

## Применение методов многомерного анализа данных при обработке полиграмм для изучения биофизических характеристик

*Н.Н. Минакова, Е.В. Божич*

Алтайский государственный университет (Барнаул, Россия)

## Multivariate Data Analysis of Polygraph Charts for Assessment of Biophysical Characteristics

*N.N. Minakova, E.V. Bozhich*

Altai State University (Barnaul, Russia)

Предложен инструментарий, дающий возможность изучать биофизические характеристики человека по результатам обработки полиграмм. Представлены результаты применения двух методов многомерного анализа данных: метода главных компонент и метода проекции на латентные структуры. Данные о биофизическом состоянии человека снимались датчиками в процессе полиграфного тестирования. Анализировались кожногальваническая реакция, фотоплетизмограмма, артериальное давление и другие биофизические характеристики.

Показано, что для выявления взаимосвязи между биофизическими характеристиками человека целесообразно применять график нагрузок. Применение метода главных компонент дает возможность оценить незначительные изменения биофизических характеристик человека, учесть ситуационно значимые символы, способные свидетельствовать о том, что в организме идет физико-химический процесс, который может быть значимым.

Метод проекции на латентные структуры целесообразно использовать для разработки регрессионной модели, которую можно постоянно применять при полиграфических обследованиях. Сделан вывод о том, что с помощью выбранных методов многомерного анализа данных можно исследовать биофизические характеристики человека по данным полиграфа, выявлять возможные сбои в процессах, их формирующих. Полученные результаты позволяют решать практические задачи, например, выявлять инсайдера при решении вопросов защиты информации.

**Ключевые слова:** биофизические характеристики, полиграммы, электропроводность, сигналы, фотоплетизмограмма, многомерный анализ данных, метод главных компонент.

In this paper, we propose tools for studying biophysical characteristics of a person by processing polygraph charts. We present the results produced by two methods of multivariate data analysis: the principal component analysis (PCA) and projection on latent structures (PLS). Data of biophysical state of a person during the polygraph testing - galvanic skin response (GSR), photoplethysmogram, blood pressure, and other biophysical characteristics – are processed and analyzed.

It is shown that load diagrams are the preferable method to identify the relationship between the biophysical characteristics of a person. The principal component analysis provides an opportunity to assess minor changes in biophysical characteristics of a person and to take into account situational significant identifiers. They can attest to the fact that a possibly significant physical-chemical process runs in a human body.

It is demonstrated that the projection on latent structures should be used to develop regression models that can constantly be used in polygraph testing.

It is concluded that proposed methods of multivariate data analysis can help in an investigation of biophysical characteristics of a person using polygraph data. Also, the proposed methods are able to detect possible malfunctions in the responsible physical-chemical processes. The obtained results enable solving practical problems of information security matters, for example, the identification of an insider.

**Key words:** biophysical characteristics, polygraph charts, conductivity, signals, photoplethysmogram, multivariate data analysis, principal component analysis.

В современных условиях изучение биофизических характеристик, связанных с жизнедеятельностью человека, используется для решения многих прикладных задач, например, задач защиты информации [1–3]. Многие фирмы при приеме новых сотрудников нередко прибегают к полиграфным проверкам [4]. Полиграф все чаще применяется при решении задач обеспечения информационной безопасности, так как технические средства защиты, в том числе и от несанкционированного доступа, не способны обеспечить защиту от инсайдеров (внутренних нарушителей безопасности информации) [4, 5].

Поэтому актуальна разработка математических методов для изучения биофизических характеристик, в том числе повышающих точность оценки результатов полиграфного тестирования. Такие методы математической (метрической) оценки разрабатываются. Используются, например, регрессионный анализ, методы статистической оценки [5]. Оценивалась возможность использования физико-математических методов, применяемых для изучения неоднородных структур [6, 7].

Была поставлена задача поиска математических методов, позволяющих получить дополнительную информацию по результатам исследования биофизических характеристик, полученных по полиграммам, для изучения протекающих процессов, анализа данных о психофизическом состоянии человека.

Распознавание биофизических характеристик человека по полиграммам, снятым датчиками в процессе полиграфного тестирования, выполняется на основе анализа изменений физико-химических процессов в организме человека при воздействии раздражающего фактора. Полиграмма формируется из нескольких сигналов, снятых с различных датчиков.

В работе ставится задача обработки полиграмм с применением методов многомерного анализа данных. Анализ многомерных данных (АМД) позволяет выделять в больших массивах данных скрытые переменные и анализировать связи, существующие в изучаемой системе [8].

Физическая основа применения полиграфа известна [4]. Биологическая система имеет собственные временные электрические и электромагнитные ритмы. При действии раздражающего фактора они изменяются. Например, дыхательные сигналы в состоянии тревоги могут убыстряться. При возбуждении обычно повышается величина кровного давления, частота сердечных сокращений. Фотоплетизмограммы регистрируют кровенаполнение в периферических сосудах [4]. Волны отражают движение объема крови в измеряемой точке во время систолы и диастолы [4, 9, 10]. На полиграммах фиксируется физическая и тоническая составляющие кожно-гальванической реакции. Ответ на короткий (ситуационный) раздражитель — медленное изменение кожного сопротивления, свя-

занного с отражением глубинных процессов реагирования в центральной нервной системе [4].

Данные для исследования были получены с помощью полиграфа КРИС со следующими датчиками: двигательной активности или тремора (TRM); фотоплетизмограммы (FPG); верхнего дыхания (VDH); нижнего дыхания (NDH); кожно-гальванической реакции (KGR); артериального давления (AD). Применялся метод контрольных вопросов.

Методами многомерного анализа изучаются данные, которые являются переменными состояниями какого-либо контролируемого процесса. В данном случае это процесс изменения физиологического состояния человека в ответ на заданный ему стимул.

Метод главных компонент (МГК или английская версия PCA) использует матрицу данных, которая состоит из  $n$  образцов и  $p$  переменных [8]. Рассматривались график счетов и график нагрузок. График счетов отражает взаимосвязь между объектами, график нагрузок показывает взаимосвязь между переменными.

Изучались данные тестирования на полиграфе при трудоустройстве, которые обезличивались. Задаваемые вопросы делились на нейтральные (Н), контрольные (К) и значимые (З) [4]. Эксперименты показали, что для испытуемого значимыми являлись вопросы №3, 5, 11 и 13 (или 31, 33, 38 и Н3). Было важно оценить возможность применения метода главных компонент (МГК) для анализа полиграмм. Набор данных для дополнительной обработки с использованием многомерного анализа данных преобразовывался в матричный вид. В качестве осей выбраны первая и вторая главные компоненты. На рисунке 1 представлен график счетов.

По левую сторону от оси ординат на графике счетов (см. рис. 1) оказались объекты, которые являются наиболее значимыми для испытуемого. Выделяется V5, который согласно данным полиграфа отнесен к значимым. График нагрузок показывает, какая переменная оказывает наибольшее влияние. Для конкретного испытуемого это переменная FPG (канал фотоплетизмограммы) — ситуационное изменение кровяного потока в сосудах.

Представленные результаты показывают, что применение многомерного анализа данных позволяет получить дополнительную информацию о биофизических изменениях в организме.

По графику счетов выявлены пики биофизических характеристик V1 и V8. При детальном рассмотрении было отмечено, что при ответе на первый вопрос у испытуемого возникли значительные изменения в канале кожно-гальванической реакции. Это можно связать с возможным волнением из-за предстоящего тестирования. Анализ характеристики V8 показывает, что у испытуемого наблюдались значительные изменения в канале ФПГ.

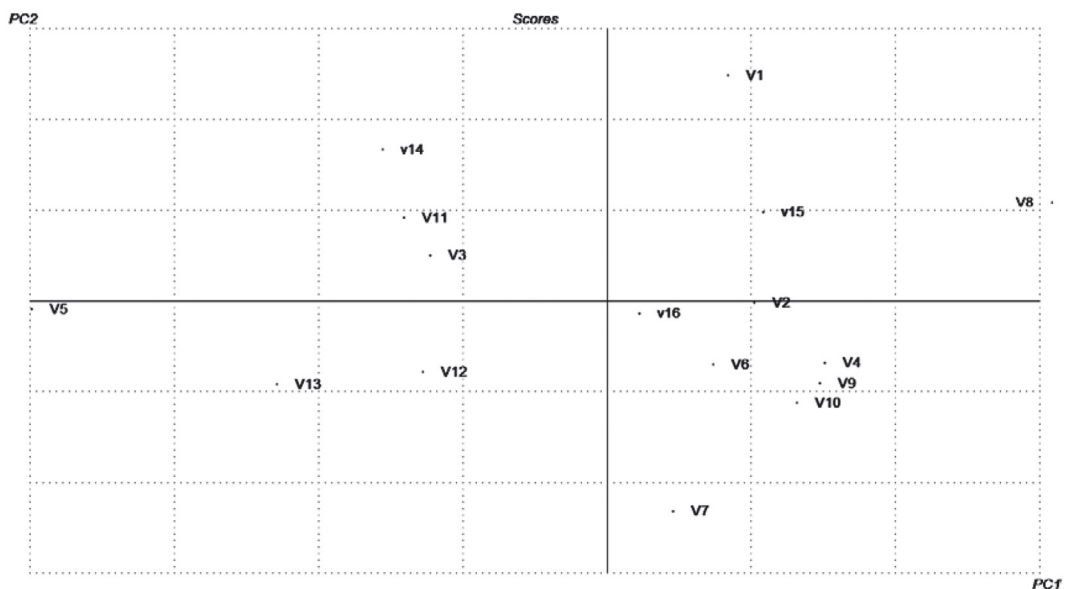


Рис. 1. Результат работы метода главных компонент. График счетов

Результаты, представленные на графике счетов, подтверждаются данными полиграмм. На рисунках 2 и 3 представлены участки полиграмм для 1-го и 8-го вопросов соответственно. Каналы расположены в сле-

дующем порядке (начиная с верхнего): верхнего дыхания, нижнего дыхания, кожно-гальванической реакции, двигательной активности, фотоплетизмограммы.

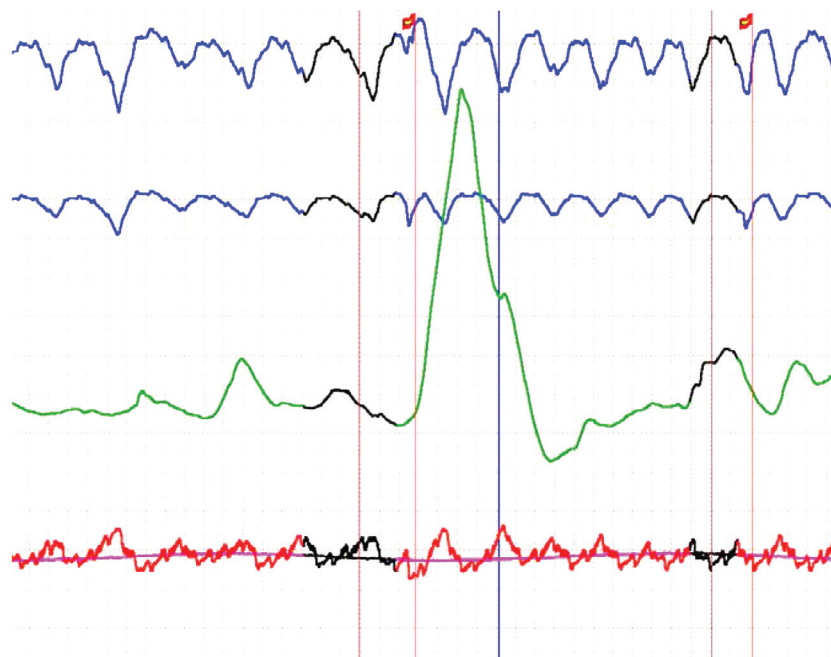


Рис. 2. Участок полиграммы для первого вопроса

Измерение реакции проводилось от первой вертикальной линии. На рисунке 2 наблюдаются значительные изменения в канале кожно-гальванической реакции (вторая линия снизу). На рисунке 3 изменения происходят в канале фотоплетизмограммы (нижняя линия). В состоянии покоя изображе-

ние канала FPG представляет собой линию, близкую к прямой.

Полученные данные позволяют полагать, что рассматриваемый метод дает возможность детально рассмотреть выбросы — получить информацию о дополнительно появляющихся изменениях в био-

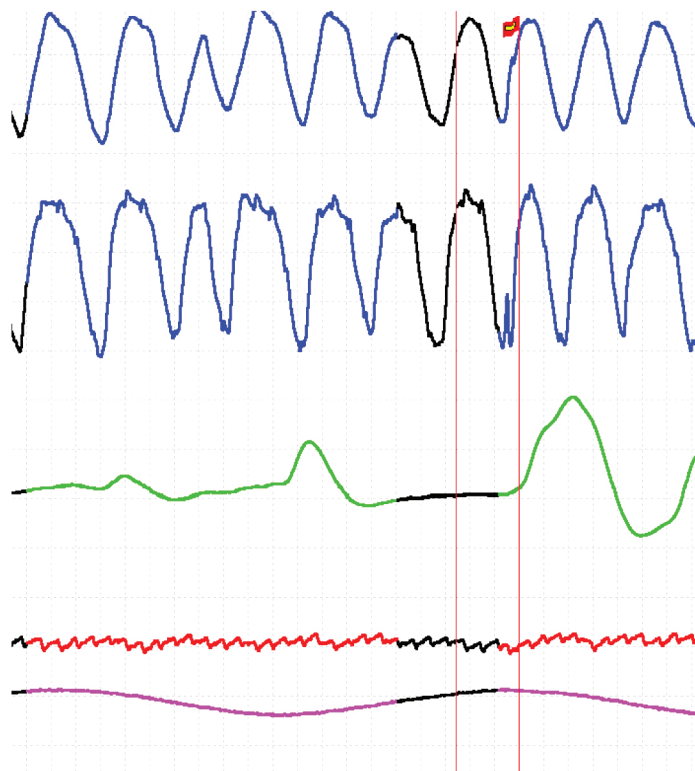


Рис. 3. Участок полиграммы для 8-го вопроса

физических процессах, что повышает точность обработки полиграмм.

При проведении эксперимента с применением метода проекции на латентные структуры (ПЛС) был использован тот же набор данных, что и при применении метода главных компонент. Полученный в результате работы ПЛС-модели график сравнивали с данными полиграфа. Выявлено совпадение анализируемых данных.

Сравнение метода проекций на латентные структуры с методом главных компонент показывает, что первый не позволяет детально рассматривать пики биофизических характеристик, на которые может повлиять хотя бы один канал. Метод удобно использовать, например, для создания определенной модели с фиксированными вопросами в тестовом режиме.

Метод главных компонент позволяет оценить незначительные изменения в биофизических характеристиках человека. Он не сглаживает выбросы, а позволяет отследить их. При детальном рассмотрении можно выявить, почему произошел выброс и как это связать с реакцией испытуемого. Любое явное изменение хотя бы в одном канале может означать, что в организме идет какой-либо физико-химический процесс, который может быть значимым.

Полученные результаты позволяют сделать вывод о том, что многомерный анализ данных (метод главных компонент и метод проекции на латентные структуры) целесообразно применять для изучения биофизических характеристик по данным полиграмм.

### Библиографический список

1. Минакова Н.Н., Поляков В.В., Толстошеев С.Н. Методы технической и правовой защиты информации в сети Интернет. — Барнаул, 2015.
2. Марков А.С., Цирлов В.Л., Барабанов А.В. Методы оценки несоответствия средств защиты информации / под ред. А. С. Маркова. — М., 2012.
3. Поляков В.В., Трушин В.А. и др. Региональные аспекты технической и правовой защиты информации. — Барнаул, 2013.
4. Журавлева Е.В., Салита Д.С. Методы анализа многомерных данных : учебное пособие. — Барнаул, 2016.
5. Ogloblin S.I., Molchanov A.Yu. Инструментальная «детекция лжи»: академический курс. — Ярославль, 2004.

6. Леонтьев К.А., Панин С.Д., Холодный Ю.И. Оценка результатов тестирования на полиграфе методами регрессионного анализа // Наука и Образование : электронный журнал / МГТУ им. Н.Э. Баумана. — 2014. — №10.

7. Абраменко Е.А., Минакова Н.Н., Ушаков В.Я. Исследование свойств полиэтилена с наноразмерными наполнителями специальной обработкой изображения макроструктуры // Известия высших учебных заведений. Физика. — 2008. — Т 51, № 7.

8. Эсбенсен К., Кучерявский С.В. Анализ многомерных данных : избранные главы. — Барнаул, 2003.

9. Ben-Shkufiar G., Furedy J. Theories and applications in (he detection of deception. A psychophysiological and international perspective. N. Y. Inc.: Springer-Verlag, 1990.

10. Сошников А.П., Пеленицын А.Б. Сравнение различных систем количественного анализа полиграмм с помощью алгоритма ChanceCalc [Электронный ресурс]. URL: <http://www.antey-group.ru/jurnal13.html> (дата обращения: 05.12.17).