

СВОЙСТВА КОДИРОВАНИЯ ЧИСЛОВЫХ ВЕЛИЧИН СЛУЧАЙНЫМИ ГИПЕРПРЯМОУГОЛЬНЫМИ РЕЦЕПТИВНЫМИ ПОЛЯМИ

Abstract: Properties of coarse coding with random hyperrectangle receptive fields are considered for two schemes: random subspace coding RSC and Prager. Following codes characteristics are provided: effective receptive fields dimensionality, code density distribution in the input space, code overlap, and others. The results of theoretical analysis are confirmed by experiments.

Key words: neural networks, coarse coding, distributed encoding, receptive fields, hyperrectangle, CMAC, RSC.

Анотація: Розглянуто характеристики грубого кодування за допомогою випадково розподілених гіперпрямокутних рецептивних полів для двохпорогової схеми випадкових підпросторів RSC та однопорогової схеми Prager. Отримано розподіли ймовірностей розмірності рецептивних полів, вирази щільності розподілу у різних точках вхідного простору, характеристики перекриття кодів та ін. Результати порівняльного аналізу підтверджені експериментально.

Ключові слова: нейронні мережі, грубе кодування, розподілені представлення, рецептивні поля, гіперпрямокутник, CMAC, RSC.

Аннотация: Рассматриваются характеристики грубого кодирования с помощью случайных гиперпрямоугольных рецептивных полей для двухпороговой схемы случайных подпространств RSC и однопороговой схемы Prager. Получены распределения эффективной размерности рецептивных полей, выражения для плотности кода в разных точках входного пространства, характеристики перекрытия кодов, и др. Результаты сравнительного анализа подтверждены экспериментально.

Ключевые слова: нейронные сети, грубое кодирование, распределенные представления, рецептивные поля, гиперпрямоугольник, CMAC, RSC.

1. Введение

Кодирование состояний или величин указанием их принадлежности некоторым интервалам (рецептивным полям признаков) называют грубым или популяционным [1]. Схемы кодирования с рецептивными полями широко известны – это кодирование в первом слое персептрона Розенблатта, радиальные базисные функции, CMAC и многие другие схемы [1–3]. Причины интереса к таким моделям кодирования состоят в том, что получаемые коды успешно применялись в задачах классификации, аппроксимации, поиска, оценки плотности вероятности и др. ([3–11] и содержащиеся там ссылки), а также их нейробиологическая релевантность [1].

Будем рассматривать грубое бинарное кодирование векторов с действительными элементами $D_i \in [0,1]$. Такие векторы D удобно представлять как точки A -мерного входного пространства первичных признаков. Результатом кодирования является N -мерный бинарный кодвектор вторичных признаков. Каждый его бит представляет выход "нейрона", который устанавливается в единичное состояние при попадании входного вектора в его рецептивное поле. Некоторые из рецептивных полей могут пересекаться, и D одновременно активирует множество полей. Его малые изменения приводят к выходу из зоны действия некоторых рецептивных полей и попаданию в другие, что изменяет соответствующий кодвектор. Однако близкие входные векторы активируют много одинаковых полей, поэтому их кодвекторы имеют много совпадающих единиц.

Кодирующие схемы с рецептивными полями, ограниченными (гипер)плоскостями, параллельными осям координат, вычислительно эффективны, так как реализуются на основе простых операций сравнения. Кроме того, такая "метрика гиперпрямоугольников" хорошо подходит для работы в пространствах большой размерности. К таким схемам относятся CMAC [2, 3], схема

Prager [4–7], разрабатываемые нами схемы кодирования со случайными порогами и случайными подпространствами (RTC и RSC, Random Subspace Coding) [12–15] и их модификации [16, 17].

В схеме RSC рецептивное поле нейрона состоит из совокупности S одномерных рецептивных полей в разных измерениях. Одномерное рецептивное поле каждого измерения определяется его границами – верхним и нижним порогом. Значение верхнего порога определяется прибавлением к координате центра поля величины его полустороны, а нижнего – вычитанием. Центр каждого одномерного поля выбирается случайно, в пределах интервала изменения входных координат – входного интервала $[0, R]$ (обычно $R=1$). Длина каждой полустороны g также выбирается случайно в интервале $[0, G]$. Для каждого из N S -мерных полей номера S входных измерений из A также выбираются случайно и без повторения. В каждом из этих S измерений выбираются два порога, определяющие длину стороны гиперпрямоугольника в этом измерении:

$$h_{i,j} = x_{i,j} + g_{i,j}; \quad l_{i,j} = x_{i,j} - g_{i,j}; \quad (1)$$

где h – координата верхнего порога; l – нижнего; i – номер нейрона; j – номера проверяемых нейроном измерений, общее число которых для одного нейрона не превышает S .

Процедура генерации порогов повторяется для всех нейронов-признаков, сгенерированные пороги запоминаются. При кодировании точки D бит, соответствующий i -му нейрону, устанавливается в 1, если выполняется $l_{i,j} \leq D_j \leq h_{i,j}$ для всех S проверяемых измерений i -го нейрона и координат D_j . Аналогично устанавливаются биты, соответствующие всем N нейронам, в результате чего формируется бинарный кодвектор X вторичных признаков, содержащий M единиц – результат кодирования D . Матожидание M/N называют плотностью кода p .

Prager и др. [4–7] независимо разработали схему кодирования, которую можно рассматривать как однопороговый вариант RSC. Однако это различие приводит к существенно разным свойствам производимых кодов. Задача данной работы состоит в получении и исследовании характеристик и свойств кодов схемы стохастического грубого кодирования Prager и их сравнительном анализе со схемой RSC. Рассматриваются распределения размерностей рецептивных полей, характеристики плотности и перекрытия кодов в разных точках пространства и др. Все результаты подтверждены экспериментами на кодах большой размерности.

2. Реальная размерность рецептивных полей

При $G > 0,5$ одномерные рецептивные поля в некоторых из S измерений могут быть сгенерированы так, что накроют входной интервал $[0, R]$, то есть $\forall D_j : l_{i,j} \leq D_j \leq h_{i,j}$. Соответственно, эти $S - s$ измерений не будут влиять на активацию S -мерного рецептивного поля. Активация будет определяться s одномерными полями, пересекающимися, но не накрывающими входной интервал. Найдем закон распределения реальной размерности полей.

Рассмотрим схему кодирования Prager [7], где входной интервал $[0, R]$, а центры рецептивных полей равномерно распределены на интервале $[0 - G, R + G]$ с плотностью

$1/(R + 2G)$ (рис. 1). Полурасстояние между порогами для всех полей фиксировано: $g_{i,j} = G > 0,5R; i = 1, N; j = 1, A$. При этом в $[0, R]$ оказывается либо один порог (одномерное рецептивное поле пересекает входной интервал), либо ни одного (поле накрывает его).

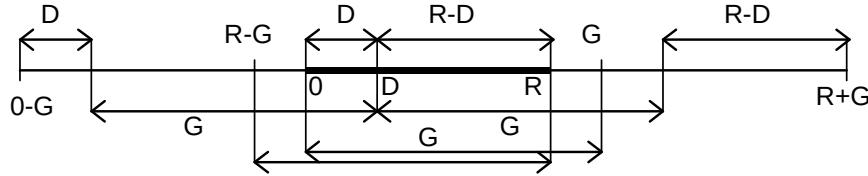


Рис. 1. Схема кодирования Prager

Определим вероятность p_{in} пересечения входного интервала $[0, R]$ (одним порогом, так как $G > 0,5R$). Так как поля с центрами в $[R - G, G]$ накрывают входной интервал,

$$p_{in} = 2R / (2G + R). \quad (2)$$

Длина полустороны рецептивных полей, обеспечивающая заданную p_{in} :

$$G = R / p_{in} - 0,5R. \quad (3)$$

Такая схема генерации соответствует последовательности S независимых испытаний по схеме Бернулли. Поэтому доля рецептивных полей, в которых s одномерных полей из S пересекают входной интервал, имеет биномиальное распределение

$$P_{in}(s; S, p_{in}(G)) = \beta(s; S, p_{in}) \equiv C_S^s p_{in}^s (1 - p_{in})^{S-s}. \quad (4)$$

Графики $P_{in}(s)$ для разных значений параметров $S < A$ и p_{in} приведены на рис. 2а. Так как нет полей с большой "реальной размерностью" s , нет смысла хранить и проверять все A порогов. В [6, 7] было предложено генерировать рецептивные поля по следующему алгоритму. По распределению Пуассона $\pi(s; p_{in}A)$ выбирается доля полей с реальной размерностью s . Для каждого поля случайно выбираются номера s измерений из множества $\{1, 2, \dots, A\}$. В каждом из выбранных полей случайно выбирается координата порога в $[0, R]$ и с какой стороны от него будет находиться рецептивное поле. Далее будет использоваться биномиальное распределение реальной размерности полей.

В схеме RSC одномерные поля являются двухпороговыми, с центрами $x_{i,j} \in [0, 1]$ и длиной полусторон $g_{i,j} \in [0, G]$, $i = 1, N$, $j = 1, A$. Поэтому в S -мерных рецептивных полях нейронов возможны комбинации различных типов одномерных рецептивных полей:

$$S = s_1 + s_2 + s_0, \quad (5)$$

где s_1 , s_2 , – число полей, пересекающих входной интервал одним и двумя порогами соответственно; s_0 – число полей, накрывающих входной интервал. Реальная размерность $s = s_1 + s_2$ имеет биномиальное распределение $\beta(s; S, p_{in}(G))$.

Размеры (g – полурасстояние между его порогами) и положение (x – координата центра поля) одномерных рецептивных полей в схеме RSC образуют двумерный случайный вектор (X, G) . Вероятность попадания двумерного случайного вектора в двумерную область S^* , задающую ограничения на допустимые сочетания x и g и соответствующую определенным событиям (пересечение входного интервала, накрывание точки и т.п.), вычисляется как [18]:

$$p((X, Y) \in S^*) = \iint_{S^*} p_{XY}(x, y) dx dy, \quad (6)$$

где $p_{XY}(x, y)$ – плотность распределения двумерного случайного вектора (в нашем случае – вектора независимых равномерно распределенных случайных величин).

Все одномерные рецептивные поля со всеми случайными сочетаниями $X \in [0, 1]$ и $G \in [0, G]$ равномерно распределены в прямоугольной области $[0, 1] \times [0, G]$. Поэтому $p_{XG}(x, g) = 1/G$. Конкретизация области S^* , то есть конкретного вида пределов интегрирования, производится для конкретной задачи. В частности, для значений p_{in} при разных G :

$$p_{in} = 1 \text{ при } 0 < G \leq 0,5;$$

$$p_{in} = \int_0^{1-G} \int_0^G \frac{1}{G} dg dx + \int_{1-G}^{0,5} \int_0^{1-x} \frac{1}{G} dg dx + \int_{0,5}^G \int_0^x \frac{1}{G} dg dx + \int_G^1 \int_0^G \frac{1}{G} dg dx = 2 - G - 0,25G \text{ при } 0,5 < G < 1; \quad (7)$$

$$p_{in} = \int_0^{0,5} \int_0^{1-x} \frac{1}{G} dg dx + \int_{0,5}^1 \int_0^x \frac{1}{G} dg dx = 0,75/G \text{ при } G > 1.$$

Более подробно геометрия S^* рассмотрена в [19]. Вид распределения реальной размерности $P_{in}(s)$ для разных G приведен на рис. 2б.

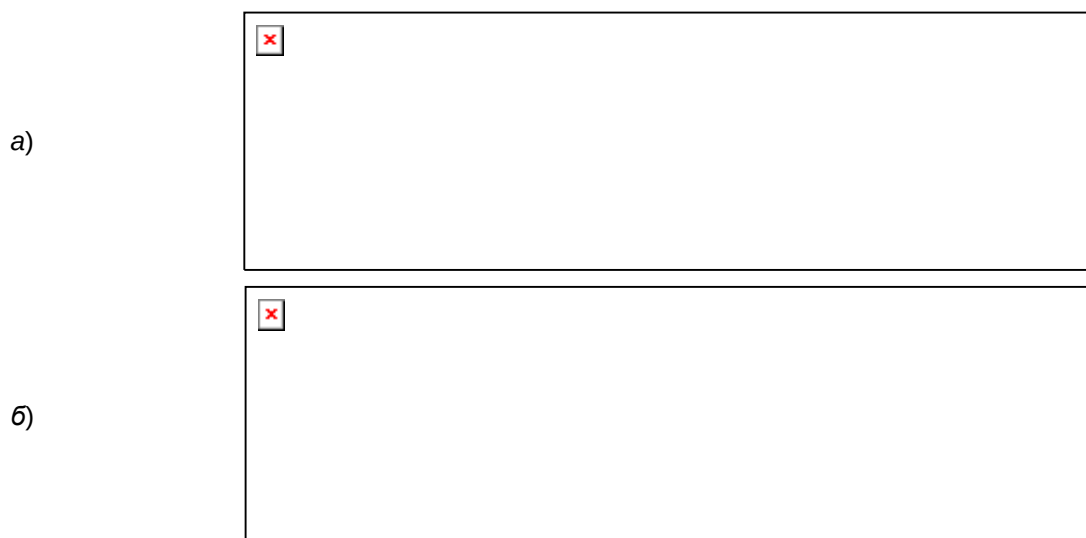


Рис. 2. Распределение «реальной» размерности рецептивных полей $P_{in}(s)$: а) схема Prager; б) схема RSC

Таким образом, в схемах RSC и Prager размерность рецептивных полей зависит от их максимальной размерности S и не зависит от размерности входного пространства A . Вероятность пересечения входного интервала одномерными полями уменьшается с ростом S и G . Соответственно растет доля одномерных полей, накрывающих входной интервал. Поэтому реальная размерность большинства полей ниже максимальной $S \leq A$. Эта тенденция особенно проявляется при увеличении G и S . При фиксированной плотности кода в центральной точке для схемы Prager снижение реальной размерности больше, чем для RSC (см. табл. 1).

Таблица 1. Примеры максимальной и средней s для схем Prager и RSC

	$p = 0,01 \quad S = 100$		$p = 0,1 \quad S = 10 - 100$		$p = 0,5 \quad S = 5 - 100$	
	max	mean	max	mean	max	mean
Prager	16	9	10	4	5	1
RSC	20	14	14	7	5	2

3. Активация рецептивных полей и плотность кода

Плотность кода p (доля $p = M/N$, активируемых входной точкой D S -мерных рецептивных полей или вероятность активности нейрона) влияет на его обобщающую способность, определяет затраты обычной памяти, емкость ассоциативной памяти, скорость операций при реализации и др. Для кодирования RSC p зависит от координат D , а для кодирования Pragera не зависит.

Определим вероятность p_{act} того, что произвольное одномерное рецептивное поле накроет точку с координатой D ($0 \leq D \leq R$). Это же можно трактовать как долю одномерных полей, в которые попадет произвольная точка с координатой D ("точка D "). Центры полей, не накрывающих точку D нижними порогами, лежат в интервале $[D + G, R + G]$. Центры полей, не накрывающих точку D верхними порогами, лежат в интервале $[D - G, 0 - G]$.

$$1 - p_{act} = \int_{D+G}^{R+G} \frac{1}{R+2G} dx + \int_{0-G}^{D-G} \frac{1}{R+2G} dx = \frac{R}{R+2G}; \quad p_{act} = \frac{2G}{R+2G}. \quad (8)$$

Для схемы Prager и других однопороговых схем p_{act} не зависят от координаты D . Сопоставляя p_{act} с $p_{in} = 2R/(2G + R)$ (разд. 2), получаем следующие соотношения:

$$p_{in} = 2(1 - p_{act}); \quad p_{act} = 1 - 0,5p_{in}. \quad \text{При } R=1 \quad p_{act} > 0,5, \text{ т.к. } G > 0,5R. \quad (9)$$

Если все одномерные поля сгенерированы с одинаковым G (т.е. с одинаковой p_{in} , приводящей к биномиальному распределению реальной размерности многомерных рецептивных полей, см. разд. 2), то p_{act} одинаковы для всех полей. Плотность кода при этом постоянна для любой D и равна

$$p = p_{act}^S. \quad (10)$$

Параметры генерации полей для получения заданной p определяются как

$$p_{in}(p) = 2(1 - p^{1/S}); G(p) = R / p_{in} - 0,5R = 0,5R / (1 - p^{1/S}) - 0,5R. \quad (11)$$

Для схемы RSC поля с сочетаниями координат центров x и длины полусторон g , позволяющими покрыть D , образуют область, форма и размер которой зависят от координаты D и ее соотношения с G . Рассмотрим различные комбинации D и G при $D \leq 0,5$ (случай $D > 0,5$ симметричен). Анализ соотношений x и g для полей, накрывающих D , позволяет определить пределы двойного интеграла для вычисления p_{act} (см. разд. 2 и [19]).

$$p_{act} = \int_{D-G}^D \int_{D-x}^G \frac{1}{G} dg dx + \int_D^{D+G} \int_{x-D}^G \frac{1}{G} dg dx = G \quad \text{при } 0 < G \leq D;$$

$$p_{act} = \int_0^D \int_{D-x}^G \frac{1}{G} dg dx + \int_D^{D+G} \int_{x-D}^G \frac{1}{G} dg dx = 0,5G + D - 0,5D^2 / G \quad \text{при } D < G \leq 1 - D; \quad (12)$$

$$p_{act} = \int_0^D \int_{D-x}^G \frac{1}{G} dg dx + \int_D^{D+G} \int_{x-D}^G \frac{1}{G} dg dx = 1 + (D - 0,5 - D^2) / G \quad \text{при } G > 1 - D.$$

Таким образом, p_{act} зависит от D . Для фиксированного D и точек на главной диагонали $p_{act}(D)$ равны для всех одномерных полей многомерного поля, поэтому

$$p(D) = (p_{act}(D))^S. \quad (13)$$

Выражения для значения плотности кода в произвольных точках пространства приведены в [19]. Заметим также, что $p_{act}(D)$ является суммой вероятностей следующих несовместных событий: $1 - p_{in}$ – вероятность того, что поле накрывает входной интервал (такие поля всегда активны), и $p_{in \& act}(D)$ – вероятность того, что поле пересекает входной интервал и накрывает D :

$$p_{act}(D) = 1 - p_{in} + p_{in \& act}(D). \quad (14)$$

Графики зависимости G и p_{in} от S и p приведены в [19]. Для получения кодов с заданной p для D на главной диагонали вычисляется $p_{act}(D) = p(D)^{1/S}$ и соответствующее G .

$$G = p_{act} \quad \text{при } p_{act} < D;$$

$$G = (p_{act} - D) + ((D - p_{act})^2 + D^2)^{1/2} \quad \text{при } p_{act} < (0,5 - D^2) / (1 - D); \quad (15)$$

$$G = (0,5 - D + D^2) / (1 - p_{act}) \quad \text{при } p_{act} > (0,5 - D^2) / (1 - D).$$

Графики отношения плотности кода в D к максимальной плотности кода в центральной точке $p(D^A) / p(0,5)$ для точек главной диагонали при разных S приведены на рис. 3.

Таким образом, для однопороговых схем плотность кода p постоянна во всех точках входного пространства и ее значением можно управлять либо выбором p_{in} для независимой генерации полей, либо подбором вида распределения P_{in} . Однако возможности управления

плотностью для однопороговых схем являются ограниченными, т.к. $p_{act} \geq 0,5$ и соответственно $p \geq 0,5^S$. Поэтому для маломерных рецептивных полей низкая плотность кода недостижима.

Для двухпороговых схем величину p можно сделать произвольно низкой путем выбора G , однако необходимо учитывать, что в углах входного пространства она меньше, чем в центре. Для $G > 0,5$ плотность кода максимальна в центральной точке, а для $G < 0,5$ постоянна в пределах A -мерной области $[G, 1-G]^A$, т.к. $0 < G \leq D \leq 0,5$. За её пределами она нелинейно уменьшается. Для некоторых задач сильная неравномерность плотности кода может быть нежелательной и требуются некоторые меры для обеспечения ее постоянства или уменьшения неравномерности. Перемасштабированием значений входного интервала $[0,1]$ при $G < 0,5$ можно обеспечить постоянную плотность кода, а при $G > 0,5$ – уменьшить ее неравномерность. Изучение того, как изменятся характеристики рецептивных полей и кодов при перемасштабировании и других вариациях кодирования, требует отдельного исследования (см. также разд. «Выводы»).

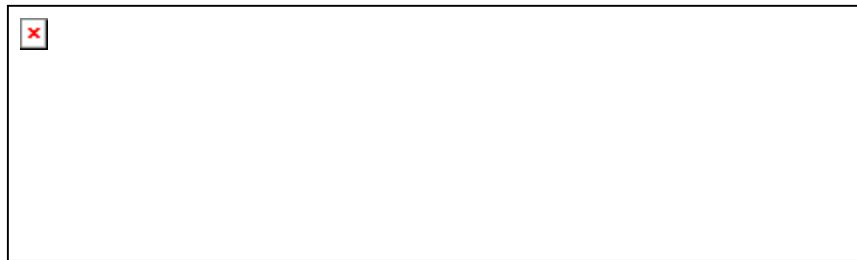


Рис. 3. Относительная плотность кода $p(D^A) / p(0,5)$ для схемы RSC

4. Размерность рецептивных полей активных нейронов

Обозначим распределение реальной размерности полей активных нейронов (доля s измерений рецептивных полей активных нейронов, в которых один или оба порога принадлежат входному интервалу) как $P_{in|act}(s)$. Оно значительно отличается от аналогичного распределения для всех нейронов $P_{in}(s)$ (разд. 2). Пусть $p_{act|in}(G, D)$ – вероятность того, что точка D попала в одномерное рецептивное поле при условии, что поле пересекает входной интервал (рецептивные поля сгенерированы с параметром G). Для схем генерации порогов, обеспечивающих равные значения $p_{act|in}(D)$ для всех измерений, вероятность активации произвольного s -мерного рецептивного поля точкой D на главной диагонали равна $p_{act|in}^s$. Закон распределения $P_{in|act}(s)$ случайной величины s для нейронов, активированных D на главной диагонали

$$P_{in|act}(s) = P_{in}(s) p_{act|in}^s / \sum_{i=0,S} P_{in}(i) p_{act|in}^i, \quad (16)$$

где $\sum_{i=0,S} P_{in}(i) p_{act|in}^i = p$ – вероятность активности нейрона (см. разд. 3).

Для биномиального распределения $P_{in}(s)$ можно показать, что

$$P_{in|act}(s) = \beta(s; S, p_{in|act}). \quad (17)$$

Если в однопороговой схеме Prager порог одномерного рецептивного поля принадлежит входному интервалу, то вероятность того, что поле накроет точку D , равна

$$p_{act|in} = p_U p_{act|U} + p_L p_{act|L}, \quad (18)$$

где $p_L = p_U = 0,5$ – вероятности того, что соответственно нижний или верхний порог принадлежит входному интервалу; $p_{act|L}$ и $p_{act|U}$ – вероятности того, что поле накроет точку D при условии, что его порог соответственно нижний или верхний.

$$p_{act|L} = \int_0^D dx / R = D; p_{act|U} = \int_D^1 dx / R = 1 - D. \quad (19)$$

Следовательно, $p_{act|in}$ не зависит от положения точки D :

$$p_{act|in} = 0,5D + 0,5(1 - D) = 0,5;$$

$$p_{in|act} = p_{in} p_{act|in} / p_{act} = 0,5 p_{in} / p_{act} = p_{in} / (2 - p_{in}) = (1 - p_{act}) / p_{act}. \quad (20)$$

Распределения $p_{in|act}(s)$ реальной размерности рецептивных полей для активных нейронов при биномиальном $P_{in}(s)$ приведены на рис. 4а. Для однопороговых схем с произвольным $P_{in}(s)$ для определения $P_{in|act}(s)$ следует использовать формулу (16).

а)



б)



в)



Рис. 4. Распределение $P_{in|act}(s)$ «реальной» размерности полей, активируемых точкой D : а) схема Prager $D = 0,5$; б) схема RSC $D = 0,5$; в) схема RSC $D = 1$

Для определения $p_{in|act} = p_{in\&act} / p_{act}$ в схеме RSC достаточно найти $p_{in\&act}(G, D)$, так как $p_{act}(G, D)$ известна (разд. 3). Выражения для $p_{in\&act}$, полученные вычислением интегралов, приведены в [19]. Также они могут быть получены из p_{in} и p_{act} по формуле (14):

$$p_{in|act}(D) = p_{in\&act} / p_{act} = 1 - (1 - p_{in}) / p_{act}(D). \quad (21)$$

Распределение $P_{in|act}(s)$ реальной размерности рецептивных полей для активных нейронов при положении точки D на диагонали приведено на рис. 4 б, в. Видно, что максимальная реальная размерность активированных рецептивных полей еще меньше, чем максимальная реальная размерность всех рецептивных полей, которая, в свою очередь, меньше S и A . Это также следует из $p_{in|act} \leq p_{in}$, так как $p_{in|act} = p_{in\&act} / p_{act} = (p_{act} + p_{in} - 1) / p_{act} = 1 - (1 - p_{in}) / p_{act}$, поэтому

$$(1 - p_{in}) \leq (1 - p_{in}) / p_{act} \text{ и } p_{in|act} = 1 - (1 - p_{in}) / p_{act} \leq 1 - (1 - p_{in}) = p_{in}. \quad (22)$$

Таким образом, признаки меньшей размерности превалируют в коде. Это особенно проявляется при увеличении p и S и удалении от центра интервала (см. табл. 2).

Таблица 2. Примеры максимальной и средней s для схем Prager и RSC

	$p = 0,01 \quad S = 100$		$p = 0,1 \quad S = 10 - 100$		$p = 0,5 \quad S = 5 - 100$	
	max	mean	max	mean	max	mean
Prager	10	4	7	2	4	0
RSC: $D = 0,5$	16	9	10	4	5	1
RSC: $D = 1$	10	4-5	7	2	4	0

5. Зависимость перекрытия кодов от расстояния между кодируемыми точками

Одной из важнейших характеристик кодирования является перекрытие кодов (зависимость величины перекрытия кодов от координат точек D_1 и D_2). Эта характеристика играет существенную роль при решении задач в разных предметных областях. Будем использовать следующую меру относительного перекрытия кодов X и Y :

$$V(X, Y) = |X \& Y| / |X|, \quad (23)$$

где $|X|$ – число единиц в X ; $\&$ – побитовая конъюнкция.

Обозначим точки главной диагонали с разными координатами как D_1 и D_2 . Если $X = X(D_1)$ а $Y = Y(D_2)$, то при увеличении числа N бит в коде значение перекрытия стремится к отношению вероятностей:

$$V(X, Y) = p(D_1, D_2) / p(D_1), \quad (24)$$

где $p(D_1, D_2) \equiv p(X(D_1) \& Y(D_2)) \approx |X \& Y| / N$ и $p(D_1) \approx |X| / N$.

Вероятности $p(D_1)$ и $p(D_1, D_2)$ параметризованы G . Для произвольных распределений размерности полей $P_{in}(i)$ с равными $p_{act|in}(D_1, D_2)$ для всех измерений плотность перекрытия кодов для точек на диагонали можно определить как

$$p(D_1, D_2) = \sum_{i=0, S} P_{act|in}(i, D_1, D_2) = \sum_{i=0, S} P_{in}(i) (p_{act|in}(D_1, D_2))^i, \quad (25)$$

где $p_{act|in}(D_1, D_2)$ – вероятность того, что одномерное рецептивное поле покрывает точки D_1 и D_2 при условии, что поле пересекает входной интервал.

Вероятность $p_{act|in}(D_1, D_2)$ может быть вычислена как

$$p_{act|in}(D_1, D_2) = p_{act \& in}(D_1, D_2) / p_{in}. \quad (26)$$

Если D_1 и D_2 не лежат на диагонали, вероятность накрыть A -мерный отрезок $[D_1, D_2]$ произвольным S -мерным рецептивным полем определяется как средняя по всем полям [19]:

$$p(D_1, D_2) = \sum_{k \in K} p_k(D_1, D_2) p(k), \quad (27)$$

где $k \in K$ индексирует все C_A^S конфигураций рецептивных полей; $p_k(D_1, D_2) = \prod_{i=1, S} p_{act}(D_{1,k(i)}, D_{2,k(i)})$ – вероятность накрыть $[D_1, D_2]$ полем с k -й конфигурацией полей (активировать его обеими этими точками); $p_{act}(D_{1,k(i)}, D_{2,k(i)})$ – вероятность накрыть одномерным рецептивным полем интервал $[D_{1,k(i)}, D_{2,k(i)}]$; $k(i)$ – номер i -го измерения поля с k -й конфигурацией, в которой у него имеется одномерное рецептивное поле.

Для точек D_1 и D_2 , лежащих на главной диагонали, для независимо сгенерированных полей с одинаковой p_{in} и одинаковой $p_{act|in}(D_1, D_2)$ для всех измерений справедливо

$$\begin{aligned} p_{act}(D_{1,k(i)}, D_{2,k(i)}) &= p_{act}(D_1, D_2); \\ p(D_1, D_2) &= (p_{act}(D_1, D_2))^S, \end{aligned} \quad (28)$$

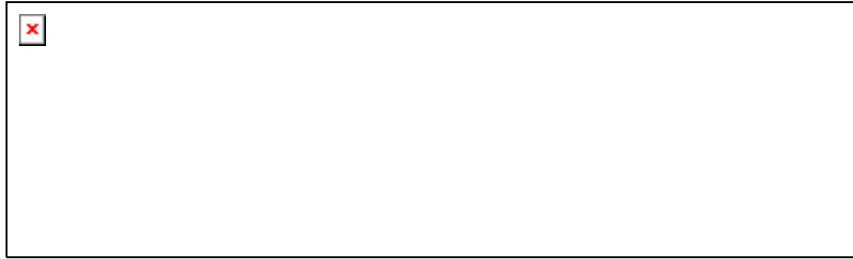
где $p_{act}(D_1, D_2)$ – вероятность одномерного рецептивного поля накрыть точки D_1 и D_2 , то есть активироваться обеими этими точками.

Для схемы Prager определим вероятность $1 - p_{act}(D_1, D_2)$ того, что интервал (D_1, D_2) не будет накрыт одномерным рецептивным полем. Центры полей, не накрывающих интервал (D_1, D_2) ($0 \leq D_1 \leq D_2 \leq R$) нижними порогами, лежат в интервале $[D_1 + G, R + G]$, а правыми – в интервале $[0 - G, D_2 - G]$.

$$\begin{aligned} 1 - p_{act}(D_1, D_2) &= \int_{D_1+G}^{R+G} \frac{1}{R+2G} dx + \int_{0-G}^{D_2-G} \frac{1}{R+2G} dx = (R + \Delta D) / (R + 2G); \\ p_{act}(\Delta D) &= (2G - \Delta D) / (R + 2G), \Delta D = D_2 - D_1. \end{aligned} \quad (29)$$

На рис. 5 а, б показаны графики относительного перекрытия кодов $V(D^A, D^*)$ (28) для точек главной диагонали D , а на рис. 6 а – для произвольных точек двумерного пространства.

а)



б)



в)



г)



Рис. 5. Перекрывание $V(D^A, D^*)$ кодов точек D^* и D на главной диагонали: а) схема Prager $D^* = 0,5$; б) схема Prager $D^* = 1$; в) схема RSC $D^* = 0,5$; г) схема RSC $D^* = 1$

Для схемы RSC вычисление вероятности $p_{act}(D_1, D_2)$ через $p_{act \& in}(D_1, D_2)$ для определения $p(D_1, D_2)$ по формулам (25), (26) рассмотрено в [19]. Значение вероятности $p_{act}(D_1, D_2)$ также можно определить непосредственным вычислением интегралов:

$$\begin{aligned}
 p_{act}(D_1, D_2) &= 0 \text{ при } G < (D_2 - D_1)/2; \\
 p_{act}(D_1, D_2) &= \int_{D_2-G}^{(D_1+D_2)/2} \int_{D_2-x}^G \frac{1}{G} dgdx + \int_{(D_1+D_2)/2}^{D_1+G} \int_{x-D_1}^G \frac{1}{G} dgdx = \\
 &= 0,25(2G + D_1 - D_2)^2 / G \text{ при } G < 1 - D_1 \text{ и } G < D_2; \\
 p_{act}(D_1, D_2) &= \int_0^{(D_1+D_2)/2} \int_{D_2-x}^G \frac{1}{G} dgdx + \int_{(D_1+D_2)/2}^{D_1+G} \int_{x-D_1}^G \frac{1}{G} dgdx = \\
 &= 0,25(D_1^2 - 2D_1D_2 - D_2^2 + 4D_1G + 2G^2) / G \text{ при } G < 1 - D_1 \text{ и } G > D_2;
 \end{aligned}
 \tag{30}$$

$$p_{act}(D_1, D_2) = \int_{D_2-G}^{(D_1+D_2)/2} \int_{D_2-x}^G \frac{1}{G} dgdx + \int_{(D_1+D_2)/2}^1 \int_{x-D_1}^G \frac{1}{G} dgdx =$$

$$= 0,25(D_2^2 - D_1^2 - 2D_1D_2 - 4D_2G + 2G^2 - 2 + 4D_1 + 4G)/G \text{ при } G > 1 - D_1, G < D_2;$$

$$p_{act}(D_1, D_2) = \int_0^{(D_1+D_2)/2} \int_{D_2-x}^G \frac{1}{G} dgdx + \int_{(D_1+D_2)/2}^1 \int_{x-D_1}^G \frac{1}{G} dgdx =$$

$$= 0,25(4D_1 + 4G - D_1^2 - 2D_1D_2 - D_2^2 - 2)/G \text{ при } G > 1 - D_1 \text{ и } G > D_2.$$

Из этих формул, полагая $D_1 = D_2 = D$, можно получить выражения для $p_{act}(D)$.

На рис. 5 в, г показаны графики относительного перекрытия кодов $V(D^A, D^*)$ для точек главной диагонали D , а на рис. 6 б – для произвольных точек двумерного пространства.

Таким образом, для схем Prager и RSC, при удалении точек друг от друга по прямой, перекрытие соответствующих кодов – нелинейная монотонно убывающая функция.

Для схемы Pragera и одномерного поля перекрытие кодов убывает линейно с ΔD , при этом существует ненулевое перекрытие для любых точек входного интервала $p_{act}(\Delta D = 1) = 2p_{act} - 1$. Для многомерных полей реальная величина перекрытия быстро приближается к нулю. Характеристика перекрытия в пределах входного интервала симметрична и одинакова для всех D .

В схеме RSC при $\Delta D > 2G$ и $G < 0,5$ из-за локальности полей коды не имеют перекрытия. При $\Delta D < 2G$ перекрытие является нелинейной убывающей функцией от ΔD . Характеристика перекрытия несимметрична и неодинакова при изменении D и фиксированных других параметров.

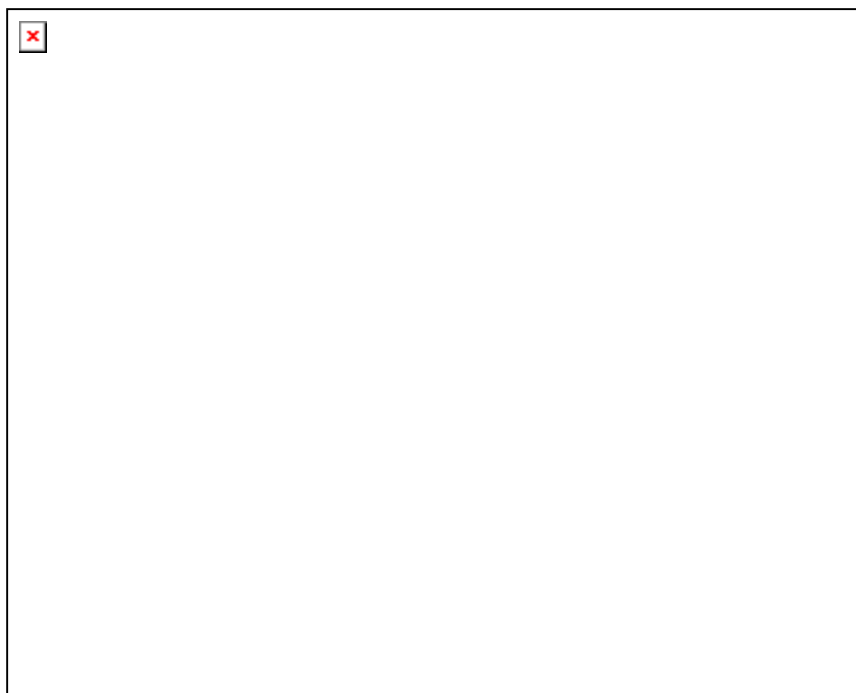
Характеристика перекрытия кодов при одинаковой плотности кода и S для схемы RSC гораздо более крутая, чем для схемы Prager. С ростом плотности кода характеристика перекрытия становится более полой. Ненулевое перекрытие кодов во всем интервале может быть явным критерием того, что код «слишком» плотен, так как имеются всегда активные поля, не несущие информации об изменении положения кодируемой точки входного пространства. Правильный выбор характеристики перекрытия кодов (в сочетании с плотностью кода и размерностью полей) очень важен для задач классификации, аппроксимации и др., так как от него во многом зависит качество решения (см. разд. «Выводы»).

7. Выводы

В данной работе исследованы характеристики кодов, которые дают схемы стохастического грубого кодирования RSC и Prager. Характеристики схемы Prager получены впервые, а для схемы RSC, подробно рассмотренной в [19], ряд характеристик удалось получить более простым способом.

Авторы обеих схем независимо пришли к идее использовать рецептивные поля, случайно расположенные во входном пространстве и ограниченные гиперплоскостями, параллельными осям координат, а также включать в каждое поле только малое, случайно выбранное и фиксированное подмножество входных измерений S из A .

а)



б)

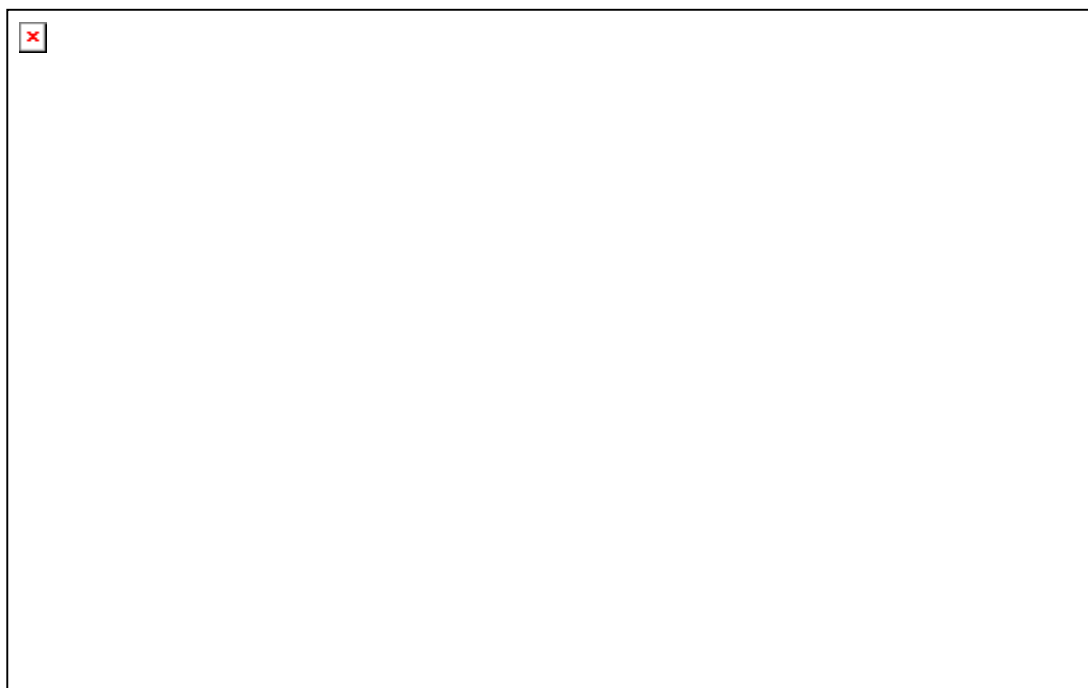


Рис. 6. График перекрытия $V(D^A, D^*)$ кода центральной точки $D = (0.5, 0.5)$ с кодами других точек двумерного пространства: а) схема Prager; б) схема RSC

Пониженная размерность рецептивных полей $S < A$ в сочетании с увеличением их размера и случайным расположением позволяет уменьшить рост числа полей по сравнению с табличной схемой и СМАС. Полученные в данной работе распределения реальной размерности полей позволяют аналитически оценить и оптимизировать затраты памяти (объем памяти, необходимый для хранения s полей, пересекающих интервал, вместо S) и вычислительную сложность кодирования (число проверок порогов).

Для маломерных рецептивных полей существуют потенциальные проблемы, связанные с тем, что они не позволяют адекватно оценить зависимости между большим, чем размерность рецептивного поля, числом входных переменных. Например, с помощью кода, полученного при $S = 1$, невозможно решить двумерную задачу исключающего ИЛИ и т.д. Теоретический анализ позволяет оценивать и управлять сложностью зависимостей между признаками, отображаемых кодом, и прогнозировать уровень сложности задач, которые можно решать с его помощью.

К счастью, реальная размерность практических задач обычно невелика. Даже если общее число информативных признаков A велико, в разных областях пространства информативны разные маломерные подпространства. Интересное развитие этих кодирующих схем состоит в выборе полей на не полностью случайной основе. Самый очевидный отбор заключается в устранении рецептивных полей, которые активны очень редко или очень часто. Более сложные методы отбора кодирующих элементов использовались в [7].

Различия между схемами RSC и Prager обусловлены различным числом порогов в одномерном поле. Как показано выше, однопороговые схемы непригодны для разреженного кодирования при использовании рецептивных полей низкой размерности, так как не могут давать коды с низкой плотностью (плотность кода не может быть меньше $0,5^A$). Двухпороговая схема RSC преодолевает этот недостаток, так как при заданном параметре S позволяет регулировать плотность кода управлением средней длиной промежутка между порогами G , а, следовательно, и средними объемами полей. Это позволяет регулировать плотность кода для малых размерностей входного пространства, что может быть важным для решения реальных задач.

Различия между схемами RSC и Prager наиболее выражены для рецептивных полей низкой размерности. Схема RSC из-за возможности наличия двух порогов имеет более локальный характер, но теряет эту локальность при увеличении среднего размера рецептивных полей G , которое требуется для поддержания нужной плотности кода при увеличении размерности полей S . Многие одномерные поля становятся в этом случае однопороговыми.

Различие между схемами проявляется также и в других свойствах кода. В однопороговой схеме Pragera и др. плотность результирующего кода однородна внутри всего входного пространства. Для двупороговой схемы RSC плотность кода уменьшается в углах входного пространства. Если это нежелательно, неоднородность можно уменьшить посредством масштабирования входного пространства так, чтобы оно находилось в центральной части пространства центров рецептивных полей.

В [17] предлагается модификация схемы RSC для решения проблемы неравномерной плотности кода. Для этого предлагается располагать центры рецептивных полей не в $[0,1]$, а на интервале $[0 - G_2, 1 + G_2]$. Длина полустороны рецептивного поля распределена при этом не на интервале $[0, G]$, а на $[G_1, G_2]$. Для $G_1 = 0, G_2 = G$ распределение длин сторон рецептивных полей становится таким же, как для RSC. При $G_1 = G_2 \equiv G > 0,5$ схема Д.В. Жоры [17] становится однопороговой. В гиперкубовой схеме кодирования В.В. Луковича [14] $S = A$ и длина стороны G фиксирована для всех измерений, однако центры полей расположены в интервале $[0,1]$.

Подобно другим схемам грубого кодирования, схемы RSC и Prager кодируют близкие точки входного пространства разреженными бинарными векторами, которые имеют большое пересечение единиц. Функция перекрытия ведет себя по-разному для "полуограниченных" схем с одним и "ограниченных" с двумя порогами: для двухпороговой схемы при одинаковых p и S она более крутая. Вероятно, это влияет на способности к обучению и обобщению систем, использующих такие коды, например, классификаторов. Это также относится и к разрешающей способности кодов. Проведенный сравнительный анализ характеристик схем грубого кодирования RSC и Prager позволит учитывать их особенности и целенаправленно соотносить их с практическими особенностями решаемых задач.

Авторы выражают благодарность за помощь при написании статьи Л.М. Касаткиной, А.С. Сычеву, Д.В. Жоре.

СПИСОК ЛИТЕРАТУРЫ

1. Arbib M. The Handbook of Brain Theory and Neural Networks. – Cambridge, MA: The MIT Press, 2003. – P. 1308.
2. Albus J.S. Data Storage in the Cerebellar Model Articulation Controller (CMAC) // Journal of Dynamic Systems, Measurement and Control, Trans. ASME. – 1975. – Т. 33, № 3. – P. 228 – 233.
3. Miller T.W., Glanz F.H., Kraft L.G. CMAC: An Associative Neural Network Alternative to Backpropagation // Proceedings of IEEE. – 1990. – Т. 78. – P. 1561 – 1567.
4. Prager R.W., Harrison T.D., Fallside F. Boltzman Machines for Speech Recognition // Computer Speech and Language. – 1986. – N 1. – P. 3 – 27.
5. Prager R.W., Fallside F. The Modified Kanerva Model for Automatic Speech Recognition // Computer Speech and Language. – 1989. – N 3. – P. 61 – 81.
6. Clarke T.J.W., Prager R.W., Fallside F. The Modified Kanerva Model: Theory and Results for Real-Time Word Recognition // IEEE. – 1991. – N 138. – P. 25 – 31.
7. Prager R. Networks Based on Kanerva's Sparse Distributed Memory: Results Showing their Strengths and Limitations, and a New Algorithm to Design the Location Matching Layer // IEEE International Conference on Neural Networks. – San Francisco: 1993. – P. 1040 – 1045.
8. Жора Д.В. Принципы построения программного агента по торговле ценными бумагами // Проблемы программирования. – 2004. – № 2. – С. 534 – 545.
9. Zhora D.V. Financial Forecasting using Random Subspace Classifier // Int. Joint Conf. Neural Networks. – 2004. – N 4. – P. 2735 – 2740.
10. Kussul N.N., Kussul M.E. Enhanced Algorithm of Nearest Neighbor Method and Its Application in Genetic Programming // International Mendel Conference "MENDEL'98". – Brno: Czech Republic, 1998. – P. 52 – 53.
11. Куссуль М.Э., Куссуль Н.Н. Ускоренный алгоритм реализации метода ближайшего соседа // Вісник Східноукраїнського національного університету. – 2000. – Т. 30, № 8. – С. 72 – 80.
12. Kussul E., Baidyk T., Lukovich V., Rachkovskij D. Adaptive High Performance Classifier Based on Random Threshold Neurons // Twelfth European Meeting on Cybernetics and Systems Research. – Austria, Vienna: World Scientific Publishing Co. Pte. Ltd, Singapore, 1994. – P. 1687 – 1694.
13. Куссуль Э.М., Байдык Т.Н., Лукович В.В., Рачковский Д.А. Нейронные классификаторы с распределенным кодированием входной информации // Нейрокомпьютер. – 1994. – № 3–4. – С. 13 – 24.
14. Лукович В.В., Рачковский Д.А. Исследование характеристик нейросетевого классификатора со случайными порогами // Нейросетевые технологии и компьютеры. – 1994. – С. 27 – 36.
15. Kussul E.M., Rachkovskij D.A., Wunsch D. The Random Subspace coarse coding scheme for real-valued vectors // International Joint Conference on Neural Networks. – Washington, DC: IEEE. – 1999. – P. 450 – 455.
16. Лукович В.В. Преобразование данных на входе нейросетевого классификатора со случайными порогами // Нейросетевые системы обработки информации. – 1996. – С. 12 – 19.
17. Жора Д.В. Анализ функционирования классификатора со случайными порогами // Кибернетика и системный анализ. – 2003. – № 3. – С. 72 – 91.
18. Вентцель Е.С. Теория вероятностей. – Москва: Наука, 1969. – 576 с.
19. Рачковский Д.А., Слипченко С.В., Куссуль Э.М., Байдык Т.Н. Свойства кодов числовых величин для схемы случайных подпространств RSC // Кибернетика и системный анализ. – 2005.