

**Алгоритм кластеризации QRS-комплексов электрокардиосигнала,
основанный на методе кластеризации k-means.**

Научный руководитель – Калининченко Александр Николаевич

Ван Геннадий Вадимович

Студент (магистр)

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина), Факультет информационно-измерительных и биотехнических систем (ФИБС), Базовая кафедра Медицинские информационные и биотехнические системы, Санкт-Петербург, Россия

E-mail: gennady.van@mail.ru

Электрокардиограмма является источником получения диагностической информации, используемой для контроля сердечной активности. Одной из актуальных проблем современной кардиологии является получение наиболее полной информации об электрическом потенциале сердца, на основе которого можно было бы расширить диагноз патологических состояний миокарда и его электрофизиологических свойств. Анализ кардиоциклов, полученных из электрокардиосигнала, может отражать симптомы, указывающие на то, что сердце нуждается в немедленном обследовании. Наиболее эффективным и перспективным подходом к анализу электрокардиосигнала является применение методов искусственного интеллекта. Важными задачами, которые требуется решить для реализации данного подхода, являются выбор метода машинного обучения или нейронной сети и выбор информативных признаков. Задача формальной кластеризации, по своей сути, сводится к задаче определения оптимального количества кластеров, используемых при кластеризации. Анализ литературы показал, что одним из наиболее эффективных подходов к решению подобных задач является использование совокупности «Elbow»-метода и метода машинного обучения k-means, которая была принята в качестве базового [1, 2]. Предложенный алгоритм формальной кластеризации был реализован в среде программирования Jupiter для языка программирования Python 3 с использованием библиотеки машинного обучения scikit-learn.

На первом этапе производится регистрация электрокардиосигнала пациента.

На втором этапе сигнал подвергается процедуре фильтрации для повышения точности алгоритма. Для преобразования используется полосовой фильтр с полосой пропускания от 5 до 35 Гц.

На третьем этапе производится расчет информативных параметров для каждого из информативных признаков. Для данного алгоритма была выбрана комбинация таких информативных признаков, как скорость нарастания сигнала, длительность QRS-комплекса; глобальный ненулевой максимум в частотном спектре, соответствующем полосе пропускания фильтра, а также площадь QRS-комплекса [3, 4].

На четвертом этапе проводится вычисление оптимального количества кластеров QRS-комплексов в ЭКГ. В данной работе используется «Elbow»-метод [2].

На пятом этапе проводится кластеризация QRS-комплексов в ЭКГ при помощи метода кластеризации k-means [1].

Для проведения исследования алгоритма кластеризации использовались фрагменты аннотированной базы данных сигналов содержащей 48 записей по трем стандартным отведениям, длительностью 60 секунд каждый, с частотой дискретизации 250 Гц. Данные сигналы были получены из базы сигналов «MGH/MF waveform database» [5].

Точность алгоритма, Асс, определенная по формуле (1) составила 97,25 [6, 7].

где P - количество правильно классифицированных QRS-комплексов; N - размер выборки.

Дальнейшее повышение точности данного алгоритма возможно за счет увеличения количества информативных признаков; корректировки процедуры предварительной обработки сигнала; изменение метода кластеризации, положенного в основу алгоритма.

Источники и литература

- 1) 1. MacQueen, J. Some methods for classification and analysis of multivariate observations. // Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1967, Vol. 1, P. 281-297.
- 2) 2. Robert L. Thorndike. Who Belong in the family? // Psychometrika, 1996, Vol.18, P. 267-276.
- 3) 3. Van G.V., Podmasteryev K.V. The analysis of informative features for R-peak detection // XII Internantional Scientific Conferense «Physics and Radioelectronics in Medicine and Ecology» with Scientific Youth session PhREME'2016 (Vladimir - Suzdal, Russia, 5-7 July, 2016), Proceedings, Book II. - P. 126-129.
- 4) 4. Hampton J.R. The ECG made easy. London: Churchill Livingstone, 2003.
- 5) 5. Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K., and Stanley, H. E. 2000. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation 101, 23, P. 215-220.
- 6) 6. ГОСТ ИЕС 60601-2-51–2011. Изделия медицинские электрические. Ч. 2-51. Частные требования безопасности с учетом основных функциональных характеристик к регистрирующим и анализирующим одноканальным и многоканальным электрокардиографам. М.: Изд-во стандартов, 2013. 104 с.
- 7) 7. Balouchestani, M. and Krishnan, S. (2016) Advanced K-Means Clustering Algorithm for Large ECG Data Sets Based on a Collaboration of Compressed Sensing Theory and K-SVD Approach. Signal, Image and Video Processing, 10, 113-120.