DOI: 10.25728/mlsd.2022.1006

МОДИФИКАЦИЯ АЛГОРИТМА ОЦЕНКИ РИСКОВОГО ПОТЕНЦИАЛА ПРОЦЕССОВ: ДЕТАЛИЗИРУЮЩИЕ КОЭФФИЦИЕНТЫ КРАТНО-МАСШТАБНОГО ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ

Абдулова Е.А.
Институт проблем управления им. В.А. Трапезникова РАН,
Россия, г. Москва ул. Профсоюзная д.65
consoft@ipu.ru

Аннотация: доклад посвящен вопросам интеллектуального анализа данных и построения на его основе мгновенных моделей прогнозирования. Рассмотрен подход к анализу информации для построения мгновенных моделей прогнозирования параметров технологического процесса с учетом детализирующих коэффициентов кратно-масштабного вейвлет-преобразования.

Ключевые слова: рисковый потенциал, мгновенная прогнозирующая модель, кратно-масштабное вейвлет-преобразование, детализирующие коэффициенты.

Введение

Интеллектуальный анализ данных является активно развивающимся направлением в области искусственного интеллекта, связанным с машинным обучением и задачами обнаружения скрытых закономерностей [1-7]. Важным классом задач, для решения которых может использоваться интеллектуальный анализ данных, являются задачи управления значимыми объектами критической информационной инфраструктуры (ЗОКИИ). Появление и развитие средств управления объектами с повышенным риском эксплуатации, относящихся к категории динамических, тесно связано с развитием интеллектуальных систем поддержки принятия решений. Принятие решений требует информации, которая основана на данных.

Имеется ряд существенных отличий Data Mining от других методов анализа данных. Традиционные методы анализа данных (статистические методы) и OnLine Analytical Processing (OLAP) [8] в основном ориентированы на проверку заранее сформулированных гипотез (verification-driven data mining) и на “грубый” разведочный анализ, составляющий основу оперативной аналитической обработки данных (OLAP), в то время как одно из основных положений Data Mining - поиск неочевидных закономерностей.

Объекты критической инфраструктуры, к которым относятся энергетические объекты, объекты нефтегазовой промышленности, транспорта, необходимо рассматривать во взаимосвязи информационной и физической инфраструктур объекта. Компьютерные атаки на ЗОКИИ могут привести не только к выходу параметров за установленные технологические пределы, но и вызвать нештатные ситуации с серьезными последствиями или остановку объекта (например, остановка трубопровода Colonial Pipeline после кибератаки). В связи с этим в состав интеллектуальных систем поддержки принятия решений [9] включают подсистему мониторинга угроз безопасности, включающую прогнозирование и диагностирование хода технологического процесса (ТП) (в том числе и на основе интеллектуального анализа и вейвлет-анализа данных) [10, 11], идентификацию и управление рисками [5, 12, 13].

1 Обобщенный алгоритм оценки рискового потенциала

Для систем, предназначенных для ЗОКИИ, все шире используются методы интеллектуального анализа данных. Отличительной особенностью таких систем является динамичность развития процессов. Поэтому эти системы требуют разработки методов, учитывающих фактор времени при комплексном анализе рисков возникновения неисправностей и аномалий.

На рис. 1 представлен обобщенный алгоритм оценки рискового потенциала процесса, протекающего в системе ЗОКИИ. К ключевым этапам оценки рискового потенциала относятся:

- подготовка данных,
- интеллектуальный анализ данных,
- мгновенная модель прогнозирования,
- оценка рискового потенциала.

Особое внимание в данном алгоритме уделяется вопросам разработки мгновенной прогнозирующей модели, а именно, определению структуры модели с учетом предыстории на основе корреляционного анализа и/или мер зависимости [11, 16-18].

Рис. 1. Обобщенный алгоритм оценки рискового потенциала процесса

2 Мгновенная прогнозирующая модель

Представим мгновенную модель прогнозирования состояния процесса в значимом объекте критической информационной инфраструктуры [15] в стандартизованном масштабе в виде:

$$\hat{y}(t) = \sum_{i=1}^{m} \hat{a}_i \hat{y}(t-i) + \sum_{s=1}^{S} \sum_{j=1}^{r_s} \hat{b}_{s,j} \hat{x}(t-j)_s \,, \qquad (1)$$

где $\hat{y} = (y - M[x])/\sigma_y$ – прогноз выхода объекта в момент времени t, $\hat{x} = (x - M[x])/\sigma_x$ – вектор входов мгновенной модели, $M[\hat{y}] = M[\hat{x}] = 0$, $\sigma_{\hat{y}} = \sigma_{\hat{x}} = 1$, m – глубина памяти по выходу, r_s – глубина памяти по входу $\left(R = \max_{s=\overline{1,S}} r_s\right)$, s – размерность вектора входов, $\hat{a}_i$, $\hat{b}_{s,j}$ – стандартизованные коэффициенты, настраиваемые в каждый момент времени, векторы $\hat{x}(t-j)_s$, и соответствующие им $\hat{y}(t-i)$ выбираются из базы данных не в хронологическом порядке. В каждый момент времени строится новая модель в соответствии с (1).

При конечном уровне разложения L любой стандартизованный входной вектор $\hat{x}(t)$ и прогнозируемый стандартизованный выход $\hat{y}(t)$ представим в виде кратно-масштабного вейвлет-преобразования:

$$\hat{x}(t_i) = \sum_{k=1}^{N} c_{L,k}^{\hat{x}}(t_i)\varphi_{L,k}(t_i) + \sum_{l=1}^{L}\sum_{k=1}^{N} d_{l,k}^{\hat{x}}(t_i)\psi_{l,k}(t_i), \quad (2)$$

$$\hat{y}(t) = \sum_{k=1}^{N} c_{L,k}^{\hat{y}}(t)\varphi_{L,k}(t) + \sum_{l=1}^{L}\sum_{k=1}^{N} d_{l,k}^{\hat{y}}(t)\psi_{l,k}(t), \quad (3)$$

где $\psi_{l,k}(t)$ – базисная вейвлет-функция; $\varphi_{L,k}(t)$ – базисная масштабирующая функция; $c_{L,k}$ и $d_{l,k}$ – масштабирующие и детализирующие коэффициенты соответственно; $l,k \in I$ – параметры масштаба (уровень детализации) и сдвига соответственно; L – глубина кратно-масштабного разложение (где $L_{\max} = \log_2 N^*$, и N^* – мощность множества состояний в базе данных). Условия устойчивости для мгновенной модели прогнозирования с учетом представления входа и выхода модели в виде кратно-масштабного вейвлет-преобразования приведены в [11].

3 Оценка рискового потенциала и детализирующие коэффициенты кратно-масштабного вейвлет-преобразования важных показателей ТЭП

АЭС относятся к значимым объектам критической информационной инфраструктуры, а в АСУ ТП АЭС выделяют системы верхнего уровня, в которых информационные задачи включают в себя мониторинг состояния объекта управления. В рамках систем верхнего уровня реализована информационная задача расчета технико-экономических показателей, анализ которых позволит выявить аномалии в работе АЭС. Расчет технико-экономических показателей (ТЭП), основанный на кратно-масштабном вейвлет-преобразовании, интеллектуальном анализе данных и моделях прогнозирования рискового потенциала, позволит эксплуатационному персоналу определять допустимые эксплуатационные пределы и местоположения аномалий. Как отмечалось ранее условия устойчивости мгновенной прогнозирующей модели определяются с использованием кратно-масштабного вейвлет-преобразования [11], которые хорошо зарекомендовали себя для анализа нестационарных сигналов различной природы.

Рассмотрим один из важных показателей ТЭП – температуру питательной воды за подогревателем высокого давления 5 (ПВД5), который подогревает питательную воду для дальнейшей подачи через систему питательной воды в парогенераторы реакторной установки во всех режимах нормальной эксплуатации. Температура питательной воды за ПВД5 является расчетным параметром, зависящим от многих параметров, которые не входят в состав важнейших параметров энергоблока. В связи с тем, что снижение температуры питательной воды приводит к снижению давления в парогенераторах, что ведет к уменьшению расхода пара на турбину и снижению выработки электроэнергии, данный параметр требует постоянного мониторинга.

Рассмотрим прогноз температуры питательной воды за подогревателем высокого давления 5 (ПВД5) энергоблока АЭС, упрощенная тепловая схема которого приведена на рис. 2.

Рассмотрим мгновенную модель прогнозирования температуры питательной воды за подогревателем высокого давления, основанную на 28 векторах [11]:

$$T_m(t) = \sum_{j=1}^{4}\left[b_j T_1(t-j) + b_{j+4}T_2(t-j) + b_{j+8}T_3(t-j) + b_{j+12}T_4(t-j) + b_{j+16}N(t-j)\right], \quad (4)$$

которая является функцией температуры питательной воды в парогенераторах ($T_i(t-j)$, $i = \overline{1,4}$ – номер парогенератора) и активной мощности генератора $N(t-j)$ которые выбираются не в хронологическом порядке, $(t-j)$ – момент времени, j – глубина предыстории.

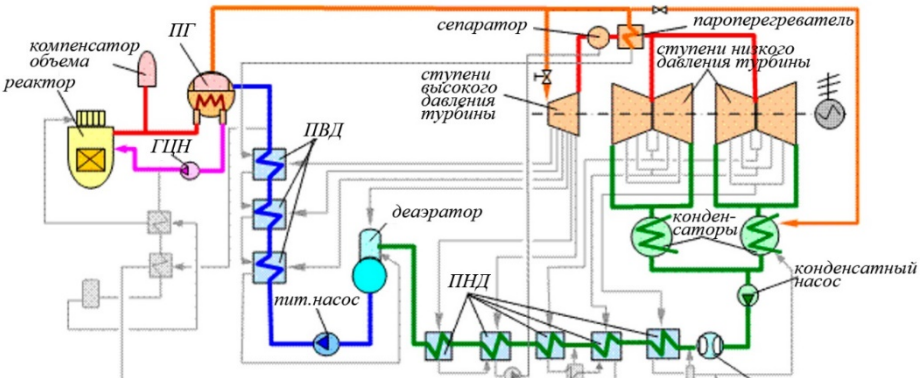

Рис. 2. Упрощенная тепловая схема энергоблока АЭС с реактором типа ВВЭР-1000

На рис. 3 представлена выборка из базы данных одного из параметров модели (температура питательной воды в парогенераторе 1ой петли) (4) и его детализирующие коэффициенты при разложении временного ряда параметра модели по вейвлетам Хаара до уровня L=9.

Анализ статистических характеристик коэффициентов детализации для компоненты входного вектора показывает, что плотность распределения вероятностей мгновенных значений коэффициентов детализации хорошо описывается гауссовским законом с параметрами $N(\mu;\sigma^2)$ для 1го уровня разложения. На рис. 4 представлены гистограммы детализирующих коэффициентов параметров модели (4).

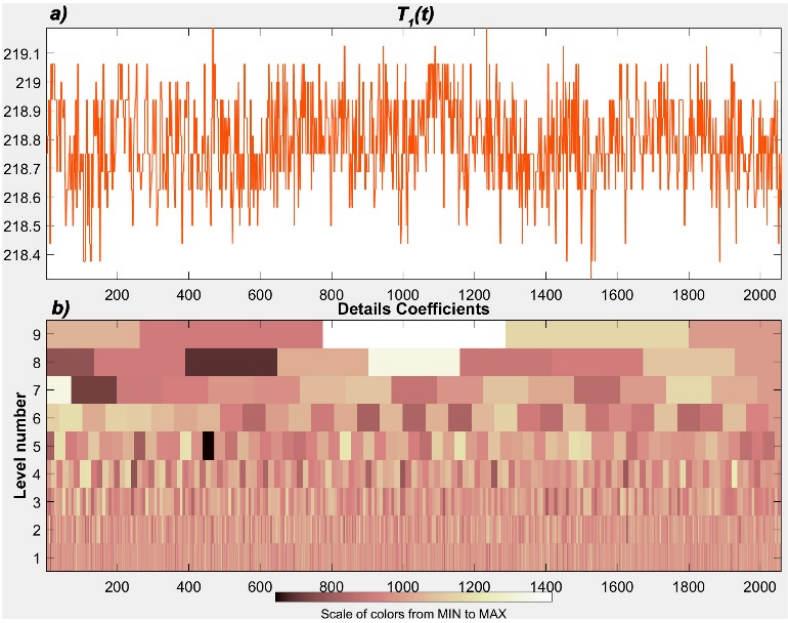

Рис. 3. Параметр $T_1(t)$ и его детализирующие коэффициенты при разложении по вейвлетам Хаара

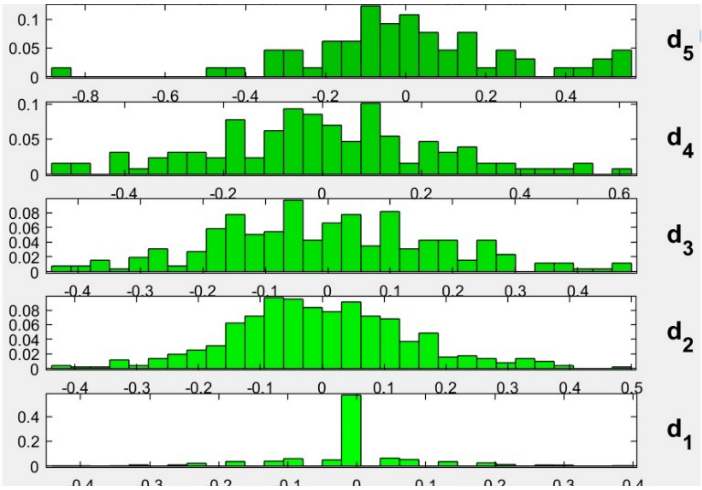

Рис. 4. Гистограмма для детализирующих коэффициентов $T_1(t)$

Вывод о нормальном законе распределения распространяется и для выбранных 28 векторов для построения модели (4). На рис. 5 представлены функции распределения $\Phi(x)$ детализирующих коэффициентов для $T_1(t)$ на 25 временных точках.

На основе проведенного анализа детализирующих коэффициентов разложения параметров модели (4) алгоритм оценки рискового потенциала (см. рис. 1) [19] с использованием вербально-числовой шкалы Харрингтона [20] на этапе «Мгновенная модель прогнозирования» дополнен оценкой метрики соответствия нормальному распределению детализирующих коэффициентов параметров модели. Применение модификации алгоритма оценки рискового потенциала позволило улучшить качество прогнозирующих моделей вида (4) и соответственно повысить точность оценки рискового потенциала (см. рис. 6).

Представленная мгновенная модель предназначена для прогнозирования аномалий технологического процесса, выхода параметров за технологические пределы, а также оценки рискового потенциала.

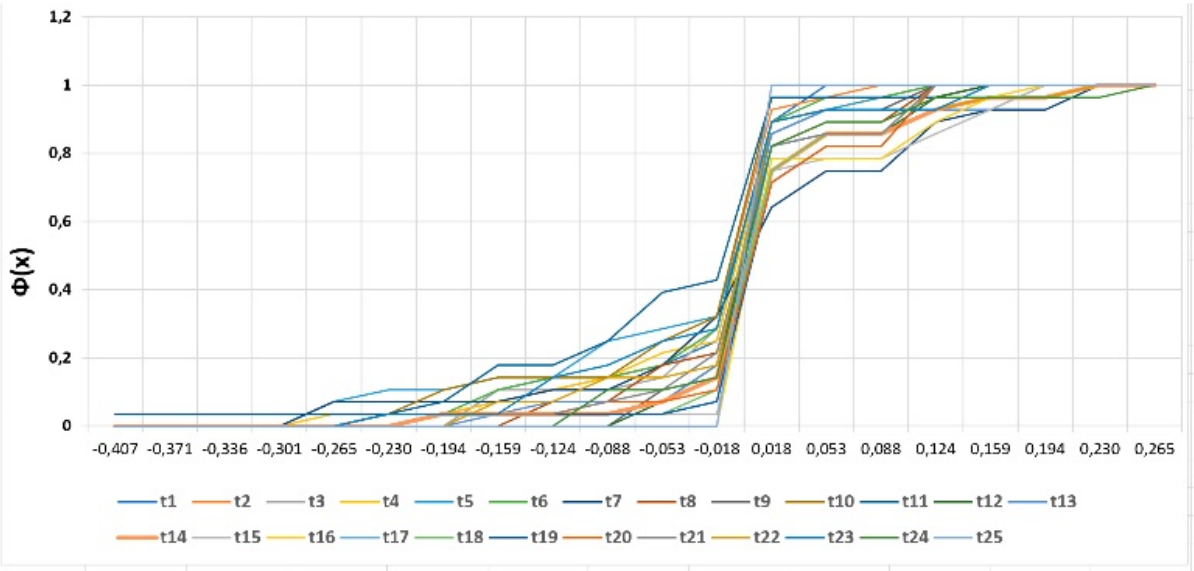

Рис. 5. Функции распределения детализирующих коэффициентов $T_1(t)$

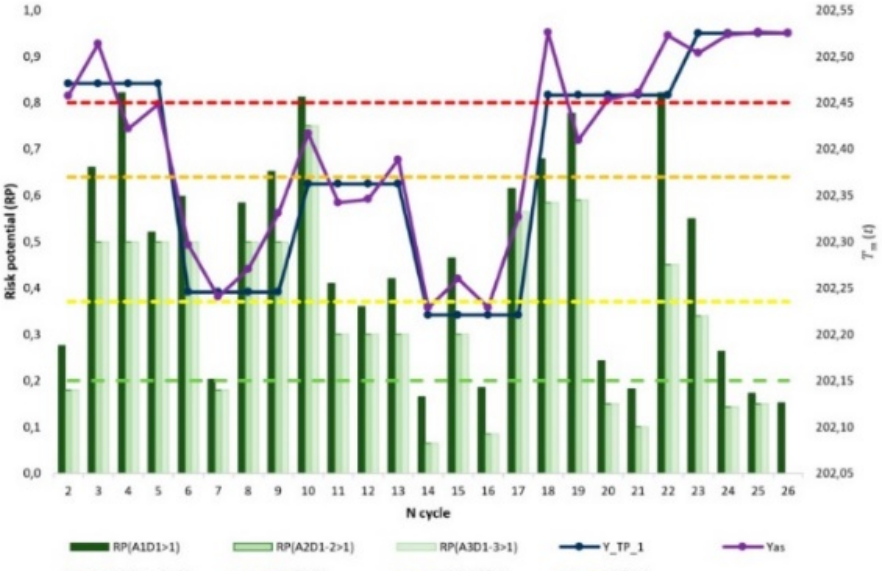

Рис. 6. Выход мгновенной модели $T_m(t)$ и рисковый потенциал процесса

Заключение

Мониторинг состояния процесса в значимом объекте критической информационной инфраструктуры позволяет выявлять изменения различного характера, оценивать риск-потенциал. Разработанный подход внедряется в комплекс информационных задач систем верхнего уровня [11, 21, 22].

Литература

1. *Dogan A., Birant D.* Machine learning and data mining in manufacturing // Expert Systems with Applications. Vol. 166:114060. 2021.
2. *Hunt I.* In-sample tests of predictability are superior to pseudo-out-of-sample tests, even when data mining // International Journal of Forecasting. 2022. (в печати)
3. *Norouziasl S., Jafari A., Zhu Y.* Modeling and simulation of energy-related human-building interaction: A systematic review // Journal of Building Engineering. Vol. 44:102928. 2021.
4. *Qiu X., Yang M., Yu Y., Dai X.* The Signal Intelligent Analysis System of Eventlog // Proceedings of the 2020 International Conference on Intelligent Computing, Automation and Systems (ICICAS). 2020. – P. 92-95.
5. *Peng Y., Zhang Y., Tang Y., Li S.* An incident information management framework based on data integration, data mining, and multi-criteria decision making // Decision Support Systems. Vol. 51, 2011, №. 2. – P.316-327.
6. *Beley O., Chaplyha V.* The application of neural networks for the intelligent analysis of multidimensional data // Proceedings of the 2017 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T). 2017. – P.400-404.
7. *Gouda A., Gomaa S., Attia A., Emara R., Desouky S.M., El-hoshoudy A.N.* Development of an artificial neural network model for predicting the dew point pressure of retrograde gas condensate // Journal of Petroleum Science and Engineering. Vol. 208, Part E:109284. 2021.
8. *Xie Z., Ying H., Yue C., Zhang M., Chen G., Ooi B.C.* Cool: a COhort OnLine analytical processing system // Proceedings of the 2020 IEEE 36th International Conference on Data Engineering (ICDE). 2020. – P.577-588.
9. *Jharko E., Chernyshov K.* Towards Basic Design Principles of Next Generation Human-Operator Decision Support Systems // Proceedings of the 2020 7th International Conference on Control, Decision and Information Technologies (CoDIT). 2020. – P.861-867.
10. *Chernyshov K., Jharko E.* Nuclear Reactor Reactivity Estimation within NPP Diagnostics Problems // Proceedings of the 2021 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM). 2021. – P.484-491.
11. *Abdulova E.* Some Issues of Assessing the Risk Potential of Technological Process at Nuclear Power Plants // Proceedings of the 2021 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM). 2021. – P.602-607.

12. *Leszczyna R.* Review of cybersecurity assessment methods: Applicability perspective // Computers and Security. Vol. 108:102376. 2021.
13. *Thangavelu M., Krishnaswamy V., Sharma M.* Impact of comprehensive information security awareness and cognitive characteristics on security incident management – an empirical study // Computers and Security. Vol. 109:102401. 2021.
14. *Kalashnikov A., Sakrutina E.* The Model of Evaluating the Risk Potential for Critical Infrastructure Plants of Nuclear Power Plants // Proceedings of the 2018 Eleventh International Conference "Management of large-scale system development. 2018. – P.1-4.
15. *Sakrutina E.* Wavelet analysis based stability conditions of a prediction model // ICINCO 2019 – Proceedings of the 16th Int. Conf. on Informatics in Control, Automation and Robotics. Vol. 1. 2019. – P.707-714.
16. *Chernyshov K.* Constructing Consistent in the Rényi Sense Measures of Dependence within System Identification // IFAC-PapersOnLine. Vol. 48. 2015, № 11. – P.825-830.
17. *Chernyshov K.R.* Towards an Extension and Normalization of the Random Vector Anisotropy Magnitude // Lecture Notes in Networks and Systems. Vol. 364. 2022. – P.200-208.
18. *Chernyshov K.R.* Non-linear Correlation Based Approach to the Identification of Maximally Stationary Systems // Lecture Notes in Networks and Systems. Vol. 364, 2022. – P.209-218.
19. *Sakrutina E.* Assessing Risk Potential of Critical Infrastructure of NPP: Fractal Analysis and Multiple-Scale Wavelet Transform // Proceedings of the 2020 International Russian Automation Conference (RusAutoCon). 2020. – P.476-481.
20. *Harrington E.C.* The Desirability Function // Industrial Quality Contol. Vol. 21. 1965, № 10. – P.494-498,
21. *Poletykin A., Jharko E., Mengazetdinov N., Promyslov V.* Some issues of creating the new generation of upper level control systems of NPP APCS // Proceedings of the 2017 5th International Conference on Control, Instrumentation, and Automation (ICCIA). 2017. – P.74-79.
22. *Jharko E.* Calculating Technical and Economical Indexes and Providing Normal NPP Operation // IFAC-PapersOnLine. Vol. 50. 2017, № 1. – P.15586-15591.