\{\gamma \leq 1\} I_M \cdot (1 - \gamma) = \mathbf{E} I\{\gamma_1 \leq \gamma_2\} I_M \cdot \left(1 - \frac{\gamma_1}{\gamma_2}\right) \leq$$

$$\leq \mathbf{E} I_M \cdot \left(1 - \frac{1}{\gamma_2}\right) \leq \mathbf{E} \Gamma_2,$$

$$\tilde{\mathbf{P}}(\overline{M}) = 1 - \tilde{\mathbf{P}}(M) = \mathbf{P}(\overline{M}) + \mathbf{E} I_M(1 - \gamma) \leq$$

$$\leq \mathbf{P}(\overline{M}) + \mathbf{E} I\{\gamma \leq 1\} I_M \cdot (1 - \gamma) \leq$$

$$\leq \mathbf{P}(\overline{M}) + \mathbf{E} \Gamma_2.$$

Отсюда находим

$$S^* \leq C \cdot \max\{S_2, S_3\} \leq$$

$$\leq C \cdot \max\{\mathbf{E} \Gamma_1 + \mathbf{E} \Gamma_2 + \mathbf{P}\{\overline{M}\}, \mathbf{E} \Gamma_2 + \mathbf{P}\{\overline{M}\}\} =$$

$$= C \cdot (\mathbf{E} \Gamma_1 + \mathbf{E} \Gamma_2 + \mathbf{P}\{\overline{M}\}).$$

□

11. Доказательство теоремы 1

Для доказательства теоремы 1 применим утверждение 6 на измеримом пространстве $(\mathcal{H}, \mathcal{F}_t)$. Покажем, что справедливы все его условия.

Пусть множество $M \in \mathcal{F}_t$ состоит из всех деревьев $H \in \mathcal{H}$, для которых выполнены условия:

$$v(H, t) \leq \beta t^2 + 1 \text{ при некотором } \beta, 0 \leq \beta < N/t^2;$$

$$v_d(H, t) \leq n_d, \quad d=0, \dots, N.$$

Заметим, что для деревьев из множества M в этом случае выполняются неравенства

$$v(H, t) - 1 \leq \frac{c}{1+c} N \quad \text{при} \quad c = \frac{c'}{1-c'} \quad c' = \frac{\beta t^2}{N}.$$

Согласно утверждению 1

$$\gamma_t(H) \cdot I\{H \in M\} = I\{H \in M\} \cdot \frac{\gamma_1(H, t)}{\gamma_2(H, t)},$$

где

$$\begin{aligned} \gamma_1(H, t) &= \prod_{i=v'(H, t)}^{v(H, t)-1} \left(1 - \frac{i}{N}\right)^{-1}, \\ \gamma_2(H, t) &= \prod_{i=1}^{v'(H, t)-1} \left(1 - \frac{i}{N}\right) \prod_{d: v_d(H, t) \geq 1} \prod_{i=1}^{v_d(H, t)-1} \left(1 - \frac{i}{n_d}\right)^{-1}, \\ v'(H, t) &= 1 + \sum_{d: v_d(H, t) \geq 1} (v_d(H, t) - 1) = v(H, t) - \bar{R}(H, t), \\ \bar{R}(H, t) &= \left(\sum_{d=0}^N I\{v_d(H, t) \geq 1\} - 1 \right). \end{aligned}$$

При $H \in M$, очевидно, $\gamma_1(H, t) \geq 1$, и по п. 2 следствия 1 выполняется оценка $\gamma_2(H, t) \geq 1$.

Рассмотрим случайные величины

$$\begin{aligned} \Gamma_1(H, t) &= \frac{(1+c) R e^{cR}}{N} v(H, t), \\ \Gamma_2(H, t) &= (1+c) \left(\sum_{d: v_d(H, t) \geq 1} \sum_{i=1}^{v_d(H, t)-1} \frac{i}{n_d} - \sum_{i=1}^{v'(H, t)-1} \frac{i}{N} \right). \end{aligned}$$

Они неотрицательны (неотрицательность $\Gamma_2(H, t)$ следует из п. 1 следствия 1) и для всех $H \in M$

$$\Gamma_1(H, t) \geq \gamma_1(H, t) - 1, \quad \Gamma_2(H, t) \geq 1 - \frac{1}{\gamma_2(H, t)}.$$

В самом деле, в первом случае верна цепочка оценок

$$\begin{aligned} \gamma_1(H, t) - 1 &= \prod_{i=v'(H, t)}^{v(H, t)-1} \left(1 + \frac{i}{N} \frac{1}{(1-i/N)}\right) - 1 \leq \\ &\leq \prod_{i=v'(H, t)}^{v(H, t)-1} \left(1 + (1+c) \frac{i}{N}\right) - 1 \leq \exp \left\{ (1+c) \sum_{i=v'(H, t)}^{v(H, t)-1} \frac{i}{N} \right\} - 1 = \\ &= \exp \left\{ (1+c) \frac{(v(H, t) - v'(H, t))(v(H, t) + v'(H, t) - 1)}{2} \right\} - 1 \leq \\ &\leq e^{\lambda(H, t)} - 1 \leq \lambda(H, t) e^{\lambda(H, t)}, \end{aligned}$$

где

$$\lambda(H, t) = (1+c) \frac{\bar{R}(H, t)v(H, t)}{N}.$$

Требуемую оценку получаем из оценок $\bar{R}(H, t) \leq R$, $v(H, t) \leq \beta t^2 = c'N$ с помощью неравенств

$$\lambda(H, t) \leq (1+c) \frac{R \beta t^2}{N} = R(1+c) c' = Rc,$$

$$\lambda(H, t) e^{\lambda(H, t)} \leq \lambda(H, t) e^{Rc} \leq \frac{(1+c) R}{N} e^{Rc} v(t).$$

По п. 3 следствия 1

$$1 - \frac{1}{\gamma_2(H, t)} \leq (1+c) \left(\sum_{d: v_d(H, t) \geq 1} \sum_{i=1}^{v_d(H, t)-1} \frac{i}{n_d} - \sum_{i=1}^{v'(H, t)-1} \frac{i}{N} \right) = \Gamma_2(H, t).$$

Таким образом, условия утверждения 6 выполнены. Следовательно,

$$|\mathbf{E} \eta(\mathcal{T}_0) - \mathbf{E} \eta(\mathcal{T}_1)| \leq \mathbf{P} \{ \mathcal{T}_1 \notin M \} + \mathbf{E} \Gamma_1(\mathcal{T}_1, t) + \mathbf{E} \Gamma_2(\mathcal{T}_1, t).$$

Вычислим слагаемые, стоящие в правой части этого неравенства.

Пусть $\mu(i), i = 0, 1, \dots$, — процесс Гальтона–Ватсона, используемый для построения меры $\mathbf{P}_1$. Как было замечено ранее, процесс μ является докритическим. Следовательно, к нему применим п. 6 утверждения 3. В соответствии с ним получаем оценку

$$\begin{aligned} \mathbf{P}(\mathcal{T}_1 \notin M) &= \mathbf{P} \left\{ \{v(t) > \beta t^2\} \bigcup_{d=0}^N \{v_d(t) > n_d\} \right\} \leq \\ &\leq \frac{t}{\beta t^2} + \sum_{d, n_d > 0} \frac{n_d}{N} \frac{t}{n_d} \leq \frac{1}{\beta t} + \frac{Rt}{N}. \end{aligned}$$

Оценим второе слагаемое. Учитывая п. 2 утверждения 3, находим, что

$$\mathbf{E} \Gamma_1(\mathcal{T}_1, t) = \frac{(1+c) R e^{cR}}{N} \mathbf{E} v(t) \leq (1+c) e^{cR} \cdot \frac{Rt}{N}.$$

Оценим третье слагаемое. Сворачивая суммы арифметических прогрессий, получаем представление

$$\mathbf{E} \Gamma_2(H, t) = (1+c) \left(\mathbf{E} \sum_{d=0}^N \frac{v_d^{(2)}(t)}{2n_d} - \mathbf{E} \frac{(v(t) - \tilde{R}(t))^{(2)}}{2N} \right),$$

где $\tilde{R}(H, t) = (\sum_{d=0}^N I\{v_d(t) \geq 1\} - 1)$. Используя п. 4 утверждения 3, перепишем последнюю формулу в следующем виде:

$$\begin{aligned} \mathbf{E} \Gamma_2(H, t) &= (1+c) \left(\mathbf{E} \frac{v^{(2)}(t)}{2N} - \mathbf{E} \frac{(v(t) - \tilde{R}(t))^{(2)}}{2N} \right) = \\ &= (1+c) \left(\mathbf{E} \frac{2v(t)\tilde{R}(t) - \tilde{R}(t)(\tilde{R}(t)+1)}{2N} \right). \end{aligned}$$

Отсюда получаем цепочку неравенств

$$\mathbf{E} \Gamma_2(H, t) \leq (1+c) \mathbf{E} \frac{v(t)R}{N} \leq (1+c) \frac{Rt}{N}.$$

Собирая полученные оценки воедино, получаем итоговую оценку

$$\begin{aligned} |\mathbf{E} \eta(\mathcal{T}_0) - \mathbf{E} \eta(\mathcal{T}_1)| &\leq C \left(\frac{1}{\beta t} + \frac{Rt}{N} + (1+c) e^{cR} \cdot \frac{Rt}{N} + (1+c) \frac{Rt}{N} \right) = \\ &= C \left(\frac{1}{\beta t} + \frac{Rt}{N} (1 + (1+c) (e^{cR} + 1)) \right). \end{aligned}$$

Положив $\beta = N/(t^2 R)$ ($c = 1/(R-1)$), находим

$$|\mathbf{E} \eta(\mathcal{T}_0) - \mathbf{E} \eta(\mathcal{T}_1)| \leq \frac{CRt}{N} \left(2 + \frac{1}{1-1/R} (1^{-1/R} e + 1) \right).$$

Теорема 1 доказана. $\square$

12. Оценка среднего значения величины $\frac{\xi(m,t,N)}{mt}$

Применим полученный результат к оценке среднего значения доли различных элементов множества X , задействованных при вычисления m цепочек длины t в процессе построения одной таблицы метода Хеллмана.

Ранее было показано, что в используемой теоретико-вероятностной модели его можно представить следующим образом:

$$\mathbf{E} \frac{\xi(m,t,N)}{mt} = \frac{1}{N} \sum_{x \in X} \mathbf{E} \eta(v(\mathcal{T}_{0,x}, t)),$$

где

$$\eta(j) = \frac{N}{mt} \left(1 - I \{j \leq N-m\} \frac{C_{N-j}^m}{C_N^m} \right),$$

а $v(\mathcal{T}_{0,x}, t)$ — число вершин высоты менее t в случайном генеалогическом дереве $\mathcal{T}_{0,x}$, распределение которого порождено случайным шлейфом элемента $x \in X$. Используем полученные результаты для оценки данной величины.

Пусть $\tilde{\mu}$ — случайная величина с распределением

$$\mathbf{P} \{\tilde{\mu}=d\} = \frac{|\tilde{S}_d|}{N}, \quad \tilde{S}_d = \{x \in X \mid |G^{-1}(x)| = d\}, \quad d=0, \dots, N.$$

Первые три факториальных момента $\tilde{\mu}$ обозначим через A , B и C . Заметим, что $A = 1$. Символом R обозначим число исходов, которое принимает (с ненулевой вероятностью) величина $\tilde{\mu}$

$$R = \sum_{d=0}^N I \{\mathbf{P} \{\tilde{\mu}=d\} > 0\}.$$

Символом $d_{\max}$ обозначим максимальное значение, которое принимает (с ненулевой вероятностью) величина $\tilde{\mu}$.

Теорема 2. Пусть

$$\alpha = \frac{mt}{N} \cdot \frac{B}{2}$$

и

$$\bar{\varepsilon} = \max \left\{ \sqrt{\frac{m}{BN}} \left(\frac{C}{2B} + \frac{B}{3} \right) + \frac{2(d_{\max}-1)}{BN} + \frac{m}{N}, \sqrt{\frac{2m}{BN}} + \sqrt{\frac{2}{mNB}} \right\}.$$

Если $\bar{\varepsilon} \leq (1 - \operatorname{th} \sqrt{\alpha})/\sqrt{\alpha}$, то справедлива оценка

$$\left| \frac{1}{N} \sum_{x \in X} \mathbf{E} \eta(v(\mathcal{T}_{0,x}, t)) - \frac{\operatorname{th} \sqrt{\alpha}}{\sqrt{\alpha}} \right| \leq \varepsilon_1 + \varepsilon_2 + \varepsilon_3,$$

где

$$\begin{aligned} \varepsilon_1 &= \frac{R}{m} \left(2 + \frac{1}{1-1/R} \left({}^{1-1/R}\sqrt{e} + 1 \right) \right) \\ \varepsilon_2 &= \frac{t^2}{(N-m)} \cdot \left(\frac{B}{3} + \frac{1}{t} \right) + \frac{N}{mt} e^{-m/2}, \\ \varepsilon_3 &= \max \left\{ \sqrt{\frac{m}{BN}} \left(\frac{C}{2B} + \frac{B}{3} \right) + \frac{2+m}{N}, \sqrt{\frac{2m}{BN}} + \sqrt{\frac{2}{mNB}} \right\}. \end{aligned}$$

Доказательство. Обозначим через $\mu_x(i)$, $i = 0, \dots$, процесс Гальтона–Ватсона, начинающийся с одной частицы, распределение числа потомков одной частицы в котором определяется формулами

$$\mathbf{P} \{ \mu_x(1) = d \} = \frac{n_{d,x}}{N}, \quad n_{d,x} = \left| \{ x' \in X \mid |G^{-1}(x') \setminus x| = d \} \right|, \quad d = 0, 1, \dots, N.$$

Первые три факториальных момента величины $\mu_x(1)$ обозначим символами A_x, B_x, C_x . Заметим, что если $x \in S_d$, $d > 0$, то верны следующие соотношения между моментами:

$$A = A_x + \frac{1}{N}, \quad B = B_x + \frac{2(d-1)}{N}, \quad C \geq C_x.$$

Через $v_x(t)$ обозначим суммарное число частиц, существовавших в процессе μ_x до момента времени t .

Определим функцию

$$\eta'(j) = \frac{N}{mt} \left(1 - \left(1 - \frac{m}{N} \right)^j \right).$$

Верна оценка

$$\left| \frac{1}{N} \sum_{x \in X} \mathbf{E} \eta(v(\mathcal{T}_{0,x}, t)) - \frac{\operatorname{th} \sqrt{\alpha}}{\sqrt{\alpha}} \right| \leq S_1 + S_2 + S_3,$$

где

$$\begin{aligned} S_1 &= \frac{1}{N} \sum_{x \in X} S_{1,x}, \quad S_{1,x} = \left| \mathbf{E} \eta(v(\mathcal{T}_{0,x}, t)) - \mathbf{E} \eta(v_x(t)) \right|, \\ S_2 &= \frac{1}{N} \sum_{x \in X} S_{2,x}, \quad S_{2,x} = \left| \mathbf{E} \eta(v_x(t)) - \mathbf{E} \eta'(v_x(t)) \right|, \\ S_3 &= \frac{1}{N} \sum_{x \in X} S_{3,x}, \quad S_{3,x} = \left| \mathbf{E} \eta'(v_x(t)) - \frac{\operatorname{th} \sqrt{\alpha}}{\sqrt{\alpha}} \right|. \end{aligned}$$

Оценим сверху слагаемые $S_{1,x}, S_{2,x}, S_{3,x}$, $x \in X$.

В силу оценки $\eta \leq N/(mt)$ и теоремы 1 при любом $x \in X$ справедливы неравенства

$$S_{1,x} \leq \frac{R}{m} \left(2 + \frac{1}{1-1/R} (1-1/\sqrt{e} + 1) \right), \quad S_1 \leq \varepsilon_1.$$

Из неравенства $B \geq B_x$ и п. 7 утверждения 3 следует, что при любом $x \in X$

$$S_{2,x} \leq \frac{t^2}{(N-m)} \cdot \left(\frac{B}{3} + \frac{1}{t} \right) + \frac{N}{mt} e^{-m/2}, \quad S_2 \leq \varepsilon_2.$$

Для оценки величины $S_{3,x}$ заметим, что величина $\mathbf{E} \eta'(v_x(t))$ совпадает с величиной, оцениваемой в утверждении 5 при $u = mt/N$ и $v = m/N$. При данных параметрах формулировка этого утверждения принимает следующий вид.

Пусть

$$\varepsilon_x = \max \left\{ \sqrt{\frac{v}{B}} \left(\frac{C}{2B} + \frac{B}{3} \right) + \left(1 - \frac{B_x}{B} \right) + v, \sqrt{\frac{2v}{B}} + \sqrt{\frac{2}{vB}} (1 - A_x) \right\}.$$

Если $\varepsilon_x \leq (1 - \text{th} \sqrt{\alpha}) / \sqrt{\alpha}$, то

$$S_{3,x} = \left| \mathbf{E} \eta'(v_x(t)) - \frac{\text{th} \sqrt{\alpha}}{\sqrt{\alpha}} \right| \leq \varepsilon_x.$$

Учитывая, что

$$1 - A_x = \frac{1}{N} \quad \text{и} \quad 1 - \frac{B_x}{B} \leq \frac{2d_{\max}}{NB},$$

при любом $x \in X$ получаем оценку $\varepsilon_x \leq \bar{\varepsilon}$. Из нее следует, что если $\bar{\varepsilon} \leq (1 - \text{th} \sqrt{\alpha}) / \sqrt{\alpha}$, то

$$\begin{aligned} S_1 &\leq \frac{1}{N} \sum_{x \in X} \max \left\{ \sqrt{\frac{v}{B}} \left(\frac{C}{2B} + \frac{B}{3} \right) + \left(1 - \frac{B_x}{B} \right) + v, \sqrt{\frac{2v}{B}} + \sqrt{\frac{2}{vB}} \frac{1}{N} \right\} \leq \\ &\leq \max \left\{ \sqrt{\frac{v}{B}} \left(\frac{C}{2B} + \frac{B}{3} \right) + \frac{1}{N} \sum_{x \in X} \left(1 - \frac{B_x}{B} \right) + v, \sqrt{\frac{2v}{B}} + \sqrt{\frac{2}{vB}} \frac{1}{N} \right\} = \\ &= \max \left\{ \sqrt{\frac{v}{B}} \left(\frac{C}{2B} + \frac{B}{3} \right) + \frac{2}{N} + v, \sqrt{\frac{2v}{B}} + \sqrt{\frac{2}{vB}} \frac{1}{N} \right\} = \varepsilon_3. \end{aligned}$$

Теорема 2 доказана. $\square$

Список литературы

- [1] Hellman M., “A cryptanalytic time-memory trade-off”, *IEEE Trans. Inf. Theory*, **IT-26**:4 (1980), 401–406.
- [2] Denning D., *Cryptography and Data Security*, Boston, Massachusetts, USA: Addison-Wesley, 1982, 100 pp.
- [3] Oechslin P., “Making a faster cryptanalytic time-memory trade-off”, *CRYPTO’03, Lect. Notes Comput. Sci.*, **2729**, 617–630.
- [4] Biryukov A., Shamir A., Wagner D., “Real time cryptanalysis of A5/1 on a PC”, *FSE 2000, Lect. Notes Comput. Sci.*, **1978**, 2001, 1–18.
- [5] Saarinen M.-J. O., “A time-memory tradeoff attack against LILI-128”, *FSE 2002, Lect. Notes Comput. Sci.*, **2365**, 2002, 231–236.
- [6] Mentens N., Batina L., Preneel B., Verbauwhede I., “Cracking Unix passwords using FPGA platforms”, *SHARCS’05 – Special Purpose Hardware for Attacking Cryptographic Systems*, 2005, 20–25.
- [7] Bono S., Green M., Stubblefield A., Juels A., Rubin A., Szydlo M., “Security analysis of a cryptographically-enabled RFID device”, *14th USENIX Security Symp.*, 2005, 1–16.
- [8] Kusuda K., Matsumoto T., “Optimization of time-memory trade-off cryptanalysis and its application to DES, FEAL-32, and Skipjack”, *IEICE Trans. Fund. Electr., Commun. and Comput. Sci.*, **E79-A(1)** (1996), 35–48.
- [9] Saran N., Doganaksoy A., “Choosing parameters to achieve a higher success rate for Hellman time memory trade off attack”, *Int. Conf. on Availability, Reliability and Security*, 2009, 504–509.
- [10] Ma D., Hong J., “Success probability of the Hellman trade-off”, *Inf. Process. Lett.*, **109**:7 (2009), 347–351.
- [11] Севастьянов Б. А., *Ветвящиеся процессы*, М.: Наука, 1971, 436 с.
- [12] Pakes A. G., “Some limit theorems for the total progeny of a branching process”, *Adv. Appl. Probab.*, **3**:1 (1971), 176–192.
- [13] Ватугин В. А., Сагитов С. М., “Разложимый критический ветвящийся процесс с двумя типами частиц”, *Труды Матем. ин-та АН СССР*, **177** (1986), 3–20.
- [14] Pilshchikov D. V., “Estimation of the characteristics of time-memory-data tradeoff methods via generating functions of the number particles and the total number of particles in the Galton-Watson process”, *Математические вопросы криптографии*, **5**:2 (2014), 103–108.
- [15] Pilshchikov D. V., “On the limiting mean values in probabilistic models of time-memory-data tradeoff methods”, *Математические вопросы криптографии*, **6**:2 (2015), 59–65.
- [16] Пильщиков Д. В., “Анализ сложности алгоритма параллельного поиска «золотой» коллизии”, *Математические вопросы криптографии*, **6**:4 (2015), 77–98.
- [17] Пильщиков Д. В., “Асимптотическое поведение мощности полного прообраза образа случайного множества при итерациях отображений конечного множества”, *Математические вопросы криптографии*, **8**:1 (2017), 95–106.
- [18] Пильщиков Д. В., “Исследование сложности метода радужных таблиц с маркерами цепочек”, *Математические вопросы криптографии*, **8**:4 (2017), 99–116.