



Д.А. РАЧКОВСКИЙ, И.С. МИСУНО, С.В. СЛИПЧЕНКО

УДК 004.22 + 004.93\*11 **РАНДОМИЗИРОВАННЫЕ ПРОЕКЦИОННЫЕ  
МЕТОДЫ ФОРМИРОВАНИЯ БИНАРНЫХ  
РАЗРЕЖЕННЫХ ВЕКТОРНЫХ ПРЕДСТАВЛЕНИЙ**

**Ключевые слова:** случайные проекции, векторные представления, разреженные бинарные представления, распределенные представления, эффективная оценка сходства.

#### **ВВЕДЕНИЕ. СЛУЧАЙНЫЕ ПРОЕКЦИИ ВЕКТОРНОЙ ИНФОРМАЦИИ**

Многие типы информации могут быть представлены в виде матриц или таблиц. Например, для поиска или классификации массивы документов часто рассматривают как матрицы документы-слова [1, 2] (строки — документы, столбцы — слова, элементы матрицы — значения некоторой функции частоты встречаемости слов в документах). Эту же информацию можно рассматривать как набор векторов или точек в многомерном пространстве. Размерность пространства — число координат (для данного примера число слов), а каждая точка соответствует объекту (документу). Координаты пространства — компоненты вектора или признаки.

Размерность пространства может составлять сотни тысяч и миллионы (например, число слов в словаре), а число точек — миллионы и миллиарды (число документов, например, веб-страниц Интернет). Это обуславливает целесообразность преобразования информационных массивов большой размерности таким образом, чтобы объекты сохранились, но повысилась эффективность их обработки и/или хранения. Преобразование может быть реализовано переходом в пространство с соответствующими свойствами — меньшей размерности, чем исходное, с вычислительно более эффективными операциями обработки, либо в пространство, для которого существуют эффективные методы обработки.

Многие методы и алгоритмы информационного поиска, классификации, кластеризации, аппроксимации, обучения и рассуждений на основе примеров, ассоциативной памяти и др. используют меры различия и сходства векторов, такие как расстояние, скалярное произведение, косинус угла, отношение расстояний и углов и др. Так, широко используемые линейные модели основаны на скалярном произведении, метод ближайшего соседа использует расстояния и т.д. Таким методам и алгоритмам для повышения эффективности полезно оперировать с преобразованными векторными представлениями и получать результаты, согласующиеся с результатами в исходном многомерном векторном пространстве.

Конструирование и реализация такого преобразования связаны с определенными трудностями. Например, распространенные методы преобразования на

© Д.А. Рачковский, И.С. Мисуну, С.В. Слипченко, 2012

основе анализа главных компонент PCA (Principal Component Analysis) или разложения по сингулярным значениям SVD (Singular Value Decomposition) не предназначены для оценивания мер сходства и различия исходных векторов. Вопросы сокращения размерности векторных представлений, продуцирующего векторы с нужными свойствами, исследуются в рамках нейросетевого подхода распределенных представлений [2–8], а также в области случайных проекций [9–12]. Здесь подход к преобразованию, который сохраняет или позволяет оценивать расстояния и углы, заключается в применении специально сконструированной проекционной матрицы  $\mathbf{R}(A \times N)$ . Исходный массив данных представлен в виде входной матрицы  $\mathbf{X}(L \times A)$ , где  $A$  — число атрибутов в векторе,  $L$  — число векторов. Результат преобразования — спроецированная матрица  $\mathbf{Y} = \mathbf{XR}$  — имеет размерность  $L \times N$ . Выходная матрица  $\mathbf{Z}$ , которая используется последующими методами обработки, получается преобразованием  $\mathbf{Y}$ :  $\mathbf{Z} = f(\mathbf{Y})$ , в частном случае  $\mathbf{Z} = \mathbf{Y}$ .

В работе [9] в качестве проекционной матрицы было предложено использовать случайную гауссову матрицу  $\mathbf{R}$  (элементы матрицы — реализации случайной гауссовой величины с нулевым средним и единичной дисперсией  $\mathcal{N}(0,1)$ ). По результатам проецирования входного массива с помощью такой матрицы можно оценить расстояния между исходными векторами: выходные векторы  $\mathbf{y} = \mathbf{R}^T \mathbf{x} / \sqrt{N}$  хорошо сохраняют попарные евклидовы расстояния между векторами  $\mathbf{x}$  во входном пространстве при  $N \ll A$ :  $E\{\|\mathbf{y}_1 - \mathbf{y}_2\|^2\} = d$ ,  $D\{\|\mathbf{y}_1 - \mathbf{y}_2\|^2\} = 2d^2 / N$ , где  $E\{\}$  — математическое ожидание,  $D\{\}$  — дисперсия,  $d = \|\mathbf{y}_1 - \mathbf{y}_2\|^2$ . Для скалярных произведений в [13] показано, что  $D\{(\mathbf{v}_1, \mathbf{v}_2)\} < 2m_1m_2 / N$ , а в [14] — что для скалярного произведения  $(\mathbf{x}_1, \mathbf{x}_2) = a$ :  $E\{(\mathbf{y}_1, \mathbf{y}_2)\} = a$ ,  $D\{(\mathbf{y}_1, \mathbf{y}_2)\} = (m_1m_2 + a^2) / N$ , где  $m_1, m_2$  — квадраты  $l_2$ -норм  $\mathbf{x}_1, \mathbf{x}_2$ .

Распределения  $y_{1,i} \sim \mathcal{N}(0,1)(m_1 / N)^{1/2}$ ,  $\|\mathbf{y}_1 - \mathbf{y}_2\|^2 \sim \chi_N^2 d / N$ , где  $\chi_N^2$  — случайная переменная с  $N$  степенями свободы и распределением хи-квадрат [10]. Анализ «хвостов» распределения хи-квадрат позволяет найти размерность  $N$ , для которой для любой пары векторов из  $L$  с заданной вероятностью и точностью  $(1 - \varepsilon)\|\mathbf{x}_1 - \mathbf{x}_2\|^2 \leq \|\mathbf{y}_1 - \mathbf{y}_2\|^2 \leq (1 + \varepsilon)\|\mathbf{x}_1 - \mathbf{x}_2\|^2$ .

К недостаткам преобразования гауссовой проекционной матрицы относятся:

- необходимость генерации большого числа ( $A \times N$ ) гауссовых случайных величин, которые представимы числами с плавающей запятой;
- необходимость умножения исходной матрицы на проекционную, что требует больших вычислительных затрат для матриц большого размера.

Конкретный вид (одинакового и независимого) распределения элементов проекционной матрицы  $\mathbf{R}$  в некоторой степени влияет на ошибки аппроксимации расстояний и гораздо сильнее — на вычислительную сложность проецирования. Вычислительная сложность проецирования зависит от вида элементов  $\mathbf{R}$  (для действительных требуется операция умножения с плавающей точкой, для дискретных  $\{-1, 0, +1\}$  — только сложение), а также от разреженности  $\mathbf{R}$  (для разреженной  $\mathbf{R}$  — с большой долей нулевых элементов — можно обрабатывать только ненулевые элементы, что значительно снижает вычислительную сложность проецирования). Для преодоления недостатков гауссовой проекционной матрицы в [11] предлагается использовать проекционные матрицы с элементами  $\{-1, 0, +1\}$ , которые генерируются с определенными и не близкими к нулю вероятностями. Такие случайные проекции с вероятностями  $q$  элементов  $r_{ij}$  матрицы

$\mathbf{R}$  ( $q_1$  для элементов 1,  $q_0$  для 0,  $q_{-1}$  для  $-1$ )

$$r_{ij} = \sqrt{s} \times \begin{cases} 1, & q_1 = 1/2s, \\ 0, & q_0 = 1-1/s, \\ -1, & q_{-1} = 1/2s \end{cases} \quad (1)$$

при  $1 \leq s \leq 3$  также хорошо сохраняют расстояния. Для  $s=3$  это позволяет в три раза повысить скорость вычислений.

В [2, 8] проводились исследования поведения контекстных векторов, полученных с помощью разреженной матрицы  $\mathbf{R}$  с очень малыми вероятностями элементов  $\{-1\}$  и  $\{+1\}$ , в задачах поиска и оценки семантической близости и получены результаты, сопоставимые по качеству с результатами без проецирования, но при существенно меньшем времени обработки. В [14] показано, что свойства проекций с  $s=1/q = A/\sqrt{A} = \sqrt{A}$  и даже с  $s = A/\log A$ , где  $q$  — доля ненулевых элементов  $\mathbf{R}$ , аналогичны свойствам проекций с  $s=1$  и  $s=3$  [11] и свойствам нормально распределенных случайных проекций [9]. Распределения и дисперсии спроецированных данных быстро асимптотически сходятся к варианту случайных проекций с гауссовой  $\mathbf{R}$ . Скорость сходимости в терминах с.к.о.:  $O(\sqrt{s/A})$ , что для  $s = \sqrt{A}$  дает  $O(1/\sqrt[4]{A})$  и для больших  $A$  приводит к малой потере точности по сравнению с проецированием гауссовой  $\mathbf{R}$ , а также дает возможность пользоваться результатами анализа гауссовой  $\mathbf{R}$  — леммой Johnson–Lindenstrauss [9] и т.п. Для такой скорости сходимости распределений требуется ограниченный третий момент компонентов входных векторов.

Для очень разреженной проекционной матрицы дополнительно снижаются вычислительные затраты на формирование и память для хранения за счет отсутствия необходимости генерации и хранения чисел с плавающей запятой. Это упрощает умножение исходной матрицы на проекционную. Однако после проецирования векторы имеют формат с плавающей точкой. Кроме того, в общем случае все значения компонента вектора не равны нулю, т.е. выходные векторы не являются разреженными. Это делает операции над такими векторами вычислительно сложными для последующих методов и алгоритмов. Поэтому перспективное направление дальнейшего повышения вычислительной эффективности — формирование бинарных выходных векторов.

В [12] предлагается трансформировать выходные векторы путем операции взятия знака  $\text{sign}()$  элементов трансформированной матрицы  $\mathbf{Y}$ . Это дает выходные бинарные векторы с компонентами  $\{0,1\}$ . Показано [15,12], что для таких векторов вероятность

$$p(\text{sign}(z_{1,i}) = \text{sign}(z_{2,i})) = 1 - \theta / \pi, \quad (2)$$

где  $z_{1,i}, z_{2,i}$  — компоненты выходных векторов  $\mathbf{z}_1$  и  $\mathbf{z}_2$ ,  $\theta$  — угол между  $\mathbf{x}_1$  и  $\mathbf{x}_2$ .

Полученные бинарные векторы, однако, являются неразреженными, плотными — вероятность единичного компонента  $p(z_i = 1) \sim 0.5$ . Хотя обработка таких бинарных векторов вычислительно более простая, чем векторов компонентов с плавающей запятой, они не применимы для ряда методов и алгоритмов, которые работают только с бинарными разреженными векторами, т.е. с векторами, в которых вероятность ненулевого (единичного) компонента  $p < 0.5$ . К такому классу относятся, например, методы обработки бинарных распределенных представлений в ассоциативно-проективных нейронных сетях [3, 16–19], распределенная ассоциативная память матричного типа [20].

В работе [2] элементы матрицы  $\mathbf{Y}$  (результата проецирования) подвергаются трансформации — пороговой операции с двумя порогами. В результате форми-

руется выходная тернарная матрица  $\mathbf{Z}$  с элементами  $\{-1, 0, +1\}$ . Вопросы количественного оценивания сходств и различий исходных векторов по таким результирующим векторам не рассматривались.

Таким образом, для повышения эффективности решения на основе векторно-матричных операций задач поиска, сравнения, классификации и другой обработки больших массивов информации, представленных в виде векторов большой размерности, актуальными являются разработка, исследование и анализ методов преобразования таких векторов случайным проецированием. В настоящей статье рассматриваются вопросы формирования бинарных векторов с регулируемой долей ненулевых компонентов, вычислительная эффективность их получения, а также исследуются их свойства, позволяющие оценить характеристики сходства–различия исходных векторов.

## 1. ФОРМИРОВАНИЕ БИНАРНЫХ ВЕКТОРОВ И ОЦЕНИВАНИЕ УГЛА МЕЖДУ ВХОДНЫМИ ВЕКТОРАМИ

Данные о нелокальном представлении информации в мозге привели к идее распределенного представления информации — форме векторного представления, где каждый объект есть совокупность компонентов вектора. В то же время для распределенных представлений семантика отдельных компонентов вектора не имеет значения. Важно, что сходным объектам соответствуют сходные распределенные представления. В качестве мер сходства используются такие меры сходства векторов, как скалярное произведение или косинус угла. Будем называть кодвекторами форму распределенного представления информации со свойствами бинарности (компоненты векторов  $\{0, 1\}$ ) и разреженности (доля ненулевых компонентов кодвектора мала).

Для формирования кодвекторов  $\mathbf{y}$  с компонентами  $\{0, 1\}$  будем использовать двухступенчатую процедуру, включающую этап изменения размерности исходных векторов  $\mathbf{x}$  путем их проецирования во вторичное пространство посредством умножения на матрицу  $\mathbf{R}$ :  $\mathbf{y} = \mathbf{x}^T \mathbf{R}$ , а также этап применения пороговых операций к результату умножения  $\mathbf{y}$ .

В качестве проекционной матрицы  $\mathbf{R}$  будем использовать  $\mathbf{R}(A \times N)$ , элементы которой — случайно расположенные элементы из множества  $\{-1, 0, +1\}$ . Для повышения вычислительной эффективности проецирования особенно интересен вариант разреженной (с малой долей ненулевых элементов) проекционной матрицы  $\mathbf{R}$ .

Каждому из  $A$  измерений входного пространства поставлен в соответствие  $N$ -мерный вектор (строка  $\mathbf{R}$ ). При размерности векторов-строк  $\mathbf{R}$   $N < A$  размерность  $\mathbf{y}$  меньше, чем  $\mathbf{x}$ . Вектор  $\mathbf{y}$  является вектором с компонентами — действительными числами, представимыми в формате с плавающей запятой.

На втором этапе процедуры применением пороговых операций к компонентам  $\mathbf{y}$  формируется бинарный кодвектор  $\mathbf{z}$  ( $\mathbf{T}$  — вектор порогов):

$$\mathbf{z} = \mathbf{T}(\mathbf{y}), \quad (3)$$

$$z_j = 1 \text{ при } y_j \geq 1 \geq T_j \geq 0, \text{ иначе } z_j = 0, j = 1, \dots, N. \quad (4)$$

Для  $T_j > 0$  получаем  $\mathbf{z}$  с вероятностью единичного компонента менее 0.5:  $p(z_j = 1) < 0.5$ . Величина  $T_j$  обеспечивает заданную вероятность единичного значения в компоненте  $z_j$ . Исследуем случай, когда в качестве  $T_j$  используется единый для всех  $j$  порог  $T$ , который обеспечивает заданную вероятность наличия единичного компонента в векторе  $\mathbf{z}$ . Для вычислительно эффективной обработки, а также для работы ряда методов и алгоритмов требуются разреженные бинарные кодвекторы  $\mathbf{z}$ .

Рассмотрим способ оценки угла между входными векторами  $\mathbf{x}$  по кодвекторам  $\mathbf{z}$ . Для  $\mathbf{R}$ , где вероятность 1 и  $-1$  равна  $q/2$ , математическое ожидание  $E\{r_{kj}x_k\} = 0$  и дисперсия

$$D\{r_{kj}x_k\} = E\{(r_{kj}x_k)^2\} = px_k^2. \quad (5)$$

Обозначим  $y = \sum_k r_{kj}x_k$ ,  $\|\mathbf{x}\|$  —  $l_2$ -норма  $\mathbf{x}$ . Тогда

$$D\{y\} = \sum_{k=1, \dots, A} D\{r_{kj}x_k\} = p\|\mathbf{x}\|^2 = \|\mathbf{x}\|^2/s. \quad (6)$$

В качестве случайной переменной рассмотрим  $y_j$ , полученную из исходной нормированием

$$y_j / (\|\mathbf{x}\|/\sqrt{s}). \quad (7)$$

Распределение  $y$  быстро сходится к гауссовому с нулевым средним и единичной дисперсией, поэтому значение порога  $t$  определяет вероятность единичного компонента  $p(\mathbf{z})$  выходного кодвектора  $\mathbf{z}$  как

$$p(\mathbf{z}) = p(y > t_p) = \frac{1}{\sqrt{2\pi}} \int_{t_p}^{\infty} e^{-y^2/2} dy = 1 - \Phi(t_p), \quad (8)$$

где  $\Phi$  — функция нормального распределения.

В [14] показано, что случайные переменные  $(y_{1,j}, y_{2,j})$  имеют совместное нормальное распределение с нулевыми средними. Для нашего случая нормированных  $y$  корреляция  $y_{1,j}$  и  $y_{2,j}$  равна  $\cos(\theta)$ , где  $\theta$  — угол между входными векторами  $\mathbf{x}_1$  и  $\mathbf{x}_2$ . Поэтому вероятность совпадения единичных компонентов  $z_{1,j} = 1$  и  $z_{2,j} = 1$  в кодвекторах  $\mathbf{z}_1$  и  $\mathbf{z}_2$  после бинаризации  $y_1$  с порогом  $t_1$  и  $y_2$  с порогом  $t_2$  можно определить как результат вычисления интеграла от двумерного гауссова распределения:

$$\begin{aligned} p(z_{1,j} = 1, z_{2,j} = 1 | \theta, t_1, t_2) &= p(y_{1,j} > t_1, y_{2,j} > t_2 | \theta) = \\ &= \frac{1}{2\pi(1 - \cos^2\theta)} \int_{t_1}^{\infty} \int_{t_2}^{\infty} e^{-\frac{y_1^2 - 2y_1y_2\cos\theta + y_2^2}{2(1 - \cos^2\theta)}} dy_1 dy_2. \end{aligned} \quad (9)$$

Для упрощения записи выражений будем использовать обозначения  $z_{1,j} \equiv a$ ,  $z_{2,j} \equiv b$ , а также опускать  $\theta$  и отделять  $t$  запятой, так что  $p(z_{1,j} = 1, z_{2,j} = 1 | \theta, t_1, t_2) \equiv p(a = 1, b = 1, t_1, t_2)$ .

Значения (9), полученные численным интегрированием с использованием MatLab для значений  $t_p \equiv \{0, 1.282, 2.054\}$ , приведены на рис. 1 (линии 1), где наблюдается «завал» характеристики из-за наличия в (9) деления на малые значения  $1 - \cos^2\theta$  (при малых  $\theta$ ) и особенностей реализации интегрирования.

Для устранения этого недостатка выражения для вычисления модифицированы следующим

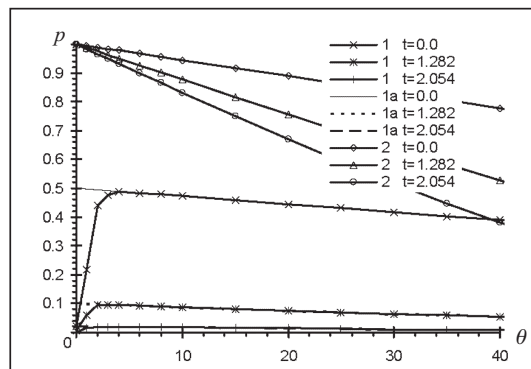


Рис. 1. Зависимость вероятности  $p$  совпадения единичных компонентов кодвекторов от входного угла  $\theta$

образом. Сделаем замену переменных:

$$w_1 = \frac{1}{2} \frac{y_1 + y_2}{\cos(\theta/2)}, w_2 = \frac{1}{2} \frac{y_1 - y_2}{\sin(\theta/2)}, w^* = \frac{w_1 \cos(\theta/2) - t}{\sin(\theta/2)}.$$

Тогда для  $t_1 = t_2 = t$

$$\begin{aligned} p(a=1, b=1|t, \theta) &= \frac{1}{2\pi} \int_{\frac{t}{\cos(\theta/2)}}^{\infty} \int_{-w^*}^{+w^*} e^{-\frac{w_1^2 + w_2^2}{2}} dw_1 dw_2 = \\ &= \frac{2}{\sqrt{2\pi}} \int_{\frac{t}{\cos(\theta/2)}}^{\infty} e^{-\frac{w_1^2}{2}} F(w^*) dw_1 dw_2 = f(\theta), \end{aligned} \quad (10)$$

где  $F(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz = \Phi(x) - \frac{1}{2}$ .

Линии 1а на рис. 1 показывают отсутствие «завала» графика  $p(a=1, b=1, t)$  при малых  $\theta$ . Это хорошо видно по линиям 2, которые соответствуют  $p(a=1|b=1, t) = p(a=1, b=1, t) / p(b=1, t)$  и поэтому исходят из точки (0,1). Рассмотрим (10) при  $t=0$ :

$$p(a=1, b=1|t=0, \theta) = \frac{2}{2\pi} \int_0^{+\infty} \int_0^{w^*} e^{-\frac{w_1^2 + w_2^2}{2}} dw_1 dw_2. \quad (11)$$

Для перехода к полярным координатам сделаем замену переменных  $w_1 = r \cos \varphi, w_2 = r \sin \varphi$ . Тогда пределы интегрирования по  $w_2$ :  $\int_0^{+w^*}$  изменятся следующим образом:  $w_2 \rightarrow \varphi = 0; w_2 = w^*(t=0) \rightarrow \varphi = \pi/2 - \theta/2$ , так как на линии

$$w_2 = w^*(t=0) = w_1 \cos(\theta/2) / \sin(\theta/2), \quad \text{tg } \varphi = \text{ctg}(\theta/2) = \text{tg}(\pi/2 - \theta/2).$$

Поэтому

$$p(a=1, b=1|t=0, \theta) = \frac{2}{2\pi} \int_0^{\pi/2 - \theta/2} \int_0^{r^2} e^{-\frac{r^2}{2}} r dr d\varphi = \frac{\pi - \theta}{2\pi} = \frac{1}{2} \left(1 - \frac{\theta}{\pi}\right). \quad (12)$$

Поскольку

$$p(a=1, b=1) = \frac{1}{2} (p(a=1) + p(b=1) - p(a \neq b)),$$

а  $p(a=1, t=0) = p(b=1, t=0) = 0.5$ , то  $2p(a=1, b=1, t=0) = p(a=b, t=0)$ , поэтому  $p(a=b, t=0) = 1 - \frac{\theta}{\pi}$ . Это выражение совпадает с результатом [16, 12], где рассматривался только случай  $t=0$ .

Для  $t=0$   $p(a=1, t=0) = p(b=1, t=0) = 0.5$ , т.е.  $t=0$  дает неразрезанные код-векторы. Выбор порога  $t > 0$  позволяет управлять степенью разреженности выходного кодвектора, см. (8). Это также видно из того, что для  $t > 0$  и  $\theta \rightarrow 0, w^* \rightarrow \infty$ , поэтому

$$F(w^*) \rightarrow \frac{1}{2} \text{ и } p(a=1, b=1|t, \theta=0) = \frac{1}{\sqrt{2\pi}} \int_t^{\infty} e^{-\frac{w_1^2}{2}} dw_1. \quad (13)$$

Для  $y$  без нормирования (7), полученного непосредственно в результате умножения входного вектора на проекционную матрицу с  $\{-1, 0, +1\}$ , ту же разреженность (долю единичных компонентов) выходного кодвектора дает порог  $T = t|\mathbf{x}|/\sqrt{s}$ .

Для того чтобы показать, что вероятность совпадения единичных компонентов монотонно уменьшается с ростом угла  $\theta$ , вычислим производную

$$\frac{dp}{d\theta} = \frac{2}{\sqrt{2\pi}} \left[ -e^{-\frac{1}{2}\left(\frac{t}{\cos\theta}\right)^2} F(0) \frac{d\left(\frac{t}{\cos\theta}\right)}{dt} + \frac{\int_0^\infty e^{-\frac{w_1^2}{2}} dw_1 \frac{dF(w^*)}{d\theta} dw_1}{\frac{t}{\cos(\theta/2)}} \right],$$

$F(0) = 0$ , поэтому первый член равен нулю. Для второго члена

$$\frac{dF(w^*)}{d\theta} = \frac{1}{\sqrt{2\pi}} e^{-\frac{w^{*2}}{2}} \frac{dw^*}{dt},$$

$$\frac{dw^*}{dt} = -\frac{w_1 \sin(\theta/2)}{2 \sin(\theta/2)} - \frac{(w_1 \cos(\theta/2) - t) \cos(\theta/2)}{2 \sin^2(\theta/2)}.$$

Последнее выражение меньше нуля, так как  $\theta \in (0, \pi)$  и  $w_1 \cos(\theta/2) - t > 0$ , поэтому подынтегральное выражение и значение интеграла, равное искомой производной  $\frac{dp}{d\theta}$ , меньше нуля.

Таким образом,  $\theta$  можно оценивать как  $\theta = f^{-1}(p(a=1, b=1, t)) = g(p(a=1, b=1, t))$ , где  $f^{-1}$  — функция, обратная к  $f$  (10),  $p$  оценивается по кодвекторам. В следующем разделе показано, что меньшую дисперсию имеет оценка угла по условной вероятности  $p(a=1|b=1) = p(a=1, b=1) / p(b=1)$  (см. также рис. 1), особенно при оценке  $p(a=1|b=1)$  как  $\frac{|z_1 \wedge z_2| p_t(z_1) N}{|z_2| |z_1|}$ , где  $\wedge$  — покомпонентная операция конъюнкции,  $|z|$  — число единичных компонентов в  $z$ ,  $p_t(z_1)$  — теоретическая оценка разреженности кодвектора, определяемая по  $1 - \Phi$  (8).

Экспериментально исследовалась зависимость оценки угла между входными векторами  $x_1, x_2$  по экспериментально полученным значениям совпадения компонентов их результирующих кодвекторов  $z_1, z_2$  для разных  $A, N, s$  и  $p(z)$ , которая определяется порогом  $t$ . На рис. 2 представлены результаты одного из экспериментов для входных векторов  $x_1, x_2$  размерностью  $A = 1000$  (компоненты вектора — положительные случайные числа в диапазоне  $[0, A]$ ),  $s = 10$ ,  $N = 1000$ ,  $p(z) = \{0.5, 0.1\}$  (т.е.  $t = \{0.0, 1.282\}$ ), усреднение результатов производилось по 1000 реализациям проекционной матрицы.

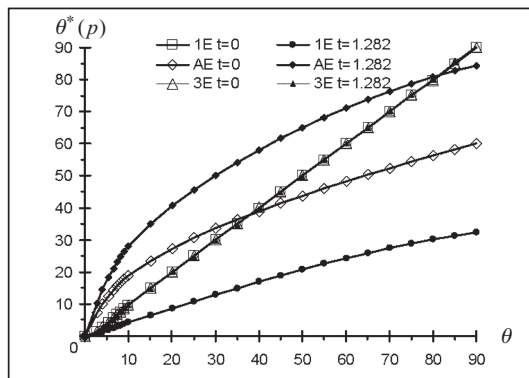


Рис. 2. Оценки угла между исходными векторами по функциям совпадения компонентов их кодвекторов

Приведены зависимости оценки  $\theta^*$  входного угла  $\theta(x_1, x_2)$  от ряда функций их кодвекторов:  $\pi(1 - p^*(z_{1,j} = z_{2,j}))$  (линии 1E);  $\arccos(z_1, z_2)$  (линии AE) и с использованием полученной выше функции  $g(p^*(a=1, b=1, t))$  (линии 3E). Видно, что все оценки имеют монотонную зависимость от входного угла. Одна-

ко оценка  $\theta^* = \pi(1 - p(a = b))$  близка к  $\theta$  только для  $t = 0$ ;  $\arccos(\mathbf{z}_1, \mathbf{z}_2)$  нелинейно (и по-разному для разных  $t$ ) зависит от  $\theta$ , а предлагаемая нами оценка  $\theta^*$  близка к  $\theta$  для разных  $t$ .

## 2. МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ И ДИСПЕРСИЯ ОЦЕНКИ ВЕЛИЧИНЫ УГЛА

Определим математическое ожидание и дисперсию оценки величины угла между исходными векторами по оценкам вероятностей совпадения компонентов выходных бинарных кодвекторов. Распределение числа  $k$  совпадающих единичных компонентов («единиц») кодвекторов для любого порога  $t$  биномиальное:  $P(k) = \beta(k, N, p(a = 1, b = 1))$  с  $E\{k\} = Np(a = 1, b = 1)$  и  $D\{k\} = Np(a = 1, b = 1)(1 - p(a = 1, b = 1))$ . Поэтому

$$E\{p^*(a = 1, b = 1)\} = E\{k / N\} = p(a = 1, b = 1);$$

$$D\{p^*(a = 1, b = 1)\} = D\{k / N\} = D\{k\} / N^2 = p(a = 1, b = 1)(1 - p(a = 1, b = 1)) / N.$$

Распределение числа  $n$  совпадающих компонентов кодвекторов (как единичных, так и нулевых) для любого порога  $t$  биномиальное:  $P(n) = \beta(n, N, p(a = b))$  с  $E\{n\} = Np(a = b)$  и  $D\{n\} = Np(a = b)(1 - p(a = b))$ . Тогда

$$E\{p^*(a = b)\} = E\{n / N\} = E\{n\} / N = p(a = b);$$

$$D\{p^*(a = b)\} = D\{n / N\} = D\{n\} / N^2 = p(a = b)(1 - p(a = b)) / N.$$

Для определения математического ожидания и дисперсии угла по зависимости  $\theta = g(p)$ , используя дельта-метод (линеаризацию функций случайного аргумента [21] при малых отклонениях  $p$  от среднего с применением одношагового разложения Тейлора), получаем

$$E\{\theta^*\} = g(E\{p^*\}), \quad (14)$$

$$D\{\theta^*\} = D\{p^*\} [g'(E\{p^*\})]^2, \quad g' = \frac{dg}{dp}. \quad (15)$$

Рассмотрим случай  $t = 0$  и  $p(a = b, t = 0)$ . Для него (см. (12))  $\theta = \pi(1 - p(a = b, t = 0)) \equiv g(p)$ ,  $\frac{d\theta}{dp} = -\pi$ , поэтому

$$\begin{aligned} D\{\theta^* | t = 0, p^*(a = b)\} &= D\{p^*(a = b)\} \pi^2 = \\ &= \pi^2 p(a = b, t = 0)(1 - p(a = b, t = 0)) / N = \\ &= \frac{\pi^2}{N} \left(1 - \frac{\theta}{\pi}\right) \frac{\theta}{\pi} = \frac{\pi}{N} \theta \left(1 - \frac{\theta}{\pi}\right) = \theta(\pi - \theta) / N. \end{aligned} \quad (16)$$

Здесь  $D\{\theta^* = 0\} = 0$ . Теперь рассмотрим оценивание величины угла на основе оценок  $p(a = 1, b = 1)$ . При  $t = 0$ , так как  $p(a = b, t = 0) = 2p(a = 1, b = 1, t = 0)$  (см. (12)), аналогично предыдущему случаю получаем

$$D\{\theta^* | t = 0, p^*(a = 1, b = 1)\} = \frac{\pi^2 - \theta^2}{N}. \quad (17)$$

При  $t > 0$  для численного определения  $[g'(E\{p\})]^2$  можно воспользоваться таблицей значений  $\theta = g(p)$ , полученной по (10), и вычислить  $g' = \frac{2\Delta\theta}{|p_{\theta+1} - p_{\theta-1}|}$ , где  $\Delta\theta$  — шаг углов в таблице, построенной по (10),  $p_{\theta+1}, p_{\theta-1}$  — значения  $p$  из следующей и предыдущей относительно  $\theta$  ячеек таблицы.



При вычислении величины угла на основе оценок  $p(a=1, b=1)$  получаем примерно одинаковые и гораздо большие, чем для оценок  $p(a=b)$ , значения дисперсии  $D\{\theta^*\}$ .

Так, при  $t=0$  значение  $ND\{\theta^*\}$  убывает от  $\pi^2$  до  $0.75\pi^2$  при изменении  $\theta$  от 0 до 90, а при  $t>0$  значения еще больше. Для уменьшения дисперсии воспользуемся условными вероятностями  $p(a=1|b=1) = p(a=1, b=1) / p(b=1)$  (см. рис. 1). Тогда

$$D\{p^*(a=1|b=1)\} = \frac{p(a=1, b=1)(1 - p(a=1, b=1))}{Np(b=1)}. \quad (18)$$

При  $\theta=0$   $p(a=1|b=1)=1$  и  $D\{p(a=1|b=1)\}=0$ , т.е.  $D\{\theta^*\}=0$ . Таким образом, в этом случае дисперсия при  $\theta=0$  равна нулю и растет с ростом угла. Заметим, однако, что вблизи  $\theta=0$  и  $\theta=\pi$  из-за высокой нелинейности линеаризация (15) дает ненадежные результаты.

На рис. 3 представлены теоретические зависимости дисперсии оценки угла (в град<sup>2</sup>) от его значения при оценке по вероятности совпадения компонентов  $p(a=b, t=0)$  (линии 1Т), совместной вероятности  $p(a=1, b=1)$  (линия 2Т), условной вероятности  $p(a=1|b=1)$  (линии 3Т). Рисунок построен для  $N=1000$ ,  $s=10$ ,  $t=\{0, 1.282\}$ , Е — эксперимент, Т — теория. Видно, что при малых углах дисперсия оценки угла по  $p(a=1, b=1)$  максимальна и незначительно уменьшается при увеличении угла до  $\pi/2$ . При оценке угла по  $p(a=1, b=1)$  дисперсия растет от нуля до максимальных значений при  $\pi/2$ , которые сравнимы с дисперсиями для  $p(a=1, b=1)$ .

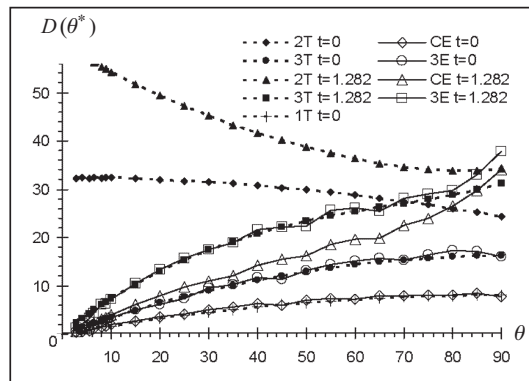


Рис. 3. Зависимость дисперсии оценки угла  $\theta$  от его величины,  $N = 1000$

Экспериментальные результаты оценок  $p(a=1|b=1)$  по кодвекторам (линия 3Е) хорошо совпадают с теоретическими. Превышение экспериментальными значениями теоретических при  $\theta \sim \pi/2$ ,  $t=1.282$  обусловлено появлением пар кодвекторов  $\mathbf{z}_1, \mathbf{z}_2$  с нулевым перекрытием, для которых оценка  $\theta$  равна  $\pi$ .

Для дальнейшего уменьшения дисперсии использовалась скорректированная оценка  $p(a=1|b=1)$ , вычисляемая как  $\frac{|\mathbf{z}_1 \wedge \mathbf{z}_2| p_t(\mathbf{z}_1) N}{|\mathbf{z}_2| |\mathbf{z}_1|}$ , где  $p_t(\mathbf{z}_1)$  — теоретическая оценка плотности кодвектора, определяемая по 1-Ф. При такой скорректированной оценке дисперсия (линия 3Е) уменьшается при  $t=0$  до значений дисперсии оценки  $\theta$  по  $p(a=b|t=0)$ .

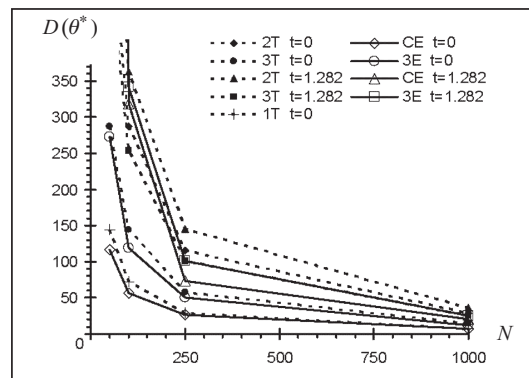


Рис. 4. Зависимость дисперсии оценки угла  $\theta$  от размерности кодвектора  $N$

угла  $\theta = 60^\circ$  при  $s=10, t = \{0,0,1,282\}, A=1000$  (обозначения такие же, как на рис. 3). Дисперсия быстро уменьшается с ростом  $N$ . При  $N > 200$  теоретические зависимости хорошо совпадают с экспериментальными. При малых  $N < 100$  экспериментальные значения дисперсии превышают теоретические как из-за нарушения гауссовости распределений, так и из-за появления пар векторов с нулевым перекрытием и нулевым числом единичных компонентов. Дисперсия оценки угла по  $p(a=1|b=1)$  2Т наибольшая.

Зная евклидовы длины исходных векторов  $\|\mathbf{x}_1\|, \|\mathbf{x}_2\|$ , можно оценить их скалярное произведение  $(\mathbf{x}_1, \mathbf{x}_2)^*$  по величине угла  $\theta^*$ , оцененного по кодвекторам:

$$(\mathbf{x}_1, \mathbf{x}_2)^* = \|\mathbf{x}_1\| \|\mathbf{x}_2\| \cos \theta^*. \quad (19)$$

Согласно дельта-методу

$$\begin{aligned} D\{(\mathbf{x}_1, \mathbf{x}_2)^*\} &= \|\mathbf{x}_1\|^2 \|\mathbf{x}_2\|^2 D\{\cos \theta^*\} = \\ &= \|\mathbf{x}_1\|^2 \|\mathbf{x}_2\|^2 D\{\theta^*\} \left( \frac{d \cos \theta}{d \theta} \right)^2 = \|\mathbf{x}_1\|^2 \|\mathbf{x}_2\|^2 D\{\theta^*\} \sin^2 \theta. \end{aligned} \quad (20)$$

Для  $t=0$   $D\{(\mathbf{x}_1, \mathbf{x}_2)^*\} = \|\mathbf{x}_1\|^2 \|\mathbf{x}_2\|^2 D\{\theta^*\} (\sin \theta)^2 (\pi - \theta) \theta / N$ .

Отметим, что это приближение справедливо при ненулевом  $\sin \theta$ , что нарушается при  $\theta = \{0, \pi\}$ . Поэтому вблизи 0 и  $\pi$  из-за высокой нелинейности оценки значений дисперсии, полученные с помощью такого приближения, ненадежны.

### 3. ГЕОМЕТРИЧЕСКАЯ ИНТЕРПРЕТАЦИЯ

Согласно (4)  $b=1$ , если скалярное произведение  $(\mathbf{x}, \mathbf{r}) \geq T = t \|\mathbf{x}\|/\sqrt{s}$ . Для больших  $A$  имеем  $E\{\|\mathbf{r}\|\} \approx \sqrt{A/s}$  и  $b=1$  при  $\cos(\mathbf{x}, \mathbf{r}) \geq t/\sqrt{A}$ . Таким образом, вероятность совпадения единичных компонентов кодвекторов  $p(a=1, b=1)$  есть отношение перекрытия площадей, отсекаемых на сфере конусами с углами раствора  $\alpha$ , где  $\alpha/2 = \arccos(t/\sqrt{A})$ , к площади сферы. Осями конусов являются векторы  $\mathbf{x}_1, \mathbf{x}_2$  с углом  $\theta = \{0, \pi\}$  между собой, а внутренняя часть конусов соответствует случайным векторам  $\mathbf{r}$  проекционной матрицы с положительными проекциями на  $\mathbf{x}_1, \mathbf{x}_2$  (рис. 5).

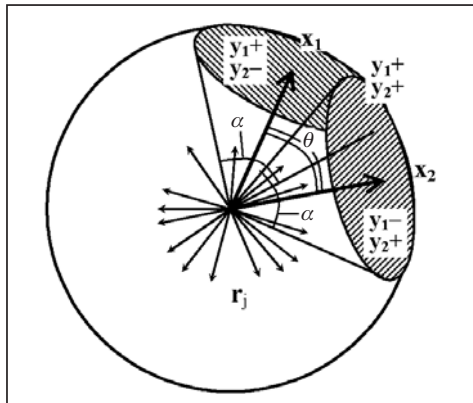


Рис. 5. Геометрическая интерпретация  $p(z_1, z_2)$  для  $t > 0$

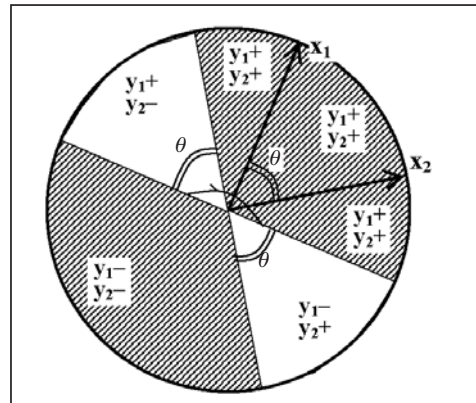


Рис. 6. Геометрическая интерпретация  $p(z_1, z_2)$  для  $t = 0$

Для порога  $t=0$  конусы имеют телесный угол  $\alpha = \pi$ , т.е. «раскрываются» в плоскости. Тогда угол между векторами есть угол между плоскостями (рис. 6).

Здесь длина векторов  $\mathbf{x}$ ,  $\mathbf{r}$  уже не имеет значения, так как для определения знака компонентов кодвекторов важен только знак  $\cos(\mathbf{x}, \mathbf{r})$ , а не его величина.

Область, где проекции  $\mathbf{x}_1, \mathbf{x}_2$  на векторы проекционной матрицы  $\mathbf{r}_j$  положительные, т.е. где выполняется условие  $(a = 1, b = 1, t = 0)$ , помечена на рис. 6 как  $y_+, y_+$ . Эта область занимает двугранный угол  $\pi - \theta$ . Поэтому вероятность совпадения единичных компонентов кодвекторов  $p(a = 1, b = 1, t = 0) = (\pi - \theta) / 2\pi$ , где  $2\pi$  — двугранный угол всего пространства (гиперсферы). Область совпадения нулевых компонентов кодвекторов такая же, поэтому  $p(a = 0, b = 0, t = 0) = (\pi - \theta) / 2\pi$ . Тогда  $p(a = b, t = 0) = (\pi - \theta) / \pi$ ,  $p(a = 1, b = 1, t = 0) = p(a = b, t = 0) / 2$ , см. (12).

Оценка  $p(a \neq b, t = 0) = \theta / \pi$  по кодвекторам есть  $H / N = |\mathbf{z}_1 \oplus \mathbf{z}_2| / N$ , где  $H = M_1 + M_2 - 2|\mathbf{z}_1 \wedge \mathbf{z}_2|$  — расстояние Хемминга,  $\oplus$  — покомпонентная операция «исключающее или»,  $|\dots|$  — число единиц в кодвекторе,  $M_1 = |\mathbf{z}_1|$ ,  $M_2 = |\mathbf{z}_2|$ . Соответственно оценка  $p(a = b, t = 0) = 1 - \theta / \pi = 1 - H / N$ .

Проецирование в сочетании с бинаризацией можно рассматривать как извлечение  $N$  признаков из входного вектора (индикаторов его положения относительно случайных гиперплоскостей), аналогичное производимому первым фиксированным слоем перцептрона Розенблатта. Заметим, что если порог равен числу единиц в  $\mathbf{r}_j$ , осуществляется извлечение признаков-предикатов (см. в [22–24] пример такого выделения признаков LIRA на изображении).

## ЗАКЛЮЧЕНИЕ

В настоящей статье исследуются свойства рандомизированных бинарных векторных представлений с регулируемой степенью разреженности (долей ненулевых компонентов), которые формируются из входных массивов векторов, имеющих формат с плавающей запятой, путем проецирования случайной матрицей с тернарными компонентами  $\{-1, 0, +1\}$ . Проанализирована точность оценивания мер сходства-различия исходных векторов, вычисляемых на основе угла, по эмпирическим вероятностям совпадения компонентов выходных векторов. Выходные бинарные векторы могут использоваться вместо входных (или в сочетании с ними) для повышения эффективности поиска и классификации [2, 5], формирования представлений сложно структурированной информации [3, 16], в нейросетевой распределенной ассоциативной памяти [20] и других приложениях интеллектуальной обработки информации.

Применение случайных разреженных проекционных матриц с компонентами  $\{-1, 0, +1\}$  и долей ненулевых компонентов в проекционной матрице  $p_1 \ll 1$  позволяет снизить вычислительную сложность проецирования по сравнению со случайными матрицами с компонентами, распределенными по нормальному закону, за счет обработки только ненулевых компонентов матрицы. Вычислительная сложность проецирования (умножения входного вектора на проекционную матрицу) составляет  $O(ANp_1p_2)$  ( $p_2$  — доля ненулевых компонентов во входном векторе,  $A$  — размерность входного вектора,  $N$  — размерность выходного вектора), что дает значительный выигрыш по сравнению с неразреженными проекционными матрицами с  $p_1 \sim 1$ . Кроме того, вычислительные затраты могут быть уменьшены за счет использования операций сложения вместо умножения.

Бинаризация с порогом обеспечивает получение выходных бинарных векторов с нужной долей ненулевых компонентов. Формирование выходного вектора выбором  $M$  максимальных компонентов вектора, полученного в результате умножения входного вектора на проекционную матрицу, требует  $O(N \log M)$  операций по сравнению с  $N$  операциями при аналитическом вычислении порога, что делает второй вариант более предпочтительным с точки зрения вычислительных затрат.

Вероятность совпадения компонентов выходных кодвекторов является монотонной функцией угла между входными векторами и позволяет оценивать угол между входными векторами. По углу также можно оценить расстояние  $d$  и ска-

лярное произведение, если известны  $l_2$ -нормы входных векторов  $\|\mathbf{x}_1\|^2, \|\mathbf{x}_2\|^2$ :  $d^2 \equiv \|\mathbf{x}_1 - \mathbf{x}_2\|^2 = \|\mathbf{x}_1\|^2 + \|\mathbf{x}_2\|^2 - 2\cos\theta\|\mathbf{x}_1\|\|\mathbf{x}_2\|$  (нормы входных векторов вычисляются один раз, а хранение одной нормы требует только одного числа).

Использование бинарных векторов позволяет сократить затраты памяти и вычислительных ресурсов на их хранение и обработку на обычных компьютерах. Это достигается использованием одного бита для хранения компонента бинарного вектора вместо 32–64 бит для формата с плавающей точкой, применением логических операций и суммирования вместо операций с плавающей запятой, а также возможностью эффективной аппаратной поддержки в виде специализированных вычислителей и нейрокомпьютеров [25]. Разреженные бинарные векторы (т.е. векторы с малым числом ненулевых компонент) дополнительно повышают эффективность обработки (например, путем использования процедур, построенных на обратном индексе, и др.).

При числе бит для представления выходных векторов, меньшем, чем для входных, получаем выигрыш в памяти для хранения и в скорости обработки. Размерность выходных векторов можно выбирать в зависимости от требуемой точности оценки сходства входных векторов. Кроме того, выходные векторы малой размерности можно использовать для быстрой огрубленной оценки сходства и выбора кандидатов на последующее более точное сравнение. Возможные области применения бинарных представлений входных векторов — повышение эффективности информационного поиска (поиск в Интернет, поиск близких к запросу и друг другу документов в хранилищах, поиск другой информации в базах данных и знаний, поиск базисных функций в больших словарях для методов машинного обучения и др.); повышение эффективности классификации и кластеризации разного рода информации в векторном формате (текстов, изображений, звуков, результатов измерений и др.); повышение эффективности других методов обработки информации и информационных технологий, требующих оперирования большими массивами векторов.

Выходные разреженные бинарные рандомизированные векторы, в которых не важна семантика отдельных компонентов, но важно отражение ими сходства входных векторов, необходимы также для ряда методов и парадигм интеллектуальной обработки информации, которые требуют именно такого формата. Примером является концепция ассоциативно-проективных нейронных сетей [3], где представления сложных структурированных объектов самой разной природы строятся из бинарных разреженных векторных рандомизированных представлений (известных как «распределенные представления информации»), и методы хранения, обобщения, и восстановления информации в распределенной ассоциативной памяти [20], а также подходы к обработке структур [16–19] и учета семантики [2], которые ориентированы именно на такой формат представления информации.

Перспективными направлениями дальнейших работ являются исследование характеристик сходства выходных векторов, обеспечивающих большую точность оценивания сходства входных векторов, анализ свойств других типов проецирования, исследование полученных векторных представлений в различных задачах.

Д.А. Рачковский выражает благодарность А.А. Фролову и И.П. Муравьеву за плодотворные обсуждения.

#### СПИСОК ЛИТЕРАТУРЫ

1. Salton G. Automatic text processing: The transformation, analysis, and retrieval of information by computer. — Reading, MA: Addison-Wesley, 1989. — 530 p.
2. Мисуно И.С., Рачковский Д.А., Слипченко С.В., Соколов А.М. Поиск текстовой информации с помощью векторных представлений // Проблемы программирования. — 2005. — № 4. — С. 50–59.
3. Куссуль Э.М. Ассоциативные нейроподобные структуры. — Киев: Наук. думка, 1992. — 144 с.
4. Рачковский Д.А. Моделирование мышления как путь повышения интеллектуальности информационных технологий и систем // Системные технологии. — 2008. — № 4(57). — С. 154–162.

5. Rachkovskij D.A. Linear classifiers based on binary distributed representations // Intern. J. Inform. Theories & Appl. — 2007. — **14**, N 3. — P. 270–274.
6. Рачковский Д.А., Мисуно И.С., Ревунова Е.Г., Слипченко С.В., Соколов А.М. Концепция и методы нейросетевого распределенного представления информации в задачах ИИ // 14-я Междунар. конф. «Проблемы нейрокибернетики». — Т. 2. — Ростов-на-Дону, Россия, 2005. — С. 30–33.
7. Рачковский Д.А., Слипченко С.В., Фролов А.А., Гусек Д. Разрешающая способность бинарного кодирования числовых векторов гиперпрямоугольными рецептивными полями // Кибернетика и системный анализ. — 2005. — № 5. — С. 3–16.
8. Kanerva P., Sjodin G., Kristoferson J., Karlsson R., Levin B., Holst A., Karlgren J. and Sahlgren M. Computing with large random patterns // Foundations of Real-World Intelligence. — CSLI Publications, Stanford, California, 2001. — P. 251–311.
9. Johnson W.B., Lindenstrauss J. Extensions of Lipschitz mapping into Hilbert space // Contemporary Mathemat. — 1984. — **26**. — P. 189–206.
10. Vempala S.S. The Random Projection Method. — Providence, RI: American Mathemat. Soc., 2004. — 105 p.
11. Achlioptas D. Database-friendly random projections: Johnson–Lindenstrauss with binary coins // J. of Comput. and System Sci. — 2003. — **66**, N 4. — P. 671–687.
12. Charikar M. Similarity estimation techniques from rounding algorithms // ACM Symp. on Theory of Comput. — Montreal, Canada: ACM, 2002. — **1**. — P. 380–388.
13. Liu K., Kargupta H., Ryan J. Random projection-based multiplicative data perturbation for privacy preserving distributed data mining // IEEE Transact. on Knowledge and Data Engineer. — 2006. — **18**, N 1. — P. 92–106.
14. Li P., Hastie T.J., Church K.W. Very sparse random projections // 12th ACM SIGKDD Intern. Conf. on Knowledge Discovery and Data Mining. — Philadelphia, PA, USA: ACM Press, 2006. — P. 287–296.
15. Michel X.G., David P.W. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming // J. of the Associat. for Comput. Machinery. — 1995. — **42**, N 6. — P. 1115–1145.
16. Rachkovskij D.A. Representation and processing of structures with binary sparse distributed codes // IEEE Transac. on Knowledge and Data Engineer. — 2001. — **13**, N 2. — P. 261–276.
17. Rachkovskij D.A. Some approaches to analogical mapping with structure sensitive distributed representations // J. of Experiment. and Theoret. Artificial Intell. — 2004. — **16**, N 3. — P. 125–145.
18. Rachkovskij D.A., Kussul E.M. Binding and normalization of binary sparse distributed representations by context-dependent thinning // Neural Comput. — 2001. — **13**, N 2. — P. 411–452.
19. Рачковский Д.А., Слипченко С.В., Куссуль Э.М., Байдык Т.Н. Процедура связывания для бинарного распределенного представления данных // Кибернетика и системный анализ. — 2005. — № 3. — С. 3–18.
20. Фролов А.А., Гусек Д., Рачковский Д.А. Время поиска по сходству в ассоциативной памяти для бинарных векторов // Там же. — 2006. — № 5. — С. 3–13.
21. Вентцель Е.С. Теория вероятностей. — М.: Наука, 1969. — 576 с.
22. Kussul E.M., Baidyk T.N., Kasatkina L.M., Lukovich V.V. Rosenblatt perceptrons for handwritten digit recognition // Intern. Joint Conf. on Neural Networks. — Washington, DC: IEEE, 2001. — **2**. — P. 1516–1521.
23. Kussul E., Baidyk T., Wunsch D., Makeyev O., Martin A. Permutation coding technique for image recognition systems // IEEE Transact. on Neural Networks. — 2006. — **17**, N 6. — P. 1566–1579.
24. Мисуно И.С., Рачковский Д.А., Слипченко С.В. Экспериментальное исследование классификации рукописных цифр // Системные технологии. — 2005. — № 4(39). — С. 110–133.
25. Амосов Н.М., Байдык Т.Н., Гольцев А.Д., Касаткин А.М., Касаткина Л.М., Куссуль Э.М., Рачковский Д.А. Нейрокомпьютеры и интеллектуальные роботы. — Киев: Наук. думка, 1991. — 272 с.

*Поступила 07.07.2010*