

Оценка тестов методики скрываемой информации методом ранжирования реакций

Летков Юрий Владимирович

Полиграфолог, разработчик СППРП «Сокол», г. Вологда, Россия

math-polygraph@yandex.ru

Аннотация. В статье представлены результаты исследования способов классификации полиграмм тестов методики скрываемой информации с использованием ранжирования реакций на стимулы ряда. Получены оценки параметров точности такого способа классификации, определены рамки его применимости. Показано, что использование линейного классификатора, который даёт возможность учитывать относительный вклад (веса) каждого физиологического канала, существенно повышает точность классификации.

Ключевые слова: алгоритм классификации полиграмм, ранжирование реакций, точность теста на знания виновного, точность поискового теста.

Введение

Ранжирование реакций для выделения значимого стимула/стимулов на полиграмме, записанной в ходе проведения полиграфного теста, является одним из способов оценки (классификации) полиграммы, которую используют полиграфологи в настоящее время. Известна, так называемая, Горизонтальная система об-счета, в основе которой лежит ранжирование реакций в каждом канале с последующим суммированием рангов, принадлежащих стимулу [1, С.110]. Автор теста знания виновного (ТЗВ) Д. Ликкен разработал систему оценки теста с использованием ранжирования реакций на стимулы тестового ряда [4, С.10]. В число признаков реакций, которые подлежали ранжированию Д.Ликкен включал: длину линии дыхания (ДЛД), амплитуду кожно-гальванической реакции (КГР), подъём артериального давления (АД), длина линии фотоплетизмограммы (ДЛФП), частота

сердечных сокращений (ЧСС). Наиболее сильной реакции присваивался ранг 1 [4, С.78]. Далее вычисленные суммарные ранги подставлялись в некоторую формулу, которая вычисляла то, что дословно звучит так: «...с её помощью вычисляется вероятность того, что полученный результат (а именно суммарный ранг «r», набранный всеми проверочными стимулами) является случайным (или неслучайным)» [4, С.44]. Невозможно представить математическую вероятностную модель, на основе которой вычислялась бы вероятность того, что результат ранжирования случайных величин – суммарный ранг является «неслучайным», так как такие вербальные конструкции находятся вне рамок теории вероятностей. А то, что он является случайной величиной следует из определения случайной величины, которая даётся в теории вероятностей. По причине того, что формула, которую использовал Д.Ликкен неизвестна [4, С.78], и видимо утрачена, её невозможно изучить и понять, что же фактически вычисляет эта формула и какая комбинаторная модель за ней стоит. Д. Ликкен провёл экспериментальную проверку точности своего теста в сочетании с определением значимого стимула при помощи ранжирования реакций и получил безошибочный результата на двадцати опрашиваемых лицах (ОЛ) [4, С.42], что говорит о высокой точности такого способа определения значимого стимула в ряду. Поэтому целью настоящей работы является изучение подобного способа оценки полиграммы, построение комбинаторной модели, определение его точности, границ применимости и способов модификации с целью повышения точности.

Построение комбинаторной ранговой модели для тестовых рядов

На рисунке 1 полиграмма ТЗВ с результатами измерений ДЛД, амплитуды КГР, минимального размаха графика ФПГ (полиграмма записана на полиграфе «Диана - 07», настройка фильтра ПГ – крайне правое положение).

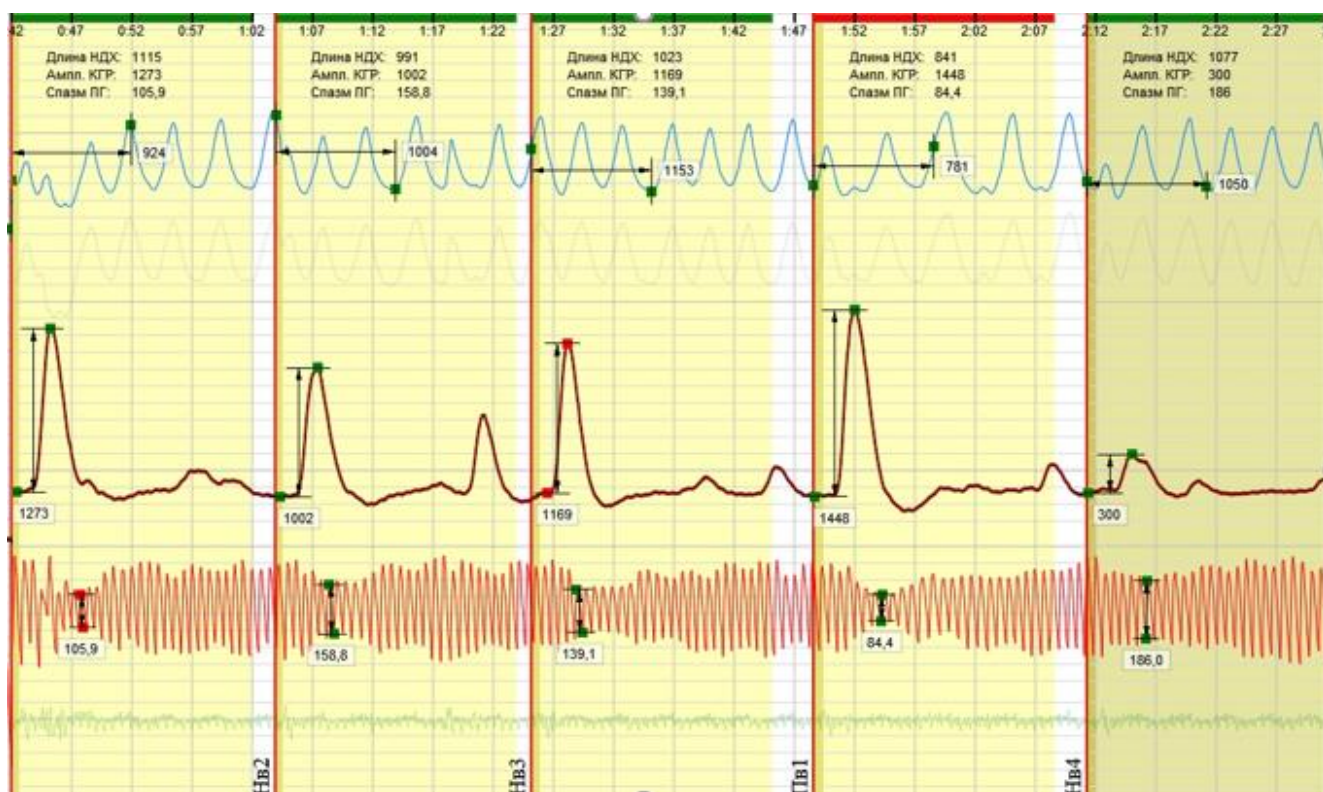


Рис. 1. Полиграмма ТЗВ с результатами измерений

В таблице 1 результаты ранжирования реакций с суммарным рангом.

Таб. 1. Ранги реакций на стимулы

	Дых.	КГР	ФПГ	Сумма
Нв1	2	2	2	6
Нв2	3	4	4	11
Нв3	5	3	3	11
Пв4	1	1	1	3
Нв5	4	5	5	14

Из таблицы можно видеть, что суммарный ранг может принимать различные значения, в том числе совпадающие. Каждое его значение может оказаться суммой различных комбинаций рангов. Рассчитаем количество возможных комбинаций суммарного ранга для одного стимула, для случая, когда значимый стимул отсутствует. Тест: $m = 5$ – проверочных вопросов и $n = 3$ – физиологических канала. Результаты расчёта в таблице 2.

Таб. 2 Количество комбинаций для суммарных рангов

Суммарный ранг	3	4	5	6	7	8	9	10	11	12	13	14	15
Кол-во возможных случаев	1	3	6	10	15	18	19	18	15	10	6	3	1

Минимальный суммарный ранг 3 может быть только в одном случае, если в каждом канале будет по одному рангу (1,1,1). Ранг 4 может сложиться в трёх случаях: (1, 1, 2) (1,2,1) (2,1,1). Ранг 5 – в 6 случаях: (1,1,3) (1,3,1) (3,1,1) (1,2,2) (2,1,2) (2,2,1). Ранг 6 – в 10 случаях: (1,1,4) (1,4,1) (4,1,1) (1,2,3) (1,3,2) (2,1,3) (2,3,1) (3,1,2) (3,1,2) (2,2,2). И так далее. Учитывая, что максимальное количество комбинаций рангов, которые могут реализоваться для каждого стимула равно $m^n = 5^3 = 125$ из таблицы 2 получим закон распределения вероятностей суммарного ранга – таблица 3. На рисунке 2 он представлен в графическом виде.

Таб. 3 Закон распределения вероятностей суммарного ранга

Сум. ранг	3	4	5	6	7	8	9	10	11	12	13	14	15
Вероятность	0.008	0.024	0.048	0.08	0.12	0.144	0.152	0.144	0.12	0.08	0.048	0.024	0.008

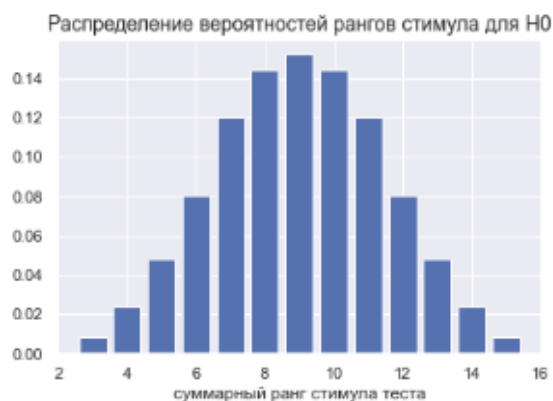


Рис. 2 Гистограмма распределения суммарного ранга

Имея закон распределения вероятностей, нетрудно посчитать вероятность того, что стимул будет иметь суммарный ранг, например, 4 **или меньше**. Она будет равна 0,032 (для ранга 5 такая вероятность равна 0,08). Это знание можно использовать для принятия решения в случае ТЗВ с одним проверочным вопросом. Для этого можно сформулировать статистическую гипотезу:

H0: Закон распределения вероятностей суммарного ранга для проверочного вопроса (Пв) описывается таблицей 2 (все стимулы теста одинаково значимы)

H1: Закон распределения вероятностей суммарного ранга для Пв отличается от закона, который даётся таблицей 1 (не все стимулы одинаково значимы)

Суммарный ранг 3 Пв1 (рисунок 1) означает, что реакция опрашиваемого лица на него сильнее, чем на остальные вопросы, и это минимально возможный для данной полиграммы суммарный ранг. Вероятность же получить такой или меньший ранг, согласно закону распределения вероятностей суммарного ранга, равна 0,008. Это значение меньше, чем 0,05. Таким образом, на уровне 0,05 мы можем отвергнуть нулевую гипотезу и принять альтернативную – проверочный вопрос является значимым. Тем самым полиграмма относится к классу полиграмм, записанных в ходе теста ТЗВ, в которых проверочный вопрос (Пв) был значим, и испытуемый отвечал на него лживо. Второй класс – Пв не значим, и ответ на него правдивый. Такой способ классификации имеет существенные недостатки. Он подходит для рядов. Для них может быть подобран уровень значимости с тем, чтобы такая классификация обладала приемлемыми чувствительностью и специфичностью, но она не подходит для тестов вопросов сравнения (ТВС). И даже в случае рядов такой способ будет давать сильно отличающуюся точность для ТЗВ и для поисковых тестов. Дело в том, что в случае ТЗВ у нас имеется один Пв вопрос, относительно которого мы проверяем верность нулевой гипотезы. А в случае поискового теста у нас несколько Пв, как правило пять - шесть. То есть, для одного теста мы должны проверить сразу пять - шесть нулевых гипотез, что, конечно, увеличивает вероятность ошибки первого рода. Такой способ для поисковых тестов будет увеличивать обвинительный уклон в сравнении с ТЗВ. Для проверки этого предположения проводился численный эксперимент. Генерировались распределения рангов

физиологических каналов для поискового теста с 6 проверочными вопросами, с тремя физиологическими каналами, для двух повторов тестов. Было сгенерировано 100 случаев. Из них в 22 случаях была ошибочно отвергнута нулевая гипотеза. В реальной ситуации было бы принято 22 ошибочных решения из 100. Таким образом, специфичность оценивается в 78%, что является довольно низкой величиной. Если же на этих данных выделить один Пв, таким образом смоделировав ТЗВ, то можно оценить специфичность этого теста. Был выделен вопрос № 4. Из 100 случаев в отношении его было принято 5 раз ошибочное решение о его значимости, при ожидании в 3,59 (вычислено из закона распределения вероятностей для данного теста). Этот результат соответствует закону распределения вероятностей суммарного ранга для отсутствия значимых стимулов. Для поискового теста из шести стимулов, в соответствии с этим законом, можно ожидать 21,54 ошибочных результатов. Полученный результат в 22 ошибки полностью соответствует рассчитанному количеству (Z тест, $p = 0,91$). Такая маленькая специфичность наблюдается при достаточно высоком пороговом уровне достигаемой значимости α (0,05). Его уменьшение увеличит специфичность, но уменьшит чувствительность, что вряд ли будет способствовать увеличению общей точности. Путь увеличения точности для такого способа классификации полиграмм методики скрываемой информации (МСИ) – это увеличение числа физиологических каналов, увеличение числа однородных стимулов в тесте, увеличение количества повторов теста. Но эти приёмы имеют естественный предел, который задаётся реальными условиями тестирования.

Для проверки соответствия вышеописанной ранговой модели ТЗВ реальным тестам использовалась выборка ТЗВ, проведённых по уголовным делам, которые были раскрыты, и по которым имеются судебные решения. Тем самым, стало известно лгали или отвечали правдиво опрашиваемые лица на Пв теста. Всего использовалось 120 ТЗВ, состоящих из пяти стимулов (шестой стимул, стоящий первым, не учитывался). Все тесты

проводились в два повтора. Используемые физиологические каналы: Дыхание, КГР, ФПГ. Полиграммы были записаны на полиграфе «Диана 04». В 51 одном тесте опрашиваемые отвечали правдиво. Для этих тестов был рассчитан **закон распределения вероятностей суммарного ранга для незначимых стимулов в тестовом ряду**. И исходя из этого закона было рассчитано ожидаемое количество ошибочных результатов для различных значений альфа-уровня – порогового уровня статистической значимости. На рисунке 3 распределение вероятностей ранга для случая $m = 5$, $n = 6$. В таблице 4 приведены числовые значения, полученные на выборке правдивых.



Рис. 3 Гистограмма распределения суммарного ранга ($m = 5$, $n = 6$)

Таб. 4 Результаты ранжирования на выборке правдивых

α	0,05 ($p = 0,0292$ 11 ранг)	0,1 ($p = 0,0991$ 13 ранг)	0,25 ($p = 0,2397$ 15 ранг)
Пв незначим	50	48	40
Пв значим	1	3	11
Специфичность	0,98	0,94	0,78
Ожидаемая ошибка	1,4892	5,0541	12,2247
p-значение (Z-тест для доли)		0,3357	0,6879

Результаты теста для доли признака в генеральной совокупности оставляют верным предположение о справедливости закона распределения вероятностей для реальных тестов с незначимыми Пв в силе. В 69 тестах опрашиваемые отвечали на Пв лживо. В таблице 5 результаты для этой выборки.

Таб. 5 Результаты ранжирования на выборке лживых

α	0,05 (p = 0,0292 11 ранг)	0,1 (p = 0,0991 13 ранг)	0,25 (p = 0,2397 15 ранг)
Пв незначим	40	25	11
Пв значим	29	44	50
Чувствительность	0,42	0,64	0,72

Как можно видеть такой алгоритм классификации для ТЗВ с пятью стимулами и двумя повторами обладает высокой специфичностью, но низкой чувствительностью. Сбалансированный вариант наблюдается при альфа-уровне равном 0,25. Увеличение количества стимулов в тесте до шести и увеличение повторов теста до трёх должны увеличить чувствительность и общую точность ТЗВ. Будет ли увеличение существенным, заранее сказать нельзя. В настоящее время этот алгоритм реализован в Системе поддержки решения полиграфолога (СППРП) «Сокол» [6] под названием CombiCalc v1 (CC v1).

При рассмотренном выше способе классификации полиграмм никак не учитывается относительный вклад, вес физиологических каналов. Но вероятнее всего, алгоритм классификации должен учитывать то, что физиологические каналы не обладают равной информативностью. Для построения такого алгоритма учтём факт, к которому привычны полиграфологи. Чем сильнее реакция в физиологическом канале, тем больше число, которое соответствует ей. Будем описывать силу реакции числом равным единице минус достигаемый уровень значимости в ранговой модели ряда в предположении нулевой гипотезы.

$$f = 1 - p \quad (1)$$

На рисунке 4 рассчитанная «сила реакции» на стимулы, рассматриваемого теста (рисунок 1).

Стимул	суммарный ранг	1 - p-value
Нв1	6	0.84
Нв2	11	0.16
Нв3	11	0.16
Пв1	3	0.992
Нв4	14	0.008

Рис. 4 Сила реакции в ранговой модели

Таким образом, для самой сильной реакции будем иметь самое большое число. На рисунке 5 привычная для полиграфологов гистограмма, отображающая «силу реакции» на стимулы теста МВСИ

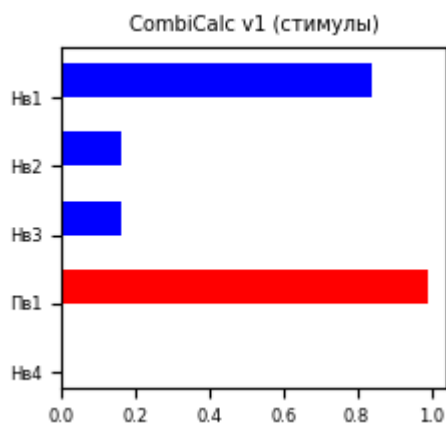


Рис.5 Гистограмма силы реакции в ранговой модели

Следующий шаг — это вычисление f отдельно по каждому физиологическому каналу. После этого результаты алгоритма можно представить в виде точки в многомерном пространстве, размерность которого равна числу используемых физиологических каналов. На рисунке 6 изображены результаты алгоритма, полученные на выборке ТЗВ, которые были проведены по раскрытым в последствии уголовным делам. Красные точки — тесты со лживым ответом, зелёные — с правдивым.

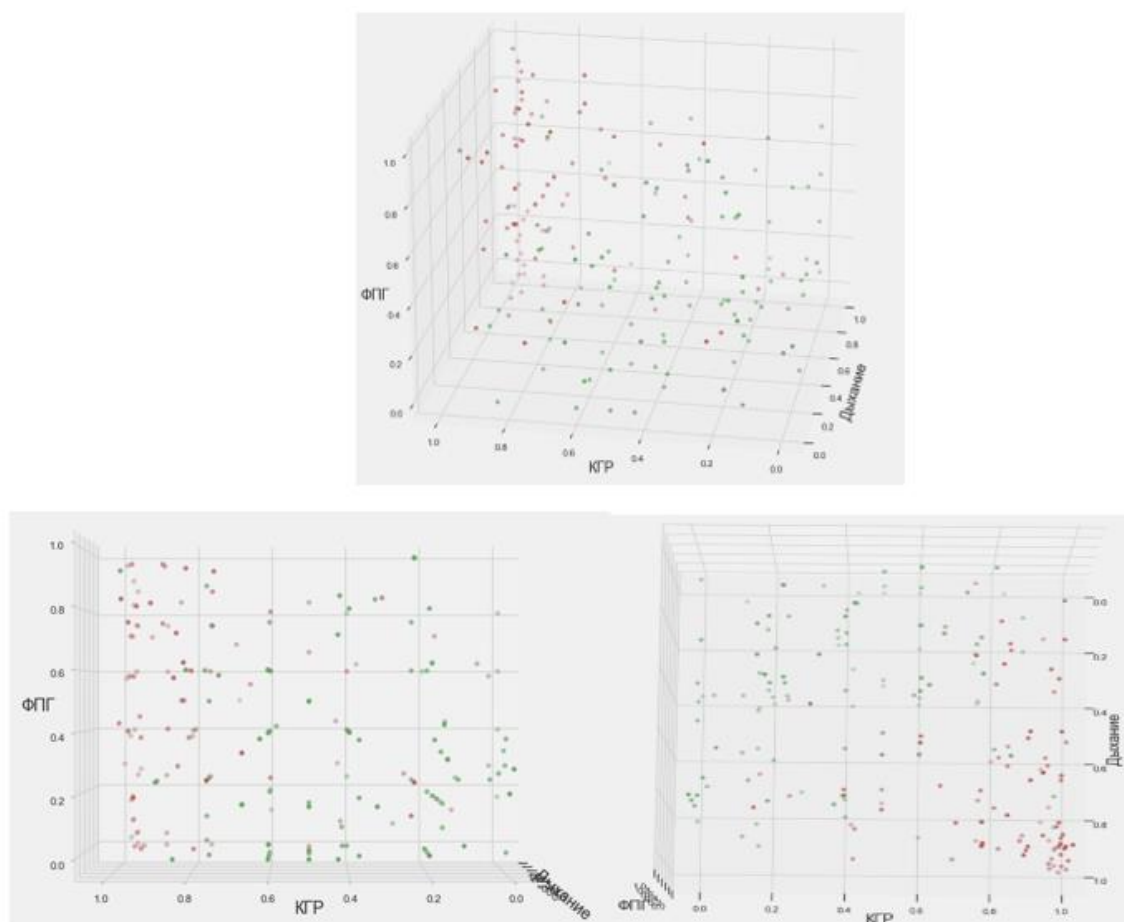


Рис. 6 Трёхмерная диаграмма рассеяния результата алгоритма SS v1 на выборке ТЗВ

Из диаграммы рассеивания можно сделать вывод о том, что классификация возможна с использованием линейного классификатора. Так как в нашем распоряжении не было полиграмм ТЗВ с известным решением, которые были записаны с использованием канала «Манжета», для построения разделяющей плоскости использовались полиграммы САТ - ов и полиграммы поисковых тестов с карточками, в которых было заранее известно, что опрашиваемые отвечали правдиво на все стимулы ряда. Описание технологии записи полиграмм на тестах с карточками можно найти в [2, С.9]. В этих тестах механически отбирались стимулы так, чтобы они находились ближе к середине ряда, так как проверочный стимул ставится в ТЗВ. Отобранные стимулы обозначались как проверочные, остальные обозначались, как стимулы для сравнения. На этих данных были построены 4-мерные и 3-мерные разделяющие плоскости для случаев использования 4-х и 3-х физиологических каналов. Условное изображение разделяющей плоскости можно

увидеть на рисунке 7. Разделяющие плоскости строятся так, чтобы минимизировать ошибку при классификации. Иначе, строятся так, чтобы по одну сторону от плоскости находилось как можно больше точек одного класса и как можно меньше точек другого класса.

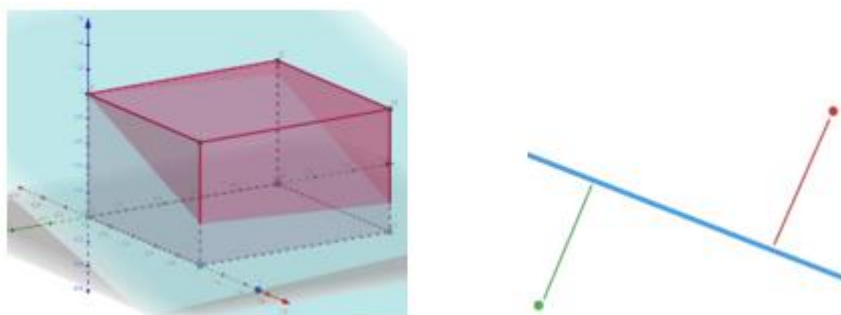


Рис. 7 Плоскость, разделяющая точки, полученные на двух классах

Чем дальше точка «силы реакции» располагается от разделяющей плоскости, тем с большей вероятностью она относится к одному из двух классов. Так как в нашем случае мы имеем ограниченное пространство, использовать логистическую регрессию с целью получения такой оценки не представляется возможным. Для этого использовалась функция вида

$$F(d) = \frac{1}{1 + e^{\frac{k \cdot d}{(a-d) \cdot (b+d)}}} \quad (2)$$

где d – расстояние до разделяющей плоскости, a и b – максимально возможные в разделённых полупространствах расстояния до разделяющей плоскости, k – множитель

На рисунке 8 график описанной функции в области определения $[-1, 1]$.

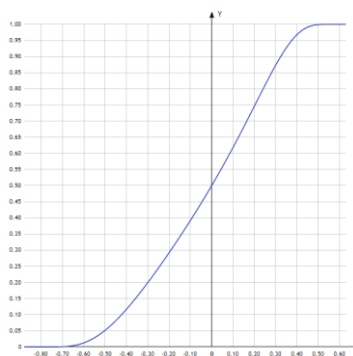


Рис. 8 График функции $F(d)$

На поисковых тестах с карточками [2, С.9], преобразованных в ТЗВ (5 стимулов, два повтора), так, чтобы стимулы, выбираемые как проверочные располагались ближе к центру ряда, были получены выборки значений, вычисленных по формуле (2) и произведена ядерная оценка плотности вероятности $F(d)$, которую можно увидеть на рисунке 9

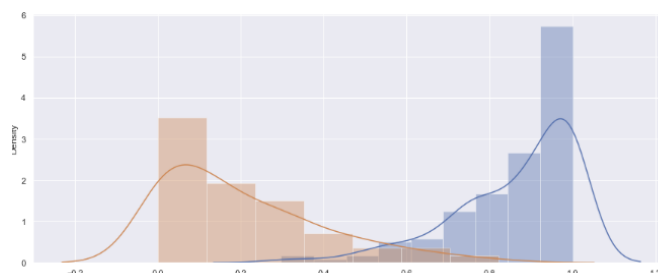


Рис. 9 Ядерная оценка плотности вероятности

Таким образом, получена возможность производить оценку вероятности принадлежности полиграммы ТЗВ к определённому классу с использованием теоремы Байеса. На рисунке 10 приведены гистограммы оценок вероятностей при априорной вероятности равной 0,5 (4 канала, 5 стимулов, два повтора, 123/165).

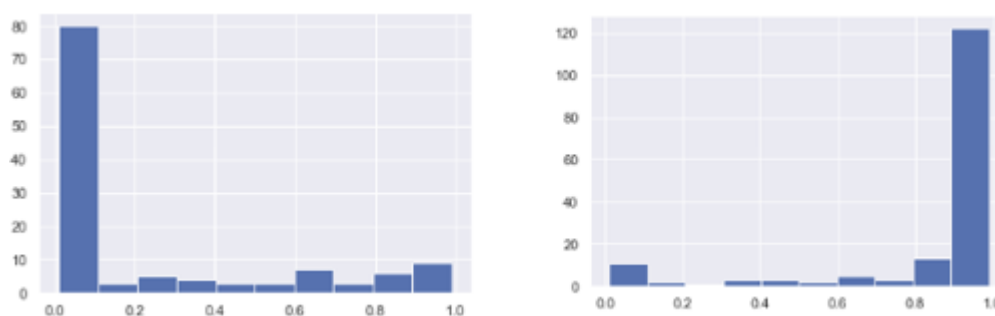


Рис. 10 Гистограммы оценок вероятностей принадлежности к классам

Исходя из этих гистограмм, получены правила принятия решений для классификации полиграмм. Решение о значимости проверочного стимула в ТЗВ принимается при оценке вероятности принадлежности к классу «лживых» от 0,9.

Решение о незначимости проверочного стимула – при оценке принадлежности к классу «лживых» менее 0,5. При оценке от 0,5 до 0,9 – отказ от классификации – неопределённый результат. Аналогичные решающие правила приняты и для случаев использования трёх физиологических каналов. Этот алгоритм классификации получил название CombiCalc v2 (СС v2) и реализован в СППРП «Сокол» [6]

Тесты на знания виновного

Правила классификации СС v2 были проверены на имеющихся выборках САТ (пять стимулов, два повтора, 65 тестов), ТЗВ по уголовным делам с известным решением (пять стимулов, два повтора, 64/49) и на полиграммах тестов с карточками. В таблице 6 результаты измерений на тестах с карточками (123/165)

Таб. 6 Результаты классификации полиграмм тестов с карточками

Тесты с карточками (с правдивым ответом на ПВ)							
Результаты Каналы	ПВ не- значим	95% дове- рительный интервал	Нет реше- ния	Пв значим (ошибка)	95% дове- рительный интервал	Специфич- ность	95% дове- рительный интервал
4	95	(86 ; 104)	19 (15%)	9	(3 ; 15)	0,91	(0,85 ; 0,97)
3 (Манжета)	93	(84 ; 102)	18 (15%)	12	(6 ; 18)	0,89	(0,82 ; 0,94)
3 (ФПГ)	91	(81 ; 101)	18 (15%)	14	(7 ; 21)	0,87	(0,79 ; 0,94)
Тесты с карточками (со лживым ответом на ПВ)							
Результаты Каналы	ПВ зна- чим	95% дове- рительный интервал	Нет реше- ния	Пв незна- чим (ошибка)	95% дове- рительный интервал	Чувстви- тельность	95% дове- рительный интервал
4	122	(111 ; 133)	23 (14%)	20	(12 ; 28)	0,86	(0,80 ; 0,92)
3 (Манжета)	125	(114 ; 136)	22 (13%)	18	(10 ; 26)	0,87	(0,81 ; 0,93)
3 (ФПГ)	132	(122 ; 142)	18 (11%)	15	(8 ; 22)	0,90	(0,85 ; 0,95)

Можно видеть, что с уменьшением числа каналов имеется тенденция в сторону обвинения, которая достигает максимального значения в третьем случае (ФПГ). Но тем не менее, статистически значимой разницы между различными вариантами не обнаруживается (χ^2 , $p > 0,1$). Возможно, для её обнаружения необходима выборка большего размера.

Оценку специфичности можно сделать, если заметить, что на гистограмме (рисунок 8) хвост имеет распределение близкое к равномерному. Исходя из этого, результат можно пересчитать и оценить как 99/19/5 (χ^2 , $p=0,52$). Где 99 – правильное решение, 5 – ошибочное. Тогда оценка чувствительности для ТЗВ с пятью стимулами, одним ПВ, с двумя повторами будет равна 0,95, что попадает в доверительный интервал. Оценённая таким же образом чувствительность (122/19/24, χ^2 $p=0,54$) будет равна 0,84. Неопределённый результат – 13%.

Полученные оценки близки к тем, которые сделали японские исследователи для четырёх признаков реакций: проводимость кожи, дыхательная активность, ЧСС и нормализованный объём пульсовой волны [3]. Для повторов теста от 3 до 5 раз (количество стимулов в тесте не указано, но все примеры тестов в статье включают в себя 5 стимулов) получены следующие значения: чувствительность – 86%, специфичность – 95%, нет решения – 20%. Алгоритм классификации не указан, но в статье [7] отмечено, что тот же тестовый материал был проанализирован с использованием «модифицированной шкалы Ликкена», что дало следующие результаты: чувствительность – 83%, специфичность – 91%, нет решения – 25%.

В таблице 7 результаты испытания алгоритма на 65 САТ – ах. Статистически значимой разницы с результатами на тестах с карточками не обнаружилось.

Таб. 7 Результаты СС v2 на САТ - ах

Результаты Каналы	ПВ значим	Нет решения	Пв не значим (ошибка)	Чувстви- тельность	Р – значе- ние (χ^2)
4	46	11 (17%)	8	0,85	0,841
3 (Манжета)	46	12 (18%)	7	0,87	0,616
3 (ФПГ)	52	6 (9%)	7	0,88	0,877

В таблице 8 указаны результаты испытания алгоритма на 113 (49/64) ТЗВ с известным решением, проведенные по реальным уголовным делам, которые в дальнейшем были раскрыты. ТЗВ с пятью стимулами (шестой, первый по счёту исключался), в два повтора, три физиологических канала (ФПГ). Статистически значимой разницы с результатами на тестах с карточками не обнаружилось.

Таб. 8 Результаты СС v2 на ТЗВ по уголовным делам

ТЗВ с правдивым ответом на ПВ					
Результаты Каналы	ПВ незна- чим	Нет реше- ния	Пв значим (ошибка)	Специфич- ность	Р – значение (χ^2)
3 (ФПГ)	35	10 (20%)	4	0,90	0,578
ТЗВ со лживым ответом на ПВ					
Результаты Каналы	ПВ значим	Нет реше- ния	Пв незна- чим (ошибка)	Чувстви- тельность	Р – значение (χ^2)
3 (ФПГ)	54	1 (2%)	9	0,86	0,567

Поисковые тесты

Выше было показано, что, производящееся в алгоритме СС v1 ранжирова-
ние реакций на стимулы «силы реакции» без учета веса физиологического канала
приводит к увеличению вероятности ошибки в поисковом тесте в 5-6 раз (на коли-
чество используемых стимулов в ряду) в случае отсутствия в нём стимулов, в от-
ношении которых опрашиваемое лицо даёт лживый ответ. Увеличение ошибки
обусловлено не спецификой алгоритма, а отражает общую закономерность поис-
ковых тестов, которая не зависит от способа принятия решения и используемого
алгоритма. Это показано в исследовании Поповичева С.В. [5]. Учитывая специ-
фику алгоритма СС v2 и крайне низкую специфичность у поисковых тестов, алго-
ритм в существующем виде можно использовать только с лицами, в отношении
которых с высокой степенью достоверности известно то, что они обладают и
скрывают необходимую информацию. Для определения значимого стимула в по-
исковом ряду производится ранжирование стимулов ряда по $F(d)$. В СППРП «Со-
кол» значение $F(d)$ обозначено как «показатель», что можно увидеть на рисунке
11

стимул	результат	показатель
R2	1	0.6734
R5	1	0.63
R6	1	0.5181
R4	-1	0.2881
R3	-1	0.0

Рис. 11 Значения $F(d)$ вычисляемые СС v2 для стимулов теста

Стимул с наибольшим показателем признаётся значимым, то есть определяется, как скрываемое опрашиваемым лицом обстоятельство, факт которые интересуют инициатора тестирования. Проверка работы алгоритма CC v1 и CC v2 с поисковыми тестами проверялась на данных, которые любезно предоставил Поповичев С.В., и которые он использовал в своём исследовании [5]. Данные представляют собой значения параметров Кирчера снятые с 36 полиграмм, записанных в ходе проведения теста «Линейка». Тест включал пять проверочных стимулов и повторялся три раза. Так же Поповичев С.В. предоставил результаты американских алгоритмов на этих данных. В таблице 9 сравниваются результаты алгоритмов.

Таб. 9 Результаты алгоритмов на полиграммах теста «Линейка»

	CC v2	CC v1	Out Lier	Polyscore	OSS(1) ¹	OSS(2) ¹	OSS(3) ¹	OSS(4) ¹	OSS(5) ¹
Правильно	31	24	31	35	24	26	31	30	30
Ошибка	5	12	5	1	12	10	5	6	6
%	86,1	66,7	86,1	97,2	66,7	72,2	86,1	83,3	83,3

В 4 из 5 случаев, когда CC v2 ошибался, действительно значимый стимул находился на втором месте, значение его показателя мало отличалось от значения показателя, который помещался на первое место в ряду.

Статистический тест не обнаружил разницу между результатами CC v1 и CC v2 (χ^2 , $p = 0,052$). Так как p – значение находится на грани принятия решения, для проверки была использована выборка большего объёма – 164 поисковых теста с карточками, пять стимулов, два повтора. Результат CC v2 на этой выборке 134/30, а CC v1 – 100/64 (χ^2 , $p = 3,3 \cdot 10^{-5}$). CC v1 – это ранжирование показателей физиологических признаков реакций на стимулы с последующим их суммированием. Решение о значимости принимается в отношении стимула, который получил минимальный суммарный ранг. CC v2 – производится ранжирование, но

¹ Расшифровка американских алгоритмов, которые используются в программе OSS3:

1. OSS(1): CIT/SPOT – Krapohl and Velez;
2. OSS(2): Rank Order Analysis – Miritello;
3. OSS(3): Ipsative-Z – Nelson;
4. OSS(4): Relative Response Magnitudes – Raskin;
5. OSS(5): Probability Analysis – Kircher and Raskin.

решение принимается с учётом весов физиологических признаков реакций. Таким образом, можно считать установленным, что для более точного выявления значимого стимула простого ранжирования показателей недостаточно. Необходимо учитывать их веса.

В исследовании Поповичева С.В. [5] получено среднее значение процента правильного определения значимого стимула экспертами равное 88%. Этот результат был получен на 43 полиграммах, 36 из которых участвовали в испытании алгоритмов, результаты которых приведены в таблице 9. Статистически значимая разница между результатами СС v2 и средним результатом экспертов отсутствует (χ^2 , $p = 0,78$). Так же не была обнаружена статистически значимая разница между результатами СС v2 на выборке из 164 тестов с карточками и его же результатами на данных Поповичева С.В. - 134/30 и 31/5 (χ^2 , $p = 0,53$). Так как вторая выборка была получена без использования канала ФПГ, в отличие от первой, но тесты записывались в три повтора, то результаты СС v2 на ней были пересчитаны, брались только первые два повтора. Результат – 27/9. При сравнении с выборкой из 164 тестов статистически значимой разницы не обнаружено (χ^2 , $p = 0,36$). Можно предположить, что результат алгоритма СС v2 незначительно зависит от ФПГ. Это было проверено на 100 полиграммах тестов с карточками (исключалось ФПГ). Статистически значимой разницы так же не обнаружено (χ^2 , $p = 0,46$). Аналогичный результат на этих же данных получен для алгоритма СС v1 (χ^2 , $p = 0,74$).

Таким образом СС v2 даёт значение чувствительности в поисковых тестах для двух повторов 0,82 и 95% доверительный интервал - (0,76 ; 0,87).

Результаты

В ходе исследования была получена формула, которая позволяет, не затрачивая больших вычислительных ресурсов, получать закон распределения вероятностей суммарного ранга для тестов МСИ, которые проводятся с ОЛ, для

которого в тестовом ряду отсутствует стимул, который он узнает, но скрывает это или на который отвечает лживо. Знание этого закона позволяет сформулировать прямую и противоположную статистическую гипотезу и, таким образом, выделить значимый вопрос и производить классификацию полиграмм. Установлено, что при использовании общепринятых параметров (ДЛД, амплитуды КГР, минимального размаха ФПГ, подъёма АД) и типичного для практики полиграфологов количества стимулов в ряду, количества повторов теста, такой способ классификации обладает невысокой точностью. В большей мере это касается чувствительности МСИ. Вместе с тем показано, что использование линейного классификатора (учёта информативности каналов) повышает точность тестов МСИ с использованием ранжирования, доводя ее до уровня не худшего, если сравнивать с другими алгоритмами и способами классификации полиграмм тестов МСИ.

Автор выражает благодарность коллегам: В. Казакову, А. Калафати, канд. психол. наук С. Поповичеву – за предоставленные ими данные и результаты исследований. Отдельное спасибо А. Калафати за ценные замечания к рукописи.

Список литературы

1. Варламов В.А., Коровин В.В., Варламов Г.В. Тесты полиграфных проверок. – Ставрополь: ГУВД Ставропольского края, 2001. — 166 с.
2. Летков Ю. В., Калафати А. Ю. 2021. Алгоритм классификации полиграмм Out_LIEr. PREPRINTS.RU. <https://doi.org/10.24108/preprints-3112294>
3. Осуги А. Проверки на полиграфе, как метод обнаружения скрываемой информации: методика и практика использования полиграфа. //Детекция лжи. – 2020. № 1. – С.98-115
4. Пеленицын А. Б., Сошников А. П. «Современные технологии применения полиграфа: подробное руководство для полиграфологов-практиков: в 4 ч.»

- /АНО ДПО "Центр прикладной психофизиологии". - Москва: Центр прикладной психофизиологии, 2015. Ч. 4. - 2015.
5. Поповичев С.В. О точности поисковых тестов. //Детекция лжи. – 2021. № 4. – С.71-78
 6. СППРП «Сокол»: [сайт] /2020. – URL: <http://www.skl-ol.ru/> (Дата обращения 11.02.2023)
 7. Ogawa, T., Matsuda, I., & Tsuneoka, M. (2013). Accuracy of Concealed Information Test as a Memory Detection Technique: A Laboratory Study. Japanese Journal of Forensic Science and Technology, 18, – С.35-44