

ОЦЕНКА ВОЗМОЖНОСТЕЙ ЛИНЕЙНЫХ РЕГРЕССИОННЫХ МОДЕЛЕЙ В ЗАДАЧАХ ПРОГНОЗА ВАРИАБЕЛЬНОСТИ СЕРДЕЧНОГО РИТМА

Туровский Я. А.^{*,**}, Авцинов И. А.^{***}, Суворцев А. С.^{***}

**Воронежский государственный университет,
Россия, г. Воронеж, Университетская площадь, 1*

***Институт проблем управления им. В. А. Трапезникова РАН,
Россия, г. Москва, ул. Профсоюзная, д. 65*

****Воронежский государственный университет инженерных технологий,
Россия, г. Воронеж, пр-т Революции, д. 19
yaroslav_turovsk@mail.ru, alexander.surovtsev.vrn@gmail.com*

Аннотация: представлена модель прогнозирования variability сердечного ритма. Используются три подхода для изучения зависимостей средних значений модулей регрессионных остатков от изменяемых переменных (величина горизонта прогнозирования и номер первого прогнозируемого RR-интервала) при максимальном значении горизонта прогнозирования равном 30 кардиоинтервалам (20-30 секунд).

Ключевые слова: прогнозирование временных рядов, линейная регрессия, variability сердечного ритма.

Введение

Значительная часть аварий и катастроф техногенного характера обусловлена человеческим фактором. Для предотвращения возникновения таких происшествий используют системы контроля деятельности человека-оператора (КДЧО). Система КДЧО эргатической системы должна действовать таким образом, чтобы возможные действия оператора, порождённые в том числе и его психофизиологическим состоянием, не приводили к аварийным ситуациям. Оценка возможных действий оператора может быть получена на основе его функционального состояния (ФС).

Для оценки ФС в системах КДЧО могут быть использованы как субъективные, так и объективные показатели. Различные самоотчеты и заполненные опросники составляют группу субъективных показателей, которые имеют весьма ограниченное применение в эргатических системах. Объективные показатели основываются на анализе физиологических параметров человека [1]. Приведем два примера систем, которые используют объективные показатели. Первый пример – это система мониторинга усталости водителей [2] на основе данных о кожно-гальванической реакции (КГР) [3]. Второй – система, производящая контроль операторов, имеющих доступ к конфиденциальной информации [4], на основе оценки variability сердечного ритма (ВСР) [5]. Variability сердечного ритма – физиологический феномен варьирования временных интервалов между сердечными ударами. В большинстве случаев, мониторинг ВСР заключается в отслеживании нахождения значений параметров внутри границ, выход за границы означает ухудшение деятельности оператора [1]. Граничные значения определяются индивидуальными особенностями конкретного оператора [6]. Очевидно, что прогнозирование ФС позволит повысить эффективность КДЧО за счет того, что система сможет заранее увидеть возможные отклонения состояния человека и заблаговременно запустить соответствующие механизмы системы безопасности для предотвращения возникновения аварийной ситуации.

Целью данной работы является оценка регрессионной модели прогнозирования ВСР человека как в состоянии покоя, так и в состоянии функциональных нагрузок.

Материалы и методы

При проведении исследования использовалась база ВСР 30 человек. Сигналы ВСР выделялись из записей электрокардиограмм с помощью программы Поли-Спектр. Запись электрокардиограмм производилась электрокардиографом Поли-Спектр-12 с частотой квантования электрокардиограммы в 1000 Гц. В этой базе для каждого человека содержится 9 коротких записей ВСР (3-5 минут),

полученных при разных функциональных нагрузках: работа с двумя вариантами дыхательного интерфейса, работа с электромиографическим (ЭМГ) интерфейсом, работа с нейрокомпьютерным интерфейсом, работа с окулографическим интерфейсом, фоновые записи перед работой с каждым видом интерфейса. Все сигналы ВСП перед сохранением в базу данных были подвергнуты процедурам удаления артефактов и экстрасистол, которые рассматривались как помехи для временного ряда ВСП.

В качестве модели для прогнозирования рассмотрен вариант использования линейной регрессии. Выбор обусловлен тем, что линейная регрессия является одним из методов, обладающим наименьшими требованиями к вычислительным ресурсам.

Очевидно, что при прогнозе значений временного ряда происходит накопление ошибки. Процедура её оценки заключается в сравнении значений, предсказанных на некоторый отрезок времени, и эмпирически полученных значений. Исходя из этого, в качестве меры ошибки в работе используется средние значения модулей регрессионных остатков.

Для прогнозирования будем составлять линейные модели, на вход которым будем подавать отрезок временного ряда определенной длины. На выход каждой модели будет подаваться одно значение. Идентификацию таких моделей будем осуществлять путем решения линейных уравнений, для чего будем применять библиотеки `numpy` и `xlsxwriter`, которые обеспечат соответственно решение уравнений линейной алгебры и формирование результатов в виде Excel-таблиц. Удаление выбросов производится кластерным анализом средствами математического пакета `Statistica`.

Результаты и обсуждение

Проведено исследование зависимости средних значений модулей ошибок прогнозирования (средних значений модулей регрессионных остатков, СЗМРО) от параметров прогнозирования при использовании моделей линейной регрессии с разным числом факторов m (30, 60 и 90 RR-интервалов). Так как в работе рассматривается задача прогнозирования временных рядов, то в качестве параметров модели прогнозирования будем использовать следующий набор параметров прогнозирования:

$$(1) \quad \text{НПП} = f(\text{startpos}, \text{forecastlength}),$$

где: *startpos* – номер первого кардиоинтервала, для которого составляется прогноз, *forecastlength* - величина горизонта прогнозирования.

Каждый подход определяется отображением множества порядковых номеров на множество элементов *НПП*. Рассмотрим принципы построения упорядоченных множеств элементов *НПП* для каждого подхода отдельно. Для наглядности построим диаграммы Ганта по каждому подходу. На рис. 1 под буквами а), б) и в) показаны диаграммы Ганта для первого, второго и третьего подходов соответственно. На каждой диаграмме изображено по 6 полос. Одна полоса (весь временной ряд) отображает всю последовательность ВСП, разделенную на обучающую (залита черным цветом) и тестовую выборки (залита серым цветом). Другие пять полос – подпоследовательности тестовой выборки. Каждая из этих пяти полос, разделена на последовательность входных данных (области, залитые серым цветом с косой штриховкой) и последовательность точек, для которых составляется прогноз (залитые серым цветом области с крестовой штриховкой). В общем случае, для каждого прогноза своя последовательность кардиоинтервалов, которая подается на вход модели для рекуррентного вычисления значений кардиоинтервалов (на графиках – прогнозируемые значения) на некоторый горизонт прогнозирования.

Исследование средних значений модулей регрессионных остатков методами непараметрической статистики позволило установить, что наилучшие результаты прогнозов дают нам модели с m равным 30 или 60, худшие результаты дают нам модели с m равным 90. Это можно объяснить тем, что вариабельность сердечного ритма представляет собой композицию волн разных периодов и при m равном 30 или 60 кардиоинтервалам обеспечивается аппроксимация восходящих или нисходящих трендов ULF и VLF частотных диапазонов, а не полных периодов волн данных частотных диапазонов. В рамках каждого подхода по исследованию зависимостей СЗМРО от изменяемых параметров (горизонта прогнозирования для первого подхода и номера первого прогнозируемого RR-интервала для второго и третьего подходов) возникают ситуации, когда СЗМРО некоторых сигналов существенно отклоняются от распределения для какой-либо из моделей (модели с некоторым числом факторов m). Эти сигналы выделены кластерным анализом как сильно отличающиеся, поэтому они не были выключены в расчеты результатов прогнозирования.

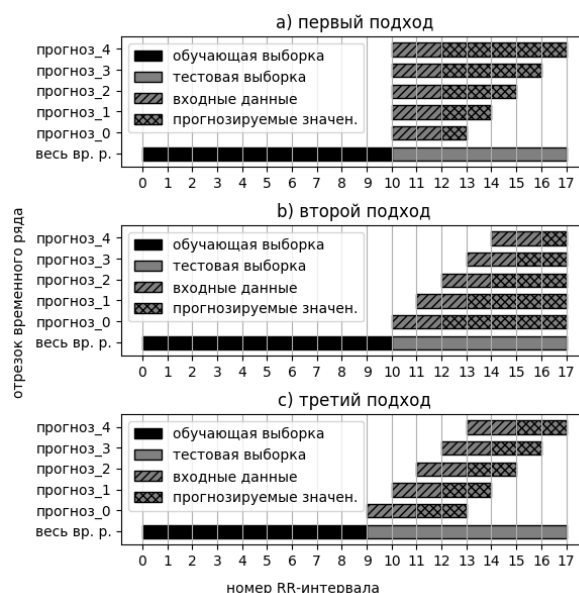


Рис. 1. Диаграмма Гантта принципов разбиения данных на обучающую и тестовую выборки для трех подходов.

Заключение

В статье представлена модель прогнозирования variability сердечного ритма. Используются три подхода для изучения зависимостей средних значений модулей регрессионных остатков от изменяемых переменных (величина горизонта прогнозирования и номер первого прогнозируемого RR-интервала) при максимальном значении горизонта прогнозирования равном 30 кардиоинтервалам (20-30 секунд). В качестве данных используется набор записей ВСП 30 человек, находящихся в разных функциональных состояниях. Было показано, что рост средних значений модулей регрессионных остатков происходит как при увеличении величины горизонта прогнозирования, так и при увеличении номера первого прогнозируемого кардиоинтервала. Установлено, что значения ошибки моделей с m равным 30 и 60 статистически не различаются, а модели с m равным 90 дают ошибки выше, чем модели с m равным 30 и 60, при прочих равных условиях. Перспективой использования данных методов обработки ВСП, является разработка информационной системы, осуществляющей прогнозирование ВСП с целью оценки возможных отклонений от нормы для человека-оператора.

Работа выполнена при поддержке гранта РФФИ (грант 16-29-08342-офи_м).

Литература

1. Машин В. А. Психологическая нагрузка, психологическое напряжение и функциональное состояние операторов систем управления / В. А. Машин // Вопросы психологии – 2007. – №6. – С. 86-96.
2. Савченко В. В. Бортовая система мониторинга функционального состояния оператора транспортного средства / В. В. Савченко // Объединенный институт машиностроения НАН Беларуси. Минск. – 2012. – №1(18). – С. 20-25.
3. Ковальчук А. С. О возможности идентификации психофизического состояния человека-оператора по variability сердечного ритма / Ковальчук А. С. // Динамика систем, механизмов и машин. Омск. – 2014. – С. 37-40.
4. Нигрей А. А. Исследование проблемы распознавания психофизиологического состояния человека-оператора / А. А. Нигрей // Скиф. Вопросы студенческой науки. – 2017. – №10(10). – С. 61-64.
5. Петрукович В. М. Оценка когнитивной деятельности операторов по нестационарным периодам variability сердечного ритма / В. М. Петрукович, В. Я. Апчел, М. В. Зотов // Вестник Российской Военно-Медицинской Академии. Санкт-Петербург. – 2011. – №2(34). – С. 138-146.
6. Баевский Р. М. Анализ variability сердечного ритма при использовании различных электрокардиографических систем / Р. М. Баевский // Вестник аритмологии. – 2001. – №24. – С.65-87.