

ИССЛЕДОВАНИЕ АЛГОРИТМОВ КОРРЕКЦИИ ЛОКАЛЬНОГО ВОЗМУЩЕНИЯ В КОНЕЧНОЙ ПОЛУМЕТРИКЕ

© Майсурадзе А.И., Громов И.А.

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМ. М.В. ЛОМОНОСОВА
Ф-Т ВМК, ЛЕНИНСКИЕ ГОРЫ, Г. МОСКВА, 119991, РОССИЯ
E-MAIL: maysuradze@cs.msu.su
ВЫЧИСЛИТЕЛЬНЫЙ ЦЕНТР ИМ. А.А. ДОРОДНИЦЫНА РАН
УЛ. ВАВИЛОВА, 40, Г. МОСКВА, 119333, РОССИЯ
E-MAIL: Igor_Gromov@mail.ru

Abstract. A three-stage scheme for synthesis of semimetric adjustment algorithms in data-mining problems is proposed. Within its framework algorithms designed to modify the distance between a single pair of objects at the given value are studied. These adjustment algorithms guarantee preservation of the semimetric properties. The proposed adjustment procedures constructed within the three-stage scheme framework are completed within the first two stages, and in the special case – within the first stage. The experimental results are considered; the estimate of the algorithms' computational complexity is given.

ВВЕДЕНИЕ

В настоящее время в интеллектуальном анализе данных широко применяются метрические методы. Эффективность их использования существенно зависит от выбора функции сходства (например, полуметрики) на объектах распознавания. Как правило, в прикладных задачах нельзя ввести некоторую единственную «объективную» полуметрику, следовательно, сами полуметрики становятся предметом анализа и настройки. В данной работе настройка полуметрики состоит в коррекции расстояний между объектами, выбранными экспертом в предметной области.

Пусть дано конечное множество объектов с заданной на нем полуметрикой ρ . Эксперт по своему усмотрению изменяет расстояние между ровно *одной* парой объектов, при этом могут быть нарушены неравенства треугольника (такое изменение расстояния будем называть *локальным возмущением*). Интерпретацию введенного изменения эксперт задает путем выбора функционала различия полуметрик. В результате внесения локального возмущения возникает новая функция расстояния – ρ' . Задача коррекции полуметрики состоит в синтезе новой полуметрики $\tilde{\rho}$, которая близка к исходной ρ , но сохраняет сделанные изменения ρ' .

В работе изучаются некоторые алгоритмы решения данной задачи. Общие требования, предъявляемые к этим алгоритмам, следующие.

- (R1). Алгоритмы должны быть *универсальны*, т. е. применимы ко всем ρ и ρ' .
- (R2). Эксперт может выбирать различные интерпретации вносимого им изменения, и алгоритмы способны учесть этот выбор. В данной работе различным интерпретациям соответствуют различные функционалы различия полуметрик.
- (R3). Алгоритмы строят полуметрику $\tilde{\rho}$, максимально «близкую» (в смысле выбранного функционала различия полуметрик) к исходной полуметрике ρ .
- (R4). Указанное экспертом значение расстояния на выбранной паре объектов сохраняется в полуметрике $\tilde{\rho}$.

Существуют различные подходы к решению поставленной задачи коррекции полуметрики ([1]-[3]). В данной работе авторами предложен новый подход – трехэтапная схема построения алгоритмов коррекции возмущенных полуметрик. В общем случае данная схема требует исследования всех троек объектов и имеет сложность $O(N^3)$, где N – общее число обрабатываемых объектов. Однако, авторами установлены достаточные условия, при выполнении которых коррекция полуметрик в рамках трехэтапной схемы завершается в ходе первых двух этапов, а в специальном случае – уже в ходе первого этапа. Таким образом, сложность понижается до $O(N^2)$ и $O(N)$ соответственно. Предложены алгоритмы, осуществляющие коррекцию в рамках данной схемы. Для анализа реальной эффективности данных алгоритмов была проведена серия экспериментов, в которых использовались три модельные полуметрики. В работе приведены результаты экспериментов, оценена вычислительная эффективность алгоритмов, проведен сравнительный анализ результатов коррекции.

В работе используется стандартное определение метрики. Пусть V – произвольное непустое множество.

Определение 1. Отображение $\rho : V \times V \rightarrow \mathbb{R}$ называется *метрикой на V* , если удовлетворяет следующим условиям:

1. $\rho(v, v) = 0, \forall v \in V$;
2. $\rho(v_1, v_2) = \rho(v_2, v_1), \forall v_1, v_2 \in V$;
3. $\rho(v_1, v_2) \geq 0, \forall v_1, v_2 \in V$;
4. $\rho(v_1, v_3) \leq \rho(v_1, v_2) + \rho(v_2, v_3), \forall v_1, v_2, v_3 \in V$;
5. $\rho(v_1, v_2) = 0 \Rightarrow v_1 = v_2, \forall v_1, v_2 \in V$.

Если отображение ρ удовлетворяет только условиям (1)–(4), то оно называется *полуметрикой на V* . Если отображение ρ удовлетворяет только условиям (1)–(3), то оно называется *расстоянием, или функцией расстояния, на V* . Условие (4) принято называть *неравенством треугольника*.

Будем рассматривать конечные множества объектов V_N мощности N , элементы которых отождествим с их индексами $1, 2, \dots, N$. Пусть задана полуметрика $\rho : V \times V \mapsto \mathbb{R}$. Матрицу попарных расстояний, определенных полуметрикой ρ , обозначим R ; $R = (r_{ij}) \in \mathbb{R}^{N \times N}$, где $r_{ij} \stackrel{\text{def}}{=} \rho(i, j)$. Когда в работе говорится, что дана функция расстояния ρ , это означает, что дана соответствующая ей матрица попарных расстояний R . Функцию расстояния, полученную в результате экспертной модификации (внесения локального возмущения), будем обозначать ρ' , а полуметрику, полученную в результате последующей коррекции – $\tilde{\rho}$. R' и \tilde{R} – матрицы попарных расстояний, соответствующие функциям ρ' и $\tilde{\rho}$. Расстояния вида r_{ii} нигде в работе рассматриваться не будут.

Во всей работе индексами i_0 и j_0 будем обозначать те два объекта, расстояние между которыми изменил эксперт: $r_{i_0 j_0} \mapsto r'_{i_0 j_0}$. Новое расстояние должно сохраниться в скорректированной полуметрике: $\tilde{r}_{i_0 j_0} = r'_{i_0 j_0}$.

Введем множество неупорядоченных пар индексов $E_N = \{(i, j) \mid i, j \in \{1, 2, \dots, N\}, i \neq j\}$ и множество неупорядоченных троек индексов $T_N = \{(i, j, k) \mid i, j, k \in \{1, 2, \dots, N\}, i, j, k \text{ попарно различны}\}$. Иными

словами, будем отождествлять пары (тройки) индексов из множества E_N (соответственно T_N), различающиеся между собой лишь порядком индексов в них. Предполагается, что в тексте статьи вновь вводимые индексы соответствуют объектам, не совпадающим с уже проиндексированными объектами.

Обозначим через P_{ij}^ρ множество (отрезок числовой оси) допустимых значений расстояния между объектами i и j в полуметрике ρ при фиксированных остальных расстояниях. Нетрудно видеть, что $P_{ij}^\rho = [\max_{k \in V_N \setminus \{i,j\}} |r_{ik} - r_{jk}|, \min_{k \in V_N \setminus \{i,j\}} (r_{ik} + r_{jk})] = [r_{ij}^{\min}, r_{ij}^{\max}]$ и $\forall (i, j) \in E_N P_{ij}^\rho \neq \emptyset$; r_{ij}^{\min} – левая граница множества P_{ij}^ρ , r_{ij}^{\max} – правая граница множества P_{ij}^ρ .

Треугольник, построенный на тройке вершин (объектов) $(i, j, k) \in T_N$ с длинами сторон, равными r_{ij}, r_{ik}, r_{jk} , будем обозначать Δijk .

В работе использована графовая интерпретация метрики, определенной на конечном множестве мощности N , как нагруженной клики на N вершинах. Множество вершин графа – это множество объектов, множество ребер – это множество попарных расстояний между объектами, а расстояния между объектами соответствуют весам (или «длинам»), приписанным ребрам.

1. ИНТЕРПРЕТАЦИЯ ИЗМЕНЕНИЯ РАССТОЯНИЯ И ФУНКЦИОНАЛЫ РАЗЛИЧИЯ ПОЛУМЕТРИК

Одновременно формализуем понятия интерпретации изменения и сходства двух полуметрик. Поскольку требуется корректировать расстояние ρ' с целью получить полуметрику $\tilde{\rho}$, минимально отличную от исходной ρ , но в то же время учитывающую изменения, внесенные экспертом, то необходимо оценить различие двух полуметрик с учетом внесенной экспертом модификации. Рассматриваются два основных подхода к интерпретации изменения одного расстояния в полуметрике: абсолютный и относительный, – а также их комбинация. Обоснование данных подходов приведено в [4]. Здесь мы лишь приведем соответствующие функционалы различия полуметрик ρ и $\tilde{\rho}$.

Абсолютный подход. «Схожими» считаются полуметрики, в которых абсолютные величины расстояний в полуметрике \tilde{R} минимально отличны от соответствующих расстояний в R : $Q_a(R, \tilde{R}) = \frac{\sum_{(k,l) \in E_N} (\tilde{r}_{kl} - r_{kl})^2}{\sum_{(k,l) \in E_N} r_{kl}^2}$.

Относительный подход применяется только для метрик. «Схожими» считаются метрики, в которых близки пропорции соответствующих расстояний. Мы будем рассматривать только случай, когда эксперт высказывает требование сохранить отношения нескорректированных расстояний в $\Delta i_0 j_0 k, k \in V_N \setminus \{i_0, j_0\}$: $\frac{\tilde{r}_{i_0 k}}{\tilde{r}_{j_0 k}} = \frac{r_{i_0 k}}{r_{j_0 k}}$. Это требование формализует функционал $Q_r(R, \tilde{R}) = \sum_{k \in V_N \setminus \{i_0, j_0\}} \left(\frac{\tilde{r}_{i_0 k}}{\tilde{r}_{j_0 k}} - \frac{r_{i_0 k}}{r_{j_0 k}} \right)^2$. Отношения величин расстояний более информативны, чем сами эти величины по отдельности, поэтому относительный подход, по-видимому, более пригоден на практике, чем абсолютный.

Взвешенный подход использует функционал: $Q_w(R, \tilde{R}) = w_a Q_a(R, \tilde{R}) + w_r Q_r(R, \tilde{R})$, где $w_a, w_r \geq 0$, $w_a + w_r = 1$. В простейшем случае значения весов w_a, w_r назначает эксперт. Однако может быть поставлена задача поиска этих весов, одно из решений которой предложено в [4]. Отметим, что на различных подмножествах множества треугольников $\{\Delta i_0 j_0 k\}$ можно применять различные значения весов w_a, w_r .

2. ТРЕХЭТАПНАЯ СХЕМА ПОСТРОЕНИЯ АЛГОРИТМОВ КОРРЕКЦИИ ПОЛУМЕТРИКИ

Пусть эксперт изменил расстояние $r_{i_0 j_0}$ и указал значение $r'_{i_0 j_0}$. В общем случае это повлечет нарушение неравенств треугольника во всех треугольниках вида $\Delta i_0 j_0 k$, $k \in V_N$, $k \notin \{i_0, j_0\}$. Будем строить алгоритмы коррекции по следующей «трехэтапной схеме».

1-й этап: коррекция тех $\Delta i_0 j_0 k$, в которых неравенства треугольника нарушены. Изменяются расстояния $r_{i_0 k}, r_{j_0 k}$. После первого этапа неравенства треугольника могут быть нарушены в $\Delta i_0 k l$ и $\Delta j_0 k l$.

2-й этап: коррекция тех $\Delta i_0 k l$ и $\Delta j_0 k l$, в которых неравенства треугольника нарушены. На втором этапе коррекции должны быть использованы такие методы, которые не вызовут новых нарушений неравенств треугольника в $\Delta i_0 j_0 k$.

3-й этап: коррекция тех $\Delta k l m$, в которых неравенства треугольника нарушены. На третьем этапе коррекции должны быть использованы такие методы, которые не вызовут новых нарушений неравенств треугольника в $\Delta i_0 j_0 k$, $\Delta i_0 k l$, $\Delta j_0 k l$, $\forall k, l, m \in T_N$, $k, l, m \notin \{i_0, j_0\}$.

Реализация каждого из этапов должна быть согласована с выбранным экспертом функционалом различия полуметрик. Тем самым будет выполнено требование (R3) о том, чтобы ρ и $\tilde{\rho}$ были максимально близки в смысле указанного функционалов.

Нарушения неравенств треугольников в процессе коррекции распространяются по полуметрике, и на каждом следующем этапе требуется преобразовать большее число расстояний, чем на предыдущем. На 1-м этапе коррекции требуется рассмотреть $O(N)$, на 2-м – $O(N^2)$, на 3-м – $O(N^3)$ треугольников.

Внесенное экспертом локальное возмущение может быть двух типов: *положительное* локальное возмущение: $r'_{i_0 j_0} > r_{i_0 j_0}$ и $r'_{i_0 j_0} > r_{i_0 k} + r_{j_0 k}$; *отрицательное* локальное возмущение: $r'_{i_0 j_0} < r_{i_0 j_0}$ и либо $r_{i_0 k} > r'_{i_0 j_0} + r_{j_0 k}$, либо $r_{j_0 k} > r'_{i_0 j_0} + r_{i_0 k}$. Каждый из этих случаев имеет свою специфику и требует применения несколько разных формул коррекции в рамках одного и того же подхода.

3. АЛГОРИТМЫ КОРРЕКЦИИ ПОЛУМЕТРИКИ \mathcal{A} И \mathcal{A}_1

В данном разделе сформулированы *универсальный* (для любого возмущения) алгоритм коррекции \mathcal{A} и алгоритм коррекции \mathcal{A}_1 для положительного локального возмущения.

Алгоритм \mathcal{A} . Пусть дана полуметрика ρ . Эксперт модифицировал в ней одно расстояние: $r_{i_0 j_0} \mapsto r'_{i_0 j_0}$ и потребовал сохранить указанное им значение $r'_{i_0 j_0}$. Кроме

того, эксперт выбрал функционал различия полуметрик (тем самым давая интерпретацию внесенного возмущения). Тогда для коррекции возникших в ρ' нарушений неравенств треугольника предлагается следующий алгоритм \mathcal{A} :

1-й этап: коррекция $\Delta i_0 j_0 k$. Формулы коррекции определяются выбором функционала различия полуметрик.

2-й этап: коррекция $\Delta i_0 k l$ и $\Delta j_0 k l$ проводится по следующему правилу: $\tilde{r}_{kl} = \alpha \tilde{r}_{kl}^{\min} + (1 - \alpha) \tilde{r}_{kl}^{\max}$, $\forall (k, l) \in E_N$, $k, l \notin \{i_0, j_0\}$, где $\tilde{r}_{kl}^{\min} = \max\{|\tilde{r}_{i_0 k} - \tilde{r}_{i_0 l}|, |\tilde{r}_{j_0 k} - \tilde{r}_{j_0 l}|\}$, $\tilde{r}_{kl}^{\max} = \min\{(\tilde{r}_{i_0 k} + \tilde{r}_{i_0 l}), (\tilde{r}_{j_0 k} + \tilde{r}_{j_0 l})\}$, $\alpha \in [0, 1]$ и α фиксировано для всех $(k, l) \in E_N$, $k, l \notin \{i_0, j_0\}$. Величину α определяет эксперт, однако может быть поставлена задача поиска оптимального α (см. [4]).

3-й этап: не требуется (доказательство см. [4]).

Алгоритм \mathcal{A}_1 . Пусть эксперт внес в полуметрику ρ положительное локальное возмущение: $r_{i_0 j_0} \mapsto r'_{i_0 j_0}$. В остальном требования аналогичны требованиям в алгоритме \mathcal{A} .

1-й этап: коррекция $\Delta i_0 j_0 k$ проводится таким образом, чтобы $\tilde{r}_{i_0 k} \geq r_{i_0 k}$, $\tilde{r}_{j_0 k} \geq r_{j_0 k}$. При этом значения $\tilde{r}_{i_0 k}$, $\tilde{r}_{j_0 k}$ должны минимизировать выбранный экспертом функционал различия полуметрик.

2-й этап: коррекция $\Delta i_0 k l$ и $\Delta j_0 k l$ осуществляется по следующему правилу:

$$\tilde{r}_{kl} = \begin{cases} r_{kl}, & \text{если } \tilde{r}_{kl}^{\min} \leq r_{kl} \leq \tilde{r}_{kl}^{\max}; \\ \tilde{r}_{kl}^{\min}, & \text{если } r_{kl} < \tilde{r}_{kl}^{\min}; \\ \tilde{r}_{kl}^{\max}, & \text{если } r_{kl} > \tilde{r}_{kl}^{\max}. \end{cases}$$

3-й этап: не требуется (доказательство см. [4]).

На 1-м этапе выполнения алгоритмов \mathcal{A} и \mathcal{A}_1 требуется рассмотреть $(N - 2)$ треугольников, на 2-м $-(N - 2)(N - 3)$ треугольников, обработка треугольников требует постоянного времени. Таким образом, сложность вычисления 2-го этапа алгоритмов \mathcal{A} , $\mathcal{A}_1 - O(N^2)$.

В формулировках алгоритмов \mathcal{A} и \mathcal{A}_1 указано, что методы коррекции на первом этапе определяются выбором функционала различия. Конкретные формулы для получения таких значений $\tilde{r}_{i_0 k}$, $\tilde{r}_{j_0 k}$, которые минимизируют функционалы Q_a , Q_r , Q_w приведены в [4]. Следует отметить, что сложность вычисления первого этапа алгоритмов \mathcal{A} и \mathcal{A}_1 в случае коррекции посредством предлагаемых формул $- O(N)$. Таким образом, алгоритмы будут иметь заявленную сложность $O(N^2)$.

Кроме того, отметим, что при $r'_{i_0 j_0} > r_{i_0 j_0}$ применение формул, минимизирующих Q_a , на первом этапе коррекции гарантировано дает полуметрику $\tilde{\rho}$. Таким образом, алгоритм, в котором на первом этапе коррекции положительного локального возмущения применяются эти формулы, имеет сложность $O(N)$, т. е. является *линейным* относительно числа объектов и не требует выполнения 2-го и 3-го этапов (см. [4]).

4. РЕЗУЛЬТАТЫ ВЫЧИСЛИТЕЛЬНЫХ ЭКСПЕРИМЕНТОВ

Вычислительные эксперименты были проведены на трех модельных полуметриках:

1. «Линия»: исходные объекты лежат на одной прямой, $\forall (k, l) \in E_N$ $r_{kl} = |k - l|$;

2. $(0, 1)$ -полуметрика: исходные объекты собраны в два кластера, $\forall(k, l) \in E_N$

$$r_{kl} = \begin{cases} 0, & k, l \in V_{N'} \\ 1, & k \in V_{N'}, l \notin V_{N'}, \text{ где } V_{N'} \subset V_N \end{cases}$$

3. «Решётка»: исходные объекты расположены в узлах квадратной решетки, $\forall(k, l) \in E_N$ $r_{kl} = \sqrt{(x_k - x_l)^2 + (y_k - y_l)^2}$, $v_k = (x_k, y_k)$, $v_l = (x_l, y_l)$, $v_k, v_l \in \mathbb{R}^2$.

При $r'_{i_0 j_0} > r_{i_0 j_0}$ алгоритм \mathcal{A}_1 дал лучшие результаты коррекции на всех модельных полуметриках, т. к. в результате его применения расстояния вида r_{kl} были минимально деформированы (в смысле функционала Q_a). Это объясняется тем, что в общем случае он не требует коррекции *всех* расстояний (в отличие от \mathcal{A}). На рис. 1 показана коррекция «линии». На первом этапе были использованы формулы, согласованные с Q_a . Данный пример хорошо иллюстрирует эффект применения алгоритма \mathcal{A}_1 : изменено лишь положение объектов i_0, j_0 относительно остальной массы объектов.

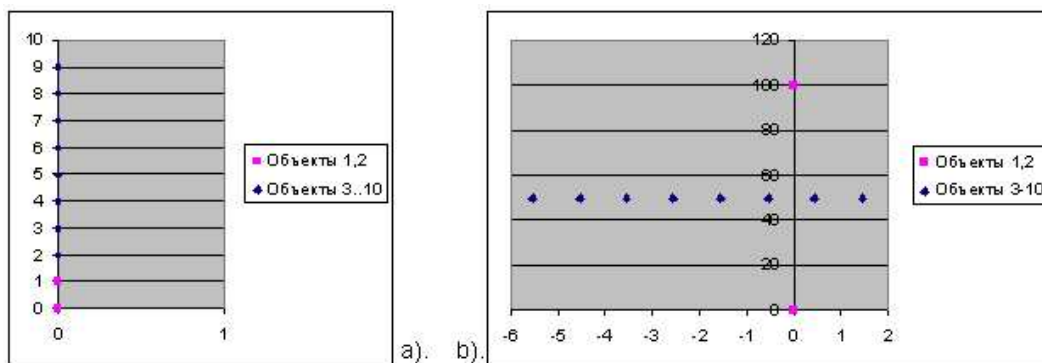


Рис. 1. Коррекция «линии» алгоритмом \mathcal{A}_1 ; $r_{1,2} = 1$, $r'_{1,2} = 100$.

На рис. 2.a – 2.d показаны результаты коррекции «решётки» алгоритмами \mathcal{A} и \mathcal{A}_1 . В обоих алгоритмах на первом этапе были использованы формулы, согласованные с Q_r . Использование на втором этапе алгоритма \mathcal{A} граничных значений параметра α может как существенно нарушить структуру полуметрики (при $\alpha = 0$, рис. 2.b), так и сохранить в целом взаимное расположение объектов (при $\alpha = 1$, рис. 2.c). Проведение оптимизации α позволяет приблизить величины \tilde{r}_{kl} к исходным r_{kl} , однако искажения расстояний все же происходят. В данном случае $\alpha_{opt} \approx 1$. Наилучший результат в смысле функционалов Q_a и Q_r дает применение алгоритма \mathcal{A}_1 (рис. 2.d).

Для иллюстрации случая $r'_{i_0 j_0} < r_{i_0 j_0}$ рассмотрим $(0, 1)$ -полуметрику. Пусть для пары объектов (i_0, j_0) $r_{i_0 j_0} = 1$, а $r'_{i_0 j_0} = 0$ и эксперт требует, чтобы объекты i_0, j_0 были выделены в отдельный кластер и при этом все остальные расстояния были минимально деформированы (в смысле Q_a).

На первом этапе коррекции были использованы формулы, согласованные с Q_a . В алгоритме \mathcal{A} значения $\alpha = 0$ и $\alpha = 1$ дают результат, неудовлетворяющий эксперта: в первом случае объекты обоих кластеров объединяются в один (рис. 3.a), а во втором –

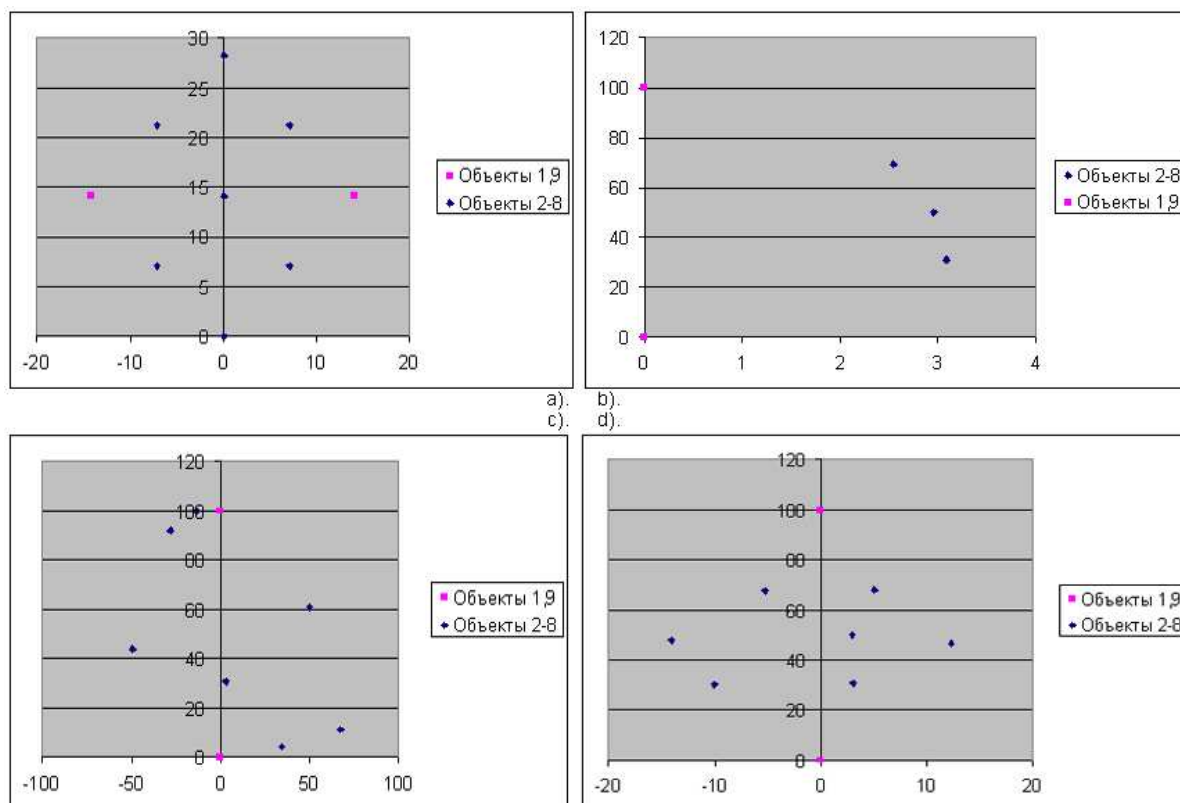


Рис. 2. Коррекция «решётки» алгоритмами \mathcal{A} и \mathcal{A}_1 ; $r_{1,9} \approx 28.28$, $r'_{1,9} = 100$.

каждый объект образует отдельный кластер (рис. 3.b). Поэтому оптимизация α не приносит ожидаемого результата (рис. 3.c). Алгоритм \mathcal{A}_1 для случая отрицательного локального возмущения, вообще говоря, не гарантирует построения полуметрики \tilde{R} . Однако в данном случае применение \mathcal{A}_1 дает полуметрику, причем она соответствует желаемому результату (рис. 3.d).

Оценим вычислительную сложность алгоритмов \mathcal{A} и \mathcal{A}_1 . На первом этапе корректируются $2(N - 2)$, а на втором этапе – $(N(N - 1)/2) - 2(N - 2) - 1$ расстояний. В алгоритмах \mathcal{A} и \mathcal{A}_1 для каждого расстояния \tilde{r}_{kl} вычисляются величины \tilde{r}_{kl}^{\min} , \tilde{r}_{kl}^{\max} , но если в \mathcal{A} $\alpha \in \{0, 1\}$, то вычисляется только одна из этих величин. Для оптимизации параметра α предлагается использовать метод «золотого сечения». Теоретические оценки сложности алгоритмов представлены в табл. 1.

На рис. 4 приведены эмпирические оценки времени работы алгоритмов \mathcal{A} и \mathcal{A}_1 при коррекции «линии» ($r_{1,2} = 1$, $r'_{1,2} = 100$). В алгоритме \mathcal{A} проводилась оптимизация параметра α методом «золотого сечения», число итераций равно 10. Этим объясняется бóльшая скорость роста времени работы с возрастанием N . Теоретические оценки показывают, что в условиях данного эксперимента алгоритм \mathcal{A}_1 должен

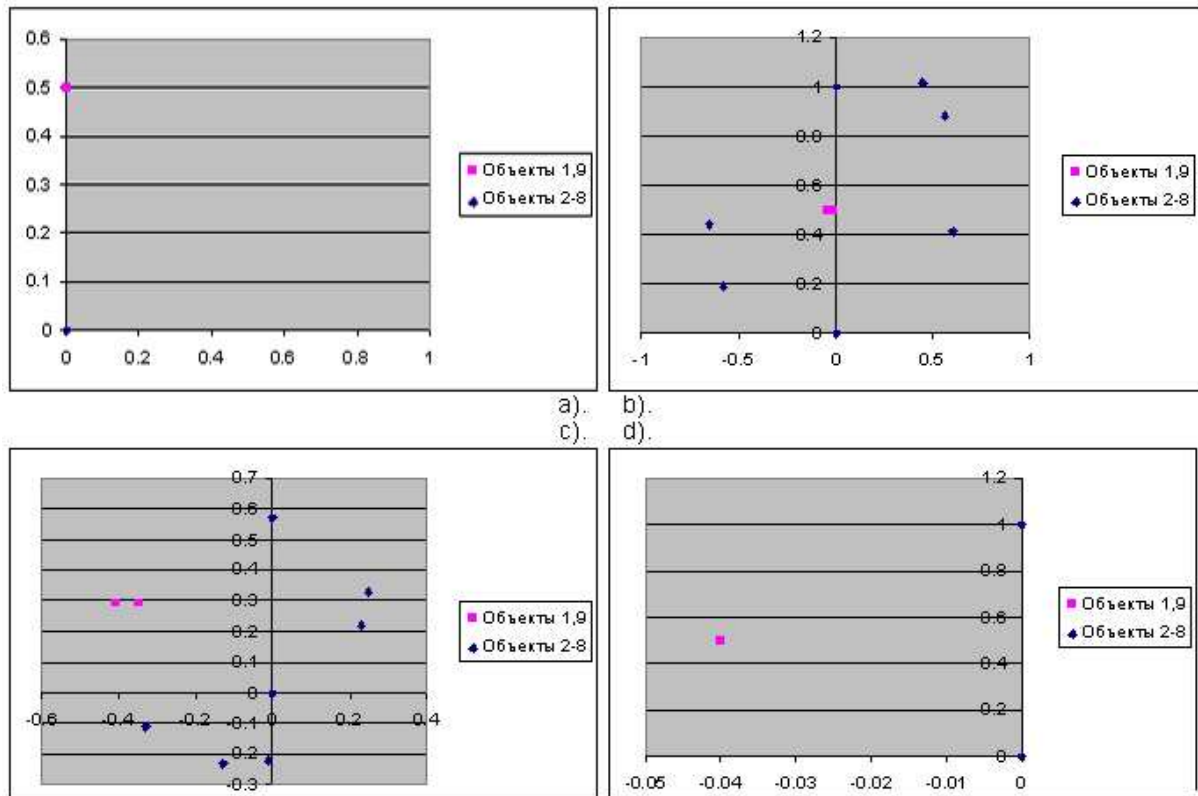


Рис. 3. Коррекция $(0, 1)$ -полуметрики алгоритмами \mathcal{A} и \mathcal{A}_1 ; $r_{1,9} = 1$, $r'_{1,9} = 0$.

работать в 12.17 раз быстрее, чем \mathcal{A} . По эмпирическим оценкам выигрыш по времени составляет 10.75 раз. Если в алгоритме \mathcal{A} не проводить оптимизацию, то \mathcal{A}_1 не дает существенного выигрыша по времени.

Эксперименты проводились на персональном компьютере (процессор AMD Athlon 64 processor 3000+ (2.0ГГц, 512КБ cache L2), оперативная память 512МБ).

Таблица 1. Теоретическая оценка сложности этапов алгоритмов \mathcal{A} и \mathcal{A}_1 .

Этап	Метод	Выч. сложность
1.	формулы, оптимизирующие Q_a	$\sim 12N$
	формулы, оптимизирующие Q_r	$\sim 13N$
	формулы, оптимизирующие Q_w	$\sim 19N$
2.	$\mathcal{A}, \alpha \in \{0, 1\}$	$\sim 1.5N^2$
	$\mathcal{A}, \alpha \in (0, 1)$	$\sim 5N^2$
	\mathcal{A} , оптимизация α (k итераций)	$\sim (3.5k + 1.5)N^2$
	\mathcal{A}_1	$\sim 3N^2$

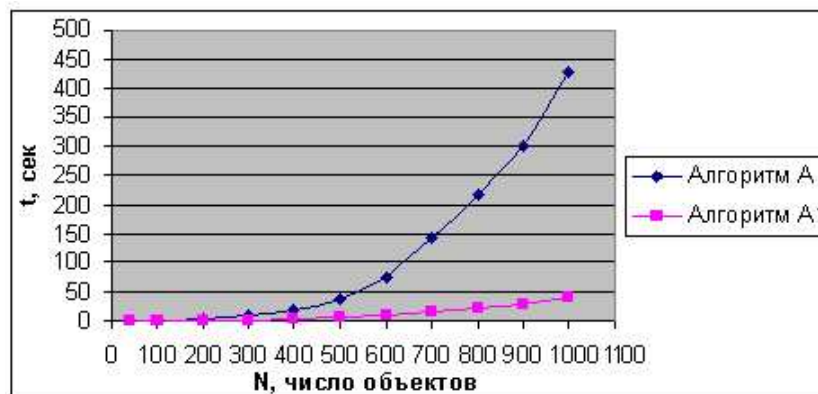


Рис. 4. Эмпирические оценки времени коррекции «линии».

ЗАКЛЮЧЕНИЕ

В работе предложена универсальная схема коррекции локальных возмущений в полуметрике, исследованы алгоритмы, построенные в ее рамках. Представлены результаты вычислительных экспериментов с использованием указанных алгоритмов. Основным направлением дальнейших исследований по данной тематике является распространение предложенного подхода на случай одновременной коррекции экспертом двух и более расстояний.

Работа выполнена при частичной финансовой поддержке РФФИ (проект № 08-07-00401), гранта Президента РФ МК-2252.2008.9.

СПИСОК ЛИТЕРАТУРЫ

1. Майсурадзе А. И. Об оптимальных разложениях конечных метрических конфигураций в задачах распознавания образов. // ЖВМ и МФ, Т. 44, №9, 2004 г., С. 1697–1707.
2. Майсурадзе А. И. Гомогенные и рагновые базисы в пространствах метрических конфигураций. // ЖВМ и МФ, Т. 46, №2, 2006 г., С. 344–361.
3. Юшманов С. В. Восстановление филогенетического древа по поддеревьям, порожденным четверками его висячих вершин. // Математическая кибернетика и ее приложения к биологии. – М.: Издательство Московского Университета. 1987 г., С. 141-147.
4. Громов И. А. Интерактивные методы коррекции полуметриков. // Сб. статей молодых ученых факультета ВМиК МГУ, Вып. 4, 2007 г., С. 27-40.

Статья поступила в редакцию 25.04.2008