УДК 004.896 + 004.056.57 Вестник УрФО № 3(41) / 2021, с. 65–87

Соколов А.Н., Рагозин А.Н., Баринов А.Е., Уфимцев М.С., Пятницкий И.А., Бухарев Д.А.

DOI: 10.14529/secur210308

РАЗРАБОТКА МОДЕЛЕЙ И МЕТОДОВ РАННЕГО ОБНАРУЖЕНИЯ КИБЕРАТАК НА ОБЪЕКТЫ ЭНЕРГЕТИКИ МЕТАЛЛУРГИЧЕСКОГО ПРЕДПРИЯТИЯ[1]

Необходимость реагирования на киберинциденты в энергетических комплексах металлургических предприятий приводит к серьезным последствиям, таким как простои и дорогостоящие перезапуски производства. Поэтому актуальной является задача реализации концепции защиты информации «предотвращать» вторжения, а не «предупреждать» о них, согласно которой поиск киберугроз становится постоянной фоновой деятельностью. В работе представлена система раннего обнаружения воздействия кибератак, разработанная с использованием технологий интеллектуальной обработки данных на основе модели системы динамических процессов, отражающих функционирование (поведение) информационной системы объекта. Разработана методика, позволяющая получать численную оценку необходимости улучшения параметров мониторинга для различных зон технологического процесса. В качестве примера использования методики приведен расчет на основе подсистемы технологического процесса металлургического предприятия, на основании которого выбраны наиболее важные информативные точки сбора данных для реализации мониторинга. Показано, что применение аппарата искусственных нейронных сетей (ИНС) автокодировщиков, генеративно-состязательных и рекуррентных ИНС полностью соответствует концепции построения диагностической поведенческой модели технологического процесса. Представленные результаты исследования моделей показали их высокую эффективность, поскольку позволяют не только обнаруживать аномалии в корпоративной сети, вызванные кибератаками, но и реализовать их локализацию. Персонал предприятия получает возможность принимать превентивные меры с целью устранения угроз информационной безопасности на производстве. В результате использования предложенных подходов ожидается значительное снижение потенциального ущерба, который могут понести металлургические предприятия в результате реализации киберинцидентов.

Ключевые слова: *автокодировщик, автоматизированная система управления технологическим процессом, генеративно-состязательная нейронная сеть, динамический процесс, зона технологического процесса, кибератака, кибервторжение, киберинцидент, киберугроза, машинное обучение, металлургическое предприятие, мони-*

[1] Исследование выполнено при финансовой поддержке РФФИ и Челябинской области в рамках научного проекта № 20-47-740006

торинг информационной системы, обнаружение аномалий, поведенческая модель, раннее обнаружение кибератак, рекуррентная нейронная сеть, энергосетевая система.

Sokolov A.N., Ragozin A.N., Barinov A.E., Ufimtcev M.S., Pyatnitskiy I.A., Bukharev D.A.

DEVELOPMENT OF MODELS AND METHODS FOR EARLY DETECTION OF CYBER ATTACKS ON ENERGY FACILITIES OF A METALLURGICAL ENTERPRISE

The need to respond to cyber incidents in the energy systems of metallurgical enterprises leads to serious consequences, such as downtime and costly restarts of production. Therefore, it is important to implement the concept of information security to "prevent" intrusions, and not "warn" about them, according to which the search for cyber threats becomes a constant background activity. This paper presents a system for the early detection of the impact of cyberattacks, developed using intelligent data processing technologies based on a model of a system of dynamic processes that reflect the behavior of an object's information system. A technique has been developed that makes it possible to obtain a numerical assessment of the need to improve the monitoring parameters for various zones of the technological process. As an example of using the methodology, a calculation based on a subsystem of the technological process of a metallurgical enterprise is given, on the basis of which the most important informative data collection points for monitoring are selected. It is shown that the use of the apparatus of artificial neural networks (ANNs) of autoencoders, generative adversarial and recurrent ANNs fully corresponds to the concept of constructing a diagnostic behavioral model of a technological process. The presented results of the study of the models showed their high efficiency, since they allow not only detecting anomalies in the corporate network caused by cyber attacks, but also their localization. The staff of the enterprise gets the opportunity to take preventive measures in order to eliminate threats to information security in production. As a result of using the proposed approaches, a significant reduction in the potential damage that can be incurred by metallurgical enterprises as a result of the implementation of cyber incidents is expected.

Keywords: *autoencoder, industrial control systems, generative adversarial neural network, dynamic process, process zone, cyber attack, cyber intrusion, cyber incident, cyber threat, machine learning, metallurgical enterprise, information system monitoring, anomaly detection, behavioral model, early detection of cyber attacks, recurrent neural network, power grid system.*

Введение.

Энергетический комплекс для металлургического производства – чрезвычайно важный элемент, от которого зависит функционирование как отдельных промышленных объектов, так и обеспечение энергоносителями

инфраструктуры, находящейся поблизости. Самая простая энергосетевая система состоит из немалого количества резервных линий, обслуживанием которых занимается большой объем оборудования, в состав которого входят коммутационные и распределительные устройства, силовые и измерительные приборы, генераторное оборудование, автоматические компенсаторы реактивной мощности и т.д. Любое из перечисленных устройств подключается к промышленной сети предприятия и управляется сложной системой, большинство действий в которой автоматизировано.

Энергетические системы, как правило, являются открытыми системами, так как взаимодействуют с другими энергосистемами, как в части учёта потреблённой электроэнергии, так и в части взаимного мониторинга нагрузки с целью её оптимального перераспределения. Поэтому площадь поверхности кибератаки на энергосетевые системы закономерно большая, и, в связи с этим, возникает необходимость мониторинга таких систем на предмет возможных кибератак. Помимо промышленного оборудования, подобные энергосетевые системы могут обеспечивать энергоснабжение ближайших промышленных объектов, жилых комплексов и социально значимых объектов. Реализация кибератаки занимает несколько минут, при этом средний показатель её идентификации составляет более ста дней. Необходимость реагирования на произошедшие инциденты в подобных энергетических комплексах приводит к серьезным последствиям, таким как простои и дорогостоящие перезапуски производства. В связи с этим для металлургических предприятий актуальной является задача реализации концепции защиты информации – «предотвращать» вторжения, а не «предупреждать» о них, согласно которой поиск киберугроз становится постоянной фоновой деятельностью. Поэтому задача разработки моделей и методов раннего обнаружения кибератак на объекты энергетики металлургического предприятия является важной.

Под «ранним» обнаружением кибератаки понимается её обнаружение в рассматриваемой системе до реализации последствий кибератаки, т.е. до необходимости реагирования на инцидент путем остановки и перезапуска системы. В предложенной работе представлена система раннего обнаружения воздействия кибератак, разработанная с использованием технологий интеллектуальной обработки данных на основе модели системы динамических процессов, отражающих функционирование (поведение) информационной системы объекта. Такая система позволяет обнаруживать ранние признаки изменения в поведении наблюдаемых динамических процессов объекта, то есть обнаруживать начало раннего развития аномалий в динамических процессах, распределённых в виде потока данных в информационной системе объекта.

В работе предложен подход, основанный на применении ИНС автокодировщиков, генеративно-состязательных и рекуррентных ИНС для анализа состояния всей автоматизированной системы управления объектами энергетики металлургических предприятий. Метод, основанный на использовании ИНС автокодировщиков, генеративно-состязательных и рекуррентных ИНС является новым в задаче обнаружения аномалий в поведении наблюдаемых динамических процессов автоматизированных систем управления при использовании машинного обучения без учителя. Модель объекта, подвергающегося воздействию кибератак, основанная на рассматриваемых в работе методах, относится к поведенческим моделям динамических процессов информационной системы. В результате использования подходов, основанных на моделях и методах раннего обнаружения кибератак, ожидается значительное снижение потенциального ущерба, который могут понести металлургические предприятия в результате реализации киберинцидентов.

Системы энергетического комплекса, обеспечивающие работу металлургического производства, обладают достаточно большим количеством контрольных точек для регистрации информативных параметров при проведении мониторинга, а также сложным зонированием исследуемой системы по важности тех или иных регистрируемых параметров. Важным является решение вопроса о том, какие данные, в каком объеме, и с какой периодичностью необходимо регистрировать, чтобы обеспечить превентивное обнаружение кибератак. Поэтому, при решении поставленной задачи важно дать описание наиболее актуальных способов регистрации информативных данных в контрольных точках исследуемой информационной системы (точки входа для мониторинга), а также разработать методику поиска наиболее приори-

тетных для мониторинга зон распределения контрольных точек исследуемой информационной системы. Как правило, при анализе имеется возможность использовать лишь ограниченный набор параметров, имеющих максимальную информативность в обеспечении безопасности технологических процессов. Задача по выбору анализируемых параметров, обладающих максимальной информативностью, сводится к необходимости выбора способа их регистрации, от которого зависит скорость получения значений, задержки фиксации параметров, стоимость получения параметров.

Задачами представленной работы являются:

1) построение концепции поведенческой модели динамических процессов информационной системы объектов энергетики металлургического предприятия, применяемой для реализации предотвращения последствий вторжений;

2) построение методики выбора контрольных точек исследуемой информационной системы и регистрации в контрольных точках параметров, обладающих максимальной информативностью, используемых при решении задачи обнаружения аномалий в наблюдаемых процессах, вызванных воздействием кибератак;

3) разработка и исследование поведенческой модели для обнаружения аномалий (обнаружения признаков кибератак) в наблюдаемых процессах информационной системы, с использованием технологий машинного обучения и ИНС автокодировщиков, а также ИНС с генеративно-состязательной и рекуррентной архитектурами.

1. Поведенческая модель динамических процессов информационной системы объектов энергетики металлургического предприятия, реализующая концепцию предотвращения последствий вторжений.

Информационная система объектов энергетики металлургического предприятия относится к сложным системам, поэтому, построение модели подобной системы базируется на принципах системного анализа [1].

В [2] отмечено, что построение адекватной модели для сложной системы невозможно, теория сложных систем должна состоять из простейших моделей нарастающей сложности, то есть грубая модель более сложной системы может быть проще точной модели более простой системы. Отмечается, что возможность построения теории сложных систем связана с возможностью построения их простых оптимизационных моделей [2]. В [3] отмечено, что при построении модели сложной системы следует руководствоваться главной целью и главным свойством исследуемой системы.

Таким образом, для построения математической модели сложной системы необходимо пойти на ослабление требований к математическому описанию исследуемой системы. При этом целесообразно использовать подход, связанный с построением диагностической модели исследуемой сложной системы, то есть модели, позволяющей достоверно оценить состояние исследуемой сложной системы. Построение диагностической модели предполагает в определённой форме идентификацию связи измеряемого вектора признаков с некоторым тестируемым свойством [4]. При этом от диагностической модели не требуется максимальной адекватности описания исследуемой системы в целом, диагностическая модель лишь предполагает описание степени отклонения технического состояния исследуемой системы от состояния, соответствующего «норме» [4].

Одной из форм построения диагностической модели является моделирование функционирования сложной системы в виде одно- или многомерных временных рядов, при этом оценка отклонения состояния исследуемой системы от «нормы» основывается на принятых допущениях о характере наблюдаемых временных рядов. В современных исследованиях оценка отклонения состояния исследуемой системы от «нормы», формируемая на основе анализа наблюдаемых временных рядов, отражающих процессы, протекающие в системе, связывается с понятием «аномалии» наблюдаемого временного ряда.

На принятых допущениях о характере наблюдаемых временных рядов, соответствующих понятию нормы исследуемой системы, строится модель временного ряда, характеризующего наблюдаемые процессы, протекающие в системы в состоянии нормы. Мера рассогласования наблюдаемого временного ряда исследуемой системы и модели временного ряда, соответствующей состоянию нормы исследуемой системы, отражает неожидаемое изменение в поведении наблюдаемых процессов, то есть аномальные изменения в динамике наблюдаемых временных рядов

данных, протекающих в исследуемой технической системе. Методы обнаружения аномалий в процессах исследуемой системы, отражаемых в виде временных рядов данных относят к поведенческим методам [5, 6].

В настоящее время методы обнаружения аномалий применяются для решения задач обнаружения кибератак как на информационном, так и на кибернетическом уровнях в сложных технических системах [7].

Кибератаки вызывают аномалии (то есть, неожидаемые изменения) в поведении наблюдаемых процессов (в динамике наблюдаемых временных рядов данных) при работе сложных технических систем. При этом задача обнаружения аномалий состоит в обнаружении расхождений между текущим (наблюдаемым) процессом работы сложной технической системы и процессом работы, который является эталонным для сложной технической системы (то есть, для системы, работающей в штатном режиме). Любое обнаруженное несоответствие наблюдаемого процесса и эталонного процесса, протекающего в сложной технической системе, рассматривается как аномалия (или кибервторжение).

Информационная система объектов энергетики металлургического предприятия относится к сложным техническим системам. Поэтому модель для раннего обнаружения кибератак и предотвращения последствий вторжений на объекты энергетики металлургического предприятия необходимо определить как диагностическую поведенческую модель процессов, протекающих в исследуемой технической системе. При этом модель процесса строится на основе временных рядов данных, наблюдаемых в исследуемой технической системе. Преимущество диагностической поведенческой модели данного типа – возможность обнаружения новых кибератак без модификации или обновления параметров модели.

Для построения диагностической поведенческой модели процессов, протекающих в информационных системах объектов энергетики металлургического предприятия целесообразно использовать интеллектуальный анализ данных, а также технологии машинного обучения, позволяющие выделять новую значимую информацию из большого объёма данных.

Значимую роль в технологическом обеспечении интеллектуального анализа данных с использованием технологии машинного обучения, в настоящее время играет аппарат искусственных нейронных сетей (ИНС). ИНС для обнаружения аномалий обучаются в течение некоторого периода времени, когда всё наблюдаемое поведение исследуемой системы считается нормальным [8, 9, 10]. После обучения нейронная сеть запускается в режиме распознавания. В ситуации, когда во входном потоке не удается распознать нормальное поведение, фиксируется аномалия, то есть факт кибератаки. В случае использования репрезентативной обучающей выборки нейронные сети дают хорошую устойчивость в пределах заданной системы.

ИНС, используемые для обнаружения аномалий в наблюдаемых процессах, можно рассматривать в контексте обнаружения аномальных образов, то есть в качестве детектора аномалий. Реализуемость детектора аномалий базируется на таком свойстве ИНС, как умение восстанавливать входную информацию на выходе. Для создания детектора аномалий необходимо составить обучающую выборку, состоящую из нормальных данных, и обучить ИНС воспроизводить на выходе нормальные данные. Тогда, если после обучения подать на вход ИНС нормальный образ, то ошибка реконструкции для входного нормального образа будет меньше, чем для входного аномального образа. При превышении ошибки реконструкции входного образа некоторого порогового значения принимается решение о принадлежности входного образа к классу аномальных образов, то есть фиксируется факт кибератаки. Для обнаружения аномальных образов наиболее подходят ИНС – автокодировщики (автоэнкодеры) [11, 12, 13, 14, 15] и генеративно-состязательные ИНС [16, 17, 18, 19].

Применение аппарата ИНС автокодировщиков, генеративно-состязательных и рекуррентных ИНС полностью соответствует концепции построения диагностической поведенческой модели, предполагающей описание степени отклонения технического состояния исследуемой системы от состояния нормы, используемой для раннего обнаружения кибератак и предотвращения последствий вторжений на объекты энергетики металлургического предприятия.

Разработанную диагностическую поведенческую модель можно отнести к технологиям, основанным на раннем обнаружении аномалий в динамических процессах, распределенных в информационной системе ис-

следуемого объекта. Поиск киберугроз, направленных на информационные системы объектов энергетики металлургического предприятия, становится постоянной фоновой деятельностью, позволяющей экономить значительные средства за счет оперативного реагирования на возникающие киберугрозы с целью предотвращения вторжений. Экономия средств также достигается за счет исключения из «ручной» обработки событий, которые не выводят качество управления объектом за пределы заданных значений [20].

2. Методика выбора контрольных точек для регистрации параметров, обладающих максимальной информативностью при решении задачи обнаружения аномалий в наблюдаемых процессах информационной системы, вызванных воздействием кибератак.

Идеальный подходом по организации мониторинга информационной системы объекта можно назвать подход, который заключается в мониторинге каждого отдельного параметра объекта. Но, при этом требуется организация сбора огромного количества данных и их оптимизация программными средствами для последующего анализа. Подобный подход является финансово затратным, поскольку требует сформированной и отлаженной системы мониторинга, где методы регистрации информативных данных оптимизированы по степени важности. Например, для получения информации с OPC-сервера (Open Platform Communications – сервера на основе программных технологий, предоставляющих единый интерфейс для управления объектами автоматизации и технологическими процессами) не требуется больших финансовых вложений, если сервер уже функционирует. Однако этот подход создает гораздо большую задержку при регистрации параметров, чем, например, использование подхода на основе зеркалирования трафика с помощью TAP-устройств или коммутаторов со SPAN-портами. Такие решения требуют дополнительных финансовых затрат, однако создают меньшие задержки при регистрации информации. Это особенно важно, когда требуется высокая частота фиксации параметров технологического процесса.

В представленной работе предложена методика, позволяющая:

1) определять информативные точки для системы мониторинга с учетом специфики рассматриваемого параметра технологического процесса;

2) оценивать наиболее критичные контрольные точки фиксации информационных параметров автоматизированной системы управления технологическим процессом (АСУ ТП);

3) создавать, либо модернизировать систему мониторинга информационных процессов производственных объектов с целью повышения ее эффективности.

Для построения методики предварительно выделяются все считываемые параметры p_i технологического процесса. Для оценки каждого параметра p_i рассчитан следующий набор характеристик:

1. Нормированные значения среднеквадратической ошибки

$$RRMSE_{p_i} = \frac{\sqrt{\frac{1}{N} MSE_{p_i}}}{\max(p_i) - \min(p_i)},$$

где MSE_{p_i} – среднеквадратическая ошибка параметра p_i, N – количество измерений, используемых для расчета ошибки.

Параметр MSE_{p_i} определяет, насколько величина p_i отклоняется от усредненного значения на заданном временном интервале.

2. По рассчитанным значениям параметра $RRMSE_{p_i}$ вычисляется вес параметра $RRMSE_{p_i}$ – w_{RRMSE} [21] (табл. 1).

Таблица 1

Таблица весов нормированных значений среднеквадратической ошибки считываемых параметров технологического процесса

Значения $RRMSE_{p_i}$	**Вес w_{RRMSE}**
RRMSE > 0.5	0.1
0.5 > RRMSE > 0.25	0.5
0.25 > RRMSE > 0.1	0.75
0.1 > RRMSE > 0	1

3. Величина *FTTI* – промежуток времени от возникновения аномального состояния наблюдаемого информационного процесса (от выхода величины из нормального диапазона значений) до наступления аварийной ситуации в технологическом процессе [22].

Параметр *FTTI* устанавливается экспертным путем и вычисляется в секундах. Параметр *FTTI* необходимо нормализовать в пределах от 0 до 1 относительно остальных параметров p_i и вычислить итоговый вес параметра $w_{FTTI}=1\text{-}FTTI_{norm}$. В итоге получим величину, отражающую, насколько критичны быстрые изменения параметра вне нормального диапазона изменения величины p_i, характеризующего технологический процесс.

4. Величина w_r отражает релевантность параметра для модели машинного обучения. Типовая АСУ ТП на большом производстве может характеризоваться большим количеством параметров, при этом не все параметры могут быть информативными при обнаружении аномалий средствами машинного обучения. Для расчета параметра w_r рассчитывается метрика релевантности для каждого параметра p_i. Для этого используется один из методов отбора признаков с дальнейшим получением показателя важности параметра.

5. Величина w_k – доля полезной информации в информационном потоке [23]. Величина w_k отражает процентное соотношение информации, которую возможно использовать для получения сведений о структуре системы.

6. Величина w_m – соотношение генерируемых данных ко всем данным. Корректная работа моделей машинного обучения обеспечивается наборами данных, в которых отсутствуют пустые значения, появившиеся, например, из-за задержек при съеме состояния параметра технологического процесса. Недостающие значения генерируются с использованием статистических методов. Параметр w_m вычисляется как

$$w_m = 1 - \frac{d_{gen}}{d_{tot}},$$

где d_{gen} – количество сгенерированных значений, d_{tot} – общее количество данных.

7. Величина w_f – вес частота съема сигнала: отражает, насколько быстро изменяются значения параметра технологического процесса. Задается экспертным путем (табл. 2).

Таблица 2

Таблица весов частоты съема сигнала

Характеристика частоты съема сигнала	w_f
Очень медленная	0.1
Медленная	0.4
Средняя	0.7
Быстрая	0.9
Очень быстрая	1

Необходимо отметить, что оценка влияния каждого (считываемого при мониторинге) параметра p_i на технологический процесс выполняется исходя из соображений критичности с использованием различных риск-ориентированных моделей, внедряемых обычно в различных государствах на законодательном уровне. В приведенном примере расчета использована четырехуровневая модель категорирования.

8. Кроме параметров критичности с каждым технологическим объектом обычно сопоставляют меры функциональной безопасности [24], для этого используют понятие уровня полноты безопасности SIL, который представляет собой величину, отражающую способность системы обеспечивать выполнение функций безопасности. Всего существует 5 уровней SIL со значениями от 0 (минимальная функциональная безопасность) до 4 (максимальная функциональная безопасность).

9. Величина w_t – параметр влияния. Обозначим требуемый уровень защищённости как SL_T (определяемый исходя из моделей рисков [25, 26]), а достигнутый как SL_A. Уровни SL нормируются от 0 до 4, где 0 – отсутствие требуемых или применённых мер защиты. Тогда для каждого параметра p_i параметр влияния вычисляется как

$$w_i = 0{,}16 \times \frac{SIL+1}{C} \times \frac{SL_T+1}{SL_A+1},$$

где C – уровень критичности в зависимости от степени ущерба. В методике использованы категории безопасности объектов критической информационной инфраструктуры, применяемые в российском законодательстве, где 1 – самый высокий уровень критичности, 3 – самый низкий. Для удобства расчётов введен уровень 4 (отсутствие критичности) для объектов, не соответствующих критериям даже минимального уровня 3.

Коэффициент 0,16 применяется для нормирования значения w_i в диапазоне [0; 1].

10. Итоговый вес важности параметра p_i, определяемый как W_{pi} вычисляется по формуле

$$W_{p_i} = \frac{(w_{RRMSE}+w_{FTTI}+w_r+w_m+w_f+w_i+w_k)}{7}.$$

Зная итоговый вес важности W_{pi} параметра p_i, определим необходимый способ мониторинга сигнала p_i.

11. Обозначим способ получения информации индексом j, который может принимать следующие значения: 4 – отсутствие мониторинга, 3 – получение параметра технологического процесса путем выгрузки из базы данных, 2 – подключение напрямую к SCADA или OPC-серверу, 1 – получение данных с минимальными задержками с использованием SPAN-портов и TAP-устройств. Время получения выгрузки обозначим безразмерной величиной t_j.

Величина R_Z – относительное улучшение мониторинга, которое определяется как

$$R_Z = t_{Aj}/t_{Tj},$$

где t_{Aj} – текущее время, t_{Tj} – прогнозируемое время после добавления средства мониторинга.

Пусть в зоне (или тракте) Z обрабатывается Z_n параметров АСУ ТП, тогда рассчитав итоговый вес W_{pzi} для каждого параметра p_{Zi}, можно выразить общий ранг улучшения (для зоны или тракта), как

$$R_Z = \frac{t_{AjZ}}{t_{TjZ}} \times \sum_{i=1}^{n_Z} W_{pZi}.$$

В качестве примера расчет параметра W_{pi} выполнен для АСУ ТП электрометаллургического предприятия, контролирующей 6 параметров. Схема АСУ ТП приведена на рис. 1. АСУ ТП разделена на условные 4 зоны, обозначенные на схеме как *Z1, Z2, Z3* и *Z4*. Эти зоны требуются для расчета ранга улучшения. При этом параметры в зоне *Z1* собираются путем выгрузки из базы данных, в зонах *Z2* и *Z3* постоянный мониторинг отсутствует, а в зоне *Z4* параметры получаются путём сбора со SCADA-сервера.

Параметры исследуемого технологического процесса и их описание представлены в табл. 3.

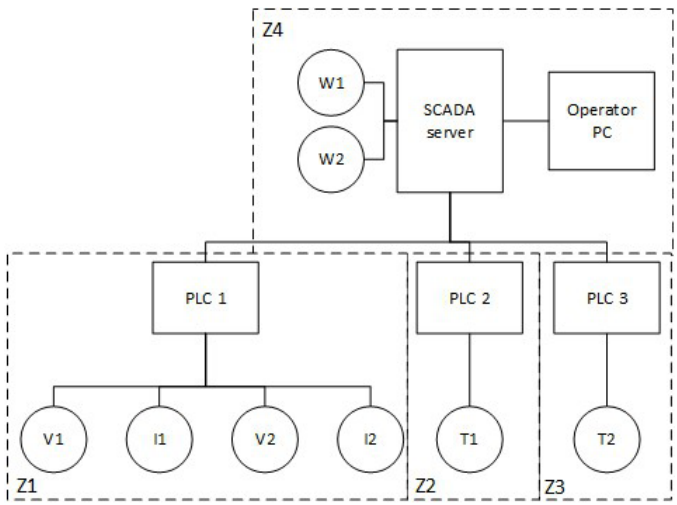

Рис. 1. Схема исследуемой АСУ ТП

Таблица 3

Параметры исследуемой АСУ ТП

Зоны	SIL	C	SL_T	SL_A	Обозначение параметра	Значение (описание) параметра
Z1	2	2	3	2	V1	Напряжение индукционной печи №1
					I1	Ток индукционной печи №1
					V2	Напряжение индукционной печи №2
					I2	Ток индукционной печи №2
Z2	2	2	4	2	T1	Температура в индукционной печи №1
Z3	2	2	4	2	T2	Температура в индукционной печи №2
Z4	2	2	3	2	W1	Мощность индукционной печи №1
					W2	Мощность индукционной печи №2

В табл. 4 приведены результаты расчетов характеристик модели. При расчете параметра w_i были использованы значения *SIL, C,* SL_T и SL_A из табл. 3.

Таблица 4

Значения характеристик модели для параметров АСУ ТП

p	w_{RRMSE}	w_{ETTI}	w_r	w_k	w_m	w_f	w_i	W_p
V1	1.00	0.08	0.04	0.40	0.85	0.70	0.32	0.48
I1	0.50	0.12	0.09	0.50	0.95	0.40	0.32	0.41
V2	1.00	0.09	0.07	0.30	0.80	0.70	0.32	0.47
I2	0.75	0.08	0.04	0.50	0.90	0.40	0.32	0.43
T1	0.75	0.29	0.49	0.80	1.00	0.70	0.40	0.63
T2	0.10	0.37	0.24	0.90	0.80	0.90	0.40	0.53
W1	0.75	0.06	0.04	0.50	1.00	1.00	0.32	0.52
W2	0.75	0.07	0.03	0.50	1.00	1.00	0.32	0.52

Далее для каждой из трех зон рассчитан прогнозируемый ранг улучшений R_Z. С учётом финансовых возможностей предприятия сделано предположение, что имеется возможность установки средств мониторинга на один уровень выше для зон *Z1 – Z3*, а в зоне *Z4* улучшения провести невозможно.

Результаты расчета ранга улучшения для каждой зоны приведены в табл. 5.

Таким образом, при распределении бюджета на выполнение соответствующих улучшений в АСУ ТП возможно применять существующее ранжирование, бюджет может распределяться пропорционально полученным рангам, или же некоторые зоны и тракты могут игнорироваться по порогу. В нашем случае зона может считаться самой приоритетной при мониторинге технологического про-

Таблица 5

Значения для зон АСУ ТП

Зона	t_{Aj}/t_{Tj}	$\sum W_{pz_i}$	R_z
Z1	1.5	1.79	2.685
Z2	1.33	0.63	0.838
Z3	1.33	0.53	0.705
Z4	1	1.04	1.040

цесса по части оптимизации программно-аппаратных средств контроля параметров технологического процесса.

Представленная в методике модель позволяет получать численную оценку необходимости улучшения мониторинга для разных зон технологического процесса. Технически это дает возможность оптимизировать систему мониторинга для существующих производств, где требуется внедрение решений по поиску аномалий, либо выбрать только наиболее критичные точки съема информации для оптимизации входного набора данных для дальнейшего исследования методами машинного обучения. В качестве примера использования методики приведен расчет на основе части технологического процесса металлургического предприятия, на основании которого выбраны наиболее важные точки регистрации данных для системы мониторинга.

3. Методы обнаружения и классификации нехарактерных состояний (аномальных изменений в динамических процессах) АСУ ТП.

Для разработки диагностической поведенческой модели исследованы методы обнаружения и классификации аномальных состояний информационных процессов с использованием ИНС автокодировщиков, генеративно-состязательных и рекуррентных ИНС [27, 28, 29].

Первый представленный метод – обнаружение аномалий с использование ИНС автокодировщиков. Автокодировщики представляют собой нейронные сети, обладающей особой архитектурой, позволяющей применять обучение без учителя, используя метод обратного распространения ошибки. Простейшая архитектура автокодировщика представляет собой сеть прямого распространения без обратных связей, и содержащая входной слой, промежуточный слой и выходной слой. Выходной слой автокодировщика должен содержать столько же нейронов, сколько и входной слой. Основ-

ной принцип работы и обучения сети автокодировщика — получить на выходном слое отклик, наиболее близкий к входному. Чтобы решение не оказалось тривиальным, на промежуточный слой автокодировщика накладывают ограничения: промежуточный слой должен быть или меньшей размерности, чем входной и выходной слои, или искусственно ограничивается количество одновременно активных нейронов промежуточного слоя — разрежённая активация. Эти ограничения заставляют ИНС искать обобщения в поступающих на вход данных, выполнять их сжатие. Таким образом, ИНС автоматически обучается выделять из входных данных общие признаки.

Автокодировщик состоят из двух частей: кодировщика и декодировщика. Кодировщик *g* переводит входной сигнал *x* в его представление (код) *h*:

$$h=g(x),$$

а декодер *f* восстанавливает сигнал *x* по его коду *h*:

$$x=f(h).$$

Автокодировщик, изменяя декодер *f* и кодировщик *g*, стремится выучить тождественную функцию $x=f(g(x))$, минимизируя некоторый функционал ошибки

$$L(x, f((g(x)))).$$

При этом семейства функций кодировщика и декодера ограничены, чтобы автокодировщик был вынужден отбирать наиболее важные свойства сигнала.

Для проведения вычислительных экспериментов был использован набор данных Power System Attack Dataset. Данный набор сгенерирован из различных наборов, содержащих 37 сценариев аномальных ситуаций, происходивших на объектах энергетики. Схема конфигурации, задействованной в данном наборе, приведена на рис. 2. Исходный набор данных был разделен на две части – тренировочный набор (80% данных) и контрольный набор (20% данных). Перед применением метода была проведена предварительная обработка данных, включая удаление строк с пустыми данными и нормализацию значений.

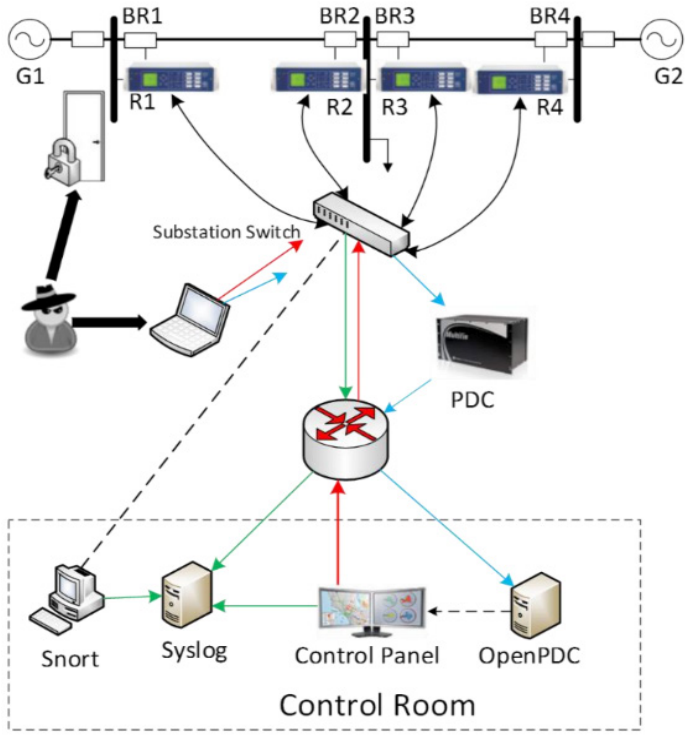

Рис. 2. Конфигурация энергетической системы Power System Attack Dataset

Архитектура ИНС автокодировщика, которая была использована для проведения вычислительных экспериментов, представлена на рис. 3.

Некоторые экспериментальные значения представлены в табл. 6.

Обозначения, использованные в табл. 6:

Bottleneck size – число нейронов в самом малом слое автокодировщика (чем меньше значение, тем больше потеря информации);

epochs – максимальное количество эпох при обучении;

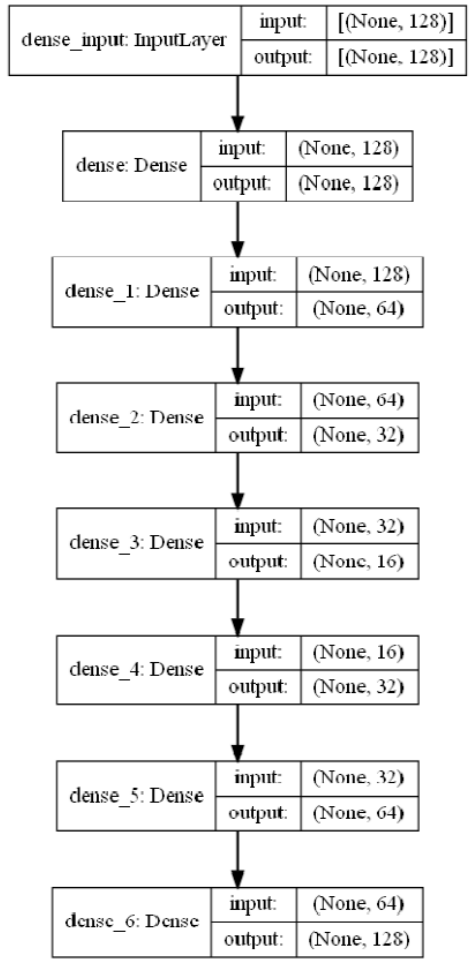

Рис. 3. Архитектура ИНС автокодировщика

Таблица 6

Результаты работы ИНС автокодировщика

Bottleneck Size	epochs	batch size	Naturals percent, %	id
16.0	50.0	16.0	95	POW1-60
16.0	50.0	16.0	90	POW1-61
16.0	50.0	16.0	85	POW1-62
16.0	50.0	16.0	80	POW1-52
16.0	50.0	16.0	75	POW1-63
16.0	50.0	16.0	70	POW1-64
16.0	50.0	16.0	65	POW1-53
16.0	50.0	16.0	60	POW1-54
16.0	50.0	16.0	55	POW1-77
16.0	50.0	16.0	50	POW1-78

batch size – число точек, используемых при обучении за момент времени (чем больше количество, тем меньше влияет на результат отдельная точка);

naturals percent – соотношение «нормальных» точек к их общему количеству;

id – идентификатор данных.

На рис. 4, 5 и 6 показаны, соответственно, зависимости метрик точности (precision), полноты (recall) и значения F_1-меры от порога отсечения. Точность (precision) интерпретируется как доля значений данных, названных автокодировщиком аномальными и при этом действительно являющимися аномальными,

а полнота (recall) показывает, какую долю аномальных данных из всего множества аномальных данных обнаружил автокодировщик. F_1-мера представляет собой агрегированный критерий качества, объединяющий precision и recall:

$$F_\beta = (1 + \beta^2) \frac{precision \cdot recall}{(\beta^2 \cdot precision) + recall}, \quad (1)$$

где β – вес точности в метрике, при $\beta = 1$ F_1 – это среднее гармоническое с множителем 2, чтобы в случае precision = 1 и recall = 1 иметь $F_1 = 1$.

Порог отсечения является настраиваемым параметром и определяет чувствительность метода по обнаружению аномальных данных. Оператор системы в зависимости от ситуации может изменять порог отсечения, получая большее количество сообщений об аномалиях, но с меньшей точностью, или наоборот, получая меньшее количество сообщений об аномалиях, но с большей точностью.

Из табл. 6 и рисунков видно, что представленный подход обладает достаточно хо-

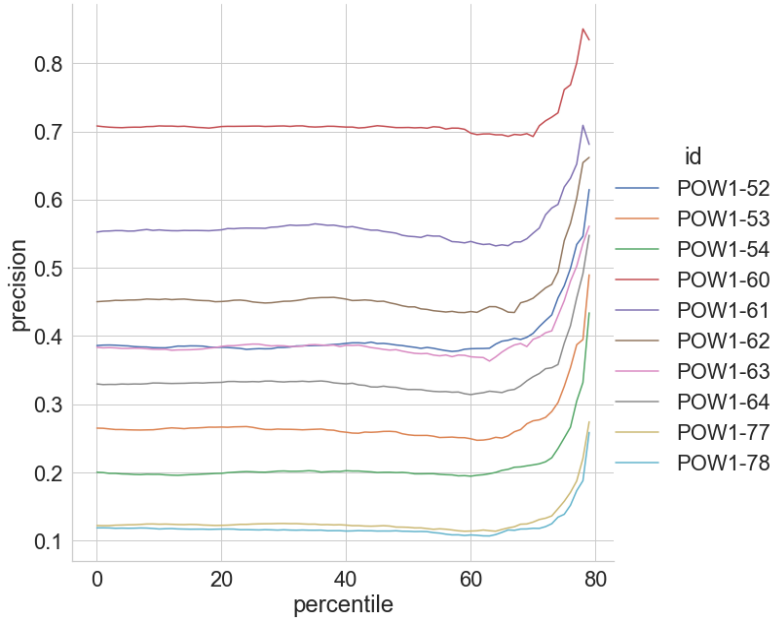

Рис. 4. Зависимость метрики точности от порога отсечения

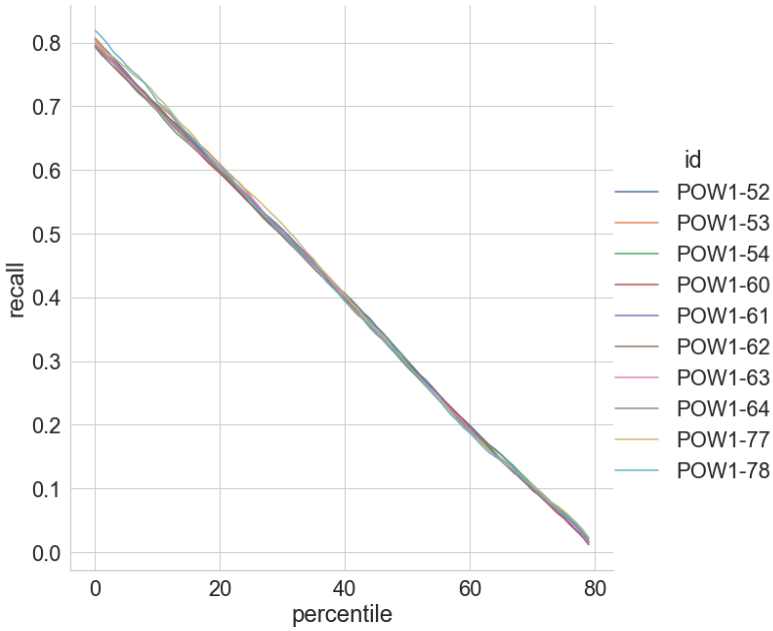

Рис. 5. Зависимость метрики полноты от порога отсечения

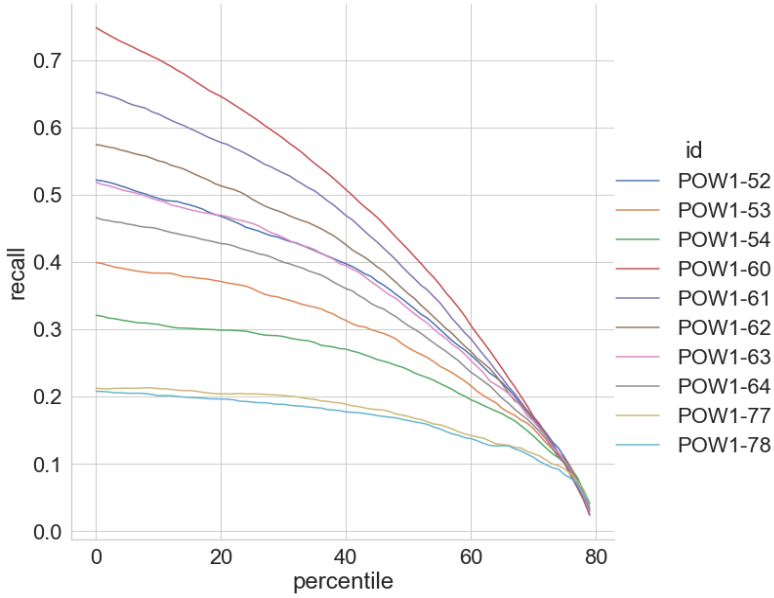

Рис. 6. Зависимость значения F1-меры от порога отсечения

рошей точностью (>0.6) при большом пороге отсечения (>95%), но наиболее оптимальные результаты достигаются тогда, когда количество аномальных состояний составляет не более 10% в общем объёме данных. При достижении данного условия, точность метода возрастает до 0.85.

Результаты, полученные при помощи автокодировщика, показывают, что данный метод возможно использовать для обнаружения аномалий в объектах энергетики, при учете того, что количество аномальных состояний в общем объёме регистрируемых данных будет не более 10%.

Второй представленный метод – обнаружение аномалий с использованием генеративно-состязательных нейронных сетей (Generative adversarial network, GAN). Генеративно-состязательные нейронные сети – это класс ИНС, который используется для генерации новых синтетических данных, основанных на имеющихся данных. Концептуально генеративно-состязательные сети основаны на идее состязательного обучения. На сегодняшний день, сети GAN используются для генерации изображений [30, 31, 32], звука [33], текста [34]. Кроме того, говоря о применении генеративно-состязательных сетей в контексте информационной безопасности, нужно отметить, что по сравнению с другими моделями глубокого обучения, GAN обладают преимуществом в случае атаки (adversarial attack) [35] со стороны злоумышленника: благодаря своей состязательной природе, обмануть такие сети сложнее, чем обычные дискриминативные модели.

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. Схема генеративно-состязательной сети представлена на рис. 7.

Генератору *G* на вход подаётся вектор, состоящий из случайного шума – вектор из некоторого пространства скрытых переменных (latent space) *Z*, на котором задано априорное распределение $p_z(z)$. На выходе генератор выдаёт новый объект из пространства данных *X*. Формально сеть-генератор можно описать в виде функции

$$G=G(z;\theta_g):Z\rightarrow X,$$

где θ_g – параметры сети-генератора. В ходе обучения генератор аппроксимирует распределение p_{data} выборки реальных данных *X*. Следовательно, по окончании обучения распределение порождаемых генератором объектов p_{gen} должно быть приближенно к распределению реальных объектов p_{data}.

Для того, чтобы генератор с каждой итерацией обучения обладал лучшей аппроксимацией распределения p_{data}, используется сеть дискриминатор. Дискриминатор, как правило, является обычным бинарным классификатором, на вход которому подаётся

Рис. 7. Схема ИНС с генеративно-состязательной архитектурой (GAN)

объект *x* из пространства данных *X*. На выходе дискриминатор выдаёт вероятность принадлежности объекта к тому или иному классу: реальный объект или объект, порождённый дискриминатором. Формально дискриминатор определяется выражением

$$D=D(x;\theta_d):X\rightarrow[0,1],$$

где θ_d – параметры сети-дискриминатора. Чтобы аппроксимировать распределение p_{data}, генератору нужно научиться обманывать дискриминатор: научиться порождать такие объекты, которые дискриминатор не в состоянии отличить от настоящих.

Для использования генеративно-состязательных сетей для решения задачи обнаружения аномалий необходимо обладать возможностью получать информацию о прообразах реальных объектов в пространстве шума, представленном найденными сетью признаками. Такая возможность реализована в двунаправленной генеративно-состязательной сети (Bidirectional GAN, BiGAN) [36]. Схема такой генеративно-состязательной сети представлена на рис. 8.

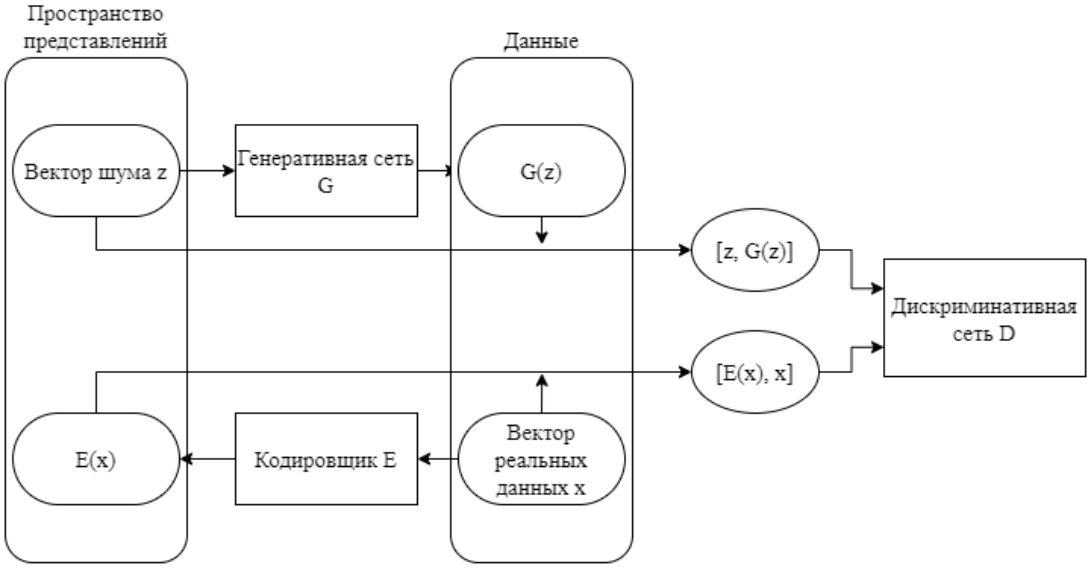

Рис. 8. Схема двунаправленной генеративно-состязательной сети (BiGAN)

По сравнению с базовой архитектурой генеративно-состязательной сети, архитектура BiGAN дополнительно использует кодирующую сеть E, которая осуществляет преобразование из пространства реальных данных в пространство скрытых переменных (из которого берётся вектор шума для генератора). Формально это действие можно представить в виде функции:

$$E=E(x;\theta_e):X\rightarrow Z,$$

где θ_e – параметры сети-кодировщика.

Таким образом, сеть-кодировщик в процессе обучения учиться выполнять преобразование, обратное генератору. Также в архитектуре BiGAN сеть-дискриминатор обучается различать не только порожденные генератором объекты и объекты из реальных дан-

ных, но и векторы из пространства скрытых переменных Z (используемые для порождения объектов) и результат отображения кодировщика $E(x)$.

Добавление сети-кодировщика в архитектуру генеративно-состязательной сети, которая каждому объекту из пространства данных x ставит в соответствие вектор z из пространства скрытых переменных, позволяет непосредственно извлекать представления этих объектов. С точки зрения семантики, вектор z репрезентирует признаки объекта , выявленные в процессе обучения BiGAN. Именно это свойство позволяет использовать BiGAN для выявления аномалий [37].

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

$$A(x)=aL_G(x)+(1-a)L_D(x),$$

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

$$L_G(x)=||x-G(E(x))||.$$

$L_D(x)$ – функция потерь дискриминатора:

$$L_D(x)=$$

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

Таким образом, обученная на данных, соответствующих нормальному состоянию, сеть BiGAN 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 применения генеративно-состязательной сети BiGAN были реализованы модели генератора, дискриминатора и кодировщика. Каждая из этих моделей является обычной полносвязной сетью. Разработанная генеративно-состязательная нейронная сеть испытана на наборе данных, полученных с лабораторного стенда Secure Water Treatment (SWaT) [38], схема которого представлена на рис. 9. Каждая запись набора состоит из 51 значения сенсора или привода, всего в наборе содержалось 964722 записи, записанных за 11 дней. Имитационные атаки проводились в течение 4 дней.

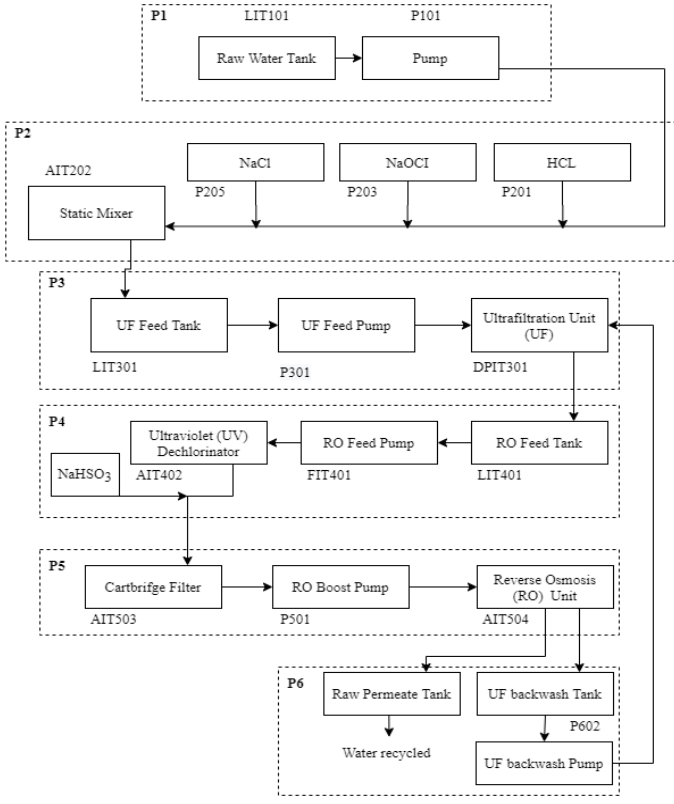

Рис. 9. Схема стенда SWaT

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 применялась обученная модель, для каждого объекта контрольной выборки строилась метрика аномальности. В ходе проведения эксперимента, модели обучались в течение 20 эпох. На рис. 10, 11 и 12 представлены графики функций потерь каждой из моделей.

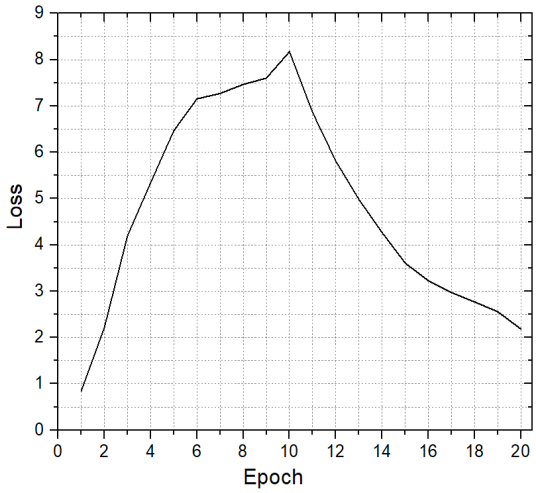

Рис. 10. Функция потерь генератора

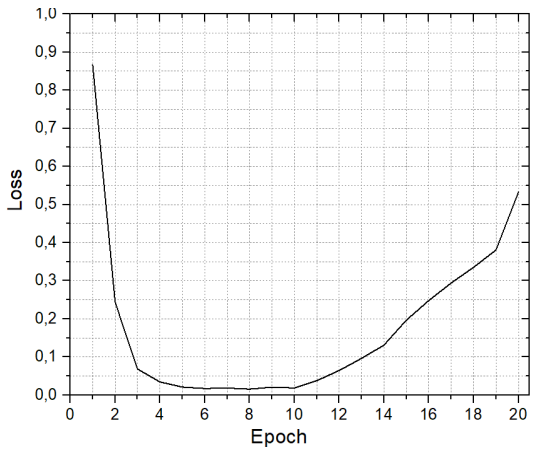

Рис. 11. Функция потерь дискриминатора

Графики функций потерь генератора, дискриминатора и кодировщика, представленные на рис. 10 – 12, отражают процесс обучения модели BiGAN: чем лучше работает дискриминатор, тем хуже работают генератор и кодировщик, и наоборот. Точка, в которой

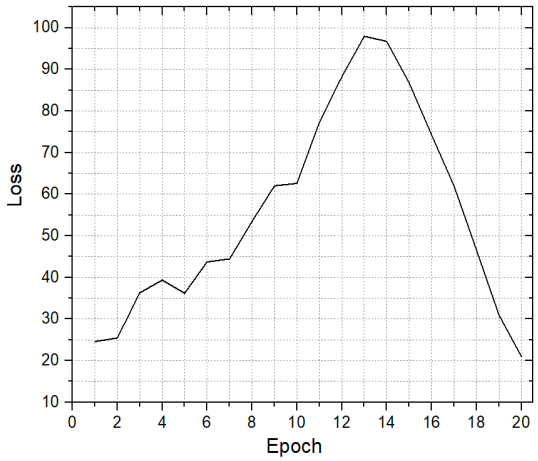

Рис. 12. Функция потерь кодировщика

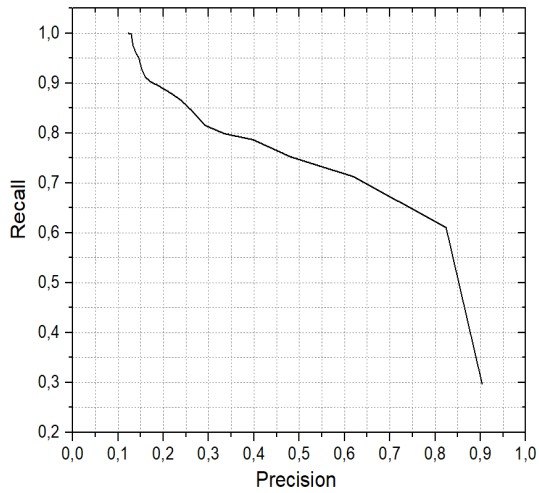

Рис. 13. Зависимость метрик точности (precision) и полноты (recall)

был остановлен момент обучения (20 эпоха), соответствует состоянию модели, при котором кодировщик и генератор работают лучше дискриминатора. Дальнейшее обучение позволяет ещё больше снизить значения функций потерь генератора и кодировщика, однако в этом случае, модель BiGAN утрачивает свою эффективность на стадии тестирования.

Качество работы обученной модели на практике во многом определяется не только тем, насколько хорошо модель научилась аппроксимировать распределение данных, соответствующих нормальной работе системы, но и конкретным значением порога метрики аномальности. Так, на рис. 13 представлена зависимость метрик точности и полноты в зависимости от порогового значения. При увеличении порогового значения увеличивается точность и уменьшается полнота. Это свой-

ство позволяет настроить желаемое поведение модели в каждом конкретном случае.

Для оценки качества работы модели использовались метрики точность (precision) и полнота (recall). Конкретные значения метрик качества определяются величиной порога для метрики аномальности. Как показано на рис. 13, при увеличении порога увеличивается значение метрики precision и уменьшается значение метрики recall.

По полученным результатам видно, что разработанная сеть может использоваться для обнаружения аномальных состояний. При этом генеративно-состязательные сети не требуют примеров данных, соответствующих аномальной работе системы, что повышает их эффективность в практическом использовании.

Третий представленный метод – обнаружение аномалий с использованием рекуррентных нейронных сетей, реализующих прогнозирование входного наблюдаемого временного ряда данных. Рекуррентные нейронные сети – это тип нейронных сетей, которые имеют соединения с обратной связью (рекуррентные соединения). Преимущества наличия таких связей состоит в том, что они позволяют запоминать предыдущее состояние и учитывать его при подсчете весов. Все рекуррентные сети имеют цепную структуру: повторяющуюся клетку. Клеткой может быть единственный нейрон или последовательность нескольких. Рассмотрим простую Recurrent Neural Network (RNN) клетку. Она имеет очень простую структуру: только один слой с функцией активации *than* (гиперболический тангенс). Пусть имеется последовательность входных данных $\{x_t\}_{t=1}^{T}$. В данном случае $x_t = (x_1^t, \dots, x_n^t)$ – векторное представление *t*-ого объекта во временном ряде. В каждый момент времени *t*, клетка анализирует объект x_t и предсказание прошлого объекта H_{t-1}. Именно в этом и проявляется рекуррентность сети.

Для целей данной работы использована ИНС Long Short Term Memory (LSTM). Сеть LSTM – особый вид рекуррентной нейронной сети, способный находить долго- и краткосрочные зависимости. LSTM также имеют цепную структуру, но повторяющаяся клетка имеет более сложное строение: она состоит из четырех нейронов, соединенных специальным образом. Структурная схема сети LSTM представлена на рис. 14.

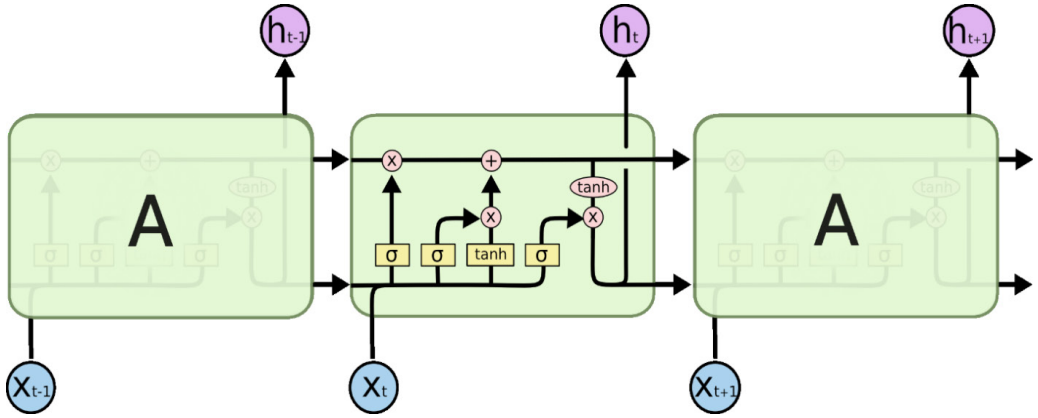

Рис. 14. Структурная схема ИНС LSTM

В работе было использована архитектура рекуррентной нейронной сети, состоящей из одномерного сверточного слоя, слоя отсева (для избегания переобучения), двух слоев LSTM и полносвязного выходного слоя. Функция Swish была использована в качестве функции активации для слоев LSTM. Предложенная сеть обрабатывает поданные на вход данные о состоянии объекта и на выходе выдает состояние объекта в следующий момент времени. Схема архитектуры ИНС представлена на рис. 15.

Разработанная нейронная сеть испытана на наборе данных, полученных с лабораторного стенда Secure Water Treatment (SWaT), который был уже описан ранее.

Результаты сравнения представленного метода обнаружения аномалий с другими методами приведены в табл. 7. Сравнение проводилось по метрикам точности (precision), отзыва (recall) и показателя F_1 (F_1-Score), определяемого формулой (1). По данным таблицам видно, что предложенный метод обладает достаточно высоким показателем F_1, усту-

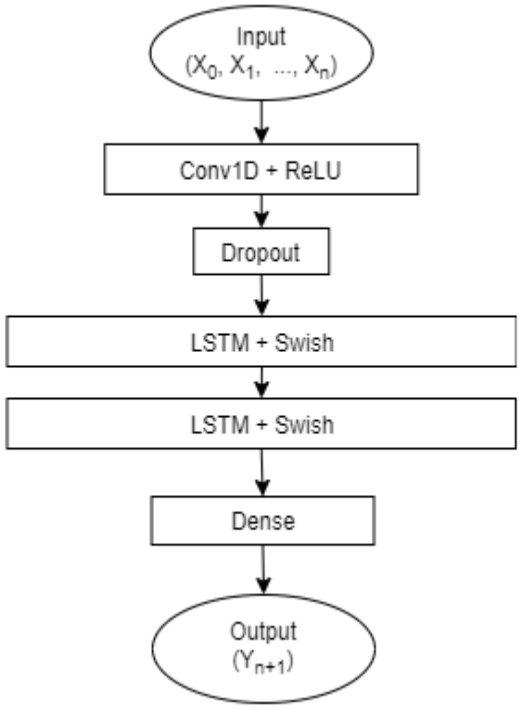

Рис. 15. Схема архитектуры ИНС LSTM

пая только методу 1D CNN. Относительно показателей точности и отзыва, предложенный метод обладает сравнимыми характеристиками и его использование возможно с целью обнаружения аномалий. Предложенный метод позволяет определить, в каком конкретном датчике произошла аномальная ситуация, что позволяет операционному персоналу объекта оперативно проанализировать ситуацию и принять необходимые превентивные меры для обеспечения информационной безопасности.

Для демонстрации качества работы представленной модели построена ROC-кривая, отображающая связь между истинно-положительными и ложноположительными результатами при разных порогах отсечки (рис. 16). О хорошей эффективности метода ИНС LSTM свидетельствует значительная отдаленность ROC-кривая от диагональной линии.

Заключение

Информационные системы объектов энергетики металлургического предприятия относится к сложным техническим системам. Поэтому модель для раннего обнаружения кибератак и предотвращения последствий вторжений на такие объекты определена как диагностическая поведенческая модель процессов, протекающих в исследуемой технической системе, которая строится на основе наблюдаемых временных рядов данных. Преимущество диагностической поведенческой модели такого типа заключается в возможности обнаружения новых кибератак без модификации или обновления параметров модели.

Таблица 7

Сравнение метода с применением ИНС LSTM с другими методами обнаружения аномалий

Метод	Precision	Recall	F_1-Score
1D CNN	0.968	0.791	0.871
MLP	0.967	0.696	0.812
CNN	0.952	0.702	0.808
RNN	0.936	0.692	0.796
DNN	0.982	0.678	0.802
OCSVM	0.925	0.699	0.796
LSTM	0.934	0.820	0.865

С целью выбора контрольных точек для регистрации параметров временных рядов данных, обладающих высокой информативностью при решении задачи обнаружения аномалий в наблюдаемых процессах информационной системы, вызванных воздействием кибератак, разработана соответствующая методика. Представленная в методике модель позволяет получить численную оценку необходимости улучшения для различных зон технологического процесса. Технически это позволяет сформировать такую систему мониторинга для производства, которую можно было бы считать достаточной для решения задач поиска аномалий наблюдаемых процессов и выбирать наиболее критичные точки регистрации данных для оптимизации входного набора данных для дальнейшего

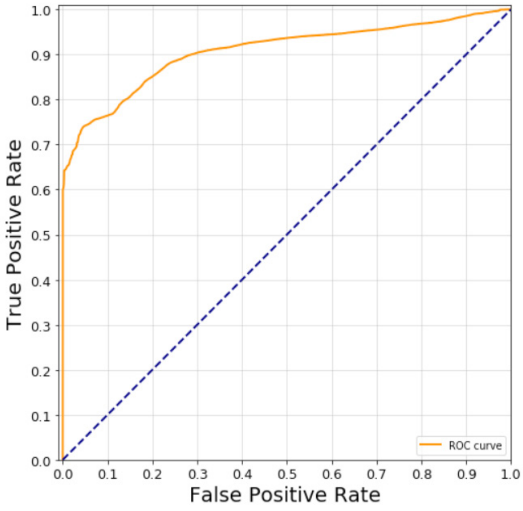

Рис. 16. ROC-кривая метода ИНС LSTM

исследования методами машинного обучения.

В ходе проведённого исследования показано, что применение аппарата ИНС автокодировщиков, генеративно-состязательных и рекуррентных ИНС полностью соответствует концепции построения диагностической поведенческой модели, предполагающей описание степени отклонения технического состояния исследуемой системы от состояния нормы, используемой для раннего обнаружения кибератак и предотвращения последствий вторжений на объекты энергетики металлургического предприятия.

Результаты, полученные авторами в ходе представленных исследований, обсуждены на международных конференциях «2021 IEEE Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology» (USBEREIT-2021, 13 – 14 мая 2021 г., г. Екатеринбург) и «International Russian Automation Conference» (RusAutoCon-2021, 5 – 11 сентября 2021 г., г. Сочи). Группой авторов выполнены также исследования, связанные с разработкой перспективных технологий иерархического кластерного анализа данных информационных процессов АСУ ТП, подвергающихся воздействию кибератак. Выполненные исследования представлены на международной конференции «International Multi-Conference on Industrial Engineering and Modern Technologie FarEastCon-2021» (FarEastCon-2021, 5 – 8 октября, 2021 г., г. Владивосток).

Литература

1. Волкова В.Н. Основы теории систем и системного анализа / В.Н. Волкова, А.А. Денисов. – СПб: Изд-во СПбГТУ, 1999. – 512 с.

2. Флейшман Б.С. Основы системологии / Б.С. Флейшман. – М.: Радио и связь. 1982. – 368 с.

3. Тырсин А.Н. О математическом моделировании сложных организационных систем // Организация и управление эффективностью и производительностью производственных и социальных систем: Матер. межд. научно-практ. конф. – Новочеркасск: ЮРГТУ (НПИ), 2005. – С. 49 – 50.

4. Биргер И.А. Техническая диагностика. – М.: Наука, 1978. – 239 с.

5. Debar H., Dacier M., Wespi A. Towards a taxonomy of intrusion-detection systems // Computer Networks. 1999. vol. 31. Issue 8. pp. 805 – 822.

6. Браницкий А.А., Котенко И.В. Анализ и классификация методов обнаружения сетевых атак. – Труды СПИИРАН. – 2016. – Вып. 45 – С. 207 – 244.

7. Абдулин А.А., Соколов А.Н. Исследование программных решений для обеспечения информационной безопасности промышленных сетей автоматизированных систем управления технологическими процессами //Вестник УрФО. Безопасность в информационной сфере. – 2021. № 1(39). – С. 43 – 53.

8. Herve Debar, Monique Becker, and Didier Siboni. A neural network omponent for an intrusion detection system //Proceedings of the 1992 IEEE Computer Sociecty Symposium on Research in Security and Privacy, pages 240–250, Oakland, CA, USA, May 1992.

9. Rasool Jalili, Fatemeh Imani-Mehr, Morteza Amini, Hamid Reza Shahriari. Detection of Distributed Denial of Service Attacks Using Statistical Pre- Processor and Unsupervised Neural Networks // Lecture notes in computer science, 2005.

10. Смелянский Р.Л., Качалин А.И. Применения нейросетей для обнаружения аномального поведения объектов в компьютерных сетях //Факультет Вычислительной Математики и Кибернетики, МГУ им. М. В. Ломоносова, Москва, 2004.

11. Neural network and artificial immune systems for malware and network intrusion detection / V. Golovko [et al.] // Studies in computational intelligence. – Heidelberg, 2010. – Vol. 263 : Advances in machine learning II. – P. 485 – 513.

12. Muna A.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. – № 41. – P. 1 – 11.

13. Sung Jin Kim, Woo Yeon Jo, Taeshik Shon. APAD: Autoencoder-based Payload Anomaly Detection for industrial IoE. December 2019 Applied Soft Computing 88:106017. DOI: 10.1016/j.asoc.2019.106017.

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

15. Schmidhuber J. Deep learning in neural networks: An overview //Neural networks. – 2015. – № 61. – P. 85 – 117.

16. Goodfellow I. et al. Generative adversarial nets //Advances in neural information processing systems. – 2014. – P. 2672 – 2680.

17. Madry A. et al. Towards deep learning models resistant to adversarial attacks //arXiv preprint arXiv:1706.06083. – 2017.

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

19. Алабугин С.К., Соколов А.Н. Использование генеративно-состязательных нейронных сетей при выявлении аномалий технологического процесса //Вестник УрФО. Безопасность в информационной сфере. – 2020. № 4(38). – С. 64 – 68.

20. Гарбук С.В., Правиков Д.И., Полянский А.В., Самарин И.В. Обеспечение информационной безопасности АСУ ТП с использованием метода предиктивной защиты //Вопросы кибербезопасности. – 2019. – № 3(31). – С. 30 – 36.

21. Quilot B., Génard M., Lescourret F., Kervella J. Simulating genotypic variation of fruit quality in an advanced peach×Prunus davidiana cross // Journal of Experimental Botany, Volume 56, Issue 422, December 2005, Pages 3071–3081, https://doi.org/10.1093/jxb/eri304.

22. Martin H., Tschabuschnig K., Bridal O., Watzenig D. (2017) Functional Safety of Automated Driving Systems: Does ISO 26262 Meet the Challenges?. In: Watzenig D., Horn M. (eds) Automated Driving. Springer, Cham. https://doi.org/10.1007/978-3-319-31895-0_16.

23. 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.

24. Parts 1 – 7 IEC 61508, Functional safety of electrical/electronic/ programmable electronic safety-related systems.

25. Braband J. Towards an IT security risk assessment framework for rail- way automation //CoRR abs/1704.01175, http://arxiv.org/abs/1704.01175, 2017.

26. Maidl M., Kröselberg D., Christ J. and Beckers K. A Comprehensive Framework for Security in Engineering Projects - Based on IEC 62443. // 2018 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW), 2018, pp. 42-47, doi: 10.1109/ISSREW.2018.00.

27. Pyatnitsky I.A., Sokolov A.N. Determination of the Optimal Ratio of Normal to Anomalous Points in the Problem of Detecting Anomalies in the Work of Industrial Control Systems // 2021 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT), 2021, pp. 0478-0480, doi: 10.1109/USBEREIT51232.2021.9455010.

28. Alabugin S.K., Sokolov A.N. Applying of Generative Adversarial Networks for Anomaly Detection in Industrial Control Systems // 2020 Global Smart Industry Conference (GloSIC), 2020, pp. 199-203, doi: 10.1109/GloSIC50886.2020.9267878.

29. Alabugin S.K., Sokolov A.N. Applying of Recurrent Neural Networks for Industrial Processes Anomaly Detection // 2021 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT), 2021, pp. 0467-0470, doi: 10.1109/USBEREIT51232.2021.9455060.

30. Radford A., Metz L., Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks //arXiv preprint arXiv:1511.06434. – 2015.44.

31. Ledig C. et al. Photo-realistic single image super-resolution using a generative adversarial network //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2017. – С. 4681-4690.

32. Odena A., Olah C., Shlens J. Conditional image synthesis with auxiliary classifier gans //Proceedings of the 34th International Conference on Machine Learning-Volume 70. – JMLR.org, 2017. – С. 2642-2651.

33. Engel J. et al. Gansynth: Adversarial neural audio synthesis //arXiv preprint arXiv:1902.08710. – 2019.

34. Subramanian S. et al. Towards text generation with adversarially learned neural outlines //Advances in Neural Information Processing Systems. – 2018. – P. 7551 – 7563.

35. Madry A. et al. Towards deep learning models resistant to adversarial attacks //arXiv preprint arXiv:1706.06083. – 2017.

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

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

38. J. Goh, S. Adepu, K. N. Junejo and A. Mathur, "A dataset to support research in the design of secure water treatment systems", Proc. International Conference on Critical Information Infrastructures Security (CRITIS 2016), 2016.

References

1. Volkova V.N. Osnovy teorii sistem i sistemnogo analiza / V.N. Volkova, A.A. Denisov. – SPb: Izd-vo SPbGTU, 1999. – 512 s.

2. Fleyshman B.S. Osnovy sistemologii / B.S. Fleyshman. – M.: Radio i svyaz'. 1982. – 368 s.

3. Tyrsin A.N. O matematicheskom modelirovanii slozhnykh organizatsionnykh sistem // Organizatsiya i upravlenie effektivnost'yu i proizvoditel'nost'yu proizvodstvennykh i sotsial'nykh sistem: Mater. mezhd. nauchno-prakt. konf. – Novocherkassk: YuRGTU (NPI), 2005. – S. 49 – 50.

4. Birger I.A. Tekhnicheskaya diagnostika. – M.: Nauka, 1978. – 239 s.

5. Debar H., Dacier M., Wespi A. Towards a taxonomy of intrusion-detection systems // Computer Networks. 1999. vol. 31. Issue 8. pp. 805 – 822.

6. Branitskiy A.A., Kotenko I.V. Analiz i klassifikatsiya metodov obnaruzheniya setevykh atak. – Trudy SPIIRAN. – 2016. – Vyp. 45 – S. 207 – 244.

7. Abdulin A.A., Sokolov A.N. Issledovanie programmnykh resheniy dlya obespecheniya informatsionnoy bezopasnosti promyshlennykh setey avtomatizirovannykh sistem upravleniya tekhnologicheskimi protsessami //Vestnik UrFO. Bezopasnost' v informatsionnoy sfere. – 2021. № 1(39). – C. 43 – 53.

8. Herve Debar, Monique Becker, and Didier Siboni. A neural network omponent for an intrusion detection system //Proceedings of the 1992 IEEE Computer Sociecty Symposium on Research in Security and Privacy, pages 240–250, Oakland, CA, USA, May 1992.

9. Rasool Jalili, Fatemeh Imani-Mehr, Morteza Amini, Hamid Reza Shahriari. Detection of Distributed Denial of Service Attacks Using Statistical Pre- Processor and Unsupervised Neural Networks // Lecture notes in computer science, 2005.

10. Smelyanskiy R.L., Kachalin A.I. Primeneniya neyrosetey dlya obnaruzheniya anomal'nogo povedeniya ob"ektov v komp'yuternykh setyakh //Fakul'tet Vychislitel'noy Matematiki i Kibernetiki, MGU im. M. V. Lomonosova, Moskva, 2004.

11. Neural network and artificial immune systems for malware and network intrusion detection / V. Golovko [et al.] // Studies in computational intelligence. – Heidelberg, 2010. – Vol. 263 : Advances in machine learning II. – P. 485 – 513.

12. Muna A.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. – № 41. – P. 1 – 11.

13. Sung Jin Kim, Woo Yeon Jo, Taeshik Shon. APAD: Autoencoder-based Payload Anomaly Detection for industrial IoE. December 2019 Applied Soft Computing 88:106017. DOI: 10.1016/j.asoc.2019.106017.

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

15. Schmidhuber J. Deep learning in neural networks: An overview //Neural networks. – 2015. – № 61. – P. 85 – 117.

16. Goodfellow I. et al. Generative adversarial nets //Advances in neural information processing systems. – 2014. – P. 2672 – 2680.

17. Madry A. et al. Towards deep learning models resistant to adversarial attacks //arXiv preprint arXiv:1706.06083. – 2017.

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

19. Alabugin S.K., Sokolov A.N. Ispol'zovanie generativno-sostyazatel'nykh neyronnykh setey pri vyyavlenii anomaliy tekhnologicheskogo protsessa //Vestnik UrFO. Bezopasnost' v informatsionnoy sfere. – 2020. № 4(38). – C. 64 – 68.

20. Garbuk S.V., Pravikov D.I., Polyanskiy A.V., Samarin I.V. Obespechenie informatsionnoy bezopasnosti ASU TP s ispol'zovaniem metoda prediktivnoy zashchity //Voprosy kiberbezopasnosti. – 2019. – № 3(31). – S. 30 – 36.

21. Quilot B., Génard M., Lescourret F., Kervella J. Simulating genotypic variation of fruit quality in an advanced peach×Prunus davidiana cross // Journal of Experimental Botany, Volume 56, Issue 422, December 2005, Pages 3071–3081, https://doi.org/10.1093/jxb/eri304.

22. Martin H., Tschabuschnig K., Bridal O., Watzenig D. (2017) Functional Safety of Automated Driving Systems: Does ISO 26262 Meet the Challenges?. In: Watzenig D., Horn M. (eds) Automated Driving. Springer, Cham. https://doi.org/10.1007/978-3-319-31895-0_16.

23. 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. № 3(25). – C. 34 – 42.

24. Parts 1 – 7 IEC 61508, Functional safety of electrical/electronic/ programmable electronic safety-related systems.

25. Braband J. Towards an IT security risk assessment framework for rail- way automation //CoRR abs/1704.01175, http://arxiv.org/abs/1704.01175, 2017.

26. Maidl M., Kröselberg D., Christ J. and Beckers K. A Comprehensive Framework for Security in Engineering Projects - Based on IEC 62443. // 2018 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW), 2018, pp. 42-47, doi: 10.1109/ISSREW.2018.00.

27. Pyatnitsky I.A., Sokolov A.N. Determination of the Optimal Ratio of Normal to Anomalous Points in the Problem of Detecting Anomalies in the Work of Industrial Control Systems //2021 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT), 2021, pp. 0478-0480, doi: 10.1109/USBEREIT51232.2021.9455010.

28. Alabugin S.K., Sokolov A.N. Applying of Generative Adversarial Networks for Anomaly Detection in Industrial Control Systems // 2020 Global Smart Industry Conference (GloSIC), 2020, pp. 199-203, doi: 10.1109/GloSIC50886.2020.9267878.

29. Alabugin S.K., Sokolov A.N. Applying of Recurrent Neural Networks for Industrial Processes Anomaly Detection // 2021 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT), 2021, pp. 0467-0470, doi: 10.1109/USBEREIT51232.2021.9455060.

30. Radford A., Metz L., Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks //arXiv preprint arXiv:1511.06434. – 2015.44.

31. Ledig C. et al. Photo-realistic single image super-resolution using a generative adversarial network //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2017. – C. 4681-4690.

32. Odena A., Olah C., Shlens J. Conditional image synthesis with auxiliary classifier gans //Proceedings of the 34th International Conference on Machine Learning-Volume 70. – JMLR.org, 2017. – C. 2642-2651.

33. Engel J. et al. Gansynth: Adversarial neural audio synthesis //arXiv preprint arXiv:1902.08710. – 2019.

34. Subramanian S. et al. Towards text generation with adversarially learned neural outlines //Advances in Neural Information Processing Systems. – 2018. – P. 7551-7563.

35. Madry A. et al. Towards deep learning models resistant to adversarial attacks //arXiv preprint arXiv:1706.06083. – 2017.

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

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

38. J. Goh, S. Adepu, K. N. Junejo and A. Mathur, "A dataset to support research in the design of secure water treatment systems", Proc. International Conference on Critical Information Infrastructures Security (CRITIS 2016), 2016.

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

РАГОЗИН Андрей Николаевич, кандидат технических наук, доцент кафедры защиты информации, ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». Россия, 454080, г. Челябинск, проспект Ленина, д. 76. E-mail: ragozinan@susu.ru.

БАРИНОВ Андрей Евгеньевич, старший преподаватель кафедры защиты информации, ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследователь-

ский университет)». Россия, 454080, г. Челябинск, проспект Ленина, д. 76. E-mail: barinovae@susu.ru.

УФИМЦЕВ Максим Сергеевич, преподаватель кафедры защиты информации, ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». Россия, 454080, г. Челябинск, проспект Ленина, д. 76. E-mail: ufimtcevms@susu.ru.

ПЯТНИЦКИЙ Илья Альбертович, аспирант кафедры защиты информации, ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». Россия, 454080, г. Челябинск, проспект Ленина, д. 76. E-mail: iamKane@mail.ru.

БУХАРЕВ Дмитрий Александрович, аспирант кафедры защиты информации, ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». Россия, 454080, г. Челябинск, проспект Ленина, д. 76. E-mail: bukharevdmitrii@gmail.com.

SOKOLOV Alexander Nikolaevich, Ph.D., Associate professor, Head of the Department of Information Security, South Ural State University (national research university). 76, Lenin prospect, Chelyabinsk, Russia, 454080. E-mail: sokolovan@susu.ru.

RAGOZIN Andrey Nikolaevich, Ph.D., Associate Professor of the Department of Information Security, South Ural State University (national research university). 76, Lenin prospect, Chelyabinsk, Russia, 454080. E-mail: ragozinan@susu.ru.

BARINOV Andrey Evgenievich, Senior Lecturer of the Department of Information Security, South Ural State University (national research university). 76, Lenin prospect, Chelyabinsk, Russia, 454080. E-mail: barinovae@susu.ru.

UFIMTCEV Maxim Sergeevich, Lecturer of the Department of Information Security, South Ural State University (national research university). 76, Lenin prospect, Chelyabinsk, Russia, 454080. E-mail: ufimtcevms@susu.ru.

PYATNITSKIY Ilya Albertovich, Post-graduate student of the Department of Information Security, South Ural State University (national research university). 76, Lenin prospect, Chelyabinsk, Russia, 454080. E-mail: iamKane@mail.ru.

BUKHAREV Dmitriy Aleksandrovich, Post-graduate student of the Department of Information Security, South Ural State University (national research university). 76, Lenin prospect, Chelyabinsk, Russia, 454080. E-mail: bukharevdmitrii@gmail.com.