О функциональной сложности двумерной задачи интервального поиска *

Э.Э.Гасанов, И.В.Кузнецова
Московский государственный университет

Аннотация

Предлагается модификация алгоритма Бентли-Маурера, решающего двумерную задачу интервального поиска. Эта модификация позволяет снизить исходное логарифмическое среднее время поиска до константного при сохранении логарифмического времени поиска в худшем случае. При этом этот алгоритм зависит от параметра, при вариации которого объем памяти, необходимый алгоритму, изменяется от $O(k^3)$ до $O(k \log k)$, при этом среднее время поиска (без учета времени на перечисление ответа) изменяется от $O(1)$ до $O(\log k)$. В частности, для любого $\varepsilon > 0$ при объеме памяти $O(k^{1+\varepsilon})$ достигается среднее время поиска $O(1)$. На основе этих результатов получены верхние оценки функциональной сложности двумерной задачи интервального поиска.

1 Введение

Данная работа посвящена исследованию сложности алгоритмов решения двумерной задачи интервального поиска, которая состоит в поиске в конечном подмножестве евклидовой плоскости всех тех точек, которые попадают в прямоугольник-запрос.

Типичная задача интервального поиска для двух измерений заключается в выявлении в базе данных, содержащей записи о служащих некоторой компании, всех служащих, чьи возраст и жалованье находятся в заданных интервалах. Это пример взят из [1]. Другой пример можно найти у Д. Кнута [2]. Он рассматривал базу данных городов США с координатами в виде широты и долготы. К такой базе естественен вопрос о перечислении всех городов, попадающих в некоторой прямоугольник-запрос. Подобные задачи возникают также в статистике [3] и автоматизации проектирования [4].

Исследованию задачи интервального поиска (в другом переводе с английского — регионального поиска) посвящено большое количество работ [1, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22]. Приведем результаты некоторых из них. Алгоритм решения двумерной задачи интервального поиска путем сведения к решению четырех двумерных задач о доминировании можно найти в [1, 6]. Бентли в [7] предложил метод многомерного двоичного дерева (k-D-дерева), исходя из того, что бинарное дерево для одномерной задачи дает хороший результат. Этот метод

*Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант 98-01-00130)

имеет линейные затраты по памяти, но Ли и Вонг [17] показали, что в худшем случае этим алгоритмом двумерная задача решается за время $\underline{O}(\sqrt{k})$. Здесь и далее, k равно мощности библиотеки, а оценки приводятся без времени перечисления ответа. В работе [8] Бентли предложил метод прямого доступа, который решает задачу за время $\underline{O}(\log k)$, но требует затрат по памяти порядка k^3. Чтобы уменьшить требуемую память в работе [10] Бентли и Маурером был предложен многоэтапный метод прямого доступа. Он позволяет снижать порядок требуемой памяти, но при этом возрастает константа при логарифме в оценке времени. С помощью метода дерева интервалов (или дерева регионов) Бентли и Шеймос [11] получили оценку времени поиска $\underline{O}(\log^2 k)$ при затратах памяти $\underline{O}(k \log k)$. Уиллард [22] и Люкер [19] независимо предложили модификацию дерева интервалов, которая позволила снизить время поиска до $\underline{O}(\log k)$ при тех же затратах памяти. Чазелле в [14] смог снизить затраты памяти до $\underline{O}(k \log k / \log \log k)$, но при этом возросла константа при логарифме в оценке времени. Во всех этих работах оценивается время поиска в худшем случае. Болоур в [13] описал метод хеширования, который он использовал вместо поиска по древовидным структурам в задаче интервального поиска. Применение этого метода позволило получить довольно быстрые в среднем решения задачи интервального поиска при условии, что область запросов находится в заранее определенных границах.

При оценке алгоритмов поиска обычно используют три основные характеристики: объем памяти, требуемый алгоритмом, время поиска в худшем случае и среднее время поиска. Естественно хотелось бы находить алгоритмы, "хорошие"в соответствии с той или иной характеристикой, или, возможно, в соответствии с комбинацией характеристик. Выбор критерия оптимизации является одним из основных моментов в теории синтеза управляющих систем (в частности, при синтезе оптимальных алгоритмов поиска). Условимся обозначать через Q объем требуемой памяти, а через T — время вычислений. В литературе в первую очередь встречаются такие критерии оптимизации: Q, T, $Q \cdot T$, $Q \cdot T^2$. Выбор конкретного критерия оптимизации осуществляется на практике с учетом специфики разрабатываемой вычислительной системы и особенностей ее использования, и многообразие таких критериев подсказывает необходимость проведения исследований глобальных зависимостей $Q(T)$ или $T(Q)$. Такая мера сложности носит название функциональной и была предложена А.С. Подколзиным [23].

Понятно, что любой алгоритм, который имеет оценки объема памяти и времени работы, дает верхнюю оценку функциональной сложности. В данной работе предлагается алгоритм решения двумерной задачи интервального поиска, зависящий от некоторого параметра. При вариации этого параметра получаются разные соотношения объема памяти и среднего времени поиска, что позволяет оценить функциональную сложность в некоторой последовательности точек. Данный алгоритм является модификацией алгоритма Бентли–Маурера, изложенного в [10], но в отличие от исходного алгоритма предлагаемый алгоритм при сохранении порядка оценок объема памяти и времени в худшем случае имеет константные оценки времени в среднем.

Результаты данной работы были анонсированы в [24, 25].

Авторы выражают благодарность В.Б.Кудрявцеву и А.С.Подколзину за поддержку в работе.

2 Основные понятия и формулировка результата

Мы будем использовать терминологию и обозначения из работы [26].

Пусть X — множество запросов с заданным на нем вероятностным пространством $\langle X, \sigma, \mathbf{P} \rangle$, где σ — алгебра подмножеств множества X, $\mathbf{P}$ — вероятностная мера на σ; Y — множество записей (объектов поиска); ρ — бинарное отношение на $X \times Y$, называемое отношением поиска. Пятерку $S = \langle X, Y, \rho, \sigma, \mathbf{P} \rangle$ будем называть *типом.* Тройку $I = \langle X, V, \rho \rangle$, где V — некоторое конечное подмножество множества Y, в дальнейшем называемое *библиотекой*, будем называть задачей информационного поиска (ЗИП) типа S, и будем считать, что ЗИП $I = \langle X, V, \rho \rangle$ содержательно состоит в перечислении для произвольно взятого запроса $x \in X$ всех тех и только тех записей $y \in V$ таких, что $x \rho y$.

Пусть f — одноместный предикат, определенный на X , то есть $f : X \to \{0, 1\}$. Множество $N_f = \{x \in X : f(x) = 1\}$ назовем *характеристическим множеством* предиката f.

Множество $O(y, \rho) = \{x \in X : x \rho y\}$ назовем *тенью* записи $y \in Y$.

Функцию $\chi_{y,\rho} : X \to \{0, 1\}$ такую, что $N_{\chi_{y,\rho}} = O(y, \rho)$ назовем *характеристической функцией* записи y.

Пусть F — множество символов одноместных предикатов, определенных на множестве X, G — множество символов одноместных переключателей, определенных на множестве X. Под переключателем будем понимать функцию, областью значений которой является конечное подмножество натурального ряда. Пару $\mathcal{F} = \langle F, G \rangle$ назовем базовым множеством.

Понятие *информационного графа* (ИГ) над базовым множеством $\mathcal{F} = \langle F, G \rangle$ определяется следующим образом. Берется конечная многополюсная ориентированная сеть. В ней выбирается некоторый полюс, который называется корнем. Остальные полюса называются листьями и им приписываются записи из Y, причем разным листьям могут быть приписаны одинаковые записи. Некоторые вершины сети (в том числе это могут быть и полюса) называются переключательными и им приписываются переключатели из G. Ребра, исходящие из каждой из переключательных вершин, нумеруются подряд, начиная с 1, и называются переключательными ребрами. Ребра, не являющиеся переключательными, называются предикатными и им приписываются предикаты из множества F. Таким образом нагруженную многополюсную ориентированную сеть называем ИГ над базовым множеством $\mathcal{F} = \langle F, G \rangle$.

Функционирование ИГ определяется следующим образом. Скажем, что предикатное ребро проводит запрос $x \in X$, если предикат, приписанный этому ребру, принимает значение 1 на запросе x, Переключательное ребро, которому приписан номер n, проводит запрос $x \in X$, если переключатель, приписанный началу этого ребра, принимает значение n на запросе x. Ориентированная цепочка ребер проводит запрос $x \in X$, если каждое ребро цепочки проводит запрос x. Запрос $x \in X$ проходит в вершину β ИГ, если существует ориентированная цепочка, ведущая из корня в вершину β, которая проводит запрос x. Запись y, приписанная листу α, попадает в ответ ИГ на запрос $x \in X$, если запрос x проходит в лист α. Ответом ИГ U на запрос x назовем множество записей, попавших в ответ ИГ на запрос x, и обозначим его $\mathcal{J}_U(x)$. Эту функцию $\mathcal{J}_U(x)$ будем считать результатом функционирования ИГ U.

Пусть нам дана ЗИП $I = \langle X, V, \rho \rangle$. Скажем, что ИГ U *разрешает* ЗИП $I =$

$\langle X, V, \rho\rangle$, если $\mathcal{J}_U(x) = \{y \in V : x\rho y\}$. ИГ U, разрешающий ЗИП I, будем также называть *допустимым* для I.

Введем понятие *сложности ИГ*. Пусть β — некоторая вершина ИГ. Предикат, определенный на множестве запросов, который принимает значение 1 на запросе x, если запрос проходит в вершину β, и 0 — в противном случае, назовем функцией фильтра вершины β и обозначим $\varphi_\beta(x)$.

Сложностью ИГ U на запросе $x \in X$ назовем число $T(U,x) = \sum_{\beta\in\mathcal{P}} \varphi_\beta(x) + \sum_{\beta\in\mathcal{R}\setminus\mathcal{P}} \psi_\beta \cdot \varphi_\beta(x)$, где $\mathcal{R}$ — множество вершин ИГ U, $\mathcal{P}$ — множество переключательных вершин ИГ U, ψ_β — количество ребер, исходящих из вершины β.

Скажем, что базовое множество $\mathcal{F}$ *измеримое*, если каждая функция из $\mathcal{F}$ — измеримая (относительно алгебры σ). Далее всюду будем предполагать, что базовое множество измеримое. В этом случае для любого ИГ U над $\mathcal{F}$ функция $T(U,x)$ как функция от x измерима.

Сложностью ИГ U назовем математическое ожидание величины $T(U,x)$, то есть число $T(U) = \mathbf{M}_x\, T(U,x)$. *Объемом $Q(U)$ ИГ U* назовем число ребер в графе U. *Сложностью задачи I при базовом множестве $\mathcal{F}$* назовем число $T(I,\mathcal{F}) = \inf\{T(U) : U \in \mathcal{U}(I,\mathcal{F})\}$, где $\mathcal{U}(I,\mathcal{F})$ — множество всех ИГ над базовым множеством $\mathcal{F}$, разрешающих ЗИП I.

Дадим описание типа задач поиска, который соответствует двумерной задаче интервального поиска.

Пусть $Y_{int2} = [0,1)^2$ — множество записей и

$$X_{int2} = \{\tilde{x} = (u_1, v_1, u_2, v_2) : 0 \le u_i \le v_i < 1,\ i = 1,2\} \text{ —}$$

множество запросов. Пусть на множестве X_{int2} задано вероятностное пространство $\langle X_{int2}, \sigma, \mathbf{P}\rangle$, где $\mathbf{P}$ задается функцией плотности вероятности $p(\tilde{x})$.

Отношение поиска ρ_{int2} определено на $X_{int2} \times Y_{int2}$ и задается следующим соотношением:

$$(u_1, v_1, u_2, v_2)\rho_{int2}(y_1, y_2) \iff u_i \le y_i \le v_i, \quad i = 1,2.$$

Тогда тип

$$S_{int2} = \langle X_{int2}, Y_{int2}, \rho_{int2}, \sigma, \mathbf{P}\rangle$$

назовем типом двумерного интервального поиска.

Пусть

$$G_1 = \Big\{g^{1,m}_{i,a,b}(\tilde{x}) = \Big[\frac{u_i - a}{b-a}\cdot m\Big] + 1,\ i = 1,2,\ m \in \mathbb{N},\ 0 \le a < b < 1\Big\}, \tag{1}$$

$$G_2 = \Big\{g^{2,m}_{a,b}(\tilde{x}) = \Big[\frac{v_1 - a}{b-a}\cdot m\Big] + 1,\ m \in \mathbb{N},\ 0 \le a < b < 1\Big\}, \tag{2}$$

$$G_3 = \Big\{g^{3}_{i,a}(\tilde{x}) = \begin{cases} 1, & \text{если } u_i \le a \\ 2, & \text{если } u_i > a \end{cases},\ i = 1,2,\ a \in [0,1)\Big\}, \tag{3}$$

$$G_4 = \Big\{g^{4}_{a}(\tilde{x}) = \begin{cases} 1, & \text{если } v_1 < a \\ 2, & \text{если } v_1 \ge a \end{cases},\ a \in [0,1)\Big\}, \tag{4}$$

$$G_5 = \Big\{g^{5}_{m}(\tilde{x}) = \begin{cases} 1, & \text{если } 0 \le v_2 - u_2 < 1/m \\ 2, & \text{в противном случае} \end{cases},\ m \in \mathbb{N}\Big\}, \tag{5}$$

$$G_6 = \Big\{g^{6}_{a,b}(\tilde{x}) = \begin{cases} 1, & \text{если } a < u_1 \le v_1 < b \\ 2, & \text{в противном случае} \end{cases},\ 0 \le a < b < 1\Big\}, \tag{6}$$

$$F_1 = \Big\{f_a(\tilde{x}) = \begin{cases} 1, & \text{если } u_2 \le a \le v_2 \\ 0, & \text{в противном случае} \end{cases},\ a \in [0,1)\Big\}. \tag{7}$$

Определим базовое множество

$$\mathcal{F} = \langle F_1, G_1 \cup G_2 \cup G_3 \cup G_4 \cup G_5 \cup G_6 \rangle. \tag{8}$$

Для удобства будем считать, что функция плотности вероятности $p(\tilde{x}) = p(u_1, v_1, u_2, v_2)$, определяющая вероятностную меру $\mathbf{P}$ вероятностного пространства над X_{int2}, задана на $[0,1)^2$, но вне X_{int2} она принимает значение 0.

Обозначим

$$X_{int} = \{(u,v) \mid 0 \leq u \leq v < 1\},$$

$$p_1(u_1, v_1) = \int_{X_{int}} p(\tilde{x})\, d\tilde{z}_2,$$

$$p_2(u_1, v_1, u_2, v_2) = p(\tilde{x}),$$

где $\tilde{z}_i = (u_i, v_i)$, $i = 1, 2$.

Пусть b, d, e, $f \in [0,1)$ и $b < d$, $e < f$, $b \leq e$.

Обозначим

$$p^1(u_1, v_1) = p_1(u_1, v_1),$$

$$p^2_{b,d,e,f}(u_2, v_2) = \int_b^d du_1 \int_e^f p_2(u_1, v_1, u_2, v_2)\, dv_1 \Big/ \int_b^d du_1 \int_e^f p_1(u_1, v_1)\, dv_1, \tag{9}$$

$$p^{1,b,d,e,f}(u_1) = \int_e^f p^1(u_1, v_1)\, dv_1 \Big/ \int_b^d du_1 \int_e^f p^1(u_1, v_1)\, dv_1, \tag{10}$$

$$p^{2,b,d,e,f}(v_1) = \int_b^d p^1(u_1, v_1)\, du_1 \Big/ \int_b^d du_1 \int_e^f p^1(u_1, v_1)\, dv_1, \tag{11}$$

Пусть c — некоторая вещественная константа. Скажем, что функция плотности вероятности $p(\tilde{x}) = p(u_1, v_1, u_2, v_2)$ обладает свойством c-ограниченности, если для любых b, d, e, $f \in [0,1)$, таких, что $b < d$, $e < f$, $b \leq e$ выполняются условия

$$(d-b)(f-e)p^2_{b,d,e,f}(u_2, v_2) \leq c, \tag{12}$$

$$(d-b)(f-e)p^{1,b,d,e,f}(u_1) \leq c, \tag{13}$$

$$(d-b)(f-e)p^{2,b,d,e,f}(v_1) \leq c. \tag{14}$$

Теорема 1 *Пусть $I = \langle X_{int2}, V, \rho_{int2} \rangle$ — двумерная задача интервального поиска, где $|V| = k$. Пусть $\mathcal{F}$ — базовое множество, определяемое соотношениями (1)–(8). Пусть $R(I) = \sum_{\tilde{y} \in V} \mathbf{P}(O(\tilde{y}, \rho_{int2}))$. Тогда если для некоторой константы c функция плотности вероятности $p(\tilde{x})$, определяющая вероятностную меру $\mathbf{P}$, обладает свойством c-ограниченности, то для любого натурального M такого, что $1 \leq M \leq 2\ln k$, справедливо*

$$R(I) \leq T\Big(I, \mathcal{F}, \frac{2}{3}Mk^{1+\frac{2}{M}} + \underline{O}\Big(k^{1+\frac{1}{M}}\Big(M + \frac{\ln k}{M}\Big)\Big)\Big) \leq R(I) + 16M - 6,$$

причем для ИГ U^*, на котором достигается верхняя оценка, для любого $\tilde{x} \in X_{int2}$ выполняется $T(U^*, \tilde{x}) < 10M + 4\log_2 k + |\{\tilde{y} \in V : \tilde{x}\rho_{int2}\tilde{y}\}|$.

При вариации параметра M мы получаем серию оценок функциональной сложности задачи в различных точках. Так если $M = 2$, то при объеме памяти $Q = 4k^2/3$ мы имеем среднее время поиска (без времени перечисления ответа) $T = 26$. Для любой константы $\varepsilon > 0$ мы можем выбрать параметр $M =]2/\varepsilon[$ так, что при объеме памяти $Q = \underline{O}(k^{1+\varepsilon})$ мы имеем среднее время $T = 16M - 6 = 16]2/\varepsilon[-6 = const$. Ограничение сверху на параметр M, приведенное в теореме, объясняется тем, что функция $Q_1(M) = Mk^{1+2/M}$ имеет точку минимума $M = 2\ln k$, функция $Q_2(M) = Mk^{1+1/M}$ имеет точку минимума $M = \ln k$, функция $Q_3(M) = \frac{\ln k}{M}k^{1+1/M}$ является убывающей по M, причем $Q_3(\ln k) = ek$. Поэтому после точки $M = 2\ln k$ объем $Q(M)$ данного метода заведомо будет возрастать при одновременном возрастании среднего времени поиска $T(M)$. Поэтому при минимальном по порядку для предлагаемого метода объеме $Q = \underline{O}(k\log k)$ мы имеем среднее время $T = \underline{O}(\log k)$, что соответствует по порядку методу Уилларда [22] и Люкера [19], но в силу сложности своих конструкций предлагаемый метод в данном случае является менее предпочтительным, чем метод Уилларда-Люкера.

Нижняя оценка в теореме следует из [27, Теорема 4]. Смысл нижней оценки в том, что время поиска не может быть меньше времени, необходимого на перечисление ответа.

Здесь будет приведено доказательство верхней оценки.

3 Неформальное описание алгоритма

Пусть абсциссы точек библиотеки $V = \{\tilde{y}_1, \ldots, \tilde{y}_k\}$, $\tilde{y}_i = (y_i^1, y_i^2)$, $i = 1, \ldots, k$, упорядочены по возрастанию: $y_1^1 \le y_2^1 \le \cdots \le y_k^1$. Для упрощения изложения будем считать, что все точки y_i^1 различны, поскольку при их совпадении сложность получаемых ИГ только уменьшается.

Пусть $M \in \mathbb{N}$, $M \le k$, и $\alpha_1, \ldots, \alpha_M$, $0 < \alpha_i \le 1$, $i = 1, \ldots, M$ — некоторые рациональные числа такие, что $\sum_{i=1}^M \alpha_i = 1$.

Для любого $q \in \{1, \ldots, M\}$ положим $l_0^q = 0$, $l^q_{k^{\alpha_1+\cdots+\alpha_q}+1} = 1$, причем здесь и везде далее, если некоторое действительное число используется как целое, то подразумевается целая часть этого числа. Множество

$$\mathcal{N}^q = \{l_j^q \mid l_j^q = y^1_{(j-1)k^{1-(\alpha_1+\cdots+\alpha_q)}+1},\ j = 1, \ldots, k^{\alpha_1+\cdots+\alpha_q}\}.$$

назовем q-й сеткой, а интервалы вида $[l_j^q, l_{j+1}^q)$, $j = 1, \ldots, k^{\alpha_1+\cdots+\alpha_q}$, назовем интервалами q-й сетки. Точки q-й сетки делят каждый интервал $(q-1)$-й сетки на k^{α_q} частей. В каждом интервале q-й сетки содержится $k^{1-(\alpha_1+\cdots+\alpha_q)}$ точек из множества $\{y_1^1, \ldots, y_k^1\}$. $l_j^{q-1} = l^q_{(j-1)k^{\alpha_q}+1}$. Очевидно, самая последняя сетка будет иметь вид $\mathcal{N}^M = \{l_j^M = y_j^1,\ j = 1, \ldots, k\}$.

Если $M > 1$ и $q \in \{2, \ldots, M\}$, то обозначим

$$\mathcal{N}_s^q = \{l_j^q \in \mathcal{N}^q \mid l_j^q \in (l_s^{q-1}, l_{s+1}^{q-1})\}, \quad s = 1, \ldots, k^{\alpha_1+\cdots+\alpha_{q-1}}.$$

Под одномерным интервальным поиском понимается задача типа $S_{int} = \langle X_{int}, [0,1), \rho_{int}\rangle$, где ρ_{int} определяется соотношением $(u,v)\rho_{int} y \Leftrightarrow u \le y \le v$.

Поиск по запросу $\tilde{x} = (u_1, v_1, u_2, v_2)$ предлагаемым алгоритмом проводится в M этапов.

1 этап

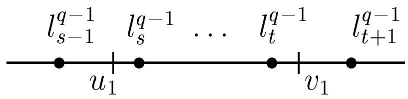

Рис. 1: Случай когда $s \le t$.

1. Среди точек множества $\mathcal{N}^1 \cup \{1\}$ ищется точка, ближайшая справа к u_1. Пусть это l_p^1, $p \in \{1, \dots, k^{\alpha_1} + 1\}$.

2. Среди точек множества $\{l_j^1 \mid j = p-1, \dots, k^{\alpha_1}\}$ ищется точка, ближайшая слева к v_1. Пусть это l_r^1, $r \in \{p-1, \dots, k^{\alpha_1}\}$.

3. Если $p < r$, то для упорядоченных ординат y_j^2 тех точек библиотеки V, абсциссы которых удовлетворяют условию $l_p^1 \le y_j^1 < l_r^1$, мы решаем одномерную задачу интервального поиска, где в качестве запроса берется отрезок $\tilde{x} = [u_2, v_2]$.

4. Если $M = 1$ и $l_r^1 = y_i^1$ для некоторого $i \in \{1, \dots, k\}$, то проверяем принадлежность ординаты y_i^2 отрезку $[u_2, v_2]$ и в зависимости от результата проверки включаем или не включаем точку $\tilde{y}_i = (y_i^1, y_i^2)$ в ответ задачи.

5. Если $M > 1$, то переходим ко второму этапу.

<u>*q этап* $(2 \le q \le M)$</u>

Пусть l_s^{q-1} — точка, ближайшая справа к u_1 на $(q-1)$-й сетке, l_t^{q-1} — точка, ближайшая слева к v_1 на $(q-1)$-й сетке. Эти точки являются левыми концами s-го и t-го интервалов $(q-1)$-й сетки соответственно.

1. Если $s \le t$ (см. рисунок 1), то

   1.1 среди точек множества $\mathcal{N}_{s-1}^q$ ищем точку, ближайшую справа к u_1; обозначим найденную точку ξ';

   1.2 среди точек множества $\mathcal{N}_t^q$ ищем точку, ближайшую слева к v_1; обозначим найденную точку ξ'';

   1.3 для упорядоченных ординат y_j^2 тех точек библиотеки V, абсциссы которых y_j^1 удовлетворяют условию $\xi' \le y_j^1 < l_s^{q-1}$, решаем одномерную задачу интервального поиска с запросом $\tilde{x}' = [u_2, v_2]$;

   1.4 для упорядоченных ординат y_j^2 тех точек библиотеки V, абсциссы которых y_j^1 удовлетворяют условию $l_t^{q-1} \le y_j^1 < \xi''$, решаем одномерную задачу интервального поиска с запросом $\tilde{x}' = [u_2, v_2]$.

2. Если $t = s - 1$ (см. рисунок 2), то

   2.1 среди точек множества $\mathcal{N}_{s-1}^q$ ищем точку, ближайшую справа к u_1; обозначим найденную точку ξ';

   2.2 пусть ξ — точка, предшествующая точке ξ' на q-й сетке; среди точек множества $\mathcal{N}_{s-1}^q$, не меньших ξ, ищем точку, ближайшую слева к v_1; обозначим ее ξ'';

   2.3 если $\xi' < \xi''$, то для упорядоченных ординат y_j^2 тех точек библиотеки V, абсциссы которых y_j^1 удовлетворяют условию $\xi' \le y_j^1 < \xi''$, решаем одномерную задачу интервального поиска, где в качестве запроса берется отрезок $\tilde{x}' = [u_2, v_2]$.

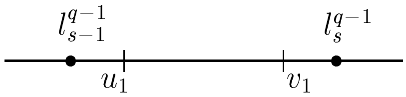

Рис. 2: Случай когда $t = s - 1$.

3. Если $q = M$ и $\xi'' = y_i^1$ для некоторого $i \in \{1, \ldots, k\}$, то проверяем принадлежность ординаты y_i^2 отрезку $[u_2, v_2]$ и в зависимости от результата проверки включаем или не включаем точку $\tilde{y}_i = (y_i^1, y_i^2)$ в ответ задачи.

4. Если $q < M$, то осуществляем переход к следующему этапу.

Для решения одномерной задачи интервального поиска и задачи о нахождении точки, ближайшей к данной (задача о близости) воспользуемся алгоритмами, приведенными в [27]. Дадим их неформальное описание.

Одномерный интервальный поиск.

Пусть $V = \{z_1, \ldots, z_l\} \subset [0, 1)$ — упорядоченное по возрастанию множество, в котором мы должны производить поиск (библиотека), X_{int} — множество запросов, $m \in \mathbb{N}$ — некоторый параметр, значение которого будет установлено позже. Построим для библиотеки V множество номеров $S = \{s_1, \ldots, s_{m-1}\}$, где s_i — номер записи из V такой, что z_{s_i} — ближайшая слева запись к точке i/m, $i = 1, \ldots, m-1$. Если такой записи нет, то $s_i = 0$. Отметим, что это множество строится только однажды.

Теперь поиск по произвольному запросу $\tilde{x} = (u, v)$ производится следующим образом.

Сначала вычисляется длина запроса $\tilde{x}$.

Если она меньше, чем $1/m$, то в множестве V дихотомическим поиском находится ближайшая справа к точке u запись. Далее, начиная с этой записи, просматриваются слева направо все записи из V и сравниваются с правым концом запроса — точкой v до тех пор, пока очередная запись не станет больше v.

Если $v - u \geq 1/m$, то вычисляем номер $j = [n \cdot m] + 1$. Понятно, что точка j/m попадает в интервал $[u, v]$. Теперь, начиная с записи с номером s_j, просматриваем справа налево записи из V и сравниваем с левым концом запроса — точкой u. Как только очередная запись окажется меньше u, мы, начиная с записи с номером $s_j + 1$, просматриваем слева направо записи из V и сравниваем с правым концом запроса — точкой v до тех пор, пока очередная запись не станет больше v.

Поиск точки, ближайшей к данной.

Пусть $V = \{z_1, \ldots, z_l\} \subset [a, b)$ — упорядоченное по возрастанию множество, в котором мы должны производить поиск (библиотека), $X = [a, b)$ — множество запросов, $m \in \mathbb{N}$ — некоторый параметр, значение которого будет установлено позже.

Разобьем множество $X = [a, b)$ на m равных частей $X_1, \ldots, X_m$, где $X_i = [a + (i-1)(b-a)/m, a + i(b-a)/m)$, $i = 1, \ldots, m$.

Пусть $V_i = X_i \cap V$, $i = 1, \ldots, m$.

Теперь для произвольной точки-запроса x поиск точки из V, ближайшей справа (слева) к x будем производить следующим образом.

Определим ту часть отрезка, которой принадлежит запрос x. Ее номер равен $i = [\frac{x-a}{b-a} \cdot m] + 1$, после чего в множестве V_i осуществим обычный дихотомический поиск.

Очевидно, что при $M = 1$ описанный выше многоэтапный алгоритм совпадает с алгоритмом решения двумерной задачи интервального поиска из [27].

4 Построение информационного графа

Прежде, чем приступить к построению ИГ над базовым множеством $\mathcal{F}$, решающего двумерную задачу интервального поиска описанным выше методом, введем вспомогательные блоки, из которых и будем впоследствии строить ИГ.

Пусть $[b, d] \subseteq [0, 1)$ — произвольный интервал.

Пусть $Z = \{z_1, \ldots, z_l\}$ — упорядоченное множество действительных чисел из интервала $[b, d]$. Через $U^1_{Z,b,d}$ обозначим ИГ с одной входной вершиной (корнем) и $(l + 1)$ выходными вершинами (листьями, которые занумерованы слева направо числами $1, \ldots, l + 1$) такой, что функция проводимости между корнем и i-м ($i = 1, \ldots, l + 1$) листом имеет вид:

$$f_i(u_1, v_1, u_2, v_2) = \begin{cases} 1, & \text{если } z_{i-1} < u_1 \leq z_i, \\ 0, & \text{в противном случае,} \end{cases} \tag{15}$$

где $z_0 = b$, $z_{l+1} = d$.

Если считать, что i-му листу ($i = 1, \ldots, l + 1$) соответствует точка z_i, то $U^1_{Z,b,d}$ позволяет находить точку из множества $\{z_1, \ldots, z_l, z_{l+1} = d\}$, ближайшую справа к u_1.

ИГ $U^1_{Z,b,d}$ строится по методу, описанному в [27, параграф 2.3] с помощью функций из $G_1 \cup G_3$. Неформальное описание алгоритма соответствующего ИГ $U^1_{Z,b,d}$ (то есть решающего задачу поиска ближайшей точки) приведено в предыдущем разделе.

Вероятностная мера вероятностного пространства над множеством запросов $u_1 \in [b, d]$ при условии, что $b \leq e \leq v_1 \leq f < 1$, в рассмотренной задаче о близости определяется функцией плотности вероятности, задаваемой соотношением (10). Поскольку справедливо (13), то мы находимся в условиях теоремы 7 из [27]. Следовательно если выбрать в качестве параметра алгоритма (см. предыдущий раздел) $m =]c \cdot l[$, то выполняется

$$Q(U^1_{Z,b,d}) = (2 + c)l + 1, \ T(U^1_{Z,b,d}) \leq 2, \ T(U^1_{Z,b,d}, x) \leq 1 +]\log_2(l + 1)[\tag{16}$$

для любого $x \in [b, d]$.

Пусть $Z = \{z_1, \ldots, z_l\}$ — упорядоченное множество действительных чисел из интервала $[e, f]$. Через $U^2_{Z,e,f}$ обозначим ИГ с одной входной вершиной (корнем) и $(l + 1)$ выходными вершинами (листьями, которые занумерованы слева направо числами $1, \ldots, l + 1$) такой, что функция проводимости между корнем и i-м ($i = 1, \ldots, l + 1$) листом имеет вид:

$$f_i(u_1, v_1, u_2, v_2) = \begin{cases} 1, & \text{если } z_{i-1} \leq v_1 < z_i, \\ 0, & \text{в противном случае,} \end{cases}$$

где $z_0 = b$, $z_{l+1} = e$.

Если считать, что i-му листу ($i = 1, \ldots, l + 1$) соответствует точка z_{i-1}, то $U^2_{Z,e,f}$ позволяет находить точку из множества $\{b = z_0, z_1, \ldots, z_l\}$, ближайшую слева к v_1.

ИГ $U^2_{Z,e,f}$ строится аналогично ИГ $U^1_{Z,e,f}$ с помощью функций из $G_2 \cup G_4$.

Вероятностная мера вероятностного пространства над множеством запросов $v_1 \in [e, f]$ при условии, что $0 \leq b \leq u_1 \leq d < 1$ и $b \leq e$, в рассмотренной задаче о близости определяется функцией плотности вероятности, задаваемой соотношением (11). Поскольку справедливо (14), то мы находимся в условиях теоремы 7 из [27]. Поэтому при $m =]c \cdot l[$ выполняется

$$Q(U^2_{Z,e,f}) = (2+c)l+1,\ T(U^2_{Z,e,f}) \le 2,\ T(U^2_{Z,e,f}, x) \le 1 +]\log_2(l+1)[\tag{17}$$

для любого $x \in [e, f]$.

Пусть $Z = \{z_1, z_2, \ldots, z_l\}$ — упорядоченное множество действительных чисел из интервала $[0, 1)$. Через U^3_Z обозначим ИГ с одной входной вершиной (корнем) и l выходными вершинами, такой, что функция проводимости между корнем и листом с номером i $(i = 1, \ldots, l)$ имеет вид:

$$f_i(u_1, v_1, u_2, v_2) = \begin{cases} 1, & \text{если } u_2 \le z_i \le v_2 \\ 0, & \text{в противном случае} \end{cases}.$$

Если считать, что i-му листу, $i = 1, \ldots, l$, соответствует точка z_i, то U^3_Z позволяет находить все точки из Z, принадлежащие отрезку $[u_2, v_2]$.

ИГ U^3_Z разрешает одномерную задачу интервального поиска для библиотеки $V = Z$ и множества запросов $X = X_{int}$. ИГ U^3_Z будем строить по методу, описанному в [27, параграф 2.4.1], с помощью функций из $G_5 \cup G_6, F_2$. Неформальное описание алгоритма, соответствующего ИГ U^3_Z, (алгоритма решения одномерной задачи интервального поиска) приведено в предыдущем разделе.

Если для некоторых b, d, e, $f \in [0, 1)$ таких, что $b < d \le e < f$ выполняется $b \le u_1 \le d$ и $e \le v_1 \le f$, то вероятностная мера вероятностного пространства над множеством запросов определяется функцией плотности вероятности, задаваемой соотношением (9). Поскольку выполняется (12), то мы находимся в условиях теоремы 8 из [27]. Следовательно если в качестве параметра алгоритма (см. предыдущий раздел) выбрать $m = 2c[\log_2 l]$, то справедливо

$$Q(U^3_Z) \le 4l - 1 + 6c[\log_2 l], \qquad T(U^3_Z) \le 5 + \sum_{z \in Z} \mathbf{P}'(O(z, \rho_{int})), \tag{18}$$

где $\mathbf{P}'$ определяется функцией плотности вероятности (9), и для любого $\tilde{x} \in X_{int}$ выполняется

$$T(U^3_Z, \tilde{x}) \le 1 +]\log_2 l[+ |\{z \in Z : \tilde{x}\rho_{int} z\}|. \tag{19}$$

Приступим к построению информационного графа U^*, решающего двумерную задачу интервального поиска методом, описанным в предыдущем разделе.

Сначала построим ИГ U_q, $q = 1, \ldots, M$.

Все полюса графа U_q, $q = 1, \ldots, M$, разделяются на несколько типов.

1. При $1 \le q < M$ в графе U_q есть полюса ϑ^q_j, $j = 2, \ldots, k^{\alpha_1 + \cdots \alpha_q}$, на которых реализуются функции

$$f^{\vartheta,q}_j(\tilde{x}) = \begin{cases} 1, & \text{если } l^q_{j-1} < u_1 \le l^q_j \le v_1, \\ 0, & \text{в противном случае.} \end{cases} \tag{20}$$

   Из вершин ϑ^q_j в дальнейшем будут выпускаться ИГ, находящие в множестве $\mathcal{N}^{q+1}_{j-1}$ точку, ближайшую справа к u_1.

2. При $1 \le q \le M$ в U_q есть полюса β_j^q, $j = 1, \dots, k^{\alpha_1+\dots+\alpha_q}$, на которых реализуются функции
$$f_j^{\beta,q}(\tilde{x}) = \begin{cases} 1, & \text{если } u_1 \le l_j^q \le v_1 < l_{j+1}^q, \\ 0, & \text{в противном случае.} \end{cases} \tag{21}$$
При $q < M$ из вершин β_j^q в дальнейшем будут выпускаться ИГ, находящие в множестве $\mathcal{N}_j^{q+1}$ точку, ближайшую слева к v_1, а при $q = M$ в этих вершинах будет выполняться проверка принадлежности отрезку $[u_2, v_2]$ ординаты точки библиотеки V с абсциссой l_j^q.

3. При $1 \le q < M$ в U_q есть полюса γ_j^q, $j = 2, \dots, k^{\alpha_1+\dots+\alpha_q}+1$. На них реализуются функции
$$f_j^{\gamma,q}(\tilde{x}) = \begin{cases} 1, & \text{если } l_{j-1}^q < u_1 \le v_1 < l_j^q, \\ 0, & \text{в противном случае.} \end{cases} \tag{22}$$
Из вершин γ_j^q будут выпускаться ИГ, которые находят ближайшую справа к u_1, и ближайшую слева к v_1 точки при условии, что $u_1, v_1 \in (l_{j-1}^q, l_j^q)$.

4. При $1 \le q \le M$ в U_q есть полюса δ_{sij}^q, $s = 1, \dots, k^{\alpha_1+\dots+\alpha_{q-1}}$, $i = 2, \dots, k^{\alpha_q} - 1$, $j = i+1, \dots, k^{\alpha_q}$. При $q = 1$ индексы меняются как $s = 1$, $i = 1, \dots, k^{\alpha_1} - 1$, $j = i+1, \dots, k^{\alpha_1}$. На них реализуются функции
$$f_{sij}^{\delta,q}(\tilde{x}) = \begin{cases} 1, & \text{если } l_{(s-1)k^{\alpha_q}+i-1}^q < u_1 \le l_{(s-1)k^{\alpha_q}+i}^q \\ & \text{и } l_{(s-1)k^{\alpha_q}+j}^q \le v_1 < l_{(s-1)k^{\alpha_q}+j+1}^q, \\ 0, & \text{в противном случае.} \end{cases} \tag{23}$$
Из вершин δ_{sij}^q будут выпускаться ИГ, решающие одномерную задачу интервального поиска для ординат тех точек из библиотеки V, абсциссы которых принадлежат интервалу $[l_{(s-1)k^{\alpha_q}+i}^q, l_{(s-1)k^{\alpha_q}+j}^q)$.

5. При $1 < q \le M$ в U_q есть полюса η_{ij}^q, $i = 2, \dots, k^{\alpha_1+\dots+\alpha_{q-1}}$, $j = 1, \dots, k^{\alpha_q} - 1$, на которых реализуются функции
$$f_{ij}^{\eta,q}(\tilde{x}) = \begin{cases} 1, & \text{если } l_{(i-2)k^{\alpha_q}+j}^q < u_1 \le l_{(i-2)k^{\alpha_q}+j+1}^q \text{ и } v_1 \ge l_i^{q-1}, \\ 0, & \text{в противном случае.} \end{cases} \tag{24}$$
Из вершин η_{ij}^q будут в дальнейшем выпускаться ИГ, решающие одномерную задачу интервального поиска для ординат тех точек библиотеки V, абсциссы которых принадлежат интервалу $[l_{(i-2)k^{\alpha_q}+j+1}^q, l_i^{q-1})$.

6. При $1 < q \le M$ в U_q есть полюса ζ_{ij}^q, $i = 1, \dots, k^{\alpha_1+\dots+\alpha_{q-1}}$, $j = 2, \dots, k^{\alpha_q}$, на которых реализуются функции
$$f_{ij}^{\zeta,q}(\tilde{x}) = \begin{cases} 1, & \text{если } l_{(i-1)k^{\alpha_q}+j}^q \le v_1 < l_{(i-1)k^{\alpha_q}+j+1}^q \text{ и } u_1 \le l_i^{q-1}, \\ 0, & \text{в противном случае.} \end{cases} \tag{25}$$
Из вершин ζ_{ij}^q будут в дальнейшем выпускаться ИГ, решающие одномерную задачу интервального поиска для ординат тех точек библиотеки V, абсциссы которых принадлежат интервалу $[l_i^{q-1}, l_{(i-1)k^{\alpha_q}+j}^q)$.

Строить U_q, $q = 1, \dots, M$, будем индуктивно по q.

Базис индукции. Пусть $q = 1$.

Возьмем ИГ $U^1_{\mathcal{N}^1,0,1}$. Он имеет $(k^{\alpha_1}+1)$ концевых вершин. Обозначим их $\theta_1, \dots, \theta_{k^{\alpha_1}+1}$.

Если $M > 1$, то вершину $\theta_{k^{a_1}+1}$ переобозначим в $\gamma^1_{k^{\alpha_1}+1}$, так как мы попадаем в нее только в том случае, когда

$$l^1_{k^{\alpha_1}} < u_1 \leq v_1 < l^1_{k^{\alpha_1}+1} = 1.$$

Если $M = 1$, то пометим ее символом "$*$". Выход к вершине, помеченной символом "$*$"будет означать, что ответ на запрос $\tilde{x} = (u_1, v_1, u_2, v_2)$ пуст. Эта пометка делается только для наглядности описания, на функционирование ИГ она не влияет.

Из каждой вершины θ_i, $i = 1, \dots, k^{\alpha_1}$, выпустим два ребра, левому припишем 1, правому — 2, а самой вершине θ_i припишем переключатель $g^4_{l^1_i}(\tilde{x}) \in G^4$; конец левого ребра обозначим θ'_i, конец правого — θ''_i. В θ'_i мы попадаем только в том случае, когда

$$l^1_{i-1} < u_1 \leq v_1 < l^1_i, \tag{26}$$

а в θ''_i — когда $l^1_{i-1} < u_1 \leq l^1_i$, $v_1 \geq l^1_i$. Поэтому если $M > 1$, то мы можем переобозначить все θ'_i, в γ^1_i, а θ''_i — в ϑ^1_i, $i = 2, \dots, k^{\alpha_1}$.

Вершину θ'_1 пометим символом "$*$".

Если $M = 1$, то все θ'_i, $i = 1, \dots, k^{\alpha_1}$, пометим символом "$*$".

Каждую вершину θ''_i, $i = 1, \dots, k^{\alpha_1} - 1$, (включая переобозначенные в ϑ^1_i) отождествим с корнем ИГ $U^2_{Z^{1,i}, l^1_i, 1}$, где

$$Z^{1,i} = \{l^1_j \in \mathcal{N}^1 \mid j \geq i + 1\}.$$

Информационный граф $U^2_{Z^{1,i}, l^1_i, 1}$ имеет $(k^{\alpha_1} - i + 1)$ концевых вершин, которые обозначим θ_{ij}, $i = 1, \dots, k^{\alpha_1} - 1$, $j = i, \dots, k^{\alpha_1}$. Понятно, что в вершины θ_{ij} мы проходим только в том случае, когда

$$l^1_{i-1} < u_1 \leq l^1_i, \qquad l^1_j \leq v_1 < l^1_{j+1}. \tag{27}$$

Следовательно, можно переобозначить вершины θ_{ij}, $i = 1, \dots, k^{\alpha_1} - 1$, $j = i + 1, \dots, k^{\alpha_1}$, в δ^1_{1ij}.

Введем новые вершины β^1_j, $j = 1, \dots, k^{\alpha_1}$, и из каждой θ_{ij}, $i = 1, \dots, j$, (включая θ_{ij}, переобозначенные в δ^1_{1ij}), проведем ребро в вершину β^1_j, которому припишем предикат, тождественно равный 1. Очевидно, что

$$\varphi_{\beta^1_j} = \begin{cases} 1, & \text{если } l^1_j \leq v_1 < l^1_{j+1} \text{ и } u_1 \leq l^1_j, \\ 0, & \text{в противном случае,} \end{cases}$$

то есть $\varphi_{\beta^1_j} = f^{\beta,1}_j(\tilde{x})$.

Построенный граф обозначим U_1. Он изображен на рисунке 3 для случая $M > 1$.

Индуктивный переход. Пусть построен ИГ U_{q-1}. Будем строить U_q, достраивая U_{q-1} следующим образом.

Из каждой вершины ϑ^{q-1}_i, $i = 2, \dots, k^{\alpha_1 + \dots + \alpha_{q-1}}$, выпустим информационный граф $U^1_{\mathcal{N}^q_{i-1}, l^{q-1}_{i-1}, l^{q-1}_i}$. Этот граф имеет k^{α_q} листьев, которые обозначим τ'_{qij}, $j = 1, \dots, k^{\alpha_q}$.

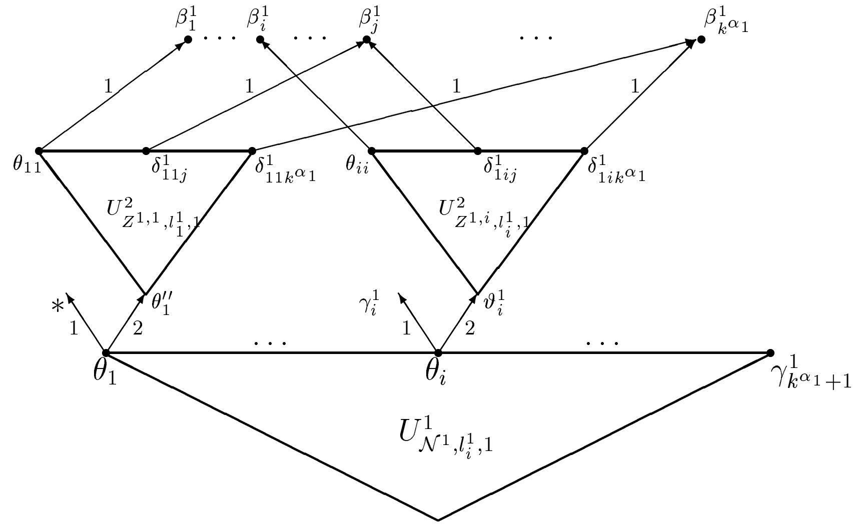

Рис. 3: Информационный граф U_1, $M > 1$.

В вершину τ'_{qij} мы попадаем только в том случае, когда

$$l^q_{(i-2)k^{\alpha_q}+j} < u_1 \le l^q_{(i-2)k^{\alpha_q}+j+1}, \quad v_1 \ge l^{q-1}_i, \tag{28}$$

поскольку $\mathcal{N}^q_{i-1} = (l^{q-1}_{i-1}, l^{q-1}_i) = (l^q_{(i-2)k^{\alpha_q}+1}, l^q_{(i-1)k^{\alpha_q}+1})$. Получаем, что $\varphi_{\tau'_{qij}} = f^{\eta,q}_{ij}(\tilde{x})$. Следовательно, можно переобозначить τ'_{qij} с индексами $i = 2, \dots, k^{\alpha_1+\dots+\alpha_{q-1}}$, $j = 1, \dots, k^{\alpha_q} - 1$, в η^q_{ij}.

Из каждой вершины β^{q-1}_i, $i = 1, \dots, k^{\alpha_1+\dots+\alpha_{q-1}}$, выпустим информационный граф $U^2_{\mathcal{N}^q_i, l^{q-1}_i, l^{q-1}_{i+1}}$. Этот граф имеет k^{α_q} листьев, которые обозначим τ''_{qij}, $j = 1, \dots, k^{\alpha_q}$. В вершину τ''_{qij} мы попадаем только в том случае, когда

$$l^q_{(i-1)k^{\alpha_q}+j} \le v_1 < l^q_{(i-1)k^{\alpha_q}+j+1}, \quad u_1 \le l^{q-1}_i. \tag{29}$$

Таким образом, $\varphi_{\tau''_{qij}} = f^{\zeta,q}_{ij}(\tilde{x})$. Следовательно, можно переобозначить τ''_{qij} с индексами $i = 1, \dots, k^{\alpha_1+\dots+\alpha_{q-1}}$, $j = 2, \dots, k^{\alpha_q}$, в ζ^q_{ij}.

Рассмотрим вершины γ^{q-1}_i, $i = 2, \dots, k^{\alpha_1+\dots+\alpha_{q-1}} + 1$. Из каждой такой вершины выпустим информационный граф $U^1_{\mathcal{N}^q_{i-1}, l^{q-1}_{i-1}, l^{q-1}_i}$. Этот граф имеет k^{α_q} листьев. Обозначим их θ_{qij}, $j = 1, \dots, k^{\alpha_q}$. Проводимость в эти вершины будет равна

$$\varphi_{\theta_{qij}} = \begin{cases} 1, & \text{если } l^q_{(i-2)k^{\alpha_q}+j} < u_1 \le l^q_{(i-2)k^{\alpha_q}+j+1},\ v_1 < l^q_{(i-1)k^{\alpha_q}+1}, \\ 0, & \text{в противном случае.} \end{cases} \tag{30}$$

Из каждой вершины θ_{qij}, $i = 2, \dots, k^{\alpha_1+\dots+\alpha_{q-1}} + 1$, $j = 1, \dots, k^{\alpha_q} - 1$, выпустим два ребра, левому припишем 1, правому — 2, а самой θ_{qij} припишем переключатель $g^4_{l^q_{(i-2)k^{\alpha_q}+j+1}}(\tilde{x})$. Конец левого ребра, выходящего из θ_{qij}, обозначим θ'_{qij}, конец правого — θ''_{qij}. Переобозначим все $\theta_{q,i,k^{\alpha_q}}$, $i = 2, \dots, k^{\alpha_1+\dots+\alpha_{q-1}} + 1$, в $\theta'_{q,i,k^{\alpha_q}}$.

Если $q < M$, то переобозначим все θ'_{qij}, $i = 2, \dots, k^{\alpha_1+\dots+\alpha_{q-1}} + 1$, $j = 1, \dots, k^{\alpha_q}$, в $\gamma^q_{(i-2)k^{\alpha_q}+j+1}$, так как в θ'_{qij} мы можем попасть только в том случае, когда $l^q_{(i-2)k^{\alpha_q}+j} < u_1 \le v_1 < l^q_{(i-2)k^{\alpha_q}+j+1}$, то есть $\varphi_{\theta'_{qij}} = f^{\gamma,q}_{(i-2)k^{\alpha_q}+j+1}(\tilde{x})$.

При изменении i от 2 до $k^{\alpha_1+\dots+\alpha_{q-1}}+1$, j от 1 до k^{α_q}, $(i-2)k^{\alpha_q}+j+1$ изменяется от 2 до $k^{\alpha_1+\dots+\alpha_q}+1$, поэтому мы получим все вершины типа γ^q_s, $s = 2, \dots, k^{\alpha_1+\dots+\alpha_q} + 1$.

Если $q = M$, то все θ'_{qij}, $i = 2, \dots, k^{\alpha_1+\dots+\alpha_{q-1}} + 1$, $j = 1, \dots, k^{\alpha_q}$, пометим символом “$*$”.

В вершины θ''_{qij} мы попадаем при условии, что

$$l^q_{(i-2)k^{\alpha_q}+j} < u_1 \le l^q_{(i-2)k^{\alpha_q}+j+1} \le v_1 < l^q_{(i-1)k^{\alpha_q}+1}. \tag{31}$$

Из каждой вершины θ''_{qij}, $i = 2, \ldots, k^{\alpha_1+\cdots+\alpha_{q-1}} + 1$, $j = 1, \ldots, k^{\alpha_q} - 2$, выпустим ИГ $U^2_{Z^{q,j}_{i-1}, l^q_{(i-2)k^{\alpha_q}+j+1}, l^{q-1}_i}$, где

$$Z^{q,j}_{i-1} = \{l^q_{(i-2)k^{\alpha_q}+j+2}, \ldots, l^q_{(i-1)k^{\alpha_q}}\}.$$

Этот информационный граф имеет $(k^{\alpha_q} - j)$ листьев, которые обозначим μ^q_{ijt}, $i = 2, \ldots, k^{\alpha_1+\cdots+\alpha_{q-1}} + 1$, $j = 1, \ldots, k^{\alpha_q} - 2$, $t = j+1, \ldots, k^{\alpha_q}$. Вершины $\theta''_{q,i,k^{\alpha_q}-1}$ переобозначим в $\mu^q_{i,k^{\alpha_q}-1,k^{\alpha_q}}$

В вершину μ^q_{ijt} мы проходим только в том случае, когда

$$l^q_{(i-2)k^{\alpha_q}+j} < u_1 \le l^q_{(i-2)k^{\alpha_q}+j+1}, \quad l^q_{(i-2)k^{\alpha_q}+t} \le v_1 < l^q_{(i-2)k^{\alpha_q}+t+1}. \tag{32}$$

Следовательно, $\varphi_{\mu^q_{ijt}} = f^{\delta,q}_{i-1,j+1,t}(\tilde{x})$, и поэтому можно переобозначить вершины μ^q_{ijt}, $i = 2, \ldots, k^{\alpha_1+\cdots+\alpha_{q-1}} + 1$, $j = 1, \ldots, k^{\alpha_q} - 2$, $t = j+2, \ldots, k^{\alpha_q}$, в $\delta^q_{i-1,j+1,t}$.

Так как $(i-1)$ изменяется от 1 до $k^{\alpha_1+\cdots+\alpha_{q-1}}$, $(j+1)$ изменяется от 2 до $k^{\alpha_q} - 1$, t изменяется от $(j+1)+1$ до k^{α_q}, мы получаем все нужные вершины типа δ^q_{spr}, $s = 1, \ldots, k^{\alpha_1+\cdots+\alpha_{q-1}}$, $p = 2, \ldots, k^{\alpha_q} - 1$, $r = p+1, \ldots, k^{\alpha_q}$.

Если $q < M$, то введем новые вершины ϑ^q_t, $t = 2, \ldots, k^{\alpha_1+\cdots+\alpha_q}$. Из каждой вершины θ''_{qij}, $i = 2, \ldots, k^{\alpha_1+\cdots+\alpha_{q-1}} + 1$, $j = 1, \ldots, k^{\alpha_q} - 1$, и каждой вершины τ'_{qij}, $i = 2, \ldots, k^{\alpha_1+\cdots+\alpha_{q-1}}$, $j = 1, \ldots, k^{\alpha_q}$ (включая переобозначенные в η^q_{ij}), проведем ребро в вершину $\vartheta^q_{(i-2)k^{\alpha_q}+j+1}$, которому припишем предикат, тождественно равный единице. Поскольку проводимость в вершины θ''_{qij} определяется соотношением (31), а в вершины τ'_{qij} — формулой (24), то на $\vartheta^q_{(i-2)k^{\alpha_q}+j+1}$ мы имеем нужную нам функцию, так как

$$\varphi_{\vartheta^q_{(i-2)k^{\alpha_q}+j+1}} = \begin{cases} 1, & \text{если } l^q_{(i-2)k^{\alpha_q}+j} < u_1 \le l^q_{(i-2)k^{\alpha_q}+j+1} \le v_1, \\ 0, & \text{в противном случае.} \end{cases}$$

При изменении i от 2 до $k^{\alpha_1+\cdots+\alpha_{q-1}}+1$, j от 1 до k^{α_q}, $t = (i-2)k^{\alpha_q}+j+1$ изменяется от 2 до $k^{\alpha_1+\cdots+\alpha_q}$, то есть в нужных нам пределах. Здесь индекс $k^{\alpha_1+\cdots+\alpha_q}+1$ не появляется поскольку при $i = k^{\alpha_1+\cdots+\alpha_{q-1}}+1$ j меняется от 1 до $k^{\alpha_q} - 1$.

Введем новые вершины β^q_s, $s = 1, \ldots, k^{\alpha_1+\cdots+\alpha_q}$. Из каждой вершины μ^q_{ijt}, $i = 2, \ldots, k^{\alpha_1+\cdots+\alpha_{q-1}} + 1$, $j = 1, \ldots, k^{\alpha_q} - 1$, $t = j+1, \ldots, k^{\alpha_q}$ (включая вершины, переобозначенные в $\delta^q_{i-1,j+1,t}$), выпустим ребро в вершину $\beta^q_{(i-2)k^{\alpha_q}+t}$, и из каждой вершины τ''_{qij}, $i = 1, \ldots, k^{\alpha_1+\cdots+\alpha_{q-1}}$, $j = 1, \ldots, k^{\alpha_q}$ (включая переобозначенные в ζ^q_{ij}), выпустим ребро в вершину $\beta^q_{(i-1)k^{\alpha_q}+j}$, ребрам припишем предикаты, тождественно равные единице.

Учитывая соотношения (29), (32), легко заметить, что функция фильтра на введенных вершинах β^q_s совпадает с нужной нам функцией, задаваемой соотношением (21), и s изменяется в пределах от 1 до $k^{\alpha_1+\cdots+\alpha_q}$.

Тем самым построение ИГ U_q завершено.

Пусть мы построили информационный граф U_M. Теперь, чтобы завершить построение окончательного ИГ U^*, решающего задачу нашим методом, нам надо добавить фрагменты, решающие одномерные задачи интервального поиска по вторым координатам (u_2, v_2).

Если q, s, i, j такие натуральные числа, что $q \in \{1, \ldots, M\}$, $s \in \{1, \ldots, k^{\alpha_1+\cdots+\alpha_{q-1}}\}$, $i \in \{1, \ldots, k^{a_q}\}$, $j \in \{i+1, \ldots, k^{a_q}+1\}$, то обозначим

$$A^q_{s,i,j} = \{y^2_p : (y^1_p, y^2_p) \in V, l^q_{(s-1)k^{\alpha_q}+i} \leq y^1_p < l^q_{(s-1)k^{\alpha_q}+j}\}, \tag{33}$$

то есть $A^q_{s,i,j}$ состоит из ординат точек библиотеки, абсциссы которых принадлежат интервалу $[l^q_{(s-1)k^{\alpha_q}+i}, l^q_{(s-1)k^{\alpha_q}+j})$.

ИГ U_M содержит полюса δ^q_{sij}, $q = 1, \ldots, M$, $s = 1, \ldots, k^{\alpha_1+\cdots+\alpha_{q-1}}$, $i = 2, \ldots, k^{\alpha_q}-1$, $j = i+1, \ldots, k^{\alpha_q}$. При $q = 1$ индексы меняются как $s = 1$, $i = 1, \ldots, k^{\alpha_1} - 1$, $j = i+1, \ldots, k^{\alpha_1}$. В эти полюса мы попадаем при условии, что

$$l^q_{(s-1)k^{\alpha_q}+i-1} < u_1 \leq l^q_{(s-1)k^{\alpha_q}+i},\ l^q_{(s-1)k^{\alpha_q}+j} \leq v_1 < l^q_{(s-1)k^{\alpha_q}+j+1}.$$

Из каждого полюса δ^q_{sij} выпустим ИГ $U^3_{A^q_{s,i,j}}$, который решает одномерную задачу интервального поиска для запроса (u_2, v_2) среди точек из $A^q_{s,i,j}$. Листья ИГ $U^3_{A^q_{s,i,j}}$, которым соответствуют точки $y^2_p \in A^q_{s,i,j}$, мы объявим листьями окончательного ИГ U^* и припишем им записи $(y^1_p, y^2_p) \in V$.

ИГ U_M содержит полюса η^q_{ij}, $q = 2, \ldots, M$, $i = 2, \ldots, k^{\alpha_1+\cdots+\alpha_{q-1}}$, $j = 1, \ldots, k^{\alpha_q}-1$, в которые мы попадаем при условии, что

$$l^q_{(i-2)k^{\alpha_q}+j} < u_1 \leq l^q_{(i-2)k^{\alpha_q}+j+1},\ v_1 \geq l^q_{(i-1)k^{\alpha_q}+1}.$$

Из каждого полюса η^q_{ij} выпустим ИГ $U^3_{A^q_{i-1,j+1,k^{\alpha_q}+1}}$. Листья ИГ $U^3_{A^q_{i-1,j+1,k^{\alpha_q}+1}}$, которым соответствуют точки $y^2_p \in A^q_{i-1,j+1,k^{\alpha_q}+1}$, мы объявим листьями окончательного ИГ U^* и припишем им записи $(y^1_p, y^2_p) \in V$.

ИГ U_M содержит полюса ζ^q_{ij}, $q = 2, \ldots, M$, $i = 1, \ldots, k^{\alpha_1+\cdots+\alpha_{q-1}}$, $j = 2, \ldots, k^{\alpha_q}$, в которые мы попадаем при условии, что

$$u_1 \leq l^q_{(i-1)k^{\alpha_q}+1},\ l^q_{(i-1)k^{\alpha_q}+j} \leq v_1 < l^q_{(i-1)k^{\alpha_q}+j+1}.$$

Из каждого полюса ζ^q_{ij} выпустим ИГ $U^3_{A^q_{i,1,j}}$. Листья ИГ $U^3_{A^q_{i,1,j}}$, которым соответствуют точки $y^2_p \in A^q_{i,1,j}$, мы объявим листьями окончательного ИГ U^* и припишем им записи $(y^1_p, y^2_p) \in V$.

ИГ U_M содержит полюса β^M_j, $j = 1, \ldots, k^{\alpha_1+\cdots+\alpha_M}$, в которые мы попадаем при условии, что

$$l^M_{j-1} < u_1 \leq l^M_j \leq v_1 < l^M_{j+1}.$$

Из каждого полюса β^M_j выпустим одно ребро, которому припишем предикат $f_{y^2_j} \in F_2$. Конец ребра объявим листом и припишем ему запись $(y^1_j, y^2_j) \in V$.

Полученный окончательный ИГ обозначим U^*. Он соответствует алгоритму решения двумерной задачи интервального поиска, описанному в предыдущем разделе.

5 Допустимость информационного графа

Покажем, что ИГ U^* решает ЗИП $I = \langle X_{int2}, V, \rho_{int2}\rangle$.

Возьмем произвольный запрос $\tilde{x} = (u_1, v_1, u_2, v_2)$. Возможны следующие два случая.

1) Для любого $q \in \{1, \dots, M\}$ существует такой номер $r = r(q) \in \{1, \dots, k^{\alpha_1 + \dots + \alpha_q} + 1\}$, что $l^q_{r-1} < u_1 \le v_1 < l^q_r$. Это означает, что $\{\tilde{y} \in V : \tilde{x}\rho_{int2}\tilde{y}\} = \emptyset$. Поскольку путь к листьям ИГ U^* лежит только через вершины типов δ^q_{sij}, η^q_{ij}, ζ^q_{ij}, β^M_j, а проводимость в эти вершины есть только в случаях когда хотя бы одна точка сетки попадает в интервал $[u_1, v_1]$, то проводимость в эти вершины на запросе $\tilde{x}$ равна нулю, и ответ ИГ U^* на запрос $\tilde{x}$ пуст.

2) Существуют такие номера $q \in \{1, \dots, M\}$, и $r = r(q) \in \{1, \dots, k^{\alpha_1 + \dots + \alpha_q}\}$, что $u_1 \le l^q_r \le v_1$. Пусть q_0 минимальный из номеров q, обладающих таким свойством, то есть либо $q_0 = 1$, либо для любого $q < q_0$ существует такой номер $r = r(q) \in \{1, \dots, k^{\alpha_1 + \dots + \alpha_q} + 1\}$, что $l^q_{r-1} < u_1 \le v_1 < l^q_r$. Возможны два подслучая.

2.1) Для любого $q \in \{q_0, \dots, M\}$ существует такой номер $r = r(q) \in \{1, \dots, k^{\alpha_1 + \dots + \alpha_q}\}$, что $l^q_{r-1} < u_1 \le l^q_r \le v_1 < l^q_{r+1}$. Это означает, что множество $\{\tilde{y} \in V : \tilde{x}\rho_{int2}\tilde{y}\}$ содержит не более одной записи, причем единственным кандидатом на принадлежность этому множеству является запись $(y^1_{r(M)}, y^2_{r(M)})$. Если $u_2 \le y^2_{r(M)} \le v_2$, то $(y^1_{r(M)}, y^2_{r(M)}) \in \{\tilde{y} \in V : \tilde{x}\rho_{int2}\tilde{y}\}$, иначе $\{\tilde{y} \in V : \tilde{x}\rho_{int2}\tilde{y}\} = \emptyset$. Непосредственно проверяется, что среди вершин типа β^M_j только проводимость в вершину $\beta^M_{r(M)}$ будет ненулевой, тогда как проводимость всех вершин типа δ^q_{sij}, η^q_{ij}, ζ^q_{ij} будет равна нулю, так как мы проходим в эти вершины только тогда, когда по крайней мере две точки сетки попадают в интервал $[u_1, v_1]$. Из вершины $\beta^M_{r(M)}$ исходит ребро с предикатом $f_{y^2_{r(M)}}$, что соответствует проверке принадлежности интервалу $[u_1, v_1]$ точки $y^2_{r(M)}$. Таким образом $\mathcal{J}_{U^*}(\tilde{x}) = \{\tilde{y} \in V : \tilde{x}\rho_{int2}\tilde{y}\}$.

2.2) Существуют такие номера $q \in \{q_0, \dots, M\}$, и $r = r(q) \in \{1, \dots, k^{\alpha_1 + \dots + \alpha_q} - 1\}$, что $u_1 \le l^q_r \le l^q_{r+1} \le v_1$. Пусть q_1 минимальный из номеров q, обладающих таким свойством, то есть либо $q_1 = q_0$, либо для любого q такого, что $q_0 \le q < q_1$, существует такой номер $r = r(q) \in \{1, \dots, k^{\alpha_1 + \dots + \alpha_q}\}$, что $l^q_{r-1} < u_1 \le l^q_r \le v_1 < l^q_{r+1}$. Для каждого $q \in \{1, \dots, M\}$ определим такие номера $r'(q)$ и $r''(q)$, что $l^q_{r'(q)-1} < u_1 \le l^q_{r'(q)}$ и $l^q_{r''(q)} \le v_1 < l^q_{r''(q)+1}$. Понятно, что если $q_0 > 1$, то для любого $q \in \{1, \dots, q_0 - 1\}$ справедливо $r'(q) = r''(q) - 1$, и если $q_0 < q_1$, то для любого $q \in \{q_0, \dots, q_1 - 1\}$ справедливо $r'(q) = r''(q)$, и для любого $q \in \{q_1, \dots, M\}$ справедливо $r'(q) < r''(q)$. Отметим, что если $q_1 > 1$, то для $q < q_1$ согласно сделанному выше замечанию проводимость всех вершин типа δ^q_{sij}, η^q_{ij}, ζ^q_{ij} будет равна нулю.

Заметим, что для того, чтобы проводимость в вершины типа δ^q_{sij} была равна 1, необходимо, чтобы либо $q = 1$, либо u_1 и v_1 принадлежали одному интервалу $(q-1)$-й сетки. Поэтому, если $q_0 < q_1$, то проводимость во все вершины типа δ^q_{sij}, будет равна 0. Если же $q_0 = q_1$, то среди вершин типа δ^q_{sij} только проводимость вершины $\delta^{q_1}_{s',i',j'}$ будет равна 1, где $s' = 1$, если $q_1 = 1$, и $s' = r''(q_1 - 1)$, если $q_1 > 1$, $i' = r'(q) - (s' - 1)k^{\alpha_q}$, $j' = r''(q) - (s' - 1)k^{\alpha_q}$. Выход к вершине $\delta^q_{s',i',j'}$ означает, что будет решаться одномерная задача интервального поиска для запроса (u_2, v_2) среди точек из множества $A^q_{s',i',j'}$, определяемого соотношением (33).

Легко видеть, что если $q_0 = q_1$, то проводимость всех вершин типа $\eta^{q_1}_{ij}$, $\zeta^{q_1}_{ij}$ будет равна нулю.

Для каждого $q \in \{\max(q_0 + 1, q_1), \dots, M\}$ среди вершин типа η^q_{ij}, проводимость только одной вершины $\eta^q_{i''(q),j''(q)}$ может быть равна 1, и это случится при условии, что $r'(q-1) > 1$ и $r'(q) < (r'(q-1) - 1)k^{\alpha_q} + 1$, при этом $i''(q) = r'(q-1)$ и $j''(q) = r'(q) - (r'(q-1) - 2)k^{\alpha_q} - 1$. Выход к вершине $\eta^q_{i''(q),j''(q)}$ означает, что будет решаться одномерная задача интервального поиска для запроса (u_2, v_2) среди точек из множества $A^q_{i''(q)-1,j''(q)+1,k^{\alpha_q}+1}$.

Для каждого $q \in \{\max(q_0+1, q_1), \ldots, M\}$ среди вершин типа ζ^q_{ij}, проводимость только одной вершины $\zeta^q_{i'''(q),j'''(q)}$ может быть равна 1, и это случится при условии, что $r''(q) > (r''(q-1)-1)k^{\alpha_q}+1$, при этом $i'''(q) = r''(q-1)$ и $j'''(q) = r''(q)-(r''(q-1)-1)k^{\alpha_q}$. Выход к вершине $\zeta^q_{i'''(q),j'''(q)}$ означает, что будет решаться одномерная задача интервального поиска для запроса (u_2, v_2) среди точек из множества $A^q_{i'''(q),1,j'''(q)}$.

И наконец, среди вершин типа β^M_j, проводимость только вершины $\beta^M_{r''(M)}$ будет равна 1. И в этой вершине делается проверка на принадлежность отрезку $[u_2, v_2]$ точки $y^2_{r''(M)}$.

Тем самым, легко видеть, что все множество записей из V, абсциссы которых принадлежат отрезку $[u_1, v_1]$, разделилось на непересекающиеся части, и для каждой части решается одномерная задача интервального поиска по вторым координатам, то есть устанавливается принадлежность ординат записей отрезку $[u_1, v_2]$. А это означает, что ответ ИГ U^* на запрос $\tilde{x}$ совпадает с множеством записей из V, удовлетворяющих запросу $\tilde{x}$.

Таким образом, в силу произвольности $\tilde{x}$ мы доказали допустимость ИГ U^*.

6 Объем информационного графа

Для вычисления объема информационного графа U^* вычислим сначала объем ИГ U_1.

ИГ U_1 содержит ИГ $U^1_{\mathcal{N}^1,0,1}$, объем которого согласно (16) равен

$$Q(U^1_{\mathcal{N}^1,0,1}) = (2+c)k^{\alpha_1} + 1.$$

Из каждой вершины τ_i $(i = 1, \ldots, k^{\alpha_1})$ исходит два ребра, что суммарно дает $2k^{\alpha_1}$ ребер.

Из каждой вершины τ''_i $(i = 1, \ldots, k^{\alpha_1} - 1)$ исходит ИГ $U^2_{Z^{1,i},l^1_i,1}$, объем которого согласно (17) равен $Q(U^2_{Z^{1,i},l^1_i,1}) = (2+c)(k^{\alpha_1} - i) + 1$.

В каждую вершину β^1_j $(j = 1, \ldots, k^{\alpha_1})$ из каждой вершины τ_{ij}, $i = 1, \ldots, j$, ведет ребро, то есть суммарное число ребер, ведущих в вершины β^1_j, равно

$$\sum_{j=1}^{k^{\alpha_1}} j = \frac{k^{\alpha_1}(k^{\alpha_1}+1)}{2}.$$

Таким образом,

$$\begin{aligned} Q(U_1) &= (2+c)k^{\alpha_1} + 1 + 2k^{\alpha_1} + \sum_{i=1}^{k^{\alpha_1}-1} \Big((2+c)(k^{\alpha_1} - i) + 1\Big) + \\ &+ \frac{k^{\alpha_1}(k^{\alpha_1}+1)}{2} = \frac{3+c}{2}k^{2\alpha_1} + \frac{9+c}{2}k^{\alpha_1}. \end{aligned}$$

Подсчитаем объем ИГ U_q при $2 \le q \le M$.

ИГ U_q содержит в себе ИГ U_{q-1}.

Из каждой вершины ϑ^{q-1}_i $(i = 2, \ldots, k^{\alpha_1+\cdots+\alpha_{q-1}})$ исходит ИГ $U^1_{\mathcal{N}^q_{i-1},l^{q-1}_{i-1},l^{q-1}_i}$, объем которого согласно (16) равен $(2+c)(k^{\alpha_q} - 1) + 1$.

Из каждой вершины β^{q-1}_i, $(i = 1, \ldots, k^{\alpha_1+\cdots+\alpha_{q-1}})$ исходит ИГ $U^2_{\mathcal{N}^q_i,l^{q-1}_i,l^{q-1}_{i+1}}$, объем которого согласно (17) равен $(2+c)(k^{\alpha_q} - 1) + 1$.

Из каждой вершины γ_i^{q-1}, $(i = 2, \dots, k^{\alpha_1+\dots+\alpha_{q-1}}+1)$ исходит ИГ $U^1_{\mathcal{N}^q_{i-1}, l^{q-1}_{i-1}, l^{q-1}_i}$. Его объем равен $(2+c)(k^{\alpha_q}-1)+1$. Из каждой вершины θ_{qij} $(j = 1, \dots, k^{\alpha_q}-1)$ исходит по 2 ребра. Из каждой вершины θ''_{qij} $(j = 1, \dots, k^{\alpha_q}-2)$ исходит ИГ $U^2_{Z^{q,j}_{i-1}, l^q_{(i-2)k^{\alpha_q}+j+1}, l^{q-1}_i}$, объем которого равен $(2+c)(k^{\alpha_q}-j-1)+1$. Таким образом, объем ИГ, растущего из вершины γ_i^{q-1}, равен

$$(2+c)(k^{\alpha_q}-1)+1+2(k^{\alpha_q}-1)+\sum_{j=1}^{k^{\alpha_q}-2}\Big((2+c)(k^{\alpha_q}-j-1)+1\Big) =$$

$$= \frac{2+c}{2}k^{2\alpha_q}+\frac{4-c}{2}k^{\alpha_q}-3.$$

Если $q < M$, то в вершины ϑ^q_t $(t = 2, \dots, k^{\alpha_1+\dots+\alpha_q})$ из каждой вершины θ''_{qij}, $i = 2, \dots, k^{\alpha_1+\dots+\alpha_{q-1}}+1$, $j = 1, \dots, k^{\alpha_q}-1$, и каждой вершины τ'_{qij}, $i = 2, \dots, k^{\alpha_1+\dots+\alpha_{q-1}}$, $j = 1, \dots, k^{\alpha_q}$, ведет по одному ребру. Следовательно суммарное число ребер, ведущих в вершины ϑ^q_t, равно

$$k^{\alpha_1+\dots+\alpha_{q-1}}(k^{\alpha_q}-1)+(k^{\alpha_1+\dots+\alpha_{q-1}}-1)k^{\alpha_q} = 2k^{\alpha_1+\dots+\alpha_q}-k^{\alpha_1+\dots+\alpha_{q-1}}-k^{\alpha_q}.$$

В вершины β^q_s $(s = 1, \dots, k^{\alpha_1+\dots+\alpha_q})$ из каждой вершины μ^q_{ijt}, $i = 2, \dots, k^{\alpha_1+\dots+\alpha_{q-1}}+1$, $j = 1, \dots, k^{\alpha_q}-1$, $t = j+1, \dots, k^{\alpha_q}$, и каждой вершины τ''_{qij}, $i = 1, \dots, k^{\alpha_1+\dots+\alpha_{q-1}}$, $j = 1, \dots, k^{\alpha_q}$, ведет по одному ребру. Следовательно суммарное число ребер, ведущих в вершины β^q_s, равно

$$k^{\alpha_1+\dots+\alpha_{q-1}}\sum_{j=1}^{k^{\alpha_q}-1}(k^{\alpha_q}-j)+k^{\alpha_1+\dots+\alpha_{q-1}}\cdot k^{\alpha_q} = k^{\alpha_1+\dots+\alpha_q}\frac{k^{\alpha_q}+1}{2}.$$

Тем самым при $2 \le q < M$

$$\begin{aligned}
Q(U_q) &= Q(U_{q-1})+\Big((2+c)(k^{\alpha_q}-1)+1\Big)(k^{\alpha_1+\dots+\alpha_{q-1}}-1)+\\
&\quad+\Big((2+c)(k^{\alpha_q}-1)+1\Big)k^{\alpha_1+\dots+\alpha_{q-1}}+\\
&\quad+\Big(\frac{2+c}{2}k^{2\alpha_q}+\frac{4-c}{2}k^{\alpha_q}-3\Big)k^{\alpha_1+\dots+\alpha_{q-1}}+\\
&\quad+2k^{\alpha_1+\dots+\alpha_q}-k^{\alpha_1+\dots+\alpha_{q-1}}-k^{\alpha_q}+k^{\alpha_1+\dots+\alpha_q}\frac{k^{\alpha_q}+1}{2} =\\
&= U(Q_{q-1})+\frac{3+c}{2}k^{\alpha_1+\dots+\alpha_{q-1}+2\alpha_q}+\underline{O}(k^{\alpha_1+\dots+\alpha_q}).
\end{aligned}$$

Учитывая, что суммарное число ребер, ведущих в вершины ϑ^q_t, равно $\underline{O}(k^{\alpha_1+\dots+\alpha_q})$, и есть еще другие слагаемые имеющие тот же порядок, получаем, что

$$Q(U_M) = \frac{3+c}{2}\sum_{q=1}^{M}k^{\alpha_1+\dots+\alpha_{q-1}+2\alpha_q}+\underline{O}(k). \tag{34}$$

Вычислим объем информационного графа U^*.

Из (33) легко видеть, что

$$|A^q_{s,i,j}| = (j-i)k^{1-(\alpha_1+\dots+\alpha_q)}. \tag{35}$$

Из каждой вершины δ^q_{sij}, $q=1,\dots,M$, $s=1,\dots,k^{\alpha_1+\dots+\alpha_{q-1}}$, $i=2,\dots,k^{\alpha_q}-1$, $j=i+1,\dots,k^{\alpha_q}$ (при $q=1$ индексы меняются как $s=1$, $i=1,\dots,k^{\alpha_1}-1$, $j=i+1,\dots,k^{\alpha_1}$) исходит ИГ $U^3_{A^q_{s,i,j}}$, объем которого согласно (18) и (35) равен

$$Q(U^3_{A^q_{s,i,j}}) \le 4(j-i)k^{1-(\alpha_1+\dots+\alpha_q)} - 1 + 6c[\log_2((j-i)k^{1-(\alpha_1+\dots+\alpha_q)})].$$

Суммируя по всем q, s, i, j, получаем, что объем ИГ, растущих из вершин δ^q_{sij}, (обозначим его Q_1) не превышает

$$\begin{aligned} Q_1 &< \sum_{q=1}^{M}\sum_{s=1}^{k^{\alpha_1+\dots+\alpha_{q-1}}}\sum_{i=2}^{k^{\alpha_q}-1}\sum_{j=i+1}^{k^{\alpha_q}} 4(j-i)k^{1-(\alpha_1+\dots+\alpha_q)} + \\ &\quad + \sum_{q=1}^{M}\sum_{s=1}^{k^{\alpha_1+\dots+\alpha_{q-1}}}\sum_{i=2}^{k^{\alpha_q}-1}\sum_{j=i+1}^{k^{\alpha_q}} 6c[\log_2((j-i)k^{1-(\alpha_1+\dots+\alpha_q)})] + \\ &\quad + \sum_{j=2}^{k^{\alpha_1}}\Big(4(j-1)k^{1-\alpha_1} + 6c[\log_2((j-1)k^{1-\alpha_1})]\Big). \end{aligned} \tag{36}$$

Оценим первое слагаемое, которое обозначим Q_{11}.

$$\begin{aligned} Q_{11} &= \sum_{q=1}^{M}\sum_{s=1}^{k^{\alpha_1+\dots+\alpha_{q-1}}}\sum_{i=2}^{k^{\alpha_q}-1}\sum_{j=i+1}^{k^{\alpha_q}} 4(j-i)k^{1-(\alpha_1+\dots+\alpha_q)} = \\ &= 2\sum_{q=1}^{M} k^{1-\alpha_q}\sum_{i=2}^{k^{\alpha_q}-1}(k^{\alpha_q}-i)(k^{\alpha_q}-i+1) = 2\sum_{q=1}^{M}k^{1-\alpha_q}\sum_{i=1}^{k^{\alpha_q}-2} i(i+1) = \\ &= 2\sum_{q=1}^{M}k^{1-\alpha_q}\Big(\frac{(k^{\alpha_q}-2)(k^{\alpha_q}-1)(2k^{\alpha_q}-4+1)}{6} + \frac{(k^{\alpha_q}-2)(k^{\alpha_q}-1)}{2}\Big) = \\ &= \frac{2}{3}\sum_{q=1}^{M}k(k^{\alpha_q}-2)(k^{\alpha_q}-1) = \frac{2}{3}\sum_{q=1}^{M}k^{1+2\alpha_q} + \underline{O}\Big(\sum_{q=1}^{M}k^{1+\alpha_q}\Big). \end{aligned} \tag{37}$$

Оценим второе слагаемое, которое обозначим Q_{12}.

$$\begin{aligned} Q_{12} &= \sum_{q=1}^{M}\sum_{s=1}^{k^{\alpha_1+\dots+\alpha_{q-1}}}\sum_{i=2}^{k^{\alpha_q}-1}\sum_{j=i+1}^{k^{\alpha_q}} 6c[\log_2((j-i)k^{1-(\alpha_1+\dots+\alpha_q)})] < \\ &< 6c\sum_{q=1}^{M}k^{\alpha_1+\dots+\alpha_{q-1}}\sum_{i=2}^{k^{\alpha_q}-1}\sum_{j=i+1}^{k^{\alpha_q}}\log_2(k^{1-(\alpha_1+\dots+\alpha_{q-1})}) = \\ &= 3c\log_2 k\sum_{q=1}^{M}k^{\alpha_1+\dots+\alpha_{q-1}}(1-(\alpha_1+\dots+\alpha_{q-1}))(k^{\alpha_q}-2)(k^{\alpha_q}-1) < \\ &< 3c\log_2 k\sum_{q=1}^{M}k^{\alpha_1+\dots+\alpha_{q-1}+2\alpha_q}(1-(\alpha_1+\dots+\alpha_{q-1})). \end{aligned} \tag{38}$$

Из каждой вершины η^q_{ij}, $q=2,\dots,M$, $i=2,\dots,k^{\alpha_1+\dots+\alpha_{q-1}}$, $j=1,\dots,k^{\alpha_q}-1$ исходит ИГ $U^3_{A^q_{i-1,j+1,k^{\alpha_q}+1}}$, объем которого согласно (18) и (35) равен

$$Q(U^3_{A^q_{i-1,j+1,k^{\alpha_q}+1}}) \le 4(k^{\alpha_q}-j)k^{1-(\alpha_1+\dots+\alpha_q)} - 1 + 6c[\log_2((k^{\alpha_q}-j)k^{1-(\alpha_1+\dots+\alpha_q)})].$$

Суммируя по всем q, i, j, получаем, что объем ИГ, растущих из вершин η^q_{ij}, (обозначим его Q_2) не превышает

$$\begin{aligned} Q_2 &< \sum_{q=2}^{M} \sum_{i=1}^{k^{\alpha_1+\cdots+\alpha_{q-1}}} \sum_{j=1}^{k^{\alpha_q}-1} (4+6c)(k^{\alpha_q}-j)k^{1-(\alpha_1+\cdots+\alpha_q)} = \\ &= (4+6c)\sum_{q=2}^{M} k^{1-\alpha_q} \sum_{j=1}^{k^{\alpha_q}-1} j < (2+3c)\sum_{q=2}^{M} k^{1+\alpha_q} \end{aligned} \tag{39}$$

Из каждой вершины ζ^q_{ij}, $q=2,\ldots,M$, $i=1,\ldots,k^{\alpha_1+\cdots+\alpha_{q-1}}$, $j=2,\ldots,k^{\alpha_q}$, исходит ИГ $U^3_{A^q_{i,1,j}}$, объем которого согласно (18) и (35) равен

$$Q(U^3_{A^q_{i,1,j}}) \le 4(j-1)k^{1-(\alpha_1+\cdots+\alpha_q)} - 1 + 6c[\log_2((j-1)k^{1-(\alpha_1+\cdots+\alpha_q)})].$$

Суммируя по всем q, i, j, получаем, что объем ИГ, растущих из вершин ζ^q_{ij}, (обозначим его Q_3) не превышает

$$\begin{aligned} Q_3 &< \sum_{q=2}^{M} \sum_{i=1}^{k^{\alpha_1+\cdots+\alpha_{q-1}}} \sum_{j=2}^{k^{\alpha_q}} (4+6c)(j-1)k^{1-(\alpha_1+\cdots+\alpha_q)} = \\ &= (4+6c)\sum_{q=2}^{M} k^{1-\alpha_q} \sum_{j=1}^{k^{\alpha_q}-1} j < (2+3c)\sum_{q=2}^{M} k^{1+\alpha_q} \end{aligned} \tag{40}$$

Из каждой вершины β^M_j, $j=1,\ldots,k^{\alpha_1+\cdots+\alpha_M}$, исходит по одному ребру, следовательно, суммарное число ребер, исходящих из вершин β^M_j, равно

$$Q_4 = k^{\alpha_1+\cdots+\alpha_M} = k. \tag{41}$$

Исходя из (37), положим

$$\alpha_1 = \alpha_2 = \cdots = \alpha_M = 1/M. \tag{42}$$

Согласно (34) и (42)

$$\begin{aligned} Q(U_M) &= \frac{3+c}{2}\sum_{q=1}^{M} k^{\frac{q+1}{M}} + \underline{O}(k) = \frac{3+c}{2}k^{1+1/M}\sum_{q=1}^{M} k^{-\frac{q}{M}} + \underline{O}(k) = \\ &= \frac{3+c}{2}k^{1+1/M}\frac{1-k^{-1}}{1-k^{-1/M}} + \underline{O}(k) = \underline{O}(k^{1+1/M}). \end{aligned} \tag{43}$$

Поскольку

$$\begin{aligned} &\sum_{i=0}^{n} d^i(i+1) = \sum_{i=0}^{n}\sum_{j=i}^{n} d^j = \sum_{i=0}^{n} d^i\frac{1-d^{n-i+1}}{1-d} = \\ &= \frac{1}{1-d}\sum_{i=0}^{n}\left(d^i - d^{n+1}\right) = \frac{1}{(1-d)^2}\Big(1-(n+2)d^{n+1}+(n+1)d^{n+2}\Big), \end{aligned}$$

то при $\alpha_1 = \alpha_2 = \dots = \alpha_q = 1/M$ согласно (38)

$$\begin{aligned}
Q_{12} &< 3c\log_2 k \sum_{q=1}^{M} k^{(q+1)/M}\Big(1 - \frac{q-1}{m}\Big) = \\
&= \frac{3c}{M} k^{1+1/M} \log_2 k \sum_{q=1}^{M} k^{(q-M)/M}(M-q+1) = \\
&= \frac{3c}{M} k^{1+1/M} \log_2 k \sum_{q=0}^{M-1} k^{-q/M}(q+1) = \\
&= \frac{3c}{M} k^{1+1/M} \log_2 k \frac{1}{(1-k^{-1/M})^2}\Big(1 - (M+1)k^{-1} + Mk^{-M-1/M}\Big) = \\
&= \underline{O}\Big(\frac{1}{M} k^{1+1/M} \ln k\Big). \qquad (44)
\end{aligned}$$

Таким образом, согласно (43), (36), (37), (44), (39), (40), (41) получим

$$\begin{aligned}
Q(U) &= Q(U_M) + Q_1 + Q_2 + Q_3 + Q_4 \le \\
&\le \frac{2}{3} M k^{1+2/M} + \underline{O}\Big(Mk^{1+1/M} + \frac{\ln k}{M} k^{1+1/M}\Big).
\end{aligned}$$

7 Сложность информационного графа

Если X — множество запросов с заданным на нем вероятностным пространством $\langle X, \sigma, \mathbf{P}\rangle$, $\mathcal{F}$ — базовое множество функций, определенных на X, U — ИГ над $\mathcal{F}$, β — вершина ИГ U, то *сложностью вершины* β называется число $\mathbf{P}(N_{\varphi_\beta})$, *сложностью предикатного ребра* (β, α) называется число $\mathbf{P}(N_{\varphi_\beta})$, то есть сложность начала ребра, *сложностью переключательного ребра* (β, α) называется число $\mathbf{P}(N_{\varphi_\beta})/\psi_\beta$, где, напомним, ψ_β — количество ребер, исходящих из вершины β.

Как показано в [27] сложность ИГ равна сумме сложностей ребер ИГ.

Прежде чем вычислять сложность ИГ U^*, докажем одно вспомогательное утверждение.

Пусть β некоторая вершина некоторого ИГ U. Скажем что подграф U' графа U является *зависимым* от β, если он содержит вершину β, является связным, состоит только из таких вершин α ИГ U, что любая цепь, ведущая из корня ИГ U в вершину α, проходит через вершину β и содержит все ребра ИГ U, исходящие из вершин U'.

Так, например, если вершина β является корнем связного ИГ U, то ИГ U будет зависимым от β. А для произвольной вершины β любой зависимый от β граф по крайней мере содержит все ребра, исходящие из вершины β.

Договоримся через $\mathbf{P}(A|B)$ обозначать условную вероятность события A при условии осуществления события B.

Пусть X — множество запросов с заданным на нем вероятностным пространством $\langle X, \sigma, \mathbf{P}\rangle$. Пусть $X' \in \sigma$. Вероятностным пространством, порожденным множеством X', назовем вероятностное пространство $\langle X', \sigma', \mathbf{P}'\rangle$, где $\sigma' = \{B \in \sigma : B \subseteq X'\}$, $\mathbf{P}'$ — такая вероятностная мера на σ', что для любого $B \in \sigma'$ $\mathbf{P}'(B) = \mathbf{P}(B|X')$.

Лемма 1 *Пусть β некоторая вершина некоторого ИГ U над базовым множеством, функции которого определены на множестве запросов X. Пусть $\langle X, \sigma, \mathbf{P}\rangle$*

— вероятностное пространство над X. Пусть $X' = N_{\varphi_\beta}$. Пусть $\langle X', \sigma', \mathbf{P}' \rangle$ — вероятностное пространство, порожденное множеством X'. Пусть U' — зависимый от β подграф ИГ U. Пусть T' — сложность ИГ U' (как отдельного ИГ с корнем в вершине β) для вероятностного пространства $\langle X', \sigma', \mathbf{P}' \rangle$. Тогда сложность U' как подграфа ИГ U для вероятностного пространства $\langle X, \sigma, \mathbf{P} \rangle$ равна $T(U') = \mathbf{P}(X') \cdot T'$.

Доказательство: Рассмотрим произвольную вершину α подграфа U'. Из определения зависимого подграфа следует, что $N_{\varphi_\alpha} \subseteq N_{\varphi_\beta} = X'$. Следовательно,

$$\mathbf{P}(N_{\varphi_\alpha}) = \mathbf{P}(N_{\varphi_\alpha} \cap X') = \mathbf{P}(X') \cdot \mathbf{P}(N_{\varphi_\alpha} | X') = \mathbf{P}(X') \cdot \mathbf{P}'(N_{\varphi_\alpha}).$$

Откуда сразу следует утверждение леммы.

Вычислим сложность ИГ U^*.

Пусть β — некоторая вершина ИГ U^*. Чтобы сократить этажность формул договоримся обозначать $N(\beta) = N_{\varphi_\beta}$.

Введем следующие множества вершин ИГ U^*

$$\mathcal{B}_1 = \{\theta_i : i = 1, 2, \ldots, k^{1/M}\},$$

$$\mathcal{B}_2 = \{\theta''_i : i = 1, 2, \ldots, k^{1/M}\},$$

$$\mathcal{B}_3 = \{\theta_{ij} : i = 1, 2, \ldots, k^{1/M} - 1; j = i + 1, \ldots, k^{1/M}\},$$

$$\mathcal{B}_4^q = \{\theta_{qij} : i = 2, \ldots, k^{(q-1)/M} + 1; j = 1, \ldots, k^{1/M} - 1\}, q = 2, \ldots, M,$$

$$\mathcal{B}^4 = \cup_{q=2}^M \mathcal{B}_4^q.$$

Обозначим через T_q ($q = 1, 2, \ldots, M$) суммарную сложность ребер ИГ U_q за исключением ребер, исходящих из вершин из множества $\mathcal{B}_4$. Через $\widehat{T}_q(\tilde{x})$ ($q = 1, 2, \ldots, M$) обозначим сложность ИГ U_q на запросе $\tilde{x}$ минус количество вычисленных переключателей, приписанных вершинам из множества $\mathcal{B}_4$.

Покажем по индукции, что

$$T_q \le 6q, \qquad \widehat{T}_q(\tilde{x}) \le 6q + \frac{2q}{M}\log_2 k$$

для любого $\tilde{x} \in X_{int2}$.

Базис индукции: $q = 1$.

Из соотношений (15), (26), (27) легко следуют следующие утверждения

$$\forall \beta_1, \beta_2 \in \mathcal{B}_i (\beta_1 \ne \beta_2 \Rightarrow N(\beta_1) \cap N(\beta_2) = \emptyset), \ i = 1, 2, 3. \tag{45}$$

Поскольку из корня ИГ U^* растет ИГ $U^1_{\mathcal{N}^{1,0,1}}$, сложность которого оценивается соотношением (16), вершинам θ_i приписаны переключатели $g^4_{l^1_i}$, из вершин θ''_i исходят ИГ $U^2_{Z^{1,i}, l^1_i, 1}$, зависимые от θ''_i и сложность которых оценивается соотношением (17), а из вершин θ_{ij} исходят ребра с предикатом тождественная 1, то с учетом (45) согласно лемме 1 имеем

$$T_1 = 2 + \sum_{\beta \in \mathcal{B}_1} \mathbf{P}(N(\beta)) + \sum_{\beta \in \mathcal{B}_2} 2\mathbf{P}(N(\beta)) + \sum_{\beta \in \mathcal{B}_3} \mathbf{P}(N(\beta)) \le 6,$$

и для любого $\tilde{x} \in X_{int2}$

$$\widehat{T}_1(\tilde{x}) \le 1 +]\log_2(k^{1/M}+1)[+1+1+]\log_2(k^{1/M}+1)[+1 \le 6 + \frac{2}{M}\log_2 k.$$

Индуктивный переход. Пусть $q \in \{2, \ldots, M\}$ и утверждение индукции выполняется для $q-1$. Обозначим

$$\begin{aligned}
\mathcal{B}_5^q &= \{\vartheta_i^{q-1} : i = 2, \ldots, k^{(q-1)/M}\} \cup \{\gamma_i^{q-1} : i = 2, \ldots, k^{(q-1)/M}+1\}, \\
\mathcal{B}_6^q &= \{\beta_i^{q-1} : i = 1, \ldots, k^{(q-1)/M}\} \cup \\
&\quad \cup \{\theta''_{qij} : i = 2, \ldots, k^{(q-1)/M}+1;\ j = 1, \ldots, k^{1/M}-1\}, \\
\mathcal{B}_7^q &= \{\tau'_{qij} : i = 2, \ldots, k^{(q-1)/M};\ j = 1, \ldots, k^{1/M}\} \cup \\
&\quad \cup \{\theta''_{qij} : i = 2, \ldots, k^{(q-1)/M}+1;\ j = 1, \ldots, k^{1/M}-1\}, \\
\mathcal{B}_8^q &= \{\tau''_{qij} : i = 1, \ldots, k^{(q-1)/M};\ j = 1, \ldots, k^{1/M}\} \cup \\
&\quad \cup \{\mu_{ijt}^q : i = 2, \ldots, k^{(q-1)/M}+1;\ j = 1, \ldots, k^{1/M}-1;\ t = j+1, \ldots, k^{1/M}\}.
\end{aligned}$$

Из соотношений (20), (21), (22), (28), (29), (31), (32) легко следуют следующие утверждения

$$\forall \beta_1, \beta_2 \in \mathcal{B}_i^q (\beta_1 \ne \beta_2 \Rightarrow N(\beta_1) \cap N(\beta_2) = \emptyset),\ i = 5, 6, 7, 8. \tag{46}$$

Поскольку из вершин ϑ_i^{q-1} и γ_i^{q-1} исходят ИГ $U^1_{\mathcal{N}_{i-1}^q, l_{i-1}^{q-1}, l_i^{q-1}}$, сложность которых оценивается соотношением (16) и зависимые от этих вершин, из вершин β_i^{q-1} исходят ИГ $U^2_{\mathcal{N}_i^q, l_i^{q-1}, l_{i+1}^{q-1}}$, а из вершин θ''_{qij} — ИГ $U^2_{Z_{i-1}^{q,j}, l^q_{(i-2)k^{\alpha_q}+j+1}, l_i^{q-1}}$, сложность которых оценивается соотношением (17) и которые зависимы от своих корневых вершин, из каждой вершины из множества $\mathcal{B}_7^q$ исходит ребро с предикатом тождественная 1, ведущее в вершины ϑ_i^q, а из каждой вершины из множества $\mathcal{B}_8^q$ исходит ребро с предикатом тождественная 1, ведущее в вершины β_i^q, то с учетом (46) согласно лемме 1 имеем

$$\begin{aligned}
T_q &= T_{q-1} + \sum_{\beta \in \mathcal{B}_5^q} 2\mathbf{P}(N(\beta)) + \sum_{\beta \in \mathcal{B}_6^q} 2\mathbf{P}(N(\beta)) + \sum_{\beta \in \mathcal{B}_7^q} \mathbf{P}(N(\beta)) + \\
&\quad + \sum_{\beta \in \mathcal{B}_8^q} \mathbf{P}(N(\beta)) \le 6(q-1) + 6 = 6q,
\end{aligned}$$

и для любого $\tilde{x} \in X_{int2}$ выполняется

$$\begin{aligned}
\widehat{T}_q(\tilde{x}) &\le \widehat{T}_{q-1}(\tilde{x}) + 1 +]\log_2(k^{1/M}+1)[+1+]\log_2(k^{1/M}+1)[+ \\
&\quad +1+1 \le 6q + \frac{2q}{M}\log_2 k.
\end{aligned}$$

При $q = M$ нам надо делать на одно вычисление меньше, так как в ИГ U_M нет вершин типа ϑ_i^M, то есть

$$T_M \le 6M - 1, \qquad \widehat{T}_M(\tilde{x}) \le 6M - 1 + 2\log_2 k$$

для любого $\tilde{x} \in X_{int2}$.

Обозначим

$$\begin{array}{rcl}\mathcal{B}_9^1 & = & \{\delta_{1ij}^1 : i=1,\dots,k^{1/M}-1;\ j=i+1,\dots,k^{1/M}\},\\ \mathcal{B}_9^q & = & \{\delta_{sij}^q : s=1,\dots,k^{(q-1)/M};\ i=2,\dots,k^{1/M}-1;\\ & & j=i+1,\dots,k^{1/M}\},\ q=2,\dots,M,\\ \mathcal{B}_{10}^q & = & \{\eta_{ij}^q : i=2,\dots,k^{(q-1)/M};\ j=1,\dots,k^{1/M}-1\},\ q=2,\dots,M,\\ \mathcal{B}_{11}^q & = & \{\zeta_{ij}^q : i=1,\dots,k^{(q-1)/M};\ j=2,\dots,k^{1/M}-1\},\ q=2,\dots,M,\\ \mathcal{B}_9 & = & \cup_{q=1}^M \mathcal{B}_9^q, \qquad \mathcal{B}_{10} \ = \ \cup_{q=2}^M \mathcal{B}_{10}^q, \qquad \mathcal{B}_{11} \ = \ \cup_{q=2}^M \mathcal{B}_{11}^q,\\ \mathcal{B}_{12} & = & \{\beta_i^M : i=1,\dots,k\}, \qquad \mathcal{B}_{13} \ = \ \mathcal{B}_9 \cup \mathcal{B}_{10} \cup \mathcal{B}_{11}.\end{array}$$

Из (21), (23), (24), (25), (30) непосредственно проверяется справедливость следующих утверждений

$$\forall \beta_1, \beta_2 \in \mathcal{B}_9 \cup \mathcal{B}_{12} (\beta_1 \neq \beta_2 \Rightarrow N(\beta_1) \cap N(\beta_2) = \emptyset), \tag{47}$$

$$\forall q \in \{2,\dots,M\} (\forall \beta_1, \beta_2 \in \mathcal{B}_i^q \cup \mathcal{B}_4^q (\beta_1 \neq \beta_2 \Rightarrow N(\beta_1) \cap N(\beta_2) = \emptyset)),\ i=10,11, \tag{48}$$

$$\forall \beta_1, \beta_2 \in \mathcal{B}_{13} (\beta_1 \neq \beta_2 \Rightarrow (N(\beta_1) \cap N(\beta_2) = \emptyset \vee A_{\beta_1} \cap A_{\beta_2} = \emptyset)).$$

Пусть $\beta \in \mathcal{B}_{13}$, тогда для некоторых q, s, i, j из вершины β исходит ИГ $U^3_{A^q_{sij}}$, зависимый от вершины β. Обозначим

$$A_\beta = \{(y_p^1, y_p^2) \in V : l^q_{(s-1)k^{\alpha_q}+i} \le y_p^1 < l^q_{(s-1)k^{\alpha_q}+j}\}.$$

Согласно лемме 1 и (18) для данной вершины β

$$\begin{array}{rcl} T(U^3_{A^q_{sij}}) & = & \mathbf{P}(N(\beta))\Big(5 + \sum\limits_{y \in A^q_{sij}} \mathbf{P}'(O(y, \rho_{int}))\Big) = \\ & = & 5\mathbf{P}(N(\beta)) + \sum\limits_{\tilde{y} \in A_\beta} \mathbf{P}(O(\tilde{y}, \rho_{int2}) \cap N(\beta)). \end{array}$$

Учитывая, что при попадании в вершину из $\mathcal{B}_{12}$ мы вычисляет один предикат, приписанный ребру, исходящему из этой вершины, а при попадании в вершину из $\mathcal{B}_4$ мы вычисляет один переключатель, приписанный этой вершине, имеем

$$\begin{array}{rcl} T(U^*) & \le & T_M + \sum\limits_{\beta \in \mathcal{B}_4} \mathbf{P}(N(\beta)) + \sum\limits_{\beta \in \mathcal{B}_{12}} \mathbf{P}(N(\beta)) + \\ & & + \sum\limits_{\beta \in \mathcal{B}_{13}} \Big(5\mathbf{P}(N(\beta)) + \sum\limits_{\tilde{y} \in A_\beta} \mathbf{P}(O(\tilde{y}, \rho_{int2}) \cap N(\beta))\Big) \le \\ & \le & T_M + 5 \sum\limits_{\beta \in \mathcal{B}_9 \cup \mathcal{B}_{12}} \mathbf{P}(N(\beta)) + 5 \sum\limits_{q=2}^{M} \sum\limits_{\beta \in \mathcal{B}_4^q \cup \mathcal{B}_{10}^q} \mathbf{P}(N(\beta)) + \\ & & +5 \sum\limits_{q=2}^{M} \sum\limits_{\beta \in \mathcal{B}_{11}^q} \mathbf{P}(N(\beta)) + \sum\limits_{\tilde{y} \in V} \sum\limits_{\beta \in \mathcal{B}_{13} : \tilde{y} \in A_\beta} \mathbf{P}(O(\tilde{y}, \rho_{int2}) \cap N(\beta)) \le \\ & \le & 16M - 6 + \sum\limits_{\tilde{y} \in V} \mathbf{P}(O(\tilde{y}, \rho_{int2})). \end{array}$$

Согласно (47) для любого запроса мы можем попасть не более чем в одну вершину из $\mathcal{B}_9 \cup \mathcal{B}_{12}$, следовательно при этом мы в худшем случае выполним $2 + \log_2 k$

вычислений плюс перечисление ответа. Согласно (48) для каждого $q = 2, \ldots, M$ и любого запроса мы можем попасть не более чем в одну вершину из $\mathcal{B}_4 \cup \mathcal{B}_{10}$ и не более чем в одну вершину из $\mathcal{B}_{11}$ и в каждом из этих случаев мы выполним не более чем $2 + \frac{1}{M}\log_2 k$ вычислений плюс перечисление ответа. Тем самым для любого запроса $\tilde{x} \in X_{int2}$

$$\begin{aligned} T(U^*, x) \quad &\leq \quad \widehat{T}_M + 2 + \log_2 k + 4(M-1) + \frac{M-1}{m}\log_2 k + \\ &+ |\{\tilde{y} \in V : \tilde{x}\rho_{int2}\tilde{y}\}| < 10M + 4\log_2 k + |\{\tilde{y} \in V : \tilde{x}\rho_{int2}\tilde{y}\}|. \end{aligned}$$

Тем самым теорема 1 полностью доказана.

Список литературы

[1] Ли Д., Препарата Ф. Вычислительная геометрия. Обзор. *Кибернетический сб.* (1987), **24**, 5–96.

[2] Кнут Д. Искусство программирования для ЭВМ. Сортировка и поиск. **3**, Мир, Москва, 1978.

[3] Loftsgaarden D.O., Queensberry C.P. A nonparametric density function. *Ann. Math. Stat.* (1965), **36**, 1049–1051.

[4] Lauter U. 4-dimensional binary search trees as a means to speed up associative searches in design verification of integrated circuits. *Jour. of Design Automation and Fault Tolerant Computing,* **2** (3), 241–247 (July 1978).

[5] Гасанов Э. Э. Мгновенно решаемые задачи поиска. *Дискретная математика* (1996) **8**, № 3, 119-134.

[6] Препарата Ф., Шеймос М. Вычислительная геометрия: Введение. Мир, Москва, 1989.

[7] Bentley J. L. Multidimensional binary search trees used for associative searching. *Commun. Ass. Comput. Mach.* (Sept. 1975), **18** 509–517.

[8] Bentley J.L. Decomposable searching problems, *Info. Proc. Lett.* (1979), **8** 244-251.

[9] Bentley J.L., Friedman J.H. Data structures for range searching. *Comput. Surveys* (1979), **11** 397–409.

[10] Bentley J.L., Maurer H.A. Efficient worst-case data structures for range searching. *Acta Informatica* (1980), **13** 155–168.

[11] Bentley J.L., Shamos M.I. A problem in multivariate statistics: Algorithms, data structure and applications. *Proc. 15th Allerton Conf. Commun., Contr., Comput.* (1977), 193–201.

[12] Bentley J.L., Stanat D.F. Analysis of range range searching in quad trees. *Inform. Processing Lett.* (1975), **3** 170–173.

[13] Bolour A. Optimal retrieval algorithms for small region queries. *SIAM J. Comput.* (1981) **10**, 721–741.

[14] Chazelle B.M. Filtering search: a new approach to query-answering. *Proc. 24th IEEE Annu. Symp. Found. Comput. Sci.* (Nov. 1983), 122–132.

[15] Fredman M. L. A lower bound of the complexity of ortogonal range queries. *J. ACM.* (1981) **28**, 696–705.

[16] Gabow H.N., Bentley J.L., Tarjan R.E. Scaling and related techniques for geometry problems. *Proc. 16th ACM Annu. Symp. Theory Comput.* (Apr. 1984) 135–143.

[17] Lee D.T., Wong C.K. Worst case analysis for region and partial region searches in multidimensional binary search trees and balansed quad trees. *Acta Informatica* (1977) **9** 23–29.

[18] Lee D.T., Wong C.K. Quintari trees: A file structures for multidimensional database system. *ACM Trans. Database Syst.* (Sept. 1980) **1**, 1, 339–353.

[19] Lueker G.S. A data structure for ortogonal range queries. *Processing of the 19th Annual IEEE Symposium on Foundations of Computer Science.* (1978), 28–34.

[20] Lueker G.S., Willard D.E. A data structure for dynamic range queries. *Inform. Processing Lett.* (Dec. 1982) **15**, 5, 209–213.

[21] Saxe J.B. On the number of range queries in k-space. *Discrete Applied Mathematics* (1979) **1**, 217–225.

[22] Willard D.E. *Predicate-oriented database search algorithms.* Ph. D. dissertation, Harvard Univ., Cambridge, MA, 1978.

[23] Абдул Саттар А.В. Функциональная мера сложности вычислений в автоматных схемах. Автореферат диссертации на соискание ученой степени кандидата физико-математических наук. Москва, 1994.

[24] Гасанов Э. Э., Кузнецова И. В. Оценки функциональной сложности двумерной задачи интервального поиска. *Тезисы докладов XII Международной конференции "Проблемы теоретической кибернетики", Нижний Новгород, (17–22 мая 1999 г.),* — 47.

[25] Gasanov E. E., Kuznetsova I. V. On one method to decrease average search time. *Abstracts of I^{st} Turkish World Mathematics Symposium Elazig, Turkey. (29 June – 2 July 1999)* — p. 135.

[26] Гасанов Э. Э. Информационно-графовая модель хранения и поиска данных. *Интеллектуальные системы* (1998) **3**, №3–4, 163–192.

[27] Гасанов Э. Э. Функционально-сетевые базы данных и сверхбыстрые алгоритмы поиска. Изд. центр РГГУ, Москва, 1997.