Math-Net.Ru

Общероссийский математический портал

Э. Э. Гасанов, Мгновенно решаемые задачи поиска, *Дискрет. матем.*, 1996, том 8, выпуск 3, 119–134

DOI: <https://doi.org/10.4213/dm532>

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: 3.80.181.102

26 апреля 2023 г., 12:46:08

Мгновенно решаемые задачи поиска

© 1996 г. Э. Э. Гасанов

Вводится понятие мгновенно решаемых задач поиска, под которыми понимаются задачи, которые могут быть решены в среднем за время, необходимое на перечисление ответа плюс некая не зависящая от размерности задачи константа. Приводятся примеры мгновенно решаемых задач и алгоритмы, обеспечивающие мгновенное решение.

1. Введение

Известным фактом в теории информационного поиска является тот факт, что решение любой задачи информационного поиска (ЗИП) требует времени не меньшего, чем время требуемое на перечисление ответа (см., например, теорему 3 в [1]), которое тем самым является тривиальной нижней оценкой сложности ЗИП.

Поэтому естественный интерес представляет проблема нахождения таких алгоритмов решения ЗИП, которые по временной сложности не существенно отличаются от тривиальной нижней оценки, в частности, отличаются на константу, не зависящую от размерности задачи. Задачи поиска для которых существуют такие алгоритмы предлагается называть мгновенно решаемыми задачами поиска.

Итак, под мгновенно решаемыми задачами поиска понимаются такие ЗИП, которые могут быть решены в среднем за время, необходимое на перечисление ответа, плюс некая не зависящая от размерности задачи константа. Здесь под размерностью задачи понимается объем информационного массива, в котором производится поиск.

Автором ранее были рассмотрены две задачи, для которых было показано, что они мгновенно решаемые. Это ЗИП с отношением поиска, являющимся отношением линейного квазипорядка [2], и одномерная задача интервального поиска [3].

В данной работе приводятся три задачи, для которых показано, что они мгновенно решаемые. Это известные задачи:

задача поиска идентичных объектов, которая состоит в поиске в информационном массиве объекта, идентичного объекту-запросу;

задачи о близости, которые состоят в поиске во множестве, в котором задан линейный порядок, объекта, ближайшего к объекту-запросу справа или слева;

n -мерная задача интервального поиска, которая состоит в поиске в конечном подмножестве n -мерного пространства всех тех точек, которые попадают в n -мерный параллелепипед-запрос.

Но если в задачах, исследованных в [2] и [3], удастся получить хорошие алгоритмы, обеспечивающие мгновенное решение, которые практически не требуют дополнительной памяти, то мгновенного решения задач, исследуемых в данной работе, удалось добиться только за счет больших затрат в объеме памяти. В тех ситуациях, когда время является очень дорогим, такие алгоритмы могут быть очень полезны.

Перечисленные выше ЗИП исследуются в рамках модели, введенной автором в [3], т.е. с помощью информационных сетей с переключателями (ИСП).

Автор выражает благодарность В. Б. Кудрявцеву и А. С. Подколзину.

2. Основные понятия и формулировка результатов

Напомним понятие ИСП и другие необходимые понятия.

Пусть X — множество запросов, причем на X определена вероятность, т.е. задано вероятностное пространство $\{X, \sigma, P\}$, где σ — алгебра подмножеств множества X , P — вероятностная мера на σ ; $V = \{y_1, y_2, \dots, y_k\}$ — конечная библиотека, причем $V \subseteq Y$, где Y — множество записей (объектов поиска); ρ — бинарное отношение на $X \times Y$, называемое отношением поиска; $I = \langle X, V, \rho \rangle$ — задача информационного поиска (ЗИП), которая состоит в перечислении для произвольно взятого запроса $x \in X$ всех тех и только тех записей из V , которые находятся в отношении ρ с запросом x ; $O(y, \rho) = \{x \in X : x\rho y\}$ — тень записи $y \in Y$; $N_f = \{x \in X : f(x) = 1\}$, где f — одноместный предикат, определенный на X , т.е. $f: X \rightarrow \{0, 1\}$; $\chi_{y, \rho}: X \rightarrow \{0, 1\}$ такая, что $N_{\chi_{y, \rho}} = O(y, \rho)$ — характеристическая функция записи y ; F — множество символов одноместных предикатов, определенных на множестве X , называемое базовым множеством предикатов; G — множество символов одноместных переключателей, определенных на множестве X . Под переключателями будем понимать функции, область значений которых являются конечные подмножества натурального ряда.

Пару $\mathcal{F} = \langle F, G \rangle$ назовем базовым множеством.

Если n — натуральное число, а $g(x)$ — некий переключатель, то через $\xi_g^n(x)$ обозначим предикат, определенный на X , такой, что

$$N_{\xi_g^n} = \{x \in X : g(x) = n\}.$$

Обозначим

$$\widehat{G} = \{\xi_g^n : g \in G, n \in \mathbb{N}\}.$$

Определение понятия ИСП можно разбить на два этапа. На первом этапе раскрывается структурная (схемная) часть этого понятия, на втором функциональная.

Начнем с определения ИСП с точки зрения ее структуры. Пусть нам дана ориентированная многополюсная сеть. Выделим в ней один полюс и назовем его корнем, а остальные полюсы назовем листьями. Выделим в сети некоторые

вершины и назовем их точками переключения (полюсы могут быть точками переключения). Если β — вершина сети, то через ψ_β обозначим полустепень исхода вершины β .

Каждой точке переключения β сопоставим некий символ из G такой, что максимальное значение переключателя, соответствующего этому символу, не превышает ψ_β . Это соответствие назовем нагрузкой точек переключения.

Для каждой точки переключения β ребрам, из нее исходящим, поставим во взаимно однозначное соответствие числа из множества $\{1, \dots, \psi_\beta\}$. Эти ребра назовем переключательными, а это соответствие нагрузкой переключательных ребер. Ребра, не являющиеся переключательными, назовем предикатными.

Каждому предикатному ребру сети сопоставим некоторый символ из множества F . Это соответствие назовем нагрузкой предикатных ребер.

Сопоставим каждому листу сети некоторую запись из множества Y . Это соответствие назовем нагрузкой листьев.

Полученную нагруженную сеть назовем информационной сетью с переключателями над базовым множеством $\mathcal{F} = \langle F, G \rangle$.

Перейдем ко второму этапу и определим функционирование ИСП. Пусть дана ИСП U . Последовательность ориентированных ребер сети

$$(\alpha_1, \alpha_2), (\alpha_2, \alpha_3), \dots, (\alpha_{m-1}, \alpha_m)$$

назовем ориентированной цепью от вершины α_1 к вершине α_m . Если c — ребро сети, то через $[c]$ обозначим его нагрузку.

Проводимостью ребра (α, β) назовем предикат, равный $[(\alpha, \beta)]$, если ребро предикатное, и равный $\xi_g^{[(\alpha, \beta)]}$, если ребро переключательное, где g — переключатель, соответствующий вершине α .

Проводимостью ориентированной цепи назовем конъюнкцию проводимостей ребер цепи.

В ИСП по аналогии с контактными схемами введем для каждой пары вершин α и β функцию проводимости $f_{\alpha\beta}$ от вершины α к вершине β , для всех $x \in X$ полагая $f_{\alpha\beta}(x) \equiv 1$ ($x \in X$), если $\alpha = \beta$; $f_{\alpha\beta}(x) \equiv 0$, если $\alpha \neq \beta$ и не существует в ИСП ориентированных цепей от α к β ; если $\alpha \neq \beta$ и множество ориентированных цепей от α к β не пусто, то $f_{\alpha\beta}(x)$ равно дизъюнкции проводимостей всех ориентированных цепей от α к β .

Функцию проводимости от корня ИСП к некоторой вершине β ИСП назовем функцией фильтра вершины β и обозначим $\varphi_\beta(x)$.

Через $\mathcal{R}(U)$, $\mathcal{P}(U)$, $\mathcal{L}(U)$ (или просто $\mathcal{R}$, $\mathcal{P}$, $\mathcal{L}$) обозначим множества вершин, точек переключения и листьев сети U соответственно.

Пусть $\mathcal{N}$ — некоторая подсеть (т.е. произвольное подмножество вершин и ребер) ИСП U . Через $\langle \mathcal{N} \rangle$ обозначим множество записей, соответствующих листьям этой подсети (в частности, если α — некоторый лист сети U , то под $\langle \alpha \rangle$ будем понимать запись, соответствующую листу α).

Будем говорить, что ИСП U реализует функцию $\mathcal{J}: X \rightarrow 2^Y$, называемую функцией ответа сети U и определяемую соотношением

$$\mathcal{J}(x) = \{ \langle \alpha \in \mathcal{L}(U) : \varphi_\alpha(x) = 1 \rangle \}.$$

Понятие ИСП полностью определено.

Скажем, что ИСП U разрешает ЗИП $I = \langle X, V, \rho \rangle$, если для всех $x \in X$

$$\mathcal{J}(x) = \{y \in V : x\rho y\}.$$

Сложностью ИСП U на запросе x назовем число

$$T(U, x) = b \sum_{\beta \in \mathcal{R} \setminus \mathcal{P}} \psi_{\beta} \varphi_{\beta}(x) + a \sum_{\beta \in \mathcal{P}} \varphi_{\beta}(x),$$

где константа a характеризует сложность вычисления одного переключателя, а константа b одного предиката.

Далее всюду будем предполагать, что мы находимся в условиях леммы 1 из [3], т.е. будем считать, что алгебра σ содержит все множества N_f , где $f \in F \cup \widehat{G}$. При этих условиях $T(U, x)$ — случайная величина по x и можно ввести понятия сложности сети и сложности ребра сети.

Сложностью ИСП U назовем математическое ожидание величины $T(U, x)$, т.е. число

$$T(U) = \mathbf{M}T(U, x).$$

Если (β, α) — ребро ИСП, то его сложностью назовем число $b\mathbf{P}(N_{\varphi_{\beta}})$, если (β, α) — предикатное ребро, и число $a\mathbf{P}(N_{\varphi_{\beta}})/\psi_{\beta}$, если это ребро переключательное.

Легко видеть, что сложность ИСП равна сумме сложностей ребер ИСП, т.е.

$$T(U) = b \sum_{\beta \in \mathcal{R} \setminus \mathcal{P}} \psi_{\beta} \mathbf{P}(N_{\varphi_{\beta}}) + a \sum_{\beta \in \mathcal{P}} \mathbf{P}(N_{\varphi_{\beta}}).$$

Далее всюду будем предполагать, что $a = b = 1$.

Пусть дана ИСП U . Объемом $Q(U)$ ИСП U назовем число ребер в сети U .

Пусть дана некая ЗИП I . Сложностью задачи I при базовом множестве $\mathcal{F}$ и заданном объеме q назовем число

$$T(I, \mathcal{F}, q) = \inf\{T(U) : U \in \mathcal{U}(I, \mathcal{F}), Q(U) \leq q\},$$

где $\mathcal{U}(I, \mathcal{F})$ — множество всех ИСП над базовым множеством $\mathcal{F}$, разрешающих ЗИП I .

Число

$$T(I, \mathcal{F}) = \min\{T(U, \mathcal{F}, q) : q \in \mathbf{N}\}$$

назовем сложностью задачи I при базовом множестве $\mathcal{F}$.

Рассмотрим следующую ЗИП, которая описывает задачу поиска идентичных объектов. Пусть Y — множество записей и на Y задано отношение линейного порядка $\preceq$. Пусть $X = Y$ — множество запросов и $V = \{y_1, \dots, y_k\} \subseteq Y$.

Пусть отношение поиска ρ есть отношение идентичности, т.е.

$$x\rho y \iff x = y.$$

При выполнении этих условий ЗИП $I = \langle X, V, \rho \rangle$ будем называть задачей поиска идентичных объектов.

Пусть задано вероятностное пространство $\{X, \sigma, \mathbf{P}\}$. Пусть $g_m^1(x)$ — переключатель такой, что

$$g_m^1(x) = i, \quad x \in X_i, \quad i = 1, \dots, m, \quad (1)$$

где $X_1, \dots, X_m$ — разбиение множества X (т. е. $X = X_1 \cup \dots \cup X_m$ и $X_i \cap X_j = \emptyset$, если $i \neq j$) такое, что $P(X_i) \leq c/m$, $i = 1, \dots, m$, где c — постоянная, не зависящая от m .

Положим

$$g_a^2(x) = \begin{cases} 1, & \text{если } x \preccurlyeq a, a \in X, \\ 2 & \text{в противном случае,} \end{cases} \quad (2)$$

$$f_a(x) = \begin{cases} 0, & \text{если } x \neq a \\ 1, & \text{если } x = a, a \in X. \end{cases} \quad (3)$$

Пусть

$$G_1 = \{g_m^1(x) : m \in \mathbf{N}\}, \quad (4)$$

$$G_2 = \{g_a^2(x) : a \in X\}, \quad (5)$$

$$F = \{f_a(x) : a \in X\}, \quad (6)$$

$$\mathcal{F} = \langle F, G_1 \cup G_2 \rangle. \quad (7)$$

Пусть $\mathbf{N}_0 = \mathbf{N} \cup \{0\}$ — множество целых неотрицательных чисел и на $\mathbf{N}_0$

$$L_1(l) = \begin{cases} 0, & \text{если } l = 0, \\ ] \log l[+ 1, & \text{если } l = 1, 2, 3, \\ \log l + 2, & \text{если } l \geq 4. \end{cases} \quad (8)$$

Справедлива следующая теорема.

Теорема 1. Пусть $I = \langle X, V, \rho \rangle$ — задача поиска идентичных объектов, где $|V| = k$, $\mathcal{F}$ — базовое множество, определяемое соотношениями (1)–(7). Пусть

$$s(k, m) = \begin{cases} m + k, & \text{если } k \leq 2m, \\ m + k + (k - 2m)2, & \text{если } 2m < k < 3m, \\ m + k + (k - [k/m]m)2^{\log([k/m]+1)[} \\ \quad + (m - k + [k/m]m)2^{\log([k/m])[}, & \text{если } k \geq 3m. \end{cases}$$

Пусть $L(l) = L_1(l)$, где $L_1(l)$ — функция, определяемая соотношением (8).

Тогда

$$1 < T(I, \mathcal{F}, s(k, m)) \leq \frac{c}{m}((k - [k/m]m)L([k/m] + 1) + (m - k + [k/m]m)L([k/m])) + 1.$$

В частности, если $c_0 = \max(c, 1)$, то

$$1 < T(I, \mathcal{F}, s(k, [c_0 k])) < 2$$

и $T(I, \mathcal{F}) \rightarrow 1$ при $k \rightarrow \infty$.

Перейдем к рассмотрению задач о близости. В отличие от ЗИП в задачах о близости (ЗБ) отношение поиска ρ задается на $X \times V$.

Рассмотрим следующую задачу о близости. Пусть на множестве записей Y задано отношение линейного порядка $\prec$. Пусть множество запросов $X = Y$ и $V = \{y_1, \dots, y_k\} \subseteq Y$. Отношение поиска ρ задается на $X \times V$ соотношением

$$x\rho y \iff (y \in V) \& (x \prec y) \& (\neg(\exists y')((y' \in V) \& (x \prec y') \& (y' \prec y))),$$

т.е. $x\rho y$, если $y \in V$ — ближайший справа к x элемент.

При выполнении этих условий ЗИП $I = \langle X, V, \rho \rangle$ назовем первой задачей о близости.

Пусть базовое множество

$$\mathcal{F} = \langle \emptyset, G_1 \cup G_2 \rangle, \quad (9)$$

где G_1 и G_2 определяются соотношениями (1), (2), (4), (5).

Пусть

$$L_2(l) = \begin{cases} l, & \text{если } 0 \leq l \leq 3, \\ \log(l+1) + 1, & \text{если } l \geq 3. \end{cases} \quad (10)$$

Теорема 2. Пусть $I = \langle X, V, \rho \rangle$ — первая задача о близости, где $|V| = k$. Пусть $\mathcal{F}$ — базовое множество, определяемое соотношением (9), $s(k, m) = 2k + m$, $L(l) = L_2(l)$, где $L_2(l)$ — функция, определяемая соотношением (10).

Тогда справедливо утверждение теоремы 1.

Аналогично решается задача, когда надо найти ближайшую к запросу слева запись в библиотеке, т.е. когда ρ определяется соотношением

$$x\rho y \iff (y \in V) \& (y \prec x) \& (\neg(\exists y')((y' \in V) \& (y' \prec x) \& (y \prec y'))).$$

ЗИП $I = \langle X, V, \rho \rangle$ с таким отношением поиска назовем второй задачей о близости.

Рассмотрим следующую ЗИП, которая описывает n -мерную задачу интервального поиска. Пусть

$$\begin{aligned} Y &= [0, 1]^n, & V &= \{\tilde{y}^1, \dots, \tilde{y}^k\} \subseteq Y, \\ X &= \{\tilde{x} = (u_1, v_1, \dots, u_n, v_n) : 0 \leq u_i \leq v_i \leq 1, i = 1, \dots, n\}. \end{aligned} \quad (11)$$

Пусть на множестве запросов X задано вероятностное распределение, т.е. определено вероятностное пространство $\langle X, \sigma, \mathbf{P} \rangle$, где $\mathbf{P}$ задается функцией плотности вероятности $p(x)$.

Отношение поиска ρ определено на $X \times Y$ соотношением

$$(u_1, v_1, \dots, u_n, v_n)\rho(y_1, \dots, y_n) \iff u_i \leq y_i \leq v_i, i = 1, \dots, n. \quad (12)$$

Пусть

$$G_1 = \{g_{i,m}^1(u_1, v_1, \dots, u_n, v_n) = \min([u_i m] + 1, m) : i \in \{1, \dots, n-1\}, m \in \mathbf{N}\}, \quad (13)$$

$$G_2 = \{g_{i,m}^2(u_1, v_1, \dots, u_n, v_n) = \min([v_i m] + 1, m) : i \in \{1, \dots, n-1\}, m \in \mathbf{N}\}, \quad (14)$$

$$G_3 = \left\{ g_{i,a}^3(u_1, v_1, \dots, u_n, v_n) = \begin{cases} 1 & \text{при } u_i \leq a, \\ 2 & \text{при } u_i > a, \end{cases} : i \in \{1, \dots, n-1\}, a \in [0, 1] \right\}, \quad (15)$$

$$G_4 = \left\{ g_{i,a}^4(u_1, v_1, \dots, u_n, v_n) = \begin{cases} 1 & \text{при } v_i \leq a, \\ 2 & \text{при } v_i > a, \end{cases} : i \in \{1, \dots, n-1\}, a \in [0, 1] \right\}. \quad (16)$$

Обозначим

$$M_{a,b} = \{\tilde{x} = (u_1, v_1, \dots, u_n, v_n) \in X : u_n \leq b, v_n \geq a\}.$$

Пусть

$$F_1 = \{f_{a,b} : N_{f_{a,b}} = M_{a,b}, 0 \leq a \leq b \leq 1\}, \quad (17)$$

$$F_2 = \{\neg f_{0,a} : a \in [0, 1], f_{0,a} \in F_1\}, \quad (18)$$

$$G_5 = \left\{ g_a^5(u_1, v_1, \dots, u_n, v_n) = \begin{cases} 1, & \text{если } u_n \leq v_n \leq u_n + a, \\ 2 & \text{в противном случае,} \end{cases} : a \in [0, 1] \right\}, \quad (19)$$

$$G_6 = \{g_m^6(u_1, v_1, \dots, u_n, v_n) = \max(1, \lfloor u_n(m+1) \rfloor) : m \in \mathbf{N}\}. \quad (20)$$

Пусть

$$\mathcal{F} = \langle F_1 \cup F_2, G_1 \cup G_2 \cup G_3 \cup G_4 \cup G_5 \cup G_6 \rangle. \quad (21)$$

Теорема 3. Пусть ЗИП $I = \langle X, V, \rho \rangle$ — n -мерная задача интервального поиска, определяемая отношениями (11)–(12), $\mathcal{F}$ — базовое множество, определяемое соотношениями (13)–(21). Пусть

$$R(I) = \sum_{y \in V} \mathbf{P}(O(y, \rho)).$$

Тогда, если существует такая постоянная c , что функция плотности вероятности $p(x) \leq c$, то

$$R(I) < T(I, \mathcal{F}, (4k + 3 + (1 + 6) \lfloor \log k \rfloor) c_0) (k(k+1)/2)^{n-1} \leq R(I) + 4n + 1,$$

где $c_0 = \max(1, c/2^{n-1})$.

Эта теорема была анонсирована в [4].

3. Поиск идентичных объектов

В этом разделе исследуется задача поиска идентичных объектов и приводится доказательство теоремы 1. Докажем сначала следующее вспомогательное утверждение.

Лемма 1. Пусть $L_1(l)$ и $L_2(l)$ — функции, определенные выше, $k, m \in \mathbb{N}$, и пусть

$$r_j(k, m) = \max \left\{ \sum_{i=1}^m L_j(l_i) : l_1, \dots, l_m \in \mathbb{N}_0, \sum_{i=1}^m l_i = k \right\}.$$

Тогда

$$r_j(k, m) = (k - [k/m]m)L_j([k/m] + 1) + [k/m]mL_j([k/m]).$$

Доказательство. Если доопределить функции $L_1(l)$ и $L_2(l)$ на положительную полуось числовой прямой, полагая, например,

$$L_1(x) = \begin{cases} x, & \text{если } 0 \leq x \leq 4, \\ \log x + 2, & \text{если } x \geq 4, \end{cases} \quad (22)$$

$$L_2(x) = \begin{cases} x, & \text{если } 0 \leq x \leq 3, \\ \log(x+1) + 1, & \text{если } x \geq 3, \end{cases} \quad (23)$$

то полученные функции будут непрерывными и вогнутыми. И, вообще говоря, вогнутость этих функций объясняет результат леммы.

Более подробное доказательство будем вести от противного. Предположим, что лемма не верна, т.е.

$$r_j(k, m) = \sum_{i=1}^m L_j(l'_i) > (k - [k/m]m)L_j([k/m] + 1) + [k/m]mL_j([k/m]), \quad j = 1, 2, \quad (24)$$

и среди чисел l'_i , $i = 1, \dots, m$, существует 2 числа, разность которых не меньше двух.

Без ограничения общности можем считать, что $l'_1 - l'_2 \geq 2$. Пусть

$$l''_1 = \lfloor (l'_1 + l'_2)/2 \rfloor, \quad l''_2 = \lceil (l'_1 + l'_2)/2 \rceil, \quad l'_1 + l''_2 = l'_1 + l'_2, \quad l''_1 - l''_2 \leq 1.$$

Так как функция $L_j(x)$ вогнутая, то по свойству вогнутых функций

$$L_j(l''_1) + L_j(l''_2) \leq L_j(l'_1) + L_j(l'_2), \quad j = 1, 2. \quad (25)$$

Если в (23) неравенство строгое, то получили противоречие, так как сумма $L_j(l'_1) + \dots + L_j(l'_m)$ не максимальна; если же нет, то обозначим $l''_i = l'_i$, $i = 3, \dots, m$ и перейдем к рассмотрению суммы

$$\sum_{i=1}^m L_j(l''_i) = \sum_{i=1}^m L_j(l'_i) = r_j(k, m), \quad j = 1, 2.$$

Если среди чисел l''_i , $i = 1, \dots, m$, нет пары чисел, разность которых превышает 1, то получим противоречие с неравенством (22), если же есть, то опять сделаем операцию, описанную выше и избавимся от такой пары, и так будем проделывать до тех пор, пока либо не получим строгое неравенство в (23), либо не придем к разбиению $l_1^{(n)}, \dots, l_m^{(n)}$, такому, что

$$\sum_{i=1}^m L_j(l_i^{(n)}) = r_j(k, m), \quad j = 1, 2,$$

и все элементы отличаются не более, чем на 1, и, тем самым, получим противоречие с неравенством (22). Лемма доказана.

Перейдем к доказательству теоремы 1. Построим сеть U_m^0 , имеющую вид дерева, следующим образом. Возьмем вершину β_0 и объявим ее корнем сети U_m^0 . Выпустим из β_0 ровно m ребер, припишем им числа от 1 до m , объявим β_0 точкой переключения и припишем ей переключатель $g_m^1(x)$. Пусть

$$V_i = X_i \cap V, \quad l_i = |V_i|, \quad i = 1, \dots, m.$$

Конец ребра с номером i обозначим β_i . Для всех i таких, что $V_i \neq \emptyset$ сделаем следующую процедуру. Выпустим из вершины β_i бинарное сбалансированное дерево высоты $\lceil \log l_i \rceil$ с l_i концевыми вершинами. Из концевых вершин, высота которых меньше $\lceil \log l_i \rceil$ (если такие есть) выпустим одно ребро. В полученном дереве D_i каждая концевая вершина будет иметь высоту $\lceil \log l_i \rceil$. Объявим концевые вершины этого дерева листьями и сопоставим им слева направо в порядке возрастания элементы из V_i .

Пусть β — произвольная внутренняя вершина дерева D_i . Обозначим через V_β множество записей, соответствующих листьям ветви, растущей из β . Пусть β' — вершина, в которую ведет левое (если оно одно, то единственное) ребро из β . Пусть

$$y_\beta = \max_{y \in V_{\beta'}} y.$$

Объявим все внутренние вершины дерева D_i с высотой, не превышающей $\lceil \log l_i \rceil - 2$, вершинами переключения и для каждой такой вершины β левому ребру, из нее выходящему, припишем 1, а правому 2, самой вершине припишем переключатель $g_{y_\beta}^2(x)$.

Все ребра дерева D_i , выходящие из вершин с высотой $\lceil \log l_i \rceil - 1$, ведут в листья; припишем каждому такому ребру, ведущему в лист с записью y , предикат $f_y(x)$.

Полученную сеть с $|V| = k$ листьями обозначим U_m^0 . Это сеть над базовым множеством $\mathcal{F}$. Покажем, что U_m^0 разрешает задачу $I = \langle X, V, \rho \rangle$. Возьмем произвольный лист α . Пусть ему приписана запись $y \in V$. Достаточно показать, что $\varphi_\alpha(x) = f_y(x)$. Так как в лист α ведет предикатное ребро с предикатом $f_y(x)$,

$$N_{\varphi_\alpha(x)} \subseteq N_{f_y(x)} = \{y\}.$$

Покажем, что $\varphi_\alpha(y) = 1$. Пусть $y \in V_i$, $i \in \{1, \dots, m\}$, т. е. лист α принадлежит дереву D_i . В него ведет единственная цепь, состоящая из $\lceil \log l_i \rceil + 1$ ребер, причем все эти ребра, кроме последнего, переключательные. Так как $g_m^2(y) = i$ и так как первое ребро в этой цепи есть i -е ребро, исходящее из корня, проводимость первого ребра в цепи равна 1. Последнее ребро в цепи предикатное с предикатом $f_y(x)$ и его проводимость равна $f_y(y) = 1$. Покажем, что проводимость остальных $\lceil \log l_i \rceil - 1$ ребер цепи равна 1. Возьмем произвольное одно из этих ребер (β, β') . Если (β, β') левое, исходящее из β , то $y \in V_{\beta'}$ и предикат $g_{y_\beta}^2(y)$, приписанный вершине β , равен 1, так как

$$y \leq y_\beta = \max_{y' \in V_{\beta'}} y'.$$

Если (β, β') — правое ребро, исходящее из β , то $y > y_\beta$ и $g_{y_\beta}^2(y) = 2$, тем самым проводимость ребра (β, β') в обоих случаях равна 1.

Таким образом, мы показали, что $\varphi_\alpha(x) = f_y(x)$. Учитывая произвольность листа α и теорему 1 из [3], получим, что сеть U_m^0 решает задачу $I = \langle X, V, \rho \rangle$.

Объем сети U_m^0

$$\begin{aligned} Q(U_m^0) &= m + \sum_{1 \leq i \leq m, l_1, \dots, l_m \leq 2} l_i + \sum_{1 \leq i \leq m, l_1, \dots, l_m > 2} (l_i + 2^{\lfloor \log l_i \rfloor} - 2) \\ &= m + k + \sum_{1 \leq i \leq m, l_1, \dots, l_m > 2} (2^{\lfloor \log l_i \rfloor} - 2). \end{aligned}$$

Подсчитаем сложность полученной сети U_m^0 . Рассмотрим произвольный запрос $x \in X_i$. При $l_i = 0$, дерево D_i пусто и $T(U_0, x) = 1$. Если $l_i > 0$, то, как мы показали выше, активные на этом запросе вершины сети U_0^m (вершина β активна на запросе, если $\varphi_\beta(x) = 1$) образуют единственную цепь, ведущую из корня к вершине с искомой записью, если искомая запись есть в библиотеке, или некую подцепь этой цепи, если искомой записи нет. Причем в этой цепи не более, чем $\lfloor \log l_i \rfloor$ переключательных вершин и одна вершина, из которой исходит не более двух ребер с предикатами. Таким образом,

$$T(U_0^m) \leq 2 + \lfloor \log l_i \rfloor \leq 1 + L_1(l_i),$$

где $L_1(l)$ — функция, определяемая соотношением (8).

Учитывая определение и лемму 1, находим, что

$$\begin{aligned} T(U_0^m) &= \mathbf{M}T(U_0^m, x) = \int_X T(U_0^m, x) \mathbf{P}(dx) \\ &= \sum_{i=1}^m \int_{X_i} T(U_0^m, x) \mathbf{P}(dx) \leq \sum_{i=1}^m (1 + L_1(l_i)) \mathbf{P}(X_i) \\ &= 1 + \sum_{i=1}^m L_1(l_i) \mathbf{P}(X_i) \leq \frac{c}{m} \sum_{i=1}^m L_1(l_i) \leq 1 + r(k, m) \frac{c}{m} \\ &= \frac{c}{m} ((k - [k/m]m)L([k/m] + 1) + (m - k + [k/m]m)L([k/m])) + 1. \end{aligned}$$

Пусть $c_0 = \max(c, 1)$. Возьмем $m = [c_0 k]$. Тогда $m \geq k$. Если $m = k$, то $k/m = 1$ и

$$r(m, k) = 0L_1(2) + kL_1(1) = kL_1(1).$$

Если $m > k$, то $[k/m] = 0$ и

$$r(m, k) = kL_1(1) + (m - k)L_1(0) = kL_1(1).$$

Следовательно,

$$T(I, \mathcal{F}) \leq T(U_0^m) \leq 1 + \frac{ck}{[c_0 k]} L_1(1) < 2.$$

Если взять $m = k\alpha(k)$, где $\alpha(k) \rightarrow \infty$ при $k \rightarrow \infty$ и $\alpha(k) > 1$ при любом k , то

$$T(I, \mathcal{F}) \leq 1 + \frac{ck}{k\alpha(k)} \approx 1.$$

С другой стороны, $T(I, \mathcal{F}) \geq 1$, так как из корня должно исходить хотя бы одно ребро, поскольку $k > 0$. Теорема доказана.

Приведем два примера задачи поиска идентичных объектов.

Пример 1. Рассмотрим следующую ЗИП. Пусть

$$Y = [0, 1], \quad X = [0, 1], \quad V = \{y_1, \dots, y_k\} \subseteq Y.$$

Отношение поиска ρ есть отношение идентичности, т. е.

$$x\rho y \iff x = y.$$

Пусть определено вероятностное пространство $\langle X, \sigma, \mathbf{P} \rangle$, и $\mathbf{P}$ задается плотностью вероятности $p(x)$. Пусть

$$\begin{aligned} X_i &= \{x \in X : (i-1)/m \leq x < i/m\}, \quad i = 1, \dots, m-1, \\ X_m &= \{x \in X : (m-1)/m \leq x \leq 1\}. \end{aligned} \quad (26)$$

Тогда переключатель $g_m^1(x) = \min([xm] + 1, m)$ удовлетворяет соотношению (1).

Если $p(x) \leq c$, где c — некоторая постоянная, то $\mathbf{P}(X_i) \leq c/m$, и мы находимся в условиях теоремы 1.

Приведем неформальное описание алгоритма поиска, приведенного в доказательстве теоремы 1, применительно к этой задаче. Разобьем отрезок $[0, 1]$ на m равных частей в соответствии с соотношениями (24). Каждой части сопоставим подмножество множества V , состоящее из точек, принадлежащих этой части. Теперь для произвольной точки-запроса x поиск точки из V , идентичной запросу, будем вести следующим образом.

Определим ту часть отрезка, которой принадлежит запрос x . Ее номер равен $\min([xm] + 1, m)$. Теперь во множестве, сопоставленном найденной части, осуществим обычный дихотомический поиск. Согласно лемме 1 наибольшее среднее время поиска будет в том случае, когда точки из множества V равномерно распределены по всем m частям отрезка. Если в качестве m взять $m = k$, то случай, когда в каждую часть попадет по одной точке из V , будет иметь наибольшую сложность. А поскольку найти или не найти объект во множестве мощности 1 можно за 1 шаг, то в среднем за 2 шага (на первом мы определили номер нужной части) мы можем решить задачу поиска идентичных объектов.

Пример 2. Пусть $X = Y = \{1, \dots, N\}$. Отношение поиска ρ есть отношение идентичности. Пусть $\langle X, \sigma, \mathbf{P} \rangle$ — вероятностное пространство над X , где σ — множество всех подмножеств, а вероятностная мера $\mathbf{P}$ определяет равномерную вероятность на множестве запросов X , т. е. $\mathbf{P}(x) = 1/N$ для любого $x \in X$. Пусть $m \in \mathbf{N}$, $r = N - m[N/m]$ и пусть $X_1, \dots, X_m$ — разбиение множества X такое, что

$$\begin{aligned} X_i &= \{x \in X : 1 + (i-1)([N/m] + 1) \leq x \leq i([N/m] + 1)\}, \\ &\quad i = 1, \dots, r, \\ X_i &= \{x \in X : r([N/m] + 1) + 1 + (i-1-r)[N/m] \leq x \leq r([N/m] + 1) + (i-r)[N/m]\}, \\ &\quad i = r+1, \dots, m, \end{aligned}$$

и $g_m^1(x) = i$, если $x \in X_i$, $i \in \{1, \dots, m\}$. Тогда

$$\mathbf{P}(X_i) \leq ([N/m] + 1)/N < 2/m,$$

и мы опять находимся в условиях теоремы 1 при $c = 2$.

4. Задачи о близости

В данном разделе приводится доказательство теоремы 2. Построим сеть U_m^1 по аналогии с сетью U_m^0 . Возьмем вершину β_0 и объявим ее корнем сети. Выпустим из β_0 ровно m ребер, припишем им числа от 1 до m , объявим β_0 точкой переключения и припишем ей переключатель $g_m^1(x)$. Пусть

$$V_i = X_i \cap V, \quad l_i = |V_i|, \quad i = 1, \dots, m.$$

Конец ребра с номером i обозначим β_i .

Для всех i таких, что $V_i \neq \emptyset$, сделаем следующую процедуру. Выпустим из вершины β_i бинарное сбалансированное дерево D_i высоты $\lceil \log(l_i + 1) \rceil$ с $l_i + 1$ концевыми вершинами. Объявим все концевые вершины этого дерева D_i , кроме последней (самой правой), листьями и сопоставим им слева направо в порядке возрастания элементы из V_i .

Для произвольной внутренней вершины β дерева D_i введем понятия V_β и y_β , определяемые так же, как и в сети U_m^0 .

Объявим все внутренние вершины дерева D_i вершинами переключения и для каждой внутренней вершины β левому ребру, из нее выходящему, припишем 1, а правому 2, самой вершине припишем переключатель $g_{y_\beta}^2(x)$.

Пусть $i \in \{1, \dots, m\}$. Обозначим $j(i)$ такой номер, что $j(i) > i$, $|V_{j(i)}| > 0$ и не существует j' такого, что $|V_{j'}| > 0$, $j' > i$ и $j' < j(i)$, т. е. $j(i)$ — индекс ближайшего сверху непустого множества $V_{j(i)}$. Если такого множества нет, то $j(i) = 0$.

Теперь для каждого дерева D_i самую правую концевую вершину дерева отождествим с самым левым листом дерева $D_{j(i)}$, если $j(i) \neq 0$. Для каждого i такого, что $l_i = 0$, вершину β_i отождествим с самым левым листом дерева $D_{j(i)}$, если $j(i) \neq 0$. Полученная сеть и будет сетью U_m^1 .

Доказательство того, что сеть U_m^1 разрешает первую задачу о близости $I = \langle X, V, \rho \rangle$ аналогично доказательству того, что сеть U_m^0 разрешает задачу поиска идентичных объектов. Объем сети U_m^1

$$Q(U_m^1) = m + \sum_{i=1}^m (2(l_i + 1) - 2) = m + 2 \sum_{i=1}^m l_i = m + 2k.$$

Подсчитаем сложность сети U_m^1 . Рассмотрим произвольный запрос $x \in X$. При $l_i = 0$ сложность $T(U_m^1, x) = 1$. Если $l_i > 0$, то активные на запросе x вершины сети U_m^1 образуют единственную цепь, ведущую из корня к вершине с искомой записью или же к самой правой вершине последнего дерева D_i . В этой цепочке не более $1 + \lceil \log(l_i + 1) \rceil$ переключательных вершин, и значит,

$$T(U_m^1, x) \leq 1 + \lceil \log(l_i + 1) \rceil.$$

Отсюда следует, что

$$T(U_m^1, x) \leq 1 + L_2(l_i).$$

Теперь утверждение теоремы 2 доказывается аналогично теореме 1 с использованием леммы 1.

5. Многомерная задача интервального поиска

Приведем доказательство теоремы 3. Обозначим

$$\begin{aligned}\tilde{x} &= (u_1, v_1, \dots, u_n, v_n), & \tilde{z}_i &= (u_i, v_i), \\ X_1 &= \{(u, v) : 0 \leq u \leq v \leq 1\}, \\ p_i(u, v) &= \int_{X_1} \dots \int_{X_1} p(\tilde{x}) d\tilde{z}_1 \dots d\tilde{z}_{i-1} d\tilde{z}_{i+1} \dots d\tilde{z}_n, \\ p_i^1(u) &= \int_u^1 p_i(u, v) dv, & p_i^2(v) &= \int_0^v p_i(u, v) du.\end{aligned}$$

Легко видеть, что

$$\begin{aligned}p_i(u, v) &\leq c \int_{X_1} \dots \int_{X_1} d\tilde{z}_1 \dots d\tilde{z}_{i-1} d\tilde{z}_{i+1} \dots d\tilde{z}_n = \frac{c}{2^{n-1}}, \\ p_i^1(u) &\leq \frac{c(1-u)}{2^{n-1}} \leq \frac{c}{2^{n-1}}, & p_i^2(v) &\leq \frac{cv}{2^{n-1}} \leq \frac{c}{2^{n-1}}.\end{aligned}$$

Решать задачу будем методом снижения размерности. Покажем, как можно перейти от решения n -мерной задачи интервального поиска к решению задачи размерности $n-1$.

Пусть $S \subseteq V$, и $1 \leq i_1 < \dots < i_l \leq n$. Обозначим

$$\Pi^{i_1, \dots, i_l}(S) = \{(y_{i_1}, \dots, y_{i_l}) : (y_1, \dots, y_n) \in S\}$$

проекцию множества S на компоненты $i_1, \dots, i_l$. Положим

$$W^i = \{(y', y'') : y', y'' \in \Pi^i(V), y' \leq y''\}, \quad i = 1, \dots, k.$$

Очевидно, что $|W^i| \leq k(k+1)/2$. Обозначим

$$Z^i = \{(y_1^1, y_2^1, \dots, y_i^1, y_i^2) : y_1^j, y_2^j \in \Pi^i(V), y_1^j \leq y_2^j, j = 1, \dots, i\}, \quad i = 1, \dots, k.$$

Для каждой пары $(y', y'') \in W^i$ определим множество

$$S_{y', y''}^i = \{\tilde{y} = (y_1, \dots, y_n) \in V : y' \leq y_i \leq y''\}.$$

Введем обозначения

$$V^1 = V, \quad M^1 = \Pi^1(V^1), \quad M_y^1 = \{y' \in V^1 : y' \geq y\},$$

$$V^i(y_1^1, y_2^1, \dots, y_{i-1}^{i-1}, y_2^{i-1}) = \bigcap_{j=1}^{i-1} S_{y_1^j, y_2^j}^j, \quad i = 2, \dots, n,$$

$$M^i(y_1^1, y_2^1, \dots, y_{i-1}^{i-1}, y_2^{i-1}) = \Pi^i(V^i(y_1^1, y_2^1, \dots, y_{i-1}^{i-1}, y_2^{i-1})), \quad i = 2, \dots, n,$$

$$M_y^i(y_1^1, y_2^1, \dots, y_{i-1}^{i-1}, y_2^{i-1}) = \{y' \in M^i(y_1^1, y_2^1, \dots, y_{i-1}^{i-1}, y_2^{i-1}) : y' \geq y\}, \quad i = 2, \dots, n.$$

Тогда n -мерную задачу интервального поиска можно решать следующим образом.

Пусть $\tilde{x} = (u_1, v_1, \dots, u_n, v_n) \in X$ — произвольный запрос. Сначала найдем такую пару (y', y'') из W^1 , что y' — ближайшая справа к u_1 точка из множества M^1 , а y'' — ближайшая слева к v_1 точка из $M_{y'}^1$. Если такой пары нет,

то ответ на запрос $\tilde{x}$ пуст, если же есть, то мы можем перейти к решению для запроса $\tilde{x}' = (u_2, v_2, \dots, u_n, v_n)$ в $(n-1)$ -мерной задаче интервального поиска в множестве $\Pi^{y_2, \dots, y_n}(V^2(y', y''))$, так как множество $V^2(y', y'')$ содержит все точки $\tilde{y} = (y_1, \dots, y_n) \in V$ такие, что $u_1 \leq y_1 \leq v_1$.

Таким способом мы за $n-1$ шагов придем к одномерной задаче интервального поиска, которую будем решать методом, описанным в [3].

Опишем ИСП, которая решает задачу I описанным выше методом. Сначала построим ИСП U_m^1 , решающую первую задачу о близости для множества M^1 так, как было описано в доказательстве теоремы 2, где в качестве параметра m возьмем $m = \lfloor c_0 k \rfloor$. Тогда $Q(U_m^1) = 2k + \lfloor c_0 k \rfloor$ и $T(U_m^1) < 2$.

Заменим в сети U_m^1 переключатель g_m^1 на $g_{1,m}^1$, а все переключатели вида $g_{y_\beta}^2$ на g_{1,y_β}^3 . Возьмем произвольный лист α сети U_m^1 . Пусть ему соответствует точка $y \in M^1$. Объявим α внутренней вершиной и уберем приписанную ему точку y .

Построим сеть $U_m^{1,y}$, решающую вторую задачу о близости для множества M_y^1 (отметим, что $M_y^1 \neq \emptyset$), где в качестве параметра m возьмем $m = \lfloor c_0 |M_y^1| \rfloor$. Тогда

$$Q(U_m^{1,y}) = 2|M_y^1| + \lfloor c_0 |M_y^1| \rfloor \leq |M_y^1|(2 + c_0), \quad T(U_m^{1,y}) < 2.$$

Заменим в сети $U_m^{1,y}$ переключатель g_m^1 на $g_{1,m}^2$, а все переключатели вида $g_{y_\beta}^2$ на g_{1,y_β}^4 . Теперь отождествим корень сети $U_m^{1,y}$ с вершиной α , т. е. сеть $U_m^{1,y}$ теперь будет расти из α , причем α не будет полюсом полученной сети. И наконец, для каждого листа α' сети $U_m^{1,y}$ заменим приписанную листу α' точку y' на пару (y, y') . Эту пару будем воспринимать как обозначение множества $V^2(y, y')$. Прделаем такую операцию для каждого листа сети $U_m^{1,y}$. Полученную сеть обозначим U_1 .

Сеть U_1 имеет не более $k(k+1)/2$ листьев, которым взаимно однозначно сопоставлены пары из Z^1 . Эта сеть для любого запроса $\tilde{x} = (u_1, v_1, \dots, u_n, v_n) \in X$ позволяет находить пару $(y', y'') \in Z^1$ такую, что y' — ближайшая справа к u_1 точка из M^1 , а y'' — ближайшая слева к v_1 точка из $M_{y'}^1$, если, конечно, такая пара существует, т. е. по сути позволяет находить множество точек $V^2(y', y'') \subseteq V$, которые удовлетворяют первой паре неравенств из (12).

Так как в сети U_m^1 только один активный путь,

$$T(U^1) = T(U_m^1) + \max_y T(U_m^{1,y}) \leq 2 + 2 = 4.$$

Легко видеть, что

$$Q(U^1) \leq (2 + c_0)k + (2 + c_0) \sum_{y \in V^1} |M_y^1| = k(k+3)(2 + c_0)/2.$$

Рассмотрим произвольный лист α сети U^1 . Пусть ему соответствует пара $(y, z) \in W^1$. Отметим, что $V^2(y, z)$ не пусто. Теперь тем же методом, которым на предыдущем шаге мы для множества V^1 строили сеть U^1 , построим сеть для множества $V^2(y, z)$. Заменим в этой сети все переключатели вида $g_{1,m}^1, g_{1,y_\beta}^3, g_{1,m}^2, g_{1,y_\beta}^4$ на переключатели $g_{2,m}^1, g_{2,y_\beta}^3, g_{2,m}^2, g_{2,y_\beta}^4$. Заменим также нагрузку листьев этой сети так, что если листу приписана пара (y', y'') , то припишем ей четверку (y, z, y', y'') . Обозначим полученную сеть $U_{y,z}^2$. Теперь объявим α внутренней вершиной и уберем приписанную ему пару (y, z) и отождествим α с корнем сети

$U_{y,z}^2$, т. е. сеть $U_{y,z}^2$ будет расти из α . Прделаем такую операцию для каждого листа сети U^1 . Полученную сеть обозначим U^2 .

Сеть U^2 имеет не более $(k(k+1)/2)^2$ листьев, которым соответствуют четверки из множества Z^2 , причем каждая четверка $(y_1^1, y_2^1, y_1^2, y_2^2)$ воспринимается как обозначение множества $V^3(y_1^1, y_2^1, y_1^2, y_2^2)$.

Легко видеть, что $T(U^2) \leq 8$ и

$$\begin{aligned} Q(U^2) &\leq Q(U^1) + (2 + c_0) \sum_{(y_1^1, y_2^1) \in Z^1} \left(|M^2(y_1^1, y_2^1)| + \sum_{y \in M^2(y_1^1, y_2^1)} |M_y^2(y_1^1, y_2^1)| \right) \\ &\leq (2 + c_0) \left(k + \frac{k(k+1)}{2} \right) + (2 + c_0) \left(k \frac{k(k+1)}{2} + \left(\frac{k(k+1)}{2} \right)^2 \right) \\ &\leq (2 + c_0) \left(\frac{k(k+1)(k+2)}{2} + \left(\frac{k(k+1)}{2} \right)^2 \right). \end{aligned}$$

Далее из U^2 аналогичным образом получим сеть U^3 и т.д.

На $(n-1)$ -ом шаге мы получим сеть U^{n-1} , которая будет иметь не более $(k(k+1)/2)^{n-1}$ листьев, которым соответствуют векторы из множества Z^{n-1} . Эта сеть позволяет для любого запроса $\tilde{x} = (u_1, v_1, \dots, u_n, v_n)$ находить вектор $(y_1^1, y_2^1, \dots, y_1^{n-1}, y_2^{n-1}) \in Z^{n-1}$ такой, что y_i^i — ближайшая справа к u_i точка из $M^i(y_1^1, y_2^1, \dots, y_1^{i-1}, y_2^{i-1})$, а y_i^i — точка из $M_{y_i^i}^1(y_1^1, y_2^1, \dots, y_1^{i-1}, y_2^{i-1})$, ближайшая слева к v_i , $i = 1, \dots, n-1$, если, конечно, такой вектор существует. Если же такой вектор не существует, то ответ на запрос будет пуст, т.е. процесс поиска прервется, не доходя листьев. Если напомнить, что каждый вектор $(y_1^1, y_2^1, \dots, y_1^{n-1}, y_2^{n-1})$ воспринимается как обозначение множества $V^n(y_1^1, y_2^1, \dots, y_1^{n-1}, y_2^{n-1})$, то сеть U^{n-1} позволяет находить для любого запроса $\tilde{x}$ подмножество точек множества V , которые удовлетворяют первым $n-1$ парам неравенств из (12).

Легко видеть, что $T(U^{n-1}) \leq 4(n-1)$ и

$$\begin{aligned} Q(U^{n-1}) &\leq Q(U^{n-2}) + (2 + c_0) \left(k \left(\frac{k(k+1)}{2} \right)^{n-2} + \left(\frac{k(k+1)}{2} \right)^{n-1} \right) \\ &\leq (2 + c_0) \left((k+2) \left(\frac{k(k+1)}{2} \right)^{n-2} + \left(\frac{k(k+1)}{2} \right)^{n-1} \right). \end{aligned}$$

Теперь осталось сделать последний шаг.

Рассмотрим произвольный лист α сети U_{n-1} . Пусть ему приписан вектор $(y_1^1, y_2^1, \dots, y_1^{n-1}, y_2^{n-1})$, который рассматривается как обозначение множества $V^n(y_1^1, y_2^1, \dots, y_1^{n-1}, y_2^{n-1})$. Построим для этого множества сеть U_α , решающую одномерную задачу интервального поиска по последней координате y_n . Эту сеть построим по методу, описанному в доказательстве теоремы 9 из [3], используя функции и предикаты из множеств F_1, F_2, G_5, G_6 . Если обозначить

$$l = |V^n(y_1^1, y_2^1, \dots, y_1^{n-1}, y_2^{n-1})|,$$

то число ребер этой сети не превышает

$$4l - 2 + 3 \cdot 2] \log l[c/2^{n-1} \leq 4k + 6c] \log k[2^{n-1}.$$

Отметим также, что $T(U_\alpha) \leq R(I) + 5$.

Объявим теперь лист α внутренней вершиной, уберем приписанный ему вектор и прикрепим к нему сеть U_α . Прделаем эту операцию для каждого листа сети U_{n-1} и полученную сеть обозначим U_n .

Так как в сети U_{n-1} не более $(k(k+1)/2)^{n-1}$ листьев,

$$Q(U_n) \leq (4k + 3 + (1 + 6] \log k]c_0)(k(k+1)/2)^{n-1}.$$

Легко видеть, что полученная сеть U_n разрешает задачу I , так как для любого запроса $\tilde{x} = (u_1, v_1, \dots, u_n, v_n)$ с помощью сети U^{n-1} мы находим подмножество точек множества V , которые удовлетворяют первым $n-1$ парам неравенств из (12), а с помощью сети, растущей из вершины подсети U^{n-1} , соответствующей этому подмножеству, мы выберем из этого подмножества точки, удовлетворяющие и последней паре неравенств из (12).

Поскольку для любого запроса $\tilde{x}$ существует только один активный путь, ведущий к конечным вершинам подсети U^{n-1} , это означает, что для любого запроса $\tilde{x}$ будет активна только одна из подсетей U_α , и значит,

$$T(U^n) \leq 4(n-1) + 5 + R(I) = 4n + 1 + R(I).$$

Согласно тривиальной нижней оценке (см. теорему 2 в [3])

$$T(I, \mathcal{F}) \geq R(I),$$

поэтому теорема полностью доказана.

Список литературы

1. Гасанов Э. Э. Об одной математической модели информационного поиска. *Дискретная математика* (1991) 3, №2, 69–76.
2. Гасанов Э. Э. Оптимальные информационные сети для отношений поиска, являющихся отношениями линейного квазипорядка. В кн.: *Конструкции в алгебре и логике*. Изд-во Тверского университета, Тверь, 1990, с. 11–17.
3. Гасанов Э. Э. Об одномерной задаче интервального поиска. *Дискретная математика* (1995) 7, №2, 40–60.
4. Гасанов Э. Э., Ерохин А. Н. О быстром в среднем решении n -мерной задачи интервального поиска. В кн.: *Методы и системы технической диагностики. Тезисы X международной конференции по проблемам теоретической кибернетики*. Изд-во Саратовского университета, Саратов, 1993, с. 51–52.