

УДК: 004.896

**АВТОМАТИЧЕСКИЙ АНСАМБЛЬ МОДЕЛЕЙ МАШИННОГО
ОБУЧЕНИЯ ДЛЯ РЕШЕНИЯ ЗАДАЧИ CHANGE POINT DETECTION ПРИ
АНАЛИЗЕ ГЕОФИЗИЧЕСКИХ ДАННЫХ СКВАЖИН.**

Самофалов А.С. (Университет ИТМО), **Чебышев И.С.** (Газпромнефть-НТЦ)
**Научный руководитель - научный сотрудник лаборатории композитного
искусственного интеллекта Ревин И. Е.** (Университет ИТМО)

Введение. Современные технологические процессы тесно взаимосвязаны и изменения происходящие в одной точки системы должно влечь изменения в остальных цепях одной производственной системы. Таким образом, эффективный производственный процесс должен быть обеспечен цифровыми системами, которые позволяют автономно реагировать на изменения состояний цепочек одной системы [1, 2].

В частности, к производственному процессу требующему своевременное детектирование таковых изменений (Change Points - CPs) относится бурение продолжительных участков горизонтальных скважин в нефтяных пластах. Автономное детектирование изменений в фиксируемых данных геофизических исследований позволит своевременно корректировать траекторию бурения горизонтальных скважин и тем самым влиять на снижение технологического времени ее строительства и на повышение эффективности вскрытия углеводородного пласта [3, 4].

Существующие подходы в задаче Change Point Detection (CPD) сводятся в основном к байесовским методам, архитектурам нейронных сетей, методам минимизации функций или методам на основе сравнение значимых статических различий [2]. Однако, существующие алгоритмы и методы имеют разную точность и разработаны под детектирование типовых изменений данных в каждой из конкретных областей промышленного применения [5, 6, 7].

Основная часть. В разрабатываемом алгоритме мы базируемся на методе анализа собственного подпространства временного ряда (Singular Sequence Transformation - SST), который позволяет в каждый момент времени оценивать пространственную удаленность текущего и предыдущего временного окна.

Для улучшения детектирующей способности SST на примере геофизических данных мы рассмотрели один из современных подходов обработки массивов. В первую очередь, временной ряд необходимо избавить от зашумленности данных и сохранить репрезентативную тенденцию сигнала [8] - для этого нами предлагается рассмотреть один из фильтров, к примеру это может быть фильтр Баттерворта или Савицкого-Голея. Во-вторую очередь, нам необходимо оптимально подобрать гиперпараметры для SST метода. Для решения такой задачи нами рассмотрены методы на основе анализа всего временного ряда (whole-series-based) и анализа подпоследовательностей временного ряда (subsequence-based) [9]. В конце, применяется метод на основе CUMSUM, который позволяет уменьшить кол-во ложноположительных срабатываний SST метода.

Более того, с целью улучшения детектирующей способности SST нами рассмотрен альтернативный подход в алгоритме базируемый на ансамблеи моделей SST и фильтра Калмана [5]. По аналогии с описанным алгоритмом на этапе анализа поступившей информации о текущем и предыдущем временном массиве мы

анализируем при помощи каждого из методов коэффициенты расстояний между временными окнами. Далее, ряды оценок стандартизируются и суммируются с некоторым базовым коэффициентом для каждого ряда в один итоговый массив.

Выводы. Рассматриваемый алгоритм позволяет с лучшим значением F1 и минимальным кол-вом пропущенных CPs обрабатывать одномерные массивы геофизических данных относительно современных подходов базирующихся на архитектурах нейронных сетей. Одной из ключевых особенностей алгоритма является автоматический подбор оптимальных гиперпараметров модели опираясь на поступивший временной ряд. Более того, нет необходимости в дообучении модели на вновь приходящих данных, так как рассматриваемый подход не предполагает использование решений на основе нейронных сетей [7].

В перспективе базируясь на алгоритме с целью инжиниринга данных решается задача регрессии и классификации для задачи онлайн траектории бурения в горизонтальном пласте [3].

Список использованных источников:

1. Samaneh Aminikhanghahi and Diane J. Cook. A survey of methods for time series change point detection. Knowledge and Information Systems, 2016.
2. Charles Truong, Laurent Oudre, and Nicolas Vayatis. Selective review of offline change point detection methods. Signal Processing, 167, 2020. 9.
3. Nikita Klyuchnikov, Alexey Zaytsev, Arseniy Gruzdev, and all. Data-driven model for the identification of the rock type at a drilling bit. Journal of Petroleum Science and Engineering, 178:506–516, 2019.
4. Dmitry Koroteev and Zeljko Tekic. Artificial intelligence in oil and gas upstream: Trends, challenges, and scenarios for the future. Energy and AI, 2021.
5. Ryan Prescott Adams and David J.C. MacKay. Bayesian online changepoint detection. 2007.
6. Tsuyoshi Ide and Koji Tsuda. Change-point detection using krylov subspace learning. Proceedings of the Seventh SIAM International Conference on Data Mining, 2007.
7. Gianluca Manca and Alexander Fay. Off-line detection of abrupt and transitional change points in industrial process signals. The 4th International Conference on Applied Automation and Industrial Diagnostics, 2022.
8. Giannis Giakas and Vasilios Baltzopoulos. A comparison of automatic filtering techniques applied to biomechanical walking data. Journal of Biomechanics, 30:847–850, 1997.
9. Arik Ermshaus, Patrick Schafer, and Ulf Leser". Window size selection in unsupervised time series analytics: A review and benchmark. The 7th International Workshop on Advanced Analytics and Learning on Temporal Data, 2022.