Построение модели атак для промышленных очистных сооружений

Е. В. Федорченко, Е. С. Новикова, И. Б. Саенко
Лаборатория проблем компьютерной безопасности
Санкт-Петербургский Федеральный исследовательский центр Российской академии наук
{doynikova, novikova, ibsaen}@comsec.spb.ru

Аннотация. **С ростом уровня цифровизации промышленных киберфизических систем увеличивается число возможных точек проникновения в них и, как следствие, растут риски информационной безопасности. Компрометация таких систем может привести к серьезным финансовым, экологическим последствиям как государственного, так и мирового уровня. В настоящей работе решается задача моделирования атак на промышленные очистные сооружения с учетом специфики технологического процесса. Разработанные модели атак будут использованы для формирования набора данных, предназначенного для изучения и исследования безопасности систем управления очисткой воды.**

Ключевые слова: очистные сооружения, промышленные киберфизические системы, модель кибератаки, испытательный стенд

I. Введение

Распространение цифровизации, а также объединение информационных и промышленных технологий, привели к расширению поверхности кибератак на критически важные инфраструктуры, такие как системы водоподготовки и очистки сточных вод. Одновременно с этим растет активность киберпреступников. Так, в 2021 г. была атакована водоочистная станция Oldsmar в США. Злоумышленник получил удаленный доступ к системе и повысил уровень гидроксида натрия в воде. Оператор водоочистной станции заметил изменения и своевременно снизил уровень. В 2022 г. в Великобритании киберпреступники осуществили атаку на компанию South Staffs Water с использованием программы-вымогателя Clop. При этом расследование показало, что они были уверены, что атакуют другую компанию – Thames Water – которая обеспечивает водоснабжение Лондона и Юго-Восточной Англии и контролирует уровень химических веществ в воде.

Успешная реализация кибератак на системы водоподготовки и очистки воды может привести к техногенной катастрофе. Таким образом, задача защиты таких систем от кибератак на основе их проактивного обнаружения и оценки рисков их успешной реализации является актуальной. Данное исследование проводится с целью разработки систем обеспечения кибербезопасности систем водоподготовки и очистки сточных вод.

Общая методология разработки системы обеспечения кибербезопасности должна включать анализ целевой системы и соответствующего технологического процесса, разработку функциональной схемы системы обеспечения кибербезопасности, исследование и разработку методов реализации выделенной функциональности, валидацию разработанных методов, разработку архитектуры системы обеспечения кибербезопасности, ее внедрение и интеграцию (рис. 1). Были выделены следующие функциональные компоненты системы обеспечения кибербезопасности: сбор данных; обнаружение кибератак и аномалий; оценка и анализ рисков; выбор мер безопасности (рис. 1). В данной работе особое внимание уделяется функционалу обнаружения кибератак и аномалий как одному из ключевых в рамках систем обеспечения кибербезопасности. Методология разработки метода обнаружения кибератак и аномалий, в свою очередь, включает следующие этапы (рис. 1): формирование набора данных для анализа, разработка метода обнаружения кибератак и аномалий, проведение экспериментов для валидации разработанного метода. Эти этапы повторяются в цикле.

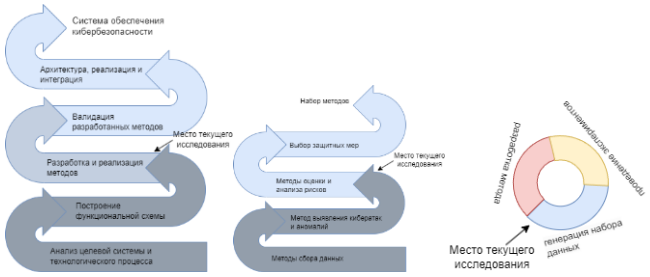

Рис. 1. Место исследования, представленного в данной статье

Данная статья посвящена этапу формирования набора данных. В [1] был описан подход к формированию набора данных. Набор данных должен содержать как нормальные, так и аномальные (полученные в результате кибератак) записи. Для проведения кибератак и генерации аномальных записей необходимо разработать модель атаки на систему водоподготовки и очистки сточных вод. В данной статье представлена разработанная модель атаки.

Статья организована следующим образом. В разделе II описаны релевантные исследования. В разделе III представлены методология и стенд, разработанные для формирования набора данных. В разделе IV приведена предлагаемая модель атак.

II. Релевантные работы

В разделе описываются существующие модели атак на промышленные киберфизические системы. Затем анализируются существующие наборы данных, а также стенды и модели атак, используемые для их генерации. Наконец, рассматриваются методы обнаружения

Работа выполнена при поддержке гранта Российского научного фонда № 23-11-20024, https://rscf.ru/project/23-11-20024/, и Санкт-Петербургского научного фонда.

кибератак и аномалий для систем водоподготовки и очистки сточных вод.

A. *Модели атак для промышленных киберфизических систем*

Существует большое количество исследований, посвященных моделям атак на промышленные киберфизические системы [2–5], включая модели теории множеств [2], цифровых двойников [3], имитационные модели [4, 5] и, в частности, на системы SCADA [6].

В качестве основы для построения модели атаки для системы водоподготовки и очистки сточных вод была выбрана абстрактная модель, предложенная в [3]. В [3] авторы вводят абстрактную модель объекта, конкретную модель объекта, намерения атакующего, модель атакующего и модель атаки.

Модель объекта DM = (*Cm*, *Pr*, *Pe*), где *Cm* – набор компонентов системы (или точек атаки), например, насос или резервуар для сырой воды в случае системы очистки воды; *Pr* – набор свойств системы, т. е. свойств, производимых или контролируемых системой продуктов, например, pH или проводимость в случае системы очистки воды; *Pe* – набор показателей производительности системы, например, галлоны фильтрованной воды, производимые блоком ультрафильтрации в минуту. *Cm* включает в себя физические, кибернетические и логические компоненты. В [3] компоненты классифицируются на исполнительные устройства, датчики, преобразователи и контроллеры.

Модель атакующего *AR* = (*I*, *DM*), где *I* – конечное множество намерений (или целей атакующего), например, повредить, узнать или изменить; *DM* – модель объекта атаки.

Модель атаки для системы *C*, AM_C = (*M*, *G*, *D*, *P*, S_0, S_e), где *M* – множество техник проведения атак, $G \subseteq I$ – конечное множество намерений атакующего, *D* – модель объекта, полученная из модели объекта *DM* системы *C*, $P \subseteq Cm$ – конечное множество точек атаки, S_0 и S_e – множества состояний системы *C*. В [3] авторы также выделяют категории атак с учетом количества компримируемых точек атаки и количества этапов воздействия: одноэтапная одноточечная атака, одноэтапная многоточечная атака, многоэтапная одноточечная атака, многоэтапная многоточечная атака.

B. *Наборы данных*

В последнее время для обнаружения кибератак и аномалий широко используются методы машинного обучения и глубокого обучения. Такие методы требуют больших массивов данных для обучения моделей, позволяющих обнаруживать кибератаки и аномалии.

Анализ существующих наборов данных позволил выделить два типа наборов данных для систем водоподготовки: наборы данных, собранные при нормальной эксплуатации системы, и наборы данных, связанные с безопасностью. Наборы данных собранные при нормальной эксплуатации системы, такие как Full Scale Waste Water Treatment Plant Data [7], Dillman Road Wastewater Treatment Plant Electricity Consumption and Influent Volume[1], Watershed Water Quality – Keypoint Qualifiers[2], Watershed Water Quality – Wastewater[3], обычно получают с реальных водоочистных сооружений. Эти наборы данных могут быть использованы для валидации наборов данных второго типа. Такая валидация необходима, так как наборы данных, связанных с безопасностью, обычно генерируются искусственно или собираются с тестовых стендов.

Анализ исследований показал, что существуют следующие наборы данных, связанные с безопасностью систем водоподготовки: набор данных Secure Water Treatment (SWaT) Сингапурского университета технологии и дизайна [8], набор данных Water Distribution (WADI) [9] и набор данных HIL-based augmented ICS (HAI) security [10]. Основные характеристики этих наборов данных приведены в табл. I. Более подробно со статистикой SWaT и HAI можно ознакомиться в [11]. Следует отметить, что протоколы оценки для этих наборов данных не приводятся, поэтому разные исследователи используют различные метрики для тестирования своих методов на этих наборах данных.

Поскольку моделирование промышленной системы управления и атак на нее в рамках тестового стенда HAI осуществляется искусственно через модификацию переменных, описывающих технологический процесс, рассмотрим модель атаки, использовавшуюся при создании набора данных SWaT (эта же модель используется для создания набора данных WADI) [3].

C. *Обнаружение кибератак и аномалий*

Исследователями были предложены различные методы обнаружения кибератак, от сигнатурных методов [12] до методов машинного обучения [13–15], и аномалий [16–19]. В настоящее время внимание исследователей переключилось на глубокие нейронные сети [20–21] и объяснение обнаруженных аномалий [22, 23]. Существующие исследования имеют ряд ограничений, в том числе ограниченность наборов данных для экспериментов, отсутствие объяснений аномалий, позднее обнаружение атак и их последствий, что ограничивает проактивное обнаружение атак и реагирование на них, отсутствие унификации метрик используемых при валидации. Таким образом, с одной стороны, тема обеспечения безопасности систем водоподготовки является высокоизученной и актуальной, а с другой – существуют проблемы, требующие решения. В настоящем исследовании рассматривается первая проблема, связанная с ограниченностью наборов данных.

III. Методология и стенд для формирования набора данных

Итоговая модель атаки сильно зависит от анализируемой системы (раздел II, A) и методологии формирования набора данных. Поэтому в данном разделе сначала описывается разработанная методология, а затем – разработанный тестовый стенд [1].

A. *Методология*

В работе [1] авторами этой статьи предложена следующая методология формирования набора данных:

[1] https://data.world/city-of-bloomington/c439b26b-90ef-44a8-9a19-94f8323ca771

[2] https://data.cityofnewyork.us/Environment/Watershed-Water-Quality-Keypoint-Qualifiers/3qwy-zqtv

[3] https://data.cityofnewyork.us/Environment/Watershed-Water-Quality-Wastewater/icbf-663g

(1) анализ и спецификация технологического процесса; (2) развертывание тестового стенда; (3) сбор данных в нормальном режиме работы стенда; (4) разработка модели атакующего; (5) разработка модели и сценариев атаки; (6) проведение атак и сбор данных; (7) валидация собранного набора данных.

ТАБЛИЦА I. ПЕРЕЧЕНЬ РЕГИОНОВ… CHARACTERISTICS OF THE SWAT, WADI AND HAI DATASETS

Набор данных	**Характеристики**			
	Технологический процесс	***Тестовый стенд***	***Размер***	***Кибератаки***
SWaT	Процесс водоподготовки и очистки воды состоит из 6 этапов: забор «сырой» воды, добавление реагентов, фильтрация, дехлорирование с помощью УФ-ламп, подача воды в систему обратного осмоса.	Модель станции очистки и обеззараживания воды.	399157 нормальных записей, 50762 аномальных записей.	36 атак, связанных с компрометацией и изменением показаний датчиков. Атаки различаются по количеству скомпрометированных датчиков или точек атаки, а также по количеству пораженных этапов: 19 одноэтапных атак с одной точкой атаки; 6 одноэтапных атак с двойной (5 атак) и тройной (1 атака) точками атаки; 4 многоэтапные атаки с одной точкой атаки на каждом этапе; 3 многоэтапные атаки с несколькими точками атаки, принадлежащими разным этапам водоподготовки.
WADI	Цикл водоподготовки: сбор, очистка и повторное распределение. Модели распределения воды путем забора части очищенной воды из системы обратного осмоса SWaT и смешивания ее с сырой водой.	Расширение SWaT	14 дней бесперебойной нормальной работы и 2 дня данных с атаками.	Атаки, связанных с компрометацией и изменением показаний датчиков. В [5] описаны 2 атаки.
HAI	Водоподготовка и выработка тепловой и насосно-аккумулирующей гидроэлектроэнергии.	Включает котел, турбину, компонент водоподготовки и симулятор аппаратного обеспечения в контуре (HIL). Параметры промышленной системы управления описываются с помощью встроенного симулятора.	HAI 20.07: обучающая выборка – 550024 нормальных записей, 776 записей в результате атак; тестовая выборка - 427073 нормальных записей, 17527 записей в результате атак. HAI 21.03: обучающая выборка – 921603 нормальных записей, 0 записей в результате атак; тестовая выборка - 393058 нормальных записей, 8947 записей в результате атак. HAI 22.04: обучающая выборка – 1004402 нормальных записей, 0 записей в результате атак; тестовая выборка - 349170 нормальных записей, 12030 записей в результате атак.	Атаки, связанные с компрометацией и изменением переменных, описывающих состояние системы. Атаки отичаются по атакуемым устройствам (Emerson Ovation, GE Mark-VIe, и Siemens S7-1500) и модифицируемым переменным (заданные значения, переменные процесса, управляющие переменные и параметры управления).

B. Тестовый стенд

За основу развернутого стенда взят учебный стенд CE 587 (G.U.N.T., GmbH, Германия), реализующий процесс флотации, включающий стадии флокуляции и флотации. Стенд и соответствующие датчики представлены на рис. 2. Конечный стенд оснащен следующими компонентами (или точками атаки):

1) *Физические.*

a) Датчики и исполнительные механизмы. Датчики потока воды, датчики потока воздуха, манометры, измерители уровня воды, pH-датчики, датчики плотности.

b) Преобразователи. Нормализаторы аналоговых датчиков.

c) Контроллеры. Программируемые логические устройства и модули, OPC (Open Platform Communication) сервер, рабочая станция SCADA, сервер Historian и человеко-машинный интерфейс (HMI).

2) *Кибер*. Связи от датчиков к контроллерам; связи от контроллера к контроллеру; связи от контроллера к SCADA.

Следует отметить, что киберфизические системы используют специфические протоколы связи, такие как RS-232, CAN, Modbus и Industrial Ethernet.

3) *Логические*. Все встроенное программное обеспечение и программное обеспечение, например, код программируемых логических устройств.

IV. МОДЕЛЬ АТАК

Используется спецификация модели атаки, представленная в [3]: $AM = (M, G, D, P, S_0, S_e)$, где M – множество техник проведения атак, $G \subseteq I$ – конечное множество намерений атакующего, D – модель объекта, полученная из модели DM системы C, $P \subseteq Cm$ – конечное множество точек атаки, S_0 и S_e – множества состояний C.

Набор техник M предлагается описать с использованием матрицы MITRE ATT&CK для промышленных систем управления[4].

[4] https://attack.mitre.org/matrices/ics/

Был определен следующий набор намерений атакующего для технологического процесса, реализуемого на стенде, описанном в разделе III:

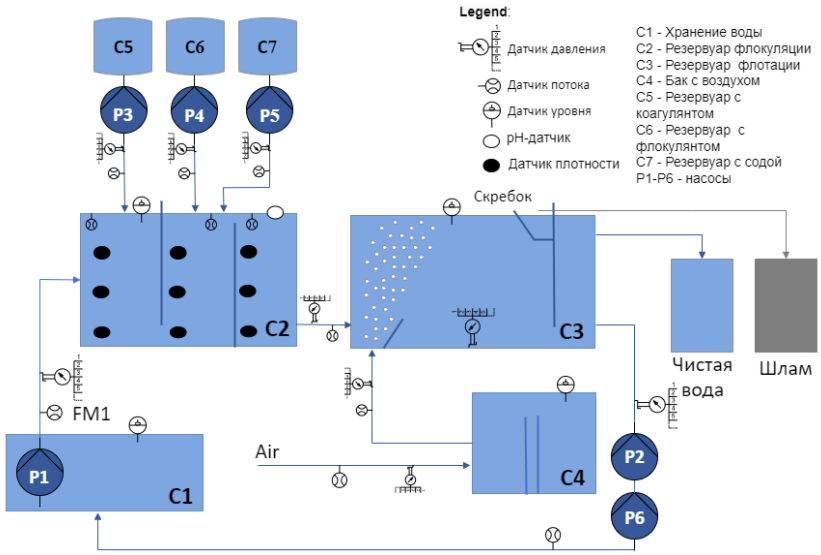

Рис. 2. Схема развернутого стенда

3 отсутствие коагулянта/флокулянта/соды в резервуаре или отказ насоса его подачи, или отсутствие подачи сжатого воздуха в систему;

4 повышение уровня воды в резервуарах флокуляции и флотации;

5 заполнение резервуара чистой воды/шлама;

6 забивка фильтра;

7 выход из строя циркуляционных насосов/скребков/смесителей/насоса сырой воды;

8 утечка жидкости из системы.

Набор точек атаки приведен в табл. II.

ТАБЛИЦА II. НАБОР ТОЧЕК АТАКИ

Точка атаки	Тип компонента	Класс компонента		
		Физи-ческий	Кибер	Логи-ческий
Насосы P1-P6	Испонительные механизмы	+	Связь с контрол-лером	-
Датчик давления	Сенсоры			
Датчик потока воды	Сенсоры			
Датчик потока воздуха	Сенсоры			
Датчик уровня воды	Сенсоры			
pH-датчик	Сенсоры			
Датчики плотности	Сенсоры			
Сервер Historian	Сенсоры			
ПЛК и модули	Контроллер	+	-	+
Нормализаторы для аналоговых датчиков	Преобразова-тель	+	-	+
OPC сервер	Преобразова-тель	+	-	+
SCADA	Контроллер	+	-	+
HMI	Контроллер	+	-	+

1. отсутствие перемешивания сырой воды;

2. некорректная работа расходомера сырой воды, насоса подачи реагентов или pH-датчика;

V. ЗАКЛЮЧЕНИЕ

В статье проанализированы существующие исследования в области моделей атак, наборов данных и методов обнаружения кибератак на промышленные киберфизические системы, а именно на системы водоподготовки и очистки сточных вод. На основе абстрактной модели атак, определенной в релевантных исследованиях, методологии формирования наборов данных и развернутого тестового стенда, была разработана модель атак на систему водоподготовки и очистки сточных вод. Предложенная модель в дальнейшем будет использоваться для определения сценариев атак и проведения атак с целью генерации аномальных записей в наборе данных.

СПИСОК ЛИТЕРАТУРЫ

[1] E. Fedorchenko, E. Novikova, A. Danilov, I. Saenko, “Towards the testbed and dataset for analysis of water treatment systems security,” Lect. Notes in Netw. and Syst., Springer, 2023 [ICDSA, 2023], in press.

[2] M. Jbair, B. Ahmad, C. Maple, R. Harrison, “Threat modelling for industrial cyber physical systems in the era of smart manufacturing,” Comp. in Ind., vol. 137, 103611, 2022. https://doi.org/https://doi.org/10.1016/j.compind.2022.103611.

[3] S. Adepu, A. Mathur, “Generalized Attacker and Attack Models for Cyber Physical Systems,” 2016. 10.1109/COMPSAC.2016.122.

[4] W.L. Duo, M.C. Zhou, and A. Abusorrah, “A survey of cyber attacks on cyber physical systems: Recent advances and challenges,” IEEE/CAA J. Autom. Sinica, vol. 9, no. 5, pp. 784–800, 2022.

[5] Y. Peng, y. Wang, C. Xiang, X. Liu, Z. Wen, D. Chen, C. Zhang, “Cyber-Physical Attack-Oriented Industrial Control Systems (ICS) Modeling, Analysis and Experiment Environment,” IIH-MSP, Adelaide, SA, Australia, pp. 322-326, 2015.

[6] M. lanazi, A. Mahmood, M.J.M. Chowdhury, “Scada vulnerabilities and attacks: A review of the state-of-the-art and open issues,” Computers Security, vol. 125, 103028, 2023. https://doi.org/https://doi.org/10.1016/j.cose.2022.103028.

[7] F. Bagherzadeh, A.S. Nouri, M.J. Mehrani, S. Thennadil, “Prediction of energy consumption and evaluation of affecting factors in a full-scale WWTP using a machine learning approach,” Process Safety and Environmental Protection, vol. 154, pp. 458–466, 2021.

[8] J. Goh, S. Adepu, K.N. Junejo, A. Mathur, “A Dataset to Support Research in the Design of Secure Water Treatment Systems,” in: Havarneanu G., Setola R., Nassopoulos H., Wolthusen S. (eds) CRITIS 2016, LNCS, vol. 10242, Springer, Cham, 2017.

[9] C. Ahmed, V. Palleti, A. Mathur, “WADI: a water distribution testbed for research in the design of secure cyber physical systems,” pp. 25-28, 2017. 10.1145/3055366.3055375.

[10] H.K. Shin, W. Lee, J.H. Yun, H. Kim, “HAI 1.0: HIL-based augmented ICS security dataset,” in Proc. of the 13th USENIX Conf. on Cyber Sec. Exp. and Test, pp. 1–1, 2020.

[11] O. Tushkanova, D. Levshun, A. Branitskiy, E. Fedorchenko, E. Novikova, I. Kotenko, “Detection of Cyberattacks and Anomalies in Cyber-Physical Systems: Approaches, Data Sources, Evaluation,” Algorithms, vol. 16, no. 2, 2023. https://doi.org/10.3390/a16020085.

[12] W. Zhu, “On the model-checking-based IDS,” 2018.

[13] D. Heckerman, “A tutorial on learning with bayesian networks,” 2022. https://doi.org/10.48550/arXiv.2002.00269.

[14] R. Ranjan, G. Sahoo, “A new clutering approach for anomaly intrusion detection,” International Journal of Data Mining & Knowledge Management Process, vol. 4, no. 2, pp. 29-38, 2014.

[15] A. Branitskiy, I. Kotenko, “Hybridization of computational intelligence methods for attack detection in computer networks,” Journal of Computational Science, vol. 23, 2016.

[16] Y. Harada, Y. Yamagata, O. Mizuno, E.H. Choi, “Log-based anomaly detection of CPS using a statistical method,” in: IWESEP2017, IEEE, 2017.

[17] J.E. Zhang, D. Wu, B. Boulet, “Time series anomaly detection for smart grids: A survey, 2021.

[18] J. Inoue, Y. Yamagata, Y. Chen, C.M. Poskitt, J. Sun, “Anomaly detection for a water treatment system using unsupervised machine learning,” in: ICDMW, pp. 1058-1065, 2017.

[19] M. Elnour, N. Meskin, K.M. Khan, R. Jain, “A dual-isolation-forests-based attack detection framework for industrial control systems,” IEEE Access, no. 8, pp. 36639-36651, 2020.

[20] D. Shalyga, P. Filonov, A. Lavrentyev, “Anomaly detection for water treatment system based on neural network with automatic architecture optimization, 2018. https://doi.org/10.48550/arXiv.1807.07282.

[21] J. Audibert, F. Guyard, S. Marti, M. Zuluaga, “USAD: Unsupervised anomaly detection on multivariate time series,” In: KDD’20, pp. 3395-3404, 2020.

[22] N. Neshenko, E. Bou-Harb, B. Furht, “A behavioral-based forensic investigation approach for analyzing attacks on water plants using gans,” For. Sci. Int.: Digital Investigation, vol. 37, 301198, 2021.

[23] C. Wang, B. Wang, H. Liu, H. Qu, “Anomaly detection for industrial control system based on autoencoder neural network,” Wirel. Commun. Mob. Comput., pp. 8897926:1-8897926:10, 2020