УДК 004.056 DOI 10.18522/2311-3103-2024-5-79-88

И.В. Машкина, А.М. Уразаева

МЕТОД РАЗРАБОТКИ БАЗЫ ЗНАНИЙ СЦЕНАРИЕВ УГРОЗ ДЛЯ СИСТЕМЫ РЕАГИРОВАНИЯ НА ИНЦИДЕНТЫ (IRP)

Цель работы – исследование возможности повышения эффективности реагирования на инциденты информационной безопасности (ИБ). Это может быть достигнуто путем разработки системы, способной быстро локализовать инцидент, обеспечивающей автоматизацию реагирования на угрозу ИБ, предпринимающей заранее заданные действия в зависимости от деталей реализуемого сценария угрозы. Предложена архитектура построения IRP-системы, основными модулями которой являются база знаний сценариев реагирования, база знаний сценариев угроз, модули определения статуса инцидента и принятия решений по формированию командной информации. Решена задача разработки сценариев угроз для создания базы знаний сценариев, на основе которой могут быть разработаны адекватные сценарии реагирования, уникальные для каждой цепочки последовательности действий киберпреступника, событий и задействованных объектов. Формализован метод разработки базы знаний сценариев угроз на основе построения EPC-диаграмм сценариев, отображающих многокомпонентные атаки с учетом тактик, техник, используемых уязвимостей, угроз безопасности информации (УБИ), приведенных в нормативных документах и базах данных. Сформулированы правила построения EPC-диаграмм сценариев угроз и методика EPC-моделирования для объектов воздействия в АСУ ТП. Рассмотрен пример сценария атаки на промышленную сеть из глобальной сети в случае, когда киберпреступник, атаковав компьютер удаленного пользователя, в первую очередь осуществляет несанкционированный доступ в корпоративный сегмент, закрепляется в нем для дальнейшего проникновения за периметр технологической сети. Приведена разработанная EPC-диаграмма сценария угрозы с указанием используемых тактик, техник, промежуточных УБИ, некоторых уязвимостей. Формализована оценка вероятности реализации сценария.

Система реагирования на инциденты; тактики; техники; уязвимости; УБИ; база знаний сценариев угроз; база знаний сценариев реагирования; EPC-диаграмма.

I.V. Mashkina, A.M. Urazaeva

METHOD OF DEVELOPMENT OF THREAT SCENARIOS KNOWLEDGE BASE FOR INCIDENT RESPONSE PLATFORM (IRP)

The objective of the work is to study the possibility of increasing the efficiency of response to information security (IS) incidents. This can be achieved by developing a system capable of quickly localizing an incident, providing automation of response to an IS threat, taking predetermined actions depending on the details of the threat scenario being implemented. An architecture for constructing an IRP system is proposed, the main modules of which are a response scenario knowledge base, a threat scenario knowledge base, modules for determining the incident status and making decisions on the formation of command information. The problem of developing threat scenarios for creating a scenario knowledge base has been solved, on the basis of which adequate response scenarios can be developed that are unique for each chain of the cybercriminal's actions, events and involved objects. The paper formalizes the method for developing a knowledge base of threat scenarios based on constructing EPC diagrams of scenarios that display multi-component attacks taking into account tactics, techniques, vulnerabilities used, and information security threats (IST) specified in regulatory documents and databases. The paper formulates the rules for constructing EPC diagrams of threat scenarios and the methodology for EPC modeling for objects of influence in ICS. An example of an attack scenario on an industrial network from a global network is considered in the case when a cybercriminal, having attacked a remote user's computer, first gains unauthorized access to the corporate segment and gains a foothold in it for further penetration beyond the perimeter of the process network. The paper presents the developed EPC diagram of a threat scenario indicating the tactics, techniques, intermediate IST, and some vulnerabilities used. The assessment of the probability of scenario implementation is formalized.

Incident response system; tactics; technics; vulnerabilities; information security threats; threat scenarios knowledge base; EPC -diagram.

Введение. Приказы ФСТЭК России №31 и №239 утверждают требования к обеспечению защиты информации в АСУ ТП на критически важных и потенциально опасных объектах: атомной энергетики, добычи и транспортировки нефти и газа, оборотно-промышленного комплекса и транспорта [1, 2]. Специалисты в области ИБ отмечают, что в 2024 году проблемы защиты АСУ ТП остаются крайне актуальными и сложными [3–5]. Среди ключевых проблем, таких как: интеграция АСУ ТП с корпоративными ИТ-сетями, увеличение числа целенаправленных атак на системы управления технологическими процессами, необходимость обеспечения совместимости средств защиты информации с импортозамещающими отечественными SCADA-системами и ПЛК, – особо отмечаются нереализованные на многих объектах процессы мониторинга и реагирования на инциденты.

Отсутствие системы мониторинга, анализа угроз и реагирования на инциденты может привести к нарушению киберустойчивости промышленной системы и, следовательно, к нарушению непрерывности технологического процесса, увеличению времени на восстановление после атаки.

В последние годы особую актуальность приобрела тематика автоматизации реагирования на угрозы ИБ. Некоторые SIEM обладают встроенной IRP-системой, способной быстро локализовать инцидент и уменьшить или исключить разрушительность последствий.

Несколько самостоятельных IRP решений российского производителя для выполнения базовых задач находятся в промышленной эксплуатации: Jet Signal компании «Инфосистемы Джет», R-Vision компании «Р-Вижн», Security Vision компании «Интеллектуальная безопасность» [6–8]. На российском рынке представлены также продукты Израильской компании CyberBit SOC 3D, а также разработка компании IBM IBM Resilient IRP [9, 10].

Архитектура IRP. IRP (Incident Response Platform) – это система автоматизации реагирования на инциденты кибербезопасности, которая выполняет функции по сбору дополнительной информации, сдерживанию, устранению угрозы либо восстановлению системы после атаки, а также по структурированию данных о расследовании инцидента [11].

На рис. 1 приведена предлагаемая архитектура построения IRP системы. Основными модулями являются база знаний сценариев реагирования и база знаний сценариев угроз. Причем база сценариев реагирования может быть разработана на основе полного перечня всех возможных типов инцидентов ИБ, то есть на основе модели угроз конкретному объекту защиты, которая, как известно [12], включает в себя список актуальных угроз. Таким образом, эффективные сценарии реагирования на киберинциденты могут быть разработаны только с учетом сценариев угроз. На основе модели сценария угрозы формируется адекватный сценарий реагирования, уникальный для каждой последовательности событий и задействованных объектов. Сценарий реагирования представляет собой совокупность правил и выполняемых действий, специфичных для индикаторов – признаков этапов реализуемого сценария угрозы.

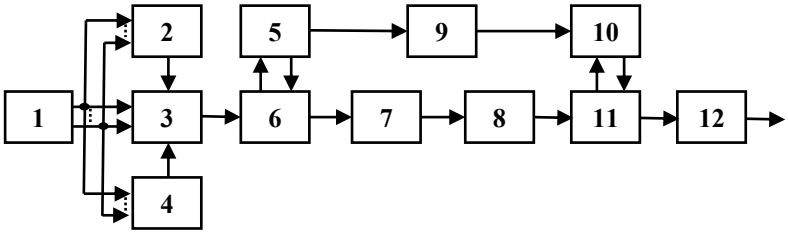

Рис. 1. Архитектура системы реагирования на инциденты

1 – источники данных о событиях безопасности (с SIEM системы);
2 – модуль определения задействованных объектов инфраструктуры;
3 – модуль анализа данных (техник, индикаторов и др.);
4 – модуль контроля привилегий;
5 – база знаний сценариев угроз целевым объектам инфраструктуры;

6 – модуль определения сценария угрозы в реальном времени;

7 – модуль численной оценки риска реализации угрозы;

8 – модуль определения статуса инцидента;

9 – модуль разработки плана реагирования;

10 – база знаний сценариев реагирования, адаптированных под конкретные сценарии угроз;

11 – модуль принятия решений;

12 – модуль формирования командной информации (Active Response) на агенты реагирования (запуск скриптов, воздействие на средства защиты и др.).

Рассмотрим, как может быть решена задача разработки сценариев угроз для создания базы знаний угроз и сценариев реагирования на их основе. В работах [13, 14] было предложено для моделирования угроз использовать принципы методологии ARIS [15]. Разработка EPC-диаграмм сценариев угроз позволяет отразить на одной схеме все значимые этапы многокомпонентной атаки. Четко обозначенные события, как результаты развития атаки, позволяют отметить и настроить точки контроля индикаторов для всех значимых этапов.

EPC-диаграмма процесса реализации угрозы целевому объекту АСУ ТП может быть представлена в виде комбинации событий и функций. При этом под функцией понимаем проводимые исполнителем-киберпреступником тактики и техники. Таким образом, функция – это действие или набор действий, выполняемых в информационной среде объекта защиты киберпреступником; функция может быть поименована соответствующей техникой из приложения 11 методики [12].

EPC-диаграмма сценария угрозы представляет собой отображение (сверху вниз) последовательности выполнения техник, начиная от исходного действия киберпреступника до достижения им цели – угрозы безопасности информации (УБИ) объекту воздействия, это может быть УБИ из банка данных [16]. Реализация на определенном этапе техники или их совокупности вызывает событие – состояние информационной среды, которое может быть оценено как некоторая промежуточная УБИ, существенная для достижения объекта воздействия угрозы. Эта промежуточная УБИ оказывает влияние на дальнейшее развитие сценария. Событие отображается на диаграмме как специальный элемент. В свою очередь событие – промежуточная УБИ – активизирует последующие тактики. Функции и события в процессе реализации сценария чередуются. Решение о развитии сценария, т.е. ходе выполнения многокомпонентной атаки, принимается киберпреступником по мере поиска уязвимостей для реализации техник.

Правила построения EPC-диаграмм сценариев многокомпонентных атак следующие:

- ♦ успешное выполнение киберпреступником техники приводит одновременно к нескольким событиям, для отображения на диаграмме используется оператор AND;
- ♦ промежуточная УБИ реализуется после одновременного выполнения двух техник, для отображения используется оператор AND перед УБИ;
- ♦ очередная техника может быть выполнена только после реализации двух УБИ, это отображается с помощью оператора AND перед техникой;
- ♦ промежуточная УБИ обеспечивает возможности выполнения двух техник, используется оператор AND перед техниками;
- ♦ выполнение техники может привести к реализации одной из двух или двух УБИ, на диаграмме это обозначается с помощью оператора OR после техники и перед $УБИ_i$ и $УБИ_j$;
- ♦ УБИ может быть реализована после выполнения одной из двух или двух техник, используется оператор OR перед УБИ;
- ♦ техника сможет быть выполнена после реализации одного из двух событий, или двух одновременно, используется оператор OR перед техникой;
- ♦ выполнение какой-либо техники может привести только к одному из событий: удалось реализовать УБИ или не удалось, это обозначается на диаграмме с помощью оператора XOR перед событиями;

♦ УБИ реализуется в результате наличия уязвимостей для выполнения только одной из двух техник, используется оператор XOR перед УБИ;

♦ техника может быть выполнена сразу после реализации одной из УБИ, тогда используется оператор XOR перед техникой.

Методика EPC-моделирования сценариев многокомпонентных атак заключается в следующем:

♦ задание целевого объекта воздействия угрозы в АСУ ТП: SCADA система, АРМ инженера-программиста по управлению процессом, OPC-сервер, АРМ оператора, PLC-контроллер, рабочая станция бизнес-контура сети с технологической сетью;

♦ определение источника угрозы: киберпреступник, удаленный пользователь-нарушитель, внутренний нарушитель;

♦ определение состояния информационной среды на начало реализации сценария;

♦ выявление возможной начальной тактики и соответствующих техник;

♦ обозначение событий, к которым могут привести реализуемые техники;

♦ оценка наличия уязвимостей информационной среды, эксплуатация которых позволит нарушителю выполнение техник в логической цепочке взаимодействия событий и функций;

♦ поиск промежуточных УБИ многокомпонентной атаки;

♦ достижение нарушителем цели реализации сценария – возможности нарушения свойства безопасности информации объекта воздействия угрозы;

♦ получение численной оценки вероятности реализации сценария угрозы.

Построение сценария атаки на корпоративный сегмент промышленного предприятия с технологической сетью. По статистике NIST [17] самыми опасными являются целенаправленные внешние атаки.

Для реализации атаки на промышленную сеть через глобальную, киберпреступник в первую очередь должен осуществить несанкционированный доступ в корпоративный сегмент и закрепиться в нем для дальнейшего проникновения за периметр промышленной сети.

Рассмотрим пример, когда удаленному пользователю предоставляется доступ в корпоративный сегмент (бизнес-контур) через VPN-соединение. В этом случае уровень защищенности сети компании, в том числе технологической сети, становится зависимым от защищенности компьютера удаленного пользователя. Если должные меры безопасности не реализованы или имеются уязвимости, киберпреступник, атаковав узел, может использовать VPN для туннелирования трафика за периметр корпоративного сегмента сети предприятия с АСУ ТП.

В результате заражения компьютера удаленного пользователя вредоносным программным обеспечением, оказываются скомпрометированы аутентификационные данные пользователя, который может иметь право доступа к каким-либо серверам бизнес-контура. Таким образом, с помощью троянской программы, которая собрала логины, пароли и другие данные, необходимые для связи с сервером бизнес-контура по каналу VPN, киберпреступник создает на своем компьютере копию скомпрометированного узла и осуществляет попытку нарушения периметра и далее подключения к серверу бизнес-контура сети, имея данные идентификации и аутентификации удаленного легитимного пользователя.

При наличии ошибок настройки прав доступа пользователей к ресурсам сервера киберпреступник после авторизации на сервере может определить директорию, в которой хранятся хэш-суммы паролей пользователей. Далее возможна процедура обратного пересчета хэш-функций, перебор собранных паролей и попытка авторизоваться на сервере под логином администратора. Права администратора позволяют ему, изменив конфигурацию сетевого оборудования, получить доступ в АСУ ТП для реализации атаки на целевой объект воздействия. На рис. 2 приведена разработанная EPC-диаграмма сценария атаки на корпоративный сегмент промышленного предприятия. В табл. 1 приведен перечень используемых киберпреступником тактик, техник, найденных уязвимостей и реализуемых УБИ.

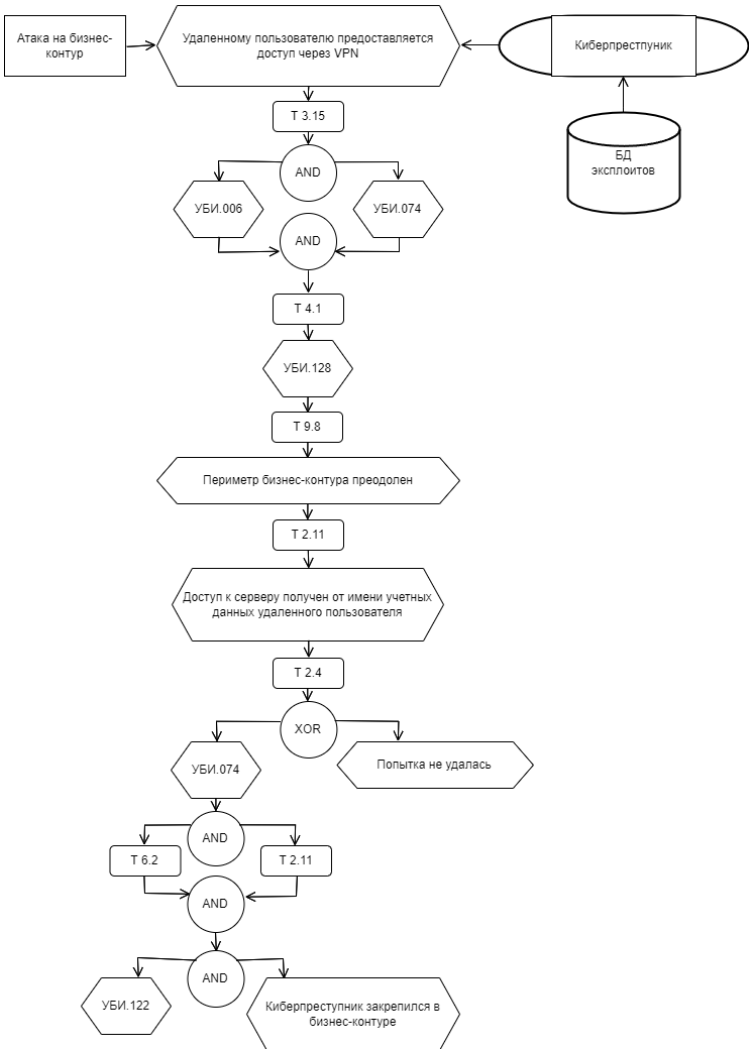

Рис. 2. EPC-диаграмма сценария атаки на корпоративный сегмент

Таблица 1

Применяемые тактики, техники, реализуемые УБИ и их описание

Номер тактики	**Применяемые техники и их описание**	**Реализуемые УБИ**	**Описание**
Т3: Внедрение и исполнение вредоносного программного обеспечения в системах и сетях	Т3.15 – Планирование запуска вредоносных программ через планировщиков задач в операционной системе, а также с использованием механизмов планирования выполнения в удаленной системе через удаленный вызов процедур. Выполнение в контексте планировщика в ряде случаев позволяет авторизовать вредоносное программное обеспечение и повысить доступные ему привилегии. (CVE-2023-21541)	УБИ.006, УБИ.074	УБИ.006: Угроза внедрения кода или данных УБИ.074: Угроза несанкционированного доступа к аутентификационной информации.

Окончание табл. 1

Т4: Закрепление (сохранение доступа) в системе или сети	Т4.1 – Несанкционированное создание учетных записей или кража существующих учетных данных. (CVE-2021-38704)	УБИ.128	УБИ.128: Угроза подмены доверенного пользователя
Т9: Сбор и вывод из системы или сети информации, необходимой для дальнейших действий при реализации угроз безопасности информации или реализации новых угроз	Т9.8 – Туннелирование трафика передачи данных через VPN	Событие	Периметр бизнес-контура преодолен
Т2: Получение первоначального доступа к компонентам систем и сетей	Т2.4 – Использование ошибок конфигурации сетевого оборудования и средств защиты, в том числе слабых паролей и паролей по умолчанию, для получения доступа к компонентам систем и сетей при удаленной атаке; Т2.11 – Несанкционированный доступ путем компрометации учетных данных сотрудника организации, в том числе через компрометацию многократно используемого в различных системах пароля (для личных или служебных нужд). (CVE-2022-23724)	УБИ.074, УБИ.122	УБИ.074: Угроза несанкционированного доступа к аутентификационной информации.
			УБИ.122: Угроза повышения привилегий
Т6: Повышение привилегий по доступу к компонентам систем и сетей	Т6.2 – Подбор пароля или другой информации для аутентификации от имени привилегированной учетной записи. (CVE-2023-27121)	УБИ.122	УБИ.122: Угроза повышения привилегий

EPC-диаграмма сценария угрозы позволяет оценить численно вероятность реализации сценария и его ранг по следующему алгоритму:

$$P_{\text{сц}} = \prod_{i=1,j=1}^{I,J}\left[W_{CVE}(T_i) + W_{CVE}(T_j) - W_{CVE}(T_i) * W_{CVE}(T_j)\right] * \prod_{x=1,y=1}^{X,Y}\left[W_{CVE}(T_x) * W_{CVE}(T_y)\right] * 0{,}5^{K} * \prod_{l=1}^{L} W_{CVE}(T_l), \quad (1)$$

где $P_{\text{сц}}$ – вероятность реализации сценария угрозы;

$W_{CVE}(T)$ – нормированные значения уязвимостей, эксплуатируемых киберпреступником при реализации техник T_i, T_j, T_x, T_y, T_l;

I = J – число техник, объединенных оператором OR;

X = Y – число техник, объединенных оператором AND;

K – число техник, для которых не найдены уязвимости в базе данных (уязвимости нулевого дня);

L – число техник, каждая из которых приводит к промежуточной УБИ.

$$R_{\text{сц}} = P_{\text{сц}} * C_a,$$

где $R_{\text{сц}}$ – величина риска реализации сценария;

C_a – ценность актива (объекта воздействия угрозы).

Ранг сценария определяется в зависимости от показателя риска.

Метод разработки базы знаний сценариев угроз включает в себя: правила построения EPC-диаграмм сценариев угроз, методику EPC-моделирования, оценку вероятности реализации сценария угрозы от значений эксплуатируемых уязвимостей и структуры EPC-диаграммы сценария (числа используемых логических операторов).

Результаты численного эксперимента. Численное значение вероятности реализации угрозы УБИ.122 рассчитывалось по формуле (1).

При этом значения эксплуатируемых при выполнении техник уязвимостей заимствованы из базы данных NIST: NVD [18–21], метрики CVSS Version 3.x. Значения уязвимостей приведены в табл. 2.

Таблица 2

Значения уязвимостей

№	Наименование уязвимости	Значение Base Score	Нормированное значение
1	CVE-2023-21541	7,8 HIGH	0,78
2	CVE-2021-38704	6,1 MEDIUM	0,61
3	CVE-2022-23724	8,1 HIGH	0,81
4	CVE-2023-27121	6,1 MEDIUM	0,61

В расчетах вероятность реализации техники T9.8 (туннелирование трафика передачи данных через VPN) принята равной единице, поскольку киберпреступнику удалось на предыдущих этапах многокомпонентной атаки получить права авторизованного удаленного пользователя; вероятность успешной для киберпреступника реализации техники T 2.4 принята равной 0,5, поскольку уязвимость не найдена, а результаты выполнения техники связаны оператором «исключающее ИЛИ».

Таким образом, $P_{УБИ.122}$ = 0,78* 0,61* 1* 0,81* 0,5* 0,61* 0,81 = 0,095.

Тогда численное значение риска ИБ для данной угрозы при ценности ресурса, принятой равной 0,02 составит 0,0019, то есть приближенно 0,2%.

После вычисления (для объекта воздействия угроз) вероятностей реализации всех возможных сценариев проводится их ранжирование.

Заключение. Предложена архитектура IRP системы, метод разработки базы знаний сценариев угроз, приведен пример построения сценария угрозы на корпоративный сегмент с промышленной сетью и результаты численного эксперимента по оценке вероятности реализации сценария. Полученные результаты позволят повысить эффективность детектирования сложных многокомпонентных атак, направленных на целевые объекты воздействия промышленной сети, будут способствовать адекватному реагированию. Функционирование модуля принятия решения по выбору варианта реагирования в составе архитектуры IRP связано с оценкой вероятности реализации сценария, его ранга; алгоритм должен обеспечивать минимизацию ущерба как от возможной атаки, так и от ответных действий.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Приказ ФСТЭК России от 14 марта 2014 № 31 «Об утверждении Требований к обеспечению защиты информации в автоматизированных системах управления производственными и технологическими процессами на критически важных объектах, потенциально опасных объектах, а также объектах, представляющих повышенную опасность для жизни и здоровья людей и для окружающей природной среды».
2. Приказ ФСТЭК России от 25 декабря 2017 № 239 «Об утверждении Требований по обеспечению безопасности значимых объектов критической информационной инфраструктуры Российской Федерации».
3. Information Security Информационная безопасность: офиц. сайт. – URL: https://www.itsec.ru/articles (дата обращения: 16.08.2024).

4. *Gaggero G.B., Armellin A., Portomauro G. and Marchese M*. Industrial Control System-Anomaly Detection Dataset (ICS-ADD) for Cyber-Physical Security Monitoring in Smart Industry Environment // IEEE Access. – 2024. – Vol. 12. – P. 64140-64149.
5. *Makrakis, Georgios Michail & Kolias, Constantinos & Kambourakis, Georgios & Rieger, Craig & Benjamin, Jacob*. Vulnerabilities and Attacks Against Industrial Control Systems and Critical Infrastructures. – 2021. – URL: https://www.researchgate.net/publication/354493711_Vulnerabilities_and_Attacks_Against_Industrial_Control_Systems_and_Critical_Infrastructures (дата обращения: 16.08.2024).
6. Jet Detective и Jet Signal внесены в реестр отечественного программного обеспечения. 12 декабря 2017. Инфосистемы Джет: офиц. сайт. – URL: https://jet.su/press-center/news/jet-detective-i-jet-signal-vneseny-v-reestr-otechestvennogo-programmnogo-obespecheniya (дата обращения: 15.08.2024).
7. Р-Вижн: офиц. сайт. – URL: https://www.rvision.ru/ (дата обращения 15.08.2024).
8. Интеллектуальная безопасность: офиц. сайт. – URL: https://www.securityvision.ru/docs/IRP.php (дата обращения: 15.08.2024).
9. Cyberbit SOC 3D. Автоматизированная консолидация всех данных управления процессом реагирования на киберинциденты на единую панель управления и повышение эффективности SOC в целом. – URL: https://www.pacifica.kz/upload/iblock/f43/f438e9f735ad1f4218e45acb9db8d706.pdf?ysclid=lzx8mckt1r891749321 (дата обращения: 16.08.2024).
10. IBM Resilient Incident Response Platform Enterprise on Cloud delivers orchestrated and automated incident response processes. IBM Asia Pacific Software Announcement AP16-0410. November 1, 2016. – URL: https://www.ibm.com/docs/en/announcements/archive/ENUSAP16-0410#abstrx (дата обращения: 15.08.2024).
11. IRP (Incident Response Platform). – URL: https://encyclopedia.kaspersky.ru/glossary/irp/ (дата обращения: 16.08.2024).
12. Методический документ ФСТЭК России от 05.02.2021. Методика оценки угроз безопасности информации. – URL: https://normativ.kontur.ru/document?moduleId=1&documentId=451500 (дата обращения: 15.08.2024).
13. *Машкина И.В., Гарипов И.Р.* Разработка EPC-моделей угроз нарушения информационной безопасности автоматизированной системы управления технологическими процессами // Безопасность информационных технологий, [S.I.]. – 2019. – Т. 26, № 4. – С. 6-20. – ISSN 2074-7136. – DOI: http://dx.doi.org/10.26583/bit.2019.4.01.
14. *Заид Алкилани М.О., Машкина И.В.* Разработка сценариев атак для оценки угроз нарушения информационной безопасности в промышленной сети // Проблемы информационной безопасности. Компьютерные системы. – 2024. – № 1 (58). – С. 96-109. – DOI 10.48612/jisp/xvkx-k619-3f2z 25.04.2024. – EDN: PDNEWN.
15. *Шеер А.В.* ARIS-моделирование бизнес-процессов. – М.: Вильямс, 2000. – 175 с.
16. Банк данных угроз безопасности информации Федеральная служба по техническому и экспортному контролю России. – URL: https://bdu.fstec.ru (дата обращения: 16.08.2024).
17. *Stouffer K., Pease M., Tang CY., Zimmerman T., Pillitteri V., Lightman S., Hahn A., Saravia S., Sherule A., Thompson M.* Title. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) NIST SP 800-82r3. – 2023. – https://doi.org/10.6028/NIST.SP.800-82r3.
18. CVE-2023-21541. – https://nvd.nist.gov/vuln/detail/CVE-2023-21541 (дата обращения: 26.10.2024).
19. CVE-2021-38704. – https://nvd.nist.gov/vuln/detail/CVE-2021-38704 (дата обращения: 26.10.2024).
20. CVE-2022-23724. – https://nvd.nist.gov/vuln/detail/CVE-2022-23724 (дата обращения: 26.10.2024).
21. CVE-2023-27121. – https://nvd.nist.gov/vuln/detail/CVE-2023-27121 (дата обращения: 26.10.2024).

REFERENCES

1. Prikaz FSTEK Rossii ot 14 marta 2014 № 31 «Ob utverzhdenii Trebovaniy k obespecheniyu zashchity informatsii v avtomatizirovannykh sistemakh upravleniya proizvodstvennymi i tekhnologicheskimi protsessami na kriticheski vazhnykh ob"ektakh, potentsial'no opasnykh ob"ektakh, a takzhe ob"ektakh, predstavlyayushchikh povyshennuyu opasnost' dlya zhizni i zdorov'ya lyudey i dlya okruzhayushchey prirodnoy sredy» [Order of the FSTEC of Russia dated March 14, 2014 No. 31 "On approval of the Requirements for ensuring information security in automated control systems for production and technological processes at critical facilities, potentially hazardous facilities, as well as facilities posing an increased danger to human life and health and to the environment"].
2. Prikaz FSTEK Rossii ot 25 dekabrya 2017 № 239 «Ob utverzhdenii Trebovaniy po obespecheniyu bezopasnosti znachimykh ob"ektov kriticheskoy informatsionnoy infrastruktury Rossiyskoy Federatsii» [Order of the FSTEC of Russia dated December 25, 2017 No. 239 "On approval of the Requirements for ensuring the security of significant facilities of the critical information infrastructure of the Russian Federation"].

3. Information Security: офиц. сайт. Available at: https://www.itsec.ru/articles (accessed 16 August 2024).
4. *Gaggero G.B., Armellin A., Portomauro G. and Marchese M*. Industrial Control System-Anomaly Detection Dataset (ICS-ADD) for Cyber-Physical Security Monitoring in Smart Industry Environment, *IEEE Access,* 2024, Vol. 12, pp. 64140-64149.
5. *Makrakis, Georgios Michail & Kolias, Constantinos & Kambourakis, Georgios & Rieger, Craig & Benjamin, Jacob*. Vulnerabilities and Attacks Against Industrial Control Systems and Critical Infrastructures, 2021. Available at: https://www.researchgate.net/publication/354493711_ Vulnerabilities_and_Attacks_Against_Industrial_Control_Systems_and_Critical_Infrastructures (accessed 16 August 2024).
6. Jet Detective i Jet Signal vneseny v reestr otechestvennogo programmnogo obespecheniya. 12 dekabrya 2017. Infosistemy Dzhet: ofits. sayt [Jet Detective and Jet Signal are included in the register of domestic software. December 12, 2017. Jet Infosystems: official website]. Available at: https://jet.su/press-center/news/jet-detective-i-jet-signal-vneseny-v-reestr-otechestvennogo-programmnogo-obespecheniya (accessed 15 August 2024).
7. R-Vizhn: ofits. sayt [R-Vision: official website]. Available at: https://www.rvision.ru/ (accessed 15 August 2024).
8. Intellektual'naya bezopasnost': ofits. sayt [Intelligent security: official website]. Available at: https://www.securityvision.ru/docs/IRP.php (accessed 15 August 2024).
9. Cyberbit SOC 3D. Avtomatizirovannaya konsolidatsiya vsekh dannykh upravleniya protsessom reagirovaniya na kiberintsidenty na edinuyu panel' upravleniya i povyshenie effektivnosti SOC v tselom [Cyberbit SOC 3D. Automated consolidation of all cyber incident response management data on a single control panel and increasing the efficiency of the SOC as a whole]. Available at: https://www.pacifica.kz/upload/iblock/f43/f438e9f735ad1f4218e45acb9db8d706.pdf?ysclid=lzx8mckt1r891749321 (accessed 16 August 2024).
10. IBM Resilient Incident Response Platform Enterprise on Cloud delivers orchestrated and automated incident response processes. IBM Asia Pacific Software Announcement AP16-0410. November 1, 2016. Available at: https://www.ibm.com/docs/en/announcements/archive/ENUSAP16-0410#abstrx (accessed 15 August 2024).
11. IRP (Incident Response Platform). Available at: https://encyclopedia.kaspersky.ru/glossary/irp/ (accessed 16 August 2024).
12. Metodicheskiy dokument FSTEK Rossii ot 05.02.2021. Metodika otsenki ugroz bezopasnosti informatsii [Methodological document of the FSTEC of Russia dated 05.02.2021. Methodology for assessing information security threats]. Available at: https://normativ.kontur.ru/document?moduleId=1&documentId=451500 (accessed 16 August 2024).
13. *Mashkina I.V., Garipov I.R.* Razrabotka ERS-modeley ugroz narusheniya informatsionnoy bezopasnosti avtomatizirovannoy sistemy upravleniya tekhnologicheskimi protsessami [Development of EPC models of threats to information security of an automated process control system], *Bezopasnost' informatsionnykh tekhnologiy* [Security of Information Technology], [S.I.], 2019, Vol. 26, No. 4, pp. 6-20. ISSN 2074-7136. DOI: http://dx.doi.org/10.26583/bit.2019.4.01.
14. *Zaid Alkilani M.O., Mashkina I.V.* Razrabotka stsenariev atak dlya otsenki ugroz narusheniya informatsionnoy bezopasnosti v promyshlennoy seti [Development of attack scenarios for assessing threats of information security breach in an industrial network], *Problemy informatsionnoy bezopasnosti. Komp'yuternye sistemy* [Problems of information security. Computer systems], 2024, No. 1 (58), pp. 96-109. DOI 10.48612/jisp/xvkx-k619-3f2z 25.04.2024. – EDN: PDNEWN.
15. *Sheer A.V.* ARIS-modelirovanie biznes-protsessov [ARIS-modeling of business processes]. Moscow: Vil'yams, 2000, 175 p.
16. Bank dannykh ugroz bezopasnosti informatsii Federal'naya sluzhba po tekhnicheskomu i eksportnomu kontrolyu Rossii [Database of information security threats Federal Service for Technical and Export Control of Russia]. Available at: https://bdu.fstec.ru (accessed 16 August 2024).
17. *Stouffer K., Pease M., Tang CY., Zimmerman T., Pillitteri V., Lightman S., Hahn A., Saravia S., Sherule A., Thompson M.* Title. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) NIST SP 800-82r3, 2023. Available at: https://doi.org/10.6028/NIST.SP.800-82r3.
18. CVE-2023-21541. Available at: https://nvd.nist.gov/vuln/detail/CVE-2023-21541 (accessed 26 October 2024).
19. CVE-2021-38704. Available at: https://nvd.nist.gov/vuln/detail/CVE-2021-38704 (accessed 26 October 2024).

20. CVE-2022-23724. Available at: https://nvd.nist.gov/vuln/detail/CVE-2022-23724 (accessed 26 October 2024).
21. CVE-2023-27121. Available at: https://nvd.nist.gov/vuln/detail/CVE-2023-27121 (accessed 26 October 2024).

Статью рекомендовал к опубликованию д.т.н., профессор А.Н. Целых.

Машкина Ирина Владимировна – Уфимский университет науки и технологий; e-mail: profmashkina@mail.ru; г. Уфа, Россия; тел.: +79279277089; кафедра вычислительной техники и защиты информации; д.т.н.; профессор.

Уразаева Айгуль Маратовна – e-mail: ajgul.urazaeva2017@yandex.ru; кафедра вычислительной техники и защиты информации; студент.

Mashkina Irina Vladimirovna – Ufa University of Science and Technology; e-mail: profmashkina@mail.ru; Ufa, Russia; phone: +79279277089; the Department of Computer Science and Information Security; dr. of eng. sc.; professor.

Urazaeva Aigul Maratovna – e-mail: ajgul.urazaeva2017@yandex.ru; the Department of Computer Science and Information Security; student.

УДК 004.853 DOI 10.18522/2311-3103-2024-5-88-102

Е.М. Герасименко, Ю.А. Кравченко, Д.А. Шаненко

АЛГОРИТМ ПОИСКА И ПРИОБРЕТЕНИЯ ЗНАНИЙ НА ОСНОВЕ ТЕХНОЛОГИЙ ОБРАБОТКИ И АНАЛИЗА ТЕКСТОВ НА ЕСТЕСТВЕННОМ ЯЗЫКЕ*

Статья посвящена решению актуальной научной проблемы повышения эффективности обработки и анализа текстовой информации при решении задач поиска и приобретения знаний. Актуальность данной задачи связана с необходимостью создания эффективных средств обработки накапливаемого огромного количества слабо структурированных данных, содержащих важные, иногда скрытые знания, необходимые для построения эффективных систем управления сложными объектами различной природы. Предлагаемый автором алгоритм поиска и приобретения знаний при обработке и анализе текстовой информации, отличается применением низкоуровневых детерминированных правил, позволяющих провести качественное упрощение текста на основе исключения из текстовой информации слов, инвариантных к смыслу. Алгоритм опирается на доменную проработку, позволяющую сформировать списки доменно-специфичных слов, что позволяет обеспечить высокое качество упрощения текста. В данной задаче исходными данными являются потоки текстовой информации (описание профилей), извлеченных из онлайн платформ для рекрутинга, выходная информация представляется предложениями, сформированными в виде тройки «субъект-глагол-объект», отражающих гранулы знаний, полученных в процессе обработки текста. Использование данного порядка единиц, составляющих предложение, обусловлено тем фактом, что данный порядок наиболее распространен в русском языке, хотя в самих текстах возможны иные вариации порядка без потери общего смысла. Основная идея алгоритма заключается в разбиении большого корпуса текста на предложения с последующей фильтрацией полученных предложений на основании введенных пользователем ключевых слов. В последствии предложения разделяются на компоненты и упрощаются в зависимости от вида поступившей компоненты (глагольная, именная). В качестве примера в данной работе использовалась сфера маркетинга, а ключевыми словами выступили «социальные сети». Автором разработан алгоритм поиска и приобретения знаний на основе технологий обработки и анализа текстов на естественном языке, а также была выполнена программная реализация предложенного алгоритма. В качестве методов оценки эффективности использовался ряд метрик: индекс Флэша-Кинкейда; индекс Кол-

* Исследование выполнено за счет гранта Российского научного фонда № 22-71-10121, https://rscf.ru/project/22-71-10121/ в Южном федеральном университете.