УДК 004.896 + 004.056.5 Вестник УрФО № 3(41) / 2021, с. 38–48

**Алабугин С.К., Соколов А.Н.**

DOI: 10.14529/secur210305

# *ОБНАРУЖЕНИЕ ВТОРЖЕНИЙ В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ УПРАВЛЕНИЯ ТЕХНОЛОГИЧЕСКИМИ ПРОЦЕССАМИ С ИСПОЛЬЗОВАНИЕМ АНСАМБЛЯ МОДЕЛЕЙ РЕКУРРЕНТНОЙ И ДВУНАПРАВЛЕННОЙ ГЕНЕРАТИВНО-СОСТЯЗАТЕЛЬНОЙ НЕЙРОННЫХ СЕТЕЙ*[1]

*В работе рассмотрены генеративно-состязательные и рекуррентные архитектуры нейронных сетей, а также практика их применения для обнаружения вторжений в автоматизированных системах управления технологическими процессами. Для проведения экспериментов использован набор данных Secure Water Treatment, описывающий работу водоочистного сооружения. В ходе экспериментальных исследований на примерах, соответствующих нормальному состоянию технологического процесса, были обучены рекуррентная и двунаправленная генеративно-состязательная нейронные сети. Для улучшения метрик качества проведено ансамблирование сетей. Применение ансамбля нейронных сетей позволило улучшить точность и полноту обнаружения вторжений.*

***Ключевые слова:*** *автоматизированная система управления технологическим*

[1] Исследование выполнено при финансовой поддержке Минобрнауки России (грант ИБ) в рамках научного проекта № 16/2020.

*процессом, ансамблирование, выявление аномалий, глубокое обучение, двунаправленная генеративно-состязательная нейронная сеть, информационная безопасность, обнаружение вторжений, реккурентная нейронная сеть.*

**Alabugin S.K., Sokolov A.N.**

# *INTRUSION DETECTION IN INDUSTRIAL CONTROL SYSTEMS USING THE ENSEMBLE OF MODELS OF RECURRENT AND BIDIRECTIONAL GENERATIVE ADVERSARIAL NEURAL NETWORKS*

*The paper considers generative adversarial and recurrent neural network architectures, as well as their application for intrusion detection in industrial control systems. For the experiments, the Secure Water Treatment dataset was used. This dataset describes the operation of the wastewater treatment plant. In the course of experimental studies, on examples corresponding to the normal state of the industrial process, recurrent and bidirectional generative-adversarial neuronal networks were trained. To improve the quality metrics, both networks were ensembled. The use of an ensemble of neural networks has improved the precision and recall.*

***Keywords:*** *industrial control systems, ensemble, anomaly detection, deep learning, bidirectional generative adversarial neural network, information security, intrusion detection, recurrent neural network.*

Обнаружение вторжений является одним из наиболее приоритетных направлений исследований в области информационной безопасности (ИБ). Важность решения этой задачи обусловлена постоянным ростом количества и разнообразия угроз ИБ, реализация которых может приводить к финансовым и репутационным потерям организации, подвергшейся атаке. В случае, когда речь идёт об информационной безопасности автоматизированных систем управления технологическими процессами (АСУ ТП), к возможным последствиям успешной атаки добавляются человеческие жертвы и ущерб экологии, вплоть до техногенной катастрофы. Это обусловлено тем, что АСУ ТП часто размещаются на индустриальных объектах, от работы которых зависит качество жизни значительного числа граждан и/или экономическое благополучие отдельного региона или страны. Примерами успешных атак на АСУ ТП являются сетевой червь Stuxnet [1], использованный для совершения диверсии на предприятиях ядерной промышленности Ирана в 2010, вредоносное программное обеспечение BlackEnergy [2], использованное для атаки на энергетический сектор Украины в 2015, атака на предприятия энергетической промышленности Израиля в 2016 [3], атаки на объекты водоснабжения и водоочистки Израиля в 2020 [4] и др.

В течение длительного времени такие системы были изолированы от внешних сетей, и, в особенности от сети Интернет. В настоящее время наблюдается тенденция к объединению индустриальных и корпоративных сетей [5]. Это позволяет более эффективно экс-

плуатировать АСУ ТП легитимным пользователям, при этом делает подобные системы более уязвимыми для действий злоумышленников. Технологии, используемые в АСУ ТП, часто разработаны без учета требований информационной безопасности, – как правило, к ним относятся проприетарные средства [6]. Для отрасли в целом характерен низкий уровень культуры информационной безопасности [7], в частности, это касается оперативной разработки обновлений безопасности производителями и своевременной установки этих обновлений на объекте. Вышеперечисленные факторы характеризуют причины, по которым АСУ ТП в целом уязвимы для злоумышленника и обосновывают необходимость повысить уровень защищенности АСУ ТП. Для своевременного обнаружения вторжений, в частности, для обнаружения таргетированных атак, предлагается использовать метод обнаружения вторжений, основанный на выявлении аномалий.

В качестве данных для анализа использованы данные состояния технологического процесса (значения сигналов сенсоров и актуаторов в конкретный момент времени). Исходя из предположения, что большая часть аномалий технологического процесса является следствием действий злоумышленника, выявленная аномалия технологического процесса позволяет установить факт вторжения. Нужно отметить, однако, что при такой постановке задачи, в случае, если аномалия является следствием программной ошибки, сбоя технического оборудования или действий оператора – она будет свидетельствовать о ложном вторжении.

При использовании методов машинного обучения для анализа состояния технологического процесса можно выделить несколько подходов. В частности, для выявления аномалий технологического процесса применимы методы классификации и кластеризации. В [8] представлены результаты работы некоторых классических алгоритмов машинного обучения (K-means, Naive Bayesian, GMM, PCA-SVD) на наборе данных Gas Pipeline. В [9, 10] проведен анализ практической применимости классических методов машинного обучения (линейная регрессия, решающие деревья, SVM) и нейронной сети классификатора для выявления аномалий технологического процесса. В [11] безопасность сетевой инфраструктуры АСУ ТП обеспечивается с помощью полносвязной нейронной сети и автокодировщика для выявления атак в сетевом трафике. В [12] нейронная сеть использована для выявления атак типа False Data Injection на основе анализа показаний сенсоров АСУ ТП. В [13] рассмотрено применение метода оптимизации Particle Swarm при обучении нейронной сети, обнаруживающей атаки. Предложенный метод позволяет обучать сеть за меньшее количество итераций и обеспечивает более высокую точность. В [14] авторы предложили систему, состоящую из нейронной сети и Lyapunov-based model predictive controller (LMPC). Нейронная сеть анализирует показания сенсоров химического процесса и производит обнаружение факта атаки. В случае атаки LMPC используется для смягчения деструктивного влияния атаки и рестабилизации системы. В [15] для выявления аномалий авторы использованы полносвязные нейронные сети, построенные при помощи генетических алгоритмов (Evolutionary based Neural Networks), в частности, они предложили применять для оптимизации весов сети алгоритм Grey Wolf Optimizer с целью увеличения скорости обучения сети. В [16] авторы применили алгоритмы Random Forest и Support Vector Machine для выявления аномалий, а также исследовали способы обработки пропусков в данных и нормализации данных для улучшения качества работы алгоритмов. В [17] авторы предложили использовать для выявления аномалий новый подход в рамках одноклассовой классификации, основанный на импульсных нейронных сетях с целью получения практически применимого алгоритма, который не нуждается в данных об аномальном состоянии АСУ ТП.

Использование нейронных сетей-классификаторов имеет несколько недостатков: на стадии обучения требуется достаточное количество примеров, репрезентирующих вторжения (как правило, получить и корректно разметить такие примеры – отдельная трудоёмкая задача). Кроме того, сеть-классификатор, при обработке принципиально новых данных, соответствующих атаке, не представленной в обучающей выборке, вполне может отработать неправильно. Методы, использующие обучение без учителя, в частности, кластеризация, зачастую не позволяют получить результат, который был бы применим на практике. Кроме того, результаты работы методов классификации и кластеризации часто не интерпретируемы: получив сообщение о выявленной аномалии, пользова-

тель не сможет определить причину ее возникновения. С целью преодоления описанных недостатков рассмотрены модели на основе двух архитектур искусственных нейронных сетей:

1. Рекуррентные нейронные сети, позволяющие анализировать последовательности данных и предсказывать их дальнейшую динамику.

2. Генеративно-состязательные сети, конструирующие из сырых данных внутреннее представление и выделяющие наиболее характерные признаки.

Для проведения экспериментальных исследований использован набор данных Secure Water Treatment (SWaT) [18], разработанный исследователями из Singapore University of Technology and Design с целью использования для создания и оценки механизмов защиты киберфизических систем. При его разработке был построен испытательный стенд, который представляет собой уменьшенную полнофункциональную копию реального водоочистного сооружения.

Стенд Secure Water Treatment включает 6 стадий технологического процесса, который реализован в водоочистном сооружении: забор воды (P1), оценка качества неочищенной воды (P2), механическая фильтрация воды (P3), дехлоризация (P4), обратный осмос (P5), перегонка очищенной воды в хранилище или на ещё один цикл очистки (P6).

Данные содержат дамп сетевого трафика SWaT и данные, описывающие показания 25 сенсоров и 26 актуаторов стенда. Набор данных содержит 964722 записи, которые разделены на две части:

• данные, собранные в течение 7 дней нормальной работы стенда, без каких-либо сбоев (первая часть);

• данные, собранные за 4 дня, во время которых проводились атаки (вторая часть).

Рекуррентные нейронные сети (recurrent neural networks, RNN) – это разновидность архитектур нейронных сетей, которые обладают обратной (рекуррентной) связью. Для анализа последовательностей использованы рекуррентные нейронные сети, так как благодаря наличию обратной связи, они могут запоминать внутреннее состояние, то есть помнить данные, которые были представлены в последовательности. С помощью рекуррентной нейронной сети прогнозируется состояние технологического процесса. Из существующих архитектур рекуррентных сетей была выбрана ячейка Long Short-Term Memory (LSTM) [19]. В модели также использован одномерный сверточный слой, позволяющий извлекать из сырого состояния технологического процесса признаки, которые подаются на вход LSTM слоям. Структура сети, используемой в работе, представлена на рис. 1.

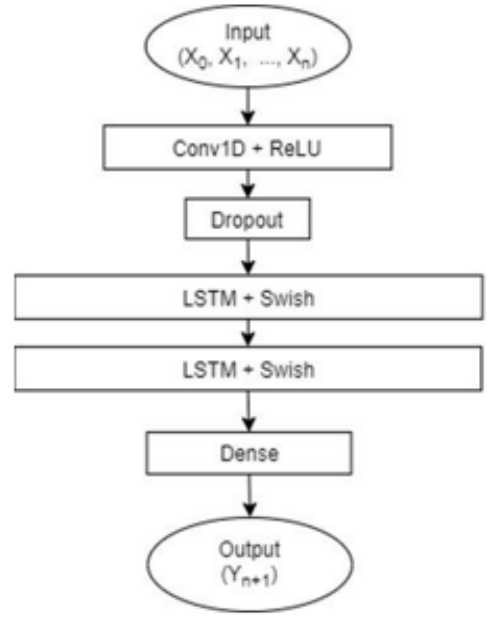


Рис. 1. Структура рекуррентной нейронной сети, используемой для выявления аномалий технологического процесса

Сеть состоит из одномерной свертки с функцией активации ReLU (Conv1D + ReLU), Dropout-слоя, который используется, чтобы избежать переобучения, двух слоев LSTM и полносвязного слоя (Dense) на выходе. В качестве функции активации LSTM-слоев была использована функция Swish [20].

Сеть принимает на вход последовательность из состояний технологического процесса SWaT, описываемого 51 признаком, в моменты времени $X_0$, $X_1$, …, $X_n$, и выдает в качестве результата работы состояние технологического процесса в следующий момент времени $Y_{n+1}$.

С помощью библиотеки машинного обучения Tensorflow реализована рекуррентная нейронная сеть. Для обучения и тестирования реализованной модели использованы данные из набора SWaT. Обучающая выборка включала данные, соответствующие нормальной работе системы. Контрольная выборка состояла из данных, полученных в период проведения атак. Данные были нормализованы и разбиты на последовательности, описывающие состояние системы в течение

30 секунд. Таким образом, на вход модели подавались многомерные тензоры, имеющие форму [M,30,51], где M – размер батча. Для обучения модели в качестве функции потерь использовалась средняя квадратичная ошибка (mean square error, MSE), а в качестве алгоритма градиентного спуска – алгоритм Adam [21]. Модель обучалась в течение 200 эпох и достигла значения MSE равного 2.7301. Показатель средней абсолютной ошибки (mean absolute error, MAE) при этом составил 0.7267.

После этого к контрольной выборке применена обученная модель. Разница вычислялась между предсказанным и фактическим состоянием технологического процесса. В качестве метрики аномальности использована максимальная разница среди всех признаков.

Для метрики аномальности был эмпирически подобран подходящий порог, обеспечивающий наилучшие результаты работы. ROC-кривая, с помощью которой был подобран порог для метрики аномальности, представлена на рис. 2.

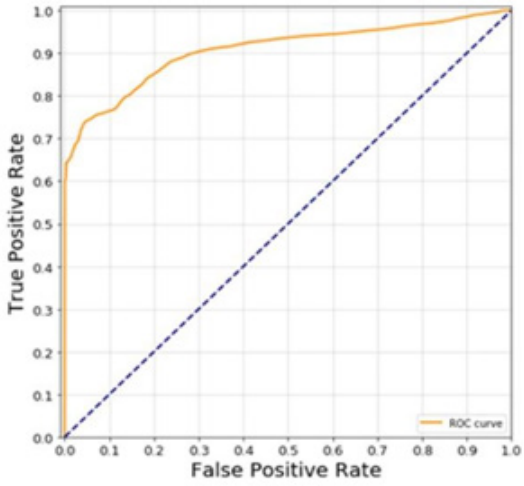


Рис.2. ROC-кривая метода выявления аномалий, основанного на применении рекуррентной сети

Генеративно-состязательные сети (Generative Adversarial Net, GAN) – это класс нейронных сетей, используемый для генерации новых синтетических данных на основе реальных. Концептуально генеративно-состязательные сети основаны на идее состязательного обучения и впервые были описаны в [22].

Архитектура GAN включает две модели: генератор G, порождающий на основе вектора шума новые, похожие на настоящие, объекты в пространстве данных, и дискриминатор D, целью которого является отличить порождаемые генератором объекты от реальных данных, – поэтому архитектура и называется генеративно-состязательной. Для применения генеративно-состязательных сетей в задаче обнаружения аномалий использована двунаправленная генеративно-состязательная сеть (Bidirectional Generative Adversarial Net, BiGAN) [17], схема которой представлена на рис. 3.

По сравнению с базовой архитектурой генеративно-состязательной сети, в архитектуре BiGAN дополнительно задействована сеть-кодировщик (encoder) E, что добавляет структуру автокодировщика (автокодировщиком является пара генератор – кодировщик) в очную генеративно-состязательную сеть.

Автокодировщик (autoencoder) [23] — это архитектура искусственной нейронной сети, позволяющая применять обучение без учителя при использовании метода обратного распространения ошибки. Наиболее простая архитектура автокодировщика — сеть прямого распространения, без наличия рекуррентных связей и содержащая входной, промежуточный и выходной слои. В отличие от перцептрона, выходной слой автокодировщика должен содержать столько же нейронов, сколько и входной слой.

Основной принцип работы и обучения сети автокодировщика — получить на выходе вектор, как можно более близкий к входному. Для того, чтобы решение не было тривиальным (идентичное преобразование вектора), на промежуточный слой сети накладываются ограничения некоторые ограничения, в частности: промежуточный слой должен быть или меньшей размерности, чем входной и выходной слои, или искусственно ограничивается количество одновременно активных нейронов промежуточного слоя (применяется разрежённая активация). Эти ограничения заставляют кодировщик искать обобщения и корреляцию в поступающих во входных данных, а также выполнять их сжатие. Таким образом, сеть обучается выделять из входных данных общие признаки, которые кодируются в значениях весов искусственной нейронной сети. Так, после обучения сети на наборе различных входных изображений, кодировщик может научиться распознавать отдельные линии и полосы под различными углами.

В архитектуре BiGAN, кодировщик E осуществляет преобразование из пространства реальных данных в пространство скрытых

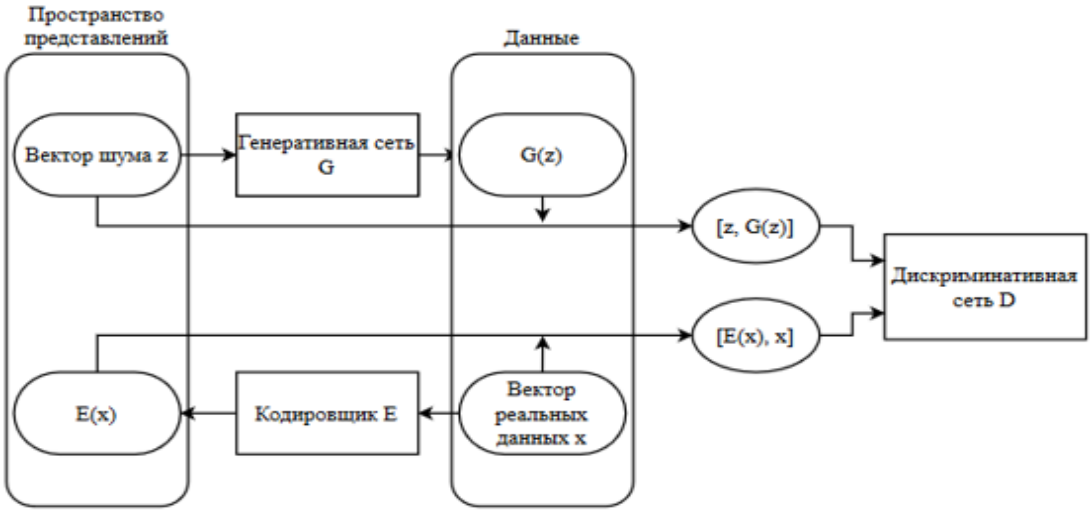


Рис.3. Схема двунаправленной генеративно-состязательной нейронной сети (BiGAN)

переменных, из которого берётся вектор шума для генератора. Формально это действие можно представить в виде функции:

$$E = E(x ; \theta_e) : X \longrightarrow Z, \quad (1)$$

где $\theta_e$ – параметры сети-кодировщика.

Таким образом, сеть-кодировщик в процессе обучения учиться делать преобразование, обратное генератору. Кроме того, в архитектуре BiGAN, сеть-дискриминатор обучается различать не только порожденные генератором объекты и объекты из реальных данных, но и векторы из пространства скрытых переменных z (используемые для порождения объектов) и результат отображения кодировщика E(x). Формальное доказательство того, что в описанной схеме кодировщик после обучения сети BiGAN, реализует функцию, обратную функции генератора, приведено в [23]. Процесс обучения сети BiGAN аналогичен процессу обучения обычной генеративно-состязательной сети: веса каждой сети обновляются попеременно.

Добавления сети-кодировщика в архитектуру генеративно-состязательной сети, который каждому объекту из пространства данных x ставит в соответствие вектор z из пространства скрытых переменных, позволяет непосредственно извлекать представления (representations) этих объектов. С точки зрения семантики, вектор z репрезентирует признаки объекта x, выявленные в процессе обучения BiGAN. Именно это свойство позволяет использовать BiGAN для выявления аномалий [24].

Один из подходов к использованию BiGAN для обнаружения аномалий состоит в построении метрики A аномальности объекта, основанной на выпуклой комбинации (convex combiantion) функции потерь реконструкции и функции потерь дискриминатора

$$A(x) = aL_G(x) + (1\text{-}a)L_D(x), \quad (2)$$

где $L_G(x)$ – функция потерь реконструкции, определяемая как модуль разности исходного вектора, репрезентирующего объект, и вектора, полученного последовательным преобразованием объекта кодировщиком и генератором:

$$L_G(x) = ||x - G(E(x))||. \quad (3)$$

$L_D(x)$ – функция потерь дискриминатора:

$$LD(x) = \sigma(D(x,(E(x)),1), \quad (4)$$

где σ – кросс-энтропия (логарифмическая функция потерь) дискриминатора, при условии, что объект x является настоящим, а не порожден генератором.

Таким образом, сеть BiGAN, обученная на данных, соответствующих нормальному состоянию, применяется к новому объекту, после чего вычисляется значение метрики аномальности A для объекта. Чем выше значение этой метрики, тем более вероятно, что объект является аномальным.

Для обучения и тестирования сети BiGAN были также использованы данные из набора SWaT. Обучающая выборка включает в себя данные соответствующие нормальной работе системы. Контрольная выборка состоит из данных, полученных в период проведения атак. Данные были приведены к одному масштабу, после чего BiGAN обучалась исключительно на данных соответствующих нормальной работе системы. После этого к контрольной выборке применена обученная модель, а для каждого объекта контрольной выборки построена метрика аномальности. Качество

работы обученной модели определяется не только ее возможностями аппроксимировать распределение данных, соответствующих нормальной работе системы, но и конкретным значением порога метрики аномальности. На рис. 4 представлена зависимость метрик точности и полноты при изменении порогового значения. Из рисунка видно, что при увеличении порогового значения увеличивается точность и уменьшается полнота. Это свойство позволяет настроить желаемое поведение модели в каждом конкретном случае.

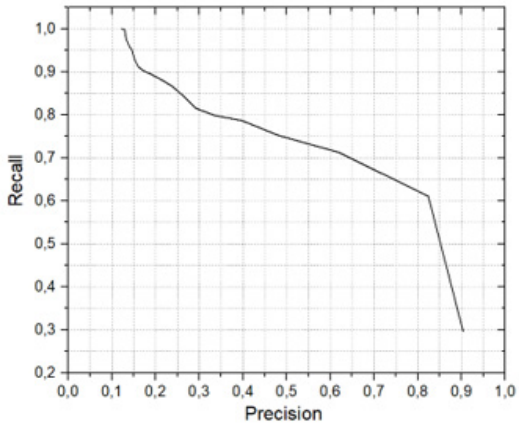


Рис. 4. Зависимость метрик точности (Precision) и полноты (Recall) при из-менении порогового значения для метода выявления аномалий технологического процесса с помощью сети BiGAN

С целью повышения результативности выявления аномалий разработана модель машинного обучения, включающая генеративно-состязательную и рекуррентную архитектуры с использованием ансамблирования моделей [25]. При таком подходе состояние технологического процесса анализируется как с помощью рекуррентной сети (прогнозирования состояния), так и с помощью двунаправленной генеративно-состязательной сети BiGAN (оценка аномальности состояния в моменте). Такой подход, как правило, используется при решении сложных задач, когда ни один из применяемых алгоритмов не показывает желаемого уровня точности. Смысл применения голосования состоит в том, чтобы взаимно компенсировать ошибки, допущенные каждой моделью. Эффективность такого подхода, как правило, определяется природой решаемой задачи, используемыми признаками и алгоритмами.

Основным решателем предложенной модели является рекуррентная сеть. Для обеспечения высокого уровня точности (Precision) для генеративно-состязательной сети архитектуры BiGAN выбран достаточно высокий порог по метрике аномальности. В предложенной ансамблевой модели оценкам генеративно-состязательной сети корректируют решения, полученные путем прогнозирования рекуррентной сети: состояние технологического процесса признается аномальным только в том случае, когда оценки, полученные с помощи обеих моделей, совпадают. Схема применения ансамбля нейронных сетей для обнаружения вторжений в АСУ ТП на основе обнаружения аномалий технологического процесса приведена на рис. 5.


Рис. 5. Схема ансамблирования моделей нейронных сетей для обнаружения вторже-ний в АСУ ТП

В таблице представлены результаты применения ансамбля моделей для набора данных SWaT в сравнении с результатами, полученными при использовании только рекуррентной сети и другими методами.

Таким образом, предложенная модель достаточно результативно справляется с выявлением аномалий технологического процесса, порожденных вторжениями злоумышленника, и может быть использована при разработке систем обнаружения вторжений. Достоинствами метода на основе ансамбля нейронных сетей являются его быстродействие и

*Таблица*

**Результаты работы методов для набора данных SWaT**

| Метод | Precision | Recall | F1-Score |
|---|---|---|---|
| 1D CNN [26] | 0.968 | 0.791 | 0.871 |
| MLP [27] | 0.967 | 0.696 | 0.812 |
| CNN [27] | 0.952 | 0.702 | 0.808 |
| RNN [27] | 0.936 | 0.692 | 0.796 |
| DNN [28] | 0.982 | 0.678 | 0.802 |
| OCSVM [28] | 0.925 | 0.699 | 0.796 |
| Метод на основе прогнозирования состояния технологического процесса с использованием рекуррентной сети | 0.934 | 0.820 | 0.865 |
| Метод на основе ансамбля нейронных сетей | 0.981 | 0.890 | 0.933 |

отсутствие необходимости использовать данные, описывающие аномальное состояние технологического процесса. В случае применения ансамблевой модели на наборе данных SWaT на компьютере с процессором Intel Core i7-9750H и видеокартой NVIDIA GeForce RTX 2070 среднее время одной итерации составило 0.65 секунды.

**Литература**


1. Kushner D. The real story of stuxnet //ieee Spectrum. – 2013. – Т. 50. – №. 3. – С. 48-53.

2. Khan R. et al. Threat analysis of blackenergy malware for synchrophasor based real-time control and monitoring in smart grid //4th International Symposium for ICS & SCADA Cyber Security Research 2016 4. – 2016. – С. 53-63.

3. Li, Zhong-wei, Weiming Tong, and Xianji Jin. "Construction of cyber security defense hierarchy and cyber security testing system of smart grid: thinking and enlightenment for network attack events to national power grid of Ukraine and Israel." Automation of Electric Power Systems 40.8 (2016): 147-151.

4. Israel Government Tells Water Treatment Companies to Change Passwords [Электронный ресурс]. — URL: https://www.zdnet.com/article/israel-says-hackers-are-targeting-its-water-supply-and-treatment-utilities/ (visited on 6/11/2020).

5. Xu H. et al. A survey on industrial Internet of Things: A cyber-physical systems perspective //IEEE Access. – 2018. – Т. 6. – С. 78238-78259.

6. Баринов А. Е., Скурлаев С. В., Соколов А. Н. Методика оценки рисков, вызванных уязвимостями в программном обеспечении автоматизированных систем управления технологическими процессами //Вестник УрФО. Безопасность в информационной сфере. – 2017. – №. 3 (25). – С. 34-42.

7. Luiijf E. Cyber (in-) security of industrial control systems: A societal challenge //International Conference on Computer Safety, Reliability, and Security. – Springer, Cham, 2014. – С. 7-15.

8. Shirazi S. N. et al. Evaluation of anomaly detection techniques for scada communication resilience //2016 Resilience Week (RWS). – IEEE, 2016. – С. 140-145.

9. Sokolov A. N., Pyatnitsky I. A., Alabugin S. K. Research of classical machine learning methods and deep learning models effectiveness in detecting anomalies of industrial control system //2018 Global Smart Industry Conference (GloSIC). – IEEE, 2018. – С. 1-6.

10. Sokolov A. N., Pyatnitsky I. A., Alabugin S. K. Applying methods of machine learning in the task of intrusion detection based on the analysis of industrial process state and ICS networking //FME Transactions. – 2019. – Т. 47. – №. 4. – С. 782-789.

11. Muna A. L. H., Moustafa N., Sitnikova E. Identification of malicious activities in industrial internet of things based on deep learning models //Journal of information security and applications. – 2018. – Т. 41. – С. 1-11.

12. Potluri S., Diedrich C., Sangala G. K. R. Identifying false data injection attacks in industrial control systems using artificial neural networks //2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). – IEEE, 2017. – С. 1-8.

13. Yang H. et al. Research on intrusion detection of industrial control system based on OPSO-BPNN algorithm //2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). – IEEE, 2017. – С. 957-961.

14. Wu Z. et al. Detecting and handling cyber-attacks in model predictive control of chemical processes //Mathematics. – 2018. – Т. 6. – №. 10. – С. 173.

15. Davahli A., Shamsi M., Abaei G. Hybridizing genetic algorithm and grey wolf optimizer to advance an intelligent and lightweight intrusion detection system for IoT wireless networks //Journal of Ambient Intelligence and Humanized Computing. – 2020. – Т. 11. – №. 11. – С. 5581-5609.

16. Perez R. L. et al. Machine learning for reliable network attack detection in SCADA systems //2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE). – IEEE, 2018. – С. 633-638.

17. Demertzis K., Iliadis L., Spartalis S. A spiking one-class anomaly detection framework for cyber-security on industrial control systems //International Conference on Engineering Applications of Neural Networks. – Springer, Cham, 2017. – С. 122-134.

18. Mathur A. P., Tippenhauer N. O. SWaT: a water treatment testbed for research and training on ICS security //2016 International Workshop on Cyber-physical Systems for Smart Water Networks (CySWater). – IEEE, 2016. – С. 31-36.

19. Hochreiter S., Schmidhuber J. Long short-term memory //Neural computation. – 1997. – Т. 9. – №. 8. – С. 1735-1780.

20. Ramachandran P., Zoph B., Le Q. V. Searching for activation functions //arXiv preprint arXiv:1710.05941. – 2017.

21. Kingma D. P., Ba J. Adam: A method for stochastic optimization //arXiv preprint arXiv:1412.6980. – 2014.

22. Goodfellow I. et al. Generative adversarial nets //Advances in neural information processing systems. – 2014. – С. 2672-2680.

23. Donahue J., Krähenbühl P., Darrell T. Adversarial feature learning //arXiv preprint arXiv:1605.09782. – 2016.

24. Baldi P. Autoencoders, unsupervised learning, and deep architectures //Proceedings of ICML workshop on unsupervised and transfer learning. – 2012. – С. 37-49.

25. Zenati H. et al. Efficient gan-based anomaly detection //arXiv preprint arXiv:1802.06222. – 2018.

26. Lueckenga J., Engel D., Green R. Weighted vote algorithm combination technique for anomaly based Smart Grid Intrusion Detection systems //2016 International Joint Conference on Neural Networks (IJCNN). – IEEE, 2016. – С. 2738-2742.

27. Kravchik M., Shabtai A. Detecting cyber attacks in industrial control systems using convolutional neural networks //Proceedings of the 2018 Workshop on Cyber-Physical Systems Security and PrivaCy. – 2018. – С. 72-83.

28. Shalyga D., Filonov P., Lavrentyev A. Anomaly detection for water treatment system based on neural network with automatic architecture optimization //arXiv preprint arXiv:1807.07282. – 2018.

29. Inoue J. et al. Anomaly detection for a water treatment system using unsupervised machine learning //2017 IEEE International Conference on Data Mining Workshops (ICDMW). – IEEE, 2017. – С. 1058-1065.

## References

1. Kushner D. The real story of stuxnet //ieee Spectrum. – 2013. – Т. 50. – №. 3. – С. 48-53.

2. Khan R. et al. Threat analysis of blackenergy malware for synchrophasor based real-time control and monitoring in smart grid //4th International Symposium for ICS & SCADA Cyber Security Research 2016 4. – 2016. – С. 53-63.

3. Li, Zhong-wei, Weiming Tong, and Xianji Jin. "Construction of cyber security defense hierarchy and cyber security testing system of smart grid: thinking and enlightenment for network attack events to national power grid of Ukraine and Israel." Automation of Electric Power Systems 40.8 (2016): 147-151.

4. Israel Government Tells Water Treatment Companies to Change Passwords [Электронный ресурс]. — URL: https://www.zdnet.com/article/israel-says-hackers-are-targeting-its-water-supply-and-treatment-utilities/ (visited on 6/11/2020).

5. Xu H. et al. A survey on industrial Internet of Things: A cyber-physical systems perspective //IEEE Access. – 2018. – Т. 6. – С. 78238-78259.

6. Barinov A. E., Skurlaev S. V., Sokolov A. N. Metodika otsenki riskov, vyzvannykh uyazvimostyami v programmnom obespechenii avtomatizirovannykh sistem upravleniya tekhnologicheskimi protsessami // Vestnik UrFO. Bezopasnost' v informatsionnoy sfere. – 2017. – no. 3. – pp. 34-42. Luiijf E. Cyber (in-) security of industrial control systems: A societal challenge //International Conference on Computer Safety, Reliability, and Security. – Springer, Cham, 2014. – С. 7-15.

7. Luiijf E. Cyber (in-) security of industrial control systems: A societal challenge //International Conference on Computer Safety, Reliability, and Security. – Springer, Cham, 2014. – С. 7-15.

8. Shirazi S. N. et al. Evaluation of anomaly detection techniques for scada communication resilience //2016 Resilience Week (RWS). – IEEE, 2016. – С. 140-145.

9. Sokolov A. N., Pyatnitsky I. A., Alabugin S. K. Research of classical machine learning methods and deep learning models effectiveness in detecting anomalies of industrial control system //2018 Global Smart Industry Conference (GloSIC). – IEEE, 2018. – С. 1-6.

10. Sokolov A. N., Pyatnitsky I. A., Alabugin S. K. Applying methods of machine learning in the task of intrusion detection based on the analysis of industrial process state and ICS networking //FME Transactions. – 2019. – Т. 47. – №. 4. – С. 782-789.

11. Muna A. L. H., Moustafa N., Sitnikova E. Identification of malicious activities in industrial internet of things based on deep learning models //Journal of information security and applications. – 2018. – Т. 41. – С. 1-11.

12. Potluri S., Diedrich C., Sangala G. K. R. Identifying false data injection attacks in industrial control systems using artificial neural networks //2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). – IEEE, 2017. – С. 1-8.

13. Yang H. et al. Research on intrusion detection of industrial control system based on OPSO-BPNN algorithm //2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). – IEEE, 2017. – С. 957-961.

14. Wu Z. et al. Detecting and handling cyber-attacks in model predictive control of chemical processes //Mathematics. – 2018. – Т. 6. – №. 10. – С. 173.

15. Davahli A., Shamsi M., Abaei G. Hybridizing genetic algorithm and grey wolf optimizer to advance an intelligent and lightweight intrusion detection system for IoT wireless networks //Journal of Ambient Intelligence and Humanized Computing. – 2020. – Т. 11. – №. 11. – С. 5581-5609.

16. Perez R. L. et al. Machine learning for reliable network attack detection in SCADA systems //2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE). – IEEE, 2018. – С. 633-638.

17. Demertzis K., Iliadis L., Spartalis S. A spiking one-class anomaly detection framework for cyber-security on industrial control systems //International Conference on Engineering Applications of Neural Networks. – Springer, Cham, 2017. – С. 122-134.

18. Mathur A. P., Tippenhauer N. O. SWaT: a water treatment testbed for research and training on ICS security //2016 International Workshop on Cyber-physical Systems for Smart Water Networks (CySWater). – IEEE, 2016. – С. 31-36.

19. Hochreiter S., Schmidhuber J. Long short-term memory //Neural computation. – 1997. – Т. 9. – №. 8. – С. 1735-1780.

20. Ramachandran P., Zoph B., Le Q. V. Searching for activation functions //arXiv preprint arXiv:1710.05941. – 2017.

21. Kingma D. P., Ba J. Adam: A method for stochastic optimization //arXiv preprint arXiv:1412.6980. – 2014.

22. Goodfellow I. et al. Generative adversarial nets //Advances in neural information processing systems. – 2014. – С. 2672-2680.

23. Donahue J., Krähenbühl P., Darrell T. Adversarial feature learning //arXiv preprint arXiv:1605.09782. – 2016.

24. Baldi P. Autoencoders, unsupervised learning, and deep architectures //Proceedings of ICML workshop on unsupervised and transfer learning. – 2012. – С. 37-49.

25. Zenati H. et al. Efficient gan-based anomaly detection //arXiv preprint arXiv:1802.06222. – 2018.

26. Lueckenga J., Engel D., Green R. Weighted vote algorithm combination technique for anomaly based Smart Grid Intrusion Detection systems //2016 International Joint Conference on Neural Networks (IJCNN). – IEEE, 2016. – С. 2738-2742.

27. Kravchik M., Shabtai A. Detecting cyber attacks in industrial control systems using convolutional neural networks //Proceedings of the 2018 Workshop on Cyber-Physical Systems Security and PrivaCy. – 2018. – С. 72-83.

28. Shalyga D., Filonov P., Lavrentyev A. Anomaly detection for water treatment system based on neural network with automatic architecture optimization //arXiv preprint arXiv:1807.07282. – 2018.

29. Inoue J. et al. Anomaly detection for a water treatment system using unsupervised machine learning //2017 IEEE International Conference on Data Mining Workshops (ICDMW). – IEEE, 2017. – С. 1058-1065.

**АЛАБУГИН Сергей Константинович,** инженер кафедры защиты информации высшей школы электроники и компьютерных наук ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». Россия, 454080, г. Челябинск, проспект Ленина, д. 76. E-mail: sergei_alabugin@mail.ru.

**СОКОЛОВ Александр Николаевич**, кандидат технических наук, доцент, заведующий кафедрой защиты информации высшей школы электроники и компьютерных наук ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». Россия, 454080, г. Челябинск, проспект Ленина, д. 76. E-mail: sokolovan@susu.ru.

**ALABUGIN Sergei,** engineer of the department of information security of the school of electrical engineering and computer science in FSAEI HE «South Ural State University (national research university)». 76, Lenin prospect, Chelyabinsk, Russia, 454080. E-mail: sergei_alabugin@mail.ru.

**SOKOLOV Alexander**, Ph.D., Associate professor, Head of the department of information security of the school of electrical engineering and computer science in FSAEI HE «South Ural State University (national research university)». 76, Lenin prospect, Chelyabinsk, Russia, 454080. E-mail: sokolovan@susu.ru.